

Additional file 6. Alignment of PoMSP1 sequences.

(A) Nucleotide substitutions and the 12-nucleotide repeat region of *PoMSP-1* (numbering based on the sequence of the *PoMSP-1* from the Thai isolate Po-7). *GenBank accession numbers: KX672045 (C0100511), KX672044 (M0500214); Cameroon (CM) isolates: FJ824670 and FJ824671; Thailand (TH) isolates: KC137340-KC137341, KC137343-KC137346 and KC137349.

A.

	421	541	1929	2814	*****	*****	2911
TH Po-7	TATCTA	CATCTT	CCTCAG	AA	GCCGCTACACAAGTCACTACGCAATCCGCTACACAAACCGCTACACATGCCTCTACACATGCCTCTACACAAGCCGCTACACAAGCCGCTACACAA	-----	GGT
C0100511	TATTTA	CATTTT	CCGCAG	AA	GCCGCTACACAAGTCACTACGCAATCCGCTACACAAACCGCTACACATGCCTCTACACATGCCTCTACACAAGCCGCTACACAAGCCGCTACACAA	GCCGCTACACAA	GGT
M0500214	TATCTA	CATCTT	CCTCAG	AA	GCCGCTACACAAGTCACTACGCAATCCGCTACACAAACCGCTACACATGCCTCTACACATGCCTCTACACAAGCCGCTACACAAGCCGCTACACAA	GCCGCTACACAA	GGT
TH Po-10	TATCTA	CATCTT	CCTCAG	AA	GCCGCTACACAAGTCACTACGCAATCCGCTACACAAACCGCTACACATGCCTCTACACATGCCTCTACACAAGCCGCTACACAAGCCGCTACACAA	-----	GGT
TH Po-4	TATCTA	CATCTT	CCTCAG	AA	GCCGCTACACAAGTCACTACGCAATCCGCTACACAAACCGCTACACATGCCTCTACACATGCCTCTACACAAGCCGCTACACAAGCCGCTACACAA	-----	GGT
CM OM1A	TATCTA	CATCTT	CCTCAG	AA	GCCGCTACACAAGTCACTACGCAATCCGCTACACAAACCGCTACACATGCCTCTACACATGCCTCTACACAAGCCGCTACACAAGCCGCTACACAA	-----	GGT
CM OM1B	TATCTA	CATCTT	CCTCAG	AA	GCCGCTACACAAGTCACTACGCAATCCGCTACACAAACCGCTACACATGCCTCTACACATGCCTCTACACAAGCCGCTACACAAGCCGCTACACAA	-----	GGT
TH Po-1	TACTTA	CATCTT	CCTCAG	---	-----TCACTACGCAAAGCAGT-----GTTCCACAGGCACTAC-----TTTCCAGTTCCTAC-----	-----	---
TH Po-2	TACTTA	CATCTT	CCTCAG	---	-----TCACTACGCAAAGCAGT-----GTTCCACAGGCACTAC-----TTTCCAGTTCCTAC-----	-----	---
TH Po-5	TACTTA	CATCTT	CCTCAG	---	-----TCACTACGCAAAGCAGT-----GTTCCACAGGCACTAC-----TTTCCAGTTCCTAC-----	-----	---
TH Po-6	TACTTA	CATCTT	CCTCAG	---	-----TCACTACGCAAAGCAGT-----GTTCCACAGGCACTAC-----TTTCCAGTTCCTAC-----	-----	---

(B) Amino acid polymorphisms in the variable domains of PoMSP-1. Protein sequence numbering was based on the sequence of the PoMSP-1 from the Thai isolate Po-7. *GenBank accession numbers: KX672045 (C0100511), KX672044 (M0500214); Cameroon (CM) isolates: FJ824670 and FJ824671; Thailand (TH) isolates: KC137340-KC137341, KC137343-KC137346 and KC137349.

B.

	Variable domain 1	Variable domain 2	Variable domain 3
	53	309	282
	1681	1691	
TH Po-7	LQKLEVTFTTGRG-IGSSTTSVSSTPNGASTHLPTDRDSSNI	YGEVAVSNGTGSAGAASGTNASASSGOENSSSESETEKYNKAKAL	EIAEKESKV-TPPSNTASOTQLOEEIN
C0100511	LQKLEVTFTTGRG-IGSSTTSVSSTPNGASTHLPTDRDSSNI	YGEVAVSNGTGSAGAASGTNASASSGOENSSSESETEKYNKAKAL	EIAEKESKV-TPPSNTASOTQLOEEIN
M0500214	LQKLEVTFTTGRG-IGSSTTSVSSTPNGASTHLPTDRDSSNI	YGEVAVSNGTGSAGAASGTNASASSGOENSSSESETEKYNKAKAL	EIAEKESKV-TPPSNTASOTQLOEEIN
TH Po-10	LQKLEVTFTTGRG-IGSSTTSVSSTPNGASTHLPTDRDSSNI	YGEVAVSNGTGSAGAASGTNASASSGOENSSSESETEKYNKAKAL	EIAEKESKV-TPPSNTASOTQLOEEIN
TH Po-4	LQKLEVTFTTGRG-IGSSTTSVSSTPNGASTHLPTDRDSSNI	YGEVAVSNGTGSAGAASGTNASASSGOENSSSESETEKYNKAKAL	EIAEKESKV-TPPSNTASOTQLOEEIN
CM OM1A	LQKLEVTFTTGRG-IGSSTTSVSSTPNGASTHLPTDRDSSNI	YGEVAVSNGTGSAGAASGTNASASSGOENSSSESETEKYNKAKAL	EIAEKESKV-TPPSNTASOTQLOEEIN
CM OM1B	LQKLEVTFTTGRG-IGSSTTSVSSTPNGASTHLPTDRDSSNI	YGEVAVSNGTGSAGAASGTNASASSGOENSSSESETEKYNKAKAL	EIAEKESKV-TPPSNTASOTQLOEEIN
TH Po-1	VDOIKTAANTAAAGCNVRSSTSSVSTTPSSAAAAAASNDADNNV	HSEDAVSGSTGNARAASGANVHANSGBSSSDIEKREKYDKDKKV	QTAEKESKVI-TPPNNTTLOTQLOEEIS
TH Po-2	VDOIKTAANTAAAGCNVRSSTSSVSTTPSSAAAAAASNDADNNV	HSEDAVSGSTGNARAASGANVHANSGBSSSDIEKREKYDKDKKV	QTAEKESKVI-TPPNNTTLOTQLOEEIS
TH Po-5	VDOIKTAANTAAAGCNVRSSTSSVSTTPSSAAAAAASNDADNNV	HSEDAVSGSTGNARAASGANVHANSGBSSSDIEKREKYDKDKKV	QTAEKESKVI-TPPNNTTLOTQLOEEIS
TH Po-6	VDOIKTAANTAAAGCNVRSSTSSVSTTPSSAAAAAASNDADNNV	HSEDAVSGSTGNARAASGANVHANSGBSSSDIEKREKYDKDKKV	QTAEKESKVI-TPPNNTTLOTQLOEEIS
		Variable domain 4	
	679		797
TH Po-7	DKEKTNPTPATAQAGALPVRGVDEILVMGNENEATAVTSPPSPSTSTSTEASEGATQSATTVOSETSVVQIG	TPVAHPGASAPPTPGVPVAPAEAPAPAPAT	EAQAQAPVQPTQGOVQA
C0100511	DKEKTNPTPATAQAGALPVRGVDEILVMGNENEATAVTSPPSPSTSTSTEASEGATQSATTVOSETSVVQIG	TPVAHPGASAPPTPGVPVAPAEAPAPAPAT	EAQAQAPVQPTQGOVQA
M0500214	DKEKTNPTPATAQAGALPVRGVDEILVMGNENEATAVTSPPSPSTSTSTEASEGATQSATTVOSETSVVQIG	TPVAHPGASAPPTPGVPVAPAEAPAPAPAT	EAQAQAPVQPTQGOVQA
TH Po-10	DKEKTNPTPATAQAGALPVRGVDEILVMGNENEATAVTSPPSPSTSTSTEASEGATQSATTVOSETSVVQIG	TPVAHPGASAPPTPGVPVAPAEAPAPAPAT	EAQAQAPVQPTQGOVQA
TH Po-4	DKEKTNPTPATAQAGALPVRGVDEILVMGNENEATAVTSPPSPSTSTSTEASEGATQSATTVOSETSVVQIG	TPVAHPGASAPPTPGVPVAPAEAPAPAPAT	EAQAQAPVQPTQGOVQA
CM OM1A	DKEKTNPTPATAQAGALPVRGVDEILVMGNENEATAVTSPPSPSTSTSTEASEGATQSATTVOSETSVVQIG	TPVAHPGASAPPTPGVPVAPAEAPAPAPAT	EAQAQAPVQPTQGOVQA
CM OM1B	DKEKTNPTPATAQAGALPVRGVDEILVMGNENEATAVTSPPSPSTSTSTEASEGATQSATTVOSETSVVQIG	TPVAHPGASAPPTPGVPVAPAEAPAPAPAT	EAQAQAPVQPTQGOVQA
TH Po-1	EKEKATPTQAAAQGVQPVQGVDEILVMGNENEITTEV----	PSTSTVPAAVVAQFTEATTSAPGAGGLPGAAGFPGV	GAAGEPGVPGVPGAQPPSSASAP
TH Po-2	EKEKATPTQAAAQGVQPVQGVDEILVMGNENEITTEV----	PSTSTVPAAVVAQFTEATTSAPGAGGLPGAAGFPGV	GAAGEPGVPGVPGAQPPSSASAP
TH Po-5	EKEKATPTQAAAQGVQPVQGVDEILVMGNENEITTEV----	PSTSTVPAAVVAQFTEATTSAPGAGGLPGAAGFPGV	GAAGEPGVPGVPGAQPPSSASAP
TH Po-6	EKEKATPTQAAAQGVQPVQGVDEILVMGNENEITTEV----	PSTSTVPAAVVAQFTEATTSAPGAGGLPGAAGFPGV	GAAGEPGVPGVPGAQPPSSASAP
		Variable domain 5	
	908		1023
TH Po-7	DKAIKAFLETQGITGAPDATPLVTTQATTOAATQVTTQASATQATATHASTHASTQAAATQAAATC	GNVPOASNAEHTPSATTVNPAATTPDKSLEESTSEGLMTOGNADDDVSEPEE	
C0100511	DKAIKAFLETQGITGAPDATPLVTTQATTOAATQVTTQASATQATATHASTHASTQAAATQAAATC	GNVPOASNAEHTPSATTVNPAATTPDKSLEESTSEGLMTOGNADDDVSEPEE	
M0500214	DKAIKAFLETQGITGAPDATPLVTTQATTOAATQVTTQASATQATATHASTHASTQAAATQAAATC	GNVPOASNAEHTPSATTVNPAATTPDKSLEESTSEGLMTOGNADDDVSEPEE	
TH Po-10	DKAIKAFLETQGITGAPDATPLVTTQATTOAATQVTTQASATQATATHASTHASTQAAATQAAATC	GNVPOASNAEHTPSATTVNPAATTPDKSLEESTSEGLMTOGNADDDVSEPEE	
TH Po-4	DKAIKAFLETQGITGAPDATPLVTTQATTOAATQVTTQASATQATATHASTHASTQAAATQAAATC	GNVPOASNAEHTPSATTVNPAATTPDKSLEESTSEGLMTOGNADDDVSEPEE	
CM OM1A	DKAIKAFLETQGITGAPDATPLVTTQATTOAATQVTTQASATQATATHASTHASTQAAATQAAATC	GNVPOASNAEHTPSATTVNPAATTPDKSLEESTSEGLMTOGNADDDVSEPEE	
CM OM1B	DKAIKAFLETQGITGAPDATPLVTTQATTOAATQVTTQASATQATATHASTHASTQAAATQAAATC	GNVPOASNAEHTPSATTVNPAATTPDKSLEESTSEGLMTOGNADDDVSEPEE	
TH Po-1	EVAVKAFLETHGITGTAP-----VTTQSSGPPATTSPP-----VTPPAATHHT-----TTTNPATPAPSK-----FEETLVAQGNADDDAPAPEE		
TH Po-2	EVAVKAFLETHGITGTAP-----VTTQSSGPPATTSPP-----VTPPAATHHT-----TTTNPATPAPSK-----FEETLVAQGNADDDAPAPEE		
TH Po-5	EVAVKAFLETHGITGTAP-----VTTQSSGPPATTSPP-----VTPPAATHHT-----TTTNPATPAPSK-----FEETLVAQGNADDDAPAPEE		
TH Po-6	EVAVKAFLETHGITGTAP-----VTTQSSGPPATTSPP-----VTPPAATHHT-----TTTNPATPAPSK-----FEETLVAQGNADDDAPAPEE		
		Variable domain 6	Variable domain 7
	1275	1344	1494 1504
TH Po-7	STGAATQTOGGEGASAAAPP-----PAAPPAASAAAPGTANGETATVAHAEDYTEDDNNVIVLPLFGKKGTHAFD		KTGDDATKATN
C0100511	STGAATQTOGGEGASAAAPP-----PAAPPAASAAAPGTANGETATVAHAEDYTEDDNNVIVLPLFGKKGTHAFD		KTGDDATKATN
M0500214	STGAATQTOGGEGASAAAPP-----PAAPPAASAAAPGTANGETATVAHAEDYTEDDNNVIVLPLFGKKGTHAFD		KTGDDATKATN
TH Po-10	STGAATQTOGGEGASAAAPP-----PAAPPAASAAAPGTANGETATVAHAEDYTEDDNNVIVLPLFGKKGTHAFD		KTGDDATKATN
TH Po-4	STGAATQTOGGEGASAAAPP-----PAAPPAASAAAPGTANGETATVAHAEDYTEDDNNVIVLPLFGKKGTHAFD		KTGDDATKATN
CM OM1A	STGAATQTOGGEGASAAAPP-----PAAPPAASAAAPGTANGETATVAHAEDYTEDDNNVIVLPLFGKKGTHAFD		KTGDDATKATN
CM OM1B	STGAATQTOGGEGASAAAPP-----PAAPPAASAAAPGTANGETATVAHAEDYTEDDNNVIVLPLFGKKGTHAFD		KTGDDATKATN
TH Po-1	PTCPTAATGAVGGGAPAEAEA-----VAAPPAAPAGTGTANGETATVAHAEDYTEDDNNVIVLPLFGKKGTHAFD		KTGDDATKATN
TH Po-2	PTCPTAATGAVGGGAPAEAEA-----VAAPPAAPAGTGTANGETATVAHAEDYTEDDNNVIVLPLFGKKGTHAFD		KTGDDATKATN
TH Po-5	PTCPTAATGAVGGGAPAEAEA-----VAAPPAAPAGTGTANGETATVAHAEDYTEDDNNVIVLPLFGKKGTHAFD		KTGDDATKATN
TH Po-6	PTCPTAATGAVGGGAPAEAEA-----VAAPPAAPAGTGTANGETATVAHAEDYTEDDNNVIVLPLFGKKGTHAFD		KTGDDATKATN