High-Throughput Generation of P. falciparum Functional Molecules by Recombinational Cloning

João Carlos Aguiar,1,3,8 Joshua LaBaer,2 Peter L. Blair,1 Victoria Y. Shamailova,1 Malvika Koundinya,2 Joshua A. Russell,1 Fengying Huang,1 Wenhong Mar,2 Robert M. Anthony,1,4 Adam Witney,1,5 Sonia R. Caruana,1 Leonardo Brizuela,2 John B. Sacci Jr.,1,6 Stephen L. Hoffman,1,7 and Daniel J. Carucci1

1Malaria Program, Naval Medical Research Center, Silver Spring, Maryland 20910, USA; 2Harvard Institute of Proteomics, Cambridge, Massachusetts 02141, USA; 3Pan American Health Organization, Washington, DC 20037, USA.

Large-scale functional genomics studies for malaria vaccine and drug development will depend on the generation of molecular tools to study protein expression. We examined the feasibility of a high-throughput cloning approach using the Gateway system to create a large set of expression clones encoding Plasmodium falciparum single-exon genes. Master clones and their ORFs were transferred en masse to multiple expression vectors. Target genes (a = 303) were selected using specific sets of criteria, including stage expression and secondary structure. Upon screening four colonies per capture reaction, we achieved 94% cloning efficiency. The genes were subcloned in parallel into three expression vectors: a DNA vaccine vector and two protein expression vectors. These transfers yielded a 100% success rate without any observed recombination based on single colony screening. The functional expression of 95 genes was evaluated in mice with DNA vaccine constructs to generate antibody against various stages of the parasite. From these, 19 induced antibody titers against the erythrocytic stages and three against sporozoite stages. We have overcome the potential limitation of producing large P. falciparum clone sets in multiple expression vectors. This approach represents a powerful technique for the production of molecular reagents for genome-wide functional analysis of the P. falciparum genome and will provide for a resource for the malaria resource community distributed through public repositories.

[Supplemental material is available online at www.genome.org. All entry clones from both Panels A and B have been deposited at the Malaria Research and Reference Reagent Resource Center (http://www.malaria.mr4.org) for distribution to researchers upon request. The following individuals kindly provided reagents, samples, or unpublished information as indicated in the paper: Y. Charoenvit and P. de la Vega.]

The entire genomic sequence of the human malaria parasite Plasmodium falciparum has been completed (Gardner et al. 2002); two of the 14 chromosomes were published earlier (Gardner et al. 1998; Bowman et al. 1999). In addition, the genome of the rodent malaria parasite Plasmodium yoelii has also been sequenced (Carlton et al. 2002), and sequencing of Plasmodium vivax, the other important human malaria species, is well underway (Carlton 2003). In addition, a significant number of genes and/or gene fragments have been sequenced through GST, EST, and whole-genome shotgun projects from other Plasmodium species including Plasmodium berghei, Plasmodium c. chabaudi, Plasmodium gallinaceum, Plasmodium knowlesi, and Plasmodium reichenowi.

For many years, research on malaria has focused on a small fraction of the Plasmodium genome, resulting in the identification of a few targets of protective immunity and fewer targets of new drug development. Despite the identification of >5200 parasite genes in the genome, new vaccines and drugs are unlikely to be developed solely from this set of data. Researchers are now focusing on large-scale functional genomics studies combined with powerful relational databases and informatics to “credential” the genome; to determine the characteristics of each encoded protein. A variety of high-throughput novel technologies are now being applied, such as DNA microarray (Ben Mamoun et al. 2001; Bozdech et al. 2003; Le Roch et al. 2003), protein microarrays (Bacarese-Hamilton et al. 2002), yeast two-hybrid (Y2H) protein interaction (University of Washington/Prolexys Pharmaceuticals/NMRC collaboration), and mass spectrometry (MS; Florens et al. 2002; Lasonder et al. 2002). A major focus of these efforts will undoubtedly be the development of new vaccines and drug targets (Hoffman et al. 1998, 2002).

Two recent genome-wide studies have elucidated the P. falciparum proteome and transcriptome of several parasite stages including sporozoites, merozoites, trophozoites, and gametocytes. By using large-scale proteomic approaches (Florens et al. 2002; Lasonder et al. 2002) and a high-density oligonucleotide array (HDA; Le Roch et al. 2003), several proteins and clusters of stage- and or function-specific have been identified for high-throughput studies.

Given that sterile protection can be conferred in humans via the bites of thousands of mosquitoes containing attenuated P. falciparum sporozoites and subunit vaccines have not yet conferred comparable protection, some hypothesize that an effective vaccine will require combining better protective antigens with vaccine delivery systems capable of eliciting the appropriate immune response(s). Together, the complexity and inaccessibility
# High-Throughput Generation of P. Falciparum Functional Molecules by Recombinational Cloning

**Naval Submarine Medical Research Laboratory**
Naval Submarine Base New London
Box 900 Bldg 148, Trout Avenue
Groton, CT 06349-5900

1. **REPORT DATE**
   - 2004

2. **REPORT TYPE**
   - N/A

3. **DATES COVERED**
   - 

4. **TITLE AND SUBTITLE**
   - High-Throughput Generation of P. Falciparum Functional Molecules by Recombinational Cloning

5. **AUTHOR(S)**
   - 

6. **PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
   - Naval Submarine Medical Research Laboratory
   - Naval Submarine Base
   - New London
   - Box 900 Bldg 148, Trout Avenue
   - Groton, CT 06349-5900

7. **SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
   - 

8. **PERFORMING ORGANIZATION REPORT NUMBER**
   - 

9. **DISTRIBUTION/AVAILABILITY STATEMENT**
   - Approved for public release, distribution unlimited

10. **SUPPLEMENTARY NOTES**
    - The original document contains color images.

11. **ABSTRACT**
    - 

12. **SECURITY CLASSIFICATION OF:**
    - a. REPORT
      - unclassified
    - b. ABSTRACT
      - unclassified
    - c. THIS PAGE
      - unclassified

13. **NUMBER OF PAGES**
    - 7

14. **NAME OF RESPONSIBLE PERSON**
    - 

---

*Standard Form 298 (Rev. 8-98)*
Prescribed by ANSI Std Z39-18
of certain malaria life cycle stages, the poor understanding of mechanisms of immunity, the lack of algorithms to predict targets of protective immune mechanisms, and the power of DNA-based vaccines have provided a foundation for a genomes-to-vaccine program and have required the development of large-scale gene cloning technologies (Doolan et al. 2003a). Hence, a high-throughput DNA cloning technology is critical for the discovery of new vaccine targets based on the reverse vaccinology approach currently being applied to various genomes (Rappuoli and Covacci 2003).

To pursue such an approach, the establishment of a repository of cloned genes would be extremely helpful and would also provide an important and valuable resource to the research community. The exploitation of the P. falciparum genome sequences has led to the identification of novel biochemical pathways and new putative targets of heretofore unrecognized antimalarial compounds, that is, drugs traditionally used as antibacterial agents (Waller et al. 1998; Jomaa et al. 1999). Using the malaria genome sequence for vaccine development is less than for drug development. Researchers undertook large-scale cloning efforts to produce 350 recombinant proteins from Neisseria meningitidis as a resource for subsequent antibody production and identification of new vaccine candidates (Pizza et al. 2000). This “reverse vaccinology” approach has shown promise in spite of the “brute force” required to generate the hundreds of recombinant molecules for screening. As the numbers of genomes are completed that contain thousands of genes each, more cost-effective and less time-consuming efforts to produce these clones will be needed.

Traditional cloning of a single gene into a vector using restriction sites is a routine procedure for most research laboratories. However, the efficiency of this process varies greatly depending on the organism. For malaria research, particularly in the case of drug and vaccine development, candidate molecules are often cloned and assessed individually. Routine screening procedures, therefore, can be limited to the number of candidates available for testing. Because the highly (80%; A+T) rich genome of P. falciparum (Triglia and Kemp 1991) often results in a large number of nonrecombinant or rearranged clones, many clones must be screened before a single correct recombinant clone is obtained. The production of 5000 or more gene fragments in one or more plasmid vectors via these traditional methods for use in functional genomics studies in malaria is daunting.

To address this technical hurdle and as a means of providing large-scale reagents to the malaria research community at large, we investigated various methods of producing the large numbers of recombinant molecules for our functional genomics and vaccine studies. The method had to be cost effective, reproducible, scalable, and able to substantially overcome the problems associated with cloning the highly AT-rich P. falciparum genome. One such potential method, the Gateway, does not involve standard nucleic acid digestion and ligation, but instead relies on the well-characterized site-specific recombination process between bacteriophage λ and Escherichia coli (Hartley et al. 2000; Walhout et al. 2000). This recombinant-based cloning provides for highly efficient and accurate directional cloning. This technique has been used recently to study protein-protein interactions in Caenorhabditis elegans (Li et al. 2004) and protein subcellular localization (Simpson et al. 2000). As part of our vaccine development program, we have successfully used the Gateway system for the high-throughput cloning of P. falciparum open reading frames (ORFs) and the production of DNA vaccines and other reagents.

RESULTS AND DISCUSSION
Recombinational Cloning and Analysis
We selected a set of genes from P. falciparum, to clone into two Gateway-compatible master vectors: the pDONR/Zeo and pDONR207 plasmids, and three Gateway-compatible destination vectors: a modified DNA vaccine destination vector VR1020-DV and two protein expression destination vectors, pDEST17 and pMAL-2c. These studies were designed to determine the feasibility of developing a robust, high-throughput cloning system for P. falciparum with the intention of expanding the scope to all identified ORFs, cDNA sequences, or gene models. Figure 1 illustrates the overall flow used to accomplish these tasks, indicating specific procedures as well as the cloning efficiency achieved.

The two panels of genes described here were limited to those between 200 and 5500 bp in length to maximize the cloning efficiency. All genes from Panel A were selected prior to the full annotation of the P. falciparum genome, and therefore only single-exon genes were used. The second set of genes, Panel B, was selected later using bioinformatics and comparative genomics tools aimed at identifying antigens involved in pre-erythrocytic immunity mainly expressed in sporozoite and/or liver stages (Doolan et al. 2003a). Both spliced and single-exon ORFs were selected; however, because a cDNA source for these two stages is scarce, we decided to clone the largest exon for most of the spliced genes to allow for amplification from genomic DNA. Panel B comprises 79 single-exon (50%) and 80 multixen genes (50%) as indicated in Supplemental Table 2. The actual amino acid sequences cloned for each ORF are also indicated in the same table. Future studies are planned to clone full-length
spliced ORFs with sporozoite cDNA using the optimized cloning protocol described here. We have also optimized the Gateway system in the *P. yoelii* rodent model with similar cloning efficiency and have used this method to assist in the rapid and large-scale identification and characterization of *P. yoelii* antigens that can protect against a parasite challenge (Haddad et al. 2004).

To maintain maximum throughput, we screened recombinant clones by colony PCR and DNA sequencing. Four colonies per BP reaction were picked for screening to increase the chances of obtaining clones without mutations. A typical agarose gel of the PCR screening is shown in Figure 2. The screening of Figure 2A showed that of the 111 ORFs screened, 95 were successfully cloned (86%), that is, they had at least one clone positive, with 65% of the reactions generating four or more positive clones (Table 1). Similar cloning efficiency (83%) was observed for Figure 2B (159/192). Negative clones in the *P. falciparum* system were mainly caused by no colonies produced or by internal deletions as shown in Figure 2 (lane b, clone PFC0275w) and/or no recombination. DNA sequencing of the master clones was optimized to avoid the predicted secondary structures known to occur with Gateway att recombination sequences. DNA sequencing from both 5'- and 3'-junctions were successfully obtained for all master clones confirming the identities of all cloned ORFs. To evaluate the frequency of potential mutations introduced during cloning, we have completely sequenced 23 genes from Panel B. The analysis of the double-strand sequences indicated that no nucleotide variation was observed in 18 genes when compared with the published *P. falciparum* 3D7 sequences. However, we detected mutations on the remaining five genes: one gene contained a single silent mutation whereas the other four had a single missense mutation each. The sequencing of these 23 genes totaled 14,319 nucleotides, which yielded a 1 in 2864 discrepancy rate. Therefore, the use of high-fidelity PCR enzyme, the lack of major DNA deletions observed during PCR screening of all master clones, and the absence of frameshift mutations in the completely sequenced subset suggest that this clone set is accurate. The size range of uncloned ORFs (687–5556 bp) and cloned ORFs (198–3450 bp) overlapped (Fig. 3). Statistical analysis indicated, however, that the mean size of the cloned genes was significantly smaller than the mean size of the uncloned genes (independent samples *t*-test, two-tailed, *p* < 0.0009). We surmise that the Gateway system is better at cloning genes smaller than 5 kb although feasible attention may need to be paid toward developing methods for cloning larger genes. All entry clones from both Panels A and B have been deposited at the Malaria Research and Reference Reagent Resource Center (http://www.malaria.mr4.org) for distribution to researchers upon request. Further entry clones will be subsequently deposited at MR4 as they are generated. The identity of the available clones will also be displayed at the PlasmoDB Web site (http://www.PlasmoDB.org).

![Figure 2](image-url)  
**Figure 2** PCR screening BP and LR reactions. (A) Recombinant clones in the pDONR/Zeo master plasmid were screened by colony PCR using plasmid-specific primers (M13 forward and reverse). The figure shows amplified DNA products from 12 BP reactions representing different genes in groups of four colonies (a, b, c, d) per clone analyzed on a 1% agarose gel. (B) The screening of recombinants clones in the VR1020-DV destination vector was done by PCR on DNA from 4 single colonies (a, b, c, d) as well as bulk cultures (B) from the same transformation. Clones shown were picked at random. (M) 1-kb DNA extension ladder.

### Table 1. Screening Efficiency of Gateway Cloning

<table>
<thead>
<tr>
<th>Panel of selected ORFs</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF size range (bp)</td>
<td>500–3000</td>
<td>200–5500</td>
</tr>
<tr>
<td>Number of ORFs screened</td>
<td>111</td>
<td>192</td>
</tr>
<tr>
<td>4/4 positive colonies</td>
<td>72</td>
<td>108</td>
</tr>
<tr>
<td>3/4 positive colonies</td>
<td>11</td>
<td>28</td>
</tr>
<tr>
<td>2/4 positive colonies</td>
<td>06</td>
<td>15</td>
</tr>
<tr>
<td>1/4 positive colonies</td>
<td>06</td>
<td>08</td>
</tr>
<tr>
<td>0/4 positive colonies</td>
<td>16</td>
<td>33</td>
</tr>
<tr>
<td>Cloning efficiency (%)</td>
<td>86</td>
<td>83</td>
</tr>
</tbody>
</table>

The screening for recombinant master clones was performed by colony PCR on 4 colonies per each BP (gene) reaction. A positive colony represented the correct amplified gene size analyzed on agarose gel. DNA sequencing using plasmid-specific primers later confirmed the identity of the cloned genes.

---

**Table 1. Screening Efficiency of Gateway Cloning**

<table>
<thead>
<tr>
<th>Panel of selected ORFs</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF size range (bp)</td>
<td>500–3000</td>
<td>200–5500</td>
</tr>
<tr>
<td>Number of ORFs screened</td>
<td>111</td>
<td>192</td>
</tr>
<tr>
<td>4/4 positive colonies</td>
<td>72</td>
<td>108</td>
</tr>
<tr>
<td>3/4 positive colonies</td>
<td>11</td>
<td>28</td>
</tr>
<tr>
<td>2/4 positive colonies</td>
<td>06</td>
<td>15</td>
</tr>
<tr>
<td>1/4 positive colonies</td>
<td>06</td>
<td>08</td>
</tr>
<tr>
<td>0/4 positive colonies</td>
<td>16</td>
<td>33</td>
</tr>
<tr>
<td>Cloning efficiency (%)</td>
<td>86</td>
<td>83</td>
</tr>
</tbody>
</table>

The screening for recombinant master clones was performed by colony PCR on 4 colonies per each BP (gene) reaction. A positive colony represented the correct amplified gene size analyzed on agarose gel. DNA sequencing using plasmid-specific primers later confirmed the identity of the cloned genes.

---

The analysis of the double-strand sequences indicated that no nucleotide variation was observed in 18 genes when compared with the published *P. falciparum* 3D7 sequences. However, we detected mutations on the remaining five genes: one gene contained a single silent mutation whereas the other four had a single missense mutation each. The sequencing of these 23 genes totaled 14,319 nucleotides, which yielded a 1 in 2864 discrepancy rate. Therefore, the use of high-fidelity PCR enzyme, the lack of major DNA deletions observed during PCR screening of all master clones, and the absence of frameshift mutations in the completely sequenced subset suggest that this clone set is accurate. The size range of uncloned ORFs (687–5556 bp) and cloned ORFs (198–3450 bp) overlapped (Fig. 3). Statistical analysis indicated, however, that the mean size of the cloned genes was significantly smaller than the mean size of the uncloned genes (independent samples *t*-test, two-tailed, *p* < 0.0009). We surmise that the Gateway system is better at cloning genes smaller than 5 kb although feasible attention may need to be paid toward developing methods for cloning larger genes. All entry clones from both Panels A and B have been deposited at the Malaria Research and Reference Reagent Resource Center (http://www.malaria.mr4.org) for distribution to researchers upon request. Further entry clones will be subsequently deposited at MR4 as they are generated. The identity of the available clones will also be displayed at the PlasmoDB Web site (http://www.PlasmoDB.org).

Once each of the master clones was obtained, generation of the subsequent destination clones was highly efficient and straightforward. Large-scale subcloning LR reactions were successfully set up in a high-
throughput format. A single colony of each LR reaction from each of three destination vectors (VR1020-DV, pDEST17, and pMAL-2c) was screened by PCR. Initial studies showed that 100% of the selected colonies from the destination vector reaction were positive and with no observed rearrangements or deletions, and the identity of each clone was confirmed by DNA sequencing.

To increase the throughput of the cloning process we further optimized the subcloning of the ORFs by analyzing VR1020-DV clones derived from bulk cultures of transformed DH5α competent cells. This modification eliminated the need to plate each LR-transfected reaction and reduced handling errors and the time-consuming screening step. This high-throughput screening approach was also very efficient (93%; Fig. 2) and comparable to the traditional single-colony screening (100%) method.

Gene Expression

The VR1020 plasmid chosen as an expression vector for the development of DNA vaccines, expresses genes as fusion proteins with the tPA signal peptide at the N terminus. The conversion of the traditional single-colony screening (100%) method.

We set out to generate antibody reagents to individual antigens and determine their expression in the malaria parasite using a high-throughput approach. For this purpose, we immunized groups of CD-1 outbred mice with individual Gateway VR1020-DV DNA vaccine constructs selected from Panel A. Because of the large number of sera to be analyzed, we performed the initial screen on group-pooled sera collected 2 wk after the second and third immunizations. The antibody screening was performed by the IFAT on three parasite stages, sporozoite and a mix of mature asexual erythrocytic and sexual erythrocytic stages. The screening of the pooled and individual mice sera collected after three doses resulted in 21 positives out of 95 genes (22%) examined (Fig. 4). A majority of these genes induced antibodies that reacted with erythrocytic stages (19/21). A list of the positive clones is provided as Supplemental Table 1. These results were consistent with the stage expression profiles determined by analysis of the P. falciparum transcriptome and proteome (Florens et al. 2002; Le Roch et al. 2003), where most (17/19) of the positive genes had high expression levels during the erythrocytic stages of development, both at RNA and protein profiles. All the positive genes for the gametocyte (three) and sporozoite (three) stages had high levels of RNA detected in these respective stages in the study of the transcriptome (Le Roch et al. 2003). The remaining 74 DNA constructs that failed to induce antibodies to parasites generally were not identified in the proteome analysis in the parasite stages evaluated; only 24 of these genes had peptides detected from any parasite stage examined. These results may reflect the fact that genes from Panel A were selected prior to publication of large-scale gene expression profiles. This correlation may indicate that the use of expression data as determined by proteome and transcriptome studies will be of significant value for predicting genes that will most likely induce an immune response and be useful in a reverse vaccinology approach.

Figure 3 Size distribution of cloned and noncloned ORFs. A total of 303 amplified P. falciparum ORFs were screened for cloning into the Gateway entry plasmid by PCR and DNA sequencing. The horizontal lines to the right of the data points are the mean and the upper and lower 95% confidence limits.

Figure 4 shows examples of the stage specificity and the cellular localization of novel proteins observed by IFAT. It is possible that some of these genes showing no expression in the sporozoite and erythrocytic stages could be inducing antibodies to other stages of the parasite life cycle not tested here, including those inside of the mosquito (oocysts and ookinetes) and in the vertebrate host hepatocyte.

Preliminary attempts to express genes from Panel A as recombinant proteins in GST- and MBP-fusion expression constructs had limited success; only seven out of 95 fusion proteins were expressed, and five of these were determined soluble. These findings are in contrast with protein expression efficiency shown for Gateway constructs encoding human ORFs using the same expression vectors (Braun et al. 2002). Subsequent to this effort, we have attempted to express recombinant proteins from 197 genes from Panel B, cloned into a GST-fusion expression plasmid. The screening indicated that 32 genes (16%) produced recombi-

<table>
<thead>
<tr>
<th>Construct</th>
<th>Dose 2</th>
<th>Dose 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCPSP–VR1020</td>
<td>40 (0–160)*</td>
<td>2560 (0–120)*</td>
</tr>
<tr>
<td>PCPSP–VR1020-DV</td>
<td>40 (0–160)*</td>
<td>1280 (0–2560)*</td>
</tr>
<tr>
<td>PyCSP–VR1020</td>
<td>2560 (0–10,240)*</td>
<td>5120 (0–20,480)*</td>
</tr>
<tr>
<td>PyCSP–VR1020–DV</td>
<td>1280 (0–5120)*</td>
<td>2560 (0–5120)*</td>
</tr>
</tbody>
</table>

Groups of 10 CD-1 mice were immunized ID with 100 µg of each DNA construct. Sera were collected 2 wks after doses 2 and 3.

*Sera tested by IFAT on air-dried P. falciparum sporozoites.

**Sera tested by IFA on air-dried P. yoelli sporozoites.
baculovirus (Liang et al. 2000), Pichia (Zou et al. 2003), and the use of codon optimized transgenes (Zou et al. 2003), the viability of these approaches depends on the availability of a high-throughput format. Codon optimization has routinely improved protein expression, although the cost of synthesizing large numbers of codon-optimized genes limits their utility in high-throughput cloning efforts. Improved methods of expressing P. falciparum proteins will continue to be required despite the methods of producing large numbers of recombinant P. falciparum molecules.

In conclusion, we demonstrated that the Gateway cloning system provides a highly efficient method to produce recombinant plasmids containing genes from P. falciparum. Noteworthy is the fact that most of the P. falciparum DNAs, which are known to be recombination-prone, were cloned without deletions in the ORF. Using a limited number of ORFs and the DNA vaccine strategy, we have been able to further explore the functional profile and identify the stage-specific expression of novel antigens with antisera produced in mice. These data establish the Gateway system as a platform technology, which provides for the rapid and flexible design and construction of recombinant plasmids needed for functional genomics studies and for vaccine and drug development efforts. Using this system, one will be able to perform high-throughput functional assays using a variety of expression vectors, that is, protein expression vectors to generate reagents for protein microarray experiments and animal immunization, conduct large-scale transfection experiments to assay protein localization, determine immunogenicity of candidate antigens, develop expression vectors for Y2H systems to screen for host receptors and parasite protein interactions, and develop large numbers of recombinant virus constructs that can be used in immune assays for screening protective antigens (Doolan et al. 2003a) and may allow for the functional identification of the large number of unknown and hypothetical proteins in the P. falciparum genome.

**METHODS**

**Selection and PCR Amplification of ORFs**

Different ORF selection criteria were applied to evaluate the high-throughput feasibility of the Gateway technology across multiple P. falciparum genes. The first panel of genes (Panel A) was comprised of 111 single-exon ORFs from early Chromosomes 2 and 3 annotations and ranged in size from 0.3 to 3 kb. The genes in Panel B were selected based on a set of credentialed data for their potential as vaccine targets (Supplemental Table 2). A separate functional genomics approach led to the creation of each of the six categories. The first category, AffyCluster1 (n = 38), is genes that were found to be up-regulated in a transcriptional cluster only in sporozoite stages (Le Roch et al. 2003). The High MudpIT SPZ category (n = 7) contains sporozoite-expressed genes identified by mass spectrometry found to have high peptide sequence coverage (Florens et al. 2002). The Bioinformatics category (n = 30) is comprised of genes with predicted S’-signal sequences and 3’-transmembrane domain(s) also possessing evidence of protein expression in sporozoite stages. The PyLCM category (n = 65) contains P. falciparum orthologs of P. yoelii liver-stage genes identified from an EST sequencing project (J.B. Sacci, J.M.C. Ribeiro, F. Huang, U. Alam, J. Russell, P.L. Blair, A. Witney, D.J. Carucci, A.F. Azad, J.C. Aguiar, et al., unpubl.) not identified by mass spectrometry. The Immunosense category (n = 6) is comprised of genes with encoded proteins originally identified in the sporozoite proteome and also recognized by volunteers immunized with the irradiated sporozoites (Doolan et al. 2003b). The PRBC Surface group (n = 13) contains genes whose proteins are candidates for being expressed on the surface of infected erythrocytes (Florens et al. 2004). In addition, this panel was predomin-
nantly selected to explore ORFs with unknown function (annotated as "hypothetical") for downstream determination of potential function. Panel B was comprised of either single-exon genes or the major exon from genes with predicted multixen gene structures.

The ORFs from Panel B varied in size from 0.2 to 5.5 kb. An entire list of cloned genes is provided in Supplemental Table 2. A customized computer script was used with the automated computational program Primer 3 (http://www.broad.mit.edu/genome_software/other/primer3.html) to specifically design gene-specific oligonucleotides used for the amplification of Gateway-compatible in-frame full-length ORFs. An additional 12 bases of overlapping sequence at the 5′-end to the universal outer adapter attB primers were used to design a nested PCR with a universal outer set of primers. The overlapping sequence for the forward and reverse primers were so designed so that the amplification products contained attB sequences at both 5′- and 3′-ends (forward primer overlapping sequence 5′-GCGAGGTCACC-3′ and reverse primer overlapping sequence 5′-AGAAAGCTTGG-3′). The sequence for the attB1 adapter primers was 5′-GGGGAACAGTTTTGAACATCAGAGATTTTGAGACAC-3′ and for the attB2 adapter primers was 5′-GGGGACCACTTTGAAGTTTGTACAAAAAAGCAGGCTCCACC-3′ and for the attB2 adapter primers was 5′-GGGGACCACTTTGAAGTTTGTACAAAAAAGCAGGCTCCACC-3′ and for the attB2 adapter primers was 5′-GGGGACCACTTTGAAGTTTGTACAAAAAAGCAGGCTCCACC-3′ and for the attB2 adapter primers was 5′-GGGGACCACTTTGAAGTTTGTACAAAAAAGCAGGCTCCACC-3′ and for the attB2 adapter primers was 5′-GGGGACCACTTTGAAGTTTGTACAAAAAAGCAGGCTCCACC-3′ and for the attB2 adapter primers was 5′-GGGGACCACTTTGAAGTTTGTACAAAAAAGCAGGCTCCACC-3′.

The nested PCR strategy briefly described below was used to reduce the overall length of hundred of primers, thereby significantly reducing the cost incurred with primer synthesis. PCR amplification as well as the PCR isolation protocols were done in the 96-well format. Full-length genes were amplified using a high fidelity Platinum polymerase (Invitrogen) and a CMV immediate/early gene promoter in fusion with the tPA (tissue plasminogen activator) leader sequence. The transfer of ORFs from master clones to expression vectors was performed using the LR reaction kit (Invitrogen). Two protocols were used; one recommended by the manufacturer that screened a single-colony for recombinants is described here; the other method described later was adapted to truly high throughput by screening clones in bulk. Briefly, 1 µL of each Master clone was incubated with 200 ng of Topoisomerase I-treated VR1020-DV plasmid in a 20-µl reaction containing the LR reaction agents. After proteinase K treatment, a 2-µL aliquot of each reaction was used to transform DH5α competent E. coli cells and plated on LB plates containing appropriate antibiotic. The screening was performed with one colony per LR reaction. Plasmid DNA was prepared as described above and screened by PCR and DNA sequencing using plasmid-specific primers.

Large-Scale Recombinant Production

In parallel to the single-colony approach described above, we tested an alternative screening for these LR reactions. All steps for this procedure were performed in a 96-well platform. Following the LR reaction, bacteria growth media (LB broth) and kanamycin was added to a fraction of the transformed DH5α cells and cultured overnight. DNA was prepared from these bulk cultures and screened by PCR as described above. PCR amplicons were analyzed in an agarose gel running in parallel with PCRs from single colony screening.

DNA Vaccine Construction

To generate DNA vaccine constructs we used the VR1020-DV plasmid as the destination vector. The original DNA vaccine VR1020 plasmid (Vical Inc.; Luke et al. 1997) was modified to be compatible with the Gateway cloning system. The expression of vaccine antigens is driven by a CMV immediate/early gene promoter in fusion with the tPA (tissue plasminogen activator) leader sequence. The transfer of ORFs from the master clone to VR1020-DV was performed via the LR reaction.

Recombinant Protein Expression

In an attempt to express the proteins of cloned genes in E. coli, we subcloned the ORFs into two recombinant protein expression vectors: pDEST17 and pMAL-c2. The pDEST17 vector (Invitrogen) is designed for expressing fusion proteins containing 6 histidines at the N terminus. Sequences transferred into pDEST17 from a master clone are expressed using T7 RNA polymerase under control of a salt-inducible promoter. In the pMAL-c2 vector, the gene of interest is fused to maltose-binding protein (MBP), and expressed using the strong tac promoter and translation initiation signals of MBP. Also pMAL-c2 contains the normal malE signal sequence, which directs the fusion protein through the cytoplasmic membrane. The pMAL-c2 vector was converted to be compatible with the Gateway cloning system (Braun et al. 2002). The same protocol described for the LR reaction and screening used to transfer genes to the VR1020-DV vector was used for these protein expression vectors.

Mouse Immunization and Antibody Testing

Cloning grade DNA was purified (QBIogene) for mouse immunization. Groups of five outbred CD-1 mice (Charles River Laboratories) were immunized three times with a 100-µg dose of each DNA construct at 3-wk intervals. The immunizations were done by the intradermal route (ID) at the base of the tail. Sera were collected 2 wk after the third dose, and antibody reactivity was assessed by standard immunofluorescence assay (IFAT) against air-dried sporozoites and erythrocytic stages of P. falciparum as described elsewhere (Aguir et al. 2001).
ACKNOWLEDGMENTS

We are grateful to Trevor Jones and Annabelle Wright from WRAIR-NMRC for guidance with statistics and graphics. Many thanks to Kristi Strang, Allison Halleck, Steven Lloyd, and Pascal Braun from the Harvard Institute of Proteomics (HIP) for their assistance with cloning and supplying the pMAI-2c destination plasmid. We also thank Steven Max, Dolores Ciufio, Darby Cortijo, Grazzella Pitas, and Gary Temple from Invitrogen for their guidance with the Gateway system. The work was supported by funds allocated to the Naval Medical Research Center by the US Army Medical Research Material Command (61102A.S13.C.A0078 and 6000.RAD1.C.A0305) and by the Office of Naval Research (Navy ONR-Core 61153M.M458.S18.A0242). The views of the authors expressed herein are their own and do not purport to reflect those of the US Navy or the US Department of Defense. The experiments reported herein were conducted according to the principles set forth in the Guide for the Care and Use of Laboratory Animals, Institute of Laboratory Animal Resources, National Research Council, National Academy of Sciences Press, 1996.

REFERENCES


Web Site References

http://www.broad.mit.edu/genome_software/other/prime3.html; Primer 3.

http://www.malaria.mrc.ac.uk/ Malaria Research and Reference Reagent Resource Center.

http://www.PlasmoDB.org/PlasmoDB.

Received June 22, 2004; accepted in revised form June 24, 2004.