Expressed Sequence Tag Analysis of the Erythrocytic Stage of *Plasmodium berghei*

Ji-Woong Seok[1], Yong-Seok Lee[2], Eun-Kyung Moon[1], Jung-Yub Lee[1], Bijay Kumar Jha[1], Hyun-Hee Kong[1], Dong-Il Chung[1] and Yeonchul Hong[1],*

[1]Department of Parasitology, Kyungpook National University School of Medicine, Daegu 700-422, Korea; [2]Department of Parasitology, College of Medicine and Frontier Inje Research for Science and Technology, Inje University, Busan 614-735, Korea

Abstract: Rodent malaria parasites, such as *Plasmodium berghei*, are practical and useful model organisms for human malaria research because of their analogies to the human malaria in terms of structure, physiology, and life cycle. Exploiting the available genetic sequence information, we constructed a cDNA library from the erythrocytic stages of *P. berghei* and analyzed the expressed sequence tag (EST). A total of 10,040 ESTs were generated and assembled into 2,462 clusters. These EST clusters were compared against public protein databases and 48 putative new transcripts, most of which were hypothetical proteins with unknown function, were identified. Genes encoding ribosomal or membrane proteins and purine nucleotide phosphorylases were highly abundant clusters in *P. berghei*. Protein domain analyses and the Gene Ontology functional categorization revealed translation/protein folding, metabolism, protein degradation, and multiple family of variant antigens to be mainly prevalent. The presently-collected ESTs and its bioinformatic analysis will be useful resources to identify for drug target and vaccine candidates and validate gene predictions of *P. berghei*.

Key words: *Plasmodium berghei*, expressed sequence tag, rodent malaria

INTRODUCTION

The rodent malaria parasite *Plasmodium berghei* is similar to human malaria parasites, such as *P. falciparum*, in aspects of the structure, genome organization, physiology, and life cycle [1-3]. Therefore, *P. berghei* represents a practical and relevant model organism for experimental studies of malaria [4]. To improve the utility of models, such as *P. berghei* in the development of drug target and vaccine candidates for malaria, the genome sequence and actual transcripts are required as primary sources of biological information.

The genome of *P. berghei* is organized into 14 chromosomes, with an estimated genome size of 18 Mb [4]. Partial shotgun sequencing of the *P. berghei* genome and transcription profile analysis with genome survey sequences (GSS) is having significant contribution to many fields of malaria research [4]. How-

ever, most gene prediction of *P. berghei* have been based on bioinformatic analyses using computer software. However, the high A/T contents of the *Plasmodium* genome, excluding *P. vivax*, hamper the prediction of the gene structure, resulting about 60% of the predicted genes encoding hypothetical proteins [5]. Therefore, it is necessary to verify the prediction with complementary DNA (cDNA), such as expressed sequenced tag (EST), a short contiguous subsequence of a transcribed DNA sequence, as a rapid means of gene identification to obtain useful information from a genome sequence, especially for intron-containing eukaryotes. Currently, large-scale random sequencing of ESTs is preceding concurrent with the *Plasmodium* genome project [6-9]. Previous efforts to generate ESTs by random clones of a *P. berghei* cDNA library have accelerated the gene discovery processes [7]. The present study constructed a SMART™ PCR-amplified cDNA library from mixed blood stages of *P. berghei* parasites to enrich for full-length transcripts for detection of rare transcripts and transcript isoforms and determination of the relative abundance of transcripts. Here, we report the analysis of the *P. berghei* ESTs, including abundance, prevalence of protein domains, and functional categorization.

MATERIALS AND METHODS

Parasite collection

P. berghei ANKA strain (kindly provided by Dr. Eun-Taek Han, Department of Parasitology, Kangwon National University) was used to infect 6-week-old CL57B/6 mice. The blood stage of the parasite was used for cDNA library construction. Blood was collected by heart puncture under anesthesia and leukocytes were obtained using Plasmodipur leukocyte filters (Euro-Diagnostica, Malmö, Sweden). Parasites were released from their host RBCs by 0.15% saponin (0.5 volume of packed RBCs) (Sigma-Aldrich, St. Louis, Missouri, USA) in PBS, pH 7.5 (PBS) and agitated for 1-2 min until the suspension became a clear red color. The suspension was diluted by addition of 15 volumes of PBS, and the released parasites were collected by centrifugation [10].

Construction of *P. berghei* cDNA library

For construction of the *P. berghei* cDNA library, a PCR-based cDNA library was used with total RNA purified with TRIzol reagent (Gibco BRL, Rockville, Maryland, USA) following the instructions for the SMART cDNA library construction kit (BD-Clontech, Palo Alto, California, USA). cDNA was synthesized with a specially designed oligonucleotide (SMART IV) in the first-strand synthesis to generate high yields of full-length, double-stranded cDNA and 3′ primer. Second-strand synthesis was performed by a long-distance PCR with Advantage 2 polymerase mix (Clontech). PCR products were extracted with phenol : choloroform (25 : 24) to remove the polymerases, digested with S*fi*I, and size-fractionated using a ChromaSpin-400 column (Clontech) to exclude cDNAs < 500 bp. The cDNA mixture was ligated into the λ TriplEx2 vector (Clontech) and packaged using the GigaPack III Plus packaging extract (Stratagene, La Jolla, California, USA) according to the manufacturer's inst ructions.

In vivo excision and random sequencing

The titer and percentage of recombinant phages in the library was determined to 1×10^8 plaque forming units with 95% as recombinant clones. *Escherichia coli* strain BM25.8 cells were transduced with recombinant phage, from which the massive excision of the pTriplEx2 phagemid library was accomplished according to the manufacturer's instruction (Clontech). After in vivo excision, bacterial colonies were randomly selected and grown in LB-ampicillin broth by incubation with shaking at 31˚C overnight. Then, plasmids from selected colonies were extracted using the DNA-spin Plasmid DNA Purification Kit (iNtRON Biotechnology, Seoul, Republic of Korea) and sequenced with a PE377 DNA sequencer (Perkin-Elmer, Boston, Massachusetts, USA) using the Bigdye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Foster City, California, USA).

Bioinformatic analysis

The ESTs were initially analyzed with well-established procedure for EST sequence processing and annotated using the PESTAS automated EST analysis platform (http://pestas.kribb. re.kr) [11-13]. Each EST cluster was analyzed using BLASTX against the GenBank non-redundant protein database (April 2010 release) and *Plasmodium* annotated protein database in PlasmoDB (ver. 7.1, released November 2010, http://plasmodb. org/plasmo/) with an E-value of < 10^{-5} for selection of matching [14]. After the first assignment, a BLASTN and TBLASTX search of the unmatched EST clusters was performed against the *P. berghei* EST and genome database in PlasmoDB to ascertain whether they were encoded in the *P. berghei* genome as putative new transcripts. EST cluster-associated GO terms were functionally classified based on protein-level annotation using BLAST2GO (cut-off ≤ $1e^{-10}$) [15]. Functional domains in novel clusters were assigned using InterProScan (HMMPfam, HMM-Smart, HMMTigr, HMMPanther, and Superfamily, flagged as true by InterProScan with E-value < $1e^{-2}$) [16]. All of the *P. berghei* ESTs generated from this study were submitted to the db-EST division of GenBank with accession numbers (HS576390-HS586433). Based on our ESTs, a specific *P. berghei* EST database (*P. berghei* EST DB) was constructed (http://parasite.knu. ac.kr).

RESULTS

P. berghei ESTs

The 12,000 clones containing DNA inserts were sequenced, sequence < 100 bp were removed, and the remainder was processed with bioinformatic software programs to generate high quality ESTs. First, total ESTs were aligned against a non-redundant database of mouse gene for exclusion of mouse DNA contamination and 4 ESTs displayed encoding mouse genes (0.04%). A total of 10,040 ESTs having an average length of 643 bp and 74% [A+T] content were produced (Table 1). Cluster analysis with the processed ESTs, using TGICL, assembled

10,040 ESTs into 2,462 EST clusters with 1,432 contigs containing at least 2 or more overlapping sequences and 1,030 ESTs remained as singletons. The sequences of assembled contigs could be up to 2.5 kb in length and were composed of an average of 6.3 ESTs. The 2,462 EST clusters were compared with the *P. berghei* annotated protein database in PlasmoDB; 2,043 (83%) EST clusters were annotated with *P. berghei* proteins showing significant BLASTX matching at the cutoff value of $<1e^{-5}$ with 419 EST clusters remaining unmatched. From BLASTX analysis, we found that 244 genes with predicted coding regions were fully covered by EST clusters.

After the first assignment, a BLASTN and TBLASTX search of the 419 unmatched EST clusters was performed against *P. berghei* EST databases in PlasmoDB. Of the 419 EST clusters, 371

Table 1. Transcriptome features of *Plasmodium berghei* EST

	Numbers
Total number of clones	12,000
Number of ESTs	10,040
Average length of the ESTs (nt)[a]	643
Number of EST clusters	2,462
Contigs	1,432
Singletons	1,030
Matches to *P. berghei* DB	
Clusters to proteins	2,043
Clusters to ESTs	371
Clusters to DNA	48

[a]nt, nucleotide.

(88.5%) displayed matching to *P. berghei* EST in the databases (Table 1). The 48 unmatched EST clusters were aligned using a BLASTN and TBLASTX analysis against a non-redundant protein database at the National Center for Biotechnology Information (NCBI) and the *P. berghei* genome database to ascertain whether they were encoded in the *P. berghei* genome. Corresponding sequences were apparent with 48 (11.5%) of these non-matched ESTs, most of them were hypothetical proteins with unknown function, implicating these EST clusters as putative new transcripts in *P. berghei* (Table 1). These results support the view that the *P. berghei* protein database remains incomplete.

Abundant *P. berghei* ESTs

We examined the redundancy of EST clusters, because redundant EST appears to reflect the highly expressed genes, which can highlight the importance of the genes in their respective biological pathways. The most abundantly detected transcripts (i.e., EST clusters containing more than 50 ESTs) are summarized in Fig. 1. Many of them corresponded to ribosomal, hypothetical, membrane proteins, or proteins in the purine salvage pathway.

Protein domains in *P. berghei* ESTs

We further analyzed EST clusters with Pfam (http://pfam.sanger.ac.uk) to catalog the protein domains present in the *P.*
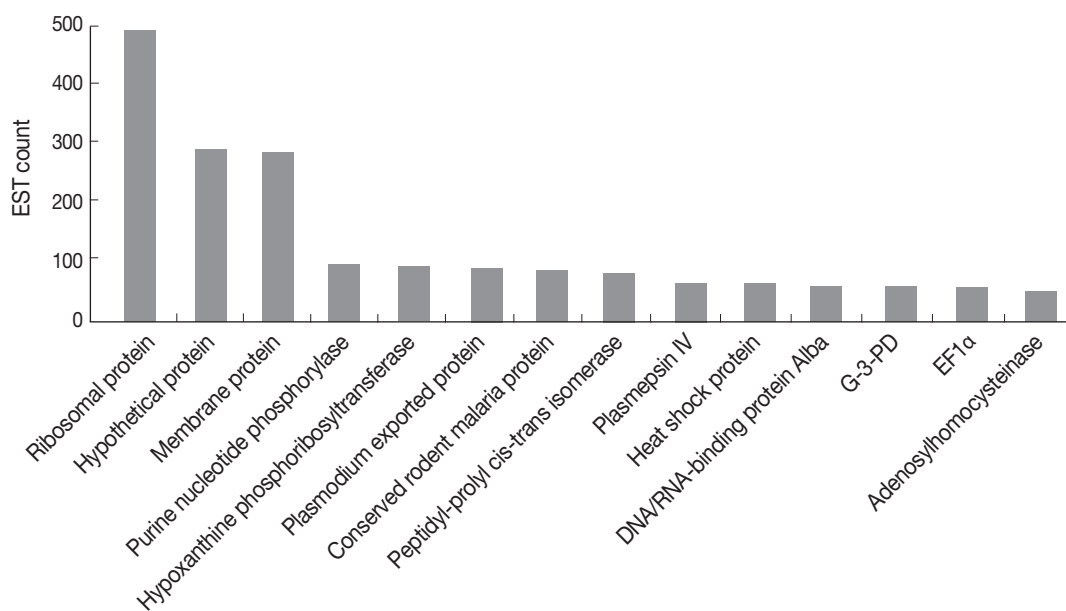

Fig. 1. The abundant transcripts in *Plasmodium berghei*.

Table 2. The prevalence of protein domains in *P. berghei* ESTs

Protein family	Pfam accession No.	Rank	No. of ESTs
RNA recognition motif domain	PF00684	1	36
Proteasome, subunit alpha/beta	PF00276	2	32
Variant antigen yir/bir/cir	PF01849	3	26
Chaperonin Cpn60/TCP-1	PF00009	4	24
Ubiquitin-conjugating enzyme, E2	PF01779	5	20
ATPase, AAA-type, core	PF11940	6	16
Heat shock protein 70	PF01020	7	14
Proteasome, alpha-subunit, conserved site	PF03939	7	14
Histone core	PF00056	8	12
DNA/RNA helicase, DEAD/DEAH box type, N-terminal	PF00137	8	12
Serine/threonine-protein kinase-like domain	PF02136	8	12
WD40 repeat, subgroup	PF08282	8	12
Protein synthesis factor, GTP-binding	PF00009	9	10
Ras	PF00333	9	10
Cytoadherence-linked asexual protein	PF01248	9	10
Pathogenesis-related transcriptional factor/ERF, DNA-binding	PF00252	9	10
Like-Sm ribonucleoprotein (LSM) domain	PF00012	9	10
Protein phosphatase 2C, N-terminal	PF01918	10	8
ABC transporter-like	PF03144	10	8
Peptidase C19, ubiquitin carboxyl-terminal hydrolase 2	PF00481	10	8
Mitochondrial substrate/solute carrier	PF02773	10	8
Heat shock protein DnaJ, N-terminal	PF00240	10	8
Helicase, C-terminal	PF00883	10	8

berghei EST datasets, because the identification of domains that contain within proteins, especially hypothetical proteins, can provide insights into their functions [17]. The prevalence of protein domains in *P. berghei* ESTs is summarized in Table 2 showing RNA recognition motifs (RRM; PF00684), which contained the RNA binding protein implicated in regulation of splicing, RNA stability, and translation, to be most prevalent. Proteasome, subunit alpha/beta (PF00276), variant antigen Yir/Bir/Cir (PF01849) and chaperonin Cpn60/TCP-1 (PF00009) are among the top 10 Pfam families in the ESTs. These results together with previous results, the redundancy of EST clusters, indicate that proteins in the asexual blood stages of *P. berghei* are mainly related in translation/protein folding and degradation. Human malaria parasites evade the host immune response through the members of multigene families, such as Var, Rif, and Stevor, encoding virulence determinants of cytoadhesion and antigenic variation. In rodent malaria parasites (*P. yoelii, P. berghei,* and *P. chabaudi*), a large paralogous multigene family of variant antigens, Yir/Bir/Cir, is also conserved [18]. Consistent with their importance, variant antigen Yir/Bir/Cir (PF01849) displayed a significant portion in the prevalence of protein domains in *P. berghei* ESTs.

Functional categorization of *P. berghei* ESTs

The EST clusters were grouped as functional categories based on GO molecular functions. GO, which consists of 3 major ontologies, i.e., biological process, molecular functions, and cellular components, is the most widely used method to predict gene families and functions of EST sequences. The BLAST-2GO program was used in the functional classification of the *P. berghei* ESTs. From this, 1,631 (66.2%) EST clusters were assigned to biological processes (523; 21.2%), cellular components (377; 15.3%), and molecular functions (731; 29.7%) (Fig. 2). Consistent with our expectation, the majority of the genes with functional assignments were related to translation/protein folding, ribosomal structure, and metabolism. In particular, EST clusters were classified into proteolysis (GO: 000-6508), including berghepain-2, a falcipain-2 homologue in *P. berghei*, plasmepsin, and many aminopeptidases has a significant proportion in biological processes. Interestingly, aminopeptidases, especially methionine aminopeptidases (MetAP), were more frequently detected compared with other proteinases, indicating the exuberant expression of MetAP in *P. berghei*. Aminopeptidases have been suggested as new targets for anti-malarial drug development [19]. In *P. falciparum*, 4 methionine aminopeptidases are expressed among the 9 identified
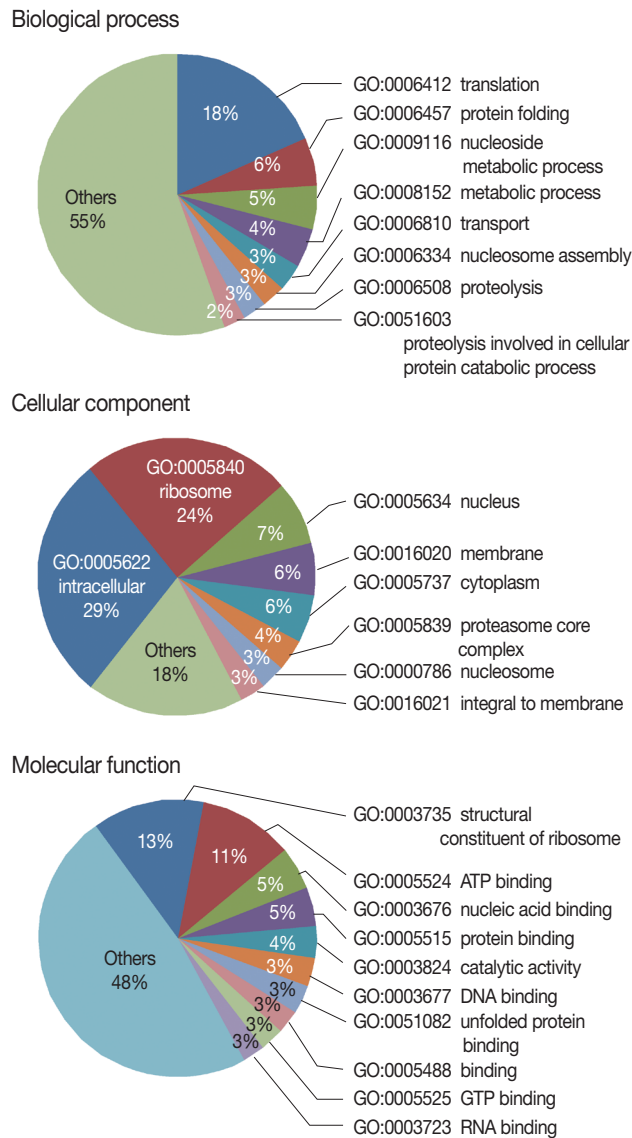
Biological process

Cellular component

Molecular function

Fig. 2. Gene ontology mapping for *P. berghei* EST clusters using BLAST2GO. The genes were functionally categorized based on the Gene Ontology Consortium. Level 3 of the assignment results are shown.

aminopeptidases [20]. An inhibitory compound against MetAP, XC11 was active against both chloroquine sensitive and resistant *P. falciparum* 3D7 in culture and *P. berghei* in mice, implicating MetAP as an important drug target for anti-chloroquine resistant malaria [21].

The substantial proportions of transport proteins (GO: 0006810, 3%) together with intracellular protein transport proteins (GO: 0006886, 1.5%, data not shown), and vesicle-mediated transport proteins (GO: 0016192, 1.3%, data not shown) were indicative of the importance of intracellular and extracel-

lular trafficking of proteins in this pathogenic parasite. Moreover, as consistent with Fig. 1, chaperones (GO: 0031072) constituted a significant proportion in the class of molecular functions (1%, data not shown) and also chaperonin (PF00009) and HSP70 (PF01020) are frequently detected domains in *P. berghei* ESTs (Table 2). Therefore, these results suggested that protein trafficking is essential for survival of *Plasmodium* in blood stages and a promising drug target to combat against human malaria.

DISCUSSION

The availability of genome, transcriptome, and proteome data of *Plasmodium* spp. has greatly advanced the understanding of the biology of these organisms. However, the high A/T content in the *P. berghei* genome hampers prediction of open reading frames or identification of target genes. Therefore, this large EST collection can provide high quality data regarding coding sequences and expressed gene profiles. As the first expression profile analysis of *P. berghei*, 5,582 ESTs and 5,482 GSSs were functionally classified [7]. Thereafter, the transcription profile of asexual stage of *P. berghei* was analyzed by hybridization to a *P. berghei* GSSs amplicon DNA microarray categorizing into the 4 strategies of gene expression, such as housekeeping, host-related expression, strategy-specific expression, and stage-specific expression. In his study, 10,040 ESTs enriched in intact 5′ ends from *P. berghei* cDNA library were assigned to functional categories based on GO using BLAST2GO revealed the expressed gene profile of *P. berghei* during asexual blood stages. The redundancy of EST clusters and the prevalence domain analysis of *P. berghei* proteins could provide clues for determining their functions and their importance in metabolic pathways. Among the highly abundant transcripts (Fig. 1), the enzymes engaged in nucleotide metabolism, purine nucleotide phosphorylase (PNP), and hypoxanthine phosphoribosyltransferase (HPRT) were well detected by multiple ESTs. *Plasmodium* spp. are unable to synthesize purine de novo and alternatively rely on the salvage pathway with host purines. Hypoxanthine, a primary source of purine, is produced by PNP or in human serum and converted into inosine monophosphate (IMP) by HPRT. Immunocillin-H, a PNP transition state analogue, inhibits *P. falciparum* growth by inhibiting PNP [22, 23]. The significant dependence on HPRT for nucleotide synthesis was presently reflected by its abundance in the ESTs. The results are consistent with the focus on HPRT as a promising

drug target for the development of anti-malarial therapies, by virtue of its different characteristics from host protein [24].

The immunosuppressants FK506 and rapamycin have anti-malarial properties by virtue of binding to the target FK506-binding protein (FKBP) having peptidyl-prolyl *cis-trans* isom-erase activity. However, their mechanisms of action against ma-laria parasites are unclear [25]. In *P. falciparum*, PfFKBP35 with peptidyl-prolyl *cis-trans* isomerase activity has been reported [26]. PfFKBP35 is inhibited by FK506, rapamycin, and calineu-rin, although the latter is independent of FK506 binding. The immunosuppressive peptide cyclosporin A also inhibits the growth of malaria parasites, presumably by binding to cyclo-philins (distinct intracellular prptidyl-prolyl *cis-trans* isomer-ase) [27]. Peptidyl-prolyl *cis-trans* isomerase activity that is com-pletely inhibited by cyclosporin A but not by FK506 or rapa-mycin has been detected in extracts of *P. falciparum* [27]. These results support the suggestion that *P. falciparum* probably con-tains more cyclophilins. Peptidyl-prolyl *cis-trans* isomerase dif-fering from the PfFBPR35 homologue was highly abundant in the *P. berghei* ESTs (Fig. 1).

Similar with the abundance of heat shock protein (HSP) from *P. vivax* as evident from EST analysis [28], HSP constitut-ed 1.34% of all *P. berghei* ESTs. HSP70 (0.8%), 1 of the 2 ma-jor HSPs (HSP90 and HSP70), was more abundant compared to HSP90 (0.3%) in the *P. berghei* library. The importance of HSP70 as a molecular chaperone concerning temperature chan-ges between vector and host, and protein trafficking, has made the protein an important potential anti-malarial drug target. The semisynthetic Hsp90 inhibitor (17-[allylamino]-17-deme-thoxygeldanamycin) that is active against *Plasmodium* HSP90 is effective in attenuating parasite growth and prolonging sur-vival in a mouse model of malaria [29].

Malaria parasites possess a relict plastid called the apicoplast that is homologous to the chloroplast of plants. The apicoplast contains the capacity for besides basic metabolic processes such as protein translation, and the biosynthesis of fatty acids, iso-prenoids, iron-sulphur clusters and heam, which are essential for parasite survival. However, fewer than 50 proteins are en-coded for in the apicoplast genome; the vast majority of meta-bolic pathway related proteins are encoded in the nuclear ge-nome and are subsequently transported to the apicoplast [30, 31]. Interestingly, the transport machinery of apicoplast target-ing proteins is similar with that in the chloroplast; the translo-con of the outer envelope of chloroplast (TOC) and translo-con of inner envelope of chloroplast (TIC) complexes are as-sumed to promote protein transport [32,33]. Analysis of the presently obtained ESTs revealed 2 TIC components, Tic20 and Tic22, and no TOC components, consistent with a previous report [34]. Because the apicoplast is non-photosynthetic, sour-ces of energy and carbon for such anabolic synthesis should be required. As an important cytosolic source of carbon, dihydro-xyacetone phosphate (DHAP) is imported and converted to glycerol-3-phosphate (G3P), which is a precursor for phos-pholipids synthesis. G3P is sequentially acylated by glycerol-3-phosphate acyltransferase (ACT1) and 1-acyl-glycerol-3-phos-phate acyltransferase (ACT2) to produce phosphatidic acid [35]. One enzyme in this pathway, ACT2, was found to be encoded for by the *P. berghei* ESTs.

In the present study, 10,040 ESTs from *P. berghei* cDNA li-brary were generated, increasing the number of *P. berghei* se-quence in public database. Moreover, the present screening method, which used ESTs enriched in genes with intact 5′ ends, provided 244 genes with predicted coding regions fully cov-ered by 254 EST clusters, showing a powerful means for con-firmation of in silico annotation and identification of the tar-get genes. Also, 48 putative new transcripts encoded in *P. ber-ghei* genome that did not match any EST and annotated pro-tein database of *P. berghei* were identified. However, many of the EST assemblies (Fig. 1) together with these putative new transcripts were assigned to the categories that encode hypo-thetical proteins with unknown functions, indicating that fur-ther studies are needed to define their functions in metabolic pathways. In addition, many EST clusters from this study are contained long 5′ and 3′ untranslated regions (UTRs). The in-formation of these regions can be useful for understanding gene regulation of *Plasmodium*. The constructed a specific *P. berghei* EST database (http://parasite.knu.ac.kr) based on our ESTs and genetic resources will be helpful for bioinformatics analysis and identification of interested genes of *Plasmodium*.

The presently-collected ESTs will be a useful resource to vali-date gene predictions, and extend our understanding of the bi-ology of *Plasmodium* spp., and screening for drug target and vaccine candidates.

REFERENCES

1. Janse CJ, Carlton JM, Walliker D, Waters AP. Conserved location of genes on polymorphic chromosomes of four species of malaria parasites. Mol Biochem Parasitol 1994; 68: 285-296.

2. Rich SM, Ayala FJ. Progress in malaria research: The case for phylogenetics. Adv Parasitol 2003; 54: 255-280.

3. Booker ML, Bastos CM, Kramer ML, Barker RH Jr, Skerlj R, Sidhu AB, Deng X, Celatka C, Cortese JF, Guerrero Bravo JE, Crespo Llado KN, Serrano AE, Angulo-Barturen I, Jimenez-Diaz MB, Viera S, Garuti H, Wittlin S, Papastogiannidis P, Lin JW, Janse CJ, Khan SM, Duraisingh M, Coleman B, Goldsmith EJ, Phillips MA, Munoz B, Wirth DF, Klinger JD, Wiegand R, Sybertz E. Novel inhibitors of *Plasmodium falciparum* dihydroorotate dehydrogenase with anti-malarial activity in the mouse model. J Biol Chem 285: 33054-33064.

4. Hall N, Karras M, Raine JD, Carlton JM, Kooij TW, Berriman M, Florens L, Janssen CS, Pain A, Christophides GK, James K, Rutherford K, Harris B, Harris D, Churcher C, Quail MA, Ormond D, Doggett J, Trueman HE, Mendoza J, Bidwell SL, Rajandream MA, Carucci DJ, Yates JR 3rd, Kafatos FC, Janse CJ, Barrell B, Turner CM, Waters AP, Sinden RE. A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. Science 2005; 307: 82-86.

5. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 2002; 419: 498-511.

6. Chakrabarti D, Reddy GR, Dame JB, Almira EC, Laipis PJ, Ferl RJ, Yang TP, Rowe TC, Schuster SM. Analysis of expressed sequence tags from *Plasmodium falciparum*. Mol Biochem Parasitol 1994; 66: 97-104.

7. Carlton JM, Muller R, Yowell CA, Fluegge MR, Sturrock KA, Pritt JR, Vargas-Serrato E, Galinski MR, Barnwell JW, Mulder N, Kanapin A, Cawley SE, Hide WA, Dame JB. Profiling the malaria genome: A gene survey of three species of malaria parasite with comparison to other apicomplexan species. Mol Biochem Parasitol 2001; 118: 201-210.

8. Watanabe J, Sasaki M, Suzuki Y, Sugano S. Analysis of transcriptomes of human malaria parasite *Plasmodium falciparum* using full-length enriched library: Identification of novel genes and diverse transcription start sites of messenger RNAs. Gene 2002; 291: 105-113.

9. Li L, Brunk BP, Kissinger JC, Pape D, Tang K, Cole RH, Martin J, Wylie T, Dante M, Fogarty SJ, Howe DK, Liberator P, Diaz C, Anderson J, White M, Jerome ME, Johnson EA, Radke JA, Stoeckert CJ Jr, Waterston RH, Clifton SW, Roos DS, Sibley LD. Gene discovery in the apicomplexa as revealed by EST sequencing and assembly of a comparative gene database. Genome Res 2003; 13: 443-454.

10. Bowman IB, Grant PT, Kermack WO. The metabolism of *Plasmodium berghei*, the malaria parasite of rodents. I. The preparation of the erythrocytic form of *P. berghei* separated from the host cell. Exp Parasitol 1960; 9: 131-136.

11. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 1998; 8: 175-185.

12. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 1998; 8: 186-194.

13. Pertea G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B, Tsai J, Quackenbush J. TIGR Gene Indices clustering tools (TGICL): A software system for fast clustering of large EST datasets. Bioinformatics 2003; 19: 651-652.

14. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: A better web interface. Nucleic Acids Res 2008; 36: W5-W9.

15. Conesa A, Götz S, Garcia-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 2005; 21: 3674-3676.

16. Mulder N, Apweiler R. InterPro and InterProScan: Tools for protein sequence classification and comparison. Methods Mol Biol 2007; 396: 59-70.

17. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A. The Pfam protein families database. Nucleic Acids Res 2010; 38: D211-222.

18. Janssen CS, Barrett MP, Turner CM, Phillips RS. A large gene family for putative variant antigens shared by human and rodent malaria parasites. Proc Biol Sci 2002; 269: 431-436.

19. Trenholme KR, Brown CL, Skinner-Adams TS, Stack C, Lowther J, To J, Robinson MW, Donnelly SM, Dalton JP, Gardiner DL. Aminopeptidases of malaria parasites: New targets for chemotherapy. Infect Disord Drug Targets 2010; 10: 217-225.

20. Zhang P, Nicholson DE, Bujnicki JM, Su X, Brendle JJ, Ferdig M, Kyle DE, Milhous WK, Chiang PK. Angiogenesis inhibitors specific for methionine aminopeptidase 2 as drugs for malaria and leishmaniasis. J Biomed Sci 2002; 9: 34-40.

21. Chen X, Chong CR, Shi L, Yoshimoto T, Sullivan DJ Jr, Liu JO. Inhibitors of *Plasmodium falciparum* methionine aminopeptidase 1b possess antimalarial activity. Proc Natl Acad Sci U S A 2006; 103: 14548-14553.

22. Kicska GA, Tyler PC, Evans GB, Furneaux RH, Kim K, Schramm VL. Transition state analogue inhibitors of purine nucleoside phosphorylase from *Plasmodium falciparum*. J Biol Chem 2002; 277: 3219-3225.

23. Kicska GA, Tyler PC, Evans GB, Furneaux RH, Schramm VL, Kim K. Purine-less death in *Plasmodium falciparum* induced by immucillin-H, a transition state analogue of purine nucleoside phosphorylase. J Biol Chem 2002; 277: 3226-3231.

24. Downie MJ, Kirk K, Mamoun CB. Purine salvage pathways in the intraerythrocytic malaria parasite *Plasmodium falciparum*. Eukaryot Cell 2008; 7: 1231-1237.

25. Ulrich P, Paul G, Perentes E, Mahl A, Roman D. Validation of immune function testing during a 4-week oral toxicity study with FK506. Toxicol Lett 2004; 149: 123-131.

26. Monaghan P, Bell A. A *Plasmodium falciparum* FK506-binding protein (FKBP) with peptidyl-prolyl cis-trans isomerase and chaperone activities. Mol Biochem Parasitol 2005; 139: 185-195.

27. Bell A, Wernli B, Franklin RM. Roles of peptidyl-prolyl cis-trans isomerase and calcineurin in the mechanisms of antimalarial action of cyclosporin A, FK506, and rapamycin. Biochem Pharmacol 1994; 48: 495-503.

28. Cui L, Fan Q, Hu Y, Karamycheva SA, Quackenbush J, Khuntirat B, Sattabongkot J, Carlton JM. Gene discovery in *Plasmodium vivax* through sequencing of ESTs from mixed blood stages. Mol Biochem Parasitol 2005; 144: 1-9.

29. Pallavi R, Roy N, Nageshan RK, Talukdar P, Pavithra SR, Reddy R, Venketesh S, Kumar R, Gupta AK, Singh RK, Yadav SC, Tatu U. Heat shock protein 90 as a drug target against protozoan infections: biochemical characterization of HSP90 from *Plasmodium falciparum* and *Trypanosoma evansi* and evaluation of its inhibitor as a candidate drug. J Biol Chem 285: 37964-37975.

30. Wilson RJ, Denny PW, Preiser PR, Rangachari K, Roberts K, Roy A, Whyte A, Strath M, Moore DJ, Moore PW, Williamson DH. Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. J Mol Biol 1996; 261: 155-172.

31. McFadden GI. Mergers and acquisitions: Malaria and the great chloroplast heist. Genome Biol 2000; 1: REVIEWS1026.

32. van Dooren GG, Schwartzbach SD, Osafune T, McFadden GI. Translocation of proteins across the multiple membranes of complex plastids. Biochim Biophys Acta 2001; 1541: 34-53.

33. Tonkin CJ, Kalanon M, McFadden GI. Protein targeting to the malaria parasite plastid. Traffic 2008; 9: 166-175.

34. Lim L, McFadden GI. The evolution, metabolism and functions of the apicoplast. Philos Trans R Soc Lond B Biol Sci 2010; 365: 749-763.

35. Ralph SA, van Dooren GG, Waller RF, Crawford MJ, Fraunholz MJ, Foth BJ, Tonkin CJ, Roos DS, McFadden GI. Tropical infectious diseases: Metabolic maps and functions of the *Plasmodium falciparum* apicoplast. Nat Rev Microbiol 2004; 2: 203-216.