Table 1. The ratio of merozoites released from HepG2 cells co-infected with either WTGreen and WTRed or Lisp1ΔGreen and WTRed sporozoites.

Experiment No.	WTGreen : WTRed	Lisp1ΔGreen : WTRed
1	1.1:1	0.20:1
2	1.0:1	0.10:1
3	1.1:1	0.24:1
4		0.12:1
5		0.059:1
6	0.8:1	0.12:1
7		0.11:1
Ave	1.0:1	0.11:1

The average (Ave) ratio from seven independent experiments is shown.

Taken together, these data indicate that in the absence of LISP1 infectious merozoites are formed normally but are not released efficiently from the LS.

Mutant LS do not rupture the PVM

We next examined the morphology of late stages of LS development by transmission electron microscopy (TEM). HepG2 cells were infected with WTGreen or Lisp1ΔGreen sporozoites, and 58 or 61 h later the infected cells were isolated by fluorescence-activated cell sorting and processed for TEM. As shown in Fig. 5A, there was no difference between the development of WT and *Lisp1* mutant parasites at 58 h. Indeed, the proportion of LS containing morphologically mature merozoites was very similar (20–30%). However, about 50% of the WT parasites were no longer surrounded by a visible PVM (Fig. 5B and C) whereas the majority of *Lisp1* mutant parasites were packed within a membrane (Fig. 5D). At 61 h, we were unable to observe WT LS containing merozoites presumably because the breakdown of the PVM rendered them too fragile to survive the sorting procedure. On the contrary, mutant-infected cells were found and the majority of these contained numerous merozoites inside a visible PVM (Fig. 5E and F). The ultrastructure of these merozoites appeared normal as assessed by the presence of a nucleus, rhoptries and a membrane coat (see Fig. S6). These observations strongly support the role of LISP1 in the PVM breakdown and merozoite egress from the host cell.

SERA1 localization is modified in LISP1-defective LS

Several lines of evidence strongly suggest that SERA cysteine proteases play an important role in the rupture of the PVM (Sturm *et al.*, 2009) so we assessed the distribution of SERA-1 protein in LISP1-defective LS. HepG2 cells were infected with WTGreen or Lisp1ΔGreen sporozoites and 48 h later the LS were stained with anti-SERA1 antibody. In WT LS the distribution of SERA gave a punctuated signal located at the PVM whereas in the mutant LS, the signal was more diffuse and possibly extended beyond the PVM (Fig. 6 and Fig. S7). These data show that in the absence of LISP1 the distribution of SERA1 is modified.

Discussion

Egress from the host cell is a fundamental and recurrent step of the *Plasmodium* life cycle. Evidence has now accumulated indicating that the process occurs in two distinct steps, with the rupture of the PVM preceding the rupture of the host cell membrane. The molecular bases of these processes, however, remain unknown.

Earlier work performed with erythrocytic forms of the parasite (reviewed in Blackman, 2008) has suggested that the breakdown of the PV and erythrocyte membranes both depend on proteases but have distinct bases, since treatment with E64, a cysteine protease inhibitor, selectively blocks PVM rupture. Several lines of evidence suggest that the cysteine proteases of the serine repeat antigen (SERA) family (Debrabant *et al.*, 1992), which share a central papain-like domain, play a major role in these processes. This gene family contains nine members in *P. falciparum* and five in rodent-infecting species (Arisue *et al.*, 2007). In *P. falciparum*, SERA proteins are expressed late during the schizogonic cycle (Hodder *et al.*, 2003; Le Roch *et al.*, 2003; 2004) and accumulate in the PV lumen (Knapp *et al.*, 1989; 1991), and the paralogues encoding SERA5 and SERA6 cannot be inactivated (Miller *et al.*, 2002a,b; McCoubrie *et al.*, 2007). In *P. berghei*, four of the five SERA-encoding genes are

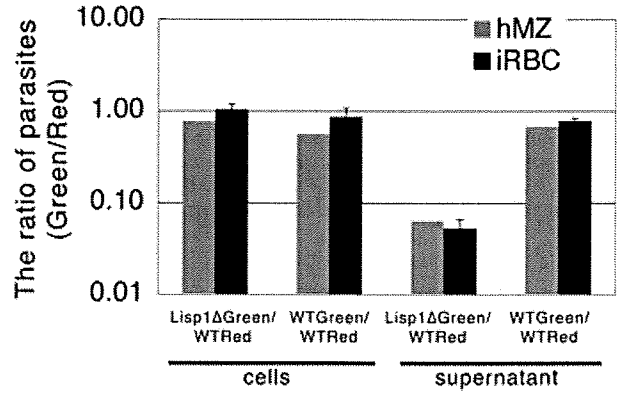

Fig. 4. *Lisp1* mutant LS parasites are defective in the release of merozoites. Graph showing the ratio of Green : Red hepatic merozoites (hMZ) collected at 65 h from adherent HepG2 cells (left) or the supernatant (right) co-infected with Lisp1ΔGreen and WTRed or WTGreen and WTRed sporozoites and the ratios found in blood stages (iRBC) after injection of these merozoites into mice. The results obtained from the supernatant and adherent cells are indicated. The average and standard deviation from four mice are shown.

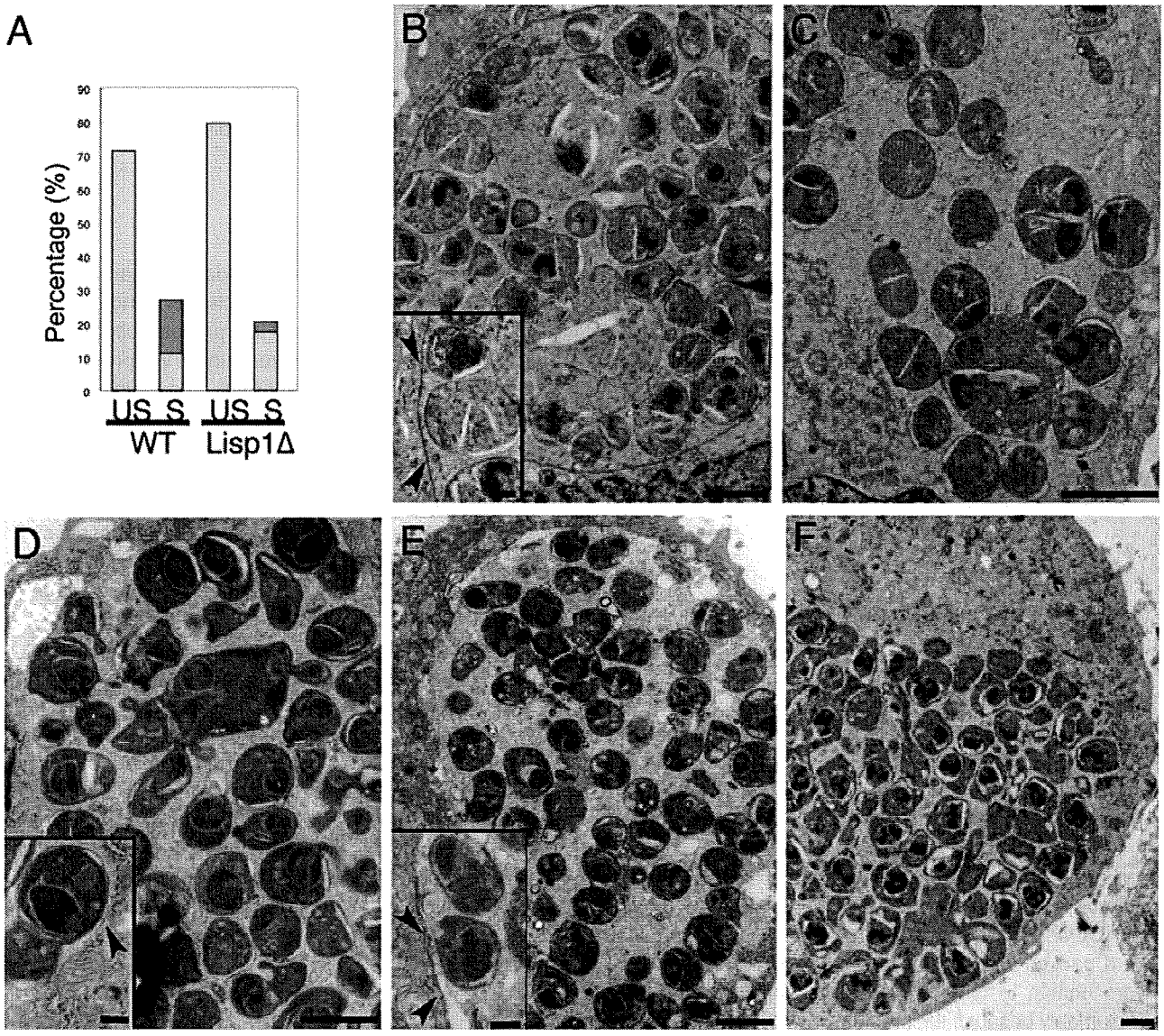

Fig. 5. *Lisp1* mutant merozoites remain within a PVM.
A. Histogram of the percentage of 58 h WTGreen and Lisp1ΔGreen unsegmented LS with no merozoites (US, lilac), and segmented LS containing merozoites (S) with (yellow) or without (orange) a visible PVM.
B–F. Representative examples of hepatic schizonts 58 h (B–D) and 61 h (E and F) post infection. (B) WTGreen with PVM; (C) WTGreen without PVM; (D) Lisp1Δ LS with PVM; (E and F) Lisp1Δ LS at 61 h containing many merozoites within a membrane boundary; scale bars are 2 μm. Insets show mature merozoites; scale bar is 500 nm and arrow heads indicate membrane.

upregulated late during LS maturation, and SERA2 and SERA3 are released in the hepatocyte cytoplasm during merozoite formation (Schmidt-Christensen *et al.*, 2008). Cysteine protease activity is known to be important for the destruction of the LS-enclosing PVM, as well as in the subsequent liberation of merosomes, since the two processes are E64-sensitive (Sturm *et al.*, 2006; Sturm and Heussler, 2007). The available data thus indicate that the SERA cysteine proteases play an important role in the disruption of the PVM, although some additional, non-cysteine protease activity is probably required for disrupting the host cell membrane. Gene targeting in *P. berghei* also showed that SERA8, called egress cysteine protease 1 (ECP1), is important for sporozoite egress from extracellular oocysts in the gut of the mosquito vector, inside which parasites extensively multiply (Aly and Matuschewski, 2005).

We have characterized here a novel protein, LISP1, which is produced by late LS and associates with their enclosing PVM. Inactivation of *Lisp1* in *P. berghei* decreased 15-fold the LS capacity to generate a blood infection in mice. *In vitro*, LISP1-deficient hepatic mero-

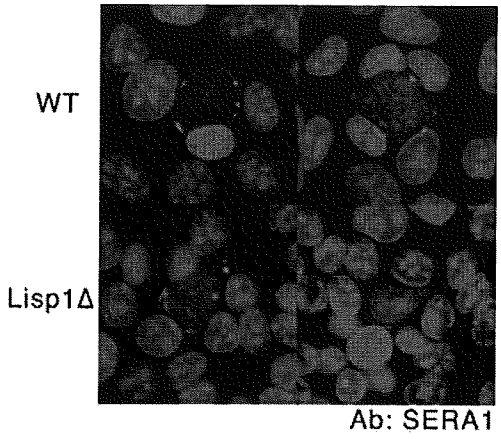

Fig. 6. Localization of SERA1 is modified in *Lisp1* mutant LS parasites. HepG2 cells were infected with wild type (WT, top) or Lisp1ΔGreen sporozoites (Lisp1Δ, bottom) and 48 h later the LS were stained an anti-SERA1 antibody (red) and nuclei stained with DAPI (blue).

zoites were normally formed inside HepG2 cells, and displayed normal infectivity to erythrocytes. This indicates that LISP1 does not play an important role in the function of the PVM, either as a structural protein or as a transporter. However, the PVM of defective LS was not degraded and the merozoites remained trapped inside hepatocytes. The timing of LISP1 production, its localization and the phenotype of the defective mutant thus suggest that LISP1 plays an important role in the rupture of the LS-enclosing PVM. Nevertheless, since LISP1-deficient parasites give rise to a delayed blood infection in mice, release of some merozoites can occur.

The molecular function of LISP1, annotated as a hypothetical protein with a predicted signal peptide but no recognizable functional domain, remains unknown. Being located at the PVM, LISP1 might act as a membrane receptor of the proteases and/or might be involved in their activation/processing at the membrane. In erythrocytic merozoites, processing of the SERA proteases depends on the subtilisin-like serine protease SUB1, which is abundant in schizont stages and is discharged from organelles named exonemes into the PV just prior to egress (Yeoh *et al.*, 2007). LISP1 might thus participate in the further processing of SERA proteases at the PVM, activating their membrane-destabilizing capacity. In the absence of LISP1, the SERA proteins might still be secreted through the PVM without being activated, hence the observed phenotype. Alternatively, LISP1 might be a target of the proteases, mediating membrane dislocation itself upon hydrolysis or processing. Finally, since LISP1 was not detected at the hepatocyte membrane, it seems unlikely that it is also directly involved in merosome formation and hepatocyte/merosome membrane disruption.

Most members of the Apicomplexa phylum, to which *Plasmodium* belongs, invade host cells and multiply within the confines of a PVM that they must eventually breach (Blackman, 2008). In spite of this conserved behaviour, it is striking that all *Plasmodium* products identified so far as being involved in the function/structure of the PVM, such as the ETRAMPs, or in its disruption, such as the SERA proteases, are specific to the genus. Moreover, these products come in stage-specific paralogues adapted to the particular host cell harbouring the PV. LISP1, which appears to be specific to the *Plasmodium* genus, is important for parasite egress from hepatocytes but dispensable for egress from erythrocytes. LISP1 thus adds to the notion of exquisitely specific mechanisms involved in the lysis of the PVM by Apicomplexa parasites.

Experimental procedures

P. berghei *strains and the infection of mice, rats and mosquitoes*

ANKAGreen is ANKA 507m6cL1 which expresses GFP from the *eef1α* promoter (Janse *et al.*, 2006a) and was obtained from MR4 (MRA-867), deposited by C.J. Janse and A.P. Waters; ANKARed (ANKA L733), expressing RFP from *eef1α* promoter, is described in Sturm *et al.* (2009). Infection of mice, *Anopheles stephensi* mosquitoes and isolation of salivary gland sporozoites were performed as previously described (Thiberge *et al.*, 2007). All studies on animals followed the guidelines on the ethical use of animals from the European Communities Council Directive of 24 November 1986 (86/609/EEC).

Real-time reverse transcription polymerase chain reaction (RT-PCR)

Two independent total RNA preparations were made from mixed infected blood stages, salivary gland sporozoites at day 21 post infection and HepG2 cells 5, 17, 40 and 50 h post infection. PCR conditions were one cycle at 95°C for 10 min, 40 cycles at 95°C for 15 s, 55°C for 15 s and 60°C for 45 s. *Lisp1* gene expression was normalized by the *hsp70* (PB001074.01.0) gene expression in each sample. Analysis was performed using the ΔCt method (User Bulletin 2, Applied Biosystems) using the blood stages as reference.

Primers used: *lisp1*F 5′-GCCAAATGCTAAACCTAATG-3′; *lisp1*R 5′-TGGGTTTGTATTGTATGCAC-3′; *hsp70*F 5′-TGCAGC AGATAATCAAACTC-3′; *hsp70*R: 5′-ACTTCAATTTGTGGAACA CC-3′.

Rabbit anti-LISP1 antibody preparation

LISP1 amino acids 1392–1437 were fused to glutathione S-transferase (GST) using the pGEX 6P-1 system (Amersham Bioscience, Uppsala, Sweden). The protein was purified on a GST column and used to immunize rabbits. Specific antibodies were affinity purified on an *N*-hydroxysuccinimide-activated column (Amersham Bioscience) coupled with recombinant protein. The antibody titre was 1.07 mg ml^{-1}.

Immunofluorescence analysis

Purified parasites were fixed in acetone for 2 min and the samples incubated with rabbit anti-LISP1 antibodies then FITC-conjugated secondary antibody (Zymed, South San Francisco, CA, USA). For nuclear staining, 0.02 μg ml^{-1} 4′, 6-diamidino-2-phenylindole (DAPI) was added to the secondary antibody solution.

For IFA of LS in infected HepG2 cells, fixation was either 4% paraformaldehyde (PFA) + 0.1% triton or acetone, and incubation was with anti-LISP1 antibodies and Alexa 647-conjugated secondary antibody. PFA (4%) fixation and 0.1% triton permeabilization were used with the anti-UIS4 antibody. Methanol fixation was used with anti-EXP1 and anti-SERA1 antibodies.

Targeted disruption of lisp1 *and genotypic analysis*

Construction of the lisp1l *targeting vector.* Two fragments of 853 and 801 bp were amplified from genomic DNA with primers: 5′-GGGGAGCTCGTCTATTTTTGATACGATATGTGCACATGC-3′ and 5′-GGGGGATCCCTTGAAGGCGATTAACTATATTGTCGC-3′; 5′-GGGCTCGAGCGAATCAGTGTCGCTTGATATTTTG-3′ and 5′-GGGGGTACCCTGGGTTTGTATTGTATGCACCTAAGG-3′ and cloned either side of the selectable marker gene in pBluescript (Stratagene, La Jolla, CA, USA). For Southern analysis ANKA genomic DNA was digested with SphI and hybridized with a probe made from amplification with primers: 5′-GGG GAGCTCGTCTATTTTTGATACGATATGTGCACATGC-3′ and 5′-GGGGGATCCCTTGAAGGCGATTAACTATATTGTCGC-3′.

Construction of the Lisp1Δ *targeting vector.* Two fragments of 630 and 600 bp were amplified from genomic DNA with primers 5′-CGATGCGGGCCCGAGAATACAACTTACTAGGAAATGCAC-3′ and 5′-AGCTGGCGCGCCCTTGCATCTTCAGACATGTTA TTTTCGA-3′; 5′-CCAGTGAGTGCGGCCGCGGCTAAAGCTCG ATGTCGTATTCAAGAA-3′ and 5′-AGCTGGCGCGCCCTGGG AATTTTTCAATTTCTTCCATTTG-3′. These primers were tailed with restriction sites for ApaI, SmaI, NotI and AscI respectively. The fragments were cloned in the vector pBC-hDHFR vector (B. Boisson, unpublished). This vector is pBC SK (+) (Stratagene), modified to introduce an AscI restriction site, and contains a human dihydrofolate reductase (hDHFR) selection cassette with the 5′ untranslated region of the *eef1α* gene and the 3′ untranslated region of the DHFR-TS gene cloned between SmaI and NotI of pBC SK (+). Transfection and selection of recombinant parasites was performed as previously described (Janse *et al.*, 2006b). Southern analysis was performed on SphI-restricted genomic DNA hybridized with the probe indicated on Fig. S2.

Evaluation of sporozoite infectivity in vivo

Sporozoites collected from mosquito salivary glands were injected intravenously into 3-week-old female Wistar rats (Japan SLC, Hamamatsu, Japan) ($n = 5$) or 4-week-old C57BL/6 mice (Janvier, France). Parasitaemias were checked at indicated time points by Geimsa-stained blood smear.

LS development assay in vivo

A total of 300 000 Lisp1l or WT sporozoites were injected into 3-week-old Wistar rats and 48 h later, the livers were perfused with phosphate-buffered saline (PBS) then 4% PFA. After inclusion in 20% sucrose, 20 μm frozen sections were prepared on a Leica cryostat. The sections were stained with anti-CS antiserum and the number of LS in six sections from each rat was counted. A total of 1619 Lisp1l, 1877 WT LS were counted.

LS development in HepG2 cells

HepG2 cells were grown in DMEM + Glutamax-1 media (Gibco) supplemented with 10% FCS (PAA laboratories GmbH) at 37°C in the presence of 5% CO_2. The day before infection $4–5.0 \times 10^4$ were plated per well in eight-well chamber slides (Nalge Nunc International, Rochester, NY, USA) and 24 h later WT and/or mutant sporozoites (5.0×10^4) were added. The medium was changed daily. The LS were detected by immunofluorescence staining as described above.

Real-time imaging of LS development in HepG2 cells

HepG2 cells were co-infected with WTRed and Lisp1ΔGreen sporozoites. The development of 27 WT and 28 mutant LS was followed from 24 to 48 h post infection by acquisition of images at 60 min intervals with a PerkinElmer spinning disc confocal microscope. The area was calculated using ImageJ software.

Detection of LISP1 by immunoblot

HepG2 cells were infected with WTGreen or Lisp1ΔGreen sporozoites as described above. After 48 h the cells from each well were lysed in Novex® Tris-Glycine sodium dodecyl sulfate (SDS) sample buffer (Invitrogen) and immediately stored at −80°C. Prior to loading on a 4–15% gradient acrylamide gel (Bio-Rad), β mercaptoethanol was added (0.7 M final) and the samples were heated for 5 min at 95°C. Molecular weight markers were Precision Plus Protein (Bio-Rad). After electrophoresis the proteins were transferred to a Hybond-ECL nitrocellulose membrane (Amersham Biosciences). The membrane was first incubated with a 1:2000 dilution of the anti-LISP1 antibody then with a 1:10 000 dilution of a chicken anti-rabbit IgG horseradish peroxidase (HRP)-conjugated antibody (Santa Cruz). The presence of antibodies was revealed with the SuperSignal® West Femto kit (Thermo Scientific). The membrane was then incubated with a 1:10 000 dilution of HRP-conjugated rabbit polyclonal anti-GAPDH (glyceraldehyde-3-phosphate dehydrogenase) (Santa Cruz) and the same kit used to reveal the presence of antibodies. The signals were scanned using Quantity One software (Bio-Rad) and quantified with NanoDrop ND-1000 software (Thermo Fischer Scientifique).

Estimation of number of hepatic merozoites

HepG2 cells were co-infected with equal numbers of either WTRed and WTGreen or WTRed and Lisp1ΔGreen sporozoites. Sixty-five hours later the supernatants were taken, centrifuged for 3 min at 12 000 r.p.m. and the pellet re-suspended in PBS. Cells were scraped from the wells and mechanically disrupted by passing through a Omnican® 30G insulin needle. The samples were placed on Ibidi slides and observed with a Zeiss Axio observer Z1 microscope.

TEM analysis of infected HepG2 cells

WTGreen or Lisp1ΔGreen sporozoites were added to HepG2 cells and 58 or 61 h later cells were collected, fixed with 0.4% PFA and GFP-positive cells enriched by sorting on a FACSAria (BD Biosciences). Sorted cells were fixed with 2.5% glutaraldehyde in 0.1 M cacodylate buffer at 4°C for 24 h. Then, cell pellets were embedded in agarose type IV (Sigma, Chemical, Saint Louis, USA). After several washes in 0.1 M cacodylate buffer, samples were post-fixed for 1 h with 1% osmium tetroxide (Merck, Darmstadt, Germany) in the same buffer. After dehydration in a graded ethanol series, the samples were embedded in Epon resin and polymerized. Contrasted ultra-thin sections (60 nm) were observed in a JEM 1010 Transmission Elecron Microscope (Jeol, Tokyo, Japan). A total of 103 Lisp1ΔGreen and 62 WT LS were observed at 58 h.

Statistical analysis

The data presented in Figs 1A, 2E and 3A were analysed by a Kruskal–Wallis rank sum test. The data presented in Fig. 4 were analysed by a Mann–Whitney test.

Data deposition

Sequences from the EST libraries were deposited in the DDBJ website 2007. The LISP1 sequence reported in this article has been deposited in the GenBank database (Accession No. AB231328).

Acknowledgements

T.I. was supported by the Japan Society for the Promotion of Science (JSPS) as JSPS Postdoctoral Fellowships for Research Abroad. B.B. was supported by the Fonds dédiés 'Combattre les maladies Parasitaires' financed by the Ministère de la Recherche and Sanofi-Aventis. We are extremely grateful to Dr Kappe for providing the anti-UIS4 antibody, Drs Thomas and Barale for the anti-AMA1 antibody, Dr Heussler for anti-EXP1 and anti-SERA1 antibodies. We thank Ms Naoko Kimura for maintenance of mosquitoes and Ms Izumi Kaneko and Ms Tomomi Kato for technical assistance. C. Bourgouin, I. Thiery and the other members of the 'Centre de Production et d'Infection des Anophèles' (Institut Pasteur) are thanked for mosquito rearing. We thank Dr Marie Nguyen-De Bernon Plate-forme de Cytometrie for cell sorting. We also thank Dr Rogerio Amino for help with real-time imaging and all members of BGP for stimulating discussions and comments. This work was supported by funds from the Institut Pasteur, the Howard Hughes Medical Institute and the European Commission (FP6 BioMalPar Network of Excellence). R.M. is a Howard Hughes Medical Institute International Scholar.

References

Aikawa, M. (1971) Parasitological review. *Plasmodium*: the fine structure of malarial parasites. *Exp Parasitol* **30:** 284–320.

Aly, A.S., and Matuschewski, K. (2005) A malarial cysteine protease is necessary for *Plasmodium* sporozoite egress from oocysts. *J Exp Med* **202:** 225–230.

Arisue, N., Hirai, M., Arai, M., Matsuoka, H., and Horii, T. (2007) Phylogeny and evolution of the SERA multigene family in the genus *Plasmodium*. *J Mol Evol* **65:** 82–91.

Baer, K., Klotz, C., Kappe, S.H., Schnieder, T., and Frevert, U. (2007) Release of hepatic *Plasmodium yoelii* merozoites into the pulmonary microvasculature. *PLoS Pathog* **3:** e171.

Bano, N., Romano, J.D., Jayabalasingham, B., and Coppens, I. (2007) Cellular interactions of *Plasmodium* liver stage with its host mammalian cell. *Int J Parasitol* **37:** 1329–1341.

Blackman, M.J. (2008) Malarial proteases and host cell egress: an 'emerging' cascade. *Cell Microbiol* **10:** 1925–1934.

Charpian, S., and Przyborski, J.M. (2008) Protein transport across the parasitophorous vacuole of *Plasmodium falciparum*: into the great wide open. *Traffic* **9:** 157–165.

Debrabant, A., Maes, P., Delplace, P., Dubremetz, J.F., Tartar, A., and Camus, D. (1992) Intramolecular mapping of *Plasmodium falciparum* P126 proteolytic fragments by N-terminal amino acid sequencing. *Mol Biochem Parasitol* **53:** 89–95.

Doolan, D.L., Hedstrom, R.C., Rogers, W.O., Charoenvit, Y., Rogers, M., de la Vega, P., *et al.* (1996) Identification and characterization of the protective hepatocyte erythrocyte protein 17 kDa gene of *Plasmodium yoelii*, homolog of *Plasmodium falciparum* exported protein 1. *J Biol Chem* **271:** 17861–17868.

Hodder, A.N., Drew, D.R., Epa, V.C., Delorenzi, M., Bourgon, R., Miller, S.K., *et al.* (2003) Enzymic, phylogenetic, and structural characterization of the unusual papain-like protease domain of *Plasmodium falciparum* SERA5. *J Biol Chem* **278:** 48169–48177.

Janse, C.J., Franke-Fayard, B., and Waters, A.P. (2006a) Selection by flow-sorting of genetically transformed, GFP-expressing blood stages of the rodent malaria parasite, *Plasmodium berghei*. *Nat Protoc* **1:** 614–623.

Janse, C.J., Franke-Fayard, B., Mair, G.R., Ramesar, J., Thiel, C., Engelmann, S., *et al.* (2006b) High efficiency transfection of *Plasmodium berghei* facilitates novel selection procedures. *Mol Biochem Parasitol* **145:** 60–70.

Knapp, B., Hundt, E., Nau, U., and Kupper, H.A. (1989) Molecular cloning, genomic structure and localization in a blood stage antigen of *Plasmodium falciparum* characterized by a serine stretch. *Mol Biochem Parasitol* **32:** 73–83.

Knapp, B., Nau, U., Hundt, E., and Kupper, H.A. (1991) A new blood stage antigen of *Plasmodium falciparum* highly homologous to the serine-stretch protein SERP. *Mol Biochem Parasitol* **44:** 1–13.

Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., *et al.* (2003) Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301:** 1503–1508.

Le Roch, K.G., Johnson, J.R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., *et al.* (2004) Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Res* **14:** 2308–2318.

Lingelbach, K., and Joiner, K.A. (1998) The parasitophorous vacuole membrane surrounding *Plasmodium* and *Toxo-*

plasma: an unusual compartment in infected cells. *J Cell Sci* **111:** 1467–1475.
McCoubrie, J.E., Miller, S.K., Sargeant, T., Good, R.T., Hodder, A.N., Speed, T.P., *et al.* (2007) Evidence for a common role for the serine-type *Plasmodium falciparum* serine repeat antigen proteases: implications for vaccine and drug design. *Infect Immun* **75:** 5565–5574.
Matuschewski, K., Ross, J., Brown, S.M., Kaiser, K., Nussenzweig, V., and Kappe, S.H. (2002) Infectivity-associated changes in the transcriptional repertoire of the malaria parasite sporozoite stage. *J Biol Chem* **277:** 41948–41953.
Meis, J.F., and Verhave, J.P. (1988) Exoerythrocytic development of malarial parasites. *Adv Parasitol* **27:** 1–61.
Miller, L.H., Baruch, D.I., Marsh, K., and Doumbo, O.K. (2002a) The pathogenic basis of malaria. *Nature* **415:** 673–679.
Miller, S.K., Good, R.T., Drew, D.R., Delorenzi, M., Sanders, P.R., Hodder, A.N., *et al.* (2002b) A subset of *Plasmodium falciparum* SERA genes are expressed and appear to play an important role in the erythrocytic cycle. *J Biol Chem* **277:** 47524–47532.
Mueller, A.K., Labaied, M., Kappe, S.H., and Matuschewski, K. (2005a) Genetically modified *Plasmodium* parasites as a protective experimental malaria vaccine. *Nature* **433:** 164–167.
Mueller, A.K., Camargo, N., Kaiser, K., Andorfer, C., Frevert, U., Matuschewski, K., *et al.* (2005b) *Plasmodium* liver stage developmental arrest by depletion of a protein at the parasite–host interface. *Proc Natl Acad Sci USA* **102:** 3022–3027.
Schmidt-Christensen, A., Sturm, A., Horstmann, S., and Heussler, V.T. (2008) Expression and processing of *Plasmodium berghei* SERA3 during liver stages. *Cell Microbiol* **10:** 1723–1734.
Simmons, D., Woollett, G., Bergin-Cartwright, M., Kay, D., and Scaife, J. (1987) A malaria protein exported into a new compartment within the host erythrocyte. *EMBO J* **6:** 485–491.
Spielmann, T., Gardiner, D.L., Beck, H.P., Trenholme, K.R., and Kemp, D.J. (2006) Organization of ETRAMPs and EXP-1 at the parasite–host cell interface of malaria parasites. *Mol Microbiol* **59:** 779–794.
Sturm, A., and Heussler, V. (2007) Live and let die: manipulation of host hepatocytes by exoerythrocytic *Plasmodium* parasites. *Med Microbiol Immunol* **196:** 127–133.
Sturm, A., Amino, R., van de Sand, C., Regen, T., Retzlaff, S., Rennenberg, A., *et al.* (2006) Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science* **313:** 1287–1290.
Sturm, A., Graewe, S., Franke-Fayard, B., Retzlaff, S., Bolte, S., Roppenser, B., *et al.* (2009) Alteration of the parasite plasma membrane and the parasitophorous vacuole membrane during exo-erythrocytic development of malaria parasites. *Protist* **160:** 51–63.
Tarun, A.S., Dumpit, R.F., Camargo, N., Labaied, M., Liu, P., Takagi, A., *et al.* (2007) Protracted sterile protection with *Plasmodium yoelii* pre-erythrocytic genetically attenuated parasite malaria vaccines is independent of significant liver-stage persistence and is mediated by CD8+ T cells. *J Infect Dis* **196:** 608–616.
Tarun, A.S., Peng, X., Dumpit, R.F., Ogata, Y., Silva-Rivera, H., Camargo, N., *et al.* (2008) A combined transcriptome and proteome survey of malaria parasite liver stages. *Proc Natl Acad Sci USA* **105:** 305–310.
Thiberge, S., Blazquez, S., Baldacci, P., Renaud, O., Shorte, S., Ménard, R., *et al.* (2007) *In vivo* imaging of malaria parasites in the murine liver. *Nat Protoc* **2:** 1811–1818.
Wickham, M.E., Culvenor, J.G., and Cowman, A.F. (2003) Selective inhibition of a two-step egress of malaria parasites from the host erythrocyte. *J Biol Chem* **278:** 37658–37663.
Yeoh, S., O'Donnell, R.A., Koussis, K., Dluzewski, A.R., Ansell, K.H., Osborne, S.A., *et al.* (2007) Subcellular discharge of a serine protease mediates release of invasive malaria parasites from host erythrocytes. *Cell* **131:** 1072–1083.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Histogram representation of real-time RT-PCR analysis of *lisp1* relative gene expression in rat livers 24, 48 and 60 h post infection. Total liver RNA was isolated 24, 48 and 60 h post infection with RNAgents Total RNA Isolation System (Promega, Madison, WI, USA) and cDNAs made. Lisp1 was amplified with primers 5′-GCCCCCTGGCGATCTTAATTT-3′ and 5′-TAAACG TTTCAGGGGGCGAT-3′. The expression was normalized to rat glyceraldehyde-3-phosphate dehydrogenase, amplified with primers 5′-TGATTCTACCCACGGCAAGTT-3′ and 5′-TGATGGG TTTCCCATTGATGA-3′. Error bars are standard deviation.
Fig. S2. Gene inactivation of *lisp1*.
A. Double-crossing-over strategy used to generate Lisp1I mutant parasites. Genomic DNA of ANKA (WT) and two clones of Lisp1I digested with SphI and hybridized with a 5′ probe indicated by a thick black line.
B. Double-crossing-over strategy used to generate Lisp1Δ clones in NK65 and WTGreen. Genomic DNA was digested with SphI and hybridized with a 5′ probe indicated by a thick black line.
Fig. S3. Infectivity of Lisp1ΔGreen sporozoites compared with WTGreen (wild type). A total of 25 000 salivary gland sporozoites were injected intravenously into C57Bl/6 mice and blood-stage parasitaemias checked daily. Bars indicate standard errors.
Fig. S4. The size of *lisp1*-defective LS is normal.
A. Histogram representation of the size of LS in HepG2 cells co-infected with WTRed (Wild type, white box) and Lisp1ΔGreen (black box) sporozoites. The LS parasites were detected by their fluorescence and imaged by spinning disk confocal microscopy and the relative area calculated with ImageJ software.
B. Sample of sorted HepG2 cells 58 h post infection with WTGreen (top) or Lisp1ΔGreen sporozoites used for TEM.
Fig. S5. Lisp1I merozoites express AMA1. Merozoites collected from the supernatant of HepG2 cells infected with ANKA and Lisp1I sporozoites were incubated with an anti-AMA1 antibody and Alexa 546-conjugated secondary antibody.
Fig. S6. Transmission electron micrographs of WT and mutant LS 59 h post infection.
A. WT LS with numerous merozoites in the host cytoplasm and absence of PVM. The aspect of the nucleus indicates that the cell is dying.
B. Lisp1ΔGreen LS with many merozoites and a visible PVM.

C and D. Detail of Lisp1ΔGreen hepatic merozoites. The merozoites appear normal at the ultrastructural level with the presence of a nucleus, rhoptry and coated membrane.

Fig. S7. Immunofluorescence analysis of LS 48 h post infection of HepG2 cells with WT sporozoites. Samples were incubated with an anti-LISP1 antibody and an anti-SERA1 antibody followed by Alexa-conjugated secondary antibodies. LISP1 labelling in green, left; SERA1 labelling in red, centre; merged image, right.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Brief report

Gene expression profiling of peripheral T-cell lymphoma including γδ T-cell lymphoma

Kana Miyazaki,[1] Motoko Yamaguchi,[1] Hiroshi Imai,[2] Tohru Kobayashi,[1] Satoshi Tamaru,[1] Kazuhiro Nishii,[1] Masao Yuda,[3] Hiroshi Shiku,[4] and Naoyuki Katayama[1]

Departments of [1]Hematology and Oncology, [2]Pathologic Oncology, [3]Medical Zoology, and [4]Cancer Vaccine and Immuno-Gene Therapy, Mie University Graduate School of Medicine, Tsu, Japan

The gene expression profile of peripheral γδ T-cell lymphoma (γδTCL) has not been investigated. Using oligonucleotide microarrays, we analyzed total RNA from 7 patients with γδTCL (4 hepatosplenic, 1 cutaneous, 1 intestinal, and 1 thyroidal) and 27 patients with αβTCL (11 peripheral TCL-unspecified, 15 angioimmunoblastic TCL, and 1 hepatosplenic). Unsupervised microarray analyses classified all hepatosplenic γδTCLs into a single cluster, whereas other γδTCLs were scattered within the αβTCL distribution. We identified a T-cell receptor signature gene set, which accurately classified γδTCL and αβTCL. A classifier based on gene expression under supervised analysis correctly identified γδTCL. One case of hepatosplenic αβTCL was placed in the γδTCL grouping. γδTCL signature genes included genes encoding killer cell immunoglobulin-like receptors and killer cell lectin-like receptors. Our results indicate that hepatosplenic γδTCL is a distinct form of peripheral TCL and suggest that nonhepatosplenic γδTCLs are heterogeneous in gene expression. (Blood. 2009;113:1071-1074)

Introduction

T cells expressing the γδ T-cell receptor (TCR) heterodimer comprise only a small fraction of the peripheral blood T-cell population and differ from those expressing the αβ TCR in terms of development, tissue distribution, and function.[1,2] Mature T-cell lymphomas (TCLs) with the γδ T-cell immunophenotype can be divided into hepatosplenic γδTCL[3] and nonhepatosplenic γδTCL.[4] The third World Health Organization (WHO) classification system describes hepatosplenic γδTCLs and hepatosplenic αβTCLs as a single disease entity (hepatosplenic TCL) as they exhibit nearly identical clinicopathologic and cytogenetic features.[4-6]

In contrast, nonhepatosplenic γδTCL occurs in only a limited number of anatomic sites, including cutaneous, nasopharyngeal, gastrointestinal, pulmonary, and thyroidal regions.[7-10] This disease has also been called mucocutaneous γδTCL because the majority of patients show some involvement of mucocutaneous sites. Among nonhepatosplenic γδTCLs, the cutaneous form is most common and overlaps with subcutaneous panniculitis-like TCL.[11,12] Whereas primary cutaneous γδTCL is categorized as a single disease entity in the new WHO scheme,[13] other nonhepatosplenic γδTCLs remains an enigma.

γδTCLs are rare lymphoid malignancies and are difficult to diagnose, resulting from the lack of available monoclonal antibodies against γδTCR for use with paraffin-embedded tissue. Several studies have elucidated the gene expression profile of peripheral TCLs (PTCLs)[14-17] but did not evaluate γδTCL. In our current study, we performed gene expression profiling in 34 PTCLs, including 7 cases of γδTCL.

Methods

Patients/samples

Our present study assessed 34 cases of PTCL, including 11 PTCL-unspecified with αβ T-cell immunophenotype, 15 angioimmunoblastic TCLs, 1 hepatosplenic αβTCL, 4 hepatosplenic γδTCLs, 1 cutaneous γδTCL, 1 intestinal γδTCL, and 1 thyroidal γδTCL. All specimens were collected between 1987 and 2002 at Mie University Hospital and diagnosed according to the third WHO classification.[6] Tumor cell expression of cell-surface antigens and TCR heterodimer (αβ or γδ) was confirmed by immunohistochemistry using frozen sections as described previously.[9] DNA microarray studies using specimens from patients with hematopoietic malignancies were approved by the Institutional Review Committee of Mie University Graduate School of Medicine. Informed consent was obtained from these patients in accordance with the Declaration of Helsinki. The clinicopathologic features of 6 of 7 cases of γδTCL have been reported previously.[8,9] The single patient we examined with thyroidal γδTCL remains alive with no evidence of disease 13 years after diagnosis. Clinical data for all cases examined are presented in Table S1 (available on the *Blood* website; see the Supplemental Materials link at the top of the online article).

Gene expression profiling and analysis

Gene expression profiles were generated and analyzed as previously reported.[18] We used the Agilent 44K human oligonucleotide microarray (Agilent 4112F; Agilent Technologies, Palo Alto, CA), and raw gene expression data are available at the Gene Expression Omnibus (accession number GSE11946).[19] For gene expression profiling supervised by TCR heterodimer phenotype, we selected genes with an average differential expression level of more than 3.0-fold and used a one-sample *t* test with a

Submitted July 10, 2008; accepted October 11, 2008. Prepublished online as *Blood* First Edition paper, October 27, 2008; DOI 10.1182/blood-2008-07-166363.

The online version of this article contains a data supplement.

Presented in part at the Tenth International Conference on Malignant Lymphoma, Lugano, Switzerland, June 4, 2008.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

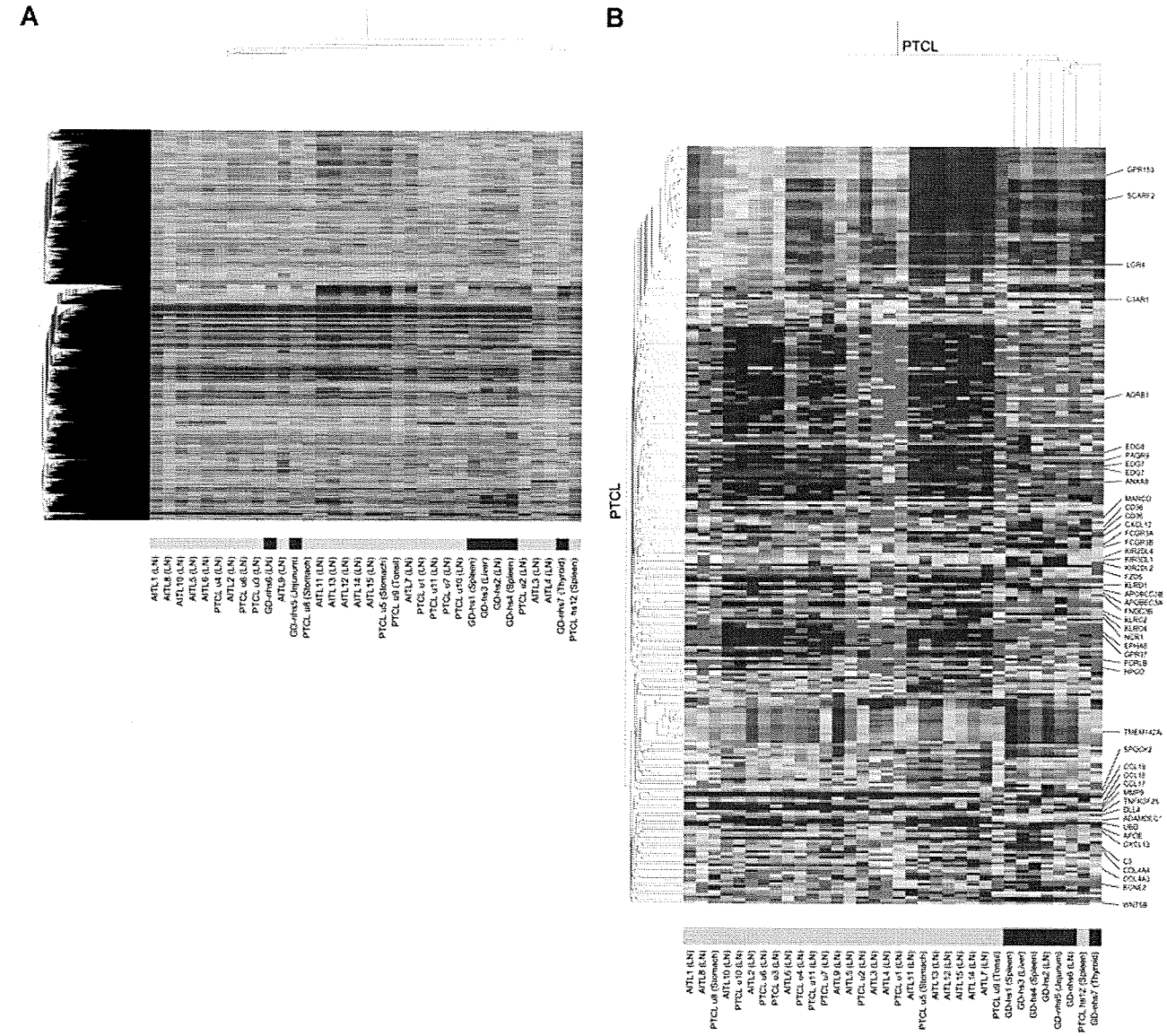

Figure 1. Hierarchical clustering of gene expression data. (A) Hierarchical clustering of 34 cases of PTCL using unsupervised analysis. Color blocks indicate the characteristics of the PTCLs; red blocks, γδTCL cases; amber blocks, αβTCL cases. The anatomic site of each specimen is shown in parentheses. AITL indicates angioimmunoblastic TCL; GD-hs, hepatosplenic γδTCL; GD-nhs, nonhepatosplenic γδTCL; LN, lymph node; PTCL hs, hepatosplenic αβTCL; PTCL u, PTCL-unspecified. Four cases of hepatosplenic γδTCL (GD-hs1-4) were classified into a single cluster. Nonhepatosplenic γδTCLs (GD-nhs5, intestinal; GD-nhs6, cutaneous; GD-nhs7, thyroidal) were scattered within the αβTCL distribution. (B) Hierarchical clustering of 34 cases of PTCL-based TCR signature gene expression. Genes also listed in Table 1 are shown on the right. γδTCL cases were correctly identified by the gene set, although one case of hepatosplenic αβTCL (PTCL hs12) was grouped with the γδTCLs.

P value cutoff of .005. Hierarchical clustering of genes was performed using the Pearson correlation, and hierarchical clustering of cases was obtained using an average linkage algorithm.

We chose WebGestalt using Gene Ontology (GO) hierarchies[20,21] for categorization of functional gene groups and the Kyoto Encyclopedia of Genes and Genomes (KEGG)[22,23] pathway for identification of signaling pathways. In both analyses, we performed a separate hypergeometric test with a *P* value cutoff of .001.

Results and discussion

Gene expression profiling is a powerful tool for establishing a molecular basis for lymphoma subtypes.[24] Unsupervised analysis of our PTCL cases classified hepatosplenic γδTCL as a single cluster, whereas other γδTCLs were scattered within the αβTCL distribution (Figure 1A). Since the 1980s, cumulative clinicopathologic evidence has demonstrated that hepatosplenic γδTCL is a distinct clinicopathologic disease entity. Our gene expression results also confirm that hepatosplenic γδTCL is distinct from other PTCLs. Conversely, nonhepatosplenic γδTCLs appeared to be more heterogeneous. The possibility that tissue differences were responsible for these data was excluded by our observation that hepatosplenic γδTCL was classified as a single cluster, irrespective of the specimen examined (Figure 1). Angioimmunoblastic TCL cases were not classified as a single cluster, consistent with prior gene expression studies.[14-17]

We next compared the gene expression profiles of γδTCL and αβTCL using 291 genes showing a greater than 3.0-fold average expression difference, which we designated the TCR signature gene set (Table S2). Of note, a classifier from supervised analysis based on gene expression identified γδTCL and hepatosplenic αβTCL (Figure 1B). This finding is supported by the notion that γδ

Table 1. GO category analysis and KEGG pathway analysis in the TCR signature gene set

Analytical tool	Gene no.	Gene	P
GO category			
γδTCL			
Cellular defense response	5	*KIR2DL4, NCR1, C3AR1, KLRC2, KLRC4*	1.17×10^{-4}
Signal transduction activity	29	*KLRD1, ANXA9, FNDC3B, GPR37, KIR2DL4, MARCO, EDG7, NCR1, MS4A5, HPGD, FCGR3A, LGR4, FCRLB, C3AR1, CXCL12, RTN4R, PAQR9, KLRC2, GPR153, FCGR3B, KIR2DL2, EDG8, ADRB1, CD36, FZD5, SCARF2, KIR3DL1, EPHA6, KLRC4*	5.09×10^{-6}
Receptor activity	28	*GPR37, KIR2DL4, ANXA9, KLRD1, FNDC3B, EDG7, HPGD, MARCO, LGR4, MS4A5, FCGR3A, NCR1, C3AR1, FCRLB, KLRC2, PAQR9, RTN4R, GPR153, EDG8, FCGR3B, KIR2DL2, ADRB1, FZD5, CD36, SCARF2, KIR3DL1, EPHA6, KLRC4*	7.10×10^{-6}
Transmembrane receptor activity	16	*ANXA9, GPR37, KIR2DL4, KLRD1, LGR4, HPGD, MARCO, EDG7, C3AR1, KLRC2, EDG8, GPR153, ADRB1, FZD5, EPHA6, KIR3DL1*	4.46×10^{-4}
IgG binding	2	*FCGR3A, FCGR3B*	3.07×10^{-4}
αβTCL			
Organismal physiologic process	14	*UBD, CXCL13, COL4A4, CCL18, KCNE2, CCL17, C3, TMEM142A, DLL4, APOE, COL4A3, TNFRSF25, MMP9, CCL19*	3.24×10^{-4}
Regulation of organismal physiologic process	4	*COL4A4, KCNE2, C3, APOE*	5.87×10^{-4}
Circulation	4	*KCNE2, DLL4, COL4A3, APOE*	3.14×10^{-4}
Regulation of neurophysiologic process	2	*COL4A4, APOE*	7.55×10^{-4}
Regulation of transmission of nerve impulse	2	*COL4A4, APOE*	7.55×10^{-4}
Regulation of synapse structure and function	2	*COL4A4, APOE*	8.69×10^{-4}
Regulation of synaptic transmission	2	*COL4A4, APOE*	7.55×10^{-4}
Inflammatory response	5	*CXCL13, CCL18, CCL17, C3, CCL19*	6.05×10^{-4}
Behavior	5	*CCL18, CXCL13, CCL17, APOE, CCL19*	2.47×10^{-4}
Locomotory behavior	4	*CXCL13, CCL18, CCL17, CCL19*	4.92×10^{-4}
Taxis	4	*CXCL13, CCL18, CCL17, CCL19*	4.22×10^{-4}
Chemotaxis	4	*CXCL13, CCL18, CCL17, CCL19*	4.22×10^{-4}
Receptor binding	9	*ADAMDEC1, CCL18, CXCL13, CCL17, C3, DLL4, APOE, COL4A3, CCL19*	2.76×10^{-5}
G-protein–coupled receptor binding	4	*CCL18, CXCL13, CCL17, CCL19*	1.57×10^{-5}
Chemokine receptor binding	4	*CCL18, CXCL13, CCL17, CCL19*	6.97×10^{-6}
Chemokine activity	4	*CCL18, CXCL13, CCL17, CCL19*	6.37×10^{-6}
Extracellular region	11	*CCL18, COL4A4, CXCL13, CCL17, C3, WNT5B, MMP9, COL4A3, APOE, SPOCK2, CCL19*	7.97×10^{-5}
Extracellular region part	9	*CCL18, COL4A4, CXCL13, CCL17, SPOCK2, APOE, COL4A3, MMP9, CCL19*	7.33×10^{-5}
Sheet-forming collagen	2	*COL4A4, COL4A3*	1.70×10^{-4}
Collagen type IV	2	*COL4A4, COL4A3*	1.21×10^{-4}
KEGG pathway			
γδTCL			
Natural killer cell–mediated cytotoxicity	5	*FCGR3A, KIR3DL1, KLRC2, KLRD1, NCR1*	8.10×10^{-4}
Antigen processing and presentation	4	*KIR3DL1, KLRC2, KLRD1, KLRC4*	7.53×10^{-4}
Atrazine degradation	2	*APOBEC3A, APOBEC3B*	4.42×10^{-4}
αβTCL			
Cytokine–cytokine receptor interaction	5	*CXCL13, CCL17, CCL18, CCL19, TNFRSF25*	2.65×10^{-4}

T cells partially share a cytotoxic immunophenotype with cytotoxic αβ T cells.[1,2] Among 30 patients for whom survival data were available, the prognosis for 8 cases with a γδTCL gene profile (7 γδTCLs and 1 hepatosplenic αβTCL) was not significantly poorer than that of 22 patients with an αβTCL gene profile ($P = .152$; Figure S1). The unusual case of thyroidal γδTCL in the γδTCL gene profile group may affect the result because the P value was .009 when we excluded this patient from the survival analysis (data not shown). Future analyses will probably reveal the relationship between our TCR signature gene set and prognostic indicators.

In γδTCL, genes of natural killer (NK) cell–associated molecules, such as killer cell immunoglobulin (Ig)–like receptor (KIR) genes (*KIR3DL1*, *KIR2DL4*, and *KIR2DL2*), and killer cell lectin-like receptors (*KLRC4*, *KLRD1*, and *KLRC2*) were found to be overexpressed (Figure 1; Table S2). These NK receptors are expressed by a subset of NK cells, γδ T cells, and $CD8^+$ αβ T cells.[25] KIR3DL1 and KIR2DL2 exhibit inhibitory functions, and KIR2DL4 has potentially both inhibitory and activating roles.[25] KIR3DL1, KIR2DL2, and KLRD1 are reported to be expressed in some cases of hepatosplenic γδTCL.[26,27] Although *KLRC4*, a top 10 feature gene that characterizes γδTCL and its protein, NKG2F, is expressed in human NK cells,[28] its expression in normal γδ T cells has not been determined. CD16 is also frequently expressed in cases of hepatosplenic γδTCL,[4,27] and its genes (*FCGR3B* and *FCGR3A*) were among the γδTCL signature genes identified in this study.

To search for functionally important genes overexpressed in γδTCL, we performed GO and pathway analysis using 139 of 204 and 53 of 87 known genes in the γδTCL and αβTCL groups, respectively. By WebGestalt, 5 and 20 GO categories were enriched in γδTCL and αβTCL, respectively (Table 1). The enriched GO categories in γδTCL were cellular defense response, signal transduction activity, receptor activity, transmembrane receptor activity, and IgG binding. Three γδTCL pathways and 1 αβTCL pathway were found to be altered in KEGG-signaling analyses (Table 1). No γδTCL and αβTCL signature genes were

shared in a GO category or KEGG pathway, indicating different functional profiles between γδTCL and αβTCL. Four of the 5 γδTCL-enriched GO categories and 2 of the 3 KEGG-signaling pathways altered in γδTCL contained genes encoding KIRs and killer cell lectin-like receptors, a finding that suggests that the expression of these genes may be important for the differential diagnosis of γδTCL and αβTCL.

In conclusion, our current gene expression data confirm that hepatosplenic γδTCL is a distinct lymphoma entity in PTCLs and reveal differences in the gene expression profiles of αβTCL and γδTCL. Further investigations of our newly identified TCR signature genes are warranted to identify novel therapeutic targets and facilitate the diagnosis of γδTCL.

Acknowledgments

We thank the following institutions for providing patient data: Suzuka Chuo General Hospital, Suzuka Kaisei Hospital, Mie University Hospital, Takeuchi Hospital, Tohyama Hospital, Nagai Hospital, Matsusaka Municipal Hospital, Matsusaka Chuo General Hospital, Matsusaka Saiseikai General Hospital, Yamada Red Cross Hospital, and Ise City General Hospital.

This work was supported in part by the Grants-in-Aid for Cancer Research (19-8, 17S-1, 20S-1) from the Ministry of Health, Labor and Welfare, Japan.

Authorship

Contribution: K.M. and M. Yamaguchi designed and performed the study, collected data and samples, interpreted data, and wrote the paper; H.I. and S.T. performed the study and collected samples; M. Yuda and H.S. contributed analytical tools and supervised the research; K.N. collected samples and wrote the paper; and T.K. and N.K. supervised the research and wrote the paper.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Motoko Yamaguchi, Department of Hematology and Oncology, Mie University Graduate School of Medicine, 2-174 Edobashi, Tsu, Mie 514-8507, Japan; e-mail: waniwani@clin.medic.mie-u.ac.jp.

References

1. Hayday AC. γδ cells: a right time and a right place for a conserved third way of protection. Annu Rev Immunol. 2000;18:975-1026.
2. Carding SR, Egan PJ. γδ T cells: functional plasticity and heterogeneity. Nat Rev Immunol. 2002; 2:336-345.
3. Farcet JP, Gaulard P, Marolleau JP, et al. Hepatosplenic T-cell lymphoma: sinusal/sinusoidal localization of malignant cells expressing the T-cell receptor γδ. Blood. 1990;75:2213-2219.
4. Vega F, Medeiros LJ, Gaulard P. Hepatosplenic and other γδ T-cell lymphomas. Am J Clin Pathol. 2007;127:869-880.
5. Macon WR, Levy NB, Kurtin PJ, et al. Hepatosplenic αβ T-cell lymphomas: a report of 14 cases and comparison with hepatosplenic γδ T-cell lymphomas. Am J Surg Pathol. 2001;25:285-296.
6. Jaffe ES, Harris NL, Stein H, Vardiman JW. World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. Lyon, France: International Agency for Research on Cancer; 2001.
7. Arnulf B, Copie-Bergman C, Delfau-Larue MH, et al. Nonhepatosplenic γδ T-cell lymphoma: a subset of cytotoxic lymphomas with mucosal or skin localization. Blood. 1998;91:1723-1731.
8. Yamaguchi M, Ohno T, Kita K. γ/δ T-cell lymphoma of the thyroid gland. N Engl J Med. 1997; 336:1391-1392.
9. Yamaguchi M, Ohno T, Nakamine H, et al. γδ T-cell lymphoma: a clinicopathologic study of 6 cases including extrahepatosplenic type. Int J Hematol. 1999;69:186-195.
10. Jaffe ES. Pathobiology of peripheral T-cell lymphomas. Hematology. 2006:317-322.
11. Toro JR, Liewehr DJ, Pabby N, et al. Gamma-delta T-cell phenotype is associated with significantly decreased survival in cutaneous T-cell lymphoma. Blood. 2003;101:3407-3412.
12. Willemze R, Jansen PM, Cerroni L, et al. Subcutaneous panniculitis-like T-cell lymphoma: definition, classification, and prognostic factors: an EORTC Cutaneous Lymphoma Group Study of 83 cases. Blood. 2008;111:838-845.
13. Swerdlow SH, Campo E, Harris NL, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Lyon, France: International Agency for Research on Cancer; 2008.
14. Ballester B, Ramuz O, Gisselbrecht C, et al. Gene expression profiling identifies molecular subgroups among nodal peripheral T-cell lymphomas. Oncogene. 2006;25:1560-1570.
15. Cuadros M, Dave SS, Jaffe ES, et al. Identification of a proliferation signature related to survival in nodal peripheral T-cell lymphomas. J Clin Oncol. 2007;25:3321-3329.
16. de Leval L, Rickman DS, Thielen C, et al. The gene expression profile of nodal peripheral T-cell lymphoma demonstrates a molecular link between angioimmunoblastic T-cell lymphoma (AITL) and follicular helper T (T_{FH}) cells. Blood. 2007;109:4952-4963.
17. Piccaluga PP, Agostinelli C, Califano A, et al. Gene expression analysis of peripheral T cell lymphoma, unspecified, reveals distinct profiles and new potential therapeutic targets. J Clin Invest. 2007;117:823-834.
18. Miyazaki K, Yamaguchi M, Suguro M, et al. Gene expression profiling of diffuse large B-cell lymphoma supervised by CD21 expression. Br J Haematol. 2008;142:562-570.
19. National Center for Biotechnology Information. GEO: Gene Expression Omnibus. http://www.ncbi.nlm.nih.gov/geo/. Accessed July 1, 2008.
20. Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res. 2005;33: W741-W748.
21. Vanderbilt University. WebGestalt. http://bioinfo.vanderbilt.edu/webgestalt/. Accessed May 2, 2008.
22. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000; 28:27-30.
23. Kanehisa Laboratories. KEGG: Kyoto Encyclopedia of Genes and Genomes. http://www.genome.ad.jp/kegg/. Accessed May 2, 2008.
24. Staudt LM, Dave S. The biology of human lymphoid malignancies revealed by gene expression profiling. Adv Immunol. 2005;87:163-208.
25. Lanier LL. NK cell recognition. Annu Rev Immunol. 2005;23:225-274.
26. Haedicke W, Ho FC, Chott A, et al. Expression of CD94/NKG2A and killer immunoglobulin-like receptors in NK cells and a subset of extranodal cytotoxic T-cell lymphomas. Blood. 2000;95: 3628-3630.
27. Morice WG, Macon WR, Dogan A, Hanson CA, Kurtin PJ. NK-cell-associated receptor expression in hepatosplenic T-cell lymphoma, insights into pathogenesis. Leukemia. 2006;20:883-886.
28. Kim DK, Kabat J, Borrego F, Sanni TB, You CH, Coligan JE. Human NKG2F is expressed and can associate with DAP12. Mol Immunol. 2004;41: 53-62.

Molecular Microbiology (2010) 75(4), 854–863 ■

doi:10.1111/j.1365-2958.2009.07005.x
First published online 27 January 2010

Transcription factor AP2-Sp and its target genes in malarial sporozoites

Masao Yuda,[1]* Shiroh Iwanaga,[1] Shuji Shigenobu,[2] Tomomi Kato[1] and Izumi Kaneko[1]

[1]*Department of Medical Zoology, Mie University School of Medicine, Mie, Tsu, 514-0001, Japan.*

[2]*Okazaki Institute for Integrative Bioscience, National Institute for Basic Biology, National Institutes of Natural Sciences, Higashiyama, Myodaiji, Okazaki, Japan.*

Summary

The malarial sporozoite is the stage that infects the liver, and genes expressed in this stage are potential targets for vaccine development. Here, we demonstrate that specific gene expression in this stage is regulated by an AP2-related transcription factor, designated AP2-Sp (APETALA2 in sporozoites), that is expressed from the late oocyst to the salivary gland sporozoite. Disruption of the AP2-Sp gene did not affect parasite replication in the erythrocyte but resulted in loss of sporozoite formation. The electrophoretic mobility-shift assay showed that the DNA-binding domain of AP2-Sp recognizes specific eight-base sequences, beginning with TGCATG, which are present in the proximal promoter region of all known sporozoite-specific genes. Promoter assays demonstrated that these sequences act as *cis*-acting elements and are critical for the expression of sporozoite-specific genes with different expression profiles. In transgenic parasites that express endogenous AP2-O (APETALA2 in ookinetes), but whose AP2 domain had been swapped with that of AP2-Sp, several target genes of AP2-Sp were induced in the ookinete stage. These results indicate that AP2-Sp is a major transcription factor that regulates gene expression in the sporozoite stage.

Introduction

During a complex life cycle, *Plasmodium* parasites invade different types of host cell and significantly change gene expression in each stage. Until recently, however, the mechanisms of stage-specific gene regulation remained unclear and transcription factors (TFs) participating in this regulation were unknown.

Recently, Apetala2 (AP2) family genes were identified in *Plasmodium* species and were suggested to encode *Plasmodium* sequence-specific TFs (Balaji *et al.*, 2005). This family is characterized by the AP2 domain, a DNA-binding domain composed of approximately 60 amino acids that was first identified in *Arabidopsis* APETALA2 protein (Jofuku *et al.*, 1994). In human and rodent *Plasmodium* parasites, this family is composed of 26 members, and the amino acid sequences of the AP2 domains are highly conserved among orthologues.

AP2-related genes have been reported to be expressed in asexual intra-erythrocytic stages of *Plasmodium* parasites. In *Plasmodium falciparum* their expression during asexual replication is co-ordinated with cell cycle progression (Balaji *et al.*, 2005; De Silva *et al.*, 2008). Among these AP2-related genes two (PlasmoDB identifier: PF14_0633 and PFF0200c) have been suggested to participate in the induction of a group of genes that are required in a cell cycle-specific manner (De Silva *et al.*, 2008). In a previous paper we reported that an AP2 family protein, designated AP2-O (Apetala2 in ookinetes), regulates stage-specific gene expression in the ookinete, which is a motile stage that invades the mosquito midgut (Yuda *et al.*, 2009). AP2-O has a single AP2 domain and binds to the proximal promoter region of target genes with this domain, thereby inducing several genes involved in mosquito midgut-invasion. These findings indicate that AP2-related TFs are widely used for gene regulation in the *Plasmodium* life cycle.

The *Plasmodium* sporozoite is the stage that infects the mammalian liver. Genes expressed in this stage are potential targets for vaccine development. Identification of TFs and elucidation of the mechanisms of gene regulation in this stage should yield novel means to discover vaccine antigens. Here, we show that an AP2-related protein, designated AP2-Sp (AP2 in sporozoites), plays a central role in gene expression in the sporozoite stage using the rodent malaria parasite *Plasmodium berghei*. We also show that AP2-Sp is not necessary for parasite

Accepted 30 November, 2009. *For correspondence. E-mail m-yuda@doc.medic.mie-u.ac.jp; Tel. (+81) 59 231 5430; Fax (+81) 59 231 5430.

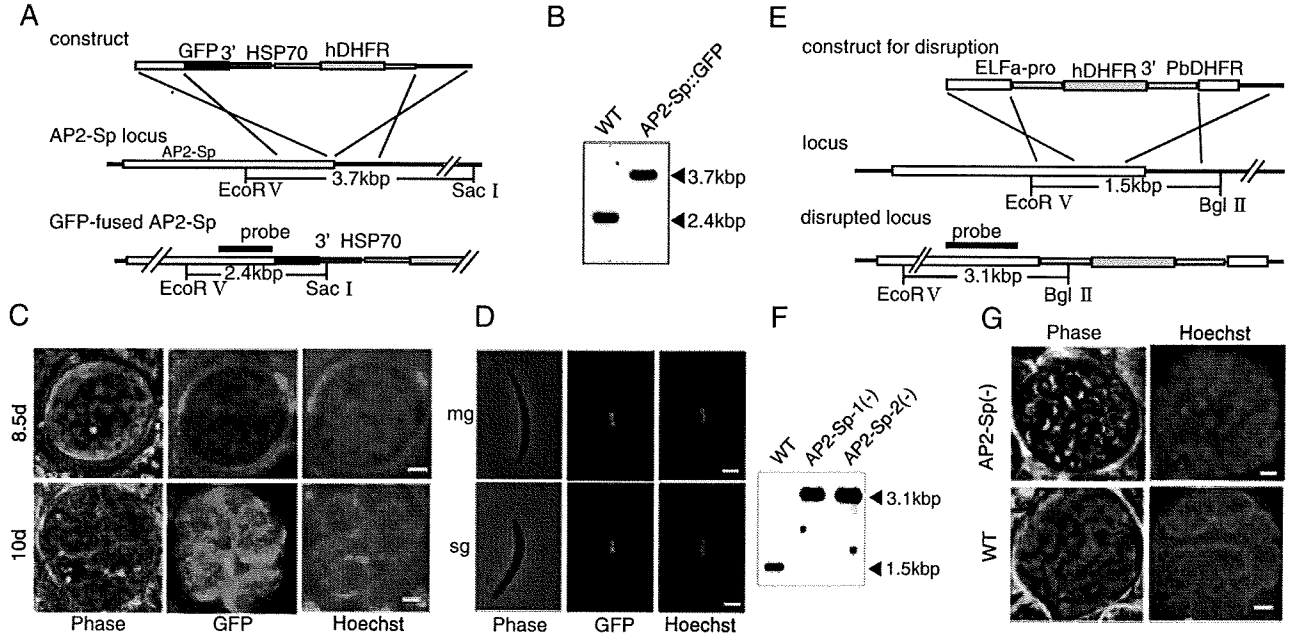

Fig. 1. AP2-Sp is necessary for sporozoite formation.
A. Preparation of parasites expressing AP2-Sp tagged with GFP. The targeting construct (upper) was inserted into the 3′-end of the endogenous AP2-Sp gene locus (middle) by homologous recombination. Transformed parasites (lower) were selected by pyrimethamine-resistance.
B. Southern blot analysis of AP2-Sp::GFP parasites. WT and AP2-Sp::GFP genomic DNA was digested with EcoRV and SacI and hybridized with the probe indicated by a solid bar in (A). WT, original wild-type parasites.
C. Upper panels: Oocyst of AP2-Sp::GFP parasites 8.5 days after an infective blood meal. Fluorescent signals of GFP were colocalized with spotted nuclei stained with Hoechst. Phase, phase-contrast image. Lower panels: Oocyst 10 days after infective blood meal. Scale bars, 5 μm.
D. Oocyst and salivary gland sporozoites of AP2-Sp::GFP parasites. GFP signals are localized in the nuclei of both stages. Scale bars, 2 μm.
E. Targeted disruption of the AP2-Sp gene. Targeting construct (upper) was inserted into the AP2-Sp gene locus (middle) by homologous recombination.
F. Southern blot analysis was carried out on AP2-Sp (–) parasite populations (1 and 2) that were independently prepared. WT and AP2-Sp (–) genomic DNA was digested with EcoRV and BglII and hybridized with the probe indicated by a solid bar in (E).
G. Oocysts of wild-type (WT) and AP2-Sp (–) 1 parasites 10 days after an infective blood meal. Midguts infected with wild-type or knockout parasites were stained with Hoechst. Budding sporozoites were not observed around sporoblasts in AP2-Sp (–) oocysts (upper panels), while numerous sporozoites budding from sporoblasts were stained with Hoechst in WT oocysts (lower panels). Scale bars, 5 μm.

proliferation in the blood stage but is essential for formation of sporozoites.

Results

AP2-Sp is expressed from the late oocyst to the salivary gland sporozoite

To search for AP2-related genes expressed in sporozoites, we screened our *P. berghei* EST (expressed sequence tag) database [available at the PlasmoDB website (http://www.plasmodb.org)] and found that ESTs of *AP2-Sp* (PlasmoDB identifier, PB000752.01.0) are present in oocyst and salivary gland sporozoites (1 and 3 ESTs, respectively). We therefore prepared transgenic parasites that expressed GFP-tagged AP2-Sp in *P. berghei* (AP2-Sp::GFP parasites, Fig. 1A and B) and investigate the expression profile of *AP2-Sp*. In AP2-Sp::GFP parasites, GFP signals were not observed in the blood stages, including gametocytes and ookinetes cultured *in vitro* (Fig. S1). In the mosquito vector, however, fluorescent signals appeared in some oocysts 6 days after an infective blood meal (or fertilization), and the proportion of oocysts expressing AP2-Sp::GFP increased with time (Table S1). In these oocysts signals were weak and localized in Hoechst-stained spotted nuclei of sporoblasts (Fig. 1C, upper panels). Around 10 days after an infective blood meal, new, strong GFP signals appeared in oocysts (Fig. 1C, lower panels). These signals were localized to the nuclei of sporozoites budding from sporoblasts and therefore displayed ring-like appearances around the core of the sporoblasts. Strong GFP signals were further observed in the nucleus of sporozoites released from oocysts (oocyst sporozoites) and in sporozoites harvested from mosquito salivary glands (salivary gland sporozoites) (Fig. 1D). These results suggested that this protein would function as a TF in the sporozoite stage.

Table 1. Formation of oocysts and sporozoites in AP2-Sp (–) parasites.

Parasite populations	Number of oocysts per mosquito (mean ± SE, $n = 20$)	Diameter (μm) of oocysts (mean ± SE, $n = 50$)	Number of oocyst sporozoites per mosquito	Number of salivary gland sporozoites per mosquito
Wild type	112.95 ± 36.17	133.75 ± 3.76	26235 ± 2095[a]	10750 ± 472[a]
AP2-Sp (–) 1	115.05 ± 43.00	130.91 ± 3.28	1.8, 0.9[b]	0, 0[b]
AP2-Sp (–) 2	107.65 ± 30.21	138.06 ± 2.59	3.6, 0.9[b]	0, 0[b]

a. Twenty mosquitoes were dissected, and sporozoites were counted using a haemocytometer. Each value is a mean of three independent experiments ± SE.
b. Fifty mosquitoes were dissected, and sporozoites were counted using a haemocytometer. Two independent experiments were performed in each population.

AP2-Sp is necessary for the formation of sporozoites in oocysts

Next, we performed targeted disruption of *AP2-Sp* in *P. berghei* (Fig. 1E and F). Two knockout populations were obtained from independent transfection experiments. The knockout populations replicated normally in the blood, developed into ookinetes and infected the mosquito midgut; the number and sizes of oocysts formed in the midgut were essentially the same as in the wild type (Table 1). In the oocysts, however, sporozoites were not observed, even 14 days after the infective blood meal and were barely detectable after harvest from the midgut (Table 1). Nuclear staining with Hoechst showed that the nucleus of each oocyst was divided into several small nuclei 10 days after the infective blood meal in knockout as in wild-type parasites (Fig. 1G). Semithin sections stained with Giemsa, however, revealed that the subsequent invagination of the plasma membrane, which is necessary for formation of sporoblasts, did not occur in knockout parasites even 14 days after infective blood meal (Fig. S2). These results suggested that AP2-SP is involved in the expression of genes necessary for sporozoite formation.

AP2-Sp binds to eight bp sequences containing TGCATG

In our previous paper (Yuda *et al.*, 2009), we reported that AP2-O binding sites in a promoter have the following two features; they are located immediately upstream of the transcription start site (TSS), and usually two or more binding sites are present. According to these observations, we explored putative *cis*-acting elements in sporozoite-specific genes. We found that the six-base sequence, TGCATG, and its reverse complimentary sequence, CATGCA, were frequently found in the upstream regions of these sporozoite-specific genes. In particular, a conjugated palindromic sequence, TGCATGCA, was found at high frequencies in them (see also Fig. 3A). The sequence is identical to the sequence that was recently reported to be a common upstream motif of sporozoite genes in *P. falciparum* (Young *et al.*, 2008). The conjugated eight-base sequence is identical to the sequence that has been reported as a binding sequence for the AP2 domain of *P. falciparum* AP2-Sp (PF14_0633) (De Silva *et al.*, 2008).

We next examined whether the AP2 domain of *P. berghei* AP2-Sp binds to the six-base sequence using EMSA (electrophoretic mobility-shift assay). Because the AP2 domain of AP2-Sp has an adjacent AT-hook motif (Balaji *et al.*, 2005) and this region might participate in binding, recombinant AP2 domains with and without the AT-hook region were tested. As a probe, a 192 bp region upstream of the *MAEBL* gene, a sporozoite-specific gene essential for salivary gland infection (Kariu *et al.*, 2002), was used. This region contained a single TGCATGCA site. As shown in Fig. 2A, EMSA generated clearly shifted bands for both recombinant proteins. However, addition of a single mutation to the putative binding sequence (TGCATGCA to TGGATGCA) virtually abolished these bands. This result indicated that the AP2 domain specifically recognizes this sequence.

Next, we designed short (29 bp) double-stranded oligonucleotide probes from the region around the TGCATG site upstream of *MAEBL*. These probes contained various point mutations, to determine the sequence necessary for binding to the AP2 domain (Fig. 2B). Only GST-fused proteins, without the AT-hook region, were used in these assays, because there seemed to be no difference in binding properties between the two recombinant proteins. The results indicated that CATG is a core sequence essential for binding (Fig. 2B, left and centre panels) and that the two bases either side of this sequence also participate in binding (Fig. 2B, left panel), i.e. the whole palindromic sequence of eight bases, TGCATGCA, is involved in the binding of the AP2 domain. The minimum sequence essential for binding the AP2 domain was TGCATG, because, binding was abolished only when both TGCATG and its reverse complementary sequence CATGCA were mutated at the same time (Fig. 2B, centre panel). A mutation to either side of the sequence had no affect on binding (Fig. 2B, right panel).

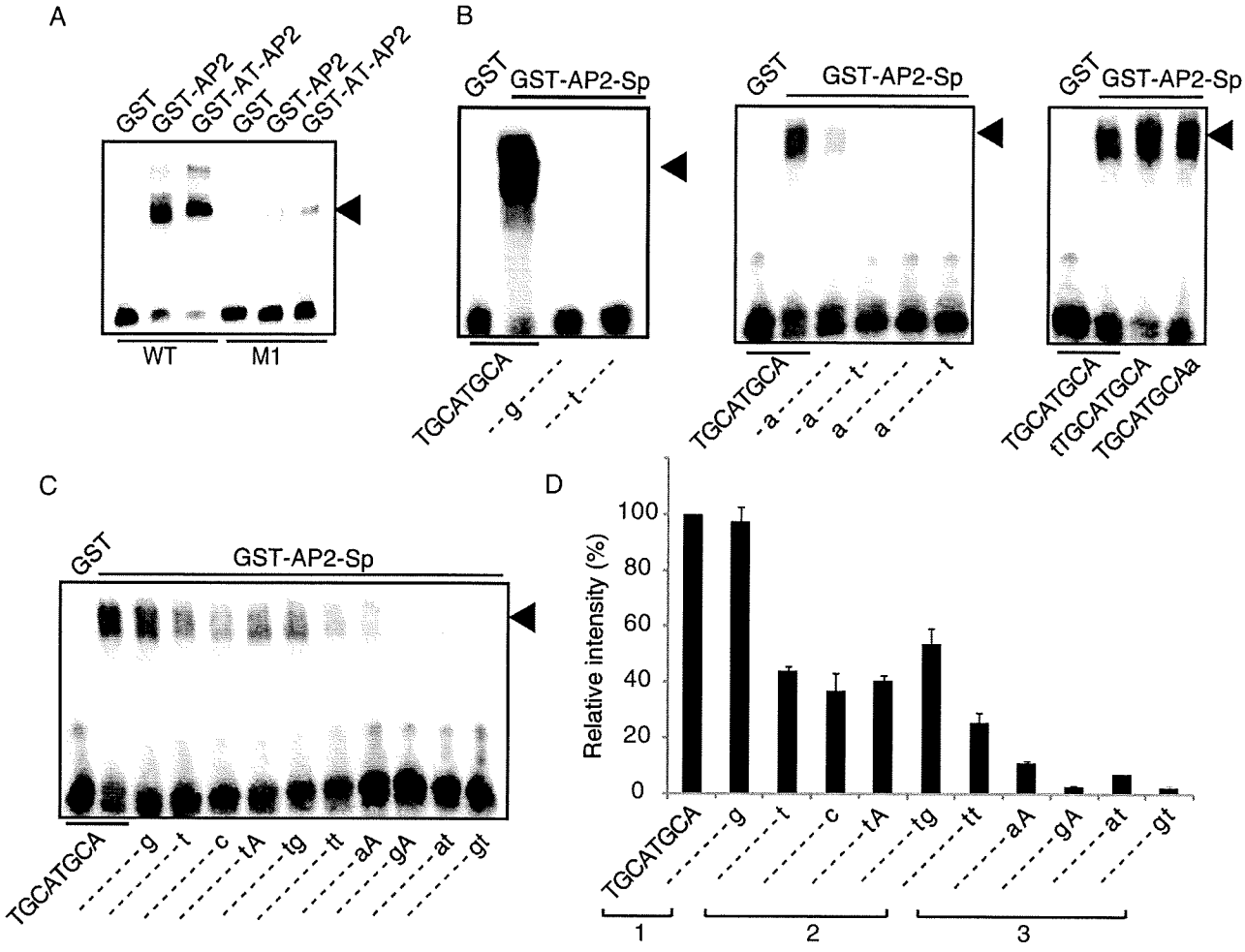

Fig. 2. AP2 domain of AP2-Sp recognizes eight-base sequences beginning with TGCATG.
A. EMSA was performed with the GST-fused AP2 domain of AP2-Sp, with and without the AT-hook region (GST-AT-AP2 and GST-AP2 respectively). The upstream region of *MAEBL* (192 bp), which contains one TGCATGCA sequence, was used as a probe (WT). A point mutation (TGCATGCA to TGGATGCA) in this probe (M1) nearly abolished the shifted band (arrow head). GST was used as a control.
B. EMSA was performed with the GST-fused AP2 domain of AP2-Sp without the AT-hook region. Oligonucleotide probes (29 bp) were designed based on the upstream region of *MAEBL* around the TGCATGCA sequence. Mutations added to this probe are indicated under each lane with small characters. A mutation to CATG at the center of TGCATGCA abolished the shifted band (left panel). A mutation to either TG or CA at the end of TGCATGCA significantly reduced the shifted band, and mutations to both ends together abolished the shifted band (centre panel). Mutation of the bases beside TGCATGCA had no affect on binding (right panel).
C and D. Three independent EMSAs were performed with oligonucleotide probes that are different from the original probe in (B) at either or both of the two bases following TGCATG. One of the electrophoresis images obtained in these assays is shown in (C). Intensities of shifted bands were measured with a densitometer, and the averages of three assays are shown in a bar graph with standard errors in (D). The sequences of oligonucleotide probes were roughly divided into three groups according to the band intensities, as indicated under the graph.

These results suggest that AP2-Sp could bind to several eight-base sequences beginning with TGCATG but that binding properties would vary among them. We therefore compared binding affinities among these sequences using 29 bp short oligonucleotide probes (Fig. 2C and D). The results showed that binding sequences for the AP2 domain could be roughly divided into three groups according to binding properties. The first group is composed of TGCATGCA and TGCATGCG, to which the AP2 domain binds most tightly. The second group contains four sequences, TGCATGCC/T and TGCATGTA/G, which are identical to the sequences of the first group at seven bases and have lower affinities. The third group includes other sequences containing TGCATG that also resulted in a shifted EMSA band, although with lower binding affinities than the former two groups.

Figure 3A illustrates these binding sequences in 1 kb upstream regions (and 1.2 kb regions in some genes) of sporozoite-specific genes and the most upstream ESTs, which indicate the putative TSS of each gene. All of these genes have one of the eight-base sequences in the upstream region (over half of them have the eight-base sequences belonging to the first group), and transcription

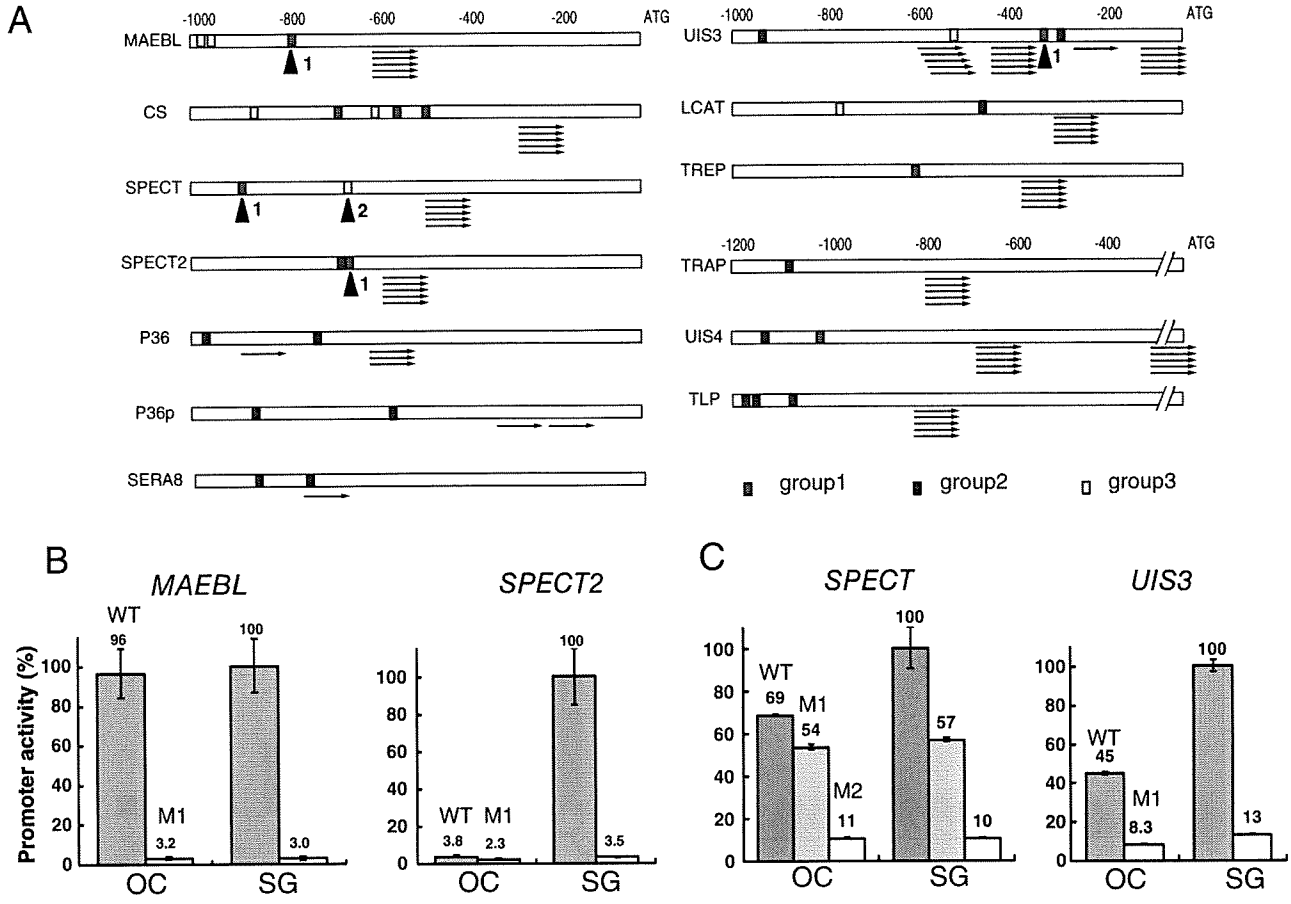

Fig. 3. Promoter activities of sporozoite-specific genes are dependent on AP2-Sp binding to sequences in the upstream region.
A. Binding sequences for AP2-Sp (rectangles) and start sites of most upstream ESTs (arrows), which indicate TSS, in the 1 kb region upstream of sporozoite-specific genes. Binding sequences are coloured according to the groups to which they belong (see also Fig. 2D). In some genes with a long 5′-untranslated region (lower right), 1.2 kb upstream regions are shown. In most genes transcription starts 100–300 bp downstream of a binding site for AP2-Sp, and two or more binding sites are usually within the 1 kb region upstream of the first ATG codon. EST data are available at the PlasmoDB website. Arrowheads indicate binding sequences that were mutated in reporter assays in (B) and (C).
B. Reporter assays of *MAEBL* and *SPECT2* were performed *in vivo* with pCen-GFP. Promoter activities in oocyst (OC) and salivary gland (SG) sporozoites were compared between original (WT) and mutated (M1) promoters. In both genes, a point mutation was added to the binding sequence for AP2-Sp (TGCATGCA to TGGATGCA) in the upstream region.
C. The reporter assays of *SPECT* and *UIS3* were performed using pCen-Luc. A mutation was added to one [numbered 1 in (A), M1] or two binding sites (1 and 2, M2) in *SPECT* and to one binding site (M1) in *UIS3*. Results were expressed as a percentage of the original activity in salivary gland sporozoites.

of these genes starts 100–300 bp downstream of the binding sequences. Computational analyses in 1 kb upstream regions of 13 genes, which have been reported sporozoite-specific, demonstrated that the six-base sequence occurs at high frequencies in these genes (Table S2). These results strongly suggest that binding sequences for AP2-Sp act as common *cis*-acting elements in the upstream regions of sporozoite-specific genes and that they are usually located in the proximal promoter region of these genes. These features are similar to those of AP2-O, which we reported previously (Yuda *et al.*, 2009). We also found that some eight-base sequences beginning with CGCATG can be binding sequences for AP2-Sp (data not shown). However, we omitted these sequences from the following analyses because the frequencies of these GC-rich sequences in the genome are very low (an example that such sequences act as *cis*-acting motifs *in vivo* is shown in the promoter analysis of PB000863.01.0 in Fig. 4B).

TGCATGCA is a cis*-acting element specific for the sporozoite stage*

We next examined whether the eight-base sequences act as *cis*-acting elements in the sporozoite stage using *in vivo* promoter assays (Fig. 3B and C). First, we tested TGCATGCA in the upstream region of *MAEBL* (Fig. 3B). For this assay we used the *P. berghei* centromere plasmid

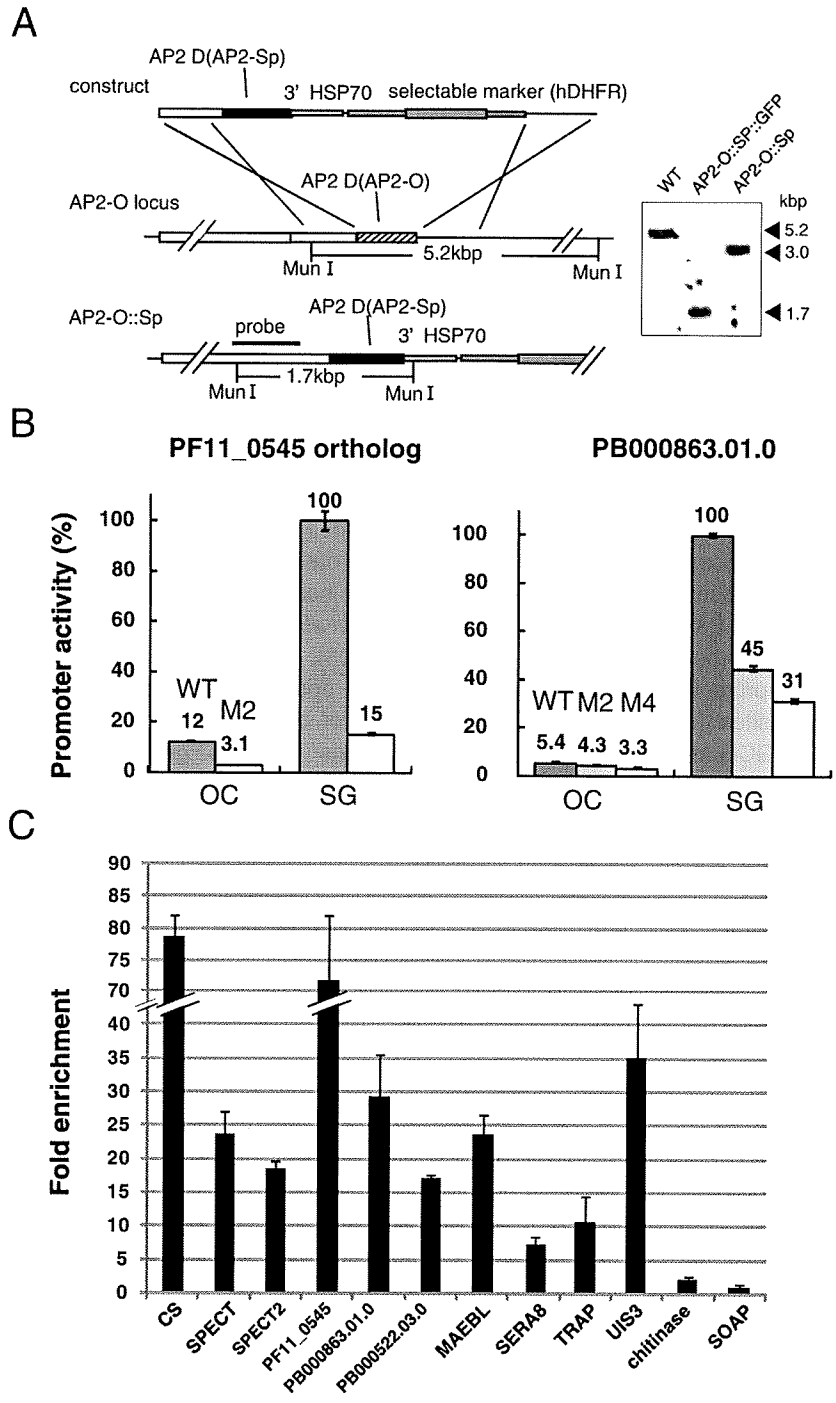

Fig. 4. Induction of AP2-Sp target genes in ookinetes.
A. Preparation of AP2-O::Sp parasites. DNA construct containing the coding region of the AP2 domain of AP2-Sp (upper) was inserted in frame into the 3'-end of the endogenous AP2-O gene (middle) by homologous recombination, resulting in the swap of the region encoding the AP2 domain of AP2-O with the region encoding the AP2 domain of AP2-Sp (lower). Preparation of AP2-O::Sp::GFP parasites was carried out with the same procedure but using a construct containing the coding region of the AP2 domain of AP2-Sp followed by the GFP gene. Results of Southern hybridization with the probe indicated by a solid bar are shown in the right panel.
B. Reporter assays were performed with the artificial chromosome pCen-Luc. Promoter activities in oocyst (OC) and salivary gland (SG) sporozoites were compared between the original upstream region (WT) and regions containing point mutations at binding sites for AP2-Sp. The binding sites mutated were TGCATGCG and TGCATGCT (M2, see also Fig. 3S) in PF11_0545 orthologue, and two TGCATGTGs (1 and 2 in Fig. 3S, M2) and CGCATGTT and CGCATGTG (3 and 4 in Fig. 3S, M4) in PB000863.01.0. Results were expressed as a percentage of the original activity in salivary gland sporozoites.
C. ChIP was performed in AP2-O::Sp::GFP parasites with anti-GFP antibodies. Precipitated DNA fragments were analysed by qPCR using primers for upstream regions of 11 genes including those induced in AP2-O::Sp parasites (*SPECT*, *SPECT2*, PF11_0545 orthologue, PB000522.03.0 and PB000863.01.0), other known sporozoite-specific genes (*TRAP*, *MAEBL*, *UIS3* and *SERA8*) and ookinete-specific genes (*chitinase* and *SOAP*) as controls. Percentage of input of each gene is shown in Table S8.

pCen-GFP, which contains the GFP reporter gene (Fig. S3A) (Yuda *et al.*, 2009). Reporter constructs made in this plasmid are retained through the sporozoite stage by approximately 90% of parasites without drug selection, having been introduced into asexual blood stages by electroporation (S. Iwanaga *et al.*, unpubl. results). In parasites transfected with pCen-GFP containing the *MAEBL* promoter, strong GFP signals were observed in oocyst sporozoites. The signal intensity reached a peak in this stage, which was in accordance with the expression profile of the *MAEBL* gene *in vivo*. Addition of a point mutation to the eight-base sequence (TGCATGCA to TGGATGCA) decreased the signals intensity to the background level, indicating that this sequence is critical for promoter activity. Next, we tested the upstream region of *SPECT2*, which also contains a single TGCATGCA site

(Fig. 3B, right). *SPECT2* is necessary for skin and liver invasion and is predominantly expressed in the salivary gland sporozoite (Ishino *et al.*, 2005; Amino *et al.*, 2008). In parasites transfected with pCen-GFP containing the *SPECT2* promoter, GFP signals were weak in oocyst sporozoites but increased approximately 20 times after salivary gland infection. This steep increase in promoter activity coincides well with the expression profile of *SPECT2*. Addition of a point mutation to the motif decreased GFP signals to background levels in both oocyst sporozoites and salivary gland sporozoites.

The promoter region of sporozoite-specific genes usually has multiple binding sites for AP2-Sp. To examine how these individual sequences contribute to the overall promoter activity, we next performed promoter assay with pCen-Luc in the *SPECT* promoter (Fig. 3C, left) (Ishino *et al.*, 2004) where two putative *cis*-acting elements different from TGCATGCA tested above (TGCATGCG and TGCATGTT) are present (Fig. 3A). In these assays luciferase was used as a reporter (pCen-Luc, Fig. S3B) instead of GFP to assess promoter activity more quantitatively. The activity of the *SPECT* promoter decreased according to the number of binding sites that were mutated. This indicated that both TGCATGCG and TGCATGTT act as *cis*-acting elements and that these elements additively affect promoter activity. The additive effect on promoter activity was similar to that reported for AP2-O binding sequences in the ookinete stage (Yuda *et al.*, 2009).

We also tested the promoter of *UIS3* (upregulated in infective sporozoites gene 3) (Mueller *et al.*, 2005) (Fig. 3C, right). *UIS3* is upregulated in salivary gland sporozoites and is also expressed in the liver stage. Therefore, the expression profile and function is different from the three genes tested above. Addition of a mutation to the single TGCATGCA site significantly decreased promoter activity of the *UIS3* upstream region in both oocyst and salivary gland sporozoites. These results together with those described above demonstrated that the eight-base sequence acts as a critical *cis*-acting element, irrespective of whether the peak activities of the promoter are present in oocyst or salivary gland sporozoites.

Induction of AP2-Sp target genes in the ookinete stage

The results presented above indicated that AP2-Sp is a *trans*-acting factor that induces sporozoite-specific genes. However, it was unknown whether this protein directly binds to the promoter of these genes *in vivo*. ChIP-qPCR (chromatin immunoprecipitation with quantitative PCR) analysis is an appropriate technique to demonstrate this. However, we were unable to obtain sufficient parasites for this analysis. Indeed, we attempted a ChIP assay with sporozoites that were harvested from 300 infected mosquitoes (approximately 6×10^6 sporozoites), but insufficient chromatin DNA for subsequent qPCR analysis was recovered (data not shown).

To conquer this problem, we prepared parasites that expressed endogenous *AP2-O*, but whose AP2 domain had been swapped with that of AP2-Sp (AP2-O::Sp parasites and AP2-O::Sp::GFP parasites, Fig. 4A), and examined (i) whether the chimeric TF might induce target genes of AP2-Sp in the ookinete stage and (ii) whether the TF would bind to the promoter regions of sporozoite-specific genes. To assess the induction, we analysed gene expression by microarray analysis of ookinetes using AP2-O::Sp parasites and selected genes whose expression increased over threefold compared with *AP2-O* (–) parasites. The microarray analysis showed that expression increased in 33 genes (Table S3). They were composed of six sporozoite-specific genes [*SPECT*, *SPECT2*, *CS*, *TLP* (TRAP-like protein), S10 and S23], whose expression in sporozoites has been demonstrated or suggested (Matuschewski *et al.*, 2002; Ishino *et al.*, 2004; 2005; Moreira *et al.*, 2008), and 27 novel genes whose expression has not been reported in the sporozoite stage. Computational analysis of upstream regions of the 30 genes (upstream regions of three genes were not obtained from the *P. berghei* genome database) showed that sequences containing TGCATGCA occurred at significantly high frequencies, indicating that direct binding of the TF to these regions could induce these genes (Table S4). Figure S3 illustrates the binding sites for AP2-Sp in the upstream regions of the 26 genes that are not in Fig. 3A. Of the 30 genes, 19 had at least one binding site belonging to the first group (TGCATGCA, TGCATGCG and their reverse complementary sequences, see also Table S3), and 17 genes had at least one binding site belonging to the second group (TGCATGTG, TGCATGTA, TGCATGCT, TGCATGCC and their reverse complementary sequences). The most upstream ESTs of each gene (Fig. S3) indicated that transcription of these genes starts approximately 100–300 bp downstream of the binding sequences, as was observed for the sporozoite-specific genes in Fig. 3A.

We performed further promoter assays to confirm that these genes really are under the control of AP2-Sp in the sporozoite stage (Fig. 4B). We tested two genes. One of the genes we tested currently has no gene ID in PlasmoDB (in *P. falciparum* PlasmoDB identifier, PF11_0545, has been attributed to the orthologue). The upstream region of this gene had a single binding sequence, TGCATGCG, and ESTs of this gene were mainly present in salivary gland sporozoites (Fig. S4 and Table S3). The promoter activity of this gene increased after salivary gland infection, and addition of a mutation (TGCATGCG to TGGATGCG) significantly decreased the activity in both oocyst and salivary gland sporozoites. Another gene

we tested (PlasmoDB identifier: PB000863.01.0) was a member of the AP2-related gene family and has an orthologue in *P. falciparum* (PlasmoDB identifier: PFD0985w). This gene has four putative *cis*-acting elements in the upstream region, two TGCATGTG and two CGCATGTG (Fig. S4), and ESTs of this gene were present mainly in salivary gland sporozoites (Table S3). Promoter activity of this gene increased steeply after salivary gland infection, and point mutations to the motifs (in two and four motifs) reduced the activity stepwise. These results indicated that expression of these two genes is controlled by AP2-Sp in the sporozoite stage.

We then examined direct binding of the chimeric TF to the promoter region of these genes by ChIP-qPCR assays using parasites expressing the chimeric TF fused with GFP at the C-terminus (AP2-O::SP::GFP parasites). In these parasites, GFP signals were detected in the nucleus of ookinetes, as was seen in AP2-O::GFP parasites (data not shown). As shown in Fig. 4C, IP with antibodies against GFP significantly enriched the upstream regions of sporozoite-specific genes, including those induced in AP2-O::SP parasites (*SPECT*, *SPECT2*, *CS*, *P. berghei* orthologue of PF11_0545, PB000863.01.0 and PB000522.03.0) and also *MAEBL*, *UIS3*, *TRAP* and *SERA* (serine repeat antigen) *8*/egress cysteine protease 1 (Aly and Matuschewski, 2005), which were not significantly induced in AP2-O::SP parasites. In contrast, binding to the upstream regions of *SOAP* (secreted ookinete adhesive protein) and *chitinase*, which are ookinete-specific genes (Dessens *et al.*, 2003), were not observed. These results demonstrated that the AP2 domain of AP2-Sp specifically binds to the upstream region of sporozoite-specific genes *in vivo*.

Discussion

Proteins expressed in sporozoites are potential targets for vaccine development, but the mechanisms of gene regulation in this stage remain poorly understood. Here we have provided evidence that the AP2 family TF, AP2-Sp, widely regulates gene expression in the sporozoite stage. This TF is expressed in the period from the late oocyst to the salivary gland sporozoite and induces the expression of genes with various functions and expression profiles. Target genes of this TF include all known genes specifically expressed in sporozoites, suggesting that this TF plays a central role in gene regulation of the sporozoite stage. We also demonstrated that AP2-Sp induces these target genes by binding to specific *cis*-acting elements in proximal promoter regions and that these elements are critical for promoter activity of each gene.

In a previous paper we reported that another AP2 family TF, AP2-O, regulates gene expression in the ookinete stage. Similar to AP2-Sp, Ap2-O directly activates several genes, including all known genes specifically expressed in the ookinete stage. AP2-O also shares some unique features of gene activation with AP2-Sp, such as binding to the proximal promoter region and additive effects of binding sites with respect to promoter activity. This indicates that mechanisms of gene regulation are closely related in these two stages.

However, a difference in gene expression between these two stages is that the sporozoite stage is composed of two forms with distinct infection ability, salivary and oocyst sporozoites, and thus target genes of AP2-Sp include those exhibiting distinct expression patterns such as *MAEBL* and *SPECT2*. Because such distinct expression patterns are difficult to explain solely by interactions between AP2-Sp and its binding sites on the promoter, there may be other TFs or co-regulators that modulate the function of AP2-Sp. Whereas at present we have not yet succeeded in identifying them, further investigation of AP2 family TFs in this stage and promoter analysis of genes exhibiting distinct expression patterns using the centromere plasmid might answer the question of how *Plasmodium* sporozoites change infection capability in the mosquito.

AP2-Sp (–) parasites form oocysts of normal size and number but they fail to generate sporozoites. This phenotype resembles to that of CS (–) parasites (Menard *et al.*, 1997), suggesting that loss of CS protein in AP2-Sp (–) parasites could be the cause of the phenotype. However, the possibility cannot be excluded that other genes involved in sporozoite formation were also regulated by AP2-Sp and that the reduced expression of these genes contributed to the phenotype. Indeed, expression of AP2-Sp starts 5 or 6 days before sporozoite budding, therefore a set of genes encoding cytoskeletal proteins, including those necessary for apical complex formation, could be under the control of AP2-Sp. In a previous paper (Yuda *et al.*, 2009), we reported that *AP2-O* (–) parasites could not form ookinetes of normal elongated shape, which suggested that AP2-O controls genes that participate in the morphogenesis of the invasive stage. Considering the analogous role of AP2-Sp and AP2-O in the respective motile stages it is possible that AP2-Sp regulates a group of genes necessary for formation of the sporozoite.

The chimeric TF induced several sporozoite-specific genes in ookinetes. ChIP-qPCR assays demonstrated that the chimeric TF activated these genes by directly binding to their promoter regions. This indicated that AP2-Sp is the *trans*-factor that interacts with the *cis*-elements. In this assay, however, the expression levels of the induced genes were not as high as those of the original ookinete-specific genes (data not shown) and significant induction was not observed in many genes, including *MAEBL*, *UIS3* and *TRAP*. It remains unclear

why this expression system did not work well on these promoters, but it is possible that co-activators of AP2-Sp, which would not be expressed in the ookinete, might be required for full activation of these promoters. Although further improvement may be required, this system could be a useful tool to determine target genes of AP2-family TFs. On the other hand, the results of the ChIP-qPCR assays suggested that the proximal promoter regions of sporozoite-specific genes are readily accessible in the ookinete stage where these genes are not primarily expressed.

Our study has demonstrated that AP2-Sp plays a central role in stage-specific gene expression in sporozoites and that *cis*-acting motifs in the proximal promoter region determine gene expression in this stage. Because the eight-base binding sequences for AP2-Sp are rarely found in the genome, these elements could mark sporozoite-specific genes on the genome. The present findings should pave the way for predicting genes expressed in this stage, based solely on genome sequence.

Experimental procedures

Parasite preparations

Female BALB/c mice (6–10 weeks old, Japan SLC, Hamamatsu, Japan) were infected with *P. berghei* ANKA strain as described previously (Yuda *et al.*, 2009). Oocyst sporozoites and salivary gland sporozoites were harvested 14 and 24 days after an infective blood meal respectively. For nuclear staining of oocysts, mosquitoes were dissected 7, 10 and 14 days after an infective blood meal, and their midguts were stained with Hoechst 34580 (Molecular Probes, Eugene, OR, USA) (0.2 μg ml^{-1} PBS final concentration) for 10 min at RT. Ookinete culture was performed as described previously (Yuda *et al.*, 2009).

Preparation of GFP-tagged AP2-Sp expressing parasites

For the targeted insertion construct (Fig. 1A), a DNA fragment containing the 3′ part of the AP2-Sp coding region was amplified by PCR, using genomic DNA as a template, and inserted into the XhoI/NheI site of the construct in frame with the GFP gene. The downstream region of the AP2-Sp gene was also amplified by PCR and inserted into the pBluescript BamHI/NotI site. Plasmids containing the construct were digested with XhoI and NotI and then used for transfection. The PCR primer pairs used for preparing the construct and Southern hybridization probe are listed in Table S5.

Targeted disruption of the AP2-Sp *gene*

The targeting construct was prepared by PCR using essentially the same procedure as described previously (Yuda *et al.*, 2009). The PCR primer pairs used for preparing the targeting construct and Southern hybridization probe are listed in Table S5.

Electrophoretic mobility-shift assays

The AP2 domain of AP2-SP was produced as a glutathione S–transferase fusion protein using the GST Gene Fusion System (Amersham Bioscience, Tokyo, Japan) essentially as described previously (Yuda *et al.*, 2009). Briefly, the coding regions were amplified from *P. berghei* genomic DNA with the primer set: 5′-CGGGATCCGCCTTTAGGGTATTTGATGT AGAC-3′ and 5′-CCGCTCGAGTTAATACTTTAGTTTCATCA TTTCGC-3′ for the AP2 domain with AT-hook region (amino acid residues 144–248) and with the primer set: 5′-CGGGATCCGCCTTTAGGGTATTTGATGTAGAC-3′ and 5′-CCGCTCGAGTTAATACTTTAGTTTCATCATTTCGC-3′ for the AP2 domain without AT-hook region (amino acid residues 157–248). Probes were prepared by PCR with 5′-biotinylated primers using the cloned promoter region of the *MAEBL* gene as template. The primer pairs used were 5′-TAAAAA TGTAAAGCATTTGAATTAAGAACC-3′ and 5′-GATGTATTT TTTGTGTAGAAAAACTGAAGG-3′. Short double-strand oligonucleotide probes were prepared from synthetic 5′-biotinylated oligonucleotides. The original oligonucleotide sequence was 5′-TAATATTATTATGCATGCATTCTTATAAG-3′. EMSA was performed as described previously (Yuda *et al.*, 2009).

Reporter assays with P. berghei *centromere plasmids*

Reporter assays were performed with the *P. berghei* centromere plasmid, pCen-GFP and pCen-Luc (Fig. S3). The upstream region of the *MAEBL* or *SPECT2* gene was inserted into the EcoRV/BamHI site, upstream of the GFP gene of pCen-GFP. For luciferase assay, the upstream promoter region was inserted into the KpnI/XhoI site upstream of the luciferase gene in pCen-Luc. The PCR primer pairs used for amplification of the upstream regions are listed in Table S6. Transfection and sporozoite preparations were performed as described previously (Yuda *et al.*, 2009). The rate of GFP-positive parasites was calculated in each sporozoite preparation. Luciferase activity was measured using 10 000 sporozoites with the Luciferase Assay System (Promega Corp., Madison, WI, USA) and normalized to rates of GFP-positive parasites.

Preparation of mutant parasites expressing AP2-O with the AP2 domain of AP2-Sp (AP2-O::Sp parasites and AP2-O::Sp::GFP parasites)

The targeted insertion construct for AP2-O::Sp parasites was prepared as follows. Part of the AP2-O coding region, just upstream of the AP2 domain, was amplified by PCR and ligated in frame to a DNA fragment encoding the AP2 domain of AP2-SP. This fragment was inserted into the XhoI/BglII site of the targeted insertion construct for AP2-O::GFP (Fig. 4A) in place of the 3′ part of the AP2-Sp coding region and the GFP gene. To prepare the targeted insertion construct for AP2-O::Sp::GFP parasites, the same fragment with XhoI/NheI