

Supplementary Material

Contrasting evolutionary trajectories of pyrimethamine resistance in human and zoonotic malaria parasites

Chaturong Putaporntip¹, Jidapha Somkuna¹, Rattanaporn Rojrung¹, Taweesak Tia¹, Napaporn Kuamsab², Somchai Jongwutiwes¹

¹Molecular Biology of Malaria and Opportunistic Parasites Research Unit, Department of Parasitology, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand;

²Community Public Health Program, Faculty of Health Science and Technology, Southern College of Technology, Nakorn Si Thammarat 80110, Thailand

List of Supplementary Data

Supplementary Table S1	Locations and collection periods of samples
Supplementary Table S2	PCR primers and amplification conditions
Supplementary Table S3	Sequence validation and clonal variation analysis of the <i>dhfr-ts</i> locus in human and zoonotic malaria isolates
Supplementary Table S4	Distribution of PfDHFR-TS haplotypes
Supplementary Table S5	Distribution of PvDHFR-TS haplotypes
Supplementary Table S6	Amino acid substitutions in PkDHFR-TS
Supplementary Table S7	Amino acid substitutions in PiDHFR-TS
Supplementary Table S8	Repeats in the junction region of PiDHFR-TS
Supplementary Table S9	Molecular diversity of <i>dhfr-ts</i> in <i>P. knowlesi</i> and <i>P. inui</i> based on all identified haplotypes and a single representative haplotype per infection
Supplementary Figure S1	Sequencing primers for <i>Pfdhfr-ts</i> , <i>Pvdhfr-ts</i> , <i>Pkdhfr-ts</i> and <i>Pidhfr-ts</i>
Supplementary Figure S2	Analytical workflow and sequence verification process
Supplementary Figure S3	Distribution of amino acid substitution in PkDHFR-TS relative to active site, binding pocket, and dimer interface
Supplementary Figure S4	Distribution of amino acid substitution in PiDHFR-TS relative to active site, binding pocket, and dimer interface
Supplementary Figure S5	Neighbor-joining trees inferred from <i>Pkdhfr-ts</i> and <i>Pidhfr-ts</i> of Thai isolates

Supplementary Table S1 Locations and collection periods of samples

Species	Sources* (n)	Collection sites (n)	Years (n)
<i>P. falciparum</i>	Humans (n = 241)	Tak (n = 165), Chanthaburi (n = 8), Ubon Ratchathani (n = 44), Narathiwat (n = 12), Yala (n = 12)	1994 (24), 1995 (25), 1996 (51), 1997 (41), 2007 (7), 2008 (29), 2012 (8), 2013 (12), 2014 (44)
<i>P. vivax</i>	Humans (n = 115)	Tak (49), Ubon Ratchathani (27), Chanthaburi (8), Kanchanaburi (6), Narathiwat (16), Yala (9)	1996 (17), 2007 (36), 2008 (6), 2010 (7), 2013 (12), 2014 (14), 2016 (13), 2017 (2), 2022 (2), 2023 (2), 2024 (4)
<i>P. knowlesi</i>	Humans (n = 4), Long-tailed macaques (n = 1), Pig-tailed macaques (n = 13))	Humans: Songkhla (n = 1), Yala (3) Macaques: Yala (n = 1), Narathiwat (n = 13)	Humans: 2018 (3), 2020 (1) Macaques: 2008 (8), 2009 (2), 2011 (2), 2018 (2)
<i>P. inui</i>	Long-tailed macaques (n = 14), Pig-tailed macaques (n = 50)	Narathiwat (n = 49), Chanthaburi (n = 9), Ranong (n = 6)	2006 (14), 2008 (13), 2009 (15), 2011 (12), 2013 (5), 2018 (4), 2019 (1)

* All human isolates were from uncomplicated malaria patients. Long-tailed macaque, *Macaca fascicularis*; pig-tailed macaque, *Macaca nemestrina*.

Supplementary Table S2 PCR primers and amplification conditions

<i>Plasmodium</i> species	Primers	<i>dhfr-ts</i> sequence (5' → 3')	Thermal cycler profile	PCR products (bp)
Primary PCR				
<i>P. falciparum</i>	PFDRF0	CCAACATTTTCAAGATTGATACATAAAG	94°C, 1 min; 20 cycles of 94°C, 40 s; 48°C, 30 s and 72°C, 2 min; a final extension at 72°C, 5 min.	1985
	PFDRR0	TTCTTAACCGTTCAGGTAATTTTGTC		
<i>P. vivax</i>	PVDRF0	GCACCTGCGCGCACCCCACCACATAAC		2022
	PVDRR0	CAAACCGCCTGAACAGCAATAAC		
<i>P. knowlesi</i>	PKDRF0	CTGCACATACATCGTCACATAAC		1950
	PKDR0	TTTTGTCCACGCTAGGCGGCCA		
<i>P. inui</i>	PIDRF0	GCGCACACGCCATCCCATAACG	1964	
	PIDRR0	CTCATTTTGGACAAGCATTAGGC		
Secondary PCR				
<i>P. falciparum</i>	PFDRF0	CCAACATTTTCAAGATTGATACATAAAG	94°C, 1 min; 25 cycles of 94°C, 40 s; 55°C, 30 s and 72°C, 2 min; a final extension at 72°C, 5 min.	1976
	PFDRR1	GTTCAGGTAATTTTGTCATCATTTGTTC		
<i>P. vivax</i>	PVDRF1	CATAACGAAGTAGGAGTAACCGCTTA		1988
	PVDRR1	CAGCAATAACACTTCGCTAGCTAGCTAT		
<i>P. knowlesi</i>	PKDRF1	CAAAGTATAAGAATAACCTTTTAAGC		1913
	PKDRR1	GCGGCCATCTCCATAGCGAT		
<i>P. inui</i>	PIDRR1	CGAAGGAGTAACCTCTTAACCCCTG	1923	
	PIDRR1	GCAGCCATGTCCATAGGAAT		

Supplementary Table S3 Sequence validation and clonal variation analysis of the *dhfr-ts* locus in human and zoonotic malaria isolates

Species	No. isolates	Validation methods, n				Clonal infections, n		No. verified sequences
		# 1	# 2	# 3	# 4	Single	Multiple	
<i>P. falciparum</i>	241	240	1 (20*)	-	(20*)	241	-	241
<i>P. vivax</i>	115	111	4 (20*)	-	(20*)	115	-	115
<i>P. knowlesi</i>	18	9	(9*)	9	(9§)	9	9 (double, n=7; triple, n=2)	29
<i>P. inui</i>	64	25	(25*)	26	13 (20§)	25	39 (double, n=26; triple, n=13)	116

Validation Protocols and Workflow after Supplementary Figure S2

1. **Electropherogram Quality:** Identification of high-quality electropherograms exhibiting recurrent nucleotide substitutions across the dataset.

2. **Independent Replication:** Consistency between direct sequencing results from two independent PCR amplicons generated using high fidelity DNA polymerase.

3. **Clone-to-Consensus Alignment:** Verification that sequences from recombinant plasmid clones matched the consensus sequences derived from concordant, superimposed signals in the electropherograms.

4. **Clone Consistency:** Confirmation of identical sequences between recombinant plasmid clones derived from separate, independent PCR amplicons.

Cross-validation Metrics

* Number of isolates selected from samples validated via Method # 1

§ Number of isolates selected from samples validated via Method # 3

Supplementary Table S4 Distribution of PfDHFR-TS haplotypes

Province	Year	Haplotypes (%)				Total
		<u>NRNI</u>	<u>NRNL</u>	<u>IRNI</u>	<u>IRNL</u>	
Tak	1994	0 (0)	5 (20.8)	3 (12.5)	16 (66.7)	24
	1995	1 (4)	1 (4)	2 (8)	21 (84)	25
	1996	1 (2)*	2 (3.9)	1 (2)	47 (92.1)	51
	1997	1 (2.4)	0 (0)	6 (14.6)	34 (83)	41
	2007	0 (0)	0 (0)	2 (28.6)	5 (71.4)	7
	2008	0 (0)	0 (0)	4 (23.5)	13 (76.5)	17
Chanthaburi	2012	1 (12.50)	0 (0)	2 (25)	5 (62.50)	8
Ubon Ratchathani	2013	0 (0)	0 (0)	5 (41.7)	7 (58.3)	12
	2014	0 (0)	0 (0)	16 (50)	16 (50)	32
Narathiwat	2008	0 (0)	0 (0)	12 (100)	0 (0)	12
Yala	2014	0 (0)	0 (0)	12 (100)	0 (0)	12
Total		4 (1.66)	8 (3.31)	65 (26.97)	164 (68.05)	241

Note: Haplotypes inferred from the amino acid positions 51, 59, 108 and 164. Mutations are underlined.

* An isolate contained a synonymous mutation c.438T>G (p.Gly) in the TS domain.

Supplementary Table S5 Distribution of PvDHFR-TS haplotypes

Haplotype	Amino acid position (codon)															Total	Origin, Year (n)
	8	1	5	5	6	6	9	9	9	1	1	1	1	1	2		
Salvador I	F	I	F	S	T	Y	N	T	H	G	G	D	S	I	A		
	(TTT)	(ATC)	(TTC)	(AGC)	(ACG)	(TAT)	(AACACACACGGTGGTGAC)	(AGC)	(ATT)	(GCC)							
Ia	.	.	.	.	.	(tac)	.	.	.	.	.	.	.	.	.	4	Ubon Ratchathani 2014 (1), 2016 (1); Kanchanaburi 2023 (1), 2024 (1)
Ib	.	.	.	.	.	(tac)	.	.	.	.	.	.	.	.	(gca)	1	Kanchanaburi 2024 (1)
II	.	.	.	.	.	.	-	-	-	-	-	-	N (AAC)	.	.	1	Tak 1996 (1)
III	.	.	L (TTG)	.	.	.	.	.	.	.	.	.	.	F (TTT)	.	1	Kanchanaburi 2024 (1)
IV	.	.	.	R (AGG)	.	.	-	-	-	-	-	-	N (AAC)	.	.	63	Tak 1996 (4), 2007 (2), 2013 (1); Chanthaburi 2007 (6), 2013 (1); Yala 2007 (3), 2008 (6); Narathiwat 2007 (8), 2010 (7); Ubon Ratchathani 2014 (13), 2016 (12)
V	S (TCT)	.	.	R (AGG)	.	.	-	-	-	-	-	-	N (AAC)	.	.	1	Narathiwat 2007 (1)
VI	.	.	L (TTG)	R (CGC)	.	.	.	.	.	.	.	.	T (ACC)	.	.	1	Tak 2023 (1)
VII	.	.	I (ATA)	R (AGG)	M (ATG)	.	.	.	.	.	.	.	T (ACC)	.	.	36	Chanthaburi 2007 (1); Tak 1996 (10), 2007 (14), 2013 (9), 2017 (2);
VIII	.	.	L (TTA)	R (AGG)	M (ATG)	.	.	.	.	.	.	.	T (ACC)	.	.	5	Tak 1996 (1), 2007 (1), 2022 (1); Kanchanaburi 2022 (1), 2024 (1)
IX	.	L (CTC)	L (TTA)	R (AGG)	M (ATG)	.	.	.	.	.	.	.	T (ACC)	.	.	2	Tak 1996 (1), 2013 (1)
Total																115	Tak (49); Ubon Ratchatani (27); Chanthaburi (8); Kanchanaburi (6); Narathiwat (16); Yala (9)

Note: Dots and dashes represent identical and deleted residues relative to the pyrimethamine-sensitive Salvador I strain, respectively.
 Synonymous and nonsynonymous substitutions are shown respectively in lowercase and uppercase letters.
 Codon 246 is in the junctional region. No mutation occurs in the TS domain.

Supplementary Table S6 Amino acid substitutions in PkDHFR-TS

Amino acid haplotype	< DHFR ><																>< TS >								n	
	Amino acid position																									
	3	3	7	8	0	0	3	0	0	4	6	6	7	7	9	0	2	4	6	1	7	7	4			
Strain H	L	G	R	E	L	T	T	I	Q	F	G	E	P	N	T	A	D	N	H	M	K	G	V	H		
I*	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.	.	.	.	.	.	4	
II	.	.	L	.	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.	.	.	.	.	.	1	
III	.	.	.	G	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.	.	.	.	.	.	1	
IV	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.	D	.	.	.	.	.	.	.	1	
V	F	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.	.	.	I	.	.	1	
VI	.	V	.	.	.	.	.	.	.	.	.	.	T	S	.	.	.	.	.	.	.	.	.	.	2	
VII	.	.	.	G	.	.	.	V	.	.	.	.	S	.	.	.	.	.	.	.	.	.	.	.	1	
VIII	.	.	.	G	.	.	.	.	R	.	.	.	S	.	.	.	.	.	.	.	.	.	.	.	1	
IX	.	.	.	.	V	.	-	.	.	.	.	.	S	.	.	.	.	.	I	.	.	.	.	.	1	
X	.	.	.	.	.	.	.	.	R	.	.	.	S	.	.	.	D	.	.	.	.	.	.	.	1	
XI	.	.	.	.	.	.	.	.	.	R	.	T	S	.	.	.	.	.	.	.	.	.	.	.	1	
XII	.	V	.	.	.	.	.	V	.	.	.	.	T	S	.	.	.	.	.	.	.	.	.	.	1	
XIII**	.	.	L	.	.	.	.	.	.	.	.	.	R	.	.	Y	.	.	.	.	.	I	.	.	2	
XIV	.	.	L	.	.	.	.	.	.	.	.	.	S	I	V	.	.	.	.	.	.	.	.	.	2	
XV	.	.	L	.	.	.	.	.	.	.	.	.	R	.	.	Y	.	.	.	.	.	I	.	.	2	
XVI	.	.	.	.	V	.	-	.	.	.	.	.	S	.	.	.	.	.	I	.	.	.	R	.	2	
XVII*	.	.	.	.	.	.	.	.	.	L	.	.	S	.	.	.	.	N	.	.	D	.	.	.	1	
XVIII	F	.	.	.	.	A	.	.	.	.	.	.	S	I	V	.	.	.	.	.	.	.	.	.	1	
XIX	.	.	L	.	.	.	.	.	.	.	.	D	.	S	I	V	.	.	.	.	.	.	.	.	1	
XX	.	.	.	.	V	.	-	.	.	.	.	.	S	.	.	.	D	.	I	E	.	.	.	.	1	
XXI	F	.	.	.	.	A	.	.	.	.	.	.	S	I	V	.	.	.	.	E	.	.	.	.	1	
Total																								29		

Boundaries of domains are marked above amino acid positions. Dots and dashes represent no substitution and deletion, respectively. Four isolates from humans are indicated by asterisks.

Supplementary Table S7 Amino acid substitutions in PiDHFR-TS

Amino acid	<	DHFR						><	Junction			><	TS	>
Haplotype		22225778	8888899999	0000011444	4556677899	1112233444	4455667777		22222222	2222233333	3333333333	3333344444	555666	
		6903891171	2357801348	1236918378	9060112437	1580215035	672329156	repeats	2346789	0258901246	8016924620	5879613625	159146	
SanAntonio	EDASNESQEQ	RESGSDTQPT	SGHGDTVLLR	QEIHSKLQNQ	VEEQRNCLSV	GEAAGVCSC			TLPVTTT	REGAPACQQN	TAEVSNVEMN	QEQFKNHVVL	YEIPFP	
I	D....Q.K..	RT.	N...MT			.G.....	VIII		.P....A	.G.....			A	
II	D....K..	RT.	T	M..K		EG.....	V		.P....	...P....			.Q...A	
III	D....K..	RT.	R...N...T			EG.....	XIII		.P....	..A...H..	K...		A	
IV	D....K..	RT.	R...N...T	K		EG.T....	VI		.P....A	.G.....			.Q...A	
V	D....K..	RT.	R...N...T	S.	V.	.G.....	XVII		.PL....	..A...S..			.QM...A	
VI	D....K..	RT.	I...T			.G....N.	III		.P....		H		A	
VII	D....K..	RT.	R...N...T	M...		.G....G..	XIII		.P....	..A...H..	K...		.Q...A	
VIII	D....K..	RT.	T			.G....G..	VI		.P....				.Q...A	
IX	D....K..	RT.	T	R.....K	V.	EG.....	VIII		.P....A	.G.....		K...	.Q.Q.A	
X	D.....H	...N..RT.	..V....T	K		.G.....	VI		.P....	S....			A	
XI	D.....	..A...RT.	N.....T			.GD.CE..	XXII		.P....			K...	A	
XII	D.....	..EN..RT.	T	M...		.G.....	XIV		.P....	...T....	I.		.Q...A	
XIII	D.....	RT.	R...N...T		V.	.G....N.	XXI		.P....	...D.P....	.P.....			
XIV	D.....	RT.	T	K		EG.....	VIII		.P....	Y...	I.....		.Q...A	
XV	D.....	RT.	T	K		EG.....	XIII		.PL....	..A.....	R....		.Q...A	
XVI	D.....	RT.	T		M	.G.....	XVI		.P....				.Q...A	
XVII	.NT.....H	...N..RT.	T	M...		.G....G..	XI		.P....	...P....		R.D..	.Q....	
XVIII *	..T...TK..	IRT.	T	K	R.	.G....G..			.P....				.Q...A	
XIX	..T...K..	RT.	I...T			.G.....	XII		.P....		P.....		A	
XX	..T...K..	RT.	T			EG.....	XIII		.PL....	..A.....	R....		.Q.QL.	
XXI	..T...K..	RT.	T	K		.G.....	XI		.P....M				.Q...A	
XXII	..T...K..	RT.	T			.G.....	V		.P....	...P....		I...	A	
XXIII	..T.....H	...N..RT.	.A.....T	K		EG.....	XIII		.PL....	..A.....			A	
XXIV	..T.....H	...N..RT.	T	M...		.G.....	V		.P....	...P..H..			.Q...A	
XXV	..T.....H	...N..RT.	T			.G.....	V		.P....				.Q...A	
XXVI	..T.....H	...N..RT.	T	M...		.G.....	V		.P....		L.	H.....L.	.Q...A	
XXVII	..T.....H	...N..RT.	T	M...		.GD..E..	XXII		.P....				A	
XXVIII	..T.....H	...N..RT.	T	K		.G.....	XV		.P....	P.....	K...		A	
XXIX	..T.....	..A...RT.	T			.G.....	XVII		.P....	...T....		..R.....	A	
XXX	..T.....	RT.	T	M...		.G.....	VI		.P.G...	H.	..KEN....	HK.....	.Q...A	
XXXI	..T.....	RT.	T	K		.G.....	V		.P....	...P....			A	
XXXII	..T.....	RT.	T			.G....G.N.	XXV		.P....	...S....		K...	.Q...A	
XXXIII	..T.....	RT.	T			.G.....	VIII		.P....	Y...	I.....I.	K...	.Q...A	
XXXIV	..T.....	T	T			.G.S....	VI		.P....			K...	A	
XXXV	..T...K..	RT.	T		..QD....	.GS...Y..	VI		.P....A	.G.....			.Q...A	
XXXVI	..T.....H	...N..RT.	T	K		.G....N.	I		.P....	...P....			.Q...A	
XXXVII	..T.....	RT.	T	K		EG.....	VIII		.P....A	.G.....	..K.....		.Q...A	
XXXVIII	...Y.....	RT.	T	M...		.G.....	V		.P....	...P....		K...	A	

Amino acid	<	DHFR						><	Junction	><	TS	>	
Haplotype		22225778	8888899999	1111111111	1111111111	2222222222	222222222		2222222	2222233333	3333333333	3333344444	555666
		6903891171	2357801348	0000011444	4556677899	1112233444	445566777		8888888	9999900000	0111122234	4488924788	366012
				1236918378	9060112437	1580215035	672329156	repeats	2346789	0258901246	8016924620	5879613625	159146
SanAntonio	EDASNESQEQ	RESGSDTQPT	SGHGDTVLLR	QEIHSKLQNG	VEEQRNCLSV	GEAAGVCSC			TLPVTTT	REGAPACQQN	TAEVSNVEMN	QEQFKNHVVL	YEIPFP
XXXIX	Q.K..	RT.	T	M...		.G.....	I		.P.....	P.....			.Q...A
XL	K.H	...N..RT.	.A.....T		F..	.G....CY	XVII		.P..M..				.Q....
XLI	K..	S...N..RT.	...N..I.T			.G.....	III		.P....	P....		K....	.Q....
XLII	K..	S...N..RT.	...N..I.T			.G.S....	XI		.P....M	...T.....		K....	.Q...A
XLIII	K..	S...N..RT.	...N..I.T			.G.....	XI		.P....M				.Q...A
XLIV	K..	.K....RT.	T	K	.Q.....	EG.....	XIII		.P.....	..A...H..	K....	K....	A
XLV	K..	.K....RT.	T	M...		.G.....	V		.P.....	...P.....			.Q...A
XLVI	K..	.K....RT.	I..T	K		.G.....	V		.P.....	...P.....			.QM..A
XLVII	K..	.K....RT.	T		.Q.....	.G.....	V		.P.....	...P.....			A
XLVIII	K..	...N..RT.	T			.G.....	III		.P.....		L.	H.....	.Q...A
XLIX	K..	RT.	N.....T	K...T....	...H.....	.G.T....	III		.P.....		M...		.Q....
L	K..	RT.	...AM..T	.K...E.K..		.G.....	XX		.PH....	...P.....			A
LI	K..	RT.	...I..T	M...		.G.....	V		.P.....	...P.....		K....	.Q.Q.A
LII	K..	RT.	T	...Q.....	I.....V..	.G.....	VI		.P.....	...P.....	Q..	K....	.Q...A
LIII	K..	RT.	T	M...	M	.G.....	V		.P.....		L.	H.....	A
LIV	K..	RT.	T	M...		.G....N.	I		.P.....	...P.....			A
LV	K..	RT.	T	M...		.G.....	VI		.P.....	...P.....		H.....	.Q...A
LVI	K..	RT.	T	M...		.G...G..	VI		.P.....		M...		.Q....
LVII	K..	RT.	T	M...		.G.....	IX		.P.....				A
LVIII	K..	RT.	T	K		EG.....	XIII		.P.....	..A...H..	K....		.Q...A
LIX	K..	RT.	T	K		EG.....	XVII		.PL....	..A.....	R....		.Q...A
LX	K..	RT.	T		.QD.....	.GS...Y..	VI		.P....A	.G.....			.Q..LA
LXI * #	K..	RT.	T		...S....	.G.....	XVIII		.P.....	P.....			A
LXII	K..	IRT.	T			.G.S....	VI		.P.....	...P.....			.Q...A
LXIII	K..	RT.	T			.G.....	VIII		.P.....	...Y....		K....	.Q...A
LXIV	Q.	S...N..RT.	T	M.S.		.G.....	V		.P.....	...P.....			A
LXV	Q.	S...N..RT.	T	M...	V.	.G.....	V		.P.....	...P.....			A
LXVI	H	...N..RTS	...V...T	M...		.G.....	VII		.P.....	P.....	H	V	A
LXVII	H	...N..RT.	T	M...	M	.G....N.	III		.P.....			..R.....	.Q...A
LXVIII		.A...RT.	T	K	.G.....	.G.....	V		.P.....	...PS....			A
LXIX		.A...RT.	...N...T	...Q.....		.G...G.N.	XXV		.P.....	...S....		K....	.Q...A
LXX		.A...RT.	...N...T	...Q.....		.G...G.N.	XXIV		.P.....	...S....		K....	H...A
LXXI		.A...RT.	T			.GS....	VI		.P.....				A
LXXII		..EN..RT.	T	K...T....	...H.....	.G.T....	III		.P.....		L.	H.....	.Q...A
LXXIII		...N..RT.	...E...T	M...		.G.....	IV		.P.....	...P.....			A
LXXIV		...N..RT.	T	...Q.....		.G.....	V		.P.....	...P.....	H		.Q...A
LXXV		...N..RT.	T			.G.....	V		.PH....	...P.....			.Q...A
LXXVI		...N..RT.	T			.G.....	VI		.P....A	.G.....		K....	.Q...A
LXXVII		...H..RT.	MT	.V.....		.G.....	VI		.P.....				.Q...A

Amino acid	<	DHFR						><	Junction				><	TS	>
Haplotype		22225778	8888899999	0000011444	4556677899	1112233444	445566777		2222222	222233333	3333333333	3333344444	555666		
		6903891171	2357801348	1236918378	9060112437	1580215035	672329156	repeats	2346789	0258901246	8016924620	5879613625	159146		
SanAntonio	EDASNESQEQ	RESGSDTQPT	SGHGDTVLLR	QEIHSKLQNQ	VEEQRNCLSV	GEAAGVCSC			TLPVTTT	REGAPACQQN	TAEVSNVEMN	QEQFKNHVVL	YEIPFP		
LXXVIII		H.RT.	T	K..	G.....			XVII	.P..M..				.Q...A		
LXXIX		RT.	N.....T	V..	G.....			XVII	.PL....	..A.....	R....		.QM..A		
LXXX		RT.	N.....T	V..	G.....			VI	.P....A	.G.....	Q..		.Q....		
LXXXI		RT.	N.....T	V..	G.....			VI	.P....		Q..		.Q....		
LXXXII		RT.	N.....T		G...CE..			XXII	.P....				.Q...A		
LXXXIII		RT.	N.....T		G...G..			X	.P....			K...	A		
LXXXIV		RT.	N.....T		G.....			VI	.P....		P.....		.Q...A		
LXXXV		RT.	.AN.N...T		G.....			XXIII	.P...N.			K...	.Q...A		
LXXXVI		RT.	...V...T	M...	G.....			VI	.P....				.Q...A		
LXXXVII		RT.	...V...T	K	EG.....			XIII	.P....	..A...H..	K...	K...	A		
LXXXVIII		RT.	T	R.....	G.....			VIII	.P....	Y...	I.....I.	K..I.	.Q...A		
LXXXIX		RT.	T	.K.....	G.....			I	NT.....				.Q...A		
XC		RT.	T	M..K	G.....			1	NT.....			K...	A		
XCI		RT.	T	M...	EG.....			VIII	.P....		L.	K...	.Q...A		
XCII		RT.	T	M...	.GD.....			V	.PH....	...P.....			.Q...A		
XCIII		RT.	T	M...	G...G..			XI	.P....	.G.....		...L.....	A		
XCIV		RT.	T	M...	G.....			XIII	.PL....	..A.....			A		
XCV		RT.	T	M...	G.....			V	.PH....	...P.....			A		
XCVI		RT.	T	M...	G.....			V	.P....	..A...H..	K...		.Q...A		
XCVII		RT.	T	K	S...	EG.....		XIII	.P....	...P.....		K...	A		
XCVIII		RT.	T	K	EG.....			XIX	.PL....	..A.....	R....		A		
XCIX		RT.	T	K	EG.....			XIII	.P....	...P.....		K...	.Q.Q.A		
C		RT.	T	K	GS...G..			XI	.P....	...P.....			.Q...A		
CI		RT.	T		...Q.Q....	.G.....		VIII	.P...A	.G.....K			.Q...A		
CII		RT.	T		S...	.G.....		VI	.P...A	.G.....			.Q...A		
CIII		RT.	T		S...	.G.....		XVIII	.P....	P.....			A		
CIV		RT.	T		V..	.G.....		XVII	.PL....	..A.....			.QM..A		
CV		RT.	T		M	.G....N.		III	.P....	..A...H..	K...		A		
CVI		RT.	T		G...G.N.			XXV	.P....		I.		.Q...A		
CVII		RT.	T		G...G.N.			XXV	.P....	...S.....		K...	.Q...A		
CVIII *		RT.	T		G.....			XVII	.P....		I.		.Q...A		
CIX		T	.AN.N...T		G.....			XXIII	.P...N.			K...	.Q...A		
CX **		T	T		G.....			II	.P....	...T.....	I.		.Q...A		

* and ** denote 2 and 4 isolates with identical amino acid sequences, respectively.

A synonymous substitution at c.153A>C, p. Ser51Ser.

Repeats are highlighted (white text on black background) and labeled with the haplotype numbers from Supplementary Table S8.

Supplementary Table S8 Repeats in the junction region of PiDHFR-TS

Haplotype	Pentapeptide repeats	n
San Antonio 1*	SVPNC	0
I	-	5
II	NTPNC NTPNC	4
III	NVPNC NTPNC	7
IV	NVPNC NVPNC	1
V	SVPNY NTPNC	20
VI	SVPNC NTPNC	19
VII	SVPNC NTLNC	1
VIII	SVSNC NTPNC	9
IX	SVPNC NAPNC	1
X	NVPNC NTPNC SVPNC	1
XI	NVPNC NTPNC NVPNC	4
XII	NVPNC NTPNC NVPKC	3
XIII	SVSNC NTPNC NVPNC	11
XIV	SVANC NTPNC NVPNC	1
XV	SVPNC NTPNC NVPKC	3
XVI	SVPNC NTPNC SVPNC	1
XVII	SVPNC NTPNC NVPNC	9
XVIII	SVPKC NTPNC NVPKC	3
XIX	SVSNC NTPNC NVSNC	1
XX	SVSNC NTPNC NVPNC	1
XXI	NTPNC NTPNC NTPNC NVTNC	1
XXII	SVPNC TTQNC NTPNC NVPKC	3
XXIII	SVPNC NTRNC NTPNC NVPNC	2
XXIV	NTPNC NTPNC NTPNC NTPNC NVPKC	1
XXV	NTPNC NTPNC NTPNC NTPNC NTPNC NTPNC NTPNC NVPKC	4
Total		116

Dash indicates no repeats. *Amino acid positions 277 – 281 (GenBank accession no. XM008818109).

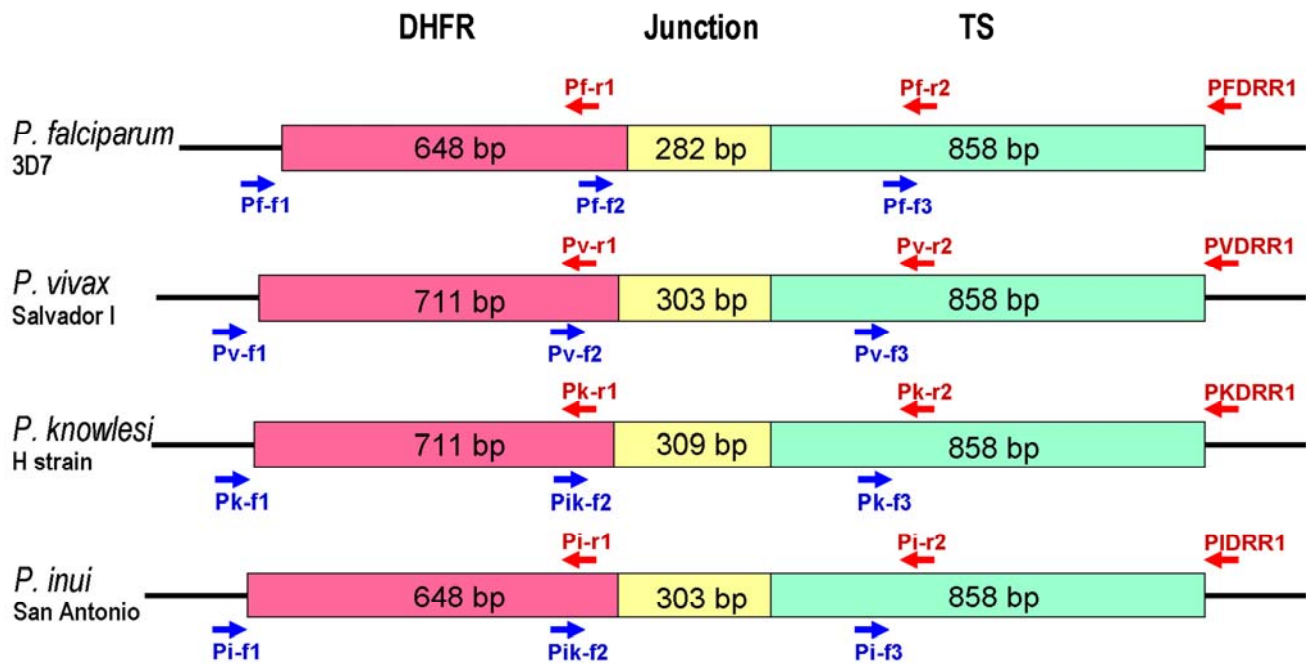
Supplementary Table S9 Molecular diversity of *dhfr-ts* in *P. knowlesi* and *P. inui* based on all identified haplotypes and a single representative haplotype per infection

Locus (n)	No. nt	M	S	H	$h \pm \text{S.D.}$	$\pi \pm \text{S.E. (\%)}$	$d_s \pm \text{S.E. (\%)}$	$d_N \pm \text{S.E. (\%)}$
<i>P. knowlesi</i> (29): All identified haplotypes								
<i>dhfr</i>	711	28	4	20	0.968 ± 0.018	0.895 ± 0.196	2.964 ± 0.724	0.314 ± 0.110
<i>junction</i>	306-309	13	5	10	0.879 ± 0.035	1.019 ± 0.329	2.256 ± 1.230	0.631 ± 0.217
<i>ts</i>	858	18	4	15	0.936 ± 0.023	0.536 ± 0.131	2.052 ± 0.240	0.137 ± 0.055
Total	1875-1878	60	13	26	0.990 ± 0.013	0.749 ± 0.108	2.437 ± 0.470	0.284 ± 0.057
<i>P. knowlesi</i> (18): One representative haplotype per infection								
<i>dhfr</i>	711	24	8	11	0.948 ± 0.030	0.806 ± 0.176	2.585 ± 0.817	0.309 ± 0.106
<i>junction</i>	306-309	13	3	9	0.895 ± 0.048	1.155 ± 0.354	2.933 ± 1.491	0.723 ± 0.281
<i>ts</i>	858	18	8	14	0.967 ± 0.030	0.505 ± 0.150	1.792 ± 0.589	0.167 ± 0.066
Total	1875-1878	55	19	16	0.987 ± 0.023	0.725 ± 0.109	2.274 ± 0.403	0.313 ± 0.073
<i>P. inui</i> (116): All identified haplotypes								
<i>dhfr</i>	732	107	18	104	0.998 ± 0.002	1.307 ± 0.179	2.987 ± 0.580	0.809 ± 0.171
<i>junction</i>	288-408	69	8	85	0.991 ± 0.003	2.691 ± 0.425	2.838 ± 0.770	1.716 ± 0.393
<i>ts</i>	858	60	18	91	0.995 ± 0.002	0.698 ± 0.125	2.590 ± 0.573	0.199 ± 0.076
Total	1878-1998	236	44	111	0.998 ± 0.002	1.259 ± 0.129	2.778 ± 0.325	0.667 ± 0.093
<i>P. inui</i> (64): One representative haplotype per infection								
<i>dhfr</i>	732	92	30	61	0.998 ± 0.003	1.297 ± 0.177	3.026 ± 0.573	0.807 ± 0.155
<i>junction</i>	288-408	66	21	51	0.988 ± 0.007	3.100 ± 0.471	3.971 ± 1.128	2.869 ± 0.680
<i>ts</i>	858	47	16	52	0.992 ± 0.005	0.671 ± 0.127	2.587 ± 0.576	0.164 ± 0.089
Total	1878-1998	205	67	61	0.988 ± 0.003	1.392 ± 0.120	2.970 ± 0.349	0.859 ± 0.120

H , number of haplotypes; M , number of mutations; S , number of singleton substitutions; π , the average number of pairwise nucleotide differences in the sample; d_s , number of synonymous nucleotide substitutions per site; d_N , number of nonsynonymous nucleotide substitutions per site.; nt, nucleotide; n, number of taxa.

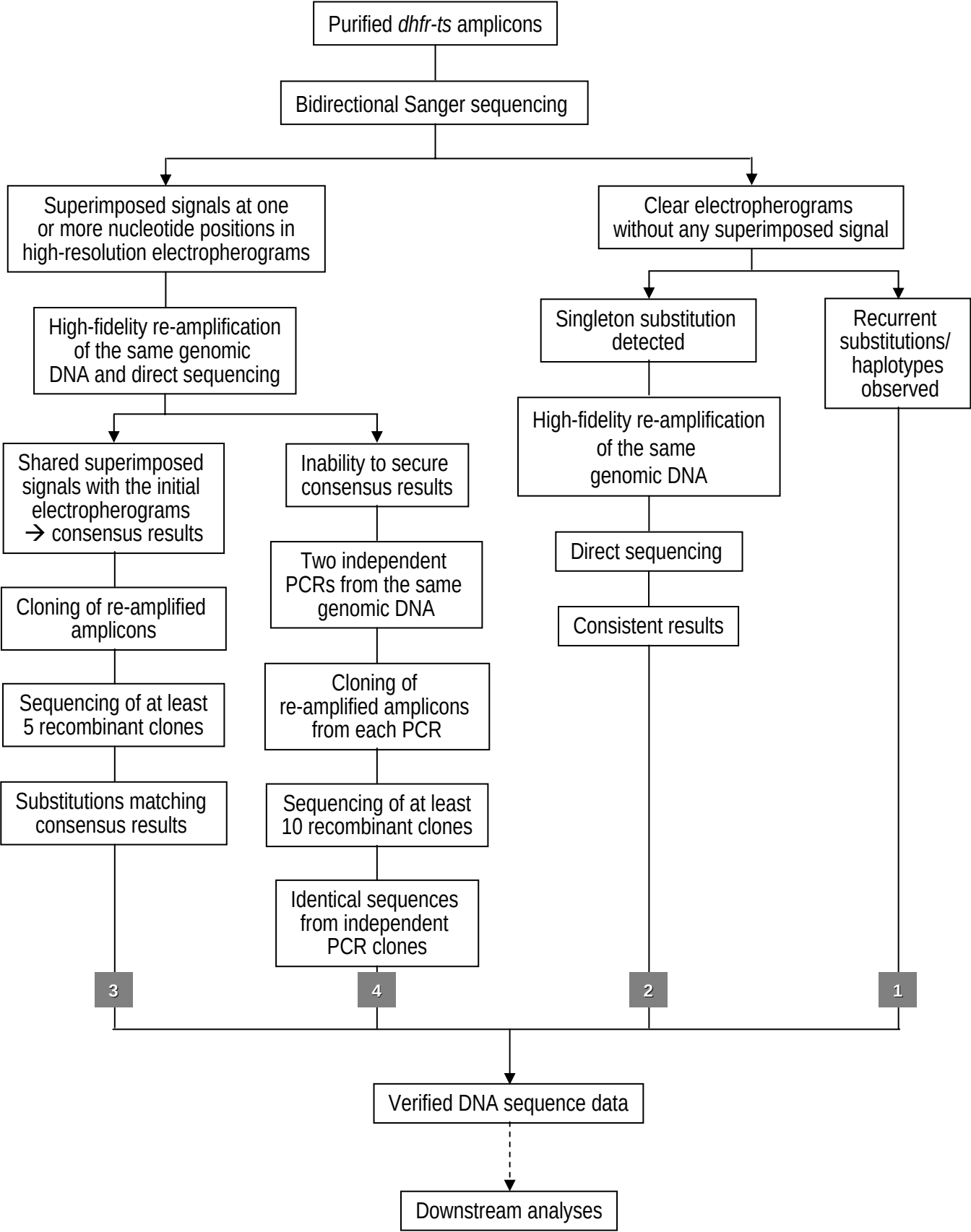
Tests of the hypotheses that π , d_N , and d_s of all identified haplotypes equals the corresponding values of a single representative haplotype per infection: all pairs yield p values > 0.05 , Z-test.

Supplementary Figure S1 Sequencing primers for *Pfdhfr-ts*, *Pvdhfr-ts*, *Pkdhfr-ts* and *Pidhfr-ts*



Locus	Primer	Sequence (5'→3')
<i>Pfdhfr-ts</i>	Pf-f1	GATTGATACATAAAGATAATATATA
	Pf-f2	GAAGATCTAATAGTTTTACTTGG
	Pf-f3	GAAGAATTGCTTTGGTTTATTAG
	Pf-r1	CATTCATATGTACTATTTATTCTAG
	Pf-r2	CCATTAGCTTCCCATATCCT
	PFDRR1	GTTCAGGTAATTTTGTTCATCATTTGTTC
<i>Pvdhfr-ts</i>	Pv-f1	GAAGTAGGAGTAACCGCTTAACC
	Pv-f2	CTCTTAAAGAAGCTGAAGTACTAC
	Pv-f3	GAACAGGTGTGGGAGTGATGA
	Pv-r1	GCTTGATTAAGTTTCTACTTAG
	Pv-r2	TTCTTTGTTGTTAATAGGGG
	PVDRR1	CAGCAATAACACTTCGCTAGCTAGCTAT
<i>Pkdhfr-ts</i>	Pk-f1	GAATAACCTTTTAAGCCCTGAGAC
	Pik-f2	TACTACAAATGCTTCATTATTGG (for <i>P. inui</i> and <i>P. knowlesi</i>)
	Pk-f3	GGGTATATGATGAAATTTAATTTAAATG
	Pk-r1	GCTTAATTAAGTTCCTATTTAG
	Pk-r2	CTCTTCAATTATTCCCCTCAG
	PKDRR1	GCGGCCATCTCCATAGCGAT
<i>Pidhfr-ts</i>	Pi-f1	GAGTAACCTCTTAACCCCTGAGAC
	Pik-f2	TACTACAAATGCTTCATTATTGG (for <i>P. inui</i> and <i>P. knowlesi</i>)
	Pi-f3	AAACAAGGAGACAGAACAGG
	Pi-r1	CCTATTTAGACATTCCTTATAAAC
	Pi-r2	TAGATCAAATCTCATCATGTA
	PIDRR1	GCAGCCATGTCCATAGGAAT

Supplementary Figure S2. Analytical workflow and sequence verification process



Supplementary Figure S3 Distribution of amino acid substitution in PkDHFR-TS relative to active site (yellow), binding pocket (red asterisks), and dimer interface (light blue). Singletons are shown as white text on black background

Supplementary Figure S4 Distribution of amino acid substitution in PiDHFR-TS relative to active site (yellow), binding pocket (red asterisks), dimer interface (light blue). Singletons are shown as white text on black background.

SanAntonio Variants	← DHFR															1	1111111111
	1234567890	1111111112	2222222223	3333333334	4444444445	5555555556	6666666667	7777777778	8888888889	9999999990	0000000001	1234567890	1234567890	1234567890	1234567890	1234567890	0000000001
	MEDLSEIFDI	YAI C ACCKVA	PPSEGIKNEP	FGPRTFRGLG	NNGTL L PWKCN	SVD M KY F SSV	TTYVDESKYE	QLKWKRERYL	QREASGGSD	TLQPGDNTNG	SGHTVGMGDD	1234567890	1234567890	1234567890	1234567890	1234567890	0000000001
	D	N	T	T	YQ	T	K	Q	HSK A	EN H	I RT	S	NAN	V	N	R	E
SanAntonio Variants	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111	1111111111
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
	TTKLQNIIVM	GRSSWES I PK	QYKPLPNRIN	VVLSKTLRQE	EVKEKIFLIH	SIEDLLLILK	KLRYYKCFI I	GGAQVYKECL	NRNLIKQIYF	TRINNAYPCD	VLFPEFDESQ	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
	I	M		I	MTKK	V	Q	T	E	K	S	K	I	Q	Q	H	D
	N			R													
SanAntonio Variants	DHFR →← Junction																
	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222	2222222222
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
	FRVTSVSEVY	NCKGCTLDL	IYSKVGEED	GAASNGSTAD	SGSENCCKVE	CGAPSCSVPN	CTLPNVTTTR	QEKWGSAPA	CQWQKNNTDA	EEDDLVYFSF	NNKVGEKNPE	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
	Q	S	F	V	MEG	ST	C	G	Y	N	Y	NPL	G	M	NAP	G	A
	G					DD		E		C		TH				S	S
				S												T	D
SanAntonio Variants	Junction →← TS																
	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333	3333333333
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
	HMHDFKIYNN	LKIQHPEYQ	YLGIIYDIIM	NGNKQGDRGT	VGMSKFGY	MRFDLSQYFP	LLTTKKLFLR	GIIEELLWFI	RGETNGNTLL	NKNVRIWEAN	GTREFLDNRK	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
	I	H	H	K		R	L			K		1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
	L																
SanAntonio Variants	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444	4444444444
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
	LFHREVNDLG	PIYGFQWRHF	GAEYTNMYDN	YEDKGVQDLK	NVIHLIKNEP	TSRRILCAW	NVKDLQDMAL	PPCHILCQFY	VFDGKLSCIM	YQRSCDLGLG	VPFNIASYSI	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
	D			L	I	V				H		1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
SanAntonio Variants	TS →																
	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555	5555555555
	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
	FTHMIAQVCN	LQPAEFIHL	GNHVVYNNHI	DSLKVQLNRI	PYPFPTLKL	PEVKNIEDFT	ISDFTIENYV	HHDKIPMDMA	A			1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
		Q	M			Q	L	A				1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

Note: Repeats from residues 277-281.

Supplementary Figure S5. Neighbor-joining trees inferred from *Pkdhfr-ts* (A), and *Pidhfr-ts* sequences (B) of Thai isolates. The trees were constructed based on the maximum composite likelihood parameter, and 1000 bootstrap pseudoreplicates; only values > 60% are shown.

