Europe PMC Funders Group
Author Manuscript
Nat Methods. Author manuscript; available in PMC 2012 August 30.

Published in final edited form as:
Nat Methods. ; 8(12): 1078–1082. doi:10.1038/nmeth.1742.

A scalable pipeline for highly effective genetic modification of a malaria parasite

Claudia Pfander[1], **Burcu Anar**[1], **Frank Schwach**[1], **Thomas D. Otto**[1], **Mathieu Brochet**[1], **Katrin Volkmann**[1], **Michael A. Quail**[1], **Arnab Pain**[1,2], **Barry Rosen**[1], **William Skarnes**[1], **Julian C. Rayner**[1], and **Oliver Billker**[1]

[1]The Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK.

[2]Computational Bioscience Research Center, Chemical Life Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Kingdom of Saudi Arabia.

Abstract

In malaria parasites the systematic experimental validation of drug and vaccine targets by reverse genetics is constrained by the inefficiency of homologous recombination and by the difficulty of manipulating adenine and thymine (AT) rich *Plasmodium* DNA in *E. coli*. We overcome these roadblocks by demonstrating that a high integrity library of *P. berghei* genomic DNA (>77% AT) in a bacteriophage N15-based vector can be modified efficiently using the lambda Red method of recombineering. We built a pipeline for generating *Plasmodium berghei* genetic modification vectors at genome scale in serial liquid cultures on 96-well plates. Vectors have long homology arms, which increase recombination frequency up to 10-fold over conventional designs. The feasibility of efficient genetic modification at scale will stimulate collaborative, genome-wide knockout and tagging programs for *P. berghei*.

Experimental genetic manipulation has had a major impact on our understanding of *Plasmodium* biology and pathogenesis[1-3]. However, technological roadblocks have prevented scale-up beyond a few dozen genes per study[4,5]. In conventional protocols the genome of *P. berghei* is modified through homologous recombination with linear DNA fragments containing a selection cassette flanked on each side by 0.4-1.0 kb of sequence homologous to the target locus[6]. In other model systems, such as mouse embryonic stem cells, the routine use of much longer (up to 10 kb) homology arms has increased recombination frequency substantially[7,8], but generating similar vectors for use in *Plasmodium* has so far been impractical, since the extreme AT content of most *Plasmodium* genomes[9,10] renders large inserts of genomic DNA unstable in conventional bacterial plasmids and difficult to manipulate by restriction-ligation cloning. As a consequence large-insert genomic DNA (gDNA) libraries in bacterial artificial chromosomes (BACs), which form the basis for the construction of complex genetic modification vectors for model organisms[11-13], are not available for *Plasmodium*.

To overcome this obstacle we here describe the construction of a *P. berghei* gDNA library in a low-copy plasmid based on bacteriophage N15, which replicates in *E. coli* as linear,

Correspondence should be addressed to O.B. (oliver.billker@sanger.ac.uk) and J.C.R. (julian.rayner@sanger.ac.uk)..
AUTHOR CONTRIBUTIONS J.C.R and O.B. initiated and directed the research. C.P., A.P., B.R., W.S., J.C.R. and O.B. designed experiments. F.S., T.D.O., M.A.Q. and A.P. generated, sequenced, mapped and quality controlled the PbG01 library. C.P. and B.A. carried out experiments to develop the recombineering pipeline. C.P., M.B. and K.V. carried out experiments to validate the vectors. C.P. and O.B. wrote the manuscript. All authors analysed data and edited the manuscript.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

double-stranded DNA molecule with covalently closed hairpin telomeres[14,15] and which can carry relatively large inserts of AT-rich and repetitive DNA[16]. We show that the genomic DNA inserts of the library can be modified by homologous recombination in *E. coli* that transiently express the recombinase complex and proofreading activity encoded by the bacteriophage lambda *redγβα* operon and bacterial *recA* under the control of the arabinose-inducible *pBAD* promoter[17,18]. This technology, termed recombineering, requires only short (<50 bp) regions of homology, which can be included in synthetic oligonucleotides. Lambda red recombineering has been used to modify BACs from model organisms, including the related apicomplexan parasite *Toxoplasma gondii*[19,20], and its robustness and independence of restriction sites have made it the method of choice to scale up targeted gene disruption in the mouse[21].

We have developed a method to convert gDNA library inserts into gene deletion and tagging vectors for the genome-wide functional analysis of *P. berghei* genes. In our two-step strategy lambda red recombineering in the *E. coli* host is first used to introduce a bacterial selection marker into the gDNA insert, such that the target gene is either deleted or prepared for 3′-tagging. We then replace the bacterial marker with a selection cassette for *P. berghei* in a Gateway LR Clonase reaction *in vitro*. The modified library insert is released from the plasmid backbone using unique restriction sites and used to transfect *P. berghei*. We demonstrate that using this method vector production can be scaled up to a 96-well plate format and show that recombineered vectors integrate into the *P. berghei* genome with increased efficiency due to their long homology arms.

RESULTS

A *P. berghei* genomic DNA library with medium size inserts

We prepared libraries of size-selected *P. berghei* genomic DNA fragments in the bacteriophage N15 derived pJAZZ-OK vector from Lucigen[16] (Fig. 1a). End-sequenced inserts of 5,109 clones could be mapped onto the *P. berghei* ANKA reference genome with high confidence, resulting in an arrayed library termed PbG01. The average insert size was 9.0 kb (range 4.8 to 28.3 kb; Fig. 1b). The library contained 76% of *P. berghei* ANKA genes in their entirety and most genomic regions were covered by multiple clones (Fig. 1c). All mapped clones can be viewed in the *P. berghei* genome browser of the PlasmoDB database[22] at http://www.plasmodb.org. The predicted GC content of the library (22.59%) showed a small bias against AT-rich genomic regions when compared to the genome assembly against which it was mapped (22.13% GC). We therefore compared actual clone coverage to that of a simulated random clone set with the same distribution of insert sizes as the actual library. This analysis confirmed gene coverage was not entirely random (Fig. 1d). A curve fit to the data plateaued at around 95%, suggesting the remaining 5% of *P. berghei* genes may be unclonable in the N15 vector. To determine the level of sequence integrity in the PbG01 library, we sequenced 39 non-overlapping clones, totalling 332 kb of insert DNA, to >10x coverage. *De novo* assemblies contained no single nucleotide polymorphisms, rearrangements, insertions or deletions when compared to the *P. berghei* ANKA reference genome, demonstrating that fragments of *Plasmodium* DNA are stable in the pJAZZ vector.

A scalable pipeline to make genetic modification vectors

We next asked if N15 derived libraries, like BACs, can be modified in *E. coli* by exploiting homologous recombination mediated by products of the *red* operon of bacteriophage lambda[17]. To convert PbG01 inserts into genetic modification vectors *E. coli* clones were first made competent for homologous recombination by introducing a temperature sensitive plasmid, *pSC101gbaA-tet*, encoding the bacteriophage *λ red* operon and *E. coli recA*[18].

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Using homologous recombination in *E. coli* we generated an intermediate vector, in which the gene of interest was replaced with a bicistronic marker cassette for positive and negative selection in *E. coli,* called *zeo-PheS*[21], that was flanked by Gateway *attR1-attR2* sites (Fig. 2a). The selection cassette conveys resistance to zeocin through the bleomycin resistance gene (*Sh ble*) from *Streptoalloteichus hindustanus* and sensitivity to *p*-chlorophenylalanine through a mutant allele of a phenylalanine tRNA synthase (*pheS*). Positive selection with zeocin yielded the intermediate vector. In the second stage an *in vitro* Gateway LR reaction was used to exchange the bacterial marker by site-specific recombination for a much larger *hdhfr-yFCU* cassette for positive and negative selection in *P. berghei*[23]. The desired recombination product was obtained after negative selection on no-tryptone medium containing *p*-chlorophenylalanine (YEG-Cl).

Recombineering can be adapted to generate vectors for multiple genetic modification applications, including tagging, allele exchange and site-directed mutagenesis. To demonstrate this versatility, a modified two-stage strategy was developed for carboxy-terminal protein tagging (Fig. 2b). In this approach, the *zeo-PheS* cassette was first inserted immediately upstream of the stop codon, without modifying the open reading frame that is to be tagged. A triple haemagglutinin (3xHA) protein tag and a generic 3′UTR from the *P. berghei dhfr-ts* gene were then introduced in frame with the upstream ORF in the Gateway step (Fig. 2b). In contrast to conventional tagging strategies, in which the target locus is duplicated and will therefore revert to wild type at a low frequency[24], recombineered tagging vectors are designed to integrate stably by an ends-out replacement mechanism.

We tested the strategies to produce deletion and tagging vectors on 16 PbG01 clones (Supplementary Table 1) and verified each step by PCR genotyping. First the presence of the gene of interest in each library insert was confirmed by PCR (Fig. 2c). Overnight liquid cultures selected with zeocin at 37°C contained recombined plasmid (Fig. 2c) and had eliminated the temperature sensitive recombinase plasmid (not shown). After the Gateway step the first colony analysed contained the desired recombination product for 14 out of the 16 test genes. For one of the missing genes, screening additional colonies yielded the desired construct (not shown). Unmodified library plasmid persisted in one culture (arrow head in Fig. 2c), presumably because N15 derived library vectors exist in about 5 copies per cell, not all of which recombine at stage 1. Counter selection on YEG-Cl had effectively eliminated the intermediate vector in all cases. As expected, the Gateway donor plasmid *pR6K-3xHA*, which could give rise to episomal resistance in *P. berghei,* was also lost (not shown), because *E. coli* TSA lack the *pir* gene required by the *R6K* origin of replication.

As with BACs[12,21], recombineering on N15 derived phage vectors allowed sequential modification steps to be carried out in continuous liquid culture, i. e. without intermediary isolation and characterization of single colonies. This suggested recombineering of *Plasmodium* DNA could be scaled up from single tubes to 96-well plates. Using an optimized 8-day protocol (Fig. 3 and Supplementary Protocol 1) we found that of 96 clones processed in parallel, 77 could be converted to genetic modification vectors (Supplementary Table 2). Retrospective analysis of all unsuccessful wells showed that one library clone failed to grow, one failed to take up the recombinase plasmid, 15 failed at the recombineering stage and 2 failed at the Gateway step. The majority of recombineering failures (9 of 15) were due to mapping errors, which we rectified for the entire PbG01 library by re-mapping against the latest assembly of the *P. berghei* genome.

Recombineered vectors modify the *P. berghei* genome

To test whether PbG01-based vectors can be used to modify *P. berghei* parasites, we released 14 modified genomic inserts from the vector backbone by digesting with *Not*I restriction endonuclease and transfected each into *P. berghei* schizonts. Pyrimethamine

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

resistant parasites were obtained within 6 days and uncloned populations were genotyped initially by Southern blot hybridisation of separated chromosomes (Fig. 4a). Each of 14 deletion and tagging constructs integrated into the predicted chromosome and recombinants were the predominant genotype in all but one population (asterisk in Fig. 4a). No false integration events or episomally maintained vector was detected. Long range PCR was used to confirm that PbG01 derived vectors had integrated by double homologous (ends-out) recombination. To assess the success of protein tagging we analysed seven tagged lines by western blotting. Mobilities of tagged proteins were in satisfactory agreement with predictions, with the exception of PBANKA_135150, which expressed as a much shorter fragment than expected (asterisk in Fig. 4b). Protein expression in some tagged lines was stage-specific (arrows in Fig 4b), and in all cases was consistent with gene expression patterns predicted by *P. falciparum* transcriptome data[25]. Immunofluorescence microscopy of genomically tagged parasite lines (Fig. 4c) showed the expected cellular localizations, for instance for the cytosolic glycolytic enzyme phosphoglycerate kinase (PBANKA_082340), or a peripheral staining pattern for an alveolin (PBANKA_143660) association with the inner membrane complex of the ookinete[26].

Long homology arms increase recombination frequency

Finally we examined whether the long homology arms of recombineered vectors increase recombination frequency. Using combinations of restriction enzymes we reduced the homology arms of a deletion vector for the phosphodiesterase δ (*pdeδ*) gene (Fig. 5a) and then transfected *P. berghei* schizonts with equimolar amounts of the digested vectors, or with a conventional *pdeδ* deletion construct[27] with short (0.5 kb) homology arms. Recombination frequency was highly reproducible and increased linearly with the total length of homology arms (Fig. 5b), suggesting that PbG01 derived, recombineered targeting vectors can boost transfection efficiency by about 10-fold over traditional designs.

DISCUSSION

Our data show that serial, liquid recombineering in a 96-well format can be applied to AT rich *Plasmodium* DNA in a phage N15 derived genomic library at a high overall efficiency (88.7%) that is sufficient to extend the production of genetic modification vectors to a large part of the genome. In designing our pipeline we opted for a 2-stage approach because the *zeo-PheS* cassette used first is sufficiently small to be amplified robustly and with good yield by PCR. As a result the recombineering step is reproducibly efficient, which is critical to minimise the amount of unrecombined low-copy plasmid that initially persists under zeocin selection. The *zeo-PheS* cassette acts as an exchange module for a *Plasmodium*-specific cassette in a Gateway LR reaction. The *in vitro* exchange is followed by retransformation of *E. coli*, which functions as a critical purifying step that helps eliminate unreacted plasmid and unwanted minor reaction products, allowing clonal selection of final vectors prior to quality control.

Importantly, since the Gateway reaction is not size-limited, the *Plasmodium* selection module can be modified to incorporate other selection markers, protein tags and reporter genes. Our modular approach will thus enable researchers to select plasmids from a genome wide resource of intermediate vectors and readily convert these into customised gene tagging and deletion alleles in a simple and scalable *in vitro* reaction. *P. berghei* is already an important *in vivo* model for the fundamental biology of malaria, in part because it enables access to the mosquito and liver stages of the life cycle, which are much less tractable in human parasites. A genome-wide library of genetic modification vectors, combined with the flexibility and genetic tractability of the *P. berghei* model system, makes the genome-wide experimental analysis of *Plasmodium* gene function by targeted gene deletion and tagging a real possibility.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

METHODS

Parasites used

All transgenic *P. berghei* parasite were generated either in strain 2.34 ANKA wild type or in selectable marker free reporter strains expressing GFP (RMgm-7) or GFP-Luc (RMgm-29)[28], generated in the same genetic background. Parasites were propagated in Theiler's Original (TO) outbred mice and were transmitted regularly through *Anopheles stephensi*. All animal research was conducted under licences issued by the United Kingdom Home Office using protocols reviewed by the ethics committee of the Wellcome Trust Sanger Institute.

Plasmodium berghei genomic DNA library

The PbG01 library was constructed using the BigEasy v2.0 Linear Cloning System (Lucigen) essentially as described but with extra end-repair and gel purification steps in order to increase cloning efficiency and tighten insert size distribution. PbG01 clones were propagated in the BigEasy TSA bacterial strain (F-*mcr*A Δ(*mrr*-*hsd*RMS-*mcr*BC) φ80*dlacZ*ΔM15 Δ*lac*X74 *end*A1 *rec*A1 *ara*D139 Δ(*ara*, *leu*)7697 *gal*U *gal*K *rps*L *nup*G λ- *tonA bla (Amp*^{R}*) sopAB telN antA,* Lucigen) in TB medium + 0.4 % glycerol + 30 μg ml^{-1} kanamycin with or without arabinose induction. 10 μg genomic DNA was sheared by 20 passages through a 30 G needle and end-repaired by incubation with 0.3 μl of mung bean nuclease (GE, 256 U μl^{-1}) at 30 °C for 10 minutes prior to ethanol precipitation. DNA was pelleted by centrifugation, resuspended in 10 μl TE buffer, then size selected through a 0.8 % agarose gel. Fragments from 6 - 8, 8 - 10, 10 - 15 and 15 - 30 kb were excised, purified and end-repaired using the Lucigen kit reagents. A second gel size selection was performed before each size class was ligated separately into pJAZZ vector overnight at 14 °C. Following phenol:chloroform extraction with subsequent ethanol precipitation, each ligation was resuspended in 10 μl water and electroporated into BigEasy TSA (Lucigen) cells as recommended by the manufacturer. Cloning efficiency was high for fragments of 6-12 kb, with 91.3 % of clones containing an insert (range 73.8-99.2%, n = 5 libraries), but this dropped to 62.7 % for fragments of 12 - 30 kb (range 57.7 - 66.4 %, n = 3 libraries).

Mapping of PbG01 clones

13,702 *P. berghei* genomic inserts were subjected to capillary sequencing from both ends. Sequences were scanned for low-quality regions (Phred score < 10 in 40 nucleotide window) and trimmed or removed accordingly using custom scripts. Sequences with significant matches to the cloning vector were removed. Remaining insert-ends were mapped to the *P. berghei* assembly version of January 2011 (ftp://ftp.sanger.ac.uk/pub/pathogens/P_berghei/January_2011/) using the SMALT mapping software (http://www.sanger.ac.uk/resources/software/smalt/), allowing only uniquely mapping paired ends with a minimum Smith-Waterman mapping score of 200 and distance between clone ends consistent with the size selection steps during library preparation to within 20 %. Artemis[29] was used for visualising clones mapping to the genome. To view clones covering a genomic region of interest in PlasmoDB (http://plasmodb.org), "http://das.sanger.ac.uk/das/pjazz_berghei" must be entered as Remote Annotation URL. Sequences of selected PbG01 clones were verified in their entirety by sequencing pooled plasmids on one lane of an Illumina Genome Analyser II. The obtained reads were assembled with velvet[30], version 0.7.63. All resulting contigs were ordered with ABACAS[31] against the reference genome, visualized for manual inspection in the Artemis Comparison Tool and analysed for differences to the reference using SAMtools and BCFtools[32] with default parameters.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Plasmids for recombineering and Gateway reaction

Plasmid pR6K attR1-zeo-PheS-attR2 (Ref. 21) for amplification of the *zeo-PheS* cassette contains the *R6K* origin of replication and was propagated in *E. coli* PIR1 (Invitrogen) bacteria (*F-Δlac169 rpoS(Am) robA1 creC510 hsdR514 endA recA1 uidA (ΔMluI)::pir-116)* in LB Broth + 10 μg ml^{-1} tetracycline. pSC101gbdA (Ref. 18) is propagated at 30 °C in DH10B cells in LB Broth + 5 μg ml^{-1} tetracycline, or together with pJAZZ library clones in TB medium + 0.4 % glycerol + 30 μg ml^{-1} kanamycin + 5 μg ml^{-1} tetracycline. These plasmids were kind gifts from Francis Stewart, Dresden.

The Gateway donor plasmid pR6K attL1-3xHA-hdhfr-yfcu-attL2 (Supplementary Fig 1) was assembled in an *R6K* plasmid backbone containing a tetracycline resistance cassette. A DNA fragment composed of 3 x HA tag and 452 bp of *Pbdhfr* 3′UTR flanked by *attL1* and *attL2* sites was synthesized by Geneart. An expression cassette for *hdhfr-yfcu* was subcloned from pL0035 (ref. 23) via *Pst*I and *Acc*65I restriction sites. Both R6K vectors were propagated in PIR1 cells. A detailed protocol for 96-well recombineering is given in Supplementary Protocol 1. Library clones, recombineering tools and genetic modification vectors generated in the course of this study are available from the authors.

P. berghei transfection

Recombineered PbG01 vectors were digested with NotI to release the insert prior to electroporation. The transfection protocol is based on a published protocol[6] with modifications as described in Supplementary Protocol 2.

Genotyping of transgenic parasites

Blood from infected mice was collected by cardiac puncture, leukocytes removed by filtration on CF-11 columns and parasites isolated by ammonium chloride lysis. Genomic DNA isolated from one quarter of parasites (blood mini kit, Qiagen) was used for genotype verification by long range PCR using primers specific for the *Plasmodium* selection cassette (5′-catactagccattttatgtg-3′ or 5′-ctttggtgacagatactac-3′) and the target gene. Remaining parasites were used to confirm integration of the selection cassette into the correct chromosome. Chromosomes were separated by FIGE as described in Supplementary Protocol 3, and blotting onto nylon membrane was followed by hybridization with a radiolabelled probe against the 3′ UTR of *Pbdhfr-ts*, end-labelled with High Prime DNA labelling kit (Roche) and ^{32}P-γATP) according to standard Southern blot protocols. Signal intensities of integrated targeting constructs containing 2 3′UTRs of *Pbdhfr-ts* were quantified using ImageJ software and compared with endogenous signals to analyze purity of parasite population.

Calculation of recombination frequency

Transfection efficiency is given as $n_2 \times (n_1 \times 10^d)^{-1}$ where n_1 denotes the parasitaemia on day 1 post infection (i.e. the surviving parasites at the beginning of drug selection), n_2 denotes the parasitaemia on day d after the beginning of drug selection[33].

METHODS – ADDITIONAL REFERENCES

28. Janse CJ, Franke-Fayard B, Waters AP. Selection by flow-sorting of genetically transformed, GFP-expressing blood stages of the rodent malaria parasite, Plasmodium berghei. Nat Protoc. 2006; 1:614–623. doi:10.1038/nprot.2006.88. [PubMed: 17406288]
29. Rutherford K, et al. Artemis: sequence visualization and annotation. Bioinformatics. 2000; 16:944–945. [PubMed: 11120685]
30. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008; 18:821–829. doi:10.1101/gr.074492.107. [PubMed: 18349386]

31. Assefa S, Keane TM, Otto TD, Newbold C, Berriman M. ABACAS: algorithm-based automatic contiguation of assembled sequences. Bioinformatics. 2009; 25:1968–1969. doi:10.1093/bioinformatics/btp347. [PubMed: 19497936]

32. Li H, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. doi:btp352 [pii]10.1093/bioinformatics/btp352. [PubMed: 19505943]

33. Janse CJ, et al. High efficiency transfection of Plasmodium berghei facilitates novel selection procedures. Mol Biochem Parasitol. 2006; 145:60–70. doi:10.1016/j.molbiopara.2005.09.007. [PubMed: 16242190]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Francis Stewart (Technische Universität Dresden), Chris Janse (Leiden University Medical Center) and Andy Waters (University of Glasgow) for sharing reagents and to Jonathan Warren (Wellcome Trust Sanger Institute) and Haiming Wang (University of Georgia, Athens) for helping to create a DAS source of clone data for PlasmoDB. This work received support from the Wellcome Trust (WT089085/Z/09/Z), the Medical Research Council (grant numbers G0501670) and the European Virtual Institute for Malaria Research (EVIMalaR) Networks of Excellence (No. 242095).

References

1. Crabb BS, Cowman AF. Characterization of promoters and stable transfection by homologous and nonhomologous recombination in Plasmodium falciparum. Proc Natl Acad Sci U S A. 1996; 93:7289–7294. [PubMed: 8692985]

2. van Dijk MR, Janse CJ, Waters AP. Expression of a Plasmodium gene introduced into subtelomeric regions of Plasmodium berghei chromosomes. Science. 1996; 271:662–665. [PubMed: 8571132]

3. Wu Y, Kirkman LA, Wellems TE. Transformation of Plasmodium falciparum malaria parasites by homologous integration of plasmids that confer resistance to pyrimethamine. Proc Natl Acad Sci U S A. 1996; 93:1130–1134. [PubMed: 8577727]

4. Tewari R, et al. The systematic functional analysis of Plasmodium protein kinases identifies essential regulators of mosquito transmission. Cell Host Microbe. 2010; 8:377–387. [PubMed: 20951971]

5. Maier AG, et al. Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. Cell. 2008; 134:48–61. [PubMed: 18614010]

6. Janse CJ, Ramesar J, Waters AP. High-efficiency transfection and drug selection of genetically transformed blood stages of the rodent malaria parasite Plasmodium berghei. Nat Protoc. 2006; 1:346–356. [PubMed: 17406255]

7. Hasty P, Rivera-Perez J, Bradley A. The length of homology required for gene targeting in embryonic stem cells. Mol Cell Biol. 1991; 11:5586–5591. [PubMed: 1656234]

8. Shulman MJ, Nissen L, Collins C. Homologous recombination in hybridoma cells: dependence on time and fragment length. Mol Cell Biol. 1990; 10:4466–4472. [PubMed: 2117699]

9. Gardner MJ, et al. Genome sequence of the human malaria parasite Plasmodium falciparum. Nature. 2002; 419:498–511. [PubMed: 12368864]

10. Hall N, et al. A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses. Science. 2005; 307:82–86. [PubMed: 15637271]

11. Ejsmont RK, Sarov M, Winkler S, Lipinski KA, Tomancak P. A toolkit for high-throughput, cross-species gene engineering in Drosophila. Nat Methods. 2009; 6:435–437. [PubMed: 19465918]

12. Poser I, et al. BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals. Nat Methods. 2008; 5:409–415. [PubMed: 18391959]

13. Sarov M, et al. A recombineering pipeline for functional genomics applied to Caenorhabditis elegans. Nat Methods. 2006; 3:839–844. [PubMed: 16990816]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

14. Ravin NV, Ravin VK. Use of a linear multicopy vector based on the mini-replicon of temperate coliphage N15 for cloning DNA with abnormal secondary structures. Nucleic Acids Res. 1999; 27:e13. [PubMed: 10446256]
15. Ravin NV, Kuprianov VV, Gilcrease EB, Casjens SR. Bidirectional replication from an internal ori site of the linear N15 plasmid prophage. Nucleic Acids Res. 2003; 31:6552–6560. [PubMed: 14602914]
16. Godiska R, et al. Linear plasmid vector for cloning of repetitive or unstable sequences in Escherichia coli. Nucleic Acids Res. 2010; 38:e88. [PubMed: 20040575]
17. Zhang Y, Buchholz F, Muyrers JP, Stewart AF. A new logic for DNA engineering using recombination in Escherichia coli. Nat Genet. 1998; 20:123–128. doi:10.1038/2417. [PubMed: 9771703]
18. Wang J, et al. An improved recombineering approach by adding RecA to lambda Red recombination. Mol Biotechnol. 2006; 32:43–53. [PubMed: 16382181]
19. Gubbels MJ, et al. Forward genetic analysis of the apicomplexan cell division cycle in Toxoplasma gondii. PLoS Pathog. 2008; 4:e36. [PubMed: 18282098]
20. Brooks CF, et al. The Toxoplasma apicoplast phosphate translocator links cytosolic and apicoplast metabolism and is essential for parasite survival. Cell Host Microbe. 2010; 7:62–73. doi:S1931-3128(09)00412-0 [pii]10.1016/j.chom.2009.12.002. [PubMed: 20036630]
21. Skarnes WC, et al. A conditional knockout resource for genome-wide analysis of mouse gene function. Nature. 2011
22. Aurrecoechea C, et al. PlasmoDB: a functional genomic database for malaria parasites. Nucleic Acids Res. 2009; 37:D539–543. doi:gkn814 [pii]10.1093/nar/gkn814. [PubMed: 18957442]
23. Braks JA, Franke-Fayard B, Kroeze H, Janse CJ, Waters AP. Development and application of a positive-negative selectable marker system for use in reverse genetics in Plasmodium. Nucleic Acids Res. 2006; 34:e39. [PubMed: 16537837]
24. Menard R, Janse CJ. Gene Targeting in Malaria Parasites. Methods. 1997; 13:148–157. [PubMed: 9405198]
25. Le Roch KG, et al. Discovery of gene function by expression profiling of the malaria parasite life cycle. Science. 2003; 301:1503–1508. [PubMed: 12893887]
26. Tremp AZ, Dessens JT. Malaria IMC1 Membrane Skeleton Proteins Operate Autonomously and Participate in Motility Independently of Cell Shape. J Biol Chem. 2011; 286:5383–5391. [PubMed: 21098480]
27. Moon RW, et al. A cyclic GMP signalling module that regulates gliding motility in a malaria parasite. PLoS Pathog. 2009; 5:e1000599. [PubMed: 19779564]

Europe PMC Funders Author Manuscripts

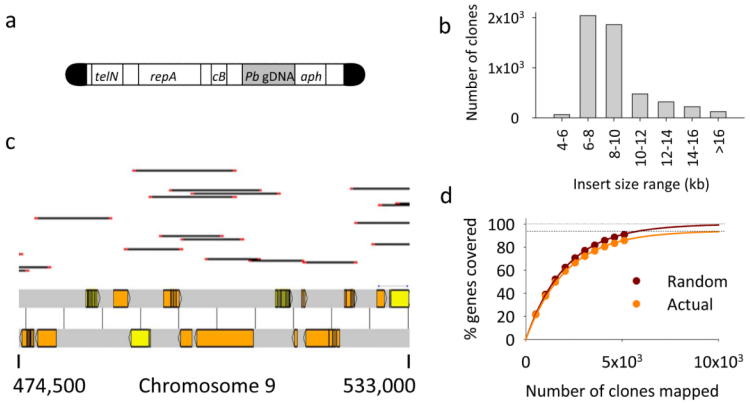

Figure 1. Characterisation of the *P. berghei* large insert genomic DNA library PbG01

(**a**) Schematic of the phage N15-derived pJAZZ vector used to generate the genomic library, showing hairpin telomeres (black), telomerase gene (*TelN*), replication factor and origin (*repA*), and kanamycin resistance gene (*aph*). (**b**) Distribution of insert sizes. (**c**) PbG01 inserts mapped on 65 kbp of chromosome 9 illustrates typical coverage. (**d**) Observed genome coverage by actual library inserts is compared with modelled coverage by random inserts. Percentage of genes covered to at least 50% is shown.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

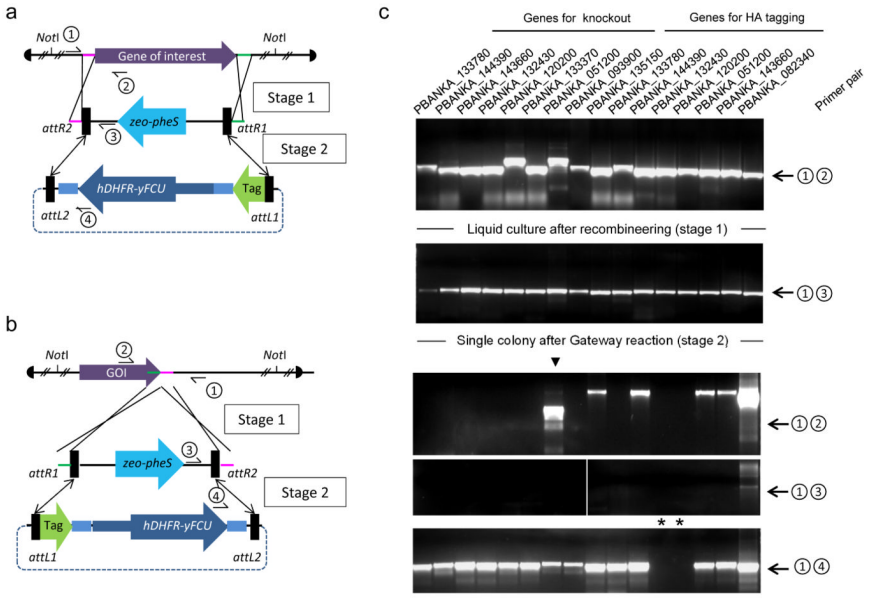

Figure 2. Modification of PbG01 inserts in *E. coli* by lambda Red recombineering and site specific recombinase

(**a**) A 2-stage strategy for gene deletion. Primer extensions homologous to 3′ and 5′ *P. berghei* target sequence are shown in magenta and green. (**b**) The strategy for 3′ tagging. (**c**) Step-by-step verification of vector product by PCR genotyping. See panels (**a**) and (**b**) for typical primer locations.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

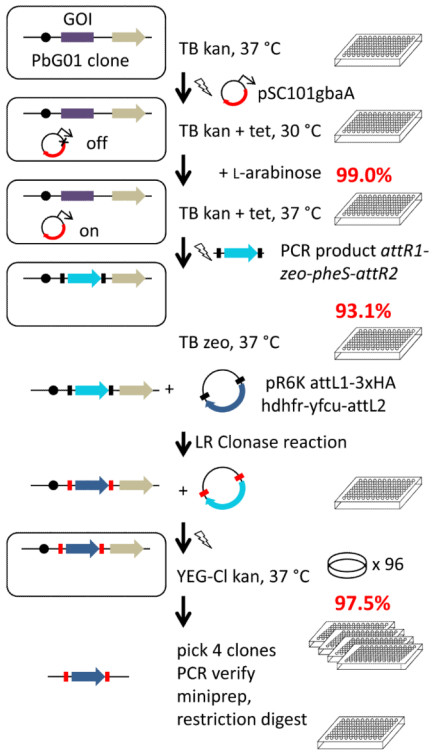

Figure 3. Knock out vector production in 96 parallel liquid cultures

Steps 1-3 and 5 take place in *E. coli*, steps 4 and 6 use purified vector *in vitro*. Cloning and genotyping is deferred until step 6. Following introduction of the recombinase plasmid, bacteria are cultured at a permissive temperature of 30 °C (step 1). Recombinase expression is induced by arabinose (step 2), and bacteria are electroporated with PCR products containing the *zeo-PheS* cassette flanked by 50 base pairs homologous to the chosen target locus (step 3). An *in vitro* Gateway reaction (step 4) switches the bacterial marker to one for *P. berghei*. Plasmids are retransformed into *E. coli* and plated on *p*-chlorophenylalanine (YEG-Cl) to select for recombination products lacking *pheS* (step 5). Colonies are picked for PCR verification (step 6). Percentages shown in red give average efficiencies of individual steps. See also Supplementary Protocol 1.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

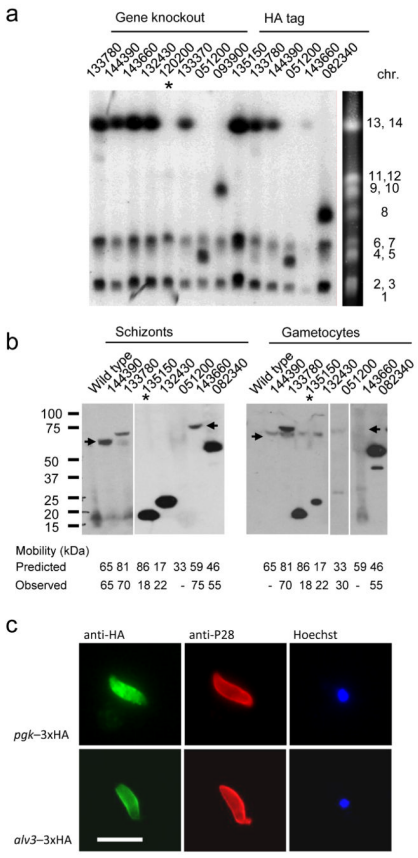

Figure 4. Validation of recombineered vectors in *P. berghei* ANKA

(**a**) Primary genotyping of resistant parasite pools by Southern hybridisation of separated chromosomes. The probe recognises two copies of the *dhfr-ts* 3′UTR in the targeting vector (variable band) and additionally highlights chromosome 7 (endogenous *dhfr-ts* gene), and chromosome 3 (*gfp* transgene integrated into the *p230p* locus, PBANKA_030600). The expected chromosomal location of target genes is given by the first two digits of the gene ID. * = recombinant genotype is not in the majority, as judged by band intensity. (**b**) Western blot analysis showing expression of HA-tagged proteins in lysates from schizonts and gametocytes. (**c**) Immunolocalisation of HA-tagged proteins showing localisation to the cytosol (PBANKA_082340, PGK), or a peripheral staining pattern consistent with localisation to the inner membrane complex (PBANKA_143660, alveolin 3, IMC1h). Fixed and permeabilised ookinetes were counter stained with Hoechst for DNA and with a monoclonal antibody against the major surface protein P28. Scale bar = 10 μm.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

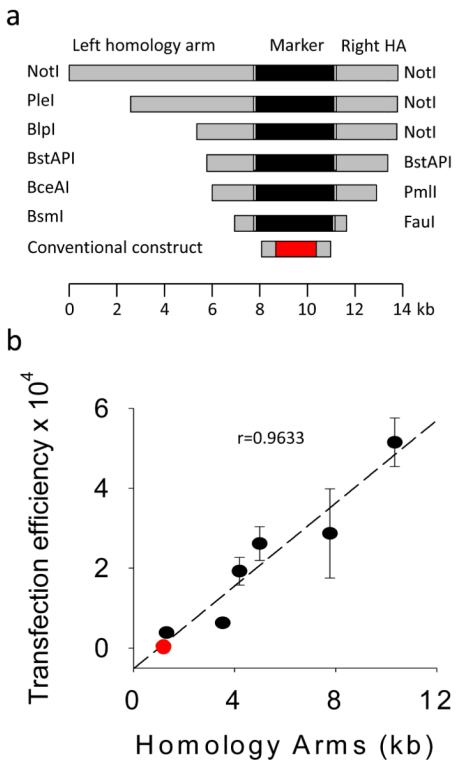

Figure 5. Effect of homology arm length on targeting frequency

(**a**) A panel of deletion vectors for the *pdeδ* gene. The restriction enzymes shown were used to modify lengths of homology arms. (**b**) Transfection efficiency is plotted against the sum of both homology arms. Error bars show standard deviations from three transfections.