

A rapid, flexible and scalable DNA assembly platform for genome engineering and regulated gene expression applications in *Plasmodium falciparum*

by

Armiyaw Sebastian Nasamu

B.S. Biology and Chemistry (2013)
Massachusetts Institute of Technology, Cambridge, MA

Submitted to the Department of Biological Engineering in partial fulfillment of the requirements for the degree of

Master of Science in Toxicology
at the
Massachusetts Institute of Technology

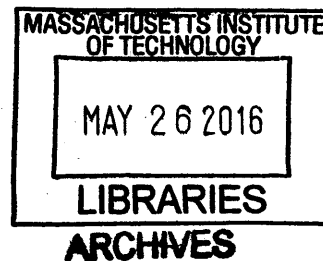
July 2015 [September 2015]

© 2015 Massachusetts Institute of Technology. All rights reserved.

Author **Signature redacted**
Armiyaw Sebastian Nasamu
Department of Biological Engineering, MIT

Certified by **Signature redacted**
Jacquin C. Niles
Associate Professor, Department of Biological Engineering, MIT
Thesis Advisor

Approved by **Signature redacted**
Forest M. White
Associate Professor, Department of Biological Engineering, MIT
Co-Chair, Course XX Graduate Program Committee





77 Massachusetts Avenue
Cambridge, MA 02139
<http://libraries.mit.edu/ask>

DISCLAIMER NOTICE

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available.

Thank you.

Some pages in the original document contain text that runs off the edge of the page.

Page 12

CONTENTS

0.1 Acknowledgements and attributions	4
0.2 Abstract.....	5
1 Introduction	
1.1 The TetR-aptamer system for conditional gene expression in <i>P. falciparum</i>	6
1.2 Genome editing in <i>P. falciparum</i> - the state of the art.....	7
1.3 Integrating CRISPR/Cas9 with the TetR-aptamer system.....	7
2 Linear plasmids for cloning [A+T] rich and repetitive <i>P. falciparum</i> DNA	
2.1 The pJAZZ-OC plasmid.....	7
2.2 Converting traditional plasmids to linear plasmids.....	9
2.3 The pSwing plasmid.....	9
2.4 Converting pSN372 to linear context.....	11
2.5 Rescuing linear plasmids into circular plasmids.....	14
2.6 Transfecting <i>P. falciparum</i> parasites with rescued plasmids.....	15
2.7 Evaluating regulatory components for controlling gene expression in BAC framework.....	17
3 Creating modular linear vectors configured to deliver pre-programmed functionalities	
3.1 pSN150	20
3.1.1 Architecture of pSN150.....	21
3.1.2 How to use pSN150.....	23
3.1.3 GSK-3.....	28
3.1.4 Hexose transporter.....	32

3.1.5 Thioredoxin Reductase	37
3.1.6 Chloroquine resistance transporter.....	39
3.1.7 Choline Kinase.....	40
3.2 PSN053.....	41
3.2.1 PF3D7_1360200 (MAL299).....	43
3.2.2 PF3D7_0925900.....	53
3.3 PSN153/154.....	55
3.3.1 PF3D7_0925900	59
3.4 Conclusions and future directions	62
3.5 Methods.....	64
3.6 References	67

Acknowledgements and attributions

I will first like to thank Professor Jacquin Niles for all the support in the last four years. From the day I met with you to discuss possible projects as a UROP, I have admired your enthusiasm for science and the critical thought processes you have. I learned so much from you and will always be indebted to you. Thanks for the financial support for the master's program as well. Next, I will like to thank the BE department for giving me the opportunity to continue to work in the lab for the last two years. I appreciate the financial support.

Some of the work described here (especially the 5' aptamer work) was done in collaboration with Bridget Wall. Thanks for being a great partner and thanks for all the hours you to put in to push this project forward. I will acknowledge Alejandra Falla for her contribution in constructing pSN154 and performing experiments to show its utility.

I will also like to thank all members of the Niles Lab, past and present for being awesome lab mates. I will specifically like to thank Jeffrey Wagner who was my mentor when I joined the lab in 2011. I learned almost all the techniques described in this document from you. I am thankful! I also want to thank Denise for all the amazing work she does to get the lab running. Sorry for all the late orders and lost packing slips. And thanks for helping with all my crazy visa and passport issues.

I had the opportunity to mentor a couple of students in the lab as well- Noah Prince and Jason Nguyen. Thanks for helping me learn how to teach what I did.

Lastly I want to thank my family, host family and close friends who have always been supportive of me. Without you, I would have never got any of this work done.

Abstract

Plasmodium falciparum is the deadliest malaria parasite. There is no approved vaccine to prevent this disease, and resistance to available antimalarial drugs is becoming widespread. Identification of parasite genes essential to survival and virulence could facilitate the development of novel therapeutics and vaccines. However, these efforts have been impeded by difficulties in manipulating the parasite's genome and functionally perturbing gene expression in a controlled way. Our lab has developed inducible systems to control *P. falciparum* gene expression, and has achieved successful editing of the *P. falciparum* genome using CRISPR/Cas9 technology. We have integrated these capabilities into a modular and scalable framework that can be used to efficiently edit, regulate and delete any target parasite gene after a single genome editing operation. This approach will accelerate studies of parasite gene function, and help prioritize potential drug and vaccine targets. A key requirement in this framework is the efficient assembly of donor vectors for modifying target loci and installing the necessary regulatory parts. This necessitates cloning several large, [A+T]-rich *P. falciparum* genomic regions that can be quite tedious and rate limiting. In this document, we show a new cloning strategy using linear vectors that facilitates rapid and accurate assembly of vectors capable of transforming *P. falciparum*. We present evidence of successful chromosomal modification of several genes via spontaneous single crossover, as well as zinc finger nuclease- and Cas9- mediated genome editing strategies using our assembled donor vectors. We also show that these modifications enable controllable expression of several previously uncharacterized genes to elicit phenotypes that we are investigating in further mechanistic detail. Importantly, these transgenic parasites can now be rapidly generated to allow identification of novel essential parasite genes in as little as a month.

Introduction

In this chapter, the various tool kits that allow for the manipulation and control of gene expression in *P. falciparum* will be briefly discussed

The TetR-aptamer system for conditional gene expression in P. falciparum

The TetR aptamer system, developed by the Niles lab, is a translational control tool for conditional protein expression in *P. falciparum* [1, 2]. An RNA sequence included in the 5' end of the regulated gene binds to the TetR repressor protein with high affinity. In this state the TetR protein prevents the translation of the targeted mRNA. In the presence of an inducer molecule, anhydrotetracycline (aTc), the TetR protein preferentially binds the inducer leaving the targeted mRNA “free” for translation (figure 1 from [3]).

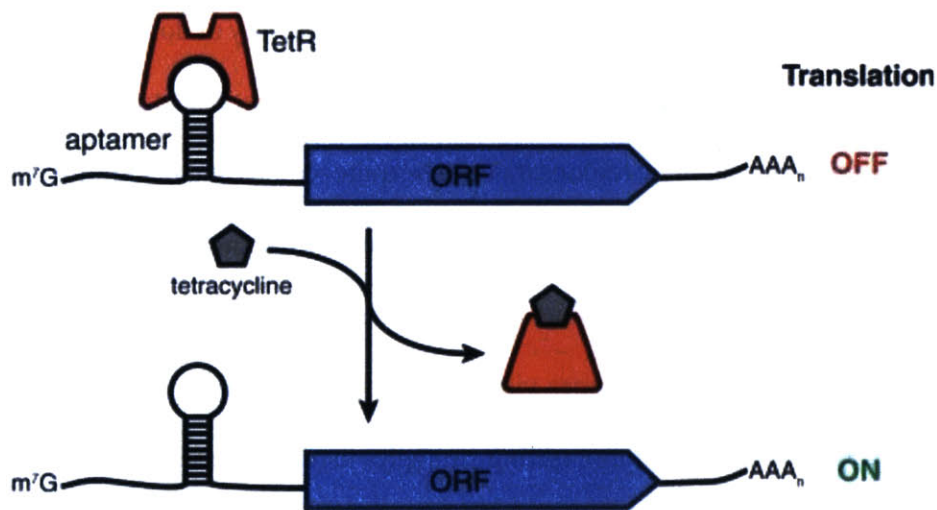


Figure 1: Figure 1A from [3]. How the TetR-RNA aptamer system works. In the absence of aTc, no protein is made. aTc induces protein expression

This TetR-aptamer system yields a 4-5 fold knockdown in protein expression of target genes. The system has since been modified in the Niles Lab to give up to 99% knockdown and an even wider dynamic range [9]. The modified system places 10 aptamers consecutively at the 3' end of target gene and employs a TetR protein fused to a native translational repressor protein from *P. falciparum*, called DOZI.

Genome editing in P. falciparum- the state of the art

As in many other organisms studied in the laboratory, the CRISPR/Cas9 system developed in the Niles lab [4] is currently the state of the art in genome editing in *P. falciparum*

Integrating CRISPR/Cas9 with the TetR-aptamer system

With the ability to regulate *Plasmodium falciparum* genes using the TetR-aptamer system and with the ability to modify the genome using CRISPR/Cas9, a modular system that allows scalable modification of hundreds of genes would open doors for the discovery of several essential genes in parasites. These could be potential therapeutic targets.

Linear plasmids for cloning [A+T] rich and repetitive P. falciparum DNA

In order to modify the *Plasmodium* genome to manipulate expression of genes, the parasites must be transfected with plasmids carrying homology to the portions of the gene of interest (these could be the coding sequence of the gene or the untranslated regions (UTRs)). The genome of *P. falciparum* was sequenced in 2002 [5] and found to be extremely [A+T] rich, usually exceeding 80% in coding sequences and as high as 97% in some UTRs.

When [A+T]-rich sequences are present in the circular plasmids typically used to generate transgenic *P. falciparum* [6], they are frequently and unpredictably deleted and/or induce host plasmid rearrangements during propagation in *E. coli* [7]. To address this challenge, we selected the linear pJAZZ plasmid vector (Lucigen) as a chassis for all routine DNA assembly operations, as it has been used to successfully manipulate very large [A+T]-rich genomic fragments, including those derived from the rodent malarial model system, *P. berghei* [8]. We reasoned that this context would permit rapid and modular DNA assembly operations to be completed with high fidelity, and significantly enable construction of complex genome-modifying vectors either through serial or parallel assembly of multiple DNA fragments in single-pot reactions.

The pJAZZ plasmid was specifically designed for cloning unstable and repetitive DNA sequences. It was shown that these plasmids could maintain up to 30kb of *P. falciparum* DNA cloned into them [7].

The pJAZZ-OC plasmid

The BigEasy v2.0 Linear Cloning Kit (pJAZZ-OC-NotI vector) from Lucigen was used as the chassis for all the work described in this document. This vector is illustrated in figure 2 and was

obtained in the NotI digested form.

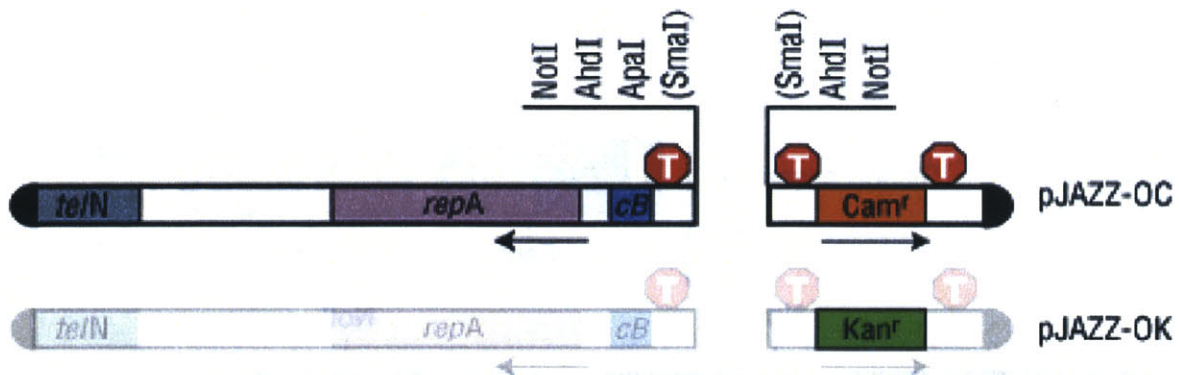


Figure 2: pJAZZ-OC from lucigen. Copied from the MA033-BigEasy-v20-Linear-Cloning-Kit manual, page 5

Converting traditional plasmids to linear plasmids

To evaluate the ability of the pJAZZ system to maintain *Plasmodium* DNA, the pJAZZ plasmid was modified to contain homologous regions to the two plasmids that were used to evaluate the TetR-aptamer system as well as the modified TetR-DOZI-aptamer system. A PCR-free transfer of these pre-assembled fragments was employed to eliminate the risk of introducing functionally deleterious mutations to the cassettes expressing the regulatory components as well as the firefly luciferase reporter. To facilitate this, a gene block (gblock) was synthesized containing regions homologous to the fragments we wished to transfer using the Gibson assembly method [24] and pre-installed this into the parental pJAZZ linear vector to create [pSwing] (Figure 3).

The pSwing plasmid

The synthesized gblock was cloned into the NotI digested pJAZZ as illustrated in figure 3.

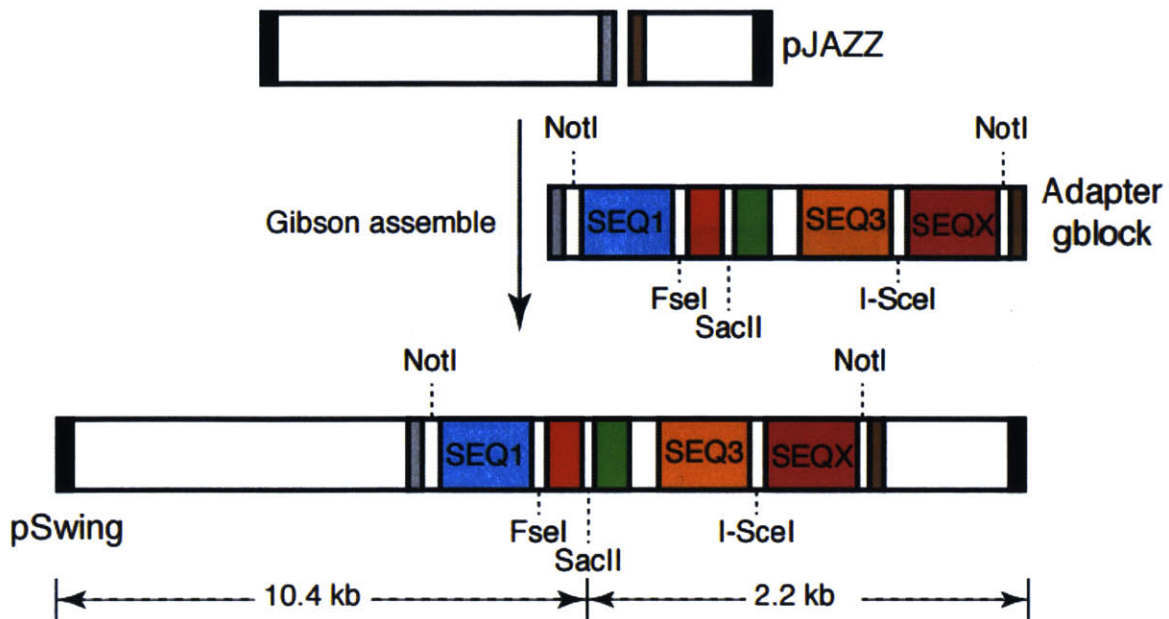


Figure 3. Creating the pSwing plasmid. NotI digested pJAZZ-OC was modified with an adapter gblock that carried homology to the two plasmids that were migrated from typical plasmids into the linear context.

Cloning of the gblock into PJAZZ-OC was 100% successful (figure 4C). The gblock contained SEQ1 and SEQ3, 45bp DNA sequences required for circularizing our linear plasmids as described later. The red and green coded DNA pieces contain homology to the two circular plasmids that were converted to the linear context.

After constructing pSwing, the next step was to examine the feasibility of rapidly transferring large, pre-existing DNA fragments into this linear plasmid format. This option is desirable as it allows directly adapting already functionally validated components to this new framework, thus facilitating straightforward transition to this new platform while preserving full access to all user-preferred features. To validate this concept, two plasmids were migrated to this linear context. These were two functionally validated versions of our previously described TetR-

aptamer regulatory systems that we intended to hardwire into the linear vector chassis to serve as integrated components in all our future designs

Converting pSN372 to linear context

pSN372 is a circular plasmid with two *P. falciparum* expression cassettes. Here, Cassette 1 encodes TetR, Renilla luciferase and blasticidin deaminase as a multi-cistronic message using the viral 2A peptide [6], and Cassette 2 encodes firefly luciferase regulated by a single TetR aptamer in its 5'UTR [2]. A 7.3 kb fragment consisting of the two head-to-head *P. falciparum* cassettes were transferred. The pSN372 plasmid was digested with NotI and ScaI and mixed with SacII digested pSwing in a single pot Gibson reaction (figure 4A). Six bacterial colonies were screened by restriction enzyme mapping, and all revealed a digestion pattern entirely consistent with correct assembly of the intended plasmid (figure 4C). The resulting plasmid was called pSN372L.

Converting JCN1847 to linear context

JCN1847 is similar to pSN372, except that the TetR component in Cassette 1 is replaced by a TetR-DOZI fusion protein, and an additional array of ten tandem TetR aptamers are included just upstream of the 3'UTR in Cassette 2, as previously described[9]. In this case, a 9.5kb fragment from JCN1847 was migrated to pSwing to create pSN1847L similarly to pSN372 (figure 4A). Again, all 6 bacterial colonies that were screened by restriction digest contained the migrated JCN1847 cargo. The resultant plasmid were labelled as pSN1847L.

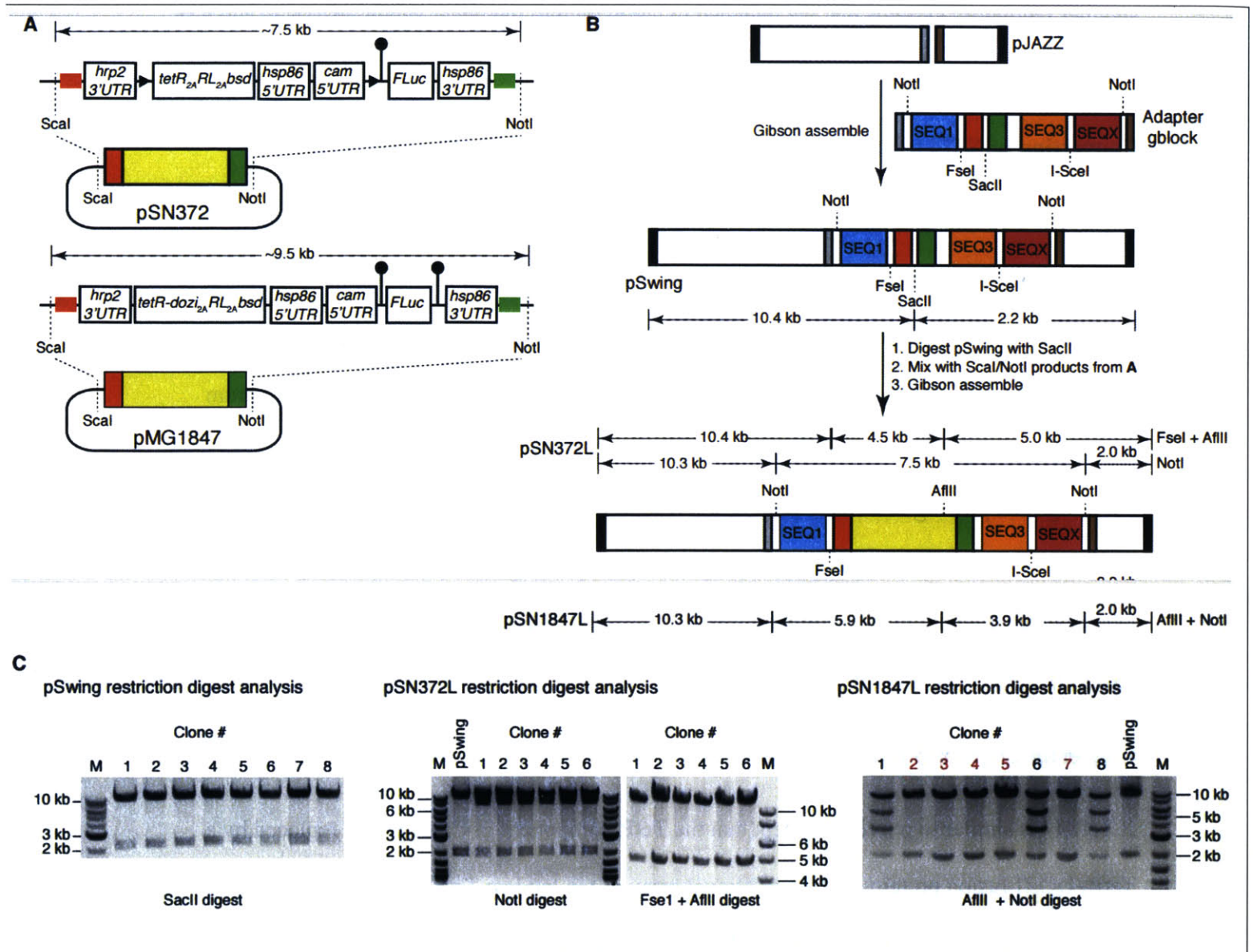


Figure 4: Converting traditional plasmids to the linear context. A) Digestion of pSN372 and pMG1847 with NotI and Scal releases the 7.3kb and 9.5kb respectively of the *P. falciparum* cassettes encoding the TetR regulatory system and Fluc, the reporter gene. B) Digestion of pSwing with SacII exposes the homologous DNA elements required for PCR free transfer of cassettes from digests in figure A. pSN372L and pSN1847L are created from the single pot Gibson reaction of the two digests. C) Screening bacterial clones by restriction digests. Left- pSwing colonies cut with SacII shows the the introduction of the gblock into the original pJazz plasmid. Middle- NotI digest of Psn372L colonies shows the release of the 7.3kb fragment that was migrated into pSwing. FseI+AfIII digest confirms that intact regulatory cassette. Right- NotI+AfIII digest shows that colonies 1,6 and 8 contain the migrated 1847 fragments.

As previously described, *P. falciparum* is not efficiently transformed using linear DNA [16]. Thus, the linear vectors generated above must be converted into circular plasmids to achieve successful

transformation of *P. falciparum*. While it seems counterintuitive to first clone into a linear vector context only to then re-circularize these for use in the parasite, several factors suggested to us that this would strategy would offer both flexibility and overall scalable efficiency.

First, linear vectors have proven utility in scalable cloning of large, [A+T]-rich sequences typical of those encountered in *P. falciparum* with high efficiency [7]. Thus, we reasoned that by selecting this framework for the vector construction phase requiring assembly of diverse components in various architectures would permit capitalization on these key proven aspects of cloning into linear vectors.

Second, we reasoned that cloning into traditional *P. falciparum* plasmid vectors is limited by both the sequence composition and overall fragment sizes that can be accommodated. Plasmid supercoiling induces single strand regions preferentially within [A+T]-rich regions, which are susceptible to nicking that can promote deleterious rearrangements that reduce supercoiling-induced torsional stress [10]. Thus, rather than rescuing linear vectors into the circular vectors typically used with *P. falciparum*, we decided to pursue strategies for rescuing these into larger bacterial artificial chromosome (BAC) vectors where supercoiling-induced torsional stress and plasmid instability are expected to be lower.

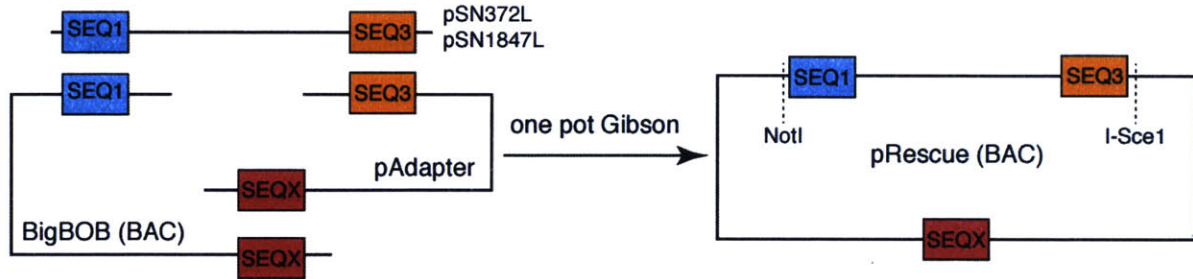
Third, a recent study by Guye *et al* demonstrated that linear vectors can be efficiently rescued into BACs in a single, standardized Gibson assembly process. Thus, in principle, libraries of linear vectors, once assembled, can be all simultaneously rescued in parallel using a common set of reagents [11].

In fact, later on in the project we discovered that even though the linear plasmids we generated could not be maintained as episomes by parasites for several generations and could not be integrated into loci by single crossover events or the use of the Bxb1 integrase system [12], they could be used as donor plasmids for CRISPR/Cas9 or zinc finger nuclease editing. This will be discussed later in the document. The ability to rescue the linear plasmids into BACs is still very useful however, for all other purposes including creation of expression vectors.

Rescuing linear plasmids into circular plasmids

Based on this rationale, we proceeded to rescue pSN372L and pSN1847L into BACs using the Guye *et al* approach (figure 5a). We used the pAdapter and pBigBOB plasmids as previously described[11], released the relevant fragments encoding components to be expressed in *P. falciparum* from either pSN372L or pSN1847L by NotI/I-SceI digestion, and assembled these in a single pot, 3-piece Gibson reaction and recovered BACs after transforming *E. coli* selected for chloramphenicol and kanamycin resistance. We isolated and analyzed 6 and 6 BACs corresponding to rescued pSN372L and pSN1847L, respectively (designated as pSN372R and pSN1847R from hereon). In 5/6 and 6/6 instances, the region derived from the linear vector had been successfully rescued into pSN372R and pSN1847R, respectively. Restriction enzyme mapping confirmed BACs with globally intact topology in 5/6 and 6/6 instances, respectively, for the pSN372R and pSN1847R analyzed clones (figure 5b and c). Altogether, these data indicate that the linear vector-encoded sequence information critical for *P. falciparum* applications can be successfully and efficiently rescued into BACs using the described framework, and furthermore, that this process can be achieved with high overall preservation of the expected BAC topology.

A



B

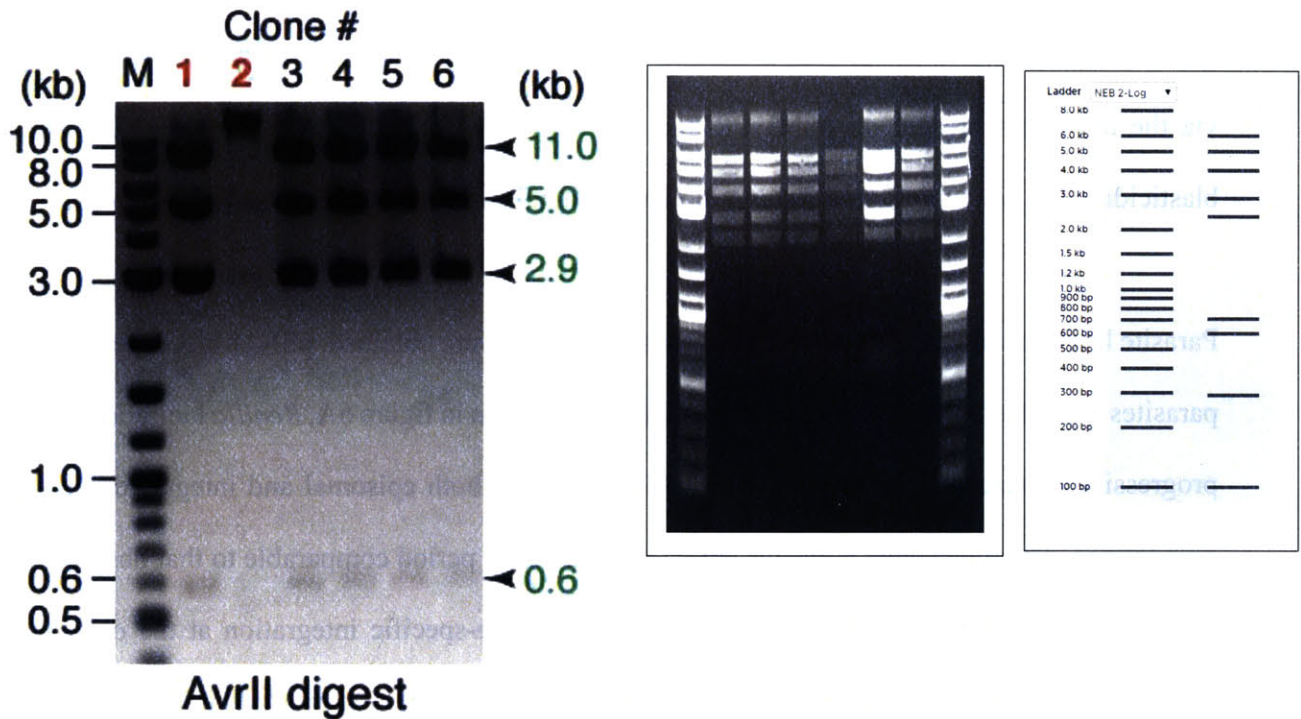


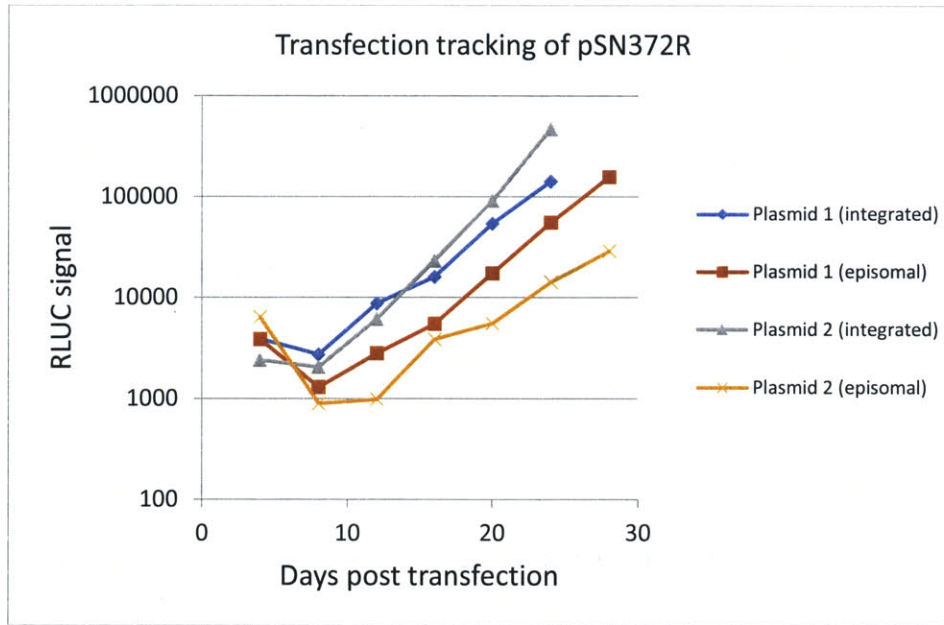
Figure 5: Rescue of linear plasmid into BACs. A) The strategy from Guye et al. The linear plasmids were digested with NotI and I-SceI to expose SEQ1 and SEQ3. BIGBOB was digested with PacI to expose SEQ1 and SEQX. pAdapter was digested with XbaI and XhoI to expose SEQ3 and SEQX. In a single pot Gibson reaction, the three pieces circularize. B) Clones 3,4,5, and 6 of rescued PSN372 showed correct topology after AvrII digest. C) FseI+XbaI+AvrII digest of 6 clones of PSN1847R showed that all 6 clones had the correct topology. A virtual map of the digest from the online vector tool Benchling is presented on the right panel

Rescued Plasmids can be successfully transfected into *P. falciparum* parasites

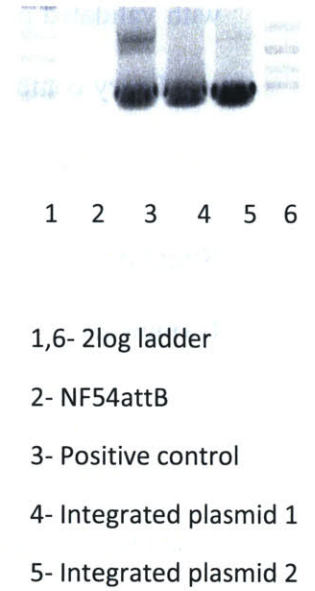
After successfully generating the pSN372R and pSN1847R BACs, it was important to know whether these would: (1) successfully transform *P. falciparum*, both in episomal and chromosomally integrated contexts; and (2) yield functional outcomes similar to those obtained using traditional *P. falciparum* expression vectors. To assess these, the commonly used NF54^{antB} *P. falciparum* parasites were transfected with either pSN372R or pSN1847R alone (episomal transfection) or together with the pINT plasmid encoding the *Mycobacterium* spp *Bxb1* integrase to promote site-specific integration at the *attB* site installed at the *cg6* locus in the NF54^{antB} line [12] via the *attP* sites present on pSN372R and pSN1847. The parasite lines were selected using blasticidin, and were monitored *Renilla* luciferase signal over the course of the transfections.

Parasite lines containing episomal or integrated pSN372R were obtained. pSN1847R is yet to yield parasites after the first transfection attempts failed. As shown in figure 6A, *Renilla* luciferase signal progressively increased over time for pSN372R, to yield both episomal and integrated parasite populations. The transfected parasites emerged over a time period comparable to that observed in transfections using typical *P. falciparum* vectors [6]. Site-specific integration at the *cg6* locus could be confirmed by PCR (figure 6B) and sequencing of the resulting products in the case where pSN372R were co-transfected with pINT. Altogether, these data indicate that the BACs generated through this process can successfully transform *P. falciparum* and that the cassettes mediating *Renilla* luciferase and blasticidin deaminase expression are both intact.

A



B



C

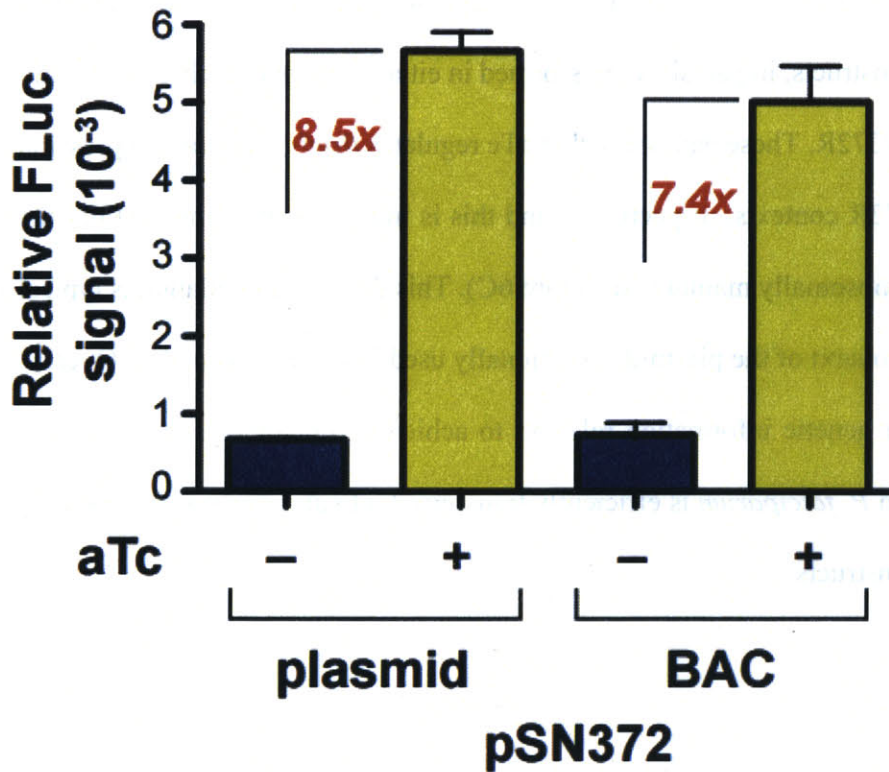


Figure 6. Rescued BAC plasmids can be transfected into *P. falciparum* parasites. A) Renilla luciferase assays track episomal and integrated plasmid transfections. Parasites are visible in

culture after 24 days just like typical plasmids. B) Integration of plasmids is confirmed by PCR with validated primers. The positive control is a parasite strain with integrated pSN372. C) The regulatory components function equally well on the BAC compared to the typical plasmid. Fold regulation of FLUC expression is not statistically significantly different.

Regulatory components for controlling gene expression function as expected within the BAC framework

Next, it was critical to evaluate if the regulated gene expression mediated by the TetR and TetR–DOZI regulatory components hardwired into our vector framework could still be achieved in this BAC context. Since only the pSN72R transfection was achieved, the experiment was performed for only this strain. Anhydrotetracycline (aTc)-regulated expression of the *firefly* luciferase reporter gene encoded by Psn372 constructs, in parasites transformed in either episomal or chromosomally integrated fashion with pSN372R. These data show that aTc regulates *firefly* luciferase expression by ~ 10- fold in the pSN372R contexts, respectively, and this is independent of the whether the BAC is episomally or chromosomally maintained (figure 6C). This degree of regulation is typical of what we observe in the context of the plasmids traditionally used in *P. falciparum* [2]. Overall, these data demonstrate that genetic information relevant to achieving outcomes related to gene expression and its control in *P. falciparum* is efficiently transferred into and remains functionally intact within these BAC constructs.

Creating modular linear vectors configured to deliver pre-programmed functionality

Having established proof-of-concept for a vector assembly pipeline that involves building relevant constructs in linear vectors and successfully rescuing them into BACs that function in *P. falciparum*, a mini-library of linear vector architectures was created. These vectors were pre-programmed to deliver the diverse outcomes routinely desired during functional genetics studies in the parasite. Four of these vectors were created and enable any form of genome editing and gene regulation practically desirable. These are described in the rest of the document. In the case of each vector, the architecture of the vector is explained and the applications are illustrated with one or more endogenous *Plasmodium falciparum* genes. The four vectors to be described are pSN150, pSN053, pSN054 and pSN154. All of these vectors share common features especially when used as donor vectors for genome editing. These features are;

1. A fixed regulatory protein, transfection tracking component and drug selection component. This cassette contains Renilla luciferase (RLUC), blasticidin resistance gene (BSD) and either TetR only or TetR-DOZI. These components have all been previously described
2. A cassette for generation of sgRNA for CRISPR/Cas9 genome editing. This cassette was previously described [4]. It contains a T7 promoter, the tracrRNA component of the sgRNA and a T7 terminator. Between the T7 promoter and tracrRNA is an Ippo-I meganuclease restriction site for cloning in the crRNA portion of the sgRNA. It is imperative to add a GG nucleotide sequence between the T7 promoter and the crRNA

being cloned as the T7 TRAP requires the GG nucleotides to initiate transcription of the sgRNA.

3. A modular Left Homologous Region (LHR) cloning site for cloning in a double strand break repair DNA element
4. A modular Right Homologous Region (RHR) cloning site for cloning in a double strand break repair DNA element and
5. Modularized affinity tags for tagging genes of interest. Tagging can be achieved in the N-terminus or C-terminus of the protein of interest depending on the strategy being used.
The tags used here are all single copy HA, FLAG or Myc tags.

These four vectors are discussed next.

pSN150

Potential applications-

1. Scaling implementation of 5'-aptamer regulated gene expression
2. Dimerizable Cre recombinase mediated deletion of target genes to study gene function and establish essentiality. This is a conditional deletion system encoded into the vector
3. Straight knockout of a non-essential gene

Architecture of pSN150

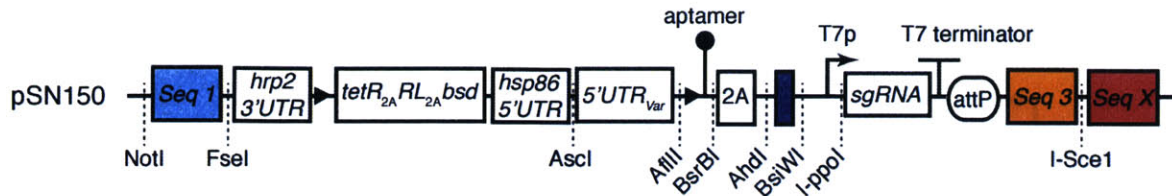


Figure 7: Architecture of pSN150. Seq1 and seq3 were previously described to be required for rescue of linear plasmids into BACs. LHR and RHR cloning sites are shown (these use FseI and AhdI/XmaI/SmaI cut sites respectively). The TetR-RLUC-BSD cassette is hardwired into the vector. The CAM 5'UTR drives the expression of the gene of interest after editing. This promoter is modular (Ascl site on the left and AflII site on the right). The sgRNA cassette is also present. The modified gene can only be tagged at the N-terminus with an HA tag if this is desirable. The HA tag is shown in purple and can be removed during cloning.

pSN150 was created by replacing the firefly luciferase gene in pSN372L with a region encoding T2A, an HA epitope tag and several unique restriction sites immediately downstream of the TetR aptamer. This architecture is designed to install a 5' aptamer to achieve TetR-dependent regulation of the target gene while concurrently swapping the native for another user-specified promoter (figures 7&8). However, a clean promoter swap with or without TetR-dependent regulation can also be achieved, as the region including the 5'UTR through to the aptamer is modular. This is illustrated in figures 7 and 8.

Regulated target gene with option to knock it out (Design 2)

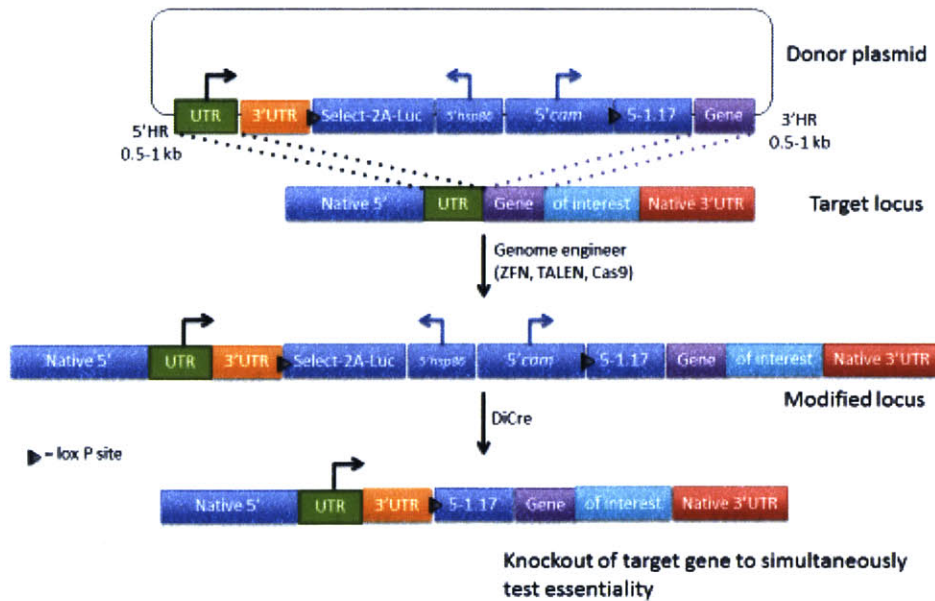


Figure 8 Regulating target genes with 5' aptamers. Conceptual native and edited loci are shown. The use of DiCre to create a conditional KO is also shown

The T2A sequence is included to prevent translational fusion of a peptide fragment arising due to initiation of translation within the aptamer element to the downstream protein, though our previous data indicate that this should not interfere with signal peptide sequences that target proteins to the secretory pathway [2]. For non-trafficked proteins, an N-terminal HA epitope can be automatically installed by cloning into the AhdI/XmaI/SmaI sites. It is critical to ensure that the coding sequence of the gene of interest is still in frame after the HA tag. As this region is modular, however, this feature can easily be excluded for proteins where the native N-terminus is required for proper subcellular trafficking. The regions of interest can be cloned into the ApaI/BsrBI sites between the 5' aptamer and the 2A tag instead.

How to use pSN150

As a 5' gene regulation donor plasmids

To facilitate modification of a target locus using this architecture, we have enabled double crossover recombination and homology directed repair of either ZFN- or CRISPR/Cas9-induced double strand DNA breaks [13, 4]. This can be achieved by cloning the relevant upstream and downstream homologous regions at the FseI and AhdI/XmaI/SmaI sites, respectively. To facilitate CRISPR/Cas9-mediated editing, a T7 promoter-driven sgRNA expression cassette is provided on the vector as described earlier.

As a straight knockout vector

This vector is also configured such that it can easily function as a knockout vector. This can be achieved by cloning the 5' UTR of the gene as LHR into the FseI site and the 3'UTR of the GOI into the AscI/ BsrBI/ApaI/AhdI/XmaI/SmaI sites and lastly, providing an sgRNA. This knockout strategy is not illustrated in this document with any *P. falciparum* genes.

DiCre-mediated deletion of target genes to study gene function and establish essentiality

Note: Data on this application of the vector is not provided here as this is still under optimization. Details of some of the optimization processes are presented here.

The pSN150 plasmid system was established to enable the regulation of genes from 5' aptamers. Because 5' aptamers only allow for ~80% knockdown in gene expression, it was necessary to incorporate other mechanisms to allow deletion of the gene of interest in order to get rid of the 20% protein still available after knockdown. To this end, a conditional gene deletion strategy

was investigated using the inducible Dimerizable Cre recombinase system previously used in *P. falciparum* [14]. This strategy was illustrated in figure 8.

In order to enable DiCre mediated conditional deletions, the pSN150 vector contains two loxP sites (blue triangles in figure 7). One loxP is between the stop codon of the BSD component of the vector and the hsp86 3'UTR used for that cassette. The second loxP is upstream of the 5' aptamer. LoxP sites are predicted to have a hair pin secondary structure [15]. The interaction of this second hairpin with the aptamer, which also folds into a hairpin, is unknown. In addition, hairpins when positioned as leaders could decrease expression of the downstream gene [15]. The effects of the loxP site on translation of ORFs and on the TetR-aptamer regulation were tested. The two loxP sites described were cloned into the AgeI and AflII sites of pSG372.5 [2] to create pSN372 which was described earlier. pSG372.5 and pSN372 (both traditional circular plasmids) were integrated into the attB site of NF54attB parasites as previously described [12]. The effects of this loxP on expression and regulation of the firefly luciferase reporter was then assayed. As shown in figure 9, the loxP had minimal effect on the function of the 5' aptamer (regulation) and the levels of firefly luciferase protein translated when inducer molecule was present (maximal expression). This was true for ring, trophozoite and schizont stages of parasites.

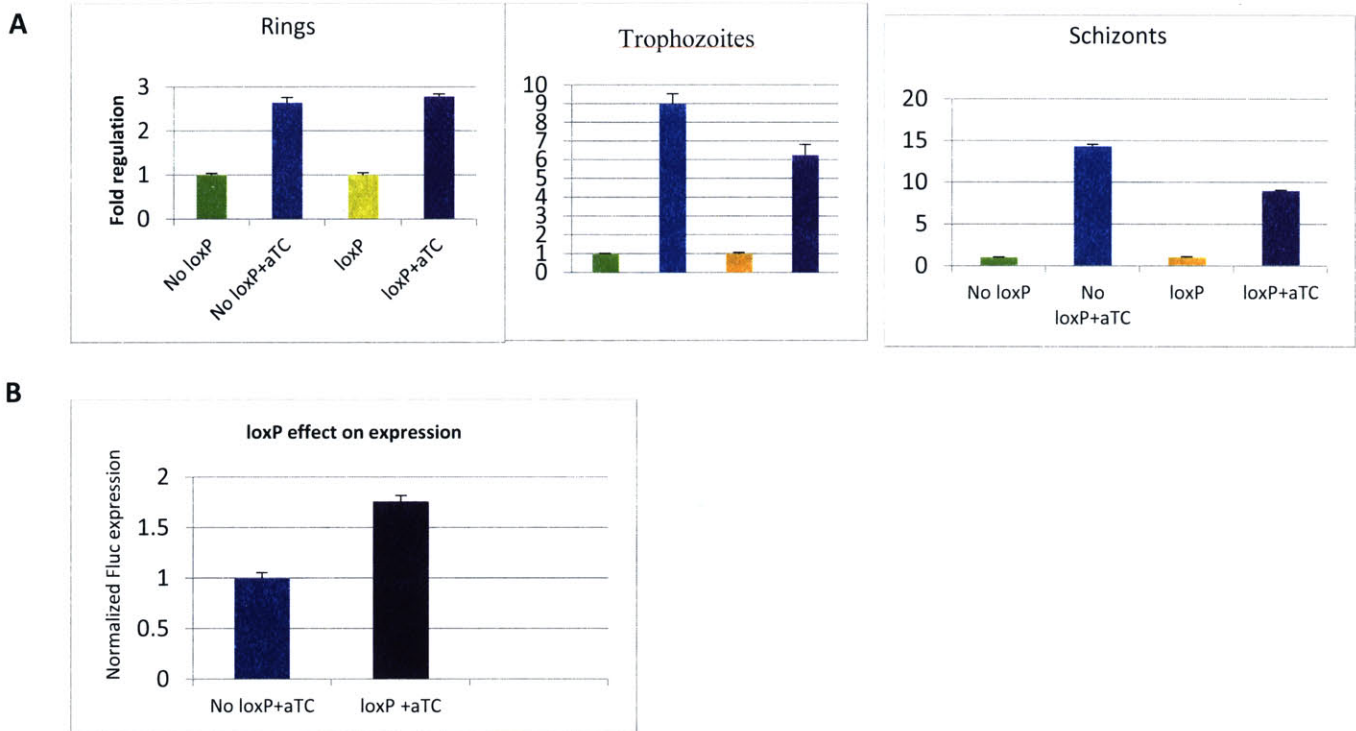


Figure 9: Effect of placing loxP upstream of 5' aptamer on A) TetR-aptamer regulation of firefly luciferase in all stages of IDC are minimally affected by the presence of loxP and B) maximal Fluc expression upon induction is unaffected by the loxP site

At this point, all components of pSN150 have been validated. To evaluate the efficiency, reliability and scalability of pSN150 as donor vector, 5 endogenous *P. falciparum* genes were picked to be edited for 5' aptamer regulation. These genes included glycogen synthase kinase-3 (PF3D7_0312400), hexose transporter (PF3D7_0204700), thioredoxin reductase (PF3D7_0923800.1), choline kinase (PF3D7_1401800) and the chloroquine resistance transporter (PF3D7_0709000).

The general donor vector design strategy is shown in figure 10.

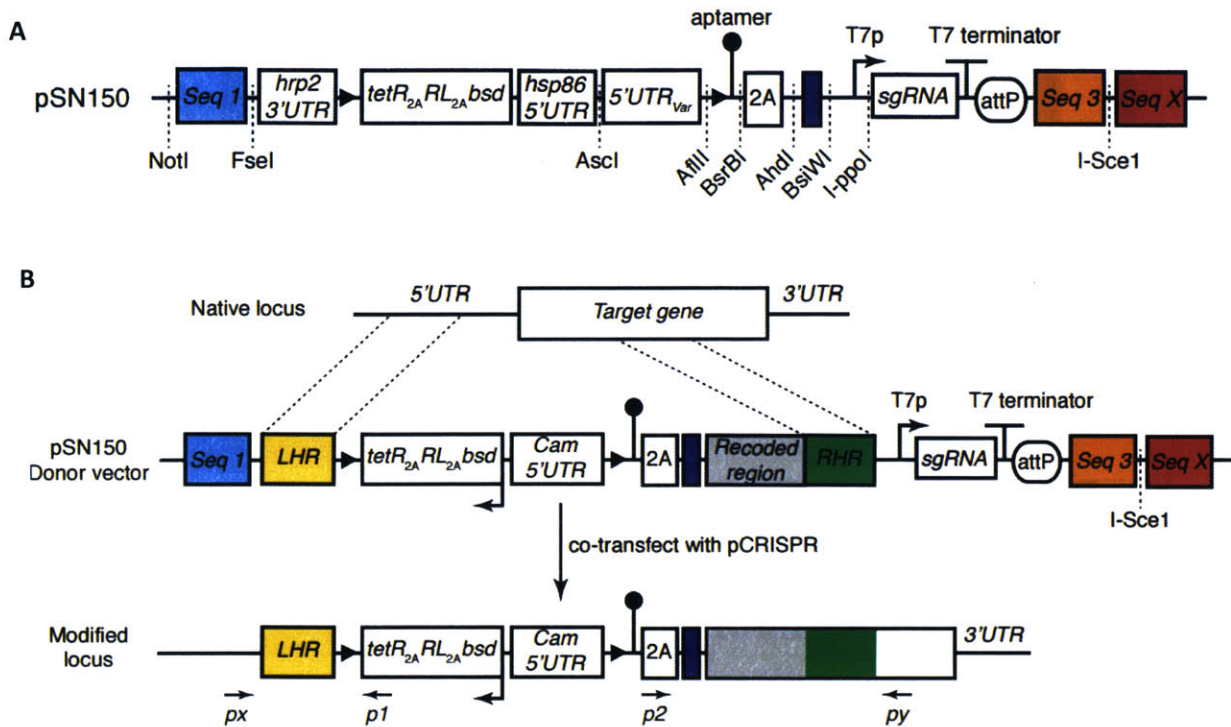


Figure 10: Donor vector architecture of 5' aptamer editing. A) Zoom in on the cloning sites of pSN150. B) A three step cloning process was used. First the LHR was cloned into the FseI site to create pSN150_LHR. Next a portion of the 5' end of the GOI was recoded around the sgRNA cut site. All recoding was done with *T. gondii* codon preferences and the synthetic DNA (gblock) was ordered from IDT. This gblock was mixed in a single reaction pot with the RHR and cloned into AhDI digested pSN150_LHR. This reaction generated pSN150_LHR_GB+RHR. Next the sgRNA was generated by a klenow reaction and cloned into I-PPOI digested pSN150_LHR_GB+RHR. This generated the final linear donor vector pSN150_LHR_GB+RHR_gRNA.

The donor vector construction was generally the same for all 5 genes in this case.

A three step cloning process was used (the gene of interest is labelled as x).

1. First the LHR was cloned into the FseI site to create pSN150_LHRx.
2. Next a portion of the 5' end of the GOI was recoded around the sgRNA cut site. All recoding was done with *Toxoplasma gondii* codon preferences and the synthetic DNA (gblock) was ordered from IDT. This gblock (GB) was mixed in a single reaction pot with the RHR and cloned into AhDI digested pSN150_LHRx. This reaction generated pSN150_LHRx_GBx+RHRx.
3. Finally the sgRNAx was generated by a klenow reaction and cloned into I-PPOI digested pSN150_LHRx_GBx+RHRx. This generated the final linear donor vector pSN150_LHRx_GBx+RHRx_gRNAx.
4. The final linear plasmid was rescued onto a BAC for transfection.

All rescued plasmids were co-transfected into NF54attB parasites with a plasmid carrying the CRISPR machinery (pCRISPR from Jeffrey Wagner). This pCRISPR plasmid confers resistance to WR). Parasites were selected with either BSD only (selection for just the donor plasmid) or BSD+WR. Atc is added to the media for the entire transfection since the essentiality of the genes investigated is unknown.

GSK-3 (PF3D7_0312400)

Glycogen synthase kinase 3 was of interest to us because of the debate in its classification. It has been deleted in *P. berghei* [19] and was also knocked out in *P. falciparum* in the Niles lab (unpublished data). However it was previously considered to be an essential gene and good therapeutic target [17,18].

Two vectors were constructed for this gene. The vectors contained the same LHR, RHR and recoded regions but differed in the sgRNAs that were used. These vectors were used to edit parasites using the co-transfection strategy previously described. As shown in figure 11, parasite populations obtained from the two donor vectors and with either drug selection tested positive for editing. The binding sites of primers p1,p2,px and py are shown in figure 10.

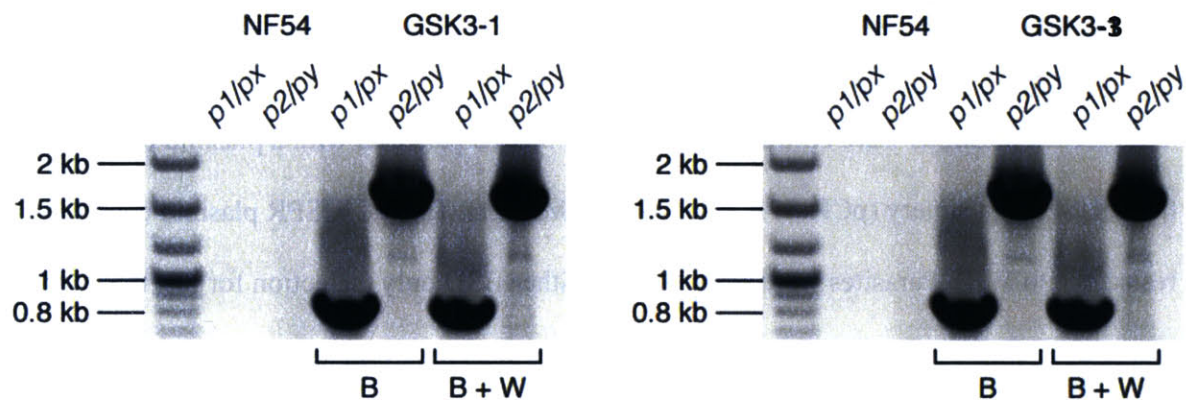
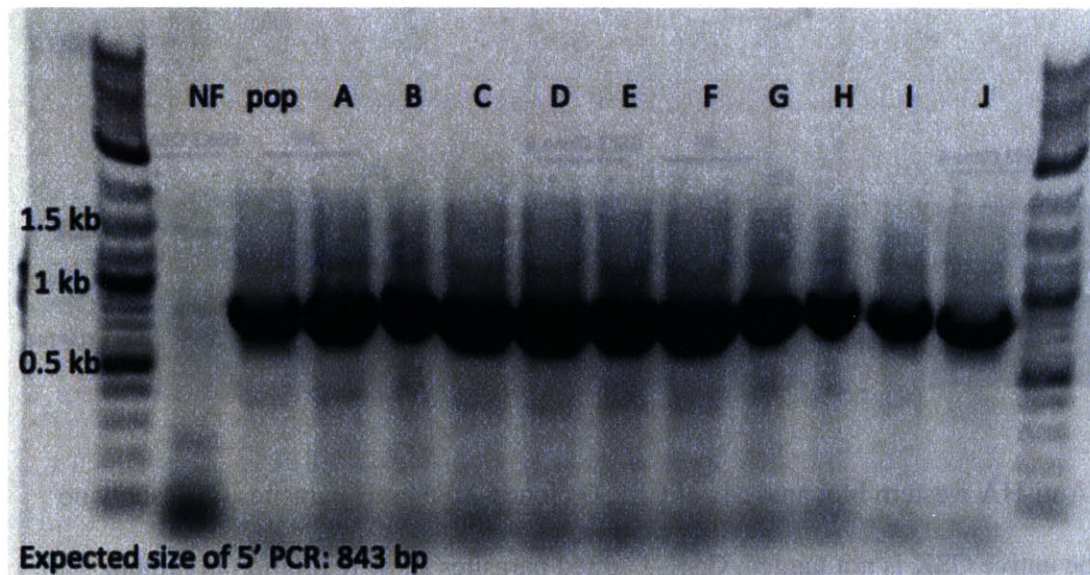


Figure 11. PCR confirmation of editing of the GSK-3 locus by CRISPR/Cas9. GSK-1 donor vector carries sgRNA1 while GSK-3 donor vector carries sgRNA3

The GSK-3 population was cloned and 10 clones were obtained. These were all positive, by PCR, for the desired editing event for both the 5' and 3' (figure 12). Three clones (A, B and C) were selected for western blot and growth assays in +/- aTc conditions. All clones showed regulation of the HA-tagged GSK-3 protein (figure 13). However, parasites with knockdown of GSK-3 protein still grew just as well as those expressing maximal levels of the protein (figure 14). The western blot shows availability of protein in the -aTc condition and this could be responsible for the unaffected fitness of parasites upon withdrawal of the inducer. A greater knockdown in protein levels is required for determining with certainty if this gene is essential or not. This is also a case where the conditional deletion of the gene with the DiCre system could be useful.



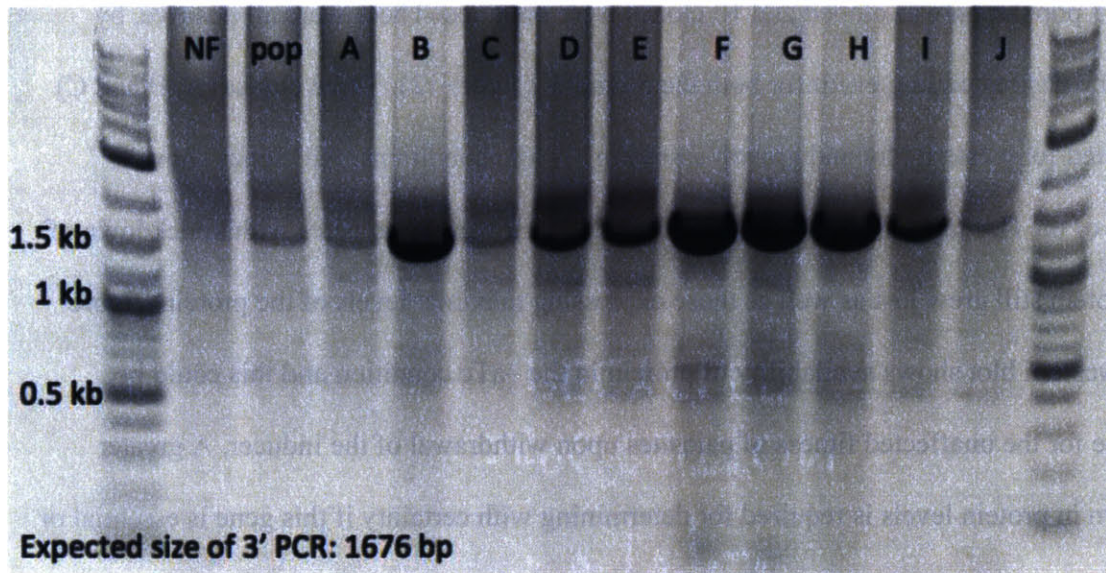


Figure 12. PCR testing of 10 clones from GSK-1 only. A) $p1+px$ confirm 5' editing in population and all clones but not in NF54attB parasites. B) $p2+py$ confirm 3' editing in population and all clones but not in NF54attB parasites

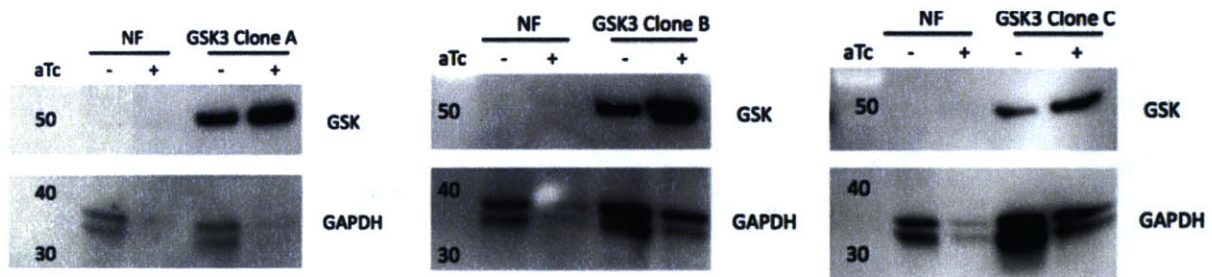


Figure 13. Anti-HA western blot analysis of GSK-3 clones. NF54attB parasites (NF) show no protein as the native GSK is not tagged. GSK-3 edited parasites show regulation in protein levels. GAPDH is used as loading control

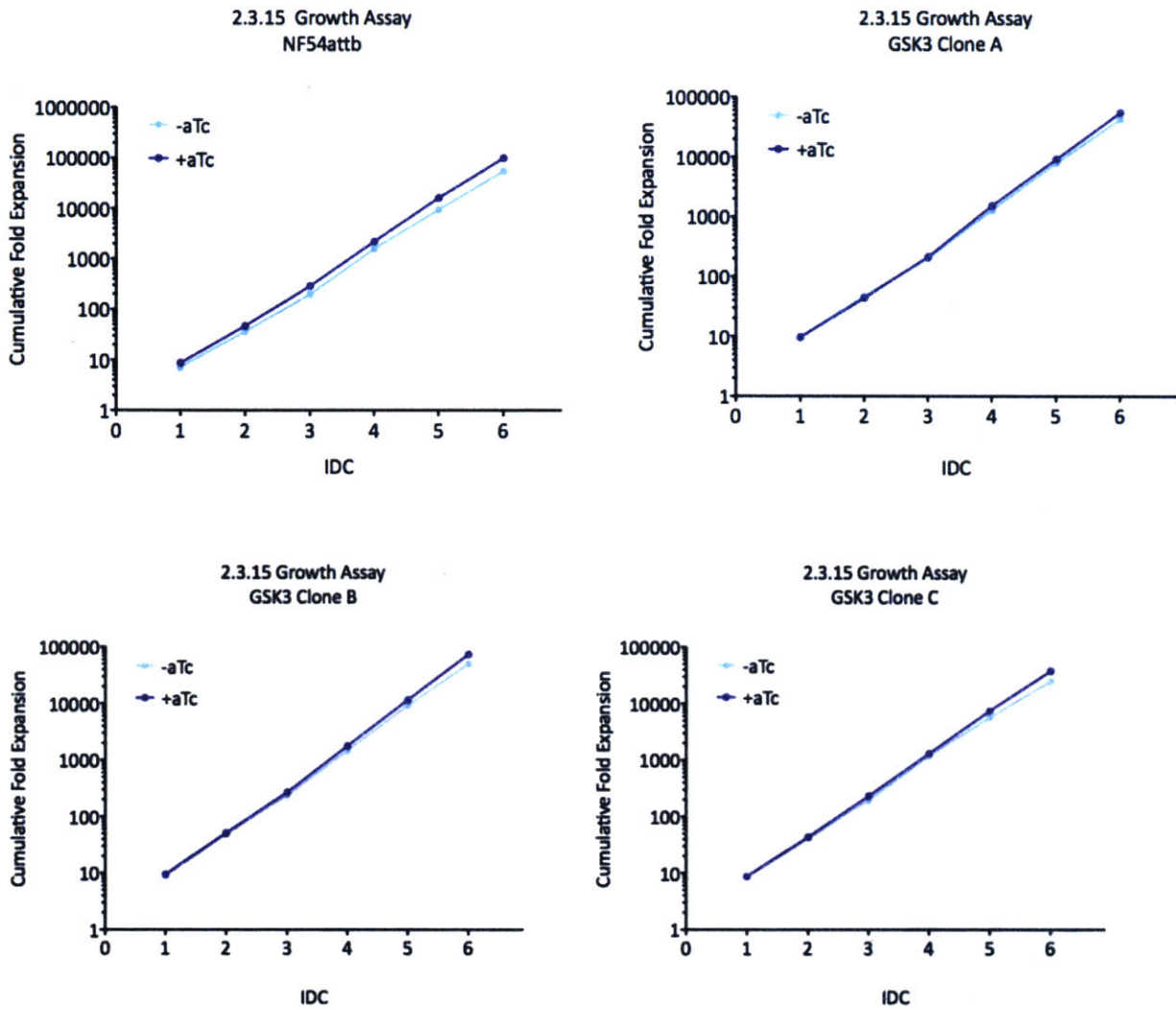


Figure 14. Growth Assay curves for NF54attB and GSK-3 edited clones. No growth defect is observed in GSK-3 edited clones even when aTc is withdrawn. NF54attB parasites are unaffected by aTc.

Hexose Transporter (HT) (PF3D7_0204700)

This gene could only be deleted from *P. falciparum* when a complimentary copy was transfected into parasites [20, 21]. It has also been touted as a therapeutic target for malaria [22]. Its essentiality has however, not been definitively shown.

A donor plasmid for 5' editing of HT was constructed similarly to the vector for GSK-3. This plasmid was co-transfected into NF54attB parasites with pCRISPR. Parasites were selected with BSD+WR. The parasite population obtained tested positive for editing and interestingly, showed a severe growth defect upon withdrawal of aTc (figure 15)

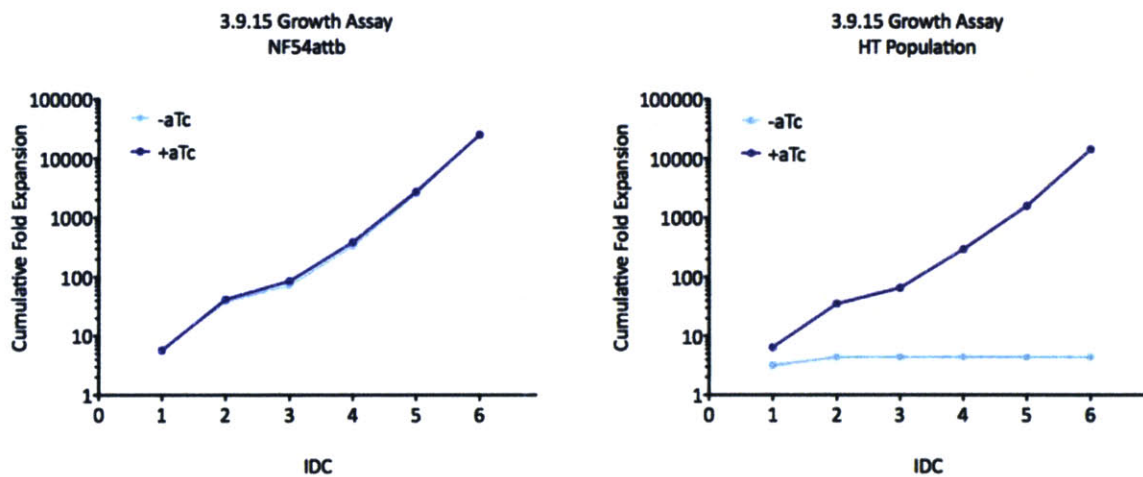
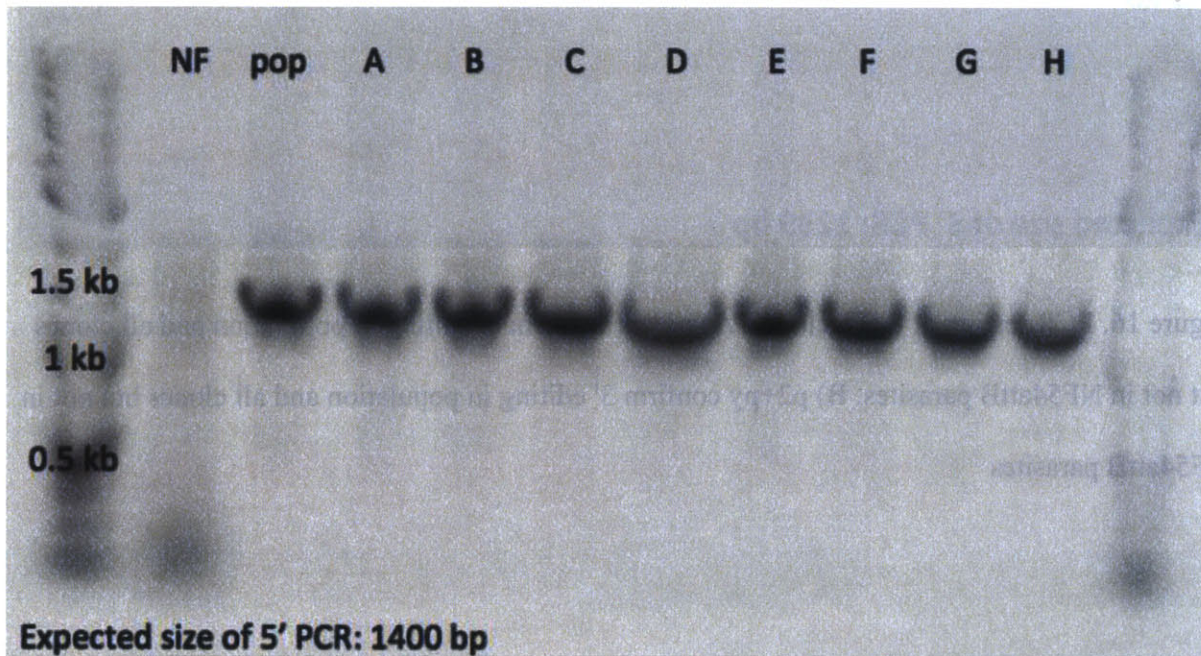


Figure 7 Growth assay shows severe growth phenotype in HT edited population when aTc is withdrawn. NF54 control grows normally

The edited HT population was then cloned out and all 8 clones tested positive for the desired editing event for both the 5' and 3' (figure 16). Four clones (A, B, C and D) were selected for western blot and growth assays in +/- aTc conditions. All clones showed regulation of the HA-tagged HT protein (figure 17). Just like the population, clonal parasites with knockdown of HT had severe growth defects after one IDC (figure 18).



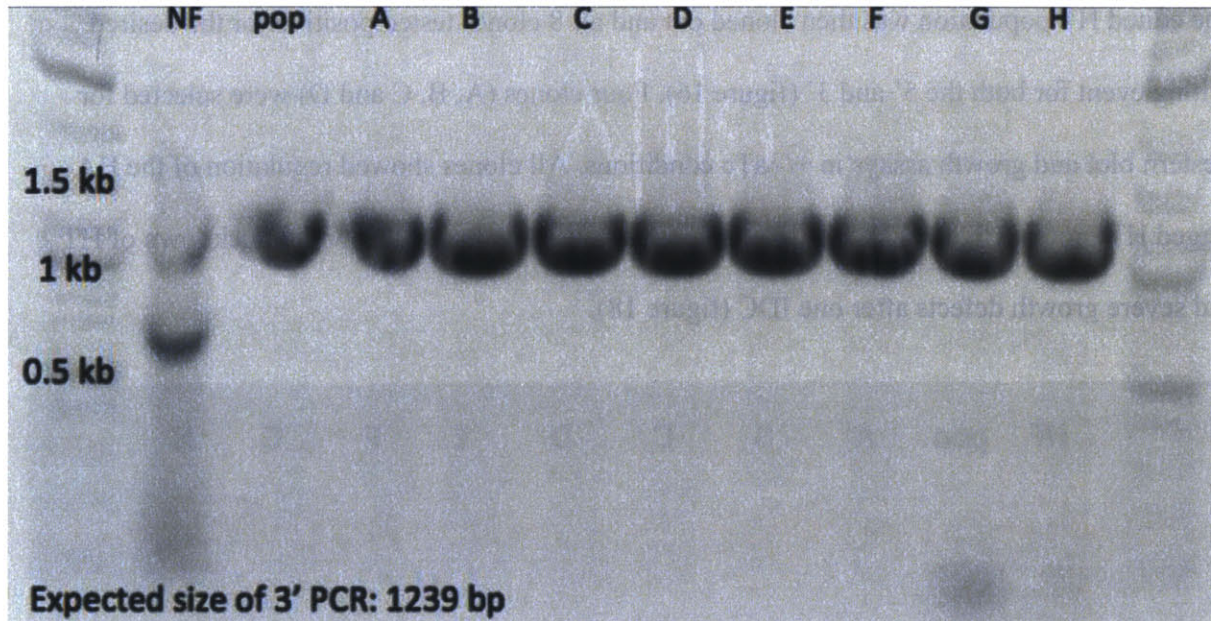


Figure 16. PCR testing of 8 HT clones. A) p1+px confirm 5' editing in population and all clones but not in NF54attB parasites. B) p2+py confirm 3' editing in population and all clones but not in NF54attB parasites

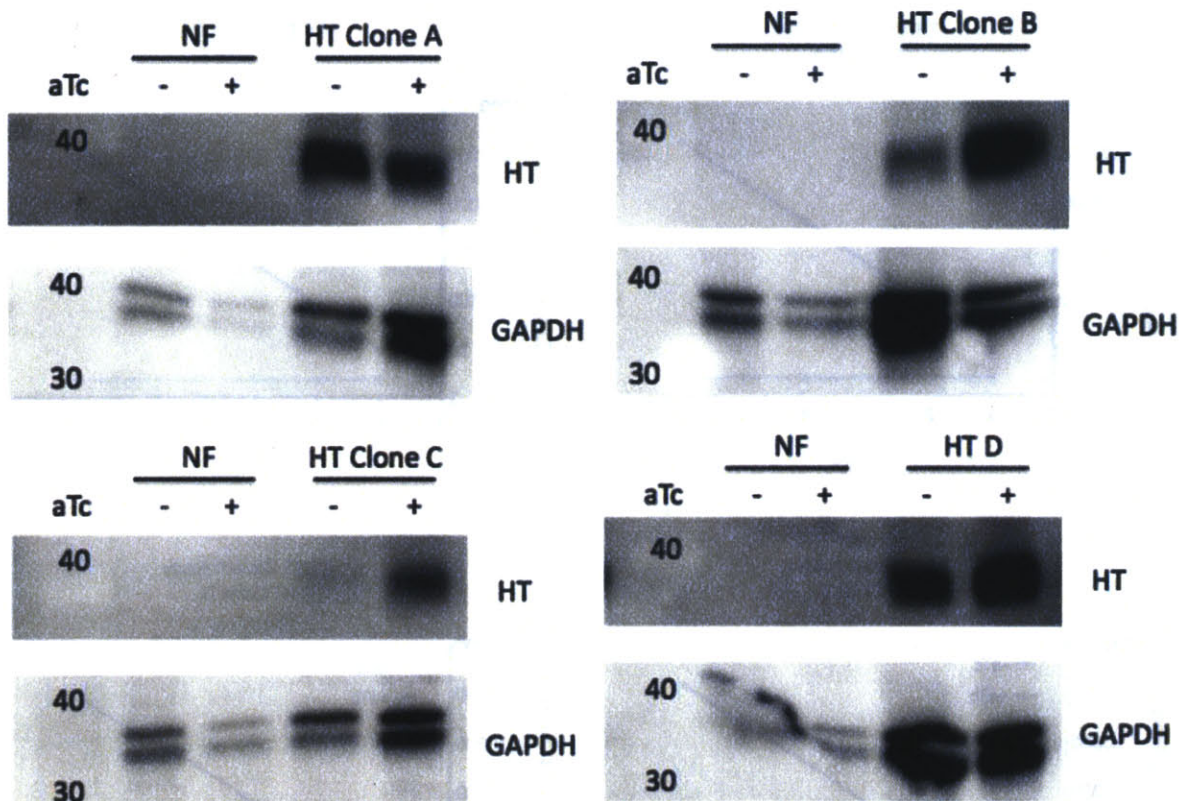


Figure 17. Anti-HA western blot analysis of HT clones. NF54attB parasites (NF) show no protein as the HT is not tagged in those parasites. HT edited parasites show regulation in protein levels

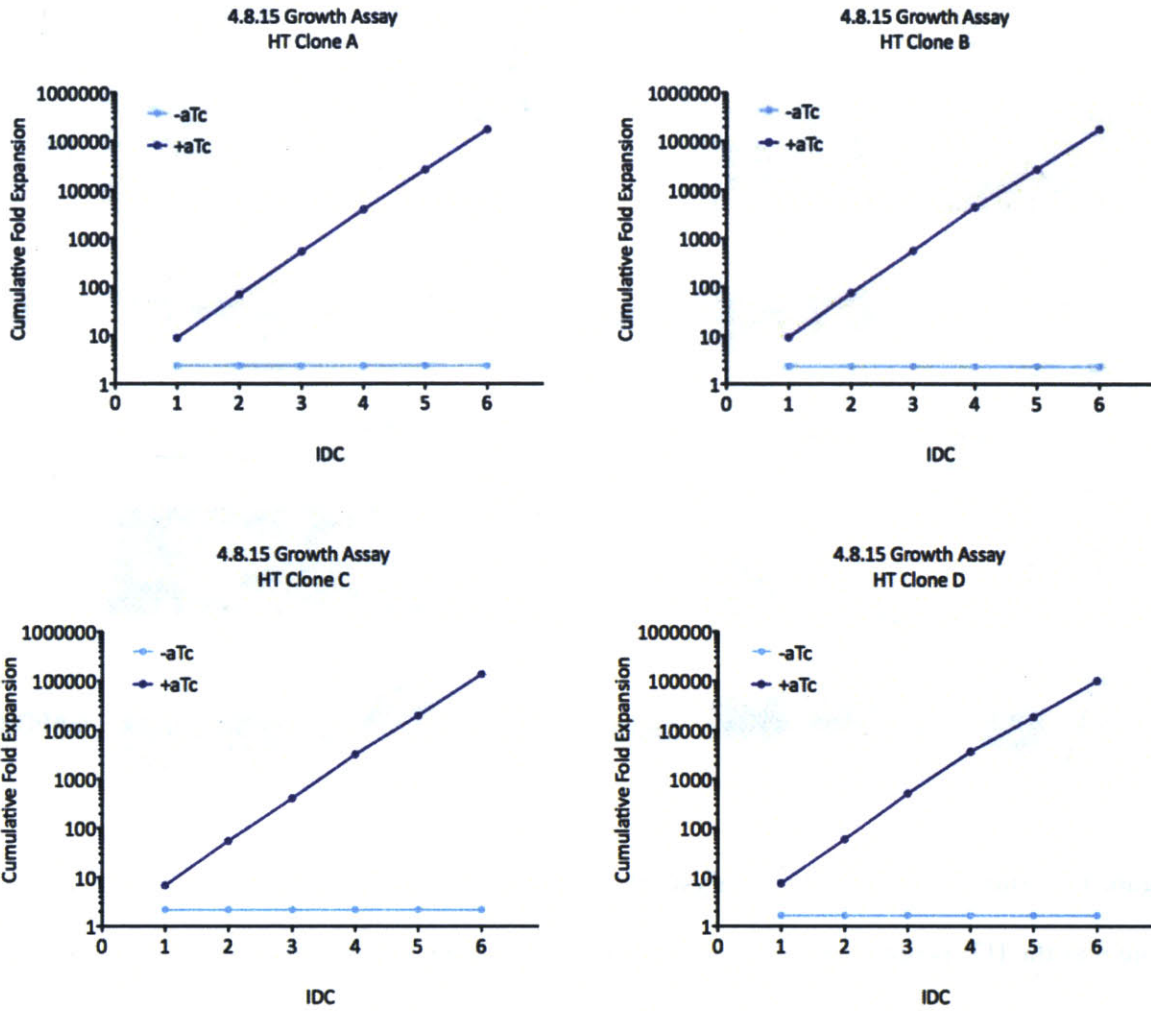


Figure 18. Growth assays confirm a severe growth defect in all HT clones starting from IDC1

Altogether, the data shown here conclusively show that HT is essential to the blood stage of *P. falciparum*.

Thioredoxin reductase (PF3D7_0923800.1)

This gene is putatively essential in parasites as it could only be knocked out when complementary copy was being expressed on an episome [23]. The TrxR donor vector also yielded a positive test for editing of the genome as shown in figure 19.

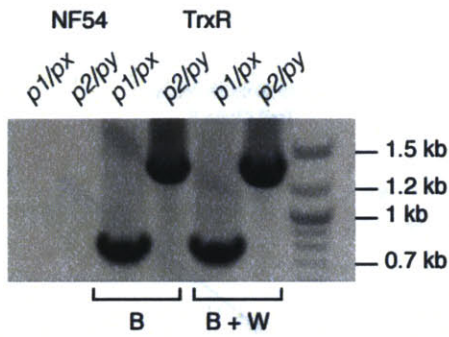


Figure 19. The population of TrxR is edited

Cloned parasites also tested positive for editing (figure 20)

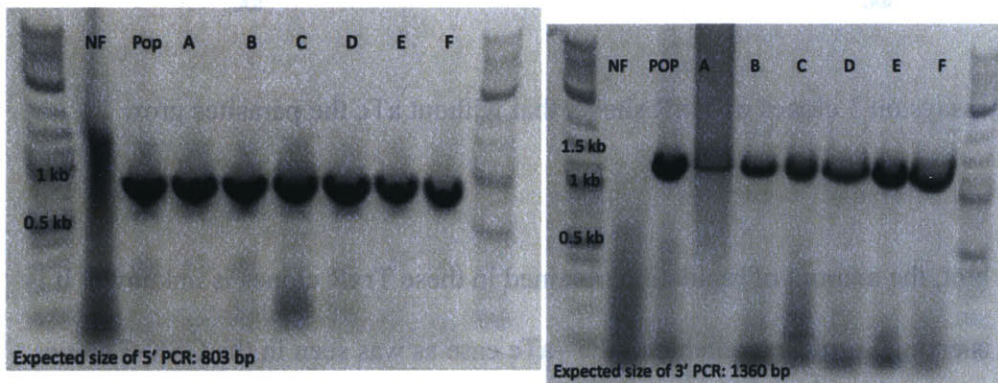


Figure 20. All clones of TrxR are edited in both 5' and 3' ends as shown by PCR with p1+px and p2+py respectively

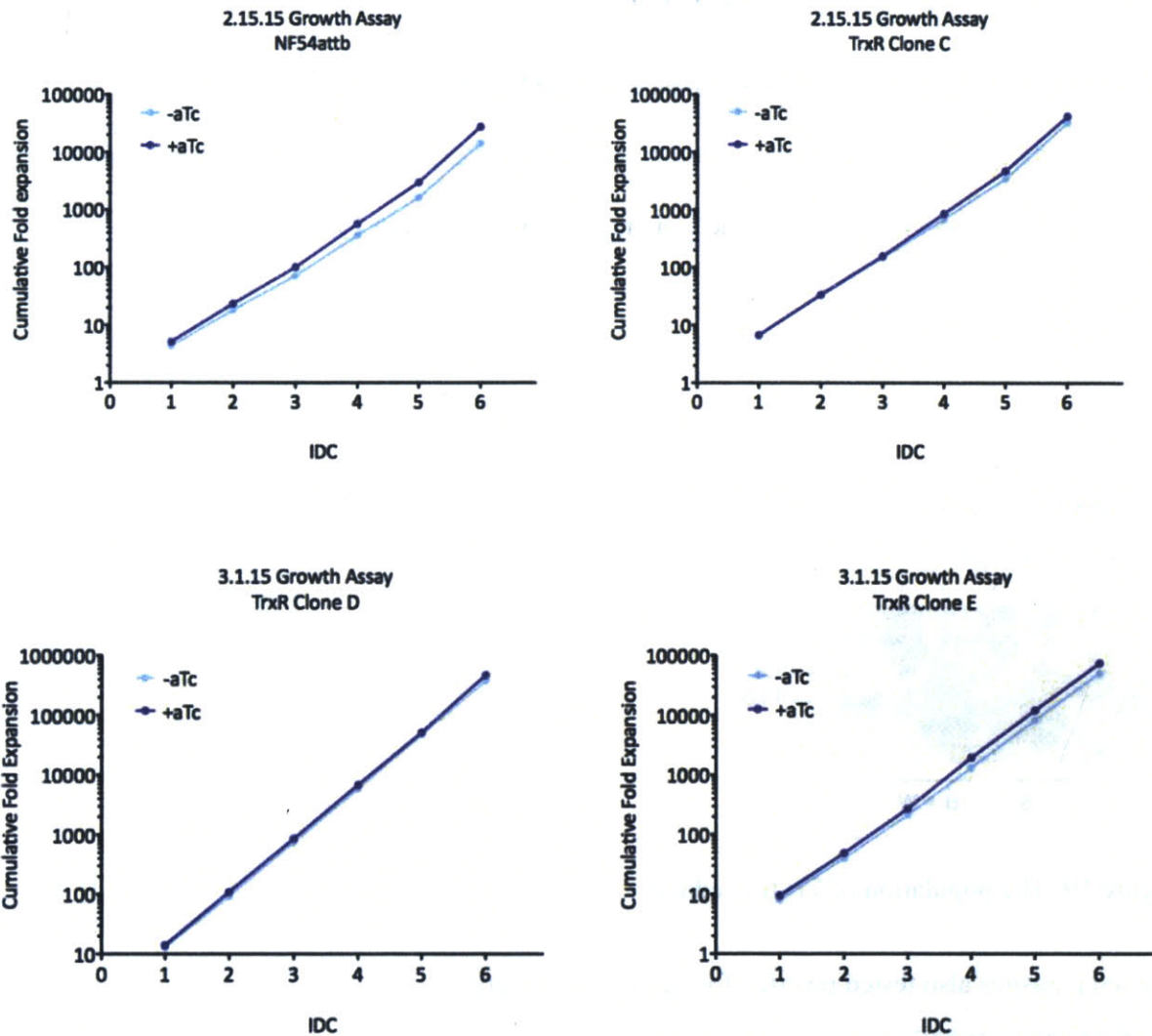


Figure 21. Growth assays on 3 clones of TrxR shows that without aTc the parasites grow normally

Without a western blot, the amount of regulation obtained in these TrxR clones is unknown. It is possible that there is enough protein available in the -aTc case as was seen in the GSK-3 clones. Since TrxR is an enzyme, it is reasonable to speculate that maybe a 20% background level of protein available in the -aTc condition can catalyze a very quick and highly evolved reaction.

The western blot is to be completed to determine how much knockdown is obtained from this 5' setup.

PfCRT (PF3D7_0709000)

Mutations in pfCRT have been shown to confer resistance to the antimalarial chloroquine in laboratory and in the field [13]. But what is the function of the native CRT gene? Is it essential to parasite survival? The pfCRT gene was edited three different ways in this project. First by using a zinc finger nuclease from the Fidock lab (confers WR resistance) [13] and a circular donor plasmid. Then the zinc finger nuclease with a linear donor plasmid. Lastly it was also edited using the CRISPR/Cas9 system.

In all of these cases parasite populations with an edited CRT locus were obtained (figure 24)

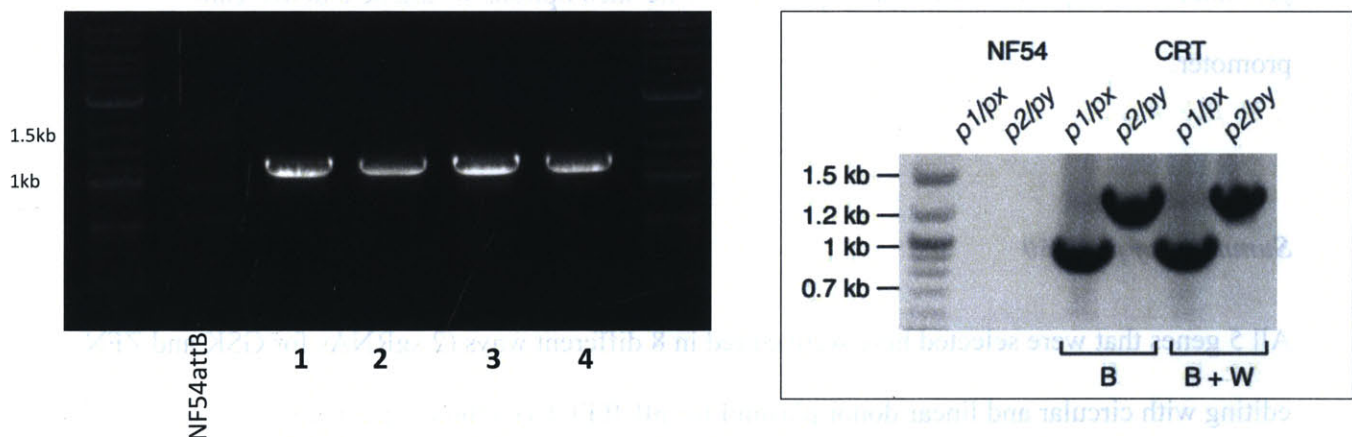


Figure 24: Editing of pfCRT locus A) p2+py PCR check of editing in ZFN edited population. NF54attB is negative. 1 is BSD +WR for 4 days, no selection ever after; 2 is parasites from 1 put back on BSD after 4 weeks off selection; 3 is BSD +WR selection for 4 days. BSD only selection afterwards; 4 is BSD only selection since day 2 after transfection

B) CRISPR/Cas9 pfCRT edited population. The BSD only and BSD+WR populations both tested positive

The HA tagged pfCRT protein has not been detected on western blots at the correct size. Growth assays became irreproducible and are not presented here.

Choline Kinase (PF3D7_1401800)

Choline kinase is another putatively essential gene with no conclusive data for essentiality. A parasite population tested positive for editing (not shown). Clones of CK were obtained and analyzed by growth assays. The assays were irreproducible but did not show any evidence of CK being essential. This result was unsurprising because CK is an enzyme and the 20% of protein available after knockdown could be enough to catalyze the phosphorylation of choline molecules. A higher knockdown would most likely elicit a phenotype for this putatively essential gene. This knockdown can be achieved by using 3' aptamers or by changing the pfCAM promoter used in pSN50 to a weaker promoter. One such option would be a uORF cam promoter.

Summary of pSN150

All 5 genes that were selected here were edited in 8 different ways (2 sgRNAs for GSK and ZFN editing with circular and linear donor plasmid for pfCRT). Even though parasite characterizations are incomplete in some cases (western blots, southern blots or growth assays missing), it is evident that this vector could be used to scale up the editing of genes for 5' regulation.

pSN053

Another vector in our linear vector family is pSN053. This vector allows for installing 3' aptamers of both 5' and 3' aptamers depending on the gene of interest and the characteristics of the gene.

Architecture of pSN053

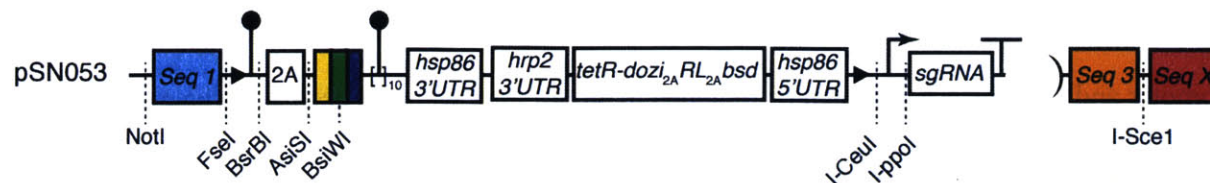


Figure 25: Zoom in of pSN053. All unique and useful cloning sites are shown. Seq1 and seq3 are required as previously described. The yellow, green and purple shaded boxes represent the FLAG, Myc and HA tags respectively

Applications of PSN053

1. 3' aptamer regulation
2. 5' and 3' dual aptamer regulation

Vector construction

3'-aptamer only regulation

1. ~500bp RHR (3' UTR of gene of interest) is cloned into Icu-I site
2. ~500bp LHR (3' end of coding sequence of GOI) is cloned between FseI+ AsiSI/BsiWI sites (AsiSI site gives access to the FLAG tag and Myc tag, BsiWI gives access to Myc tag and HA tag only). Notes: the regions around the sgRNA may need to be recoded depending on the distance from the sgRNA to the stop codon. For example, if the sgRNA

is 300bp from the stop codon, all 300bps will have to be recoded. It is also required to recode the sgRNA site itself to prevent cutting of the donor plasmid.

3. sgRNA can be cloned into I-PPOI meganuclease site. This site also contains an AflIII site which could be used if it is still unique. It is imperative to include the GG nucleotide sequence between the T7 promoter and sgRNA for T7 RNAP transcriptional initiation.

Notes: To make cloning 2 steps rather than 3 steps, the RHR and sgRNA can be cloned together. This can be achieved by digesting the PSN053 vector with Iceu-I and AflIII/I-PPOI. The T7 promoter and 20bp gRNA can then be included in an ultramer which is used as reverse primer for PCR of the RHR.

5' and 3'-aptamer regulation

1. ~500bp RHR (3' UTR of gene of interest) is cloned into Iceu-I site
2. ~500bp LHR (5' UTR of the GOI) is cloned into FseI site.
3. Recoded version of entire gene is cloned into AsiSI or BsiWI sites (AsiSI site gives access to the FLAG tag and Myc tag, BsiWI gives access to Myc tag and HA tag only). This recoding limits the length of genes that can be used for this strategy. Recoded DNA is expensive and only up to 2kb can be synthesized contiguously by IDT.
4. sgRNA can be cloned into I-PPOI meganuclease site. This site also contains an AflIII site which could be used if it is still unique. Again, it is imperative to include the GG nucleotide sequence between the T7 promoter and sgRNA for T7 RNAP transcriptional initiation.

Notes: Again, to make cloning 2 steps rather than 3 steps, the RHR and sgRNA can be cloned together. This can be achieved by digesting the PSN053 vector with Iceu-I and AflIII/I-PPOI. The

T7 promoter and 20bp gRNA can then be included in an ultramer which is used as reverse primer for PCR of the RHR.

To utilize PSN053, two conserved genes of unknown function were chosen for 5'&3' dual aptamer regulation. These were PF3D7_1360200 (MAL299) and PF3D7_0925900.

PF3D7_1360200 (MAL299)

MAL299 is a conserved protein in *Plasmodium* genus and of unknown function in all of these organisms. It is expressed during the intraerythrocytic development cycle of *Plasmodium falciparum*, according to RNA seq data from plasmodb.org

This gene fit the criteria for dual aptamer regulation, the tightest regulation we currently have in our lab. The spliced mRNA is short (<1kb) and so a recoded version of the gene can be purchased as a relatively inexpensive synthetic DNA (gblock) from IDT. The native MAL299 gene has 7 exons and 6 introns. The cDNA was recoded to *Saccharomyces cerevisiae* codon preference using the IDT codon optimization tool.

Cloning the donor vector

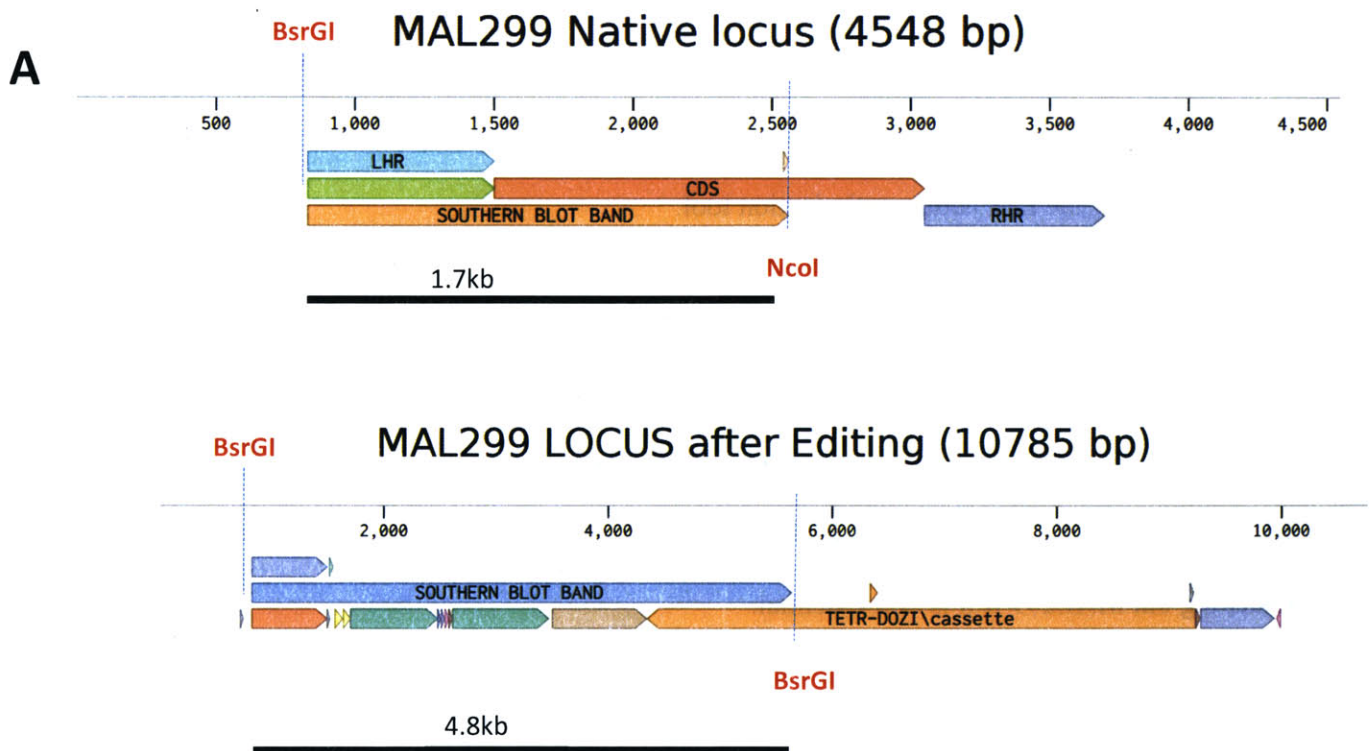
Using PSN053 as the base vector, the RHR, sgRNA, LHR and recoded gene were cloned into the Iccu-I, AflII, FseI and AsisI sites respectively in that order. The order of cloning is important because the PSN053_MA299_LHR_RHR_sgRNA vector could be used as a donor vector for knocking out MAL299 if it is non-essential.

Transfecting MAL299 donor vector

The linear pSN29_MAL299 donor plasmid was co-transfected with the pCRISPR INT plasmid into NF54attB parasites. A stable parasite strain was obtained after 24 days, 20 days after selection with only BSD and constant maintenance of the culture on aTc.

Analysis of MAL299 edited population

Editing of the MA299 locus was confirmed by both PCR and southern blotting at the population level. ~100% of the population was edited as shown by the southern blot (figure 26B) using LHR as probe and NcoI and BsrGI restriction digest of genomic DNA.



B



Figure 26: Genotyping MAL299 edited population. A) Design of southern blot experiment. The left homologous region (LHR) was used as probe. Genomic DNA was digested with BsrGI and NcoI. B) Southern blot confirms editing of MAL299 locus. NF54attB shows a native locus at 1728bp and edited locus in transfected parasites at 4805bp.

Western blot analysis on the population showed that almost no protein was present in parasites 72 hours after withdrawal of aTc (figure -----)

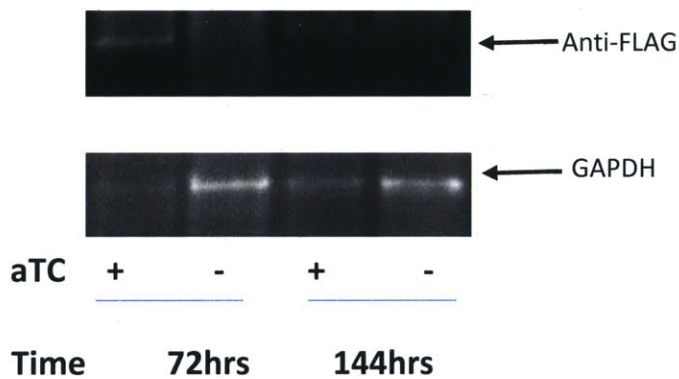
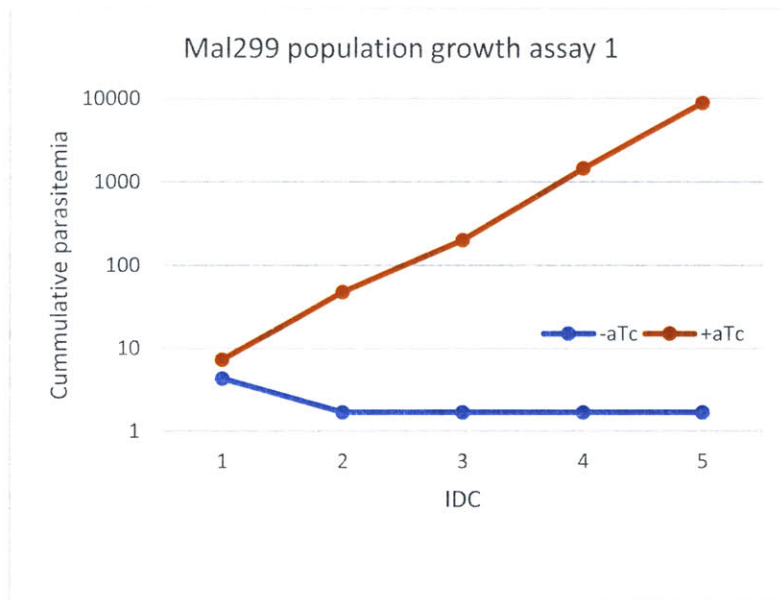


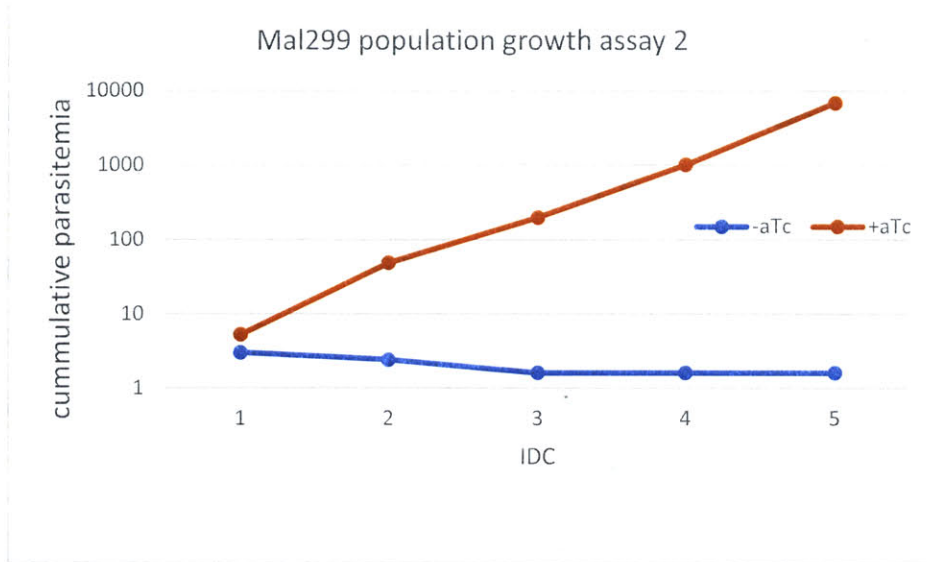
Figure 27: Anti FLAG tag western blot for the 2x-FLAG tagged MAL299 protein in MAL299 edited population. GAPDH is a loading control. After 72 hrs of aTc withdrawal, no MAL299 protein is available

Growth assay analysis suggested a severe growth phenotype in these modified parasites upon withdrawal of aTc. This phenotype is apparent from 72 hours (after one IDC), coinciding with the loss of the MAL299 protein from parasites. After the third IDC (144 hours after aTc withdrawal), the parasitemia could no longer be measured by the flow cytometer. These results are highly reproducible as shown in figure 28. Because the parasitemia fell below the detection limit of the flow cytometer, a renilla luciferase assay was used to analyze IDCs 3 and 4. The renilla luciferase signals measured were not higher than two fold above background in the samples with no aTc (figure---). The signal stayed the same for each sample even when parasites were allowed go through 1DC.

A



B



C

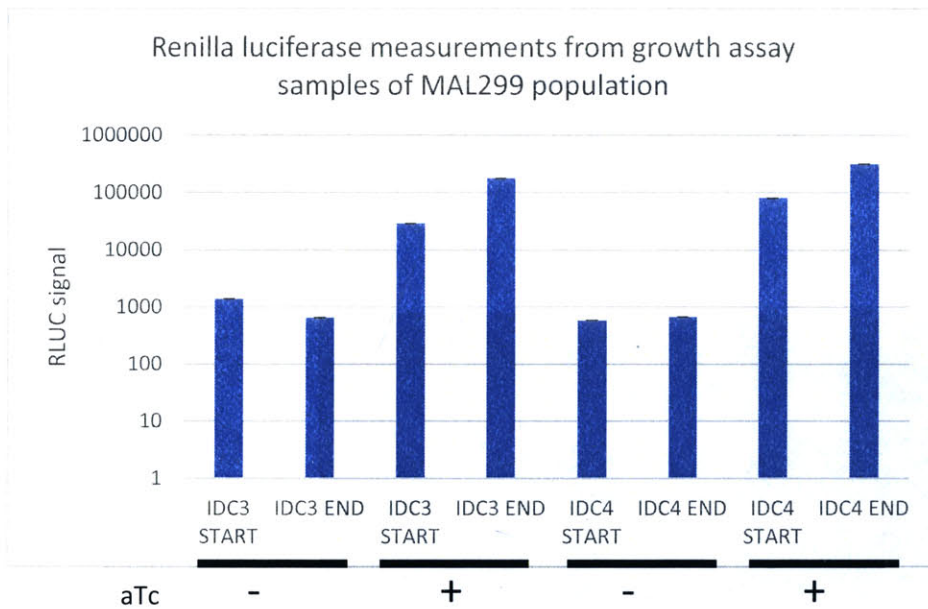


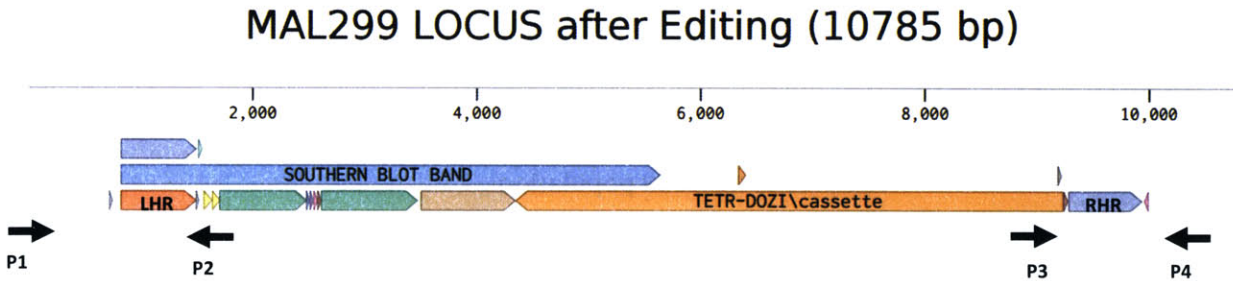
Figure 28: Growth assay and RLUC analysis of MAL299 edited populations. A) and B) reproducible growth assays show that MAL299 edited population show a severe growth defect within one IDC

C) Renilla luciferase measurements from growth assay samples show that -aTc parasites no longer expand after 2IDCs are parasitemia is very close to 0%

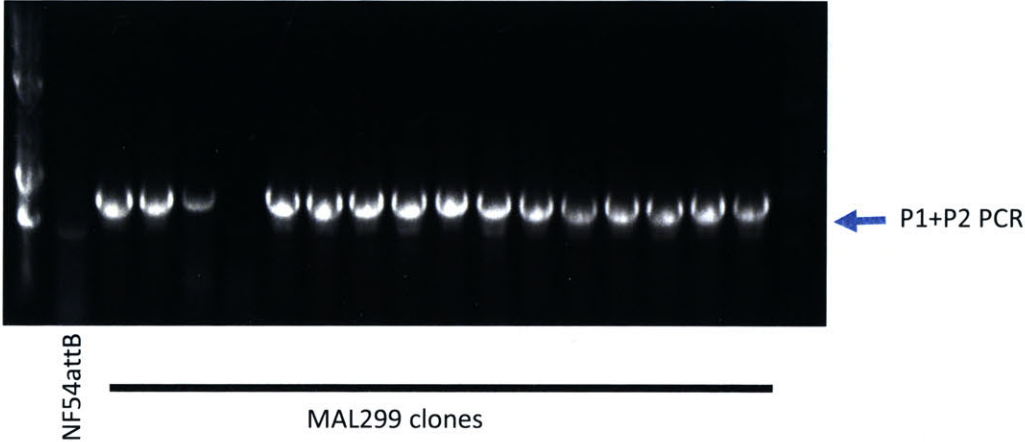
Analysis of MAL299 clones

The MAL299 population was then cloned out and 16 clones were obtained. All clones were analyzed by PCR for the integration event. Primers P1 and P2 check for integration at the 5' end of the locus. An expected product of 798bp was obtained in all but one of the 16 clones (figure 29A) with no PCR product obtained in unedited parent line. Primers P3 and P4 PCR out an 800bp fragment in all previously positive 15 clones to confirm editing in the 3' end of the locus (figure 29B). Six of these clones were further analyzed by southern blot (figure 29C) and confirmed to be edited as expected. These clones were then used for protein and phenotypic analysis.

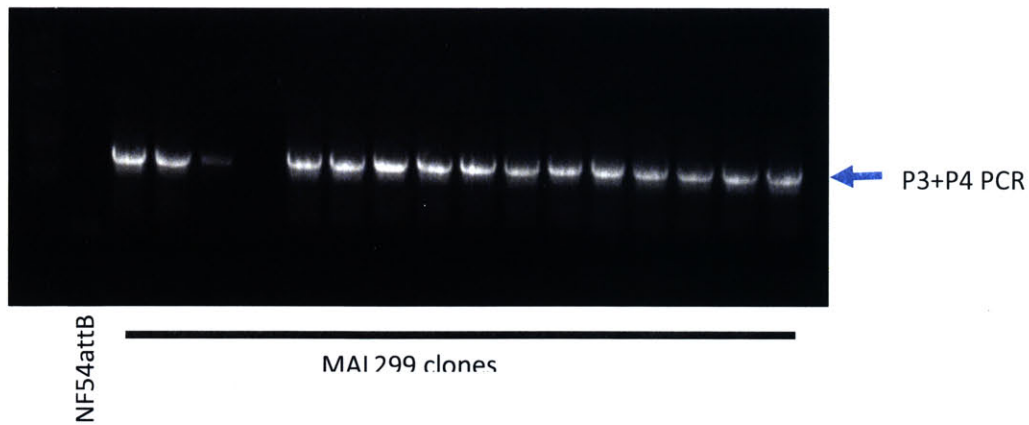
A



B



C



D

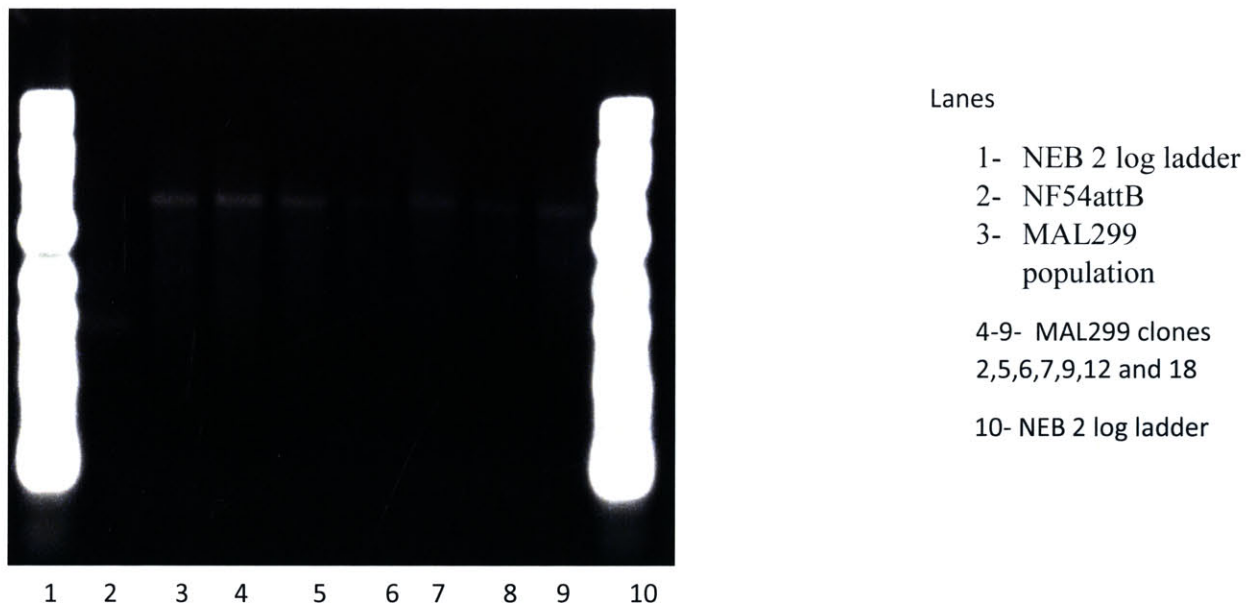


Figure 29: Genotyping MAL299 clones. A) Southern blot and PCR designs. Genomic DNA was digested with BsrGI and NcoI for southern blot. The LHR was used as probe. Primers P1,2,3 and 4 were used for PCR confirmation of editing. P1 and P4 bind on the genomic locus while P2 and P3 are standard primers that bind on the plasmid. B) 5' integration (P1+P2) is present in 15 of 16

clones and absent in parent unedited strain. C) 3' integration (P3+P4) is present in 15 of 16 clones and absent from parent strain. D) Southern blotting shows that 6 clones have the expected editing event. This is similar to what was obtained at the population level. Unedited parasites show the native MAL299 locus band

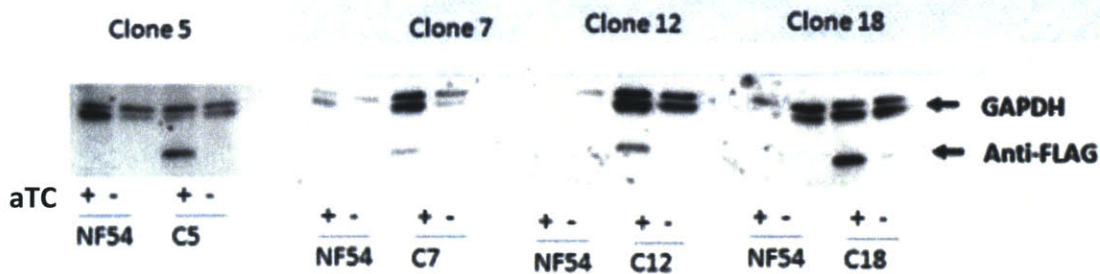


Figure 30: Anti FLAG tag western blot for the 2x-FLAG tagged MAL299 protein in MAL299 clones. GAPDH is a loading control. After 72 hrs of aTc withdrawal, no MAL299 protein is available in MAL299 clones.

Growth assays for these 4 clones showed an even more severe growth phenotype than what was observed in the population level. The clones are purer and most likely behave more similarly to each other than one would expect in a ~100% edited, albeit heterogeneous population.

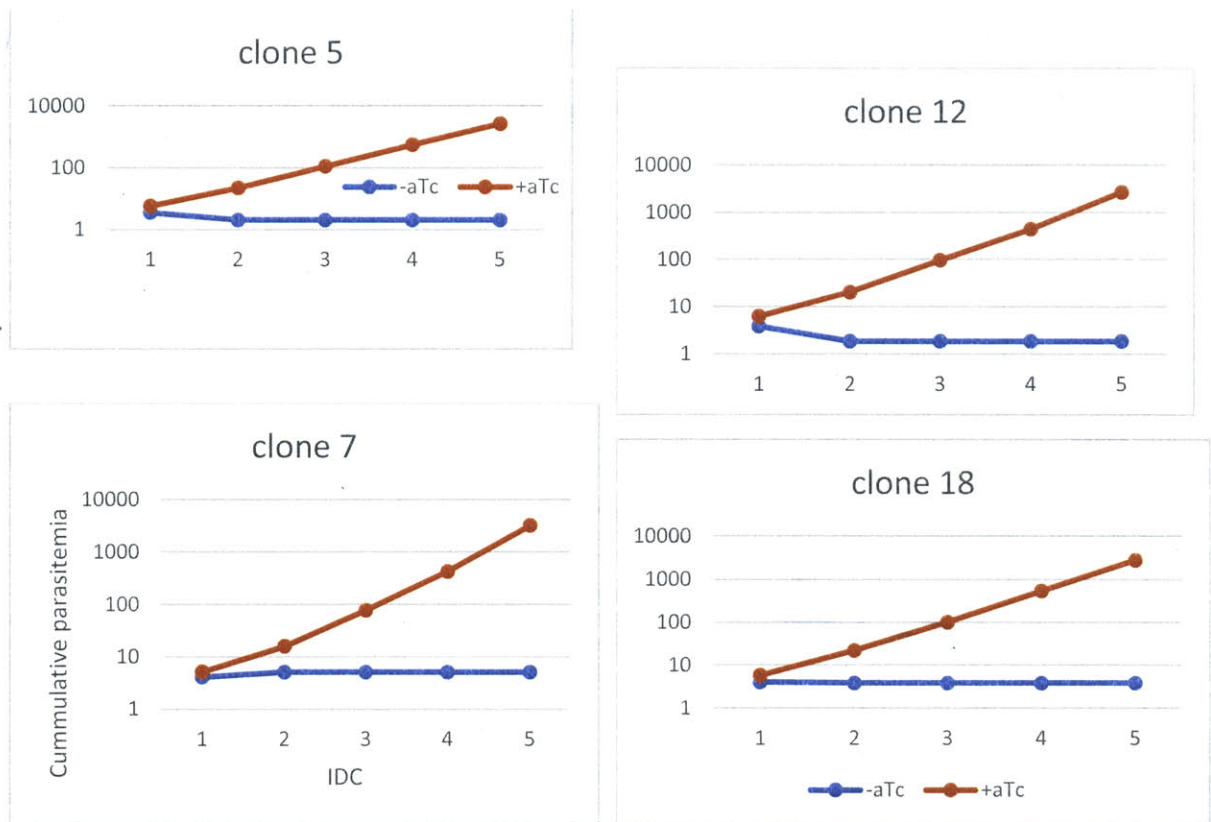


Figure 31: Growth assays show a severe growth defect in all 4 clones in which regulation of MAL299 protein was observed. Absence of MAL299 coincides with the beginning of the growth phenotype after IDC1

A second conserved protein of unknown function was investigated using the method described for MAL299. This gene, PF3D7_092500, is 654bp after splicing out the single intron it contains. This short length renders the gene easy for dual aptamer regulation via recoding of the entire coding sequence. PF3D7_0925900 was recoded to *Toxoplasma gondii* codon preferences and tagged with 2X FLAG just like MAL299.

To clone the donor vector for PF3D7_0925900, the native 5'UTR of the gene was retained by using a portion of it for left homologous region in the FseI site of pSN053. A portion of the 3' UTR was used as right homologous region. This RHR was cloned together with the sgRNA by using an ultramer that carried the sgRNA, T7 promoter and a region for Gibson cloning between

the IceuI and AflII sites of pSN053. Edited parasites were obtained in 4 weeks and confirmed to be edited by PCR. Growth assays on the population indicated that a growth phenotype was present (figure 32)

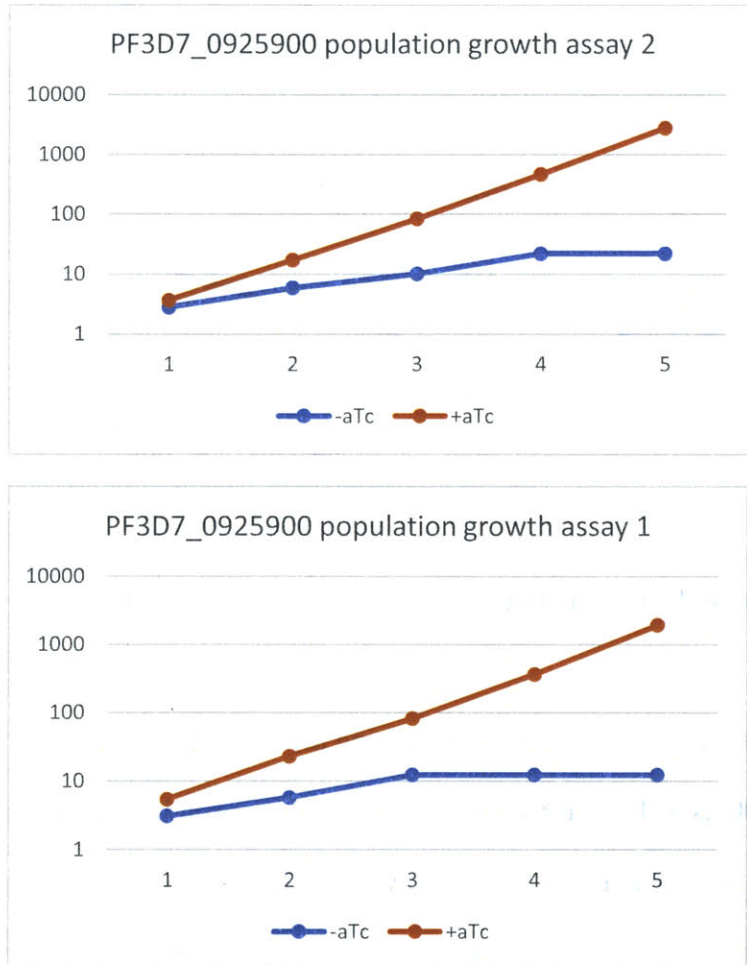
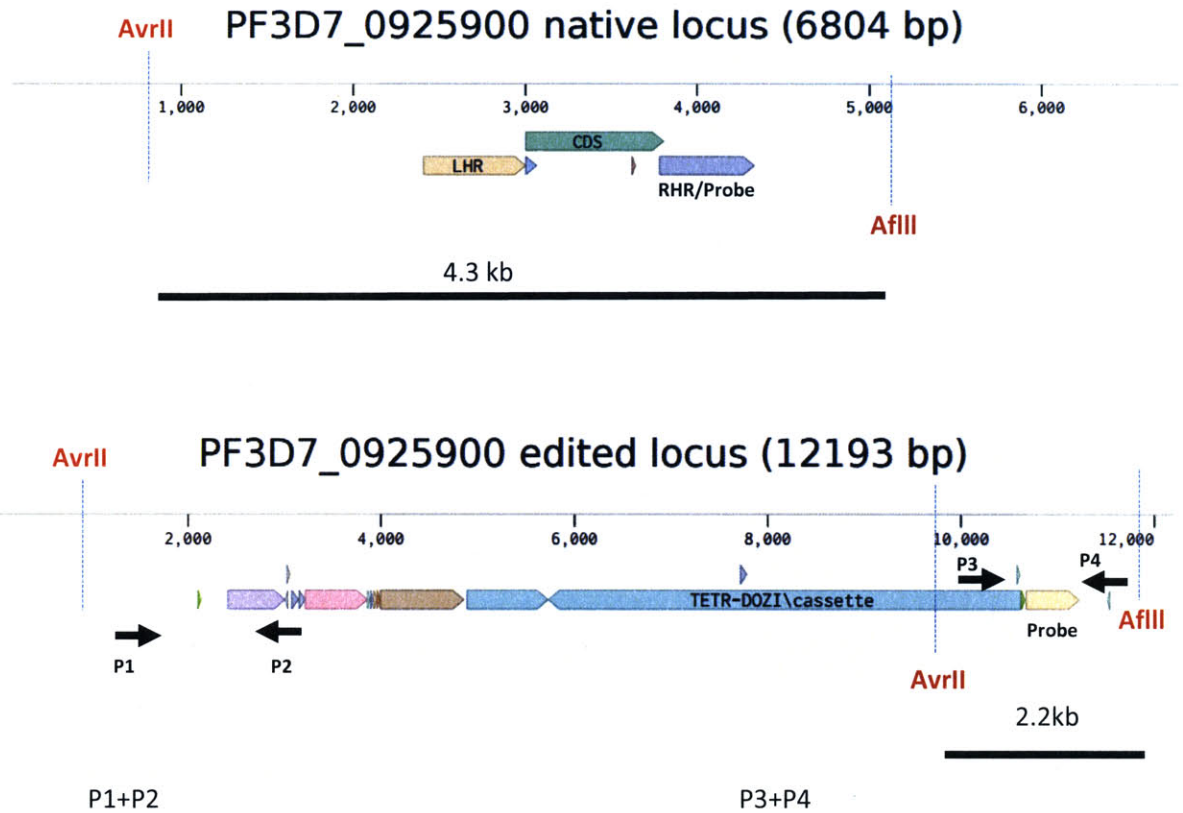


Figure 32: Growth assay analysis of edited PF3D7_092500 populations. Reproducible growth assays

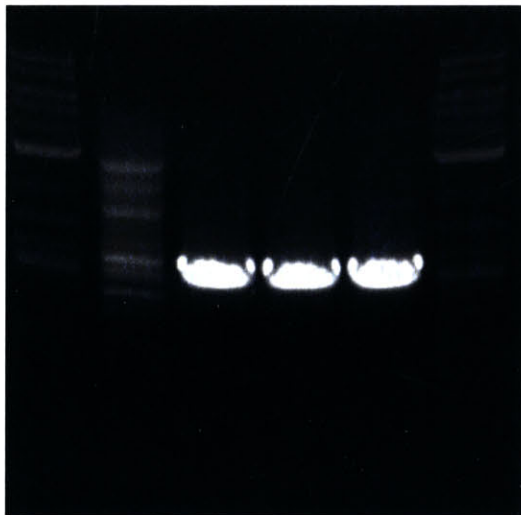
show that PF3D7_092500 edited population show a growth defect within one IDC

Analysis of PF3D7_092500clones

A



B



C

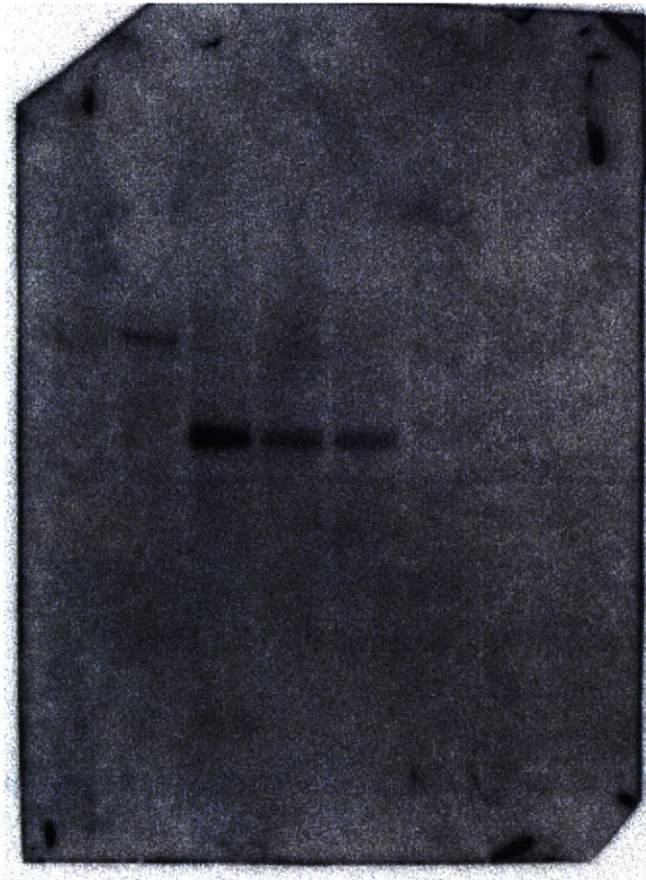


Figure 33: Genotyping PF3D7_0925900 clones. A) Southern blot and PCR designs. Genomic DNA was digested with AvrII and AflII for southern blot. The RHR was used as probe. Primers P1,2,3 and 4 were used for PCR confirmation of editing. P1 and P4 bind on the genomic locus while P2 and P3 are standard primers that bind on the plasmid. B) 5' integration (P1+P2) is present in 3 of 3 clones and absent in parent unedited strain. C) 3' integration (P3+P4) is present in 3 of 3 clones and absent from parent strain. D) Southern blotting shows that the 3 clones have the expected editing event.

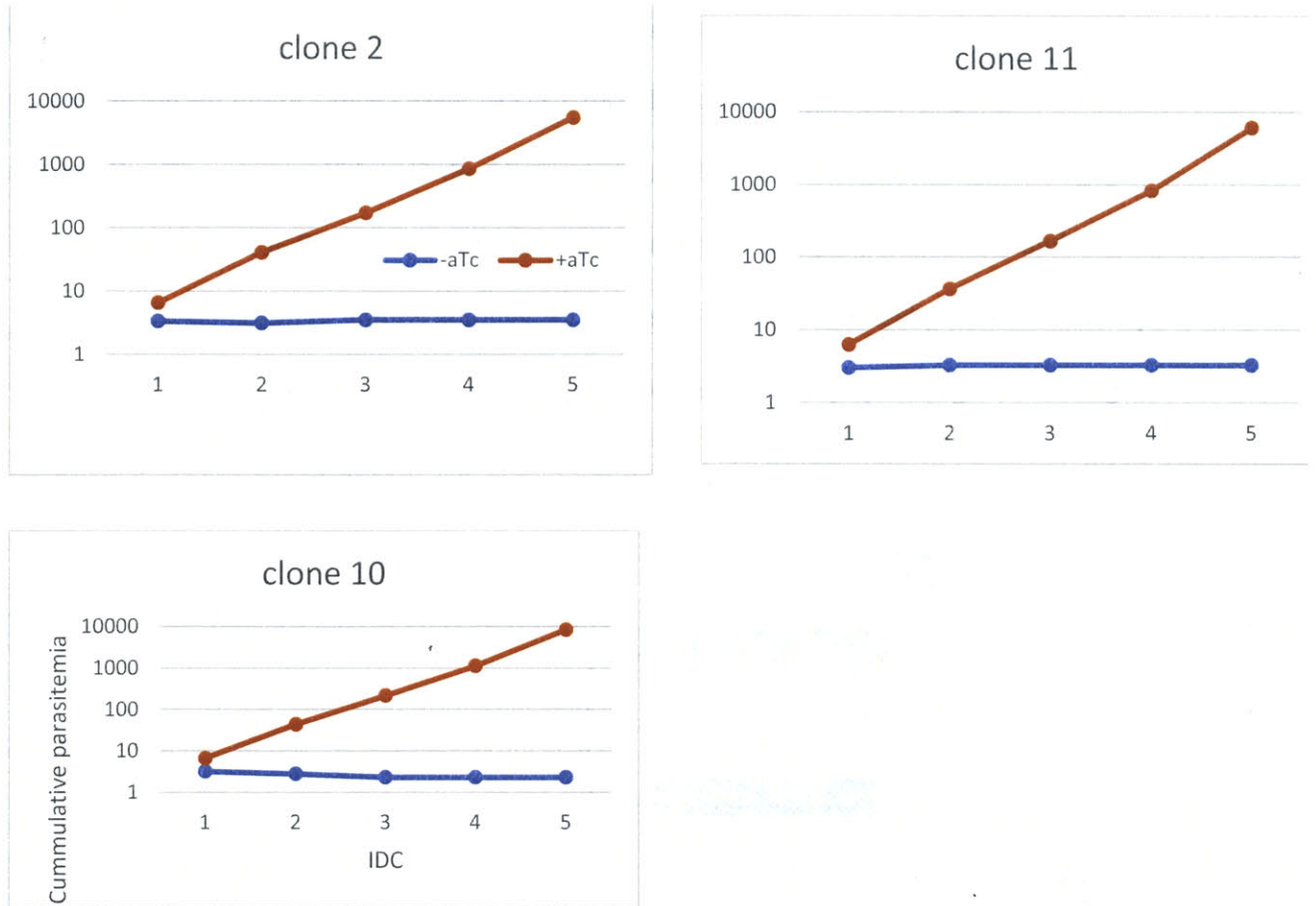


Figure 34: Growth assays show a severe growth defect in all 3 clones of *PF3D7_0925900*. This defect is apparent in IDC 1

pSN153/154

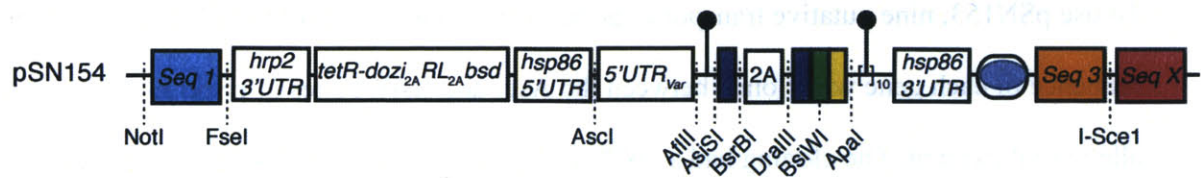


Figure 35: Zoom in of pSN154

pSN153 and pSN154 were designed and built as expression vectors. Using the multiple cloning site between a pfCAM 5'UTR and an HSP86 3'UTR, these vectors allow for the expression of any gene under the control of a CAM promoter and both 5' and 3' aptamers. The CAM promoter is modularized and subsequent versions of this plasmid contain a pfCRT promoter and a BIP promoter. These plasmids could also be used for genome editing to introduce 3' aptamer regulation for genes by spontaneous allelic replacement as illustrated in figure 29.

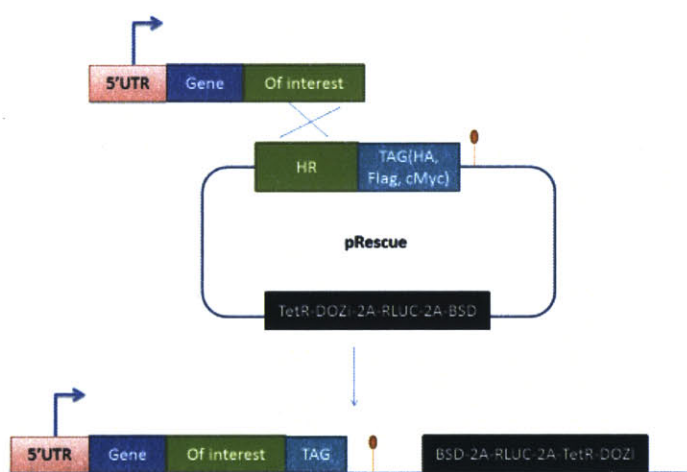


Figure 36: Spontaneous allelic for modifying genes in *P. falciparum* to install 3' aptamers

To use pSN153, nine putative transporter genes were cloned. ~1kb of the 3' end of the coding sequence of each gene was cloned between the AscI and BsiWI sites of pSN153 for spontaneous allelic replacement. The cloning had a ~67% or higher success rate. Six of the homologous regions cloned into the vector had perfect sequencing results and were rescued into circular plasmids for transfection. For spontaneous allelic replacement, drug cycling experiments were performed. Within one on-off cycle, one of the donor vectors (for (PF3D7_0210300) was

confirmed by PCR to be integrated into the genome. The integrated population was cloned and studied. One other donor plasmid (for PF3D7_0926400) yielded edited parasites but was not studied. The other 4 donor vectors did not yield any integrants after 3 on-off cycles (6 months of drug cycling)

PF3D7_092590 (HR2)

The population of parasites obtained from the integration of the HR2 donor plasmid was cloned. Twelve clones were obtained and genotyped by PCR (figure ---). Further genotyping was performed on three clones via southern blotting. These three clones were then analyzed for any growth defects upon knockdown of the HR2 protein.

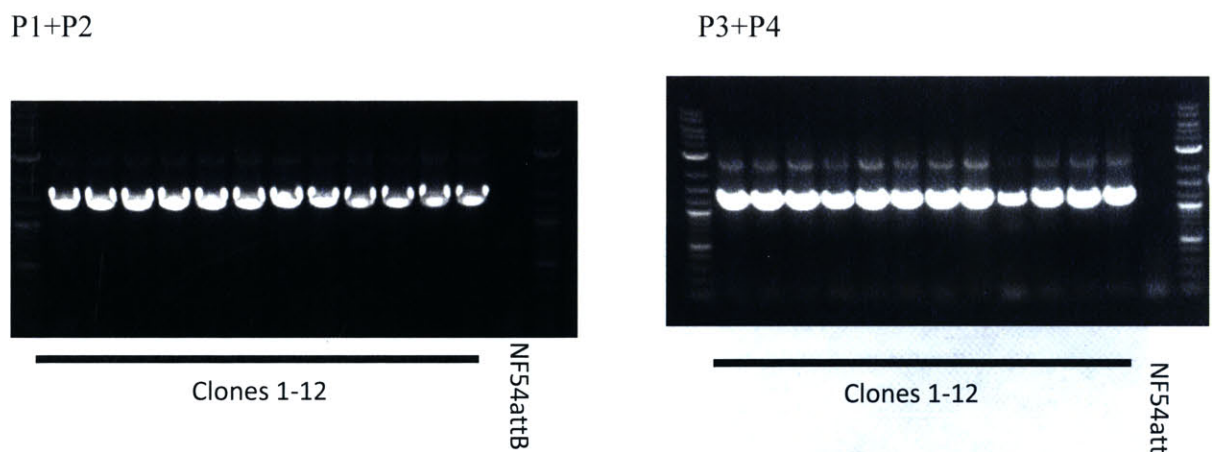
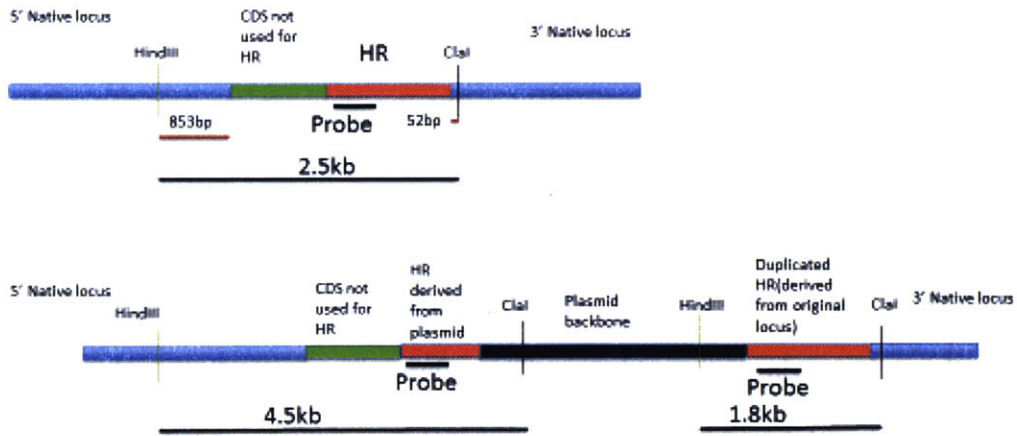


Figure 37: PCR test for integration of HR2 donor plasmid into the genome. P1+P2 confirms integration at the 5' end of the gene in all 12 clones. P3+P4 PCR confirms editing in the 3' end of the gene in all clones. NF54attb was negative for both PCR products

A

HR2 Southern Blot design 2- HindIII + Clal



Episomal rescue plasmid will be probed as well at 8.3kb

Notes: The Clal restriction site is downstream of 10x aptamer and HRPII 3'UTR on the plasmid backbone

B

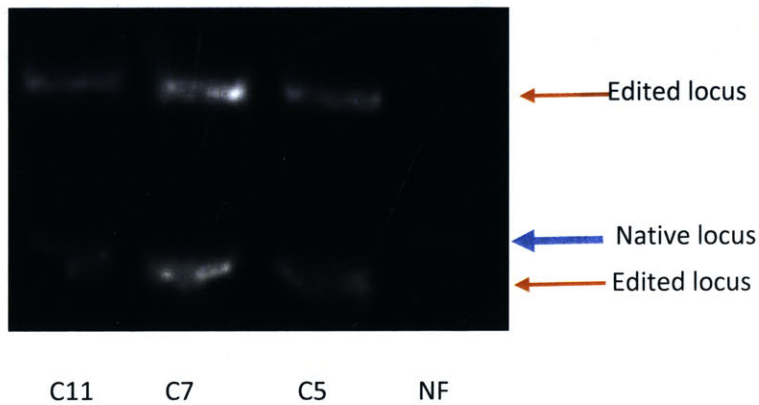
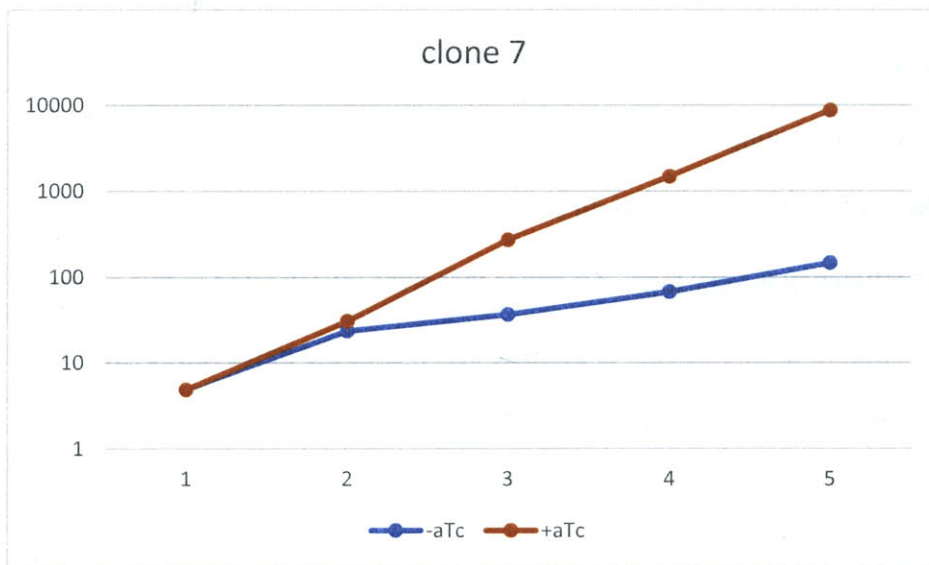
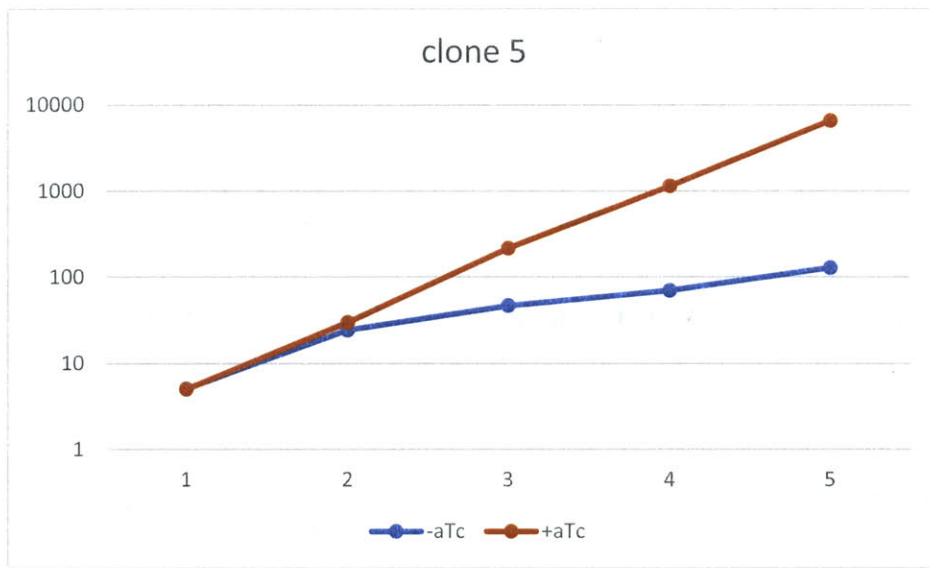


Figure 38: Southern blot analysis of HR2 clones. A) Southern blot design. The enzymes HindIII and Clal were used to digest gDNA. The probe was the 1kb HR region. NF54 parasites (NF-native locus only) are probed at 2.5kb while edited parasites are probed at 1.8kb and 4.5kb. B) Southern blot results

The three genotyped HR2 clones were then analyzed by growth assays for any growth defects.

Interesting, a reproducible growth phenotype was observed after IDC2 in all three clones (figure 28)



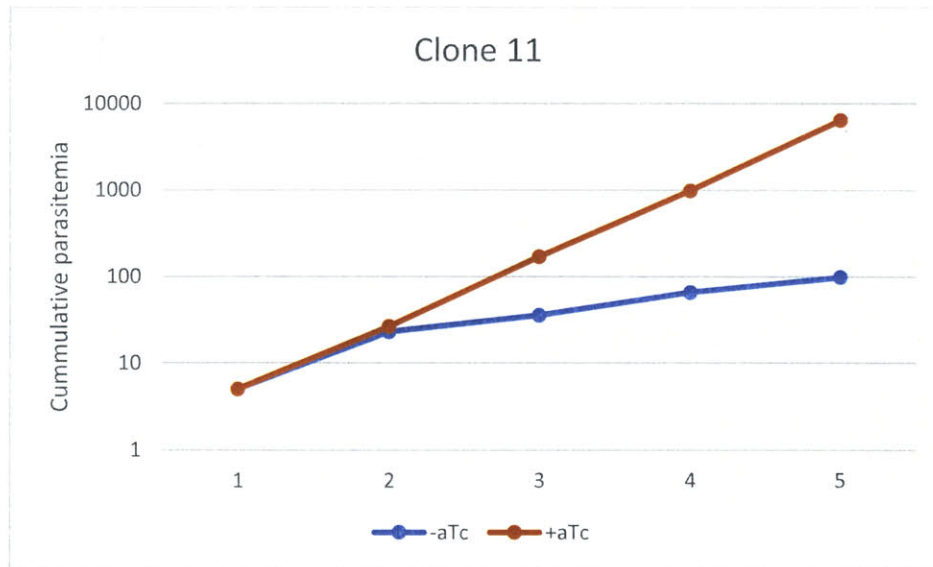


Figure 39: Growth assays for HR2 clones show a severe growth defect apparent after IDC2

Altogether, that data shown here point to essentiality of this previously unstudied HR2 protein.

Conclusions and future directions

In this document, a comprehensive set of vectors were presented that allow for any practically desirable genomic modification in *P. falciparum* to place genes under TetR-aptamer regulation. Because of the high efficiency of cloning into these vectors, donor plasmid construction time has been reduced from months to days. The system could easily be scaled up to investigate hundreds of genes since the cloning efficiency is very high. All genes that were investigated in this project were edited and some were shown to be essential to parasite survival in the intraerythrocytic developmental cycle. The hexose transporter has now been shown conclusively, through genetic manipulation, to be essential to parasites. Three conserved genes of unknown function and which have never been studied, (PF3D7_1360200, PF3D7_0962500 AND PF3D7_0210300) were also

discovered to be essential. This document also presented the first ever set of linear vectors that allow for genetic modification of *P. falciparum* without the need for typical circular plasmids.

One future modification to the pSN150 vector in particular is the creation of a promoter library to match the needs of any gene of interest. If a weaker promoter than pfCAM was used, there would be less choline kinase or TrxR protein than was shown in this document upon withdrawal of inducer. This lower background could provide more utility for discerning essentiality for enzymatic proteins in particular. An additional vector which was proposed but never built would allow for dual aptamer regulation of any gene, no matter how big, with a single plasmid.

Methods

Cloning of linear plasmids

Linear plasmids were obtained from Lucigen. They were grown in TSA BigEasy electrocompetent cells (Lucigen) according to manufacturer's instructions. For all clonings, 200ng of plasmid was digested with 1ul of enzyme in 10ul reactions. This digest was then dialyzed and 1ul of this product was mixed with 2ul of insert (of any concentration above 10ng/ul) in a 3ul 2x Gibson master mix (NEB)

Rescuing Linear Plasmids into BACs

pBIGBOB and pAdapter were obtained from the Lu lab at MIT. pBIGBOB was digested with PacI and pAdapter was digested with XbaI+XhoI. They were digested to a 7pg/ul concentration. Linear vectors cloned were also digested to this concentration with I-SceI and NotI. 7pg of the three vectors were then mixed in a Gibson reaction for rescue.

Culturing Plasmodium falciparum

The NF54attB strain of *P. falciparum* was cultured in leukocyte-free human RBCs (stored in acid-citrate-dextrose anticoagulant; from Research Blood Components, Brighton, MA) under an atmosphere of 5% O₂, 5% CO₂, and 95% N₂ at 2% hematocrit as previously described [115]. Parasites were synchronized with 0.3M alanine. To transfect parasites, uninfected red blood cells were washed twice with Wash Media (RPMI media with only RPMI and HEPES-KOH). A solution of 200μL wash media with 50μg plasmid DNA was added to 200μL packed fresh (<1 week old) uninfected red blood cells and electroporated using a 2mm cuvette and an 8

pulse/365volt square wave program. For co-transfections, 50µg of each plasmid was used, and for transfecting. Electroporated red blood cells were incubated at 37°C for 1hour and then washed twice with warm RPMI. Half was resuspended in 4 mL media in a 12 well plate. Next, 0.5 MI infected red blood cells taken directly from a 10 mL culture was added to the plate. Media was changed daily, and drug selection was added on day four. Transfections were monitored using renilla luciferase.

Western Blots

When required, parasites were lysed with a solution of 0.5% saponin, and incubated on ice for 10 minutes.

RPMI Media

- RPMI 1640 Medium (10.4 g/L)
- Hypoxanthine (30 mg/L)
- 25 mM HEPES-KOH (7.0825 g HEPES/L)
- Gentamicin (50 mg/L)
- Albumax II (5 g/L)
- Sodium bicarbonate (2 g/L)

Flow cytometry to determine parasitemia of Plasmodium falciparum

An aliquot of 200µL of parasites at 2% HCT were added to separate wells in a 96 well u-bottom plate

and stained with SYBR green 1 for 30 minutes in a 37°C incubator (0.2 μL per 1 mL RPMI).

Parasites

were washed twice with PBS and subsequently measured for SYBR green 1 expression on a flow cytometer. Uninfected red blood cells, both stained and unstained, were used for gating purposes.

Vector Design

Plasmids were designed using the online tool Benchling and vector NTI (VNTI). Primers were ordered from IDT and sequencing was done by Genewiz. In order to detect editing of the genomic locus, a set of primers was designed. Each construct shared two primers, px and py, and had two unique primers, p1 and

P4 (Table 1). After PCR products are run on a gel, they were gel-extracted and the sequence analyzed by Genewiz.

Genomic DNA Extraction

gDNA was extracted using the Qiagen Qiaamp Blood Mini Kit

ZFN plasmid

The ZFN pFCRT plasmid was obtained from the Fidock lab. The CRT HR originally on this vector was deleted and the plasmid was resealed.

References

1. Stephen J Goldfless, Brian J Belmont, Alexandra M de Paz, Jessica F Liu, and Jacquin C Niles. Direct and specific chemical control of eukaryotic translation with a synthetic rna-protein interaction. *Nucleic Acids Research*, 40(9):e64, May 2012.
2. Stephen J Goldfless, Jeffrey C Wagner, and Jacquin C Niles. Versatile control of plasmodium falciparum gene expression with an inducible protein-rna interaction. *Nat Commun*, 5:5329, Jan 2014.
3. Brian J Belmont and Jacquin C Niles. Engineering a direct and inducible protein-rna interaction to regulate rna biology. *ACS Chem Biol*, 5(9):851–61, Sep 2010.
4. Jeffrey C Wagner, Randall J Platt, Stephen J Goldfless, Feng Zhang, and Jacquin C Niles. Efficient crispr-cas9-mediated genome editing in plasmodium falciparum. *Nature Methods*, 11(9):915–8, Sep 2014.
5. Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S. et al. (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*, 419, 498–511
6. Jeffrey C Wagner, Stephen J Goldfless, Suresh M Ganesan, Marcus C S Lee, David A Fidock, and Jacquin C Niles. An integrated strategy for efficient vector construction and multi-gene expression in plasmodium falciparum. *Malar J*, 12:373, Jan 2013
7. Godiska R, Mead D, Dhodda V, et al. Linear plasmid vector for cloning of repetitive or unstable sequences in *Escherichia coli*. *Nucleic Acids Research*. 2010;38(6):e88.
8. Claudia Pfander, Burcu Anar, Frank Schwach, Thomas D Otto, Mathieu Brochet, Katrin Volkmann, Michael A Quail, Arnab Pain, Barry Rosen, William Skarnes, Julian C Rayner, and Oliver Billker. A scalable pipeline for highly effective genetic modification of a malaria parasite. *Nature Methods*, 8(12):1078–82, Dec 2011.
9. Suresh M. Ganesan, Alejandra Falla, Stephen J. Goldfless, Armiyaw S. Nasamu, and Jacquin C. Niles. Synthetic RNA-protein modules integrated with native translation mechanisms to control gene expression in malaria parasites. Unpublished
10. Benham, C.J., *PNAS*, 76(8): 3870-4; 1979

11. Patrick Guye, Yinqing Li, Liliana Wroblewska, Xavier Duportet, and Ron Weiss. Rapid, modular and reliable construction of complex mammalian gene circuits. *Nucleic Acids Research*,41(16):e156,Sep 2013.
12. Louis J Nkrumah, Rebecca A Muhle, Pedro A Moura, Pallavi Ghosh, Graham F Hatfull, William R Jacobs Jr & David A Fidock. Efficient site-specific integration in *Plasmodium falciparum* chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nature Methods* - 3, 615 - 621 (2006)
13. Judith Straimer, Marcus C S Lee, Andrew H Lee, Bryan Zeitler, April E Williams, Jocelynn R Pearl, Lei Zhang, Edward J Rebar, Philip D Gregory, Manuel Llinás, Fyodor D Urnov, and David A Fidock. Site-specific genome editing in plasmodium falciparum using engineered zinc-finger nucleases. *NatMeth*,9(10):993–8,Oct2012.
14. Yap, A., Azevedo, M. F., Gilson, P. R., Weiss, G. E., O'Neill, M. T., Wilson, D. W., Crabb, B. S. and Cowman, A. F. (2014), Conditional expression of apical membrane antigen 1 in *Plasmodium falciparum* shows it is required for erythrocyte invasion by merozoites. *Cellular Microbiology*, 16: 642–656
15. Brian Sauer. Inducible Gene Targeting in Mice Using the Cre/lox System. *METHODS: A Companion to Methods in Enzymology* 14, 381–392 (1998)
16. Kirk W. Deitsch, Casey L. Driskill and Thomas E. Wellems. Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes. *Nucleic Acids Res.* 2001; 29(3): 850-853
17. Nobutaka Kato, Tomoyo Sakata, Ghislain Breton, Karine G Le Roch, Advait Nagle, Carsten Andersen, Badry Bursulaya, Kerstin Henson, Jeffrey Johnson, Kota Arun Kumar, Felix Marr, Daniel Mason, Case McNamara, David Plouffe, Vandana Ramachandran, Muriel Spooner, Tove Tuntland, Yingyao Zhou, Eric C Peters, Arnab Chatterjee, Peter G Schultz, Gary E Ward, Nathanael Gray, Jeffrey Harper, and Elizabeth A Winzeler. Gene expression signatures and small-molecule compounds link a protein kinase to plasmodium falciparum motility. *Nat Chem Biol*,4(6):347–56,Jun2008.
18. Eliane Droucheau, Aline Primot, Virginie Thomas, Denise Mattei, Marie Knockaert, Chris Richardson, Pina Sallicandro, Pietro Alano, Ali Jafarshad, Blandine Baratte, Conrad Kunick, Daniel Parzy, Laurence Pearl, Christian Doerig, and Laurent Meijer. Plasmodium falciparum glycogen synthase kinase-3: molecular model, expression, intracellular localisation and selective inhibitors. *Biochim Biophys Acta*, 1697(1-2):181–96, Mar 2004.
19. Ana Rita Gomes, Ellen Bushell, Frank Schwach, Gareth Girling, Burcu Anar, Michael A Quail, Colin Herd, Claudia Pfander, Katarzyna Modrzynska, Julian C Rayner, and Oliver

- Billker. A genome-scale vector resource enables high-throughput reverse genetic screening in a malaria parasite. *Cell Host Microbe*, Feb 2015.
20. Ksenija Slavic, Michael J Delves, Miguel Prudêncio, Arthur M Talman, Ursula Straschil, Elvira T Derbyshire, Zhengyao Xu, Robert E Sinden, Maria M Mota, Christophe Morin, Rita Tewari, Sanjeev Krishna, and Henry M Staines. Use of a selective inhibitor to define the chemotherapeutic potential of the plasmodial hexose transporter in different stages of the parasite's life cycle. *Antimicrobial Agents and Chemotherapy*, 55(6):2824–30, Jun 2011.
 21. Ksenija Slavic, Ursula Straschil, Luc Reininger, Christian Doerig, Christophe Morin, Rita Tewari, and Sanjeev Krishna. Life cycle studies of the hexose transporter of plasmodium species and genetic validation of their essentiality. *Molecular Microbiology*, 75(6):1402–13, Mar 2010
 22. J I Penny, S T Hall, C J Woodrow, G M Cowan, A M Gero, and S Krishna. Expression of substrate specific transporters encoded by plasmodium falciparum in xenopus laevis oocytes. *Mol Biochem Parasitol*, 93(1):81–9, May 1998
 23. Zita Krnajska, Tim-Wolf Gilberger, Rolf D Walter, Alan F Cowman, and Sylke Müller. Thioredoxin reductase is essential for the survival of plasmodium falciparum erythrocytic stages. *J Biol Chem*, 277(29):25970–5, Jul 2002
 24. Gibson DG, Young L, et al. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5):343–345