EXPERIMENTAL PROCEDURES

Newly generated Plasmodium falciparum datasets

Single cell analysis of the *in vitro* strain of *Plasmodium falciparum*, 3D7 strain (1) was conducted as follows. Parasites were maintained in a complete medium at 1% haematocrit at 37°C in a 5% CO2/3% O2/balanced N2 gas mixture (2). As described, parasites were treated with chloroquine at the final concentration of 10 nM and harvested at the indicated time after the drug treatment. Next, the infected red blood cells were enriched as previously described using a purification column, MACS (Miltenyi Biotec). Enriched infected cells were further subjected to the single cell separation with the Fluidigm C1 platform following library construction by SMARTer Ultra Low RNA Kit for the Fluidigm C1 System (Clontech, Cat#634835), Advantage 2 PCR Kit (Clontech, Cat#639207), C1 Single-Cell AutoPrep Array for mRNA seq (5-10 um) (Fluidigm, Cat#100-5759) and Illumina Nextera XT DNA Sample Prep Kit (Illumina, Cat.FC-131-1096). Obtained RNA Seq libraries were subjected to Illumina HiSeq2500 platform for 101-bp paired-end sequencing. An average of 0.36 million tags were generated for each single cell library and mapped to the reference genome of *Pf* (PlasmoDB Release 6.0; http://plasmodb.org/plasmo/). Mapping was performed using BWA with default parameters, allowing two-base mismatches. Further details on the biological analysis on this dataset will be presented elsewhere.

"Interactive" RNA-Seq analyses for malaria-infected patients were conducted as previously described (3). In short, peripheral blood samples from 116 Indonesian malaria patients infected with *Pf* were collected and subsequently used for RNA extraction and RNA-Seq analysis. The obtained collections of 36-bp long single-end tags were mapped to the reference genomes of human (hg38, UCSC Genome Browser;

http://genome.ucsc.edu) and *Pf* (PlasmoDB Release 11.0; http://plasmodb.org/plasmo/; (4). By this approach, we expect to allow for simultaneous analysis of gene expression in both parasite and its host human cells. For further details, see the reference (3).

New RNA-Seq tags for apicomplexan parasites

The following eight apicomplexan species, *Plasmodium berghei* ANKA (*Pb*), *Theileria equi* USDA (*Te*), *Neospora caninum* Nc-1 (*Nc*), *Eimeria maxima* NIAH (*Em*), *Babesia caballi* USDA (*Bc*)*, Babesia divergens* (*Bd*)*, Babesia gibsoni* Oita (*Bg*) and *Babesia bigemina* Argentina (*Bb*), were subjected to the conventional RNA-Seq analysis, following the manufacturers' instructions using the TruSeq RNA-Seq kit (Illumina). The resulting products were loaded onto the Illumina HiSeq 2500 platform and subjected to 101-bp paired-end sequencing. The *Pb* samples, strain ANKA, were cultured in mice RBCs. For *Nc*, the Nc-1 strain (5) was cultured with human foreskin fibroblast (HFF) cells. The *Bg* Oita strain (6) was cultured with canine RBCs. *Bc* USDA strain, *Bd, Bb*, and *Te* USDA strain were cultured with bovine RBCs. For *Em*, the NIAH strain samples were collected from feces of infected chickens. Here we examined two developmental stages, namely oocyst and sporozoite. The obtained reads were mapped to the available respective genome sequences, which were obtained mainly from EupathDB and its integrated databases (version 21; http://eupathdb.org/eupathdb/; (7)) using Tophat2 and Bowtie2, v2.0.10 and v2.1.1, respectively (8, 9), allowing two-base mismatches. The generated read mappings were subsequently used for transcriptome reconstruction with Cufflinks v2.1.1 (10, 11). Values for the fragment length mean and standard deviation were assessed individually for each species sample (see the statistics page of the database; http://fullmal.hgc.jp/docs/statistics_2015.html). In cases when the reference

genomes were not available (*Bc*, *Bg*, *Bd, Bb*), a *de novo* transcriptome assembly was conducted using Trinity (12). To facilitate representation of the generated assemblies, the sequences were associated with transcripts of a closely related species, namely *Babesia bovis* (PiroplasmaDB Release 5.0; http://piroplasmadb.org/plasmo/). Mapping was performed with tBLASTx, with E-value<0.001. The candidate coding regions (open reading frames) for all assembled transcripts were identified using TransDecoder (http://transdecoder.sourceforge.net/, release r20131110). The longest and most upstream ORFs were selected preferably.

REFERENCES

1. Walliker, D., Quakyi, I.A., Wellems, T.E., McCutchan, T.F., Szarfman, A., London, W.T., Corcoran, L.M., Burkot, T.R. and Carter, R. (1987) Genetic analysis of the human malaria parasite Plasmodium falciparum. *Science*, **236**, 1661–6.

2. Moll, K., Ljungström, I., Perlmann, H., Scherf, A. and Wahlgren, M. (2013) Methods in Malaria Research, Sixth Edition. Malaria Research and Reference Reagent Resource Center (MR4), Manassas, Virginia, USA.

3. Yamagishi, J., Natori, A., Tolba, M.E.M., Mongan, A.E., Sugimoto, C., Katayama, T., Kawashima, S., Makalowski, W., Maeda, R., Eshita, Y., et al. (2014) Interactive transcriptome analysis of malaria patients and infecting Plasmodium falciparum. *Genome Res.*, 10.1101/gr.158980.113.

4. Aurrecoechea, C., Brestelli, J., Brunk, B.P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O.S., et al. (2009) PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Res.*, **37**, D539–43.

5. Dubey, J.P., Carpenter, J.L., Speer, C.A., Topper, M.J. and Uggla, A. (1988) Newly recognized fatal protozoan disease of dogs. *J. Am. Vet. Med. Assoc.*, **192**, 1269–85.

6. Sunaga, F., Namikawa, K. and Kanno, Y. (2002) Continuous in vitro culture of erythrocytic stages of Babesia gibsoni and virulence of the cultivated parasite. *J. Vet. Med. Sci.*, **64**, 571–5.

7. Aurrecoechea, C., Barreto, A., Brestelli, J., Brunk, B.P., Cade, S., Doherty, R., Fischer, S., Gajria, B., Gao, X., Gingle, A., et al. (2013) EuPathDB: the eukaryotic pathogen database. *Nucleic Acids Res.*, **41**, D684–91.

8. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, R36.

9. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–9.

10. Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J. and Pachter, L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.*, **28**, 511–5.

11. Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L. and Pachter, L. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.*, **7**, 562–78.

12. Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.*, **8**, 1494–512.