



Supplementary Information for

Transcriptome profiling of *Plasmodium vivax* in *Saimiri* monkeys identifies potential ligands for invasion

Karthigayan Gunalan, Juliana M. Sá, Roberto R. Moraes Barros, Sarah L. Anzick, Ramoncito L. Caleon, J. Patrick Mershon, Kishore Kanakabandi, Monica Paneru, Kimmo Virtaneva, Craig Martens, John W. Barnwell, Jose M. Ribeiro and Louis H. Miller

Karthigayan Gunalan or Louis H. Miller

Email: Karthigayan.gunalan@nih.gov, LMILLER@niaid.nih.gov

This PDF file includes:

Supplementary text

Figs. S1 to S3

Tables S1 to S5

Other supplementary materials for this manuscript include the following:

Dataset S1 and Dataset S2

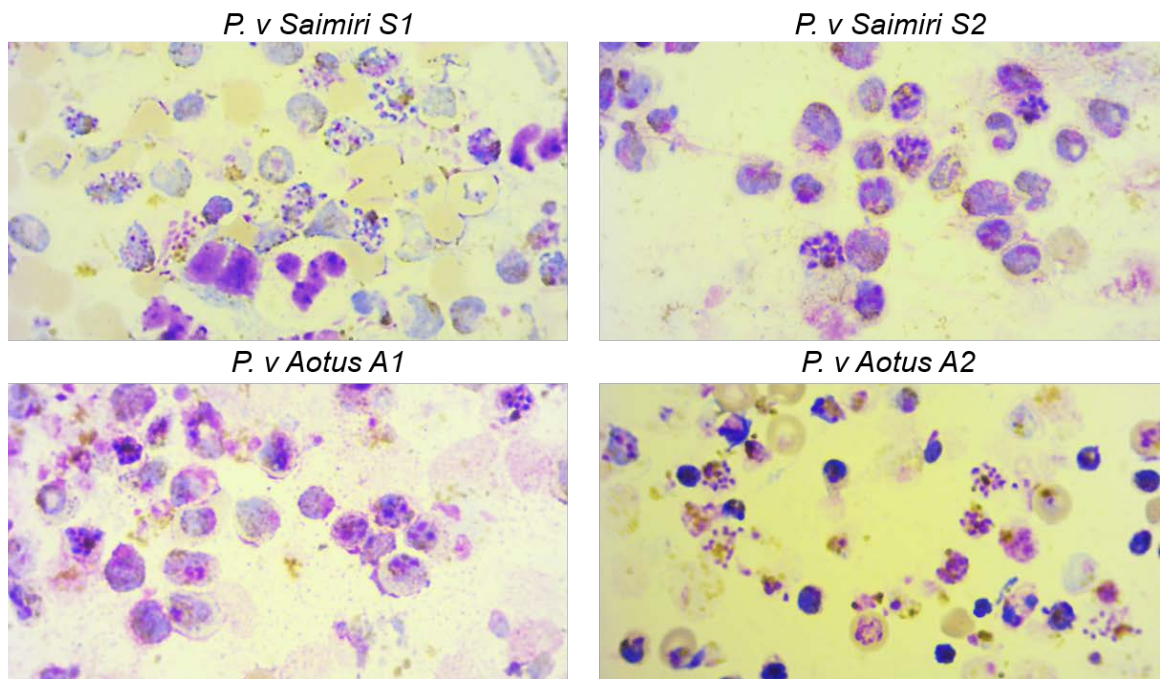


Fig. S1. *P. vivax* Sal I infection in *Saimiri* and *Aotus* monkeys. Two *Saimiri* monkeys and two *Aotus* monkeys were infected with the *P. vivax* Sal I parasites. Late schizont stage parasites were purified from monkey blood samples for RNA sequencing analysis. Following enrichment by culture, Giemsa smear was performed on two *Saimiri* and two *Aotus* infections. The panels show the representative smear images from two *Saimiri* and two *Aotus* infections.

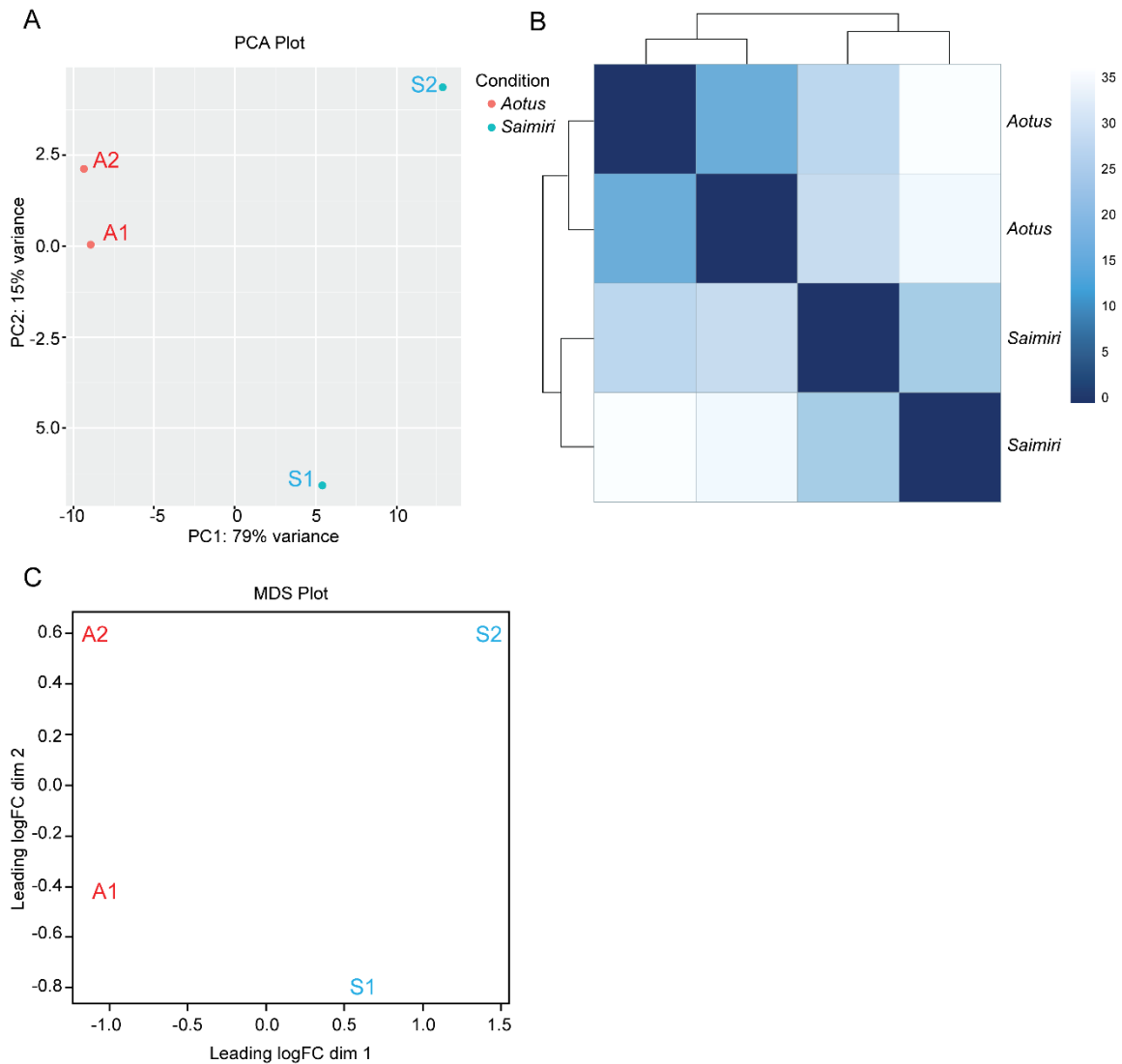


Fig. S2. Grouping of *P. vivax* infection in *Saimiri* and *Aotus* monkey using RNA seq data. A. Principal component analysis (PCA) plot shows the grouping of the two *Aotus* (Red) and two *Saimiri* (Blue) *P. vivax* Sal I infections based on the differentially expressed genes between the two groups using DESeq2. B. Heatmap of Euclidean distances between samples show the grouping of the two *Aotus* and two *Saimiri* samples using DESeq2. C. Multiple dimensional scaling (MDS) plot shows the grouping of the two *Aotus* (Red) and two *Saimiri* (Blue) *P. vivax* Sal I infections based on the differentially expressed genes between the two groups using edgeR.

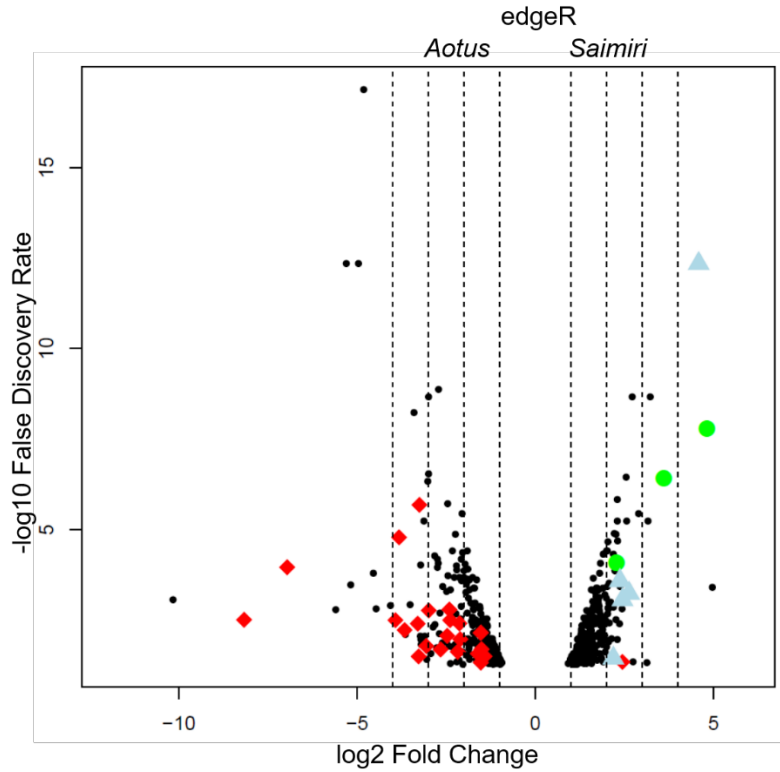


Fig. S3. Volcano plot showing the differential expression of genes in *Aotus* and *Saimiri* *P. vivax* infection by edgeR analysis. The volcano plot represents the significantly differential gene expression (False Discovery Rate < 0.05) of *P. vivax* schizonts from *Saimiri* or *Aotus* monkeys. The horizontal axis represents the expected log2 fold change as estimated by the program edgeR. The left four vertical dashed lines represents 16-, eight-, four- and twofold changes in *Aotus*-derived *P. vivax* genes, while the right four vertical lines represent two-, four-, eight- and 16-fold changes in *Saimiri*-derived *P. vivax* genes. The red diamonds represent transcripts coding for Vir proteins, the green circles represent transcripts coding for proteins of the MSP family, and the light blue triangles represent transcripts coding for proteins of the tryptophan-rich family. All genes can be identified in the Dataset S2.

Table S1. Differential expression of merozoite invasion ligands in *Saimiri* compared to *Aotus* monkey.

Entrez Gene	Description	DESeq2		edgeR	
		log2 Fold Change	Padj*	log2 Fold Change	FDR
<i>Saimiri P. vivax</i> infection					
PVX_090330	reticulocyte binding protein 2 precursor (PvRBP-2)	1.72	8.46E-02		
PVX_098585	reticulocyte-binding protein 1 (RBP1)	1.55	1.26E-02	1.44633	0.0221
PVX_125738	reticulocyte binding protein 1 precursor	0.67	5.10E-01		
PVX_098582	reticulocyte-binding protein 1 (RBP1), like	0.89	4.16E-01		
PVX_094255	reticulocyte binding protein 2 precursor	0.36	7.90E-01		
PVX_121920	reticulocyte-binding protein 2 (RBP2), like	0.11	9.23E-01		
PVX_116930	reticulocyte binding protein 2 homolog B	0.63	2.91E-01		
<i>Aotus P. vivax</i> infection					
PVX_090325	reticulocyte binding protein 2 precursor (PvRBP-2)	-0.05	9.72E-01		
PVX_101590	reticulocyte-binding protein 2 (RBP2), like	-0.85	5.71E-01		
<i>Saimiri P. vivax</i> infection					
PVX_092275	apical merozoite antigen 1	1.45	1.22E-02	1.44159	0.0173
PVX_110810	Duffy receptor precursor	1.06	1.54E-01		

The table shows the list of nine *P. vivax* reticulocyte binding proteins, Duffy binding protein 1 and Apical merozoite antigen 1 proteins that were differentially upregulated (positive) in *Saimiri* infections with the corresponding log2 fold change by DESeq2 and edgeR analyses. Italicized font indicates genes that were up-regulated above a log2 fold change cutoff of 1. * Genes with adjusted p value (padj) < 0.05. FDR: False-discovery rate.

Table S2. Differential expression of serine-repeat antigen (SERA) families in *Saimiri* and *Aotus* infection

Entrez Gene	Description	DESeq2		edgeR	
		log2 Fold Change	Padj*	log2 Fold Change	FDR
<i>Saimiri P. vivax</i> infection					
PVX_003820	serine-repeat antigen 4 (SERA)	2.44	5.51E-04	2.545943	0.000835
PVX_003840	serine-repeat antigen 3 (SERA)	2.09	5.92E-03	0	0
PVX_003825	serine-repeat antigen 4 (SERA)	2.01	1.34E-03	2.03052	0.002042
PVX_003810	serine-repeat antigen 5 (SERA)	2.01	7.13E-03	1.997015	0.017012
PVX_003830	serine-repeat antigen 5 (SERA)	<i>1.93</i>	9.65E-03	1.955192	0.022948
PVX_003800	serine-repeat antigen (SERA)	<i>1.67</i>	1.02E-02	1.493291	0.02619
PVX_003795	serine-repeat antigen (SERA)	<i>1.52</i>	3.10E-02	0	0
<i>Aotus P. vivax</i> infection					
PVX_003790	serine-repeat antigen (SERA)	<i>-1.46</i>	6.57E-03	-1.53943	0.00766

The table shows the number of serine-repeat antigens (SERA) that were upregulated (positive) in *Saimiri* infection or upregulated (negative) in *Aotus* infection with log2 fold change by DESeq2 and edgeR analyses. Bold font indicates genes that were up-regulated above a log2 fold change cutoff of 2; italicized font indicates genes that were up-regulated above a log2 fold change cutoff of 1. * Genes with adjusted p value (padj) < 0.05. FDR: False-discovery rate.

Table S3. Differential expression of new *P. vivax* gene families in *Saimiri* and *Aotus* infection.

Entrez Gene	Description	DESeq2		edgeR	
		log2 Fold Change	Padj*	log2 Fold Change	FDR
<i>Saimiri P. vivax</i> infection					
PVX_104690	Pv-fam-c protein	3.72	6.13E-03	3.128738	0.04566
PVX_001695	Phist protein (Pf-fam-b)	3.11	8.99E-08	3.158187	5.68E-06
PVX_002565	RAD protein (Pv-fam-e)	2.17	9.99E-05	2.033917	0.001434
PVX_094650	RAD protein (Pv-fam-e)	2.04	2.92E-03	2.300391	0.010668
PVX_089852	RAD protein (Pv-fam-e)	<i>1.87</i>	1.98E-02	0	0
PVX_089455	RAD protein (Pv-fam-e)	<i>1.72</i>	3.93E-02	0	0
PVX_089470	RAD protein (Pv-fam-e)	<i>1.58</i>	7.01E-03		
PVX_001710	Phist protein (Pf-fam-b)	<i>1.52</i>	4.59E-02		
<i>Aotus P. vivax</i> infection					
PVX_002515	Pv-fam-b protein	<i>-1.39</i>	7.80E-03	-1.402667	0.009692
PVX_089860	RAD protein (Pv-fam-e)	<i>-1.73</i>	1.19E-02	0	0
PVX_002525	Pv-fam-b protein	<i>-1.81</i>	3.28E-02	0	0
PVX_001705	Phist protein (Pf-fam-b)	-2.30	6.67E-05		
PVX_002530	Pv-fam-b protein	-2.38	5.44E-04	-2.437857	0.001674
PVX_101535	Phist protein (Pf-fam-b)	-2.50	2.87E-02	0	0
PVX_112710	Pv-fam-d protein	-2.66	1.10E-02	0	0
PVX_089467	RAD protein (Pv-fam-e)	-2.75	4.19E-06	0	0
PVX_003515	Phist protein (Pf-fam-b)	-2.80	9.33E-04	-2.69403	0.007233
PVX_101540	Pv-fam-d protein	-3.00	1.96E-02	-3.053298	0.037646
PVX_101580	Pv-fam-d protein	-5.00	3.37E-05	-5.186813	0.000321

The table shows the new *P. vivax* gene families that were upregulated (positive) in *Saimiri* and upregulated (negative) in *Aotus* infection with log2 fold change by DESeq2 and edgeR analyses. Bold font indicates genes that were up-regulated above a log2 fold change cutoff of 2; italicized font indicates genes that were up-regulated above a log2 fold change cutoff of 1. * Genes with adjusted p value (padj) < 0.05. FDR: False-discovery rate.

Table S4. Differential expression of ETRAMP in *Saimiri* and *Aotus* infection.

Entrez Gene	Description	DESeq2		edgeR	
		log2 Fold Change	Padj*	log2 Fold Change	FDR
<i>Saimiri P. vivax</i> infection					
PVX_003565	early transcribed membrane protein (ETRAMP)	<i>1.04</i>	4.21E-02	1.3107897	0.0411052
<i>Aotus P. vivax</i> infection					
PVX_121950	early transcribed membrane protein (ETRAMP)	<i>-1.91</i>	4.09E-03	0	0
PVX_118680	early transcribed membrane protein (ETRAMP)	-2.35	2.89E-06	-2.335045	3.688E-05
PVX_088870	early transcribed membrane protein (ETRAMP)	-5.13	1.36E-18	-1.314227	0.0224803
PVX_001715	early transcribed membrane protein (ETRAMP)	-5.48	2.13E-32	-5.312396	4.498E-13

The table shows the number of *P. vivax* early transcribed membrane proteins (ETRAMP) that were upregulated (positive) in *Saimiri* infection or upregulated (negative) in *Aotus* infection with log2 fold change by DESeq2 and edgeR analyses. Bold font indicates genes that were up-regulated above a log2 fold change cutoff of 2; italicized font indicates genes that were up-regulated above a log2 fold change cutoff of 1. * Genes with adjusted p value (padj) < 0.05. FDR: False-discovery rate.

Table S5. List of primers for qRT-PCR.

Gene	PlasmoDB ID	Oligo Type	5'-3' sequence
Apical merozoite antigen 1	PVX_092275	For	ATTGCAACAACGGCGTTATCT
		Rev	GATTCATGTTCCCTCGATTGTTTCTT
Duffy Binding Protein 1	PVX_110810	FAM-BHQ1	ACCCACAAGAAGTAGACCTGGAGTTCCCC
		For	ATGTGGACAATGCTGCTAAATCTC
		Rev	ATGCGTAGAATCTCCTGGAACCT
Phist protein (Pf-fam-b)	PVX_001695	FAM-BHQ1	CCACCAATTCAAATCCGATAAGTCAGCCT
		For	TGCTGGTGGGATTAATGTTCTT
		Rev	AGGACGCCGCTAGTTTTCTCT
Pv-fam-c protein	PVX_104690	FAM-BHQ1	CGAGCAGCCGCTTAGCACCAAAGA
		For	GATGAAGAAGCATATGAAACAGAAAATTC
		Rev	TCAGATCTTCCAGGTTAGGATCATCT
Merozoite Surface Protein 3 gamma (MSP3g)	PVX_097670	FAM-BHQ1	TCCAGGAACGCCAAACCATCATG
		For	CCCCAGTGGAAGTGATAATCT
		Rev	CCTGGGGTTCCGCATCT
Merozoite Surface Protein 1	PVX_099980	FAM-BHQ1	CCGTTTCGTTCTGCTGTTCTCCGACTT
		For	TGACGCAGAGGAGTTGGAGTAC
		Rev	CTCTTTCAGGGTTGGAAGTGTCT
Tryptophan-rich antigen (Pv-fam-a)	PVX_112690	FAM-BHQ1	CACGCTAATCATCCCAGAAAGCACCAAT
		For	GAGCATAGCTTCATAAATTTATAATTC
		Rev	TCTGTACTTTGTTATAACGATAGTTT
Hypothetical Protein	PVX_094230	FAM-BHQ1	ACGAACGCAAAGAAATAAATAATTTGCAC
		For	AACCATTGGCGCCTGTTC
		Rev	CAGACACCGCCGGTTCTT
Alanyl tRNA synthetase	PVX_115345	FAM-BHQ1	CACCCACGGCACCCCTGCACA
		For	TTCCCCGTGAGGCTTCTATTC
		Rev	CCCCTAGGTGGAAGAAAAGCTT
Fumarate Hydratase	PVX_099805	CFO-BHQ1	CATCAAGCCTAAGGGAAACAACGCC
		For	CTGTCTCCCCTGAGCATGTTCT
		Rev	GCGCACGGAGCTGTAATCT
Glutathione S-transferase	PVX_085515	CFO-BHQ1	ACGAAGAACAACCTTGCCGTGCCAGATT
		For	CGAGCTGATTCGGCTGATCT
		Rev	TCGCCGAAGCGCTTGT
		CFO-BHQ1	CGCCTACCTGGGCATCCAGTACACG