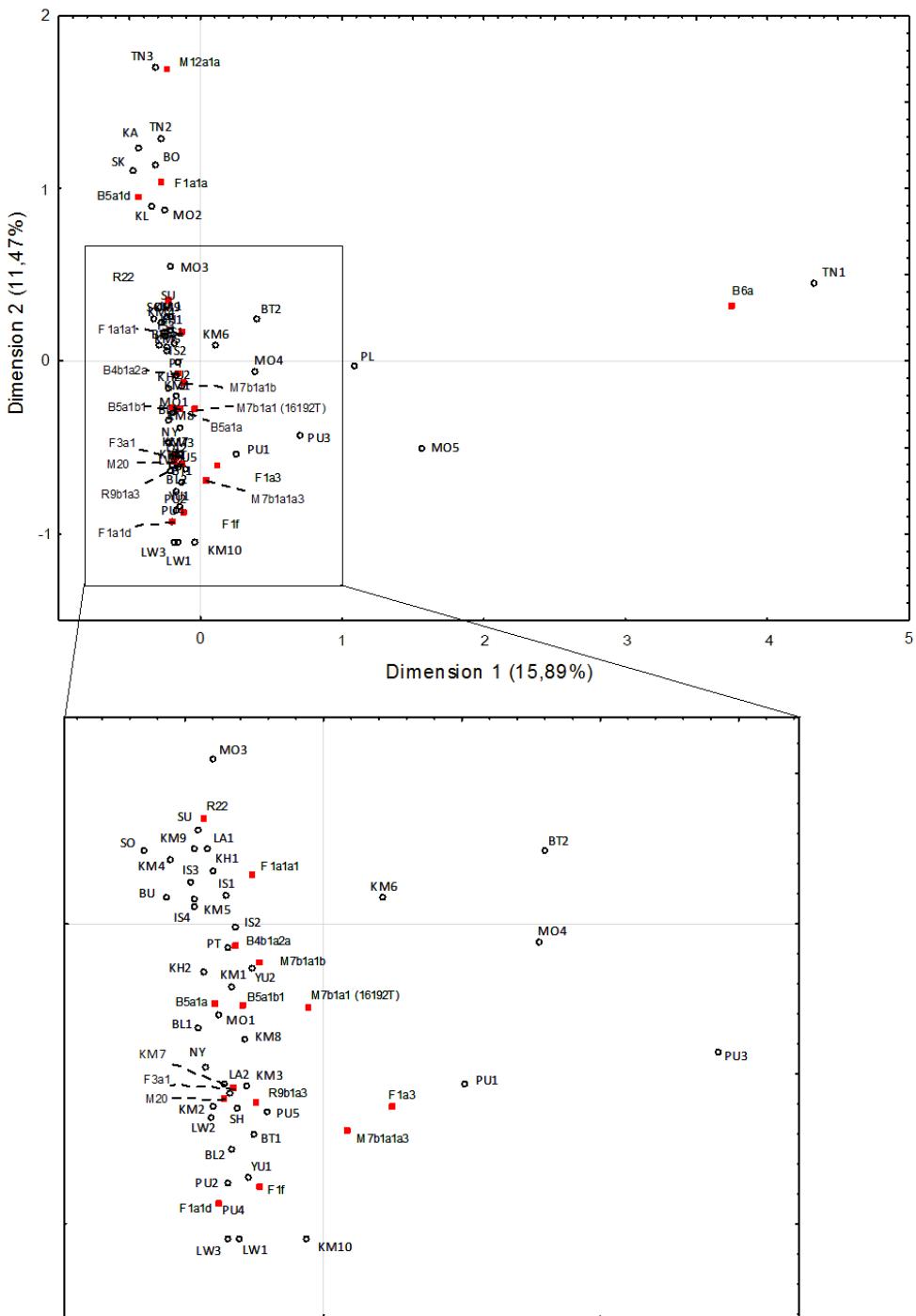
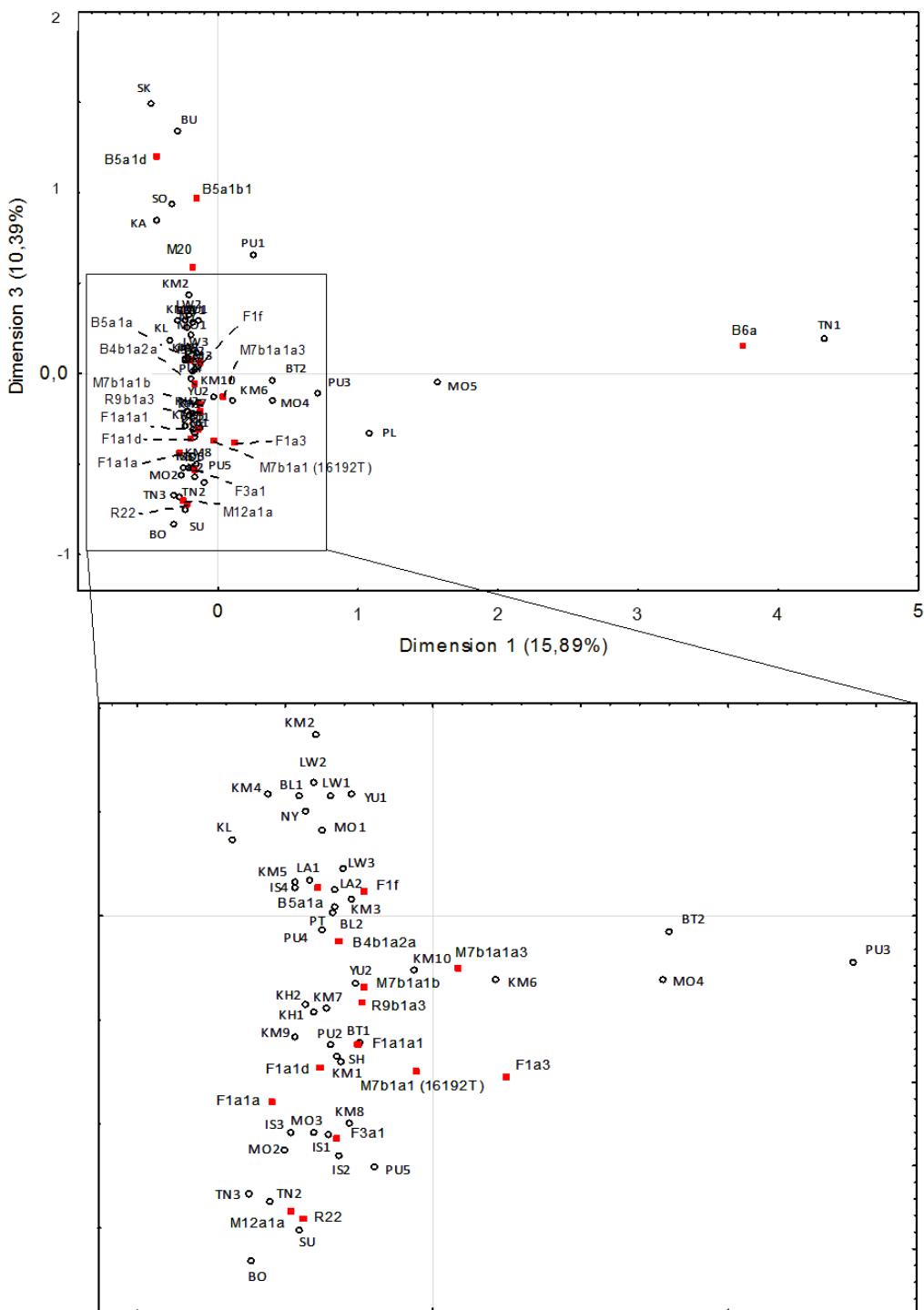
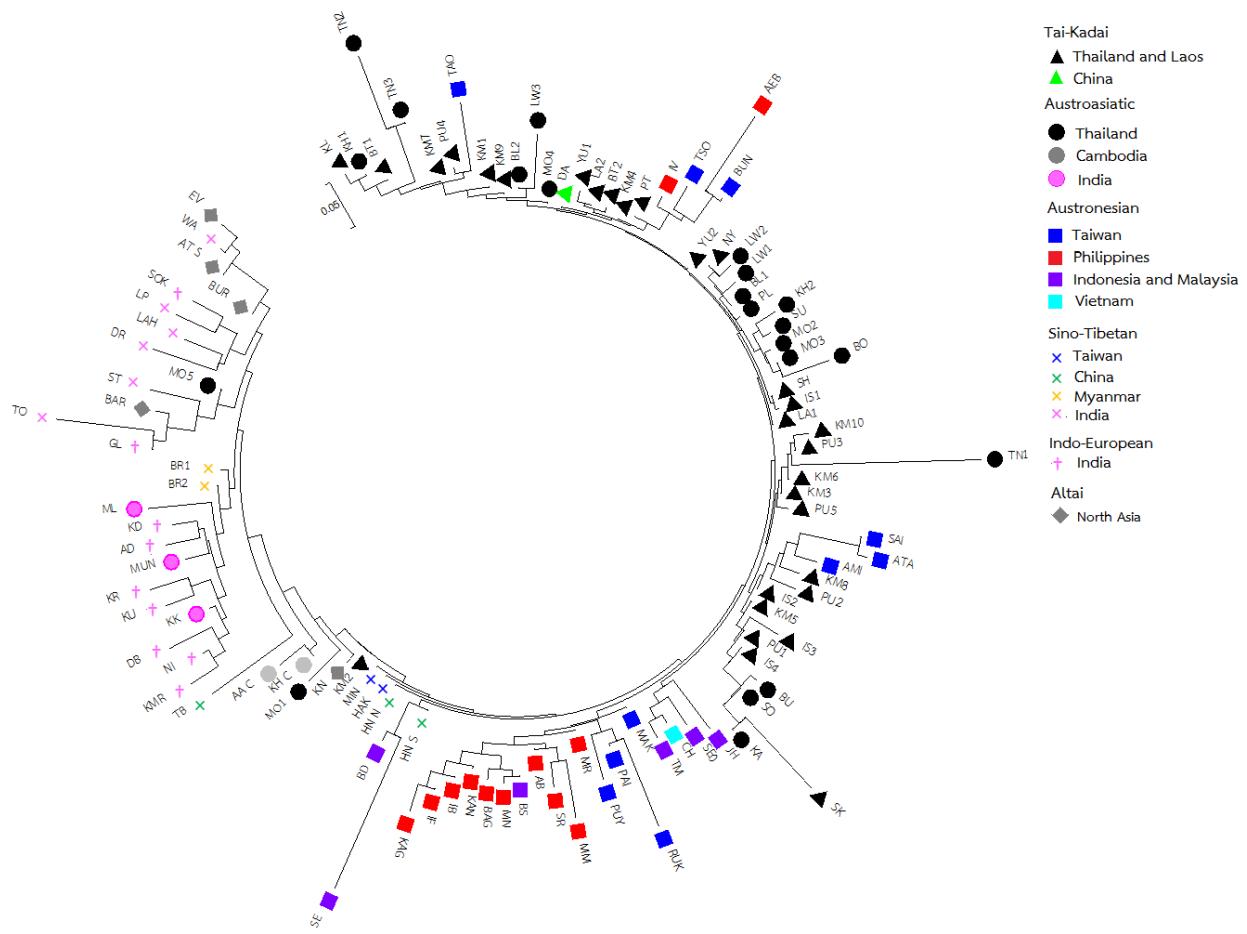


**Online Resource 3** includes Figures S1-S7 with respective legends. Figure S1: Correspondence analysis (CA) based on haplogroup counts. Figure S2: Neighbor Joining tree based on the  $\Phi_{st}$  genetic distance matrix for the populations in the present study and comparative populations. Figure S3: Bayesian maximum clade credibility (MCC) trees for 17 haplogroups (F1a, F1f, B5a, B4, B6a, R22, R9b, W3a1b, A, N9a, M7b1a, D, CZ, M20, M12-G, M24, and M74). Figure S4: Networks of 14 major haplogroups (F1a, F1f, B5a, B4, M7, D, M8, M12-G, M20, M24, M74, R9b, R22 and N9a) based on Thai/Lao mtDNA genomes and three haplogroups (B4c2, B6a and R22) based on sequences in Thai/Lao populations and other East Asian populations. Figure S5: The contour maps of the frequencies of 14 haplogroups (F1, F1f, B5, B5a1d, B4, B6a, D, M7, M12, M20, M24, M74, R9b and R22) in Thai/Lao populations. Figure S6: Bayesian skyline plots show the effective population size change over time in each studied population. Figure S7: Principal Component Analysis (PCA) plots reveal the fit between the simulated and observed data generated by the best 1,000 (or 5,000) simulations for each model.

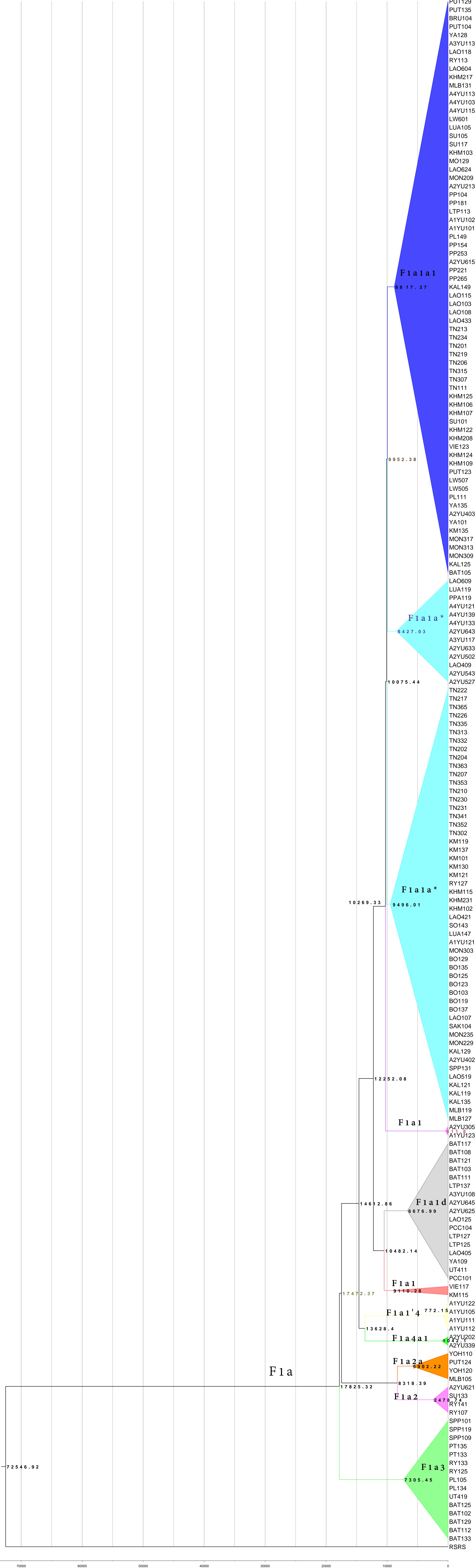





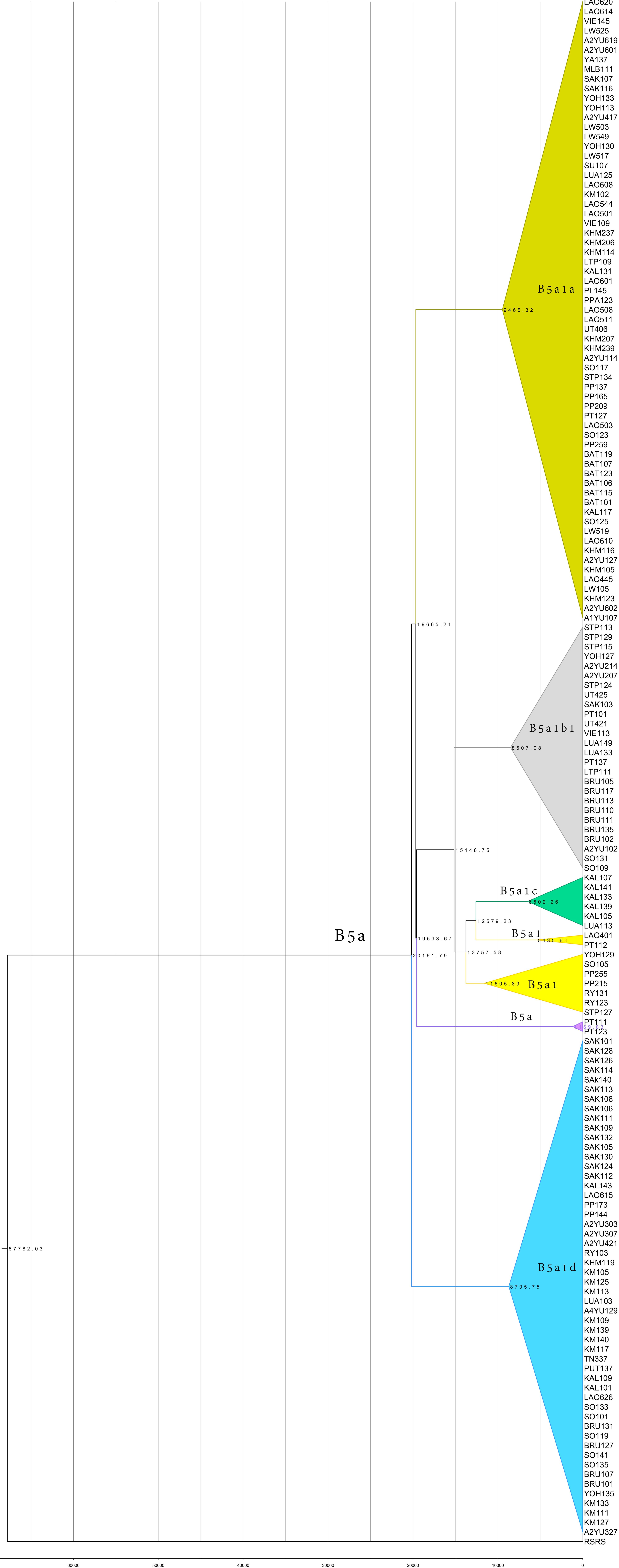
**Fig. S1** Correspondence analysis (CA) based on haplogroup counts.




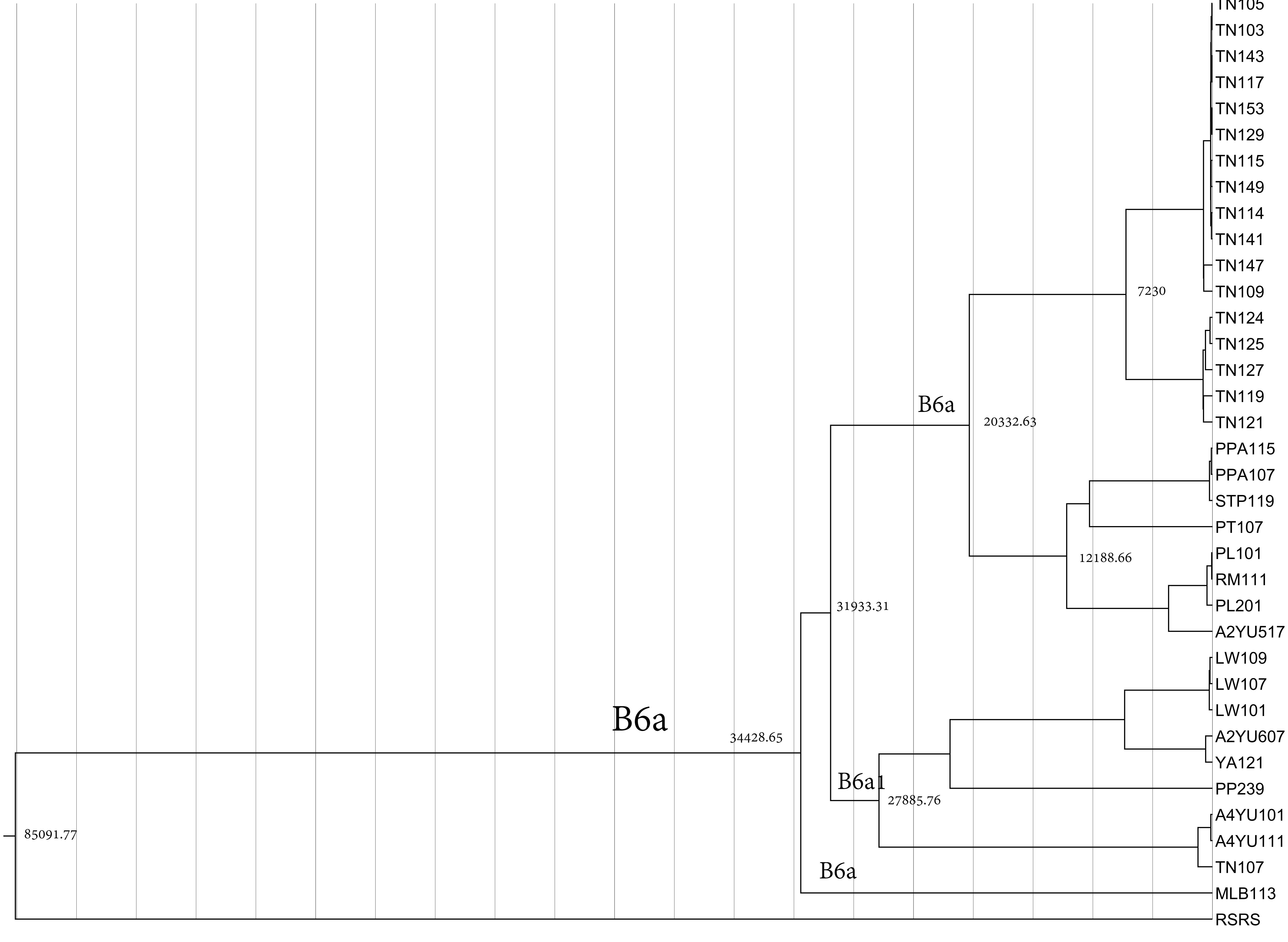
**Fig. S2** Neighbor Joining tree based on the  $\Phi_{st}$  genetic distance matrix for the populations in the present study and comparative populations. Population codes are provided in Fig. 1 and Table S2.

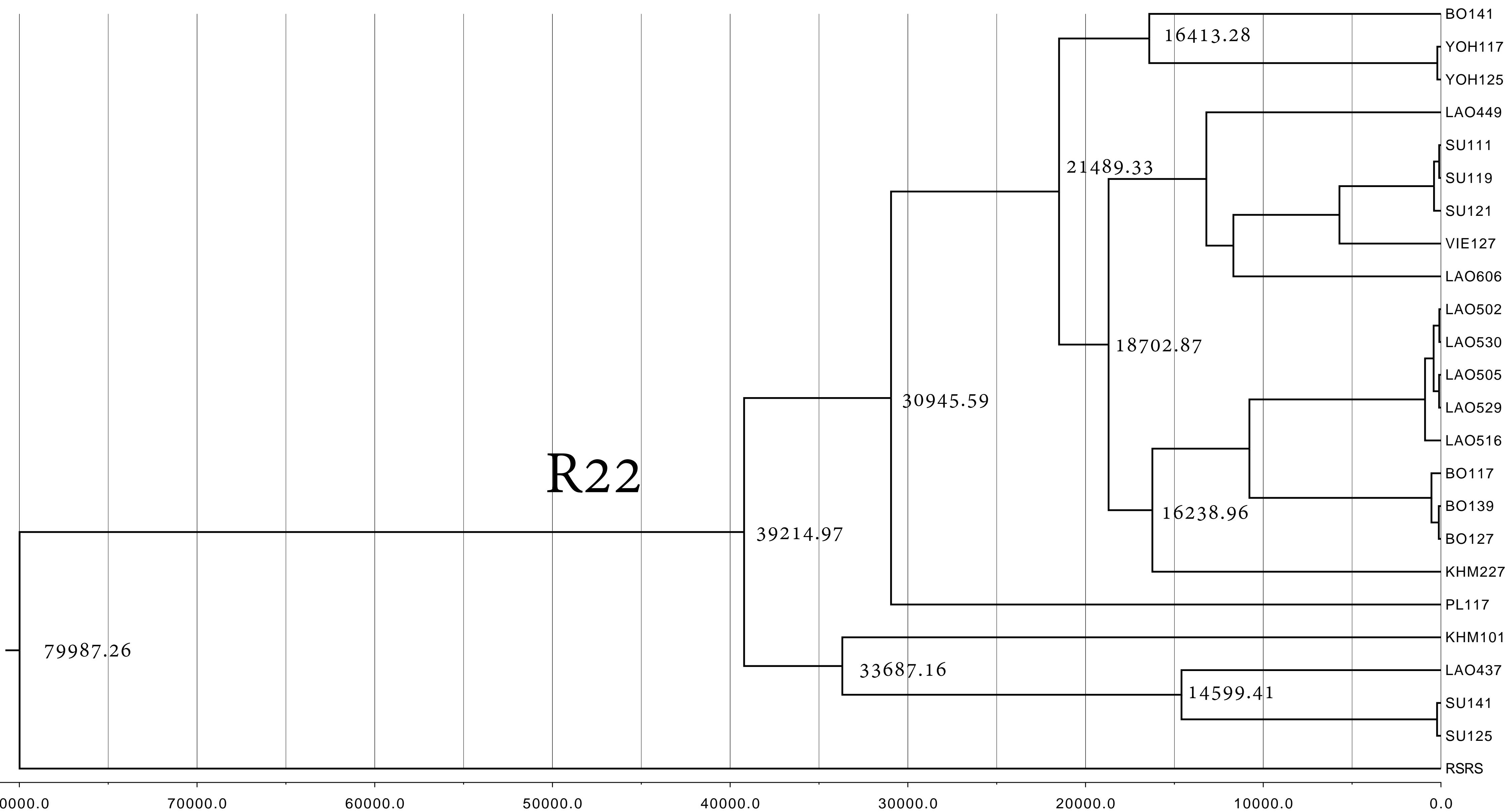


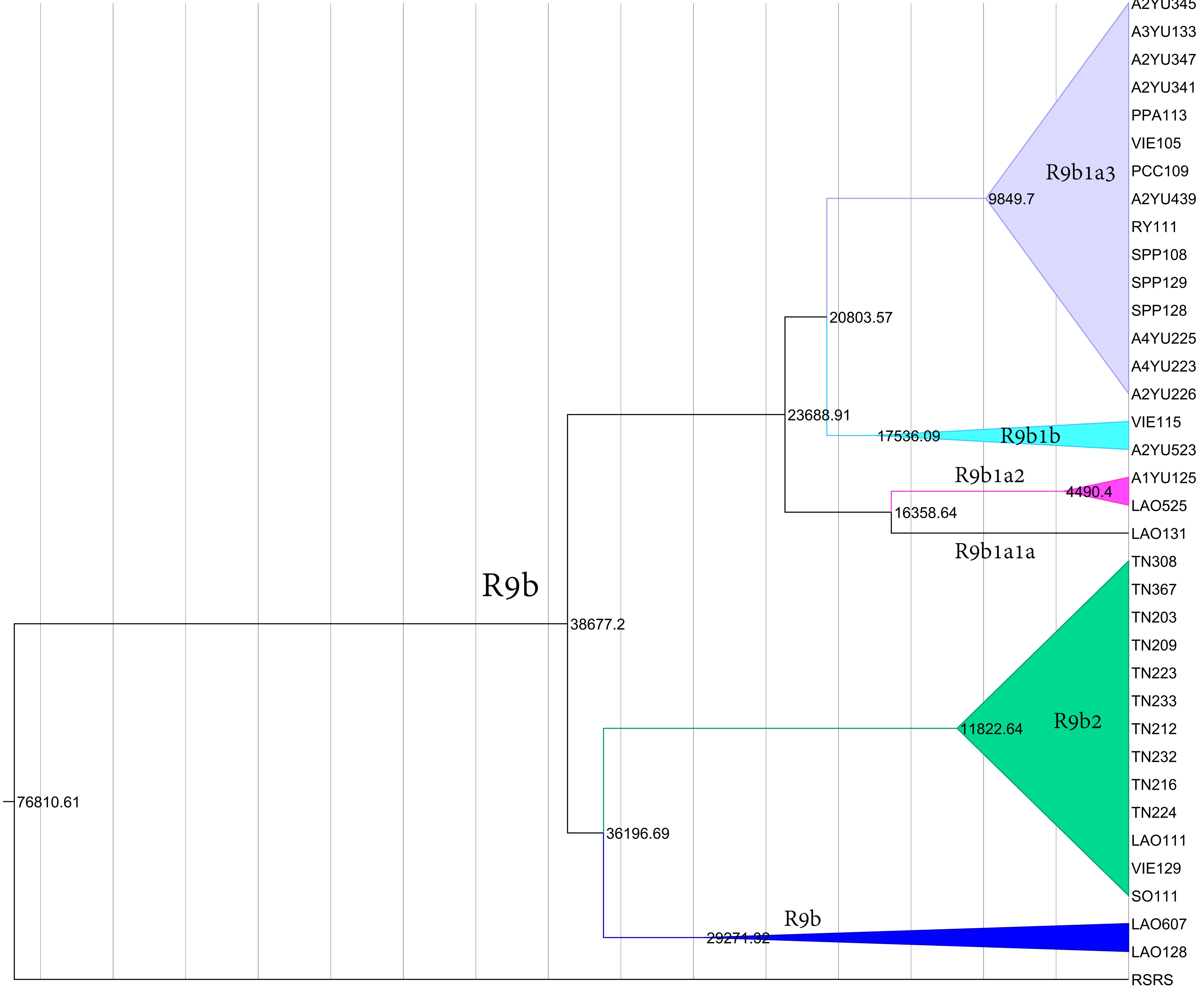












MON219

W3a1b

13418.78

MO111

MO123

MO105

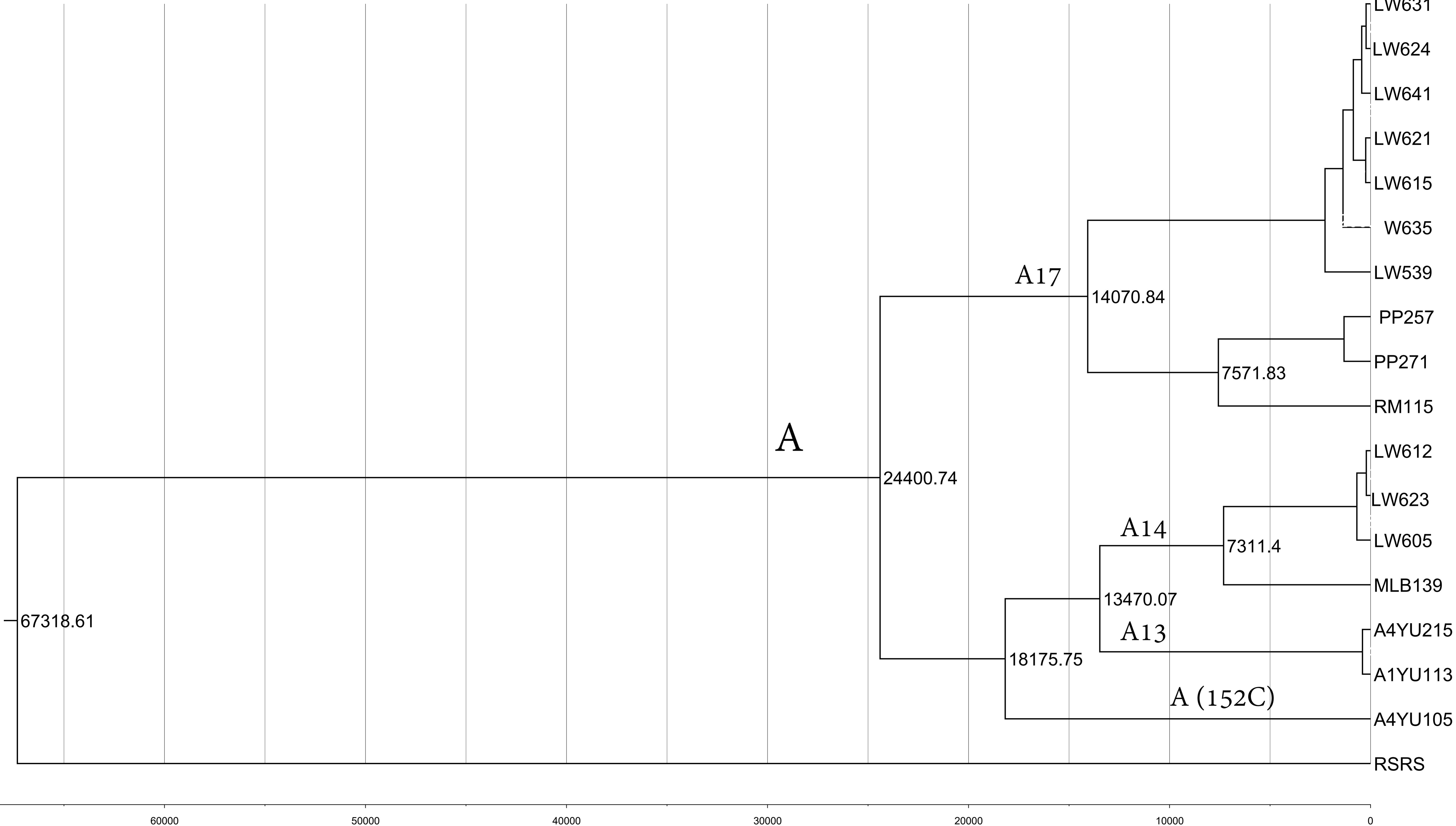
MO121

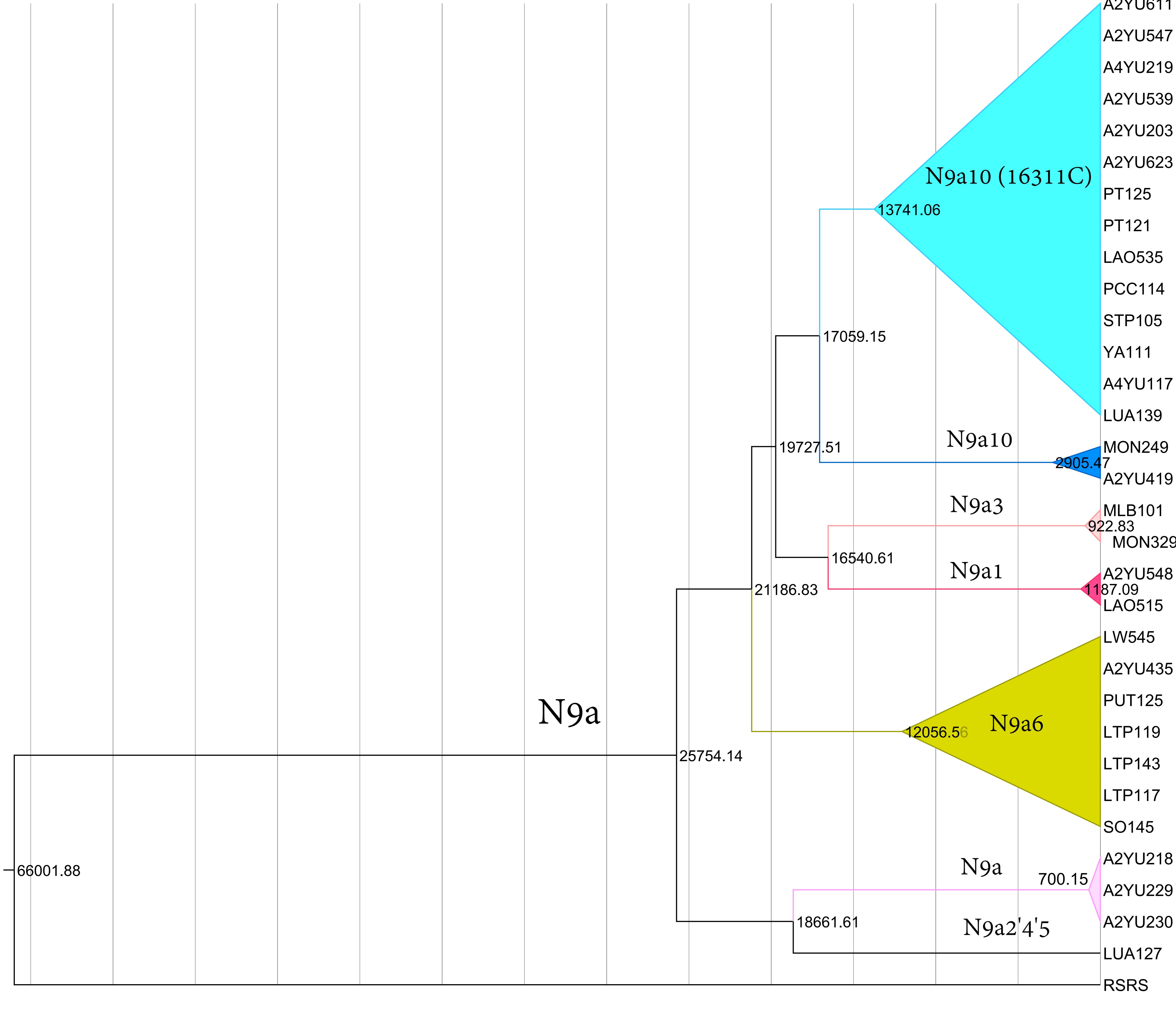
MO115

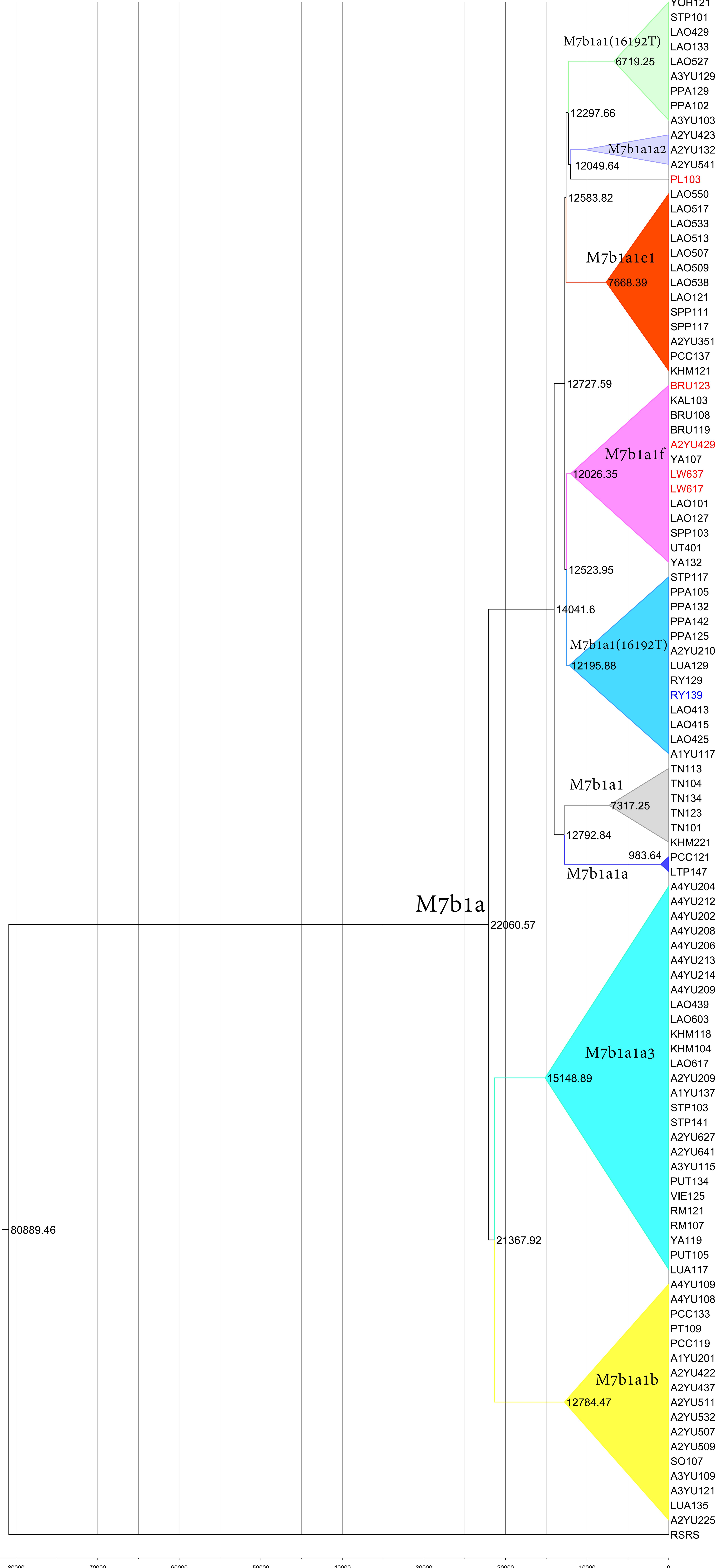
MO113


RSRS

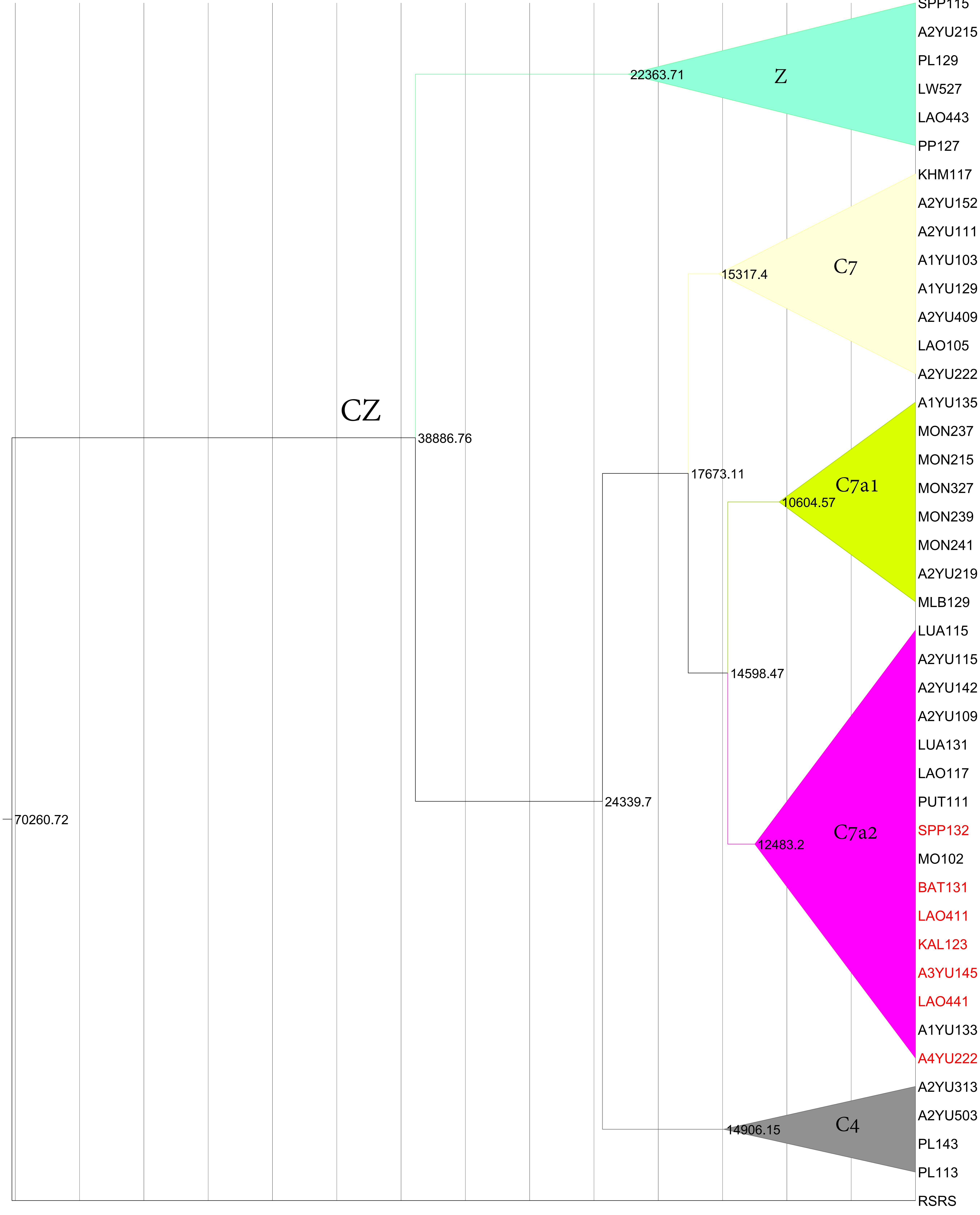
101329.99

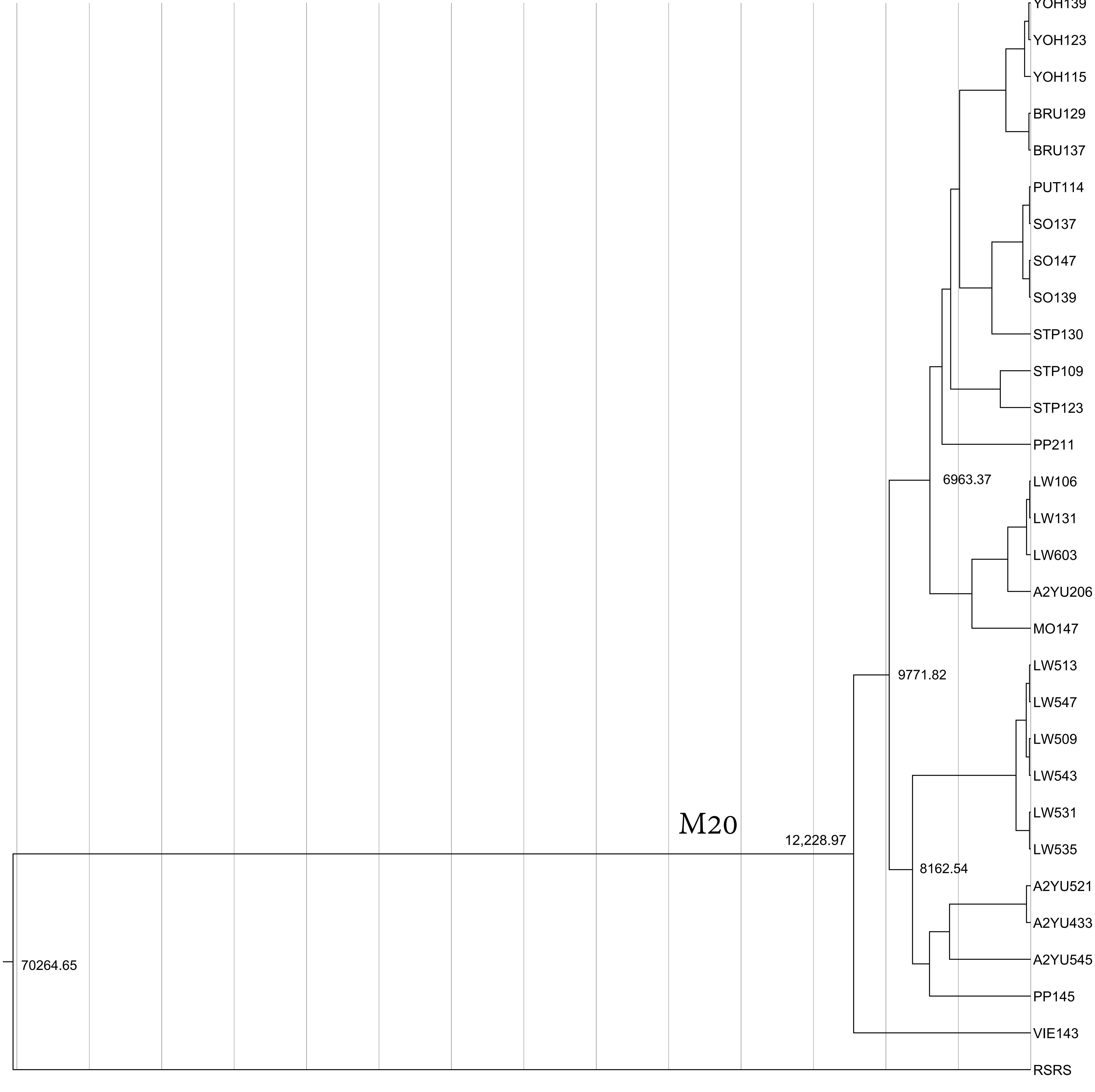


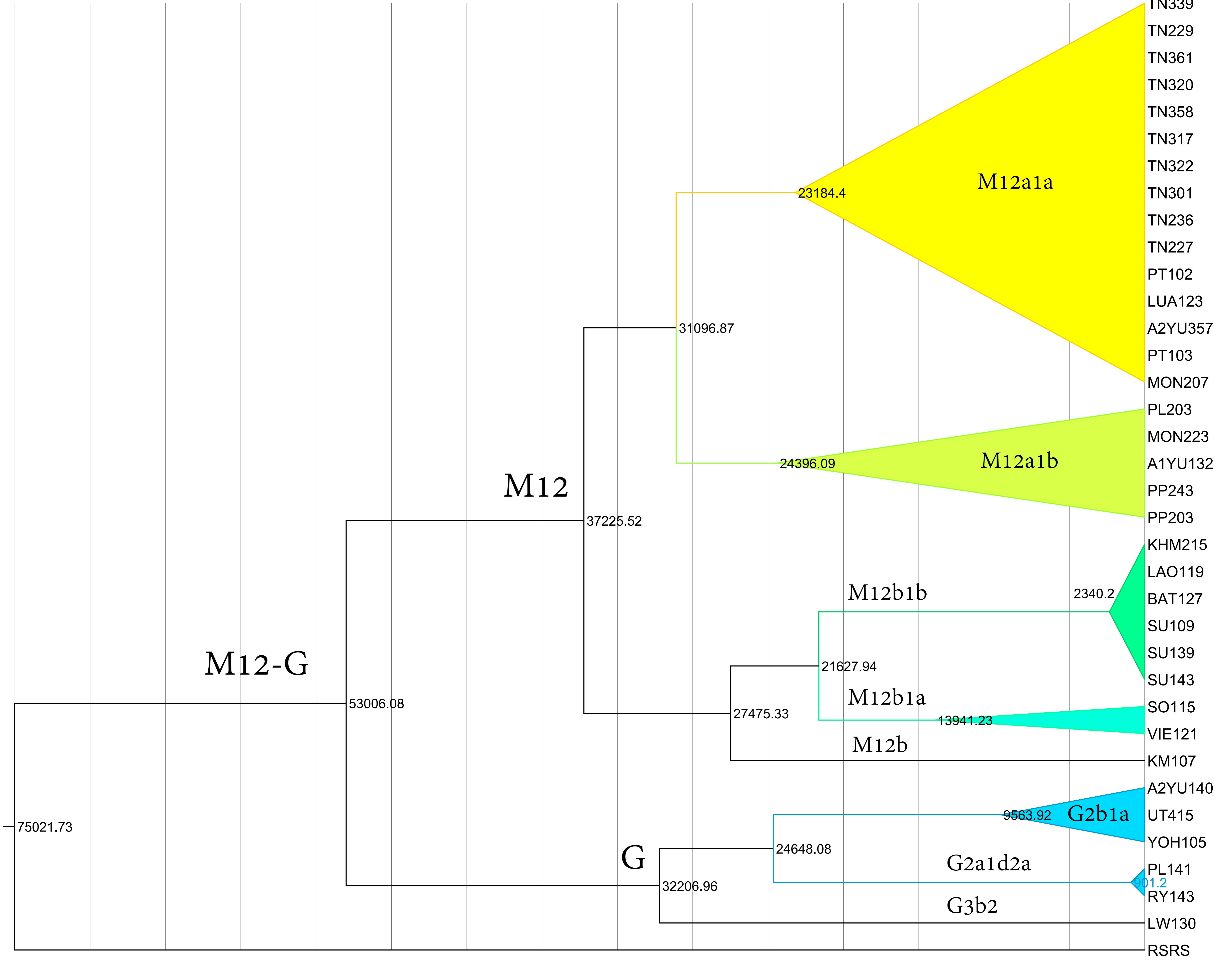


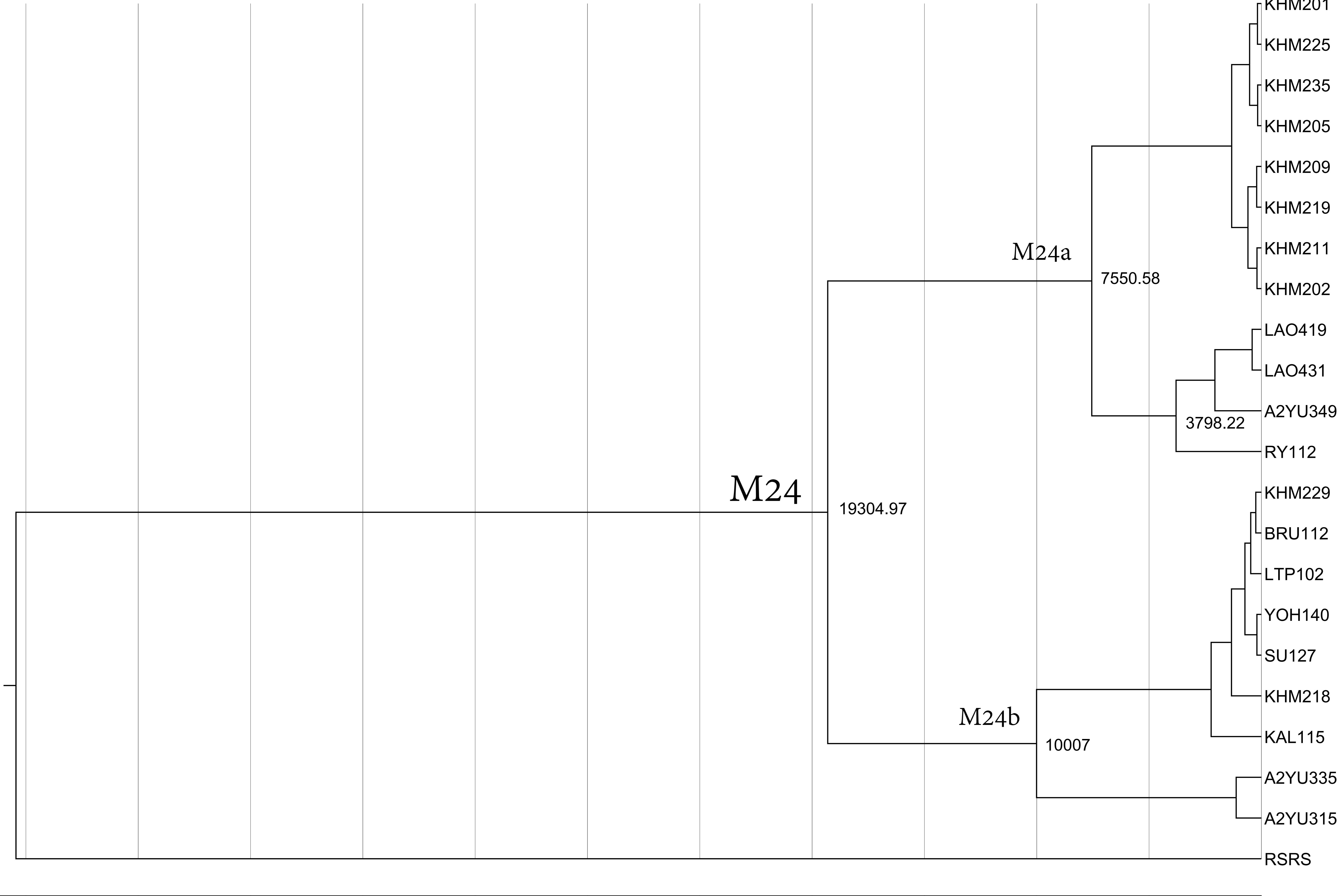


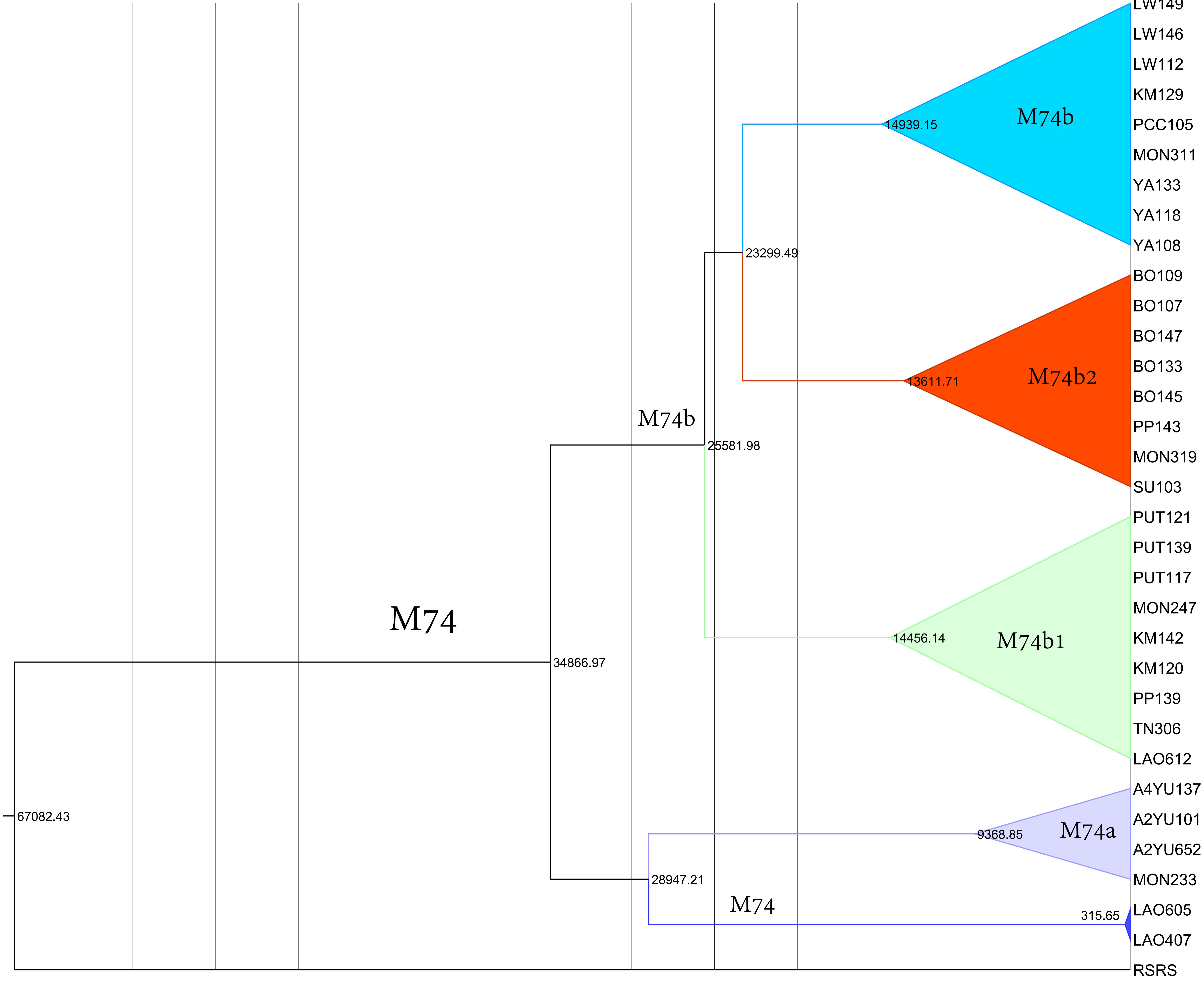






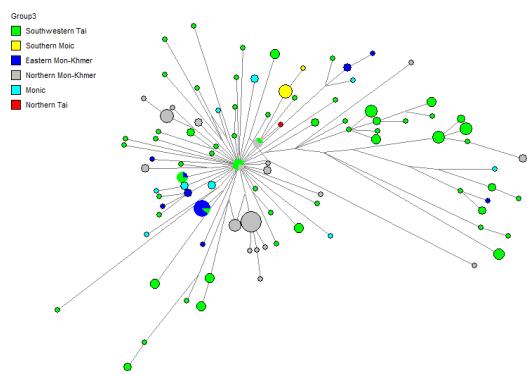
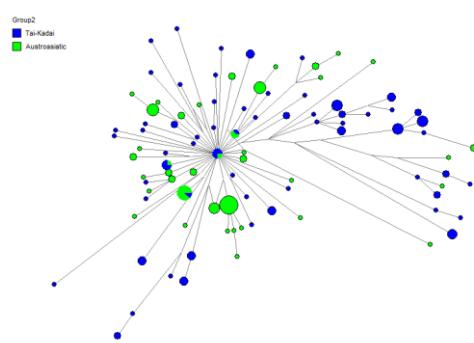
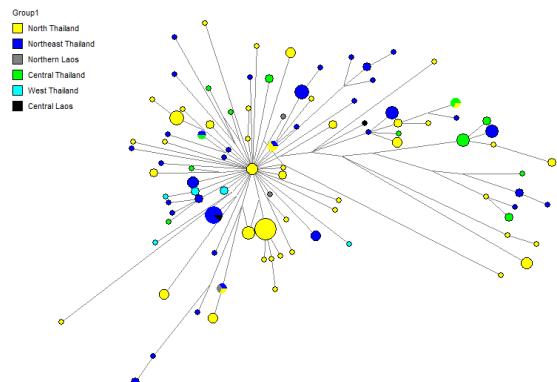
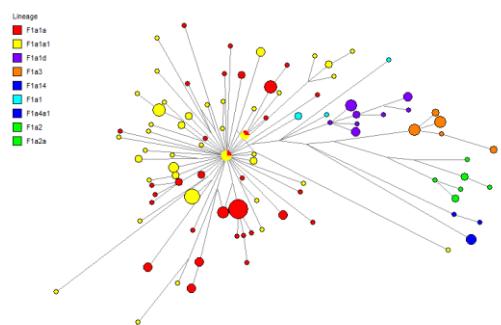




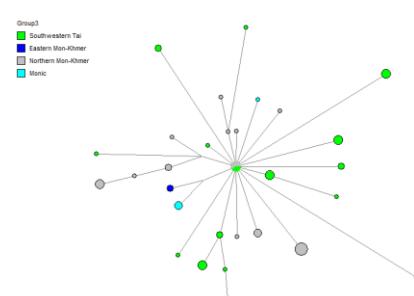
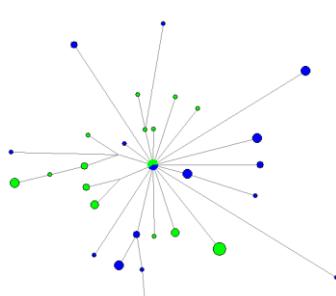
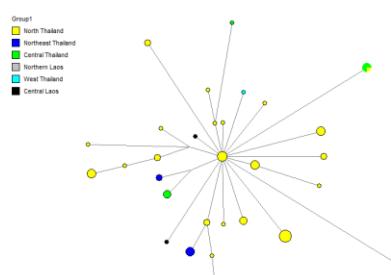


**Fig. S3** Bayesian maximum clade credibility (MCC) trees. MCC trees for 17 haplogroups (F1a, F1f, B5a, B4, B6a, R22, R9b, W3a1b, A, N9a, M7b1a, D, CZ, M20, M12-G, M24, and M74). Tree topology and coalescent time estimates were generated by BEAST and TreeAnnotator with parameters described in the Material and Methods. To aid visualization, some trees were schematically modified to combine individuals in the same lineage or sublineage by Figtree. Sample codes and RSRS are indicated at the tips of the tree.

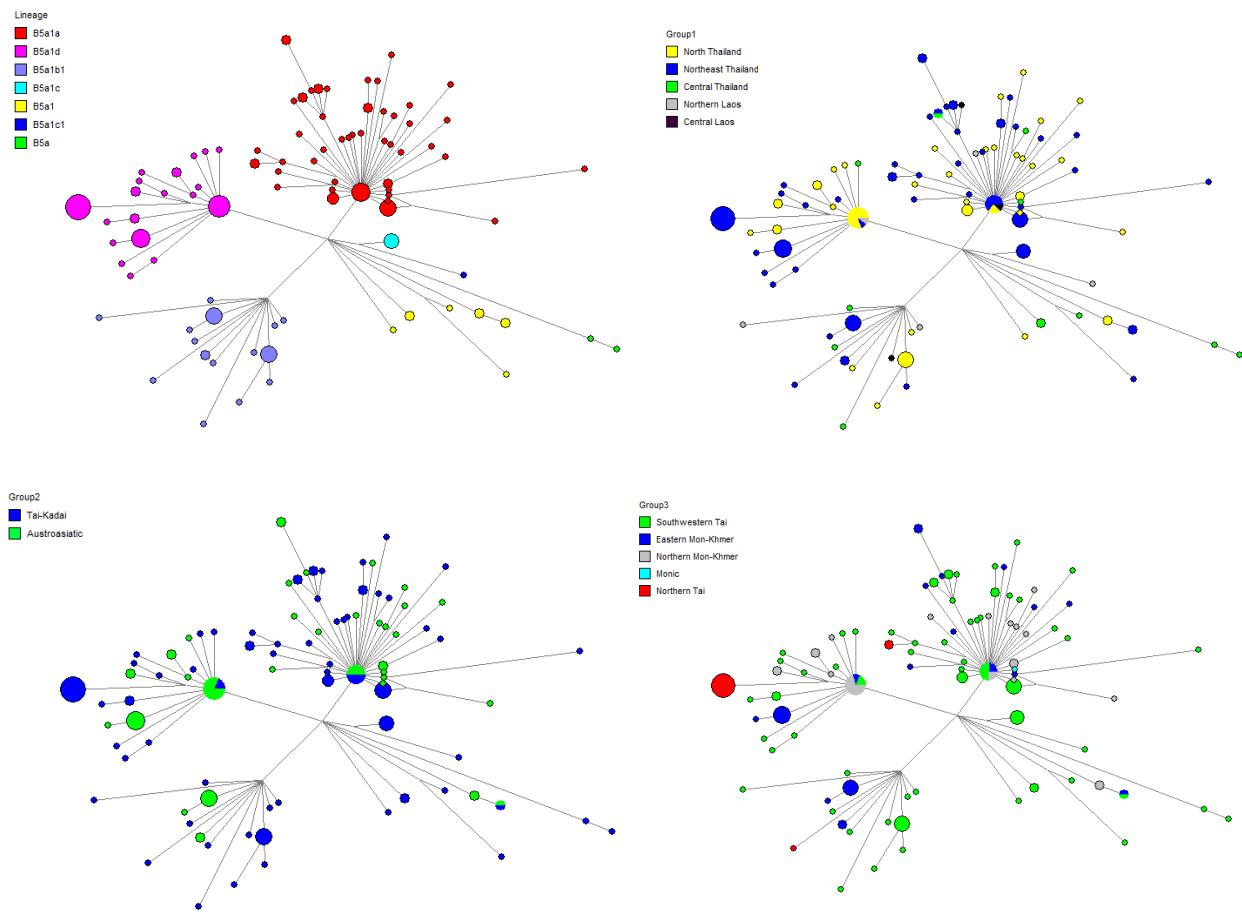
F1a



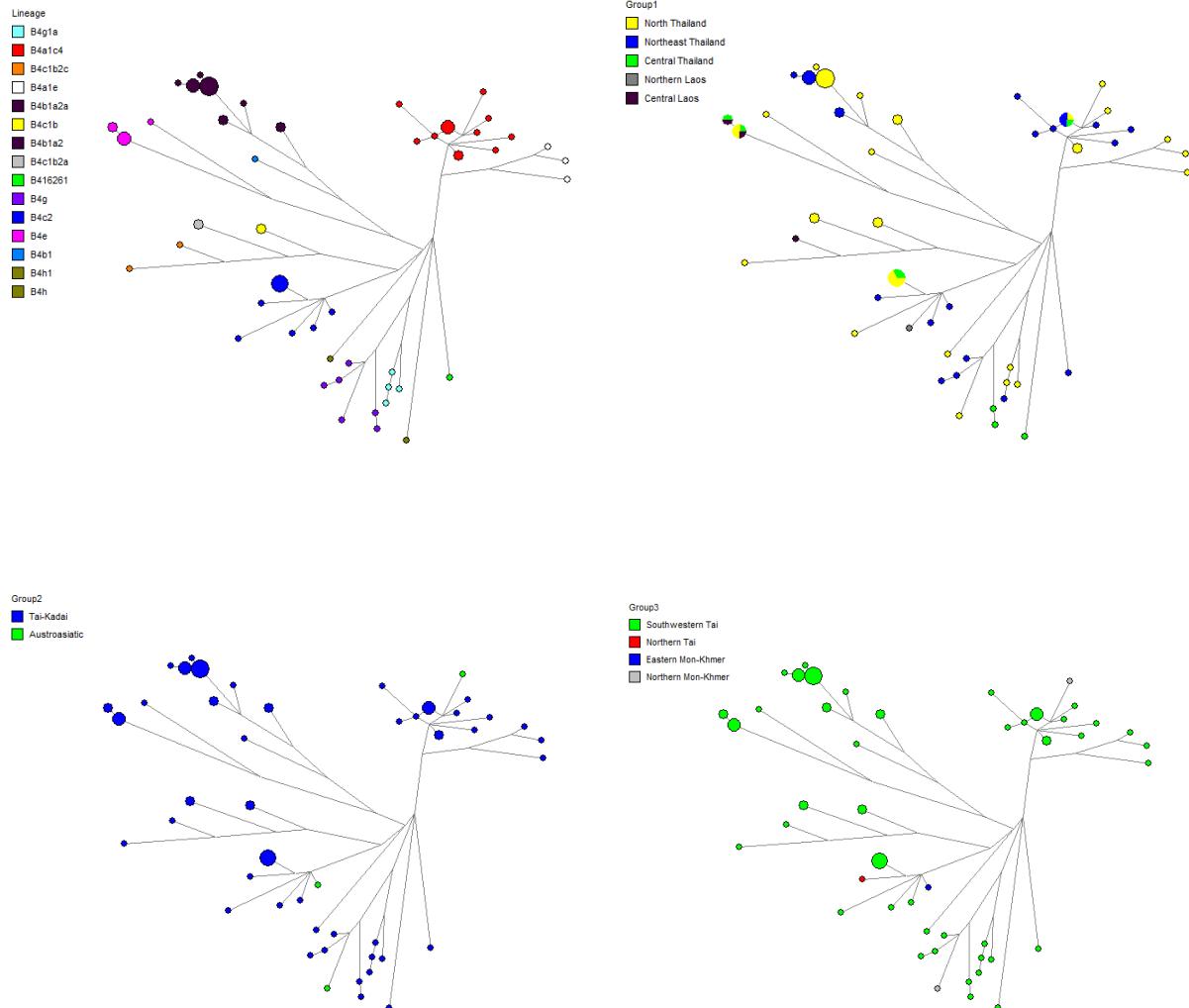
F1f



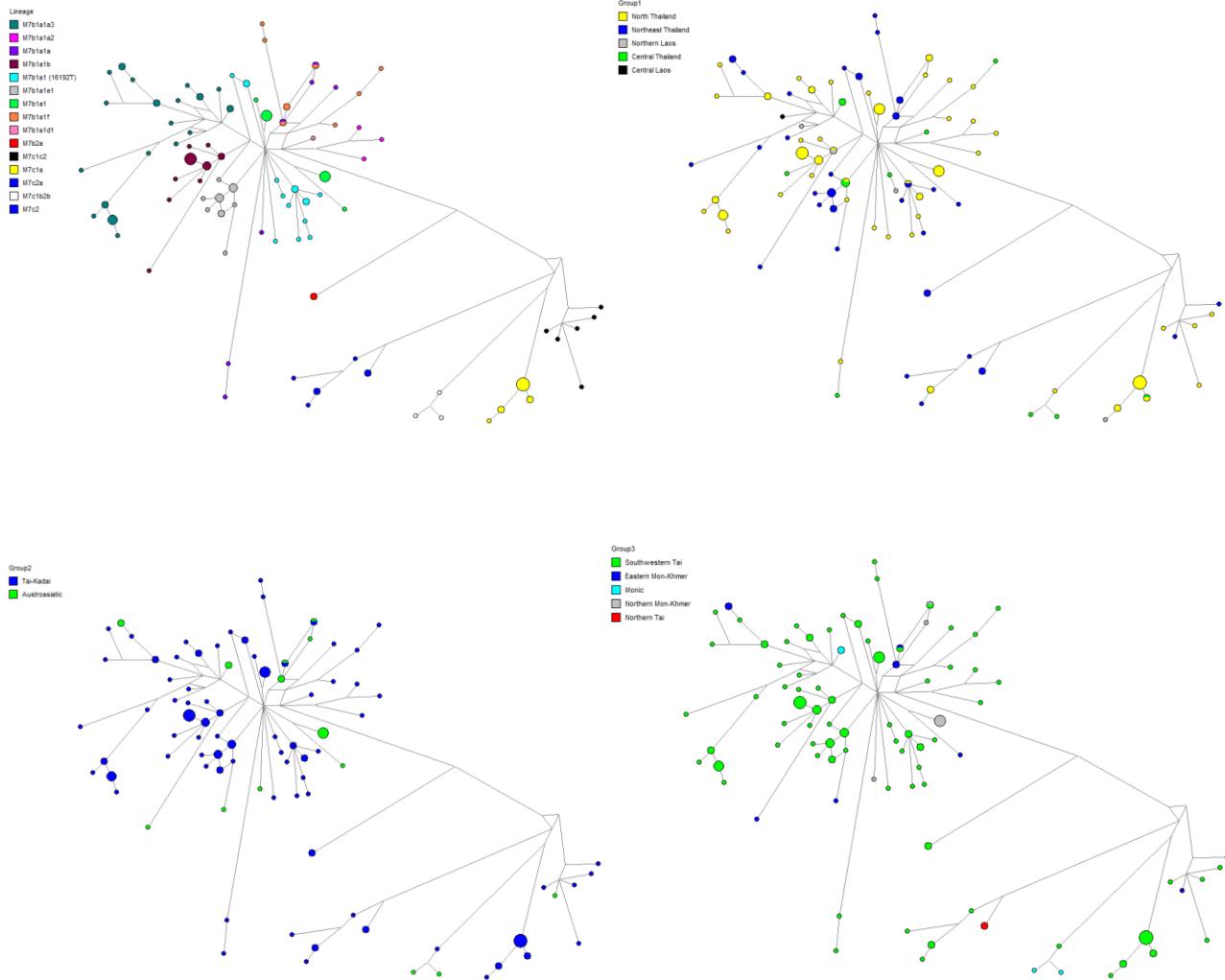
B5a



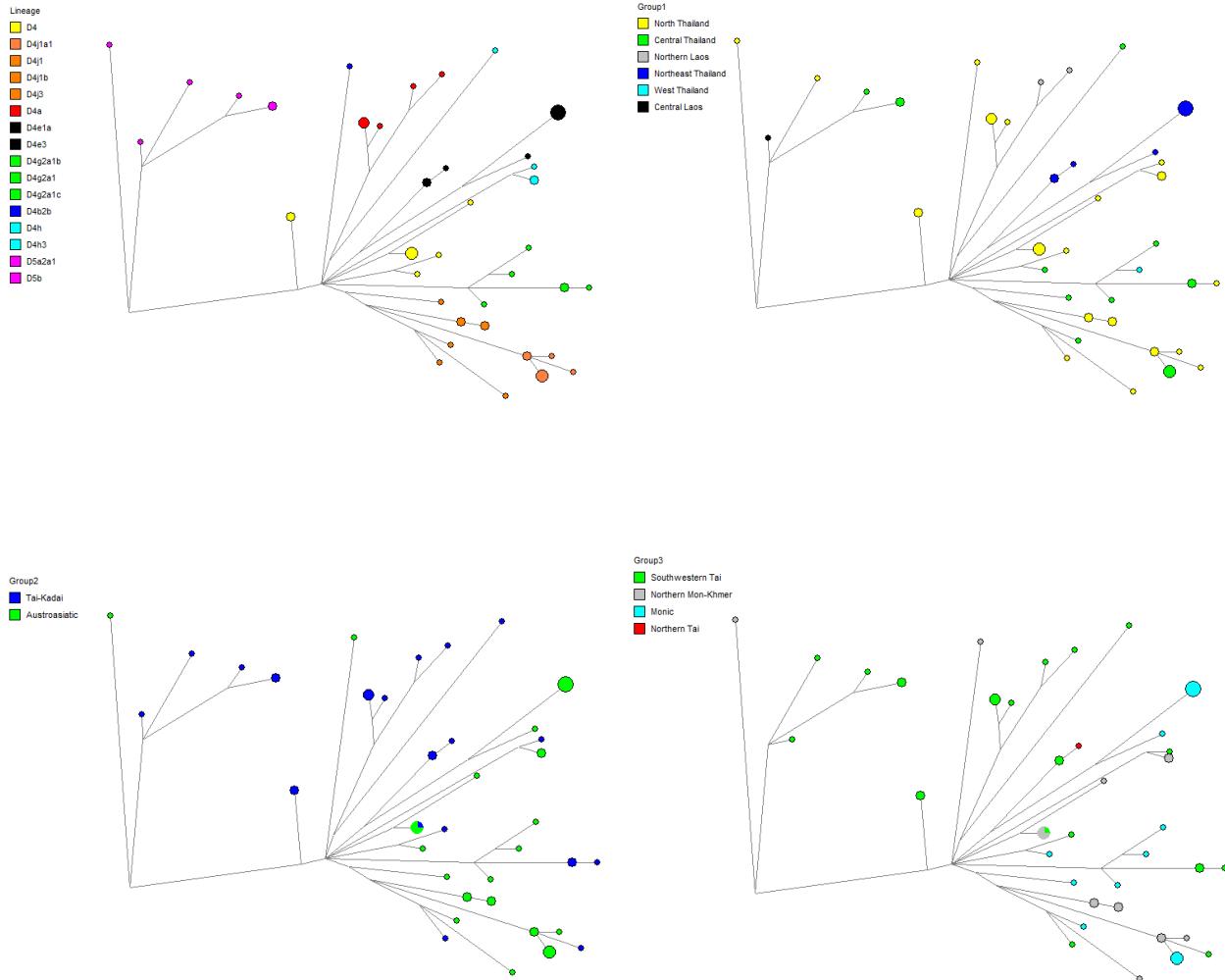
## B4



M7



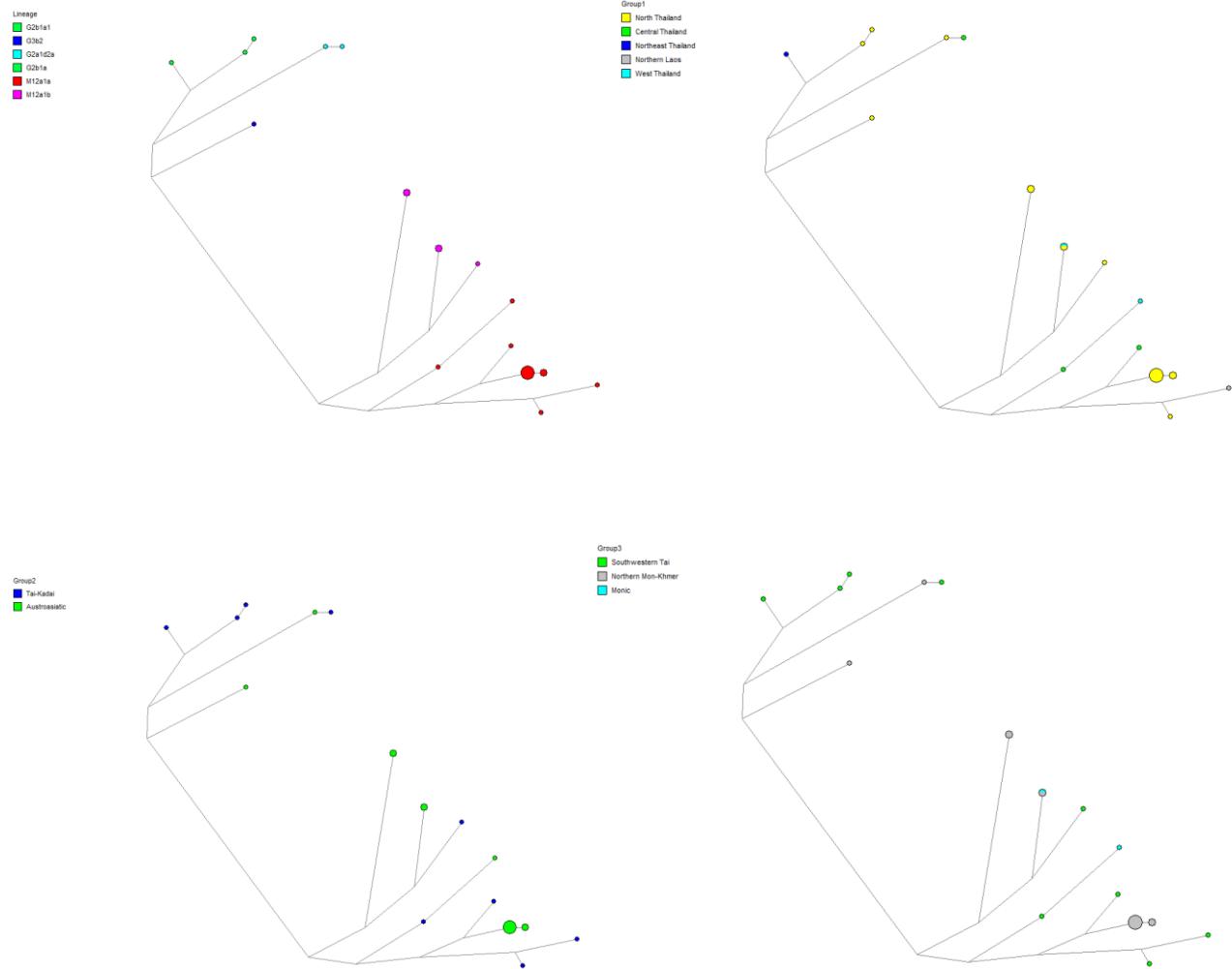
D



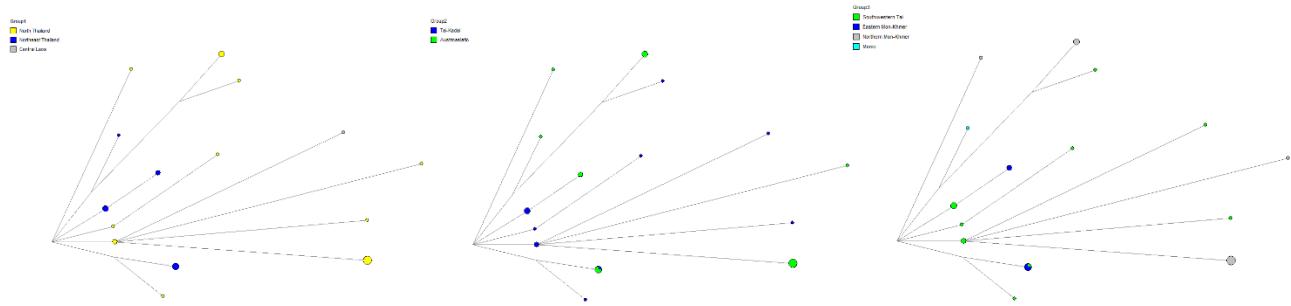
## M8



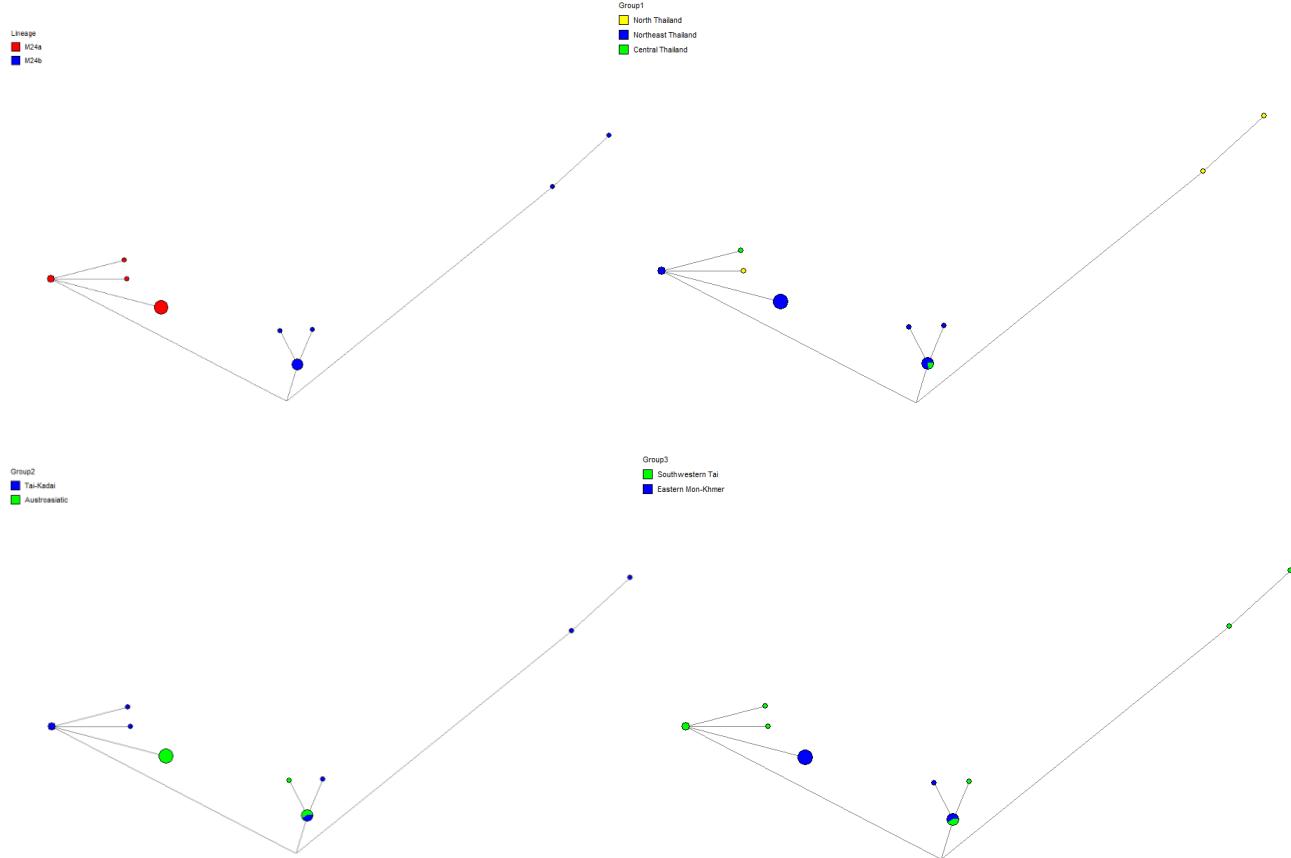
## M12-G



M20



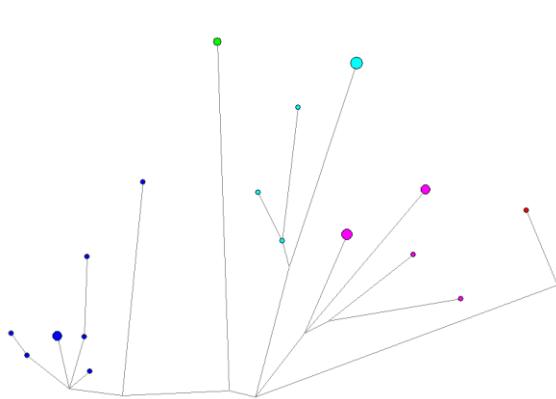
M24



# M74

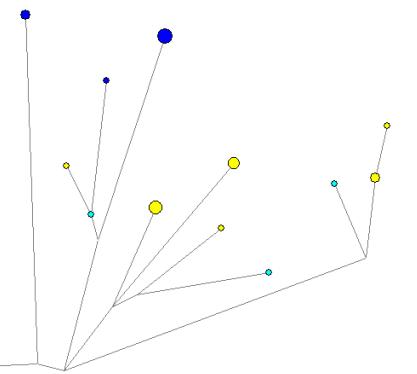
Lineage

- M74a
- M74a2
- M74a1
- M74b
- M74



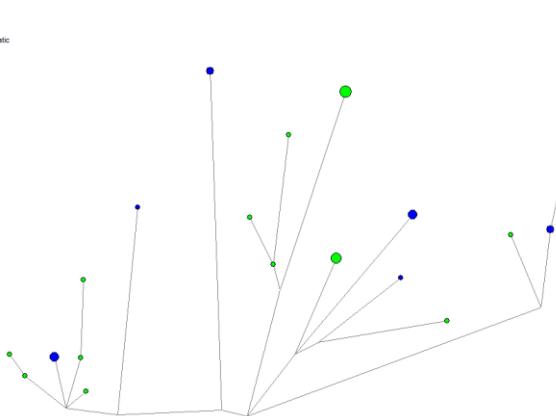
Group1

- North Thailand
- Northeast Thailand
- West Thailand



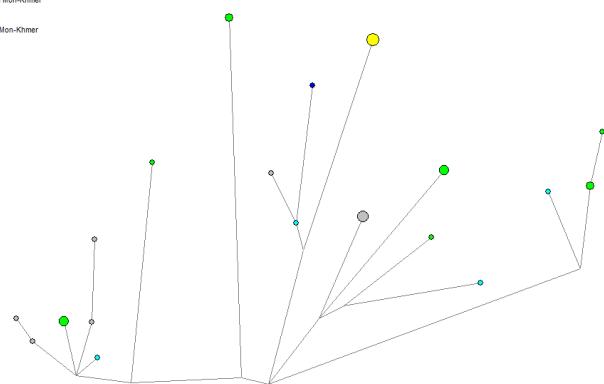
Group2

- Tai-Kadai
- Austronesian

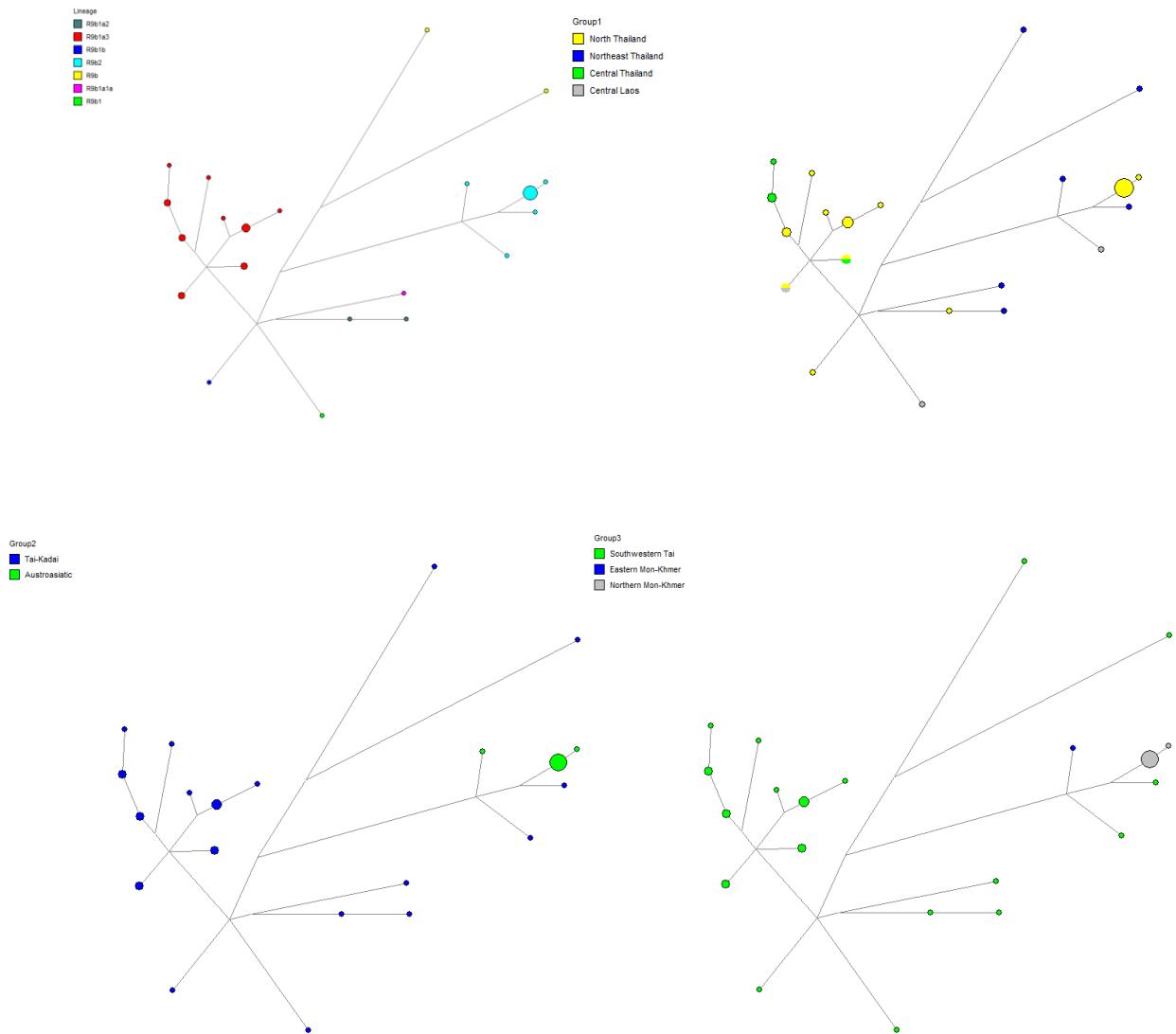


Group2

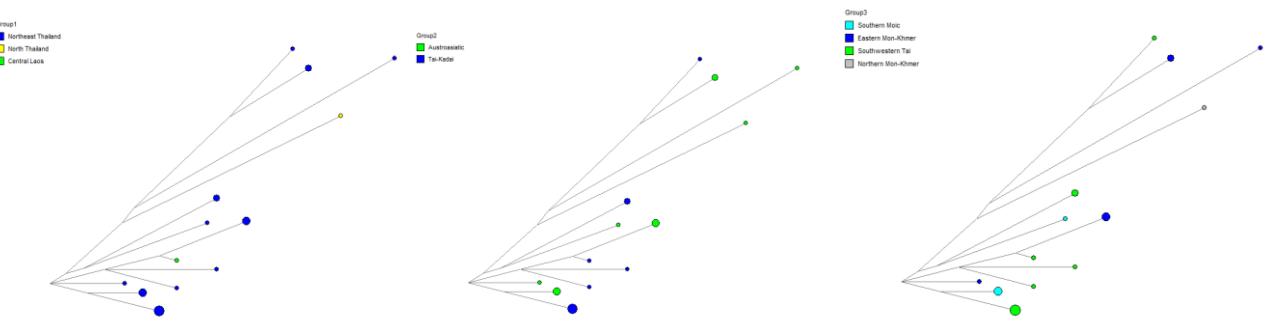
- Southwestern Tai
- Southern Mon
- Northern Mon-Khmer
- Monic
- Eastern Mon-Khmer



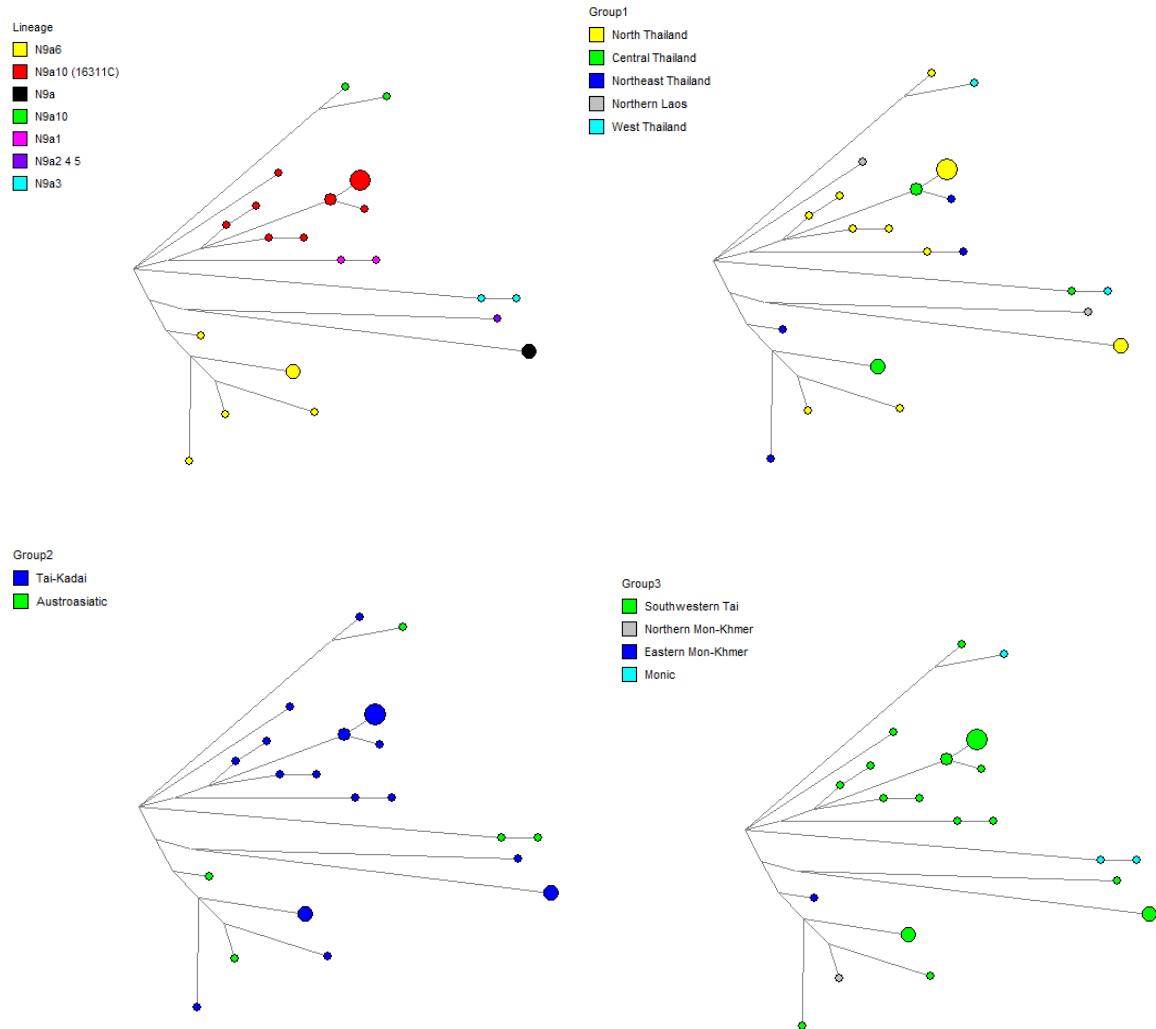
R9b



R22

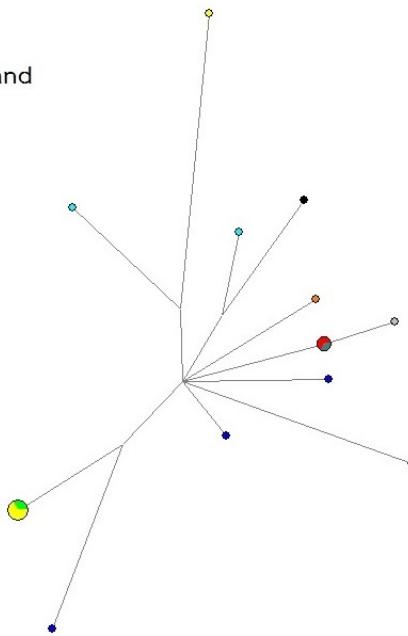


N9a



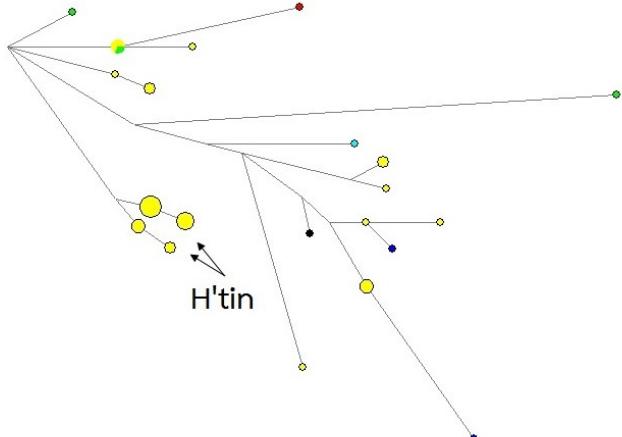
B4c2

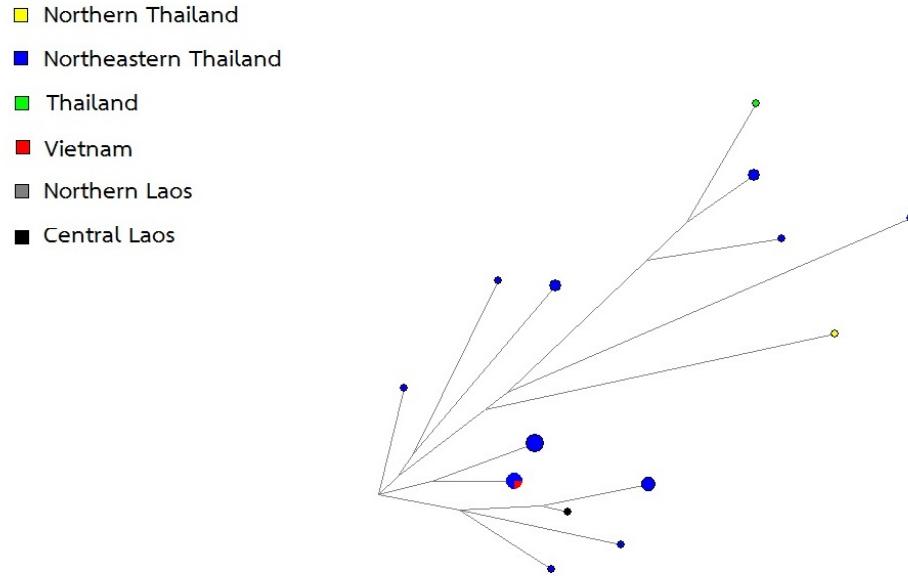
- Northern Thailand
- Northeastern Thailand
- Central Thailand
- Northern Laos
- China
- Vietnam
- Thailand
- Malaysia
- Taiwan



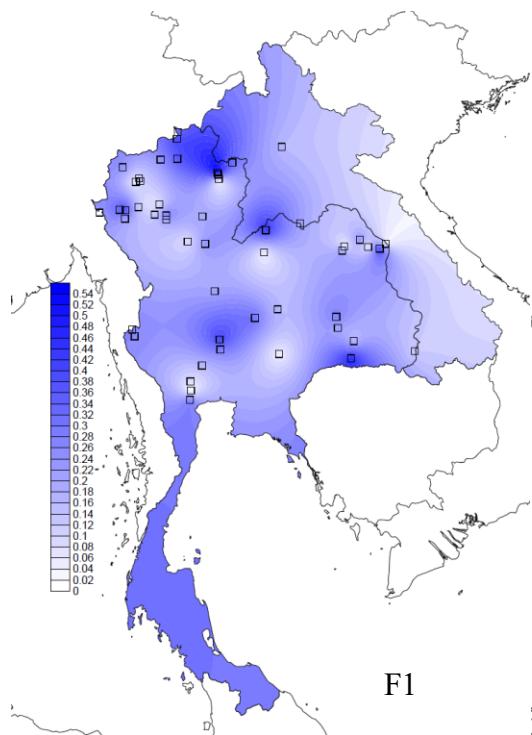
B6a

- Northern Thailand
- Central Thailand
- Philippines
- Southern China
- Malaysia
- Myanmar

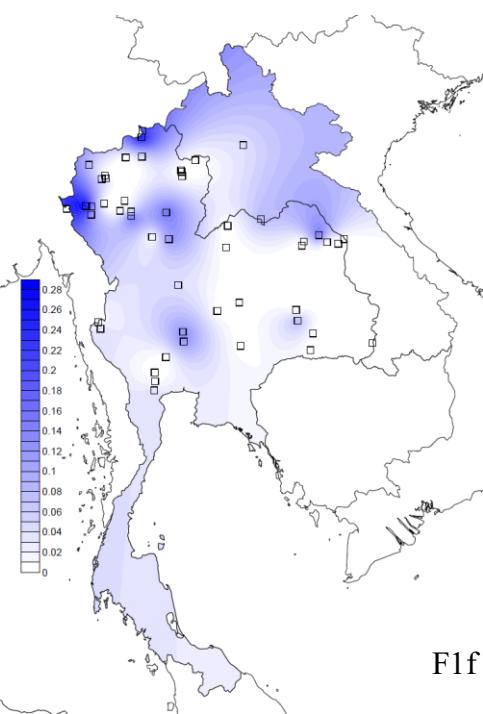




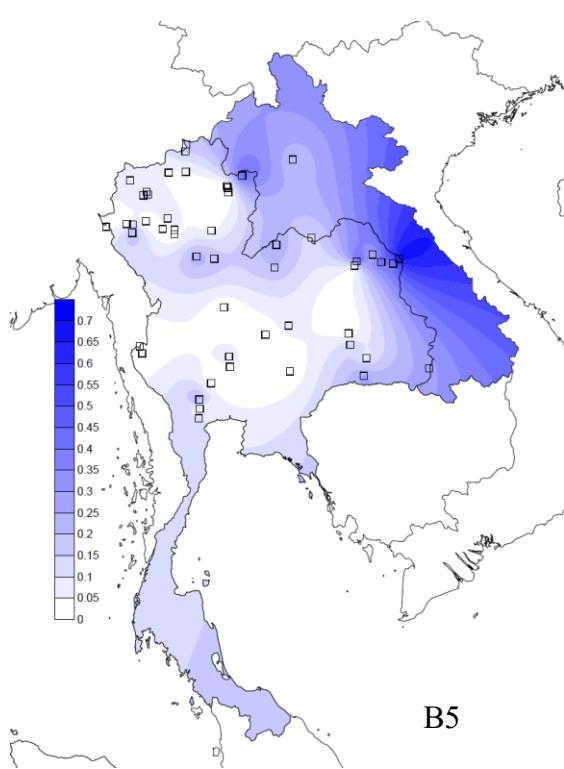
**Fig. S4** Networks of 14 major haplogroups (F1a, F1f, B5a, B4, M7, D, M8, M12-G, M20, M24, M74, R9b, R22 and N9a) based on Thai/Lao mtDNA genomes and three haplogroups (B4c2, B6a and R22) based on sequences in Thai/Lao populations and other East Asian populations. Networks were generated by the program Network and post-processed by Network Publisher. The networks were colored variously according to lineages, geographic regions and languages. Geographic and linguistic classifications are as in the AMOVA analysis (Table 1).



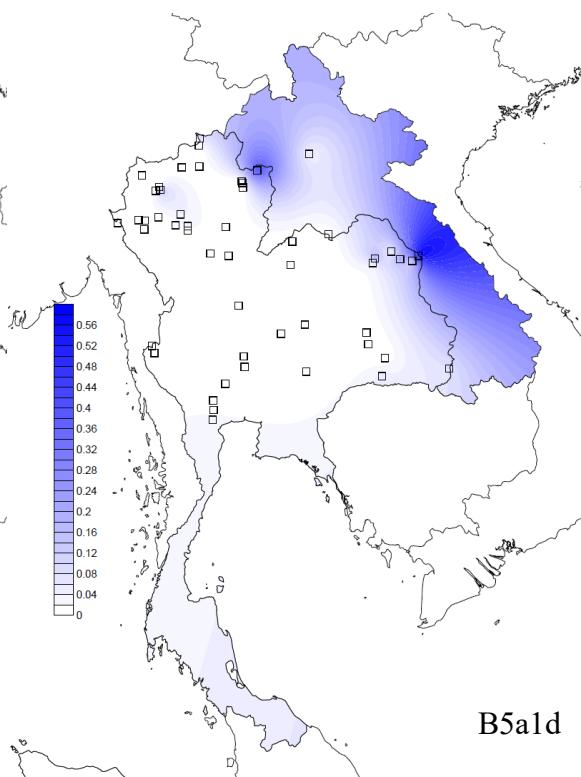
F1



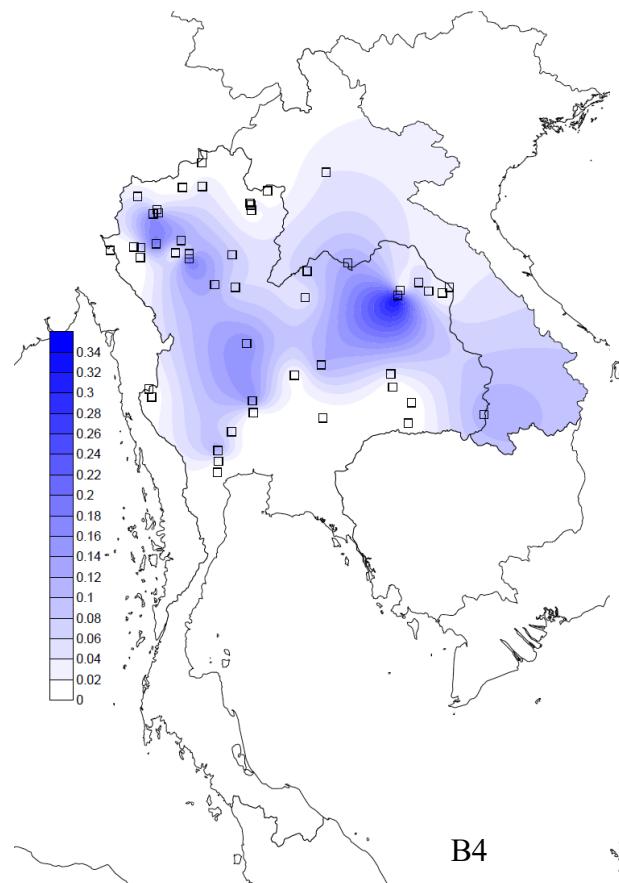
F1f



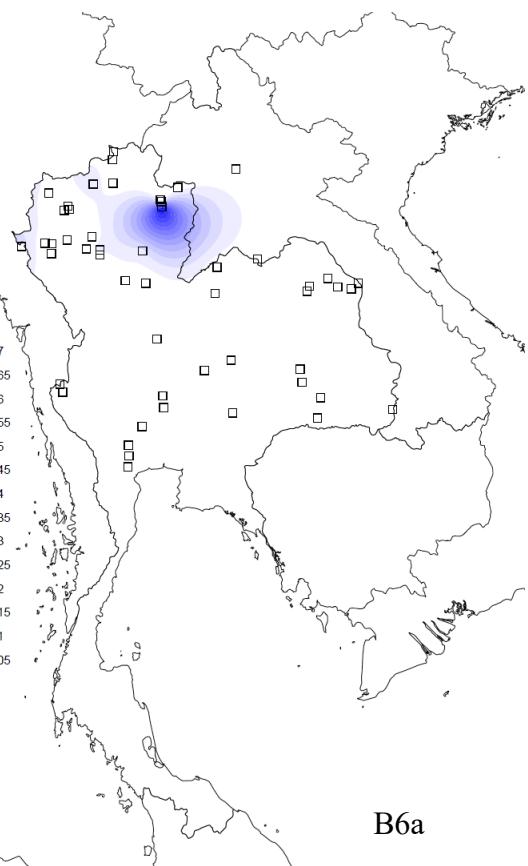
B5



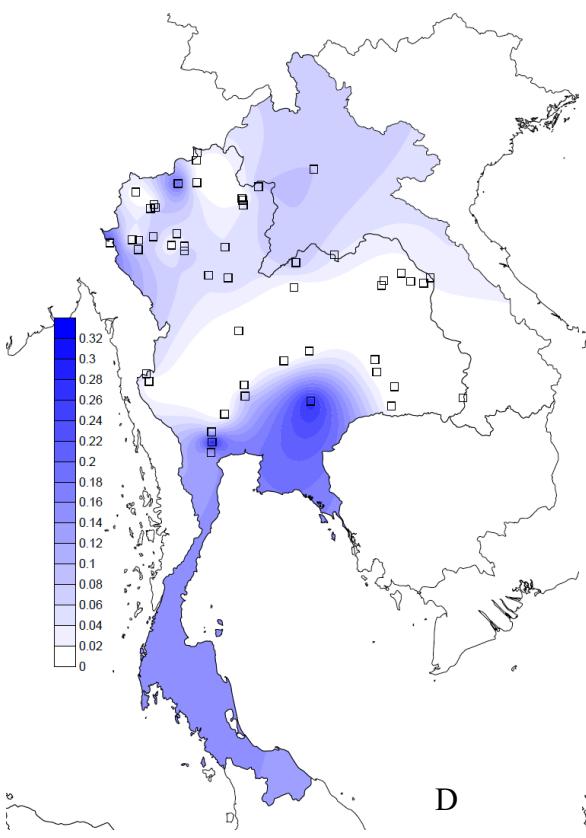
B5ald



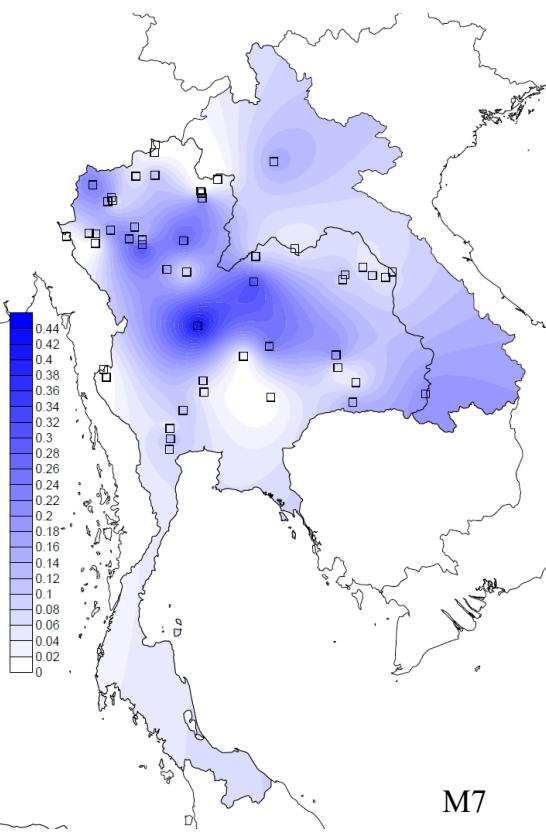
B4



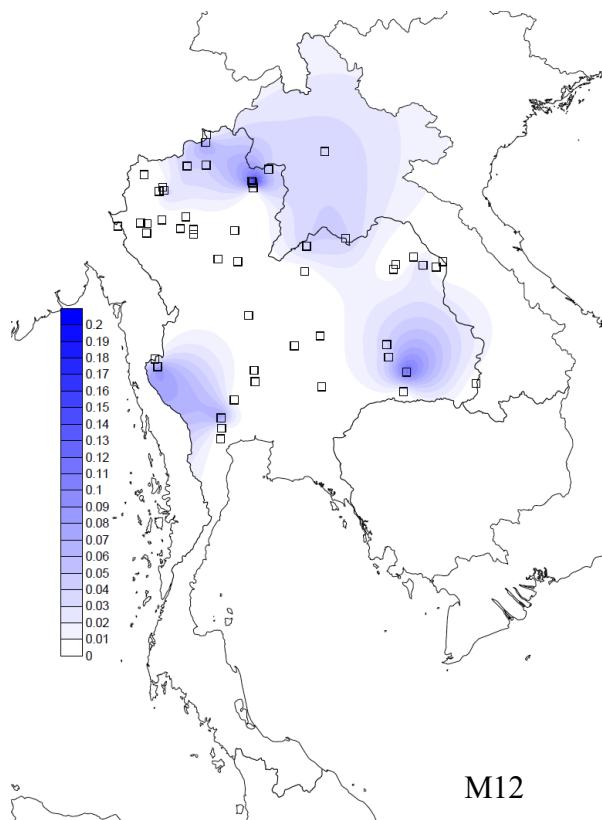
B6a



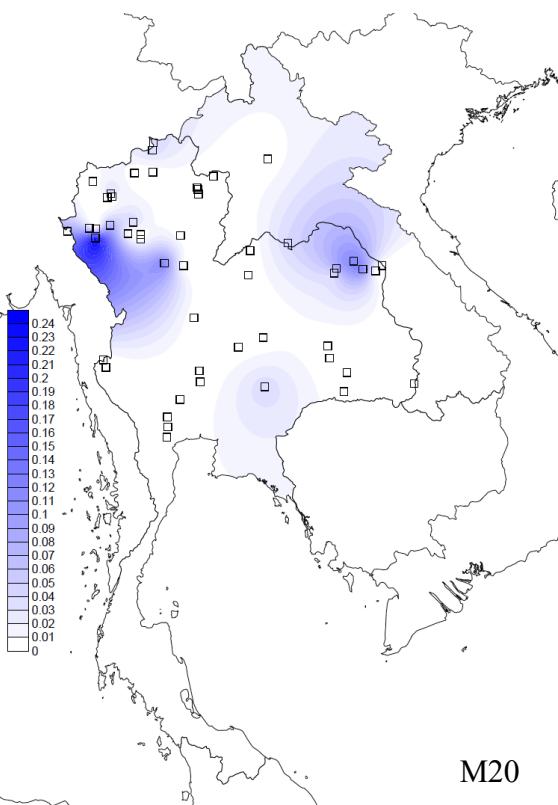
D



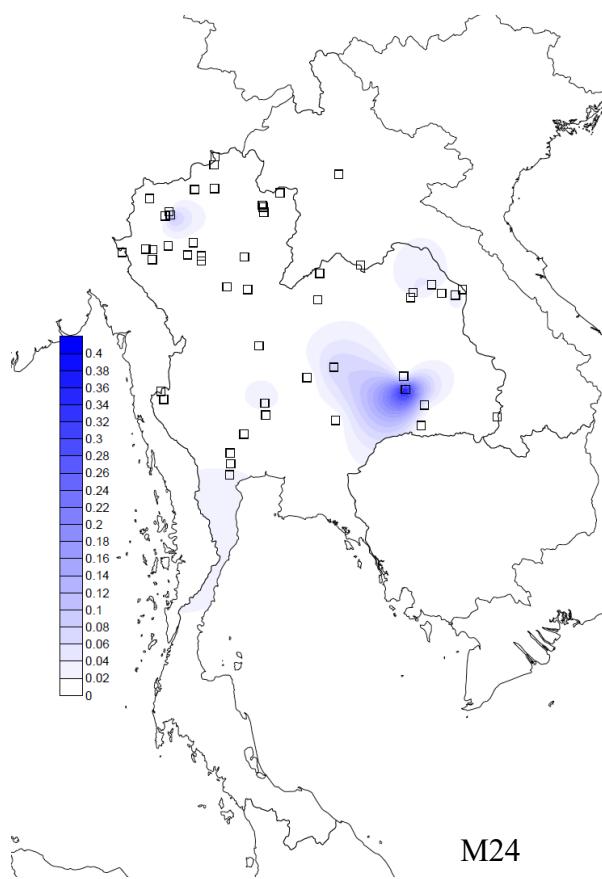
M7



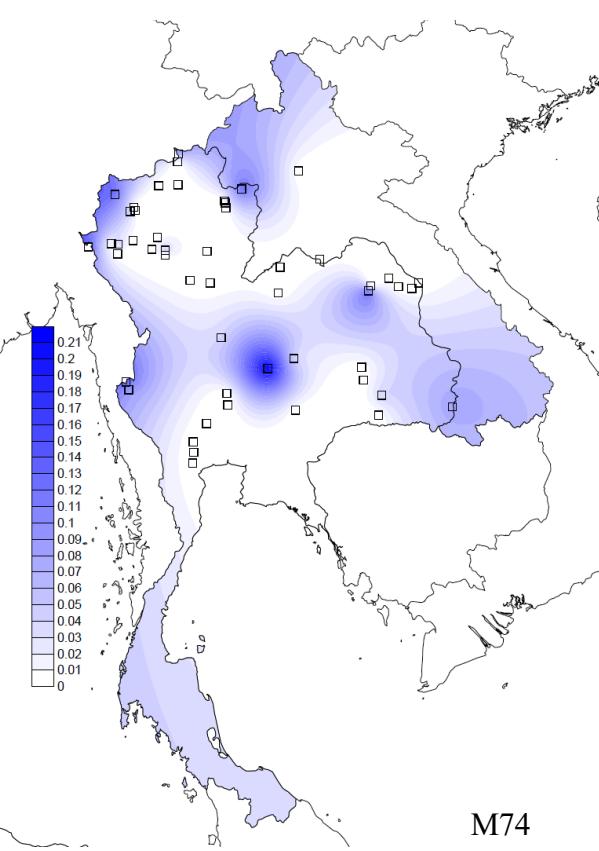
M12



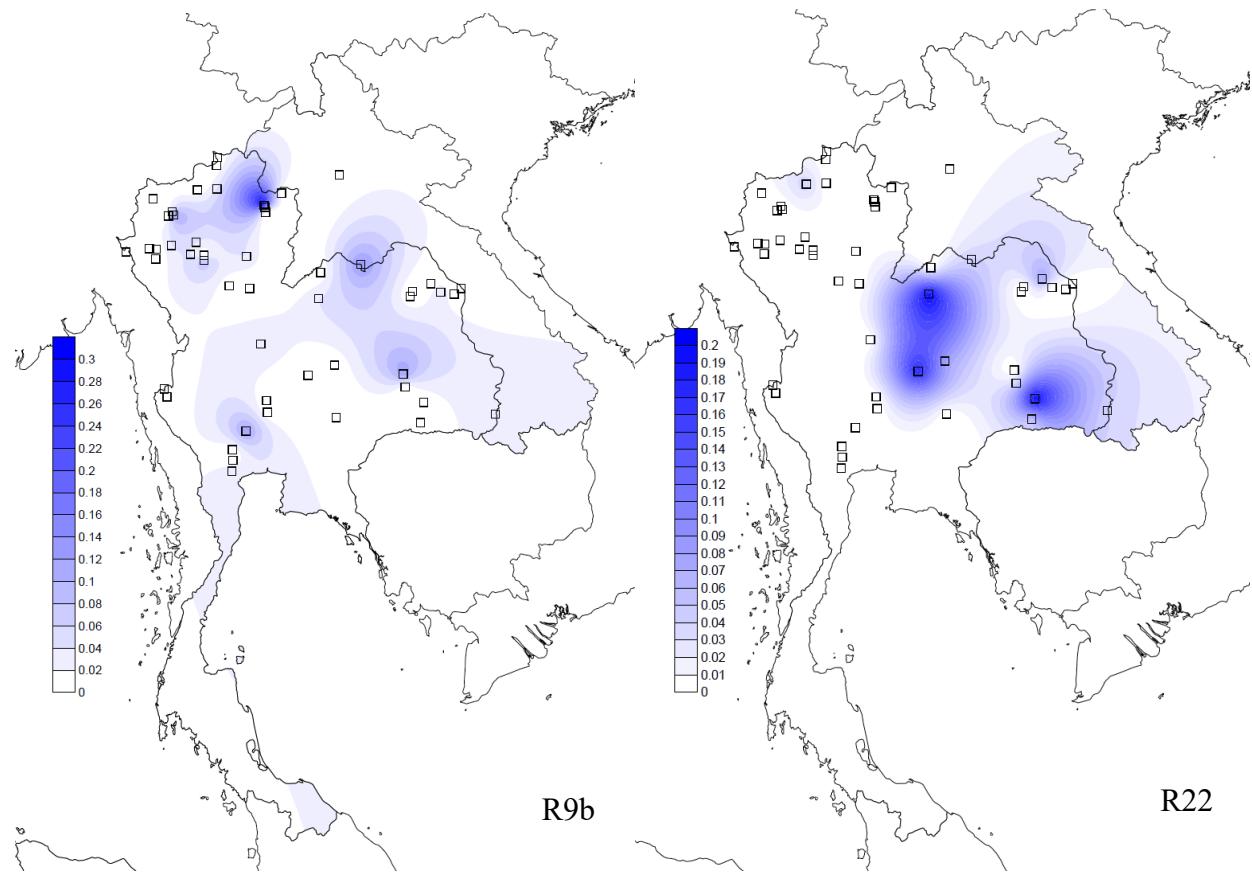
M20



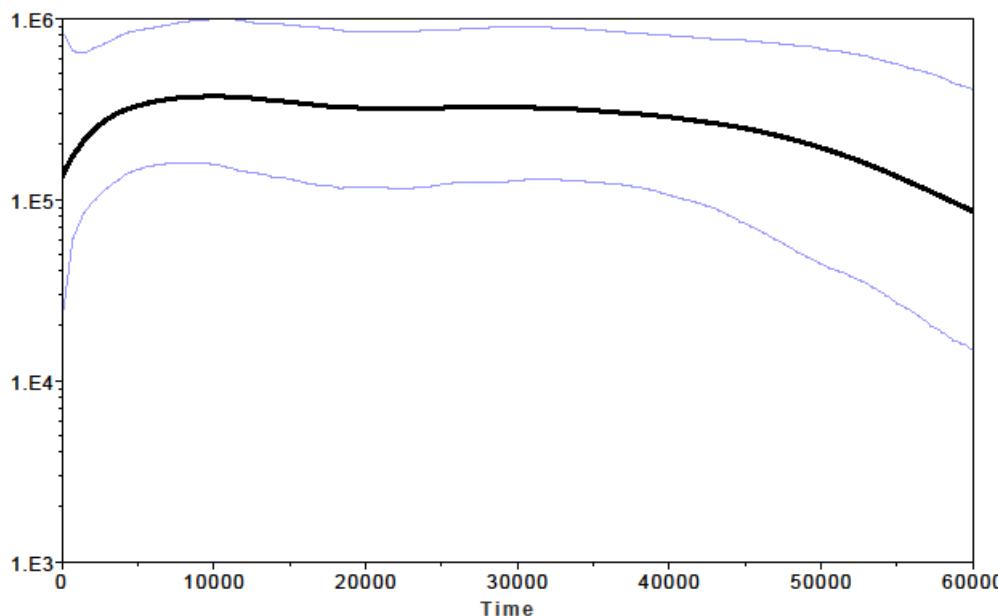
M24



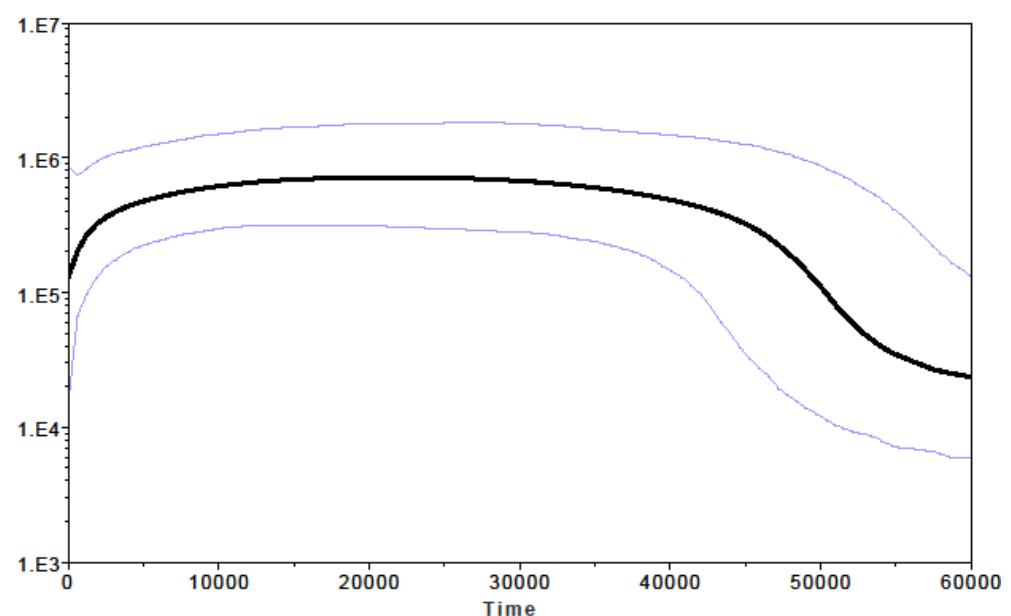
M74



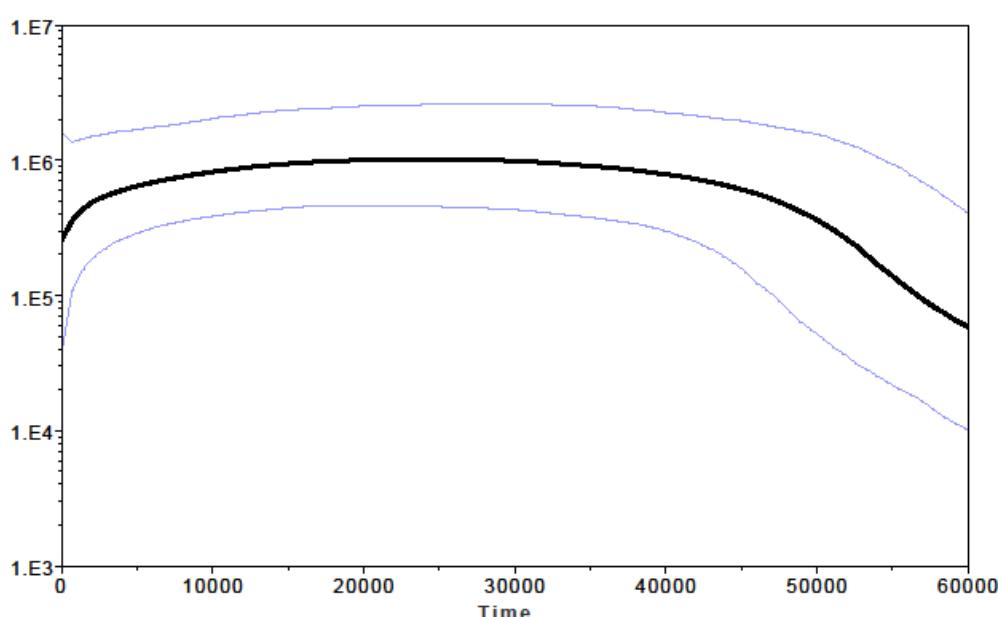
**Fig. S5** The contour maps of the frequencies of 14 haplogroups (F1, F1f, B5, B5a1d, B4, B6a, D, M7, M12, M20, M24, M74, R9b and R22) in Thai/Lao populations. The maps were generated by Golden Software Surfer 10.0.



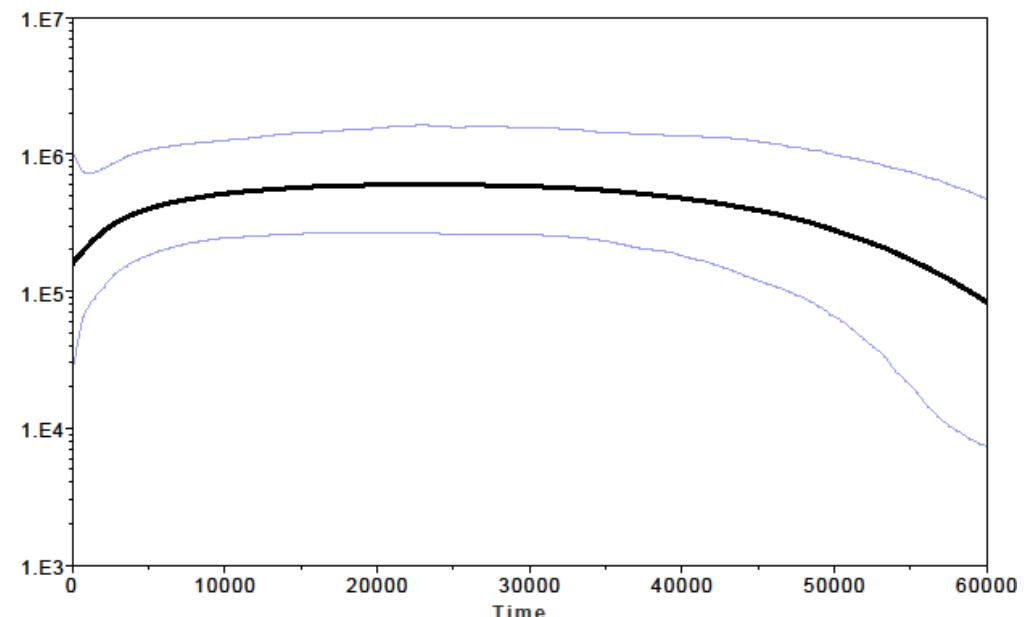
Khon Mueang1 (KM1)



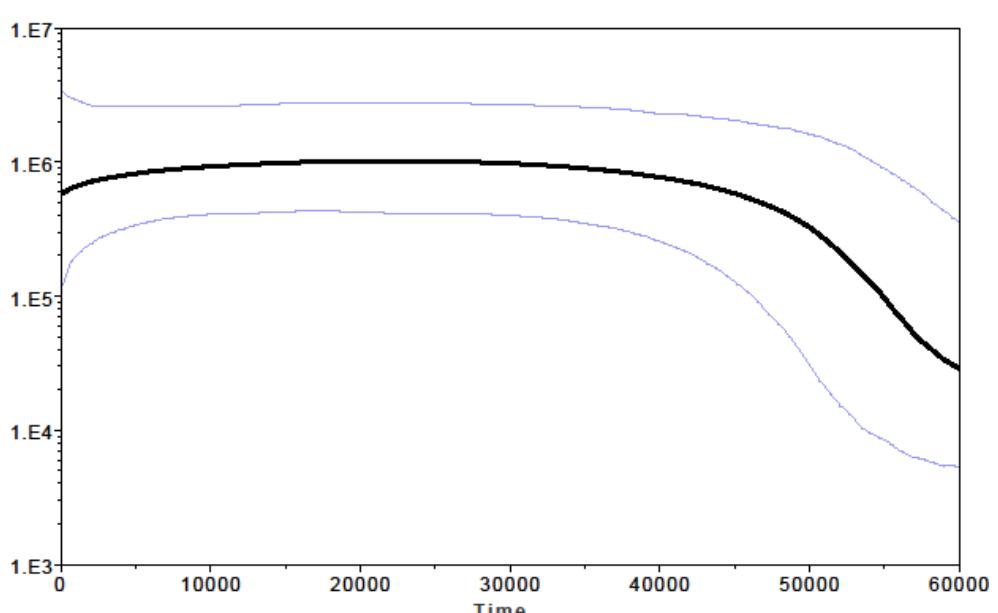
Khon Mueang2 (KM2)



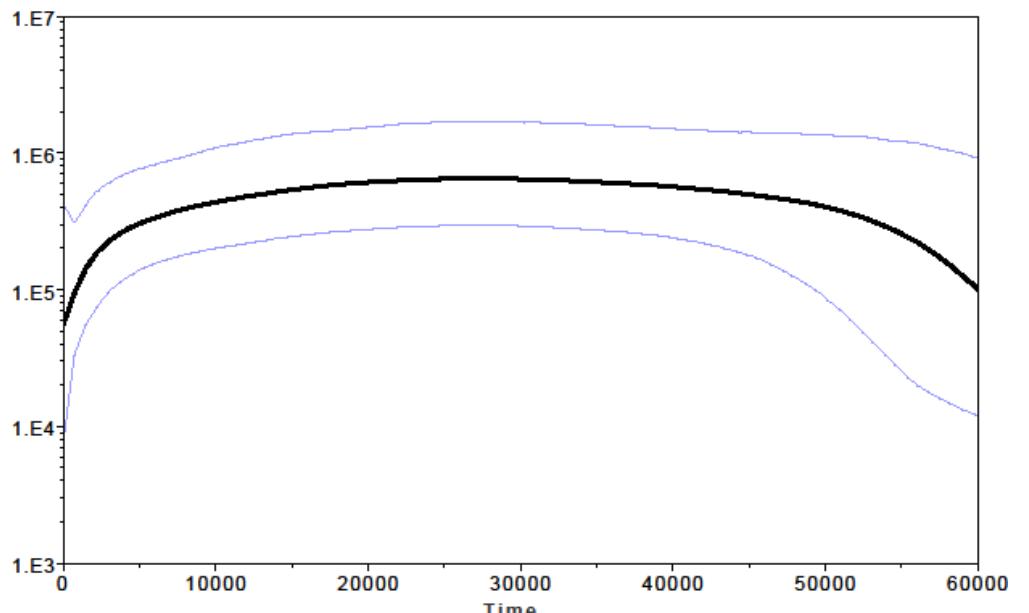
Khon Mueang3 (KM3)



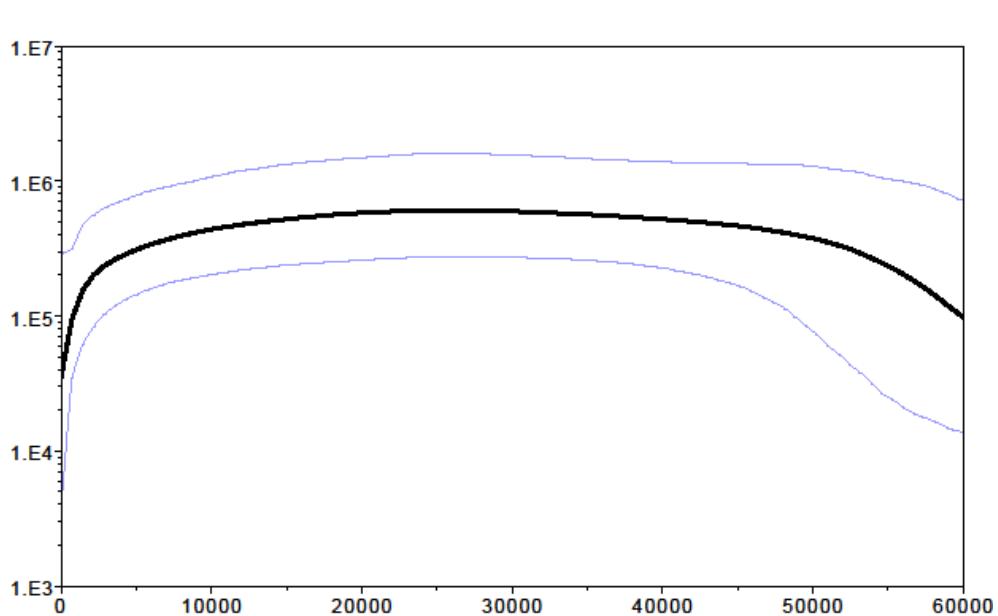
Khon Mueang4 (KM4)



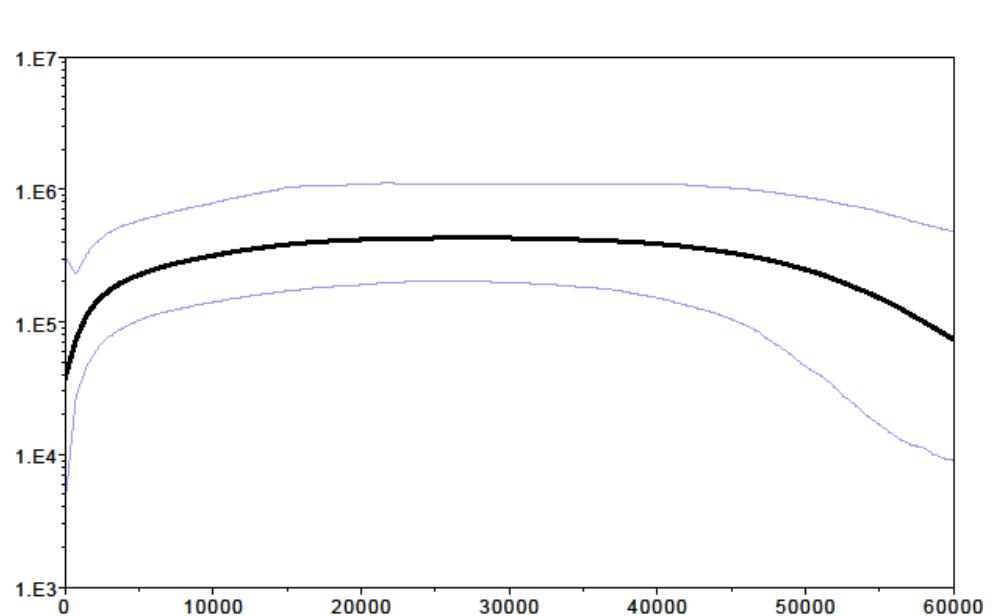
Khon Mueang5 (KM5)



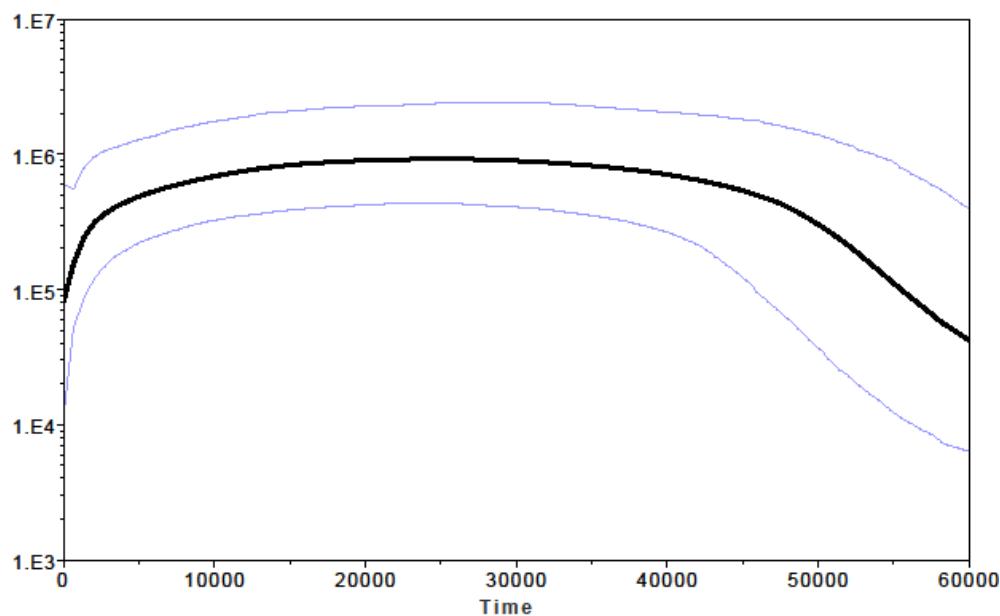
Khon Mueang6 (KM6)



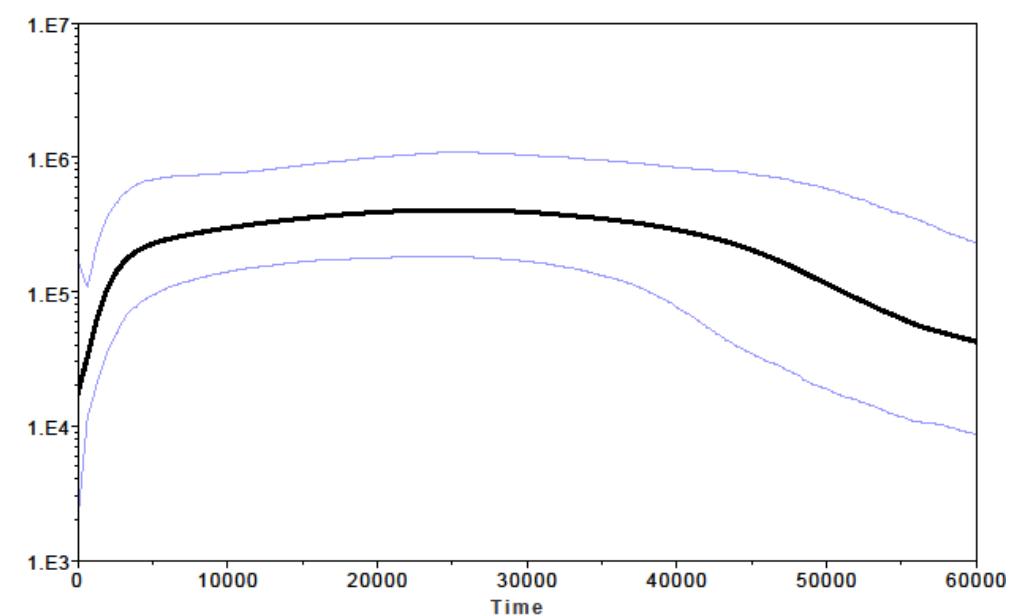
Khan Muang 7 (KM7)



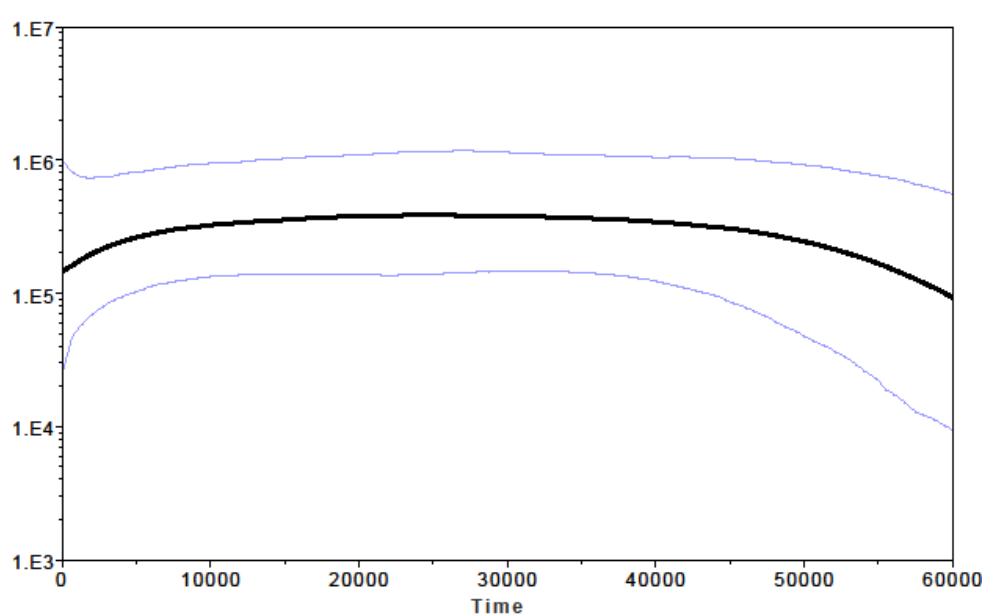
Khen Mucang8 (KM8)



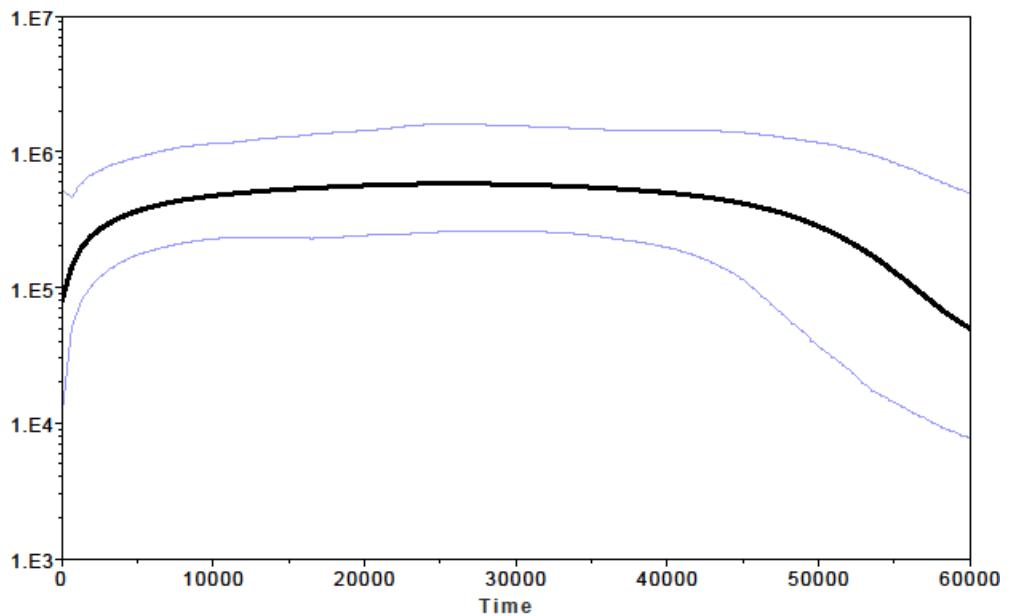
Khon Mueang9 (KM9)



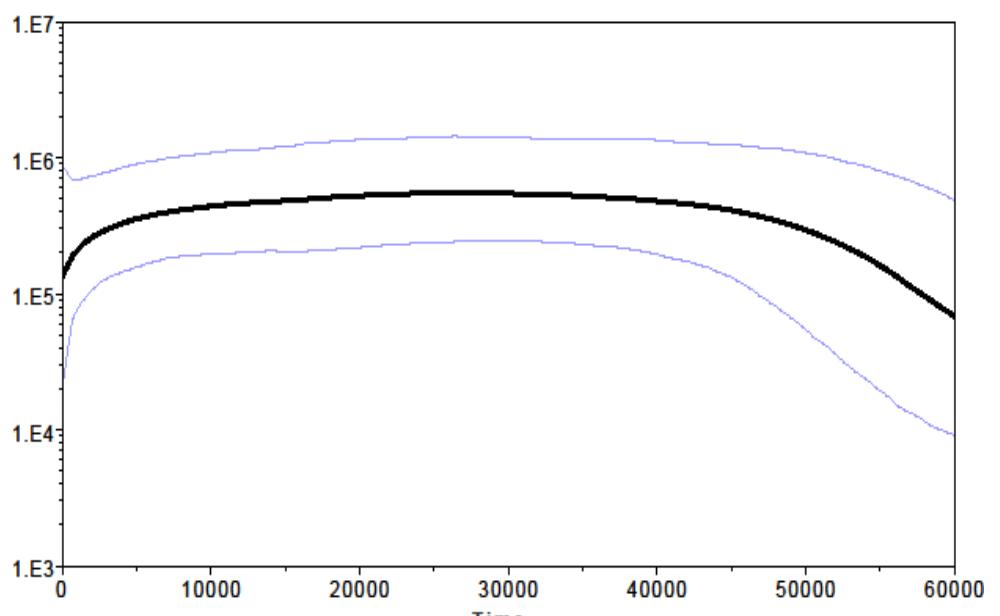
Khon Mueang10 (KM10)



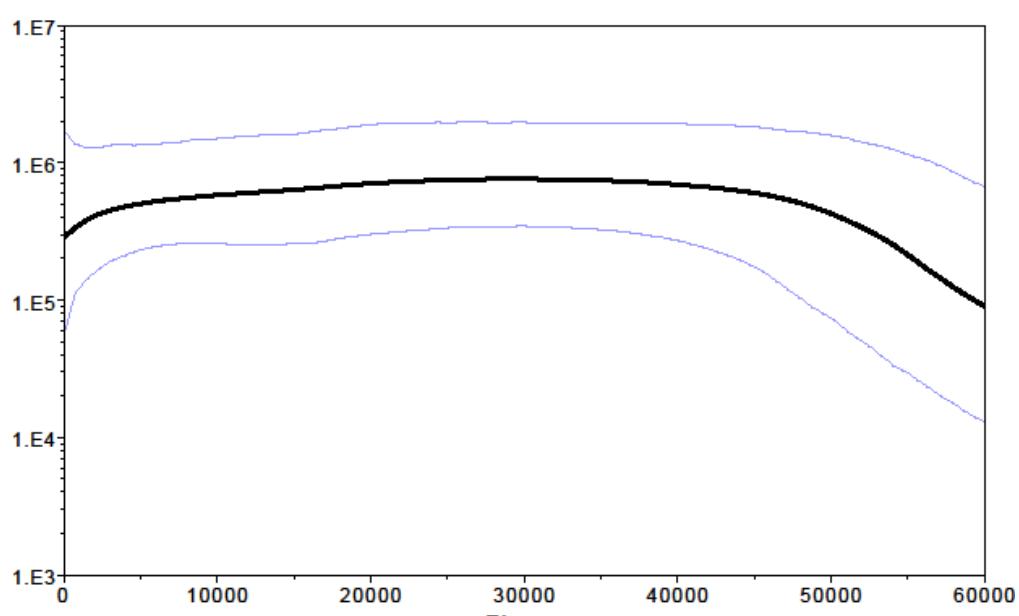
Yuan1 (YU1)



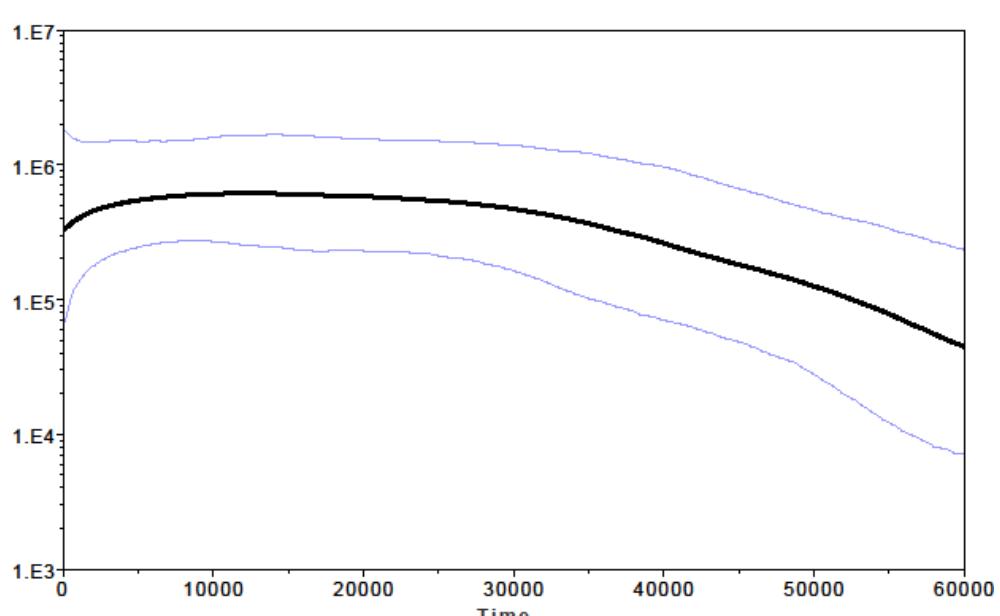
Yuan2 (YU2)



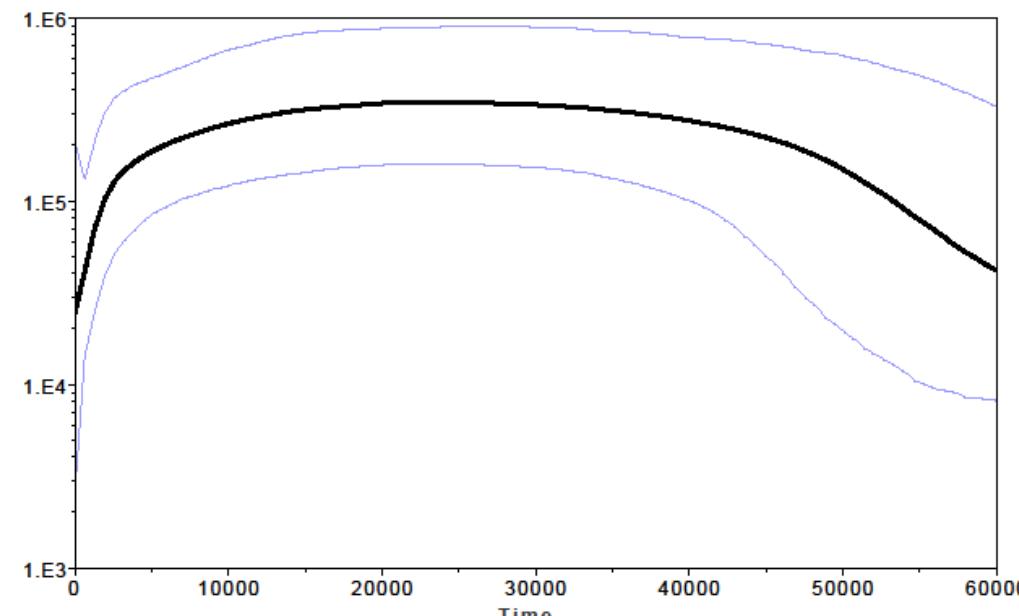
Shan (SH)



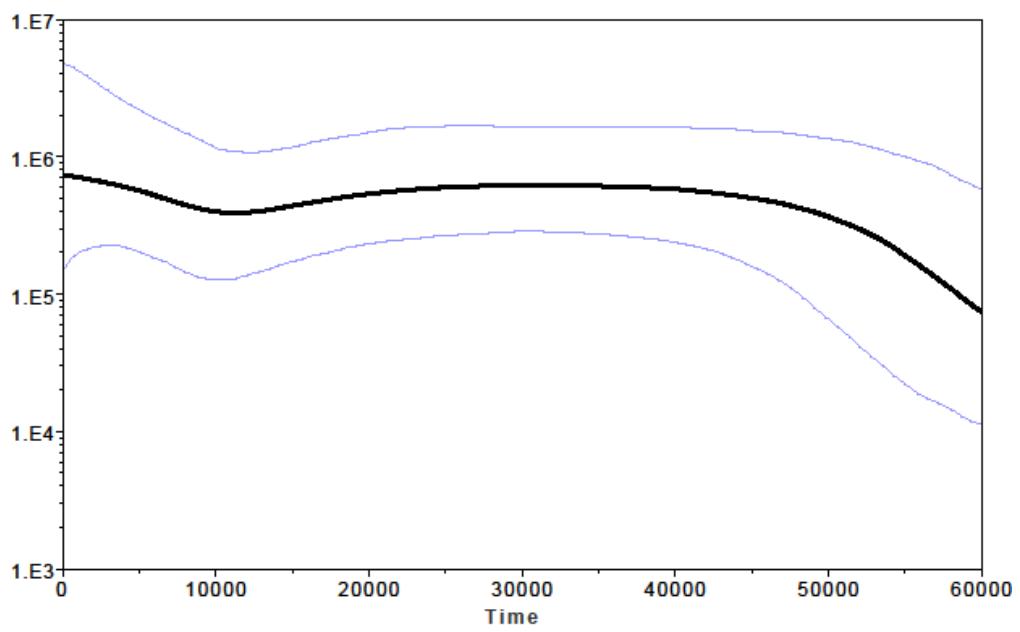
Lao Isan1 (IS1)



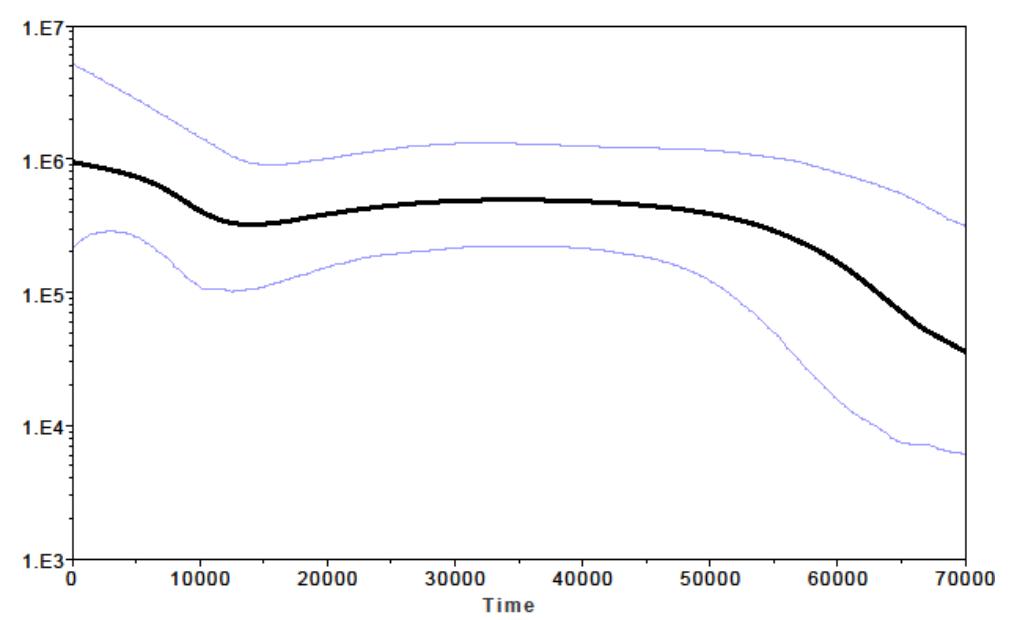
Lao Isan2 (IS2)



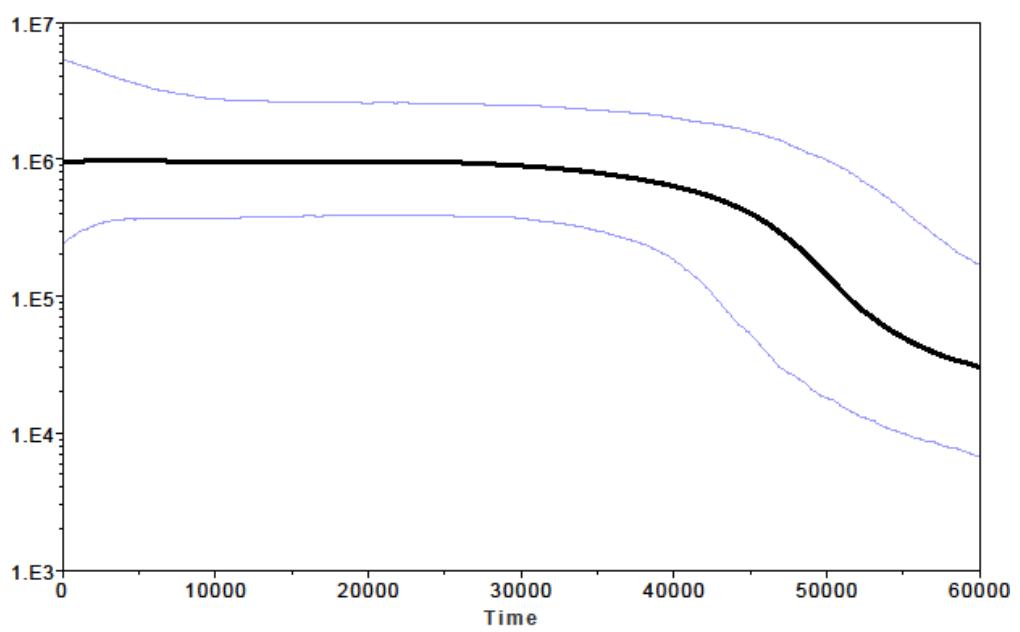
Lao Isan3 (IS3)



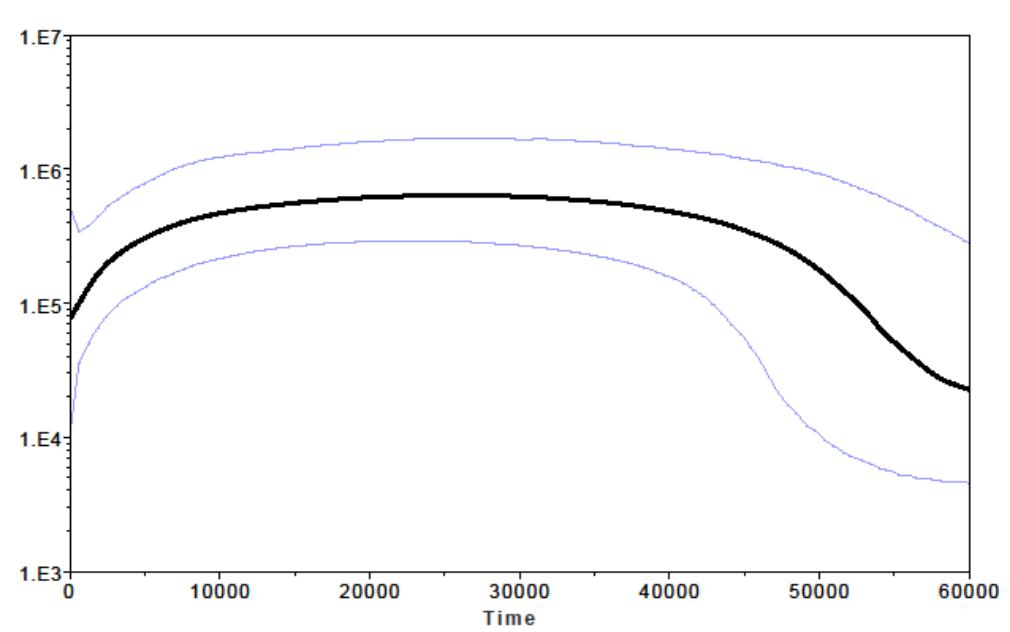
Lao Isan4 (IS4)



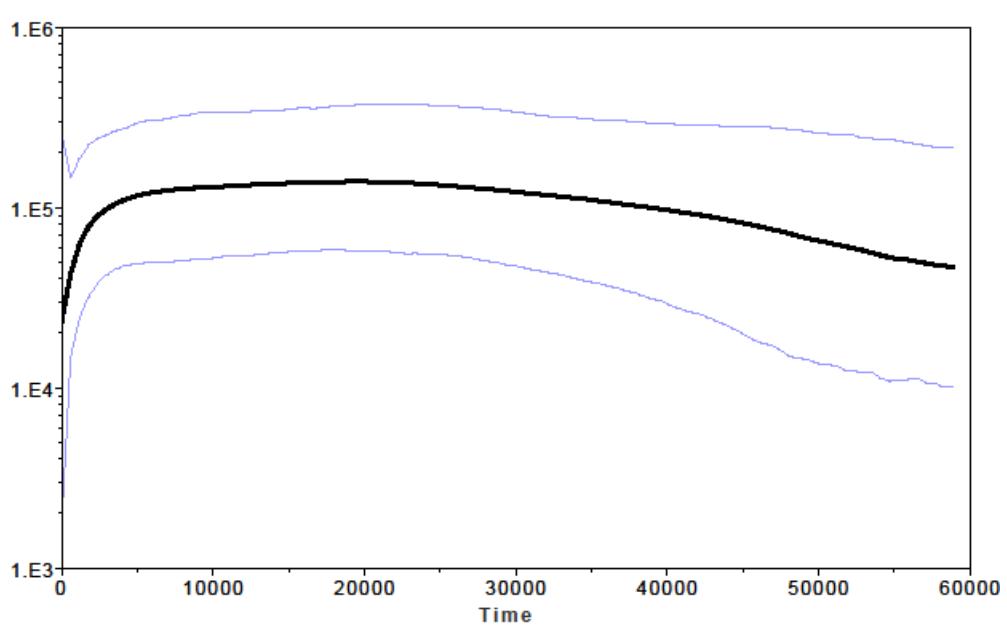
Lao Lum1 (LA1)



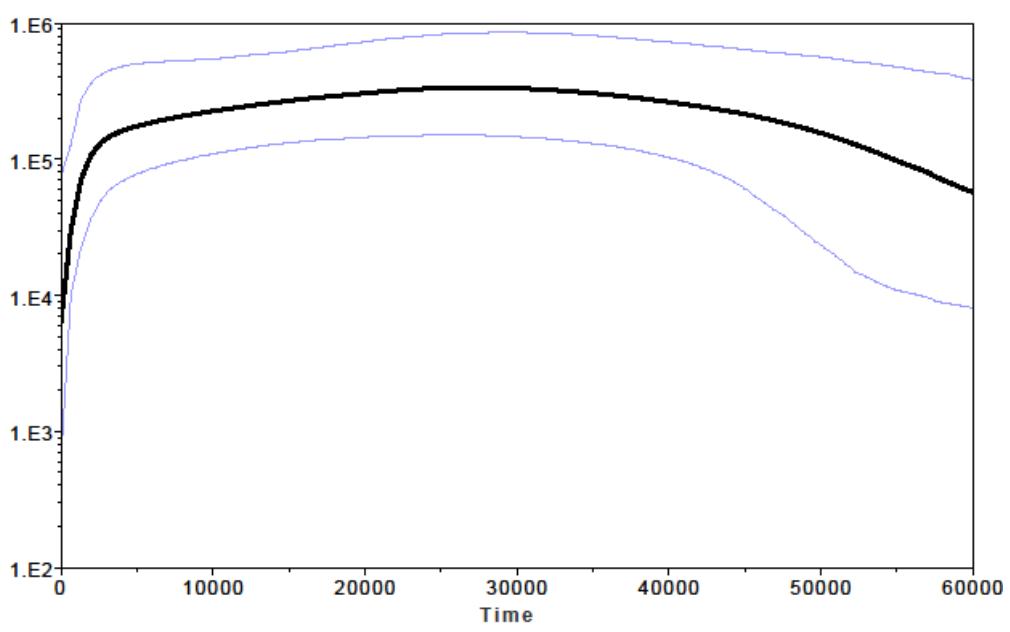
Lao Lum2 (LA2)



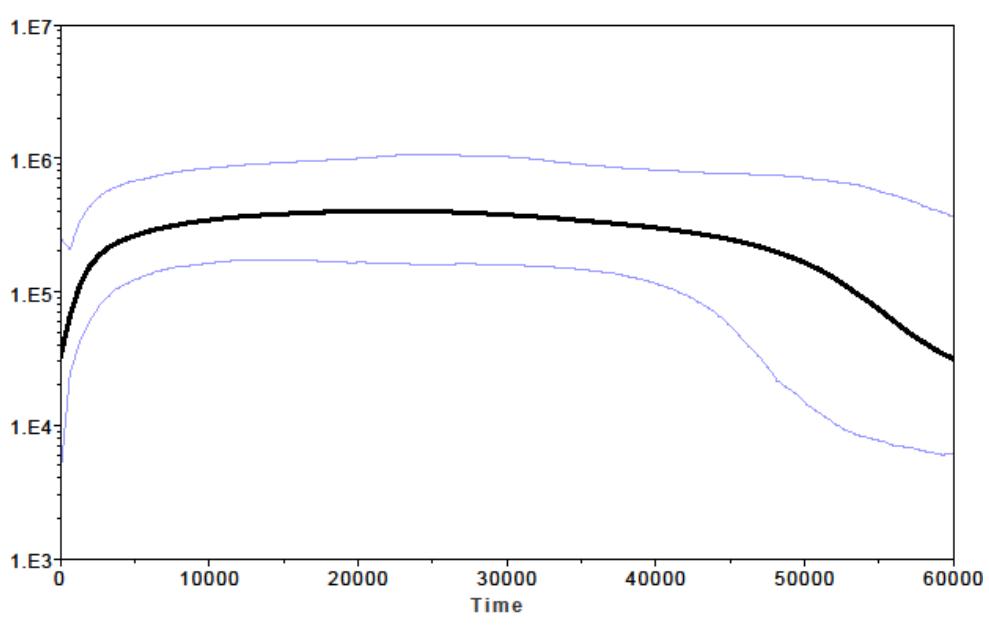
Phutai (PT)



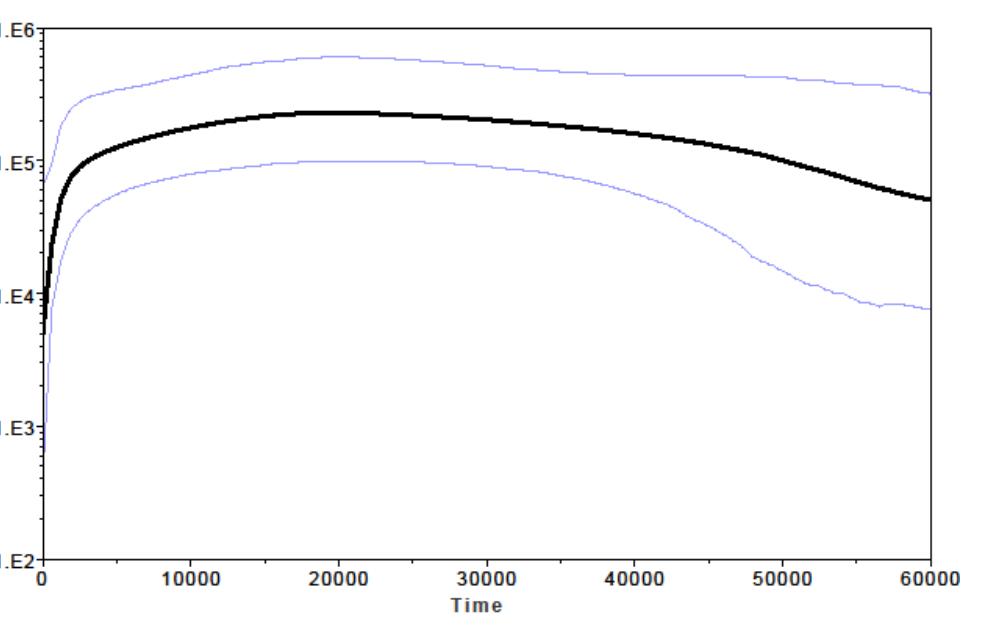
Kalueang (KL)



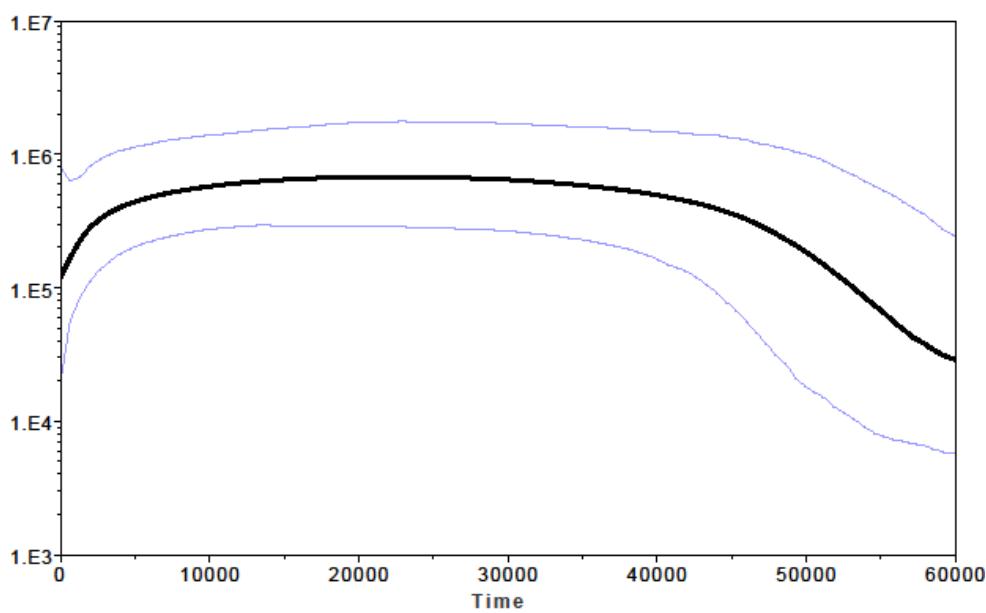
Seak (SK)



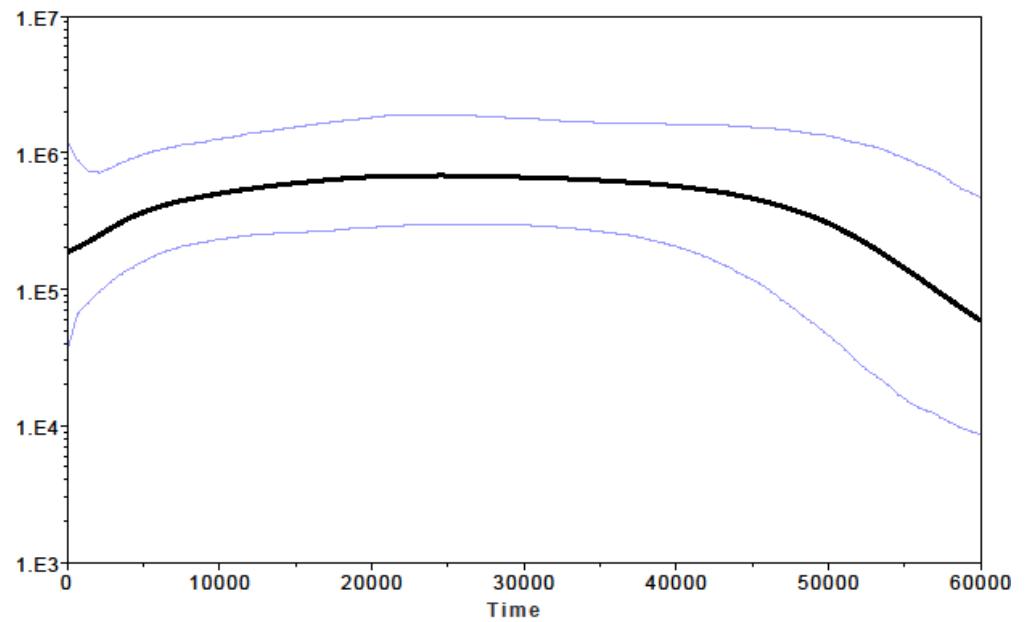
Nyaw (NY)



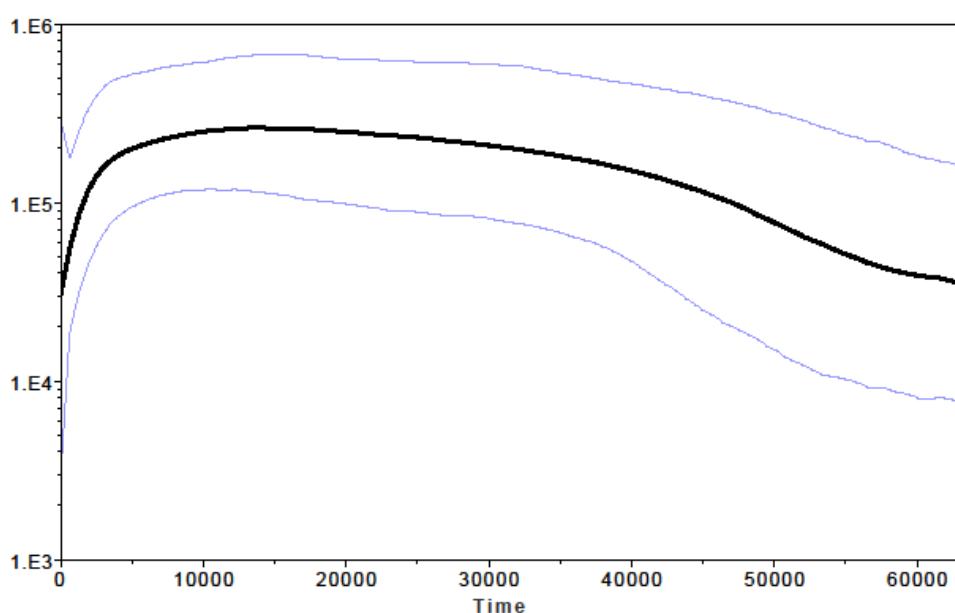
Black Tail (BT1)



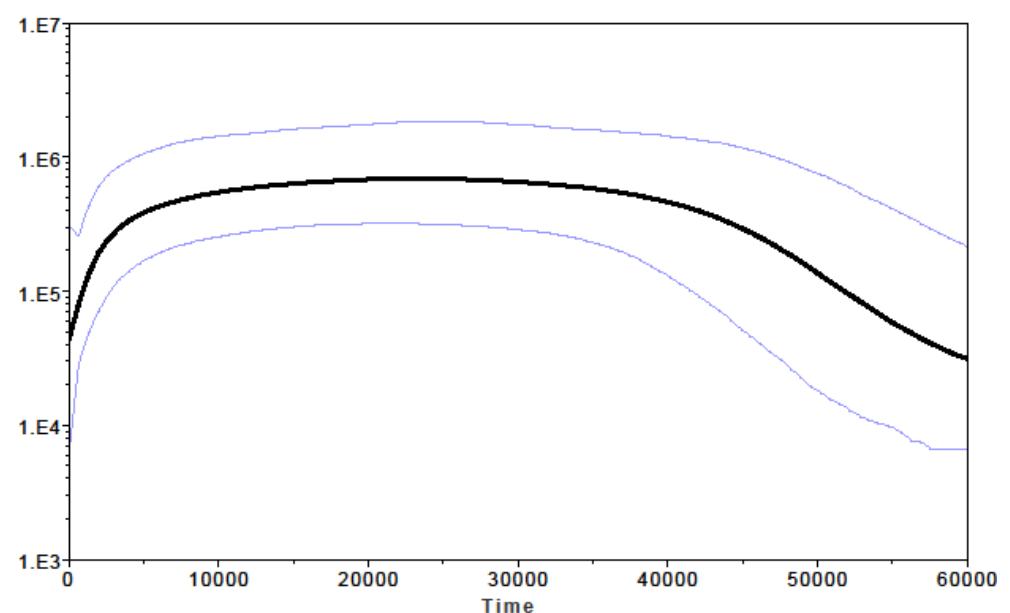
Black Tai2 (BT2)



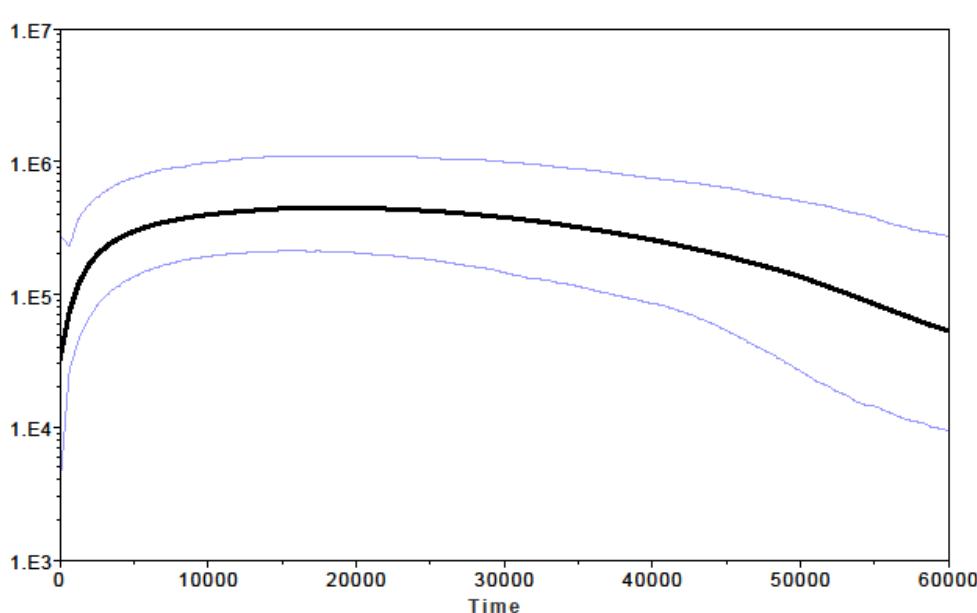
Phuan1 (PU1)



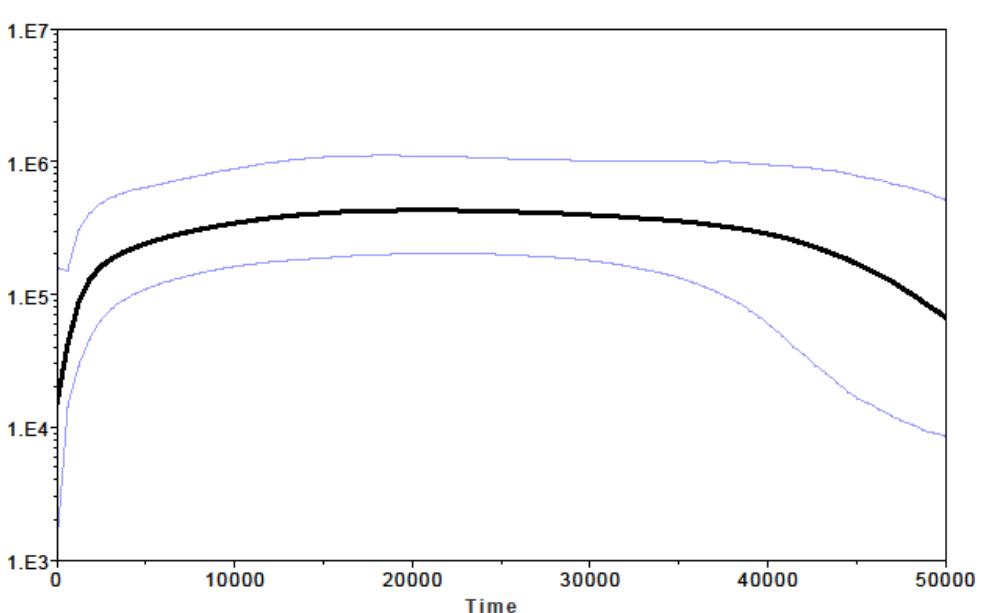
Phuan2 (PU2)



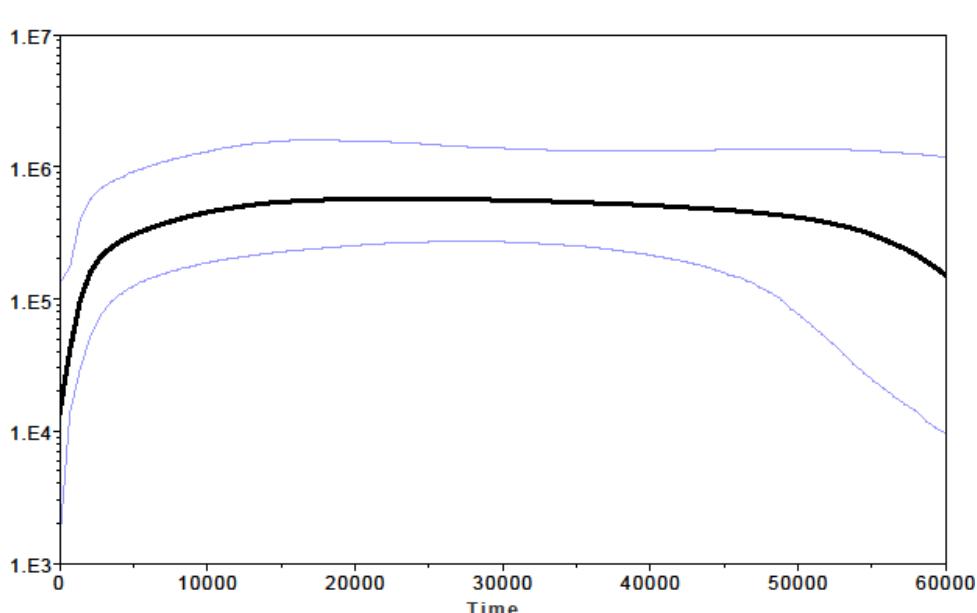
Phuan3 (PU3)



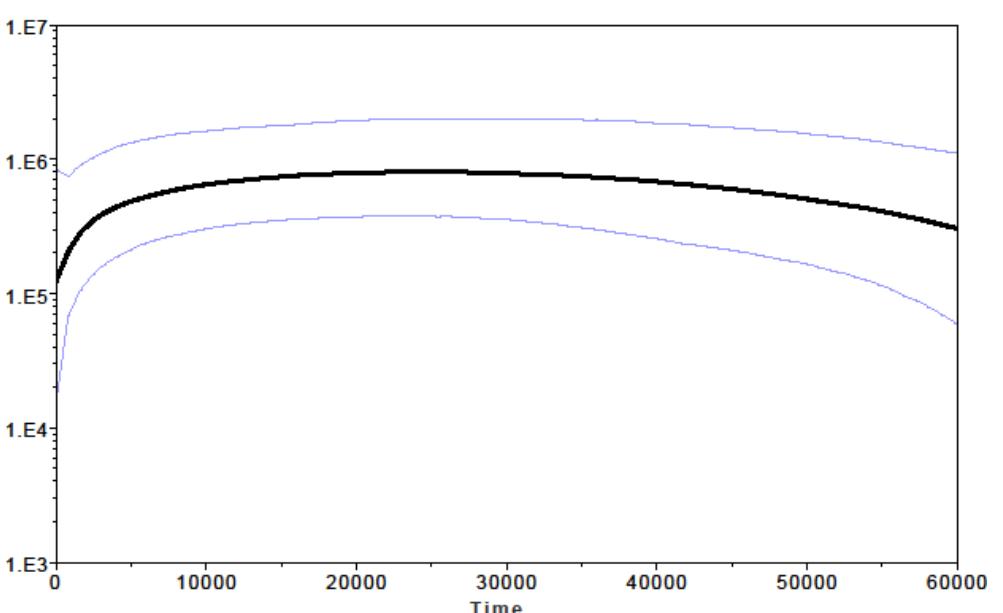
Phuan4 (PU4)



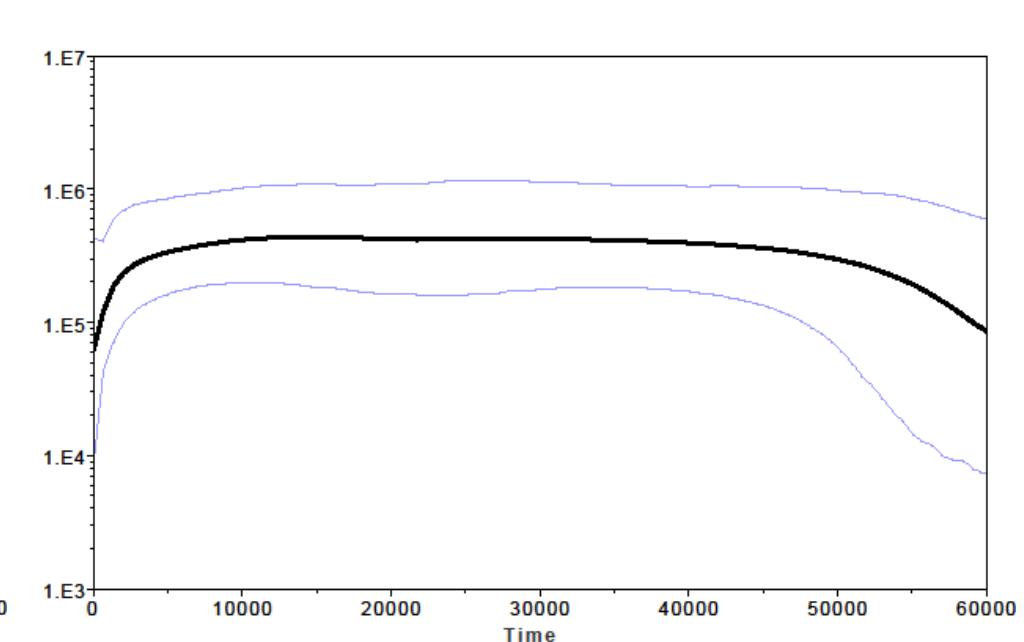
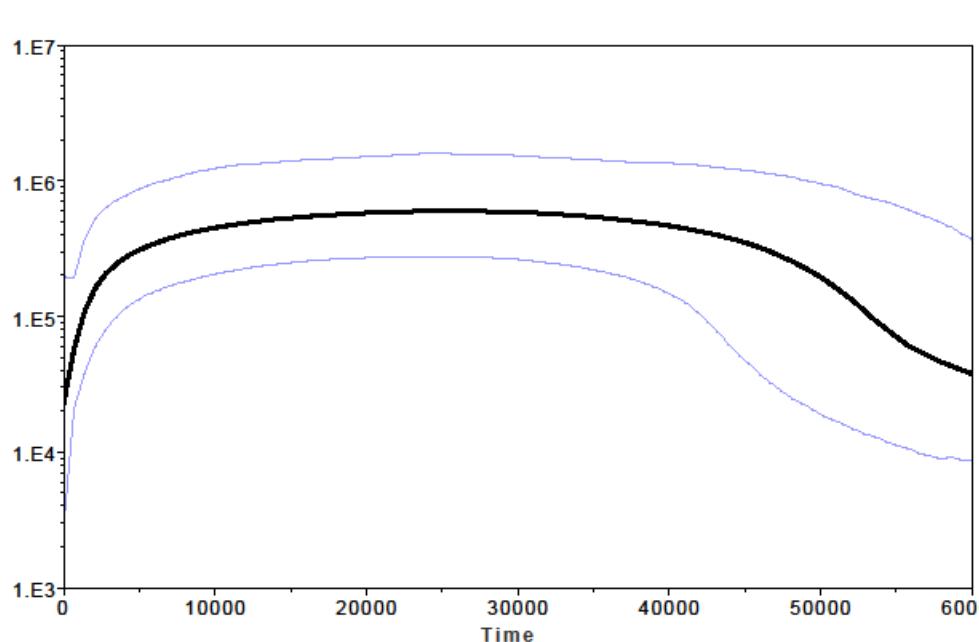
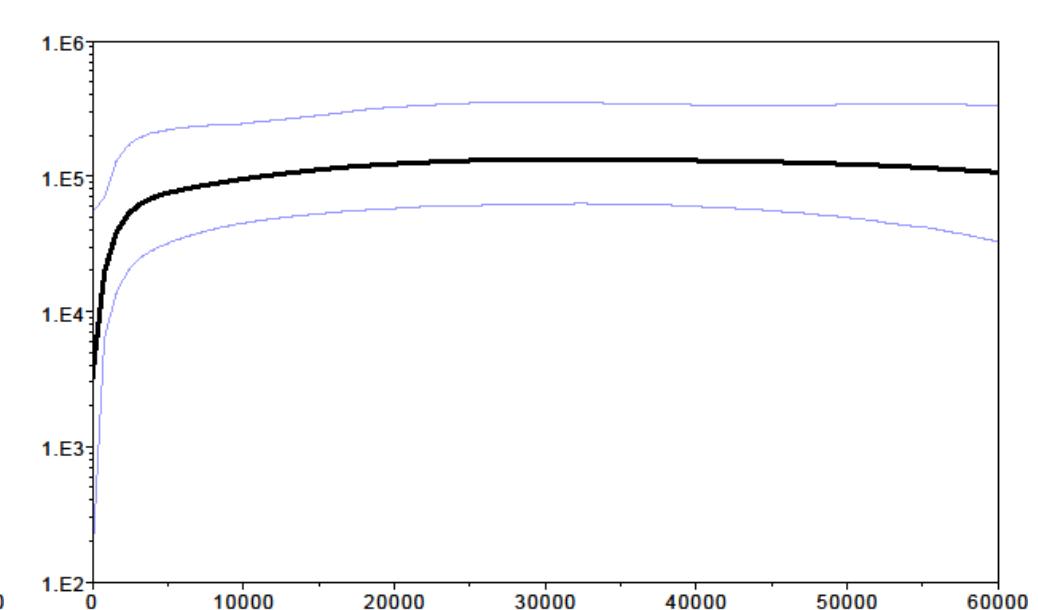
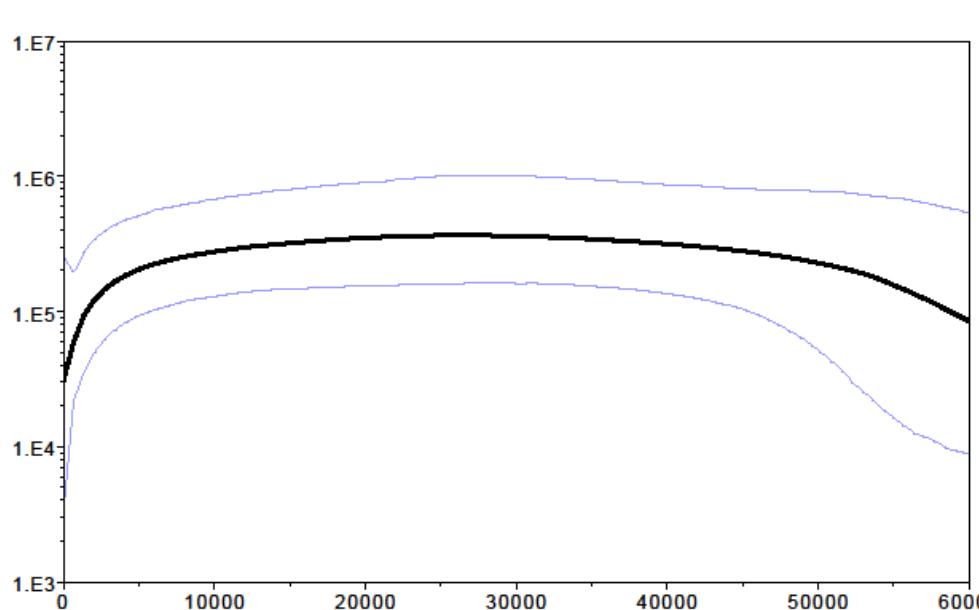
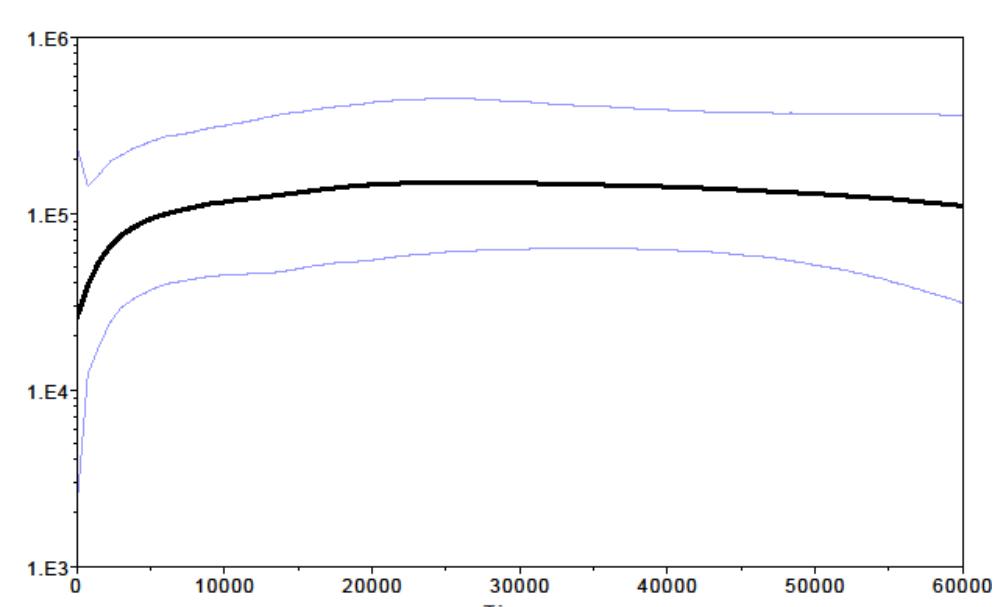
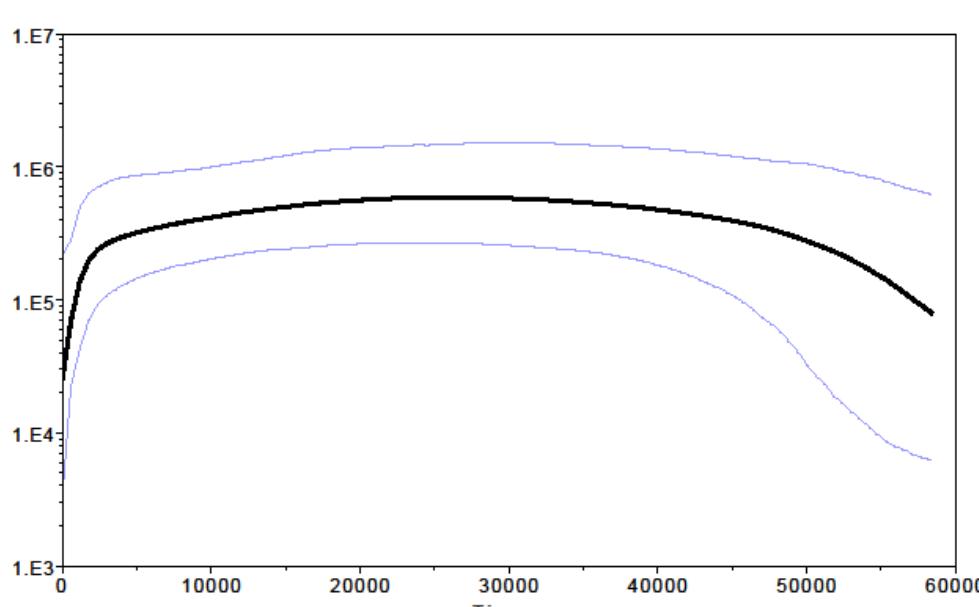
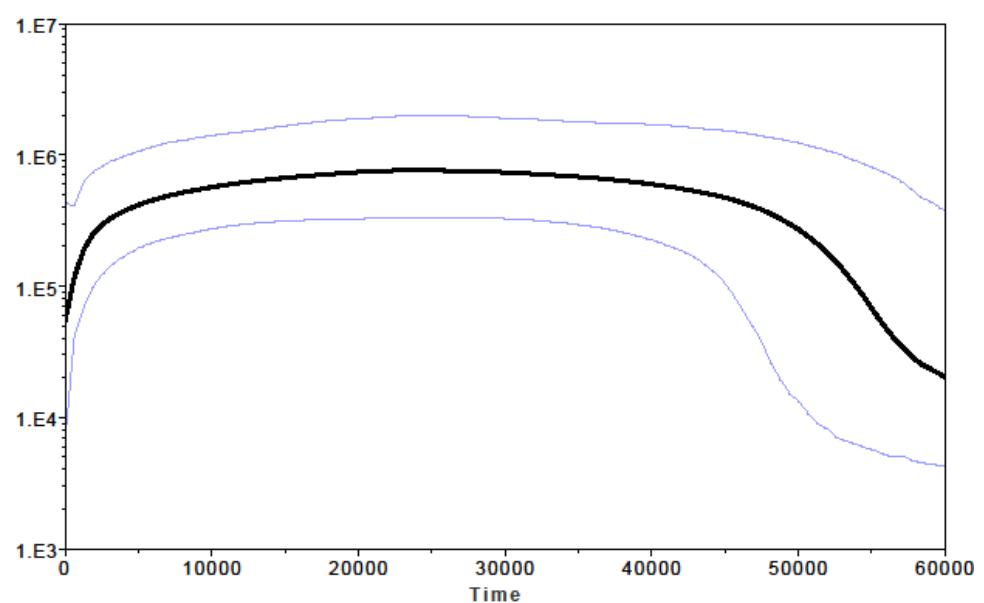
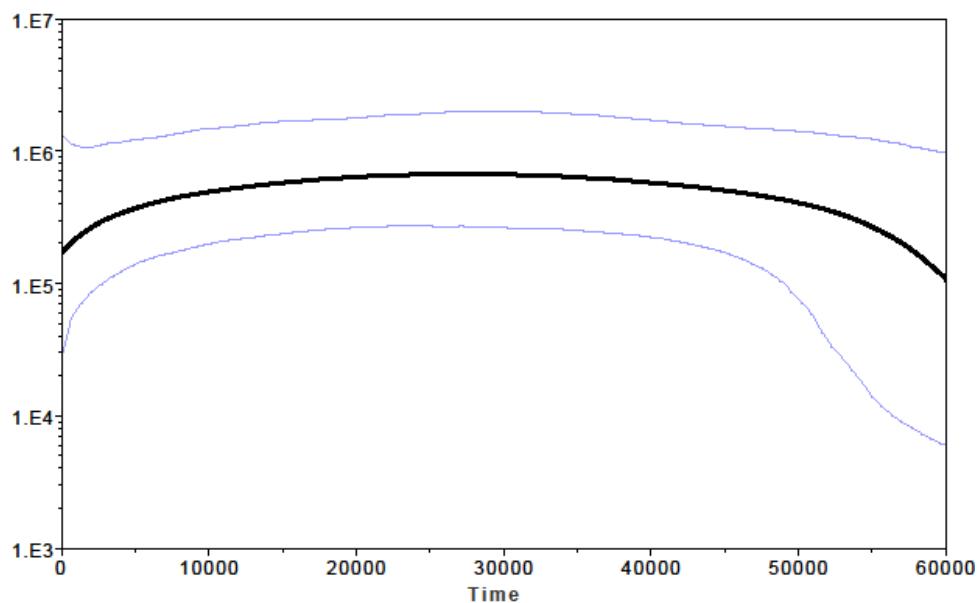
Phuan5 (PU5)

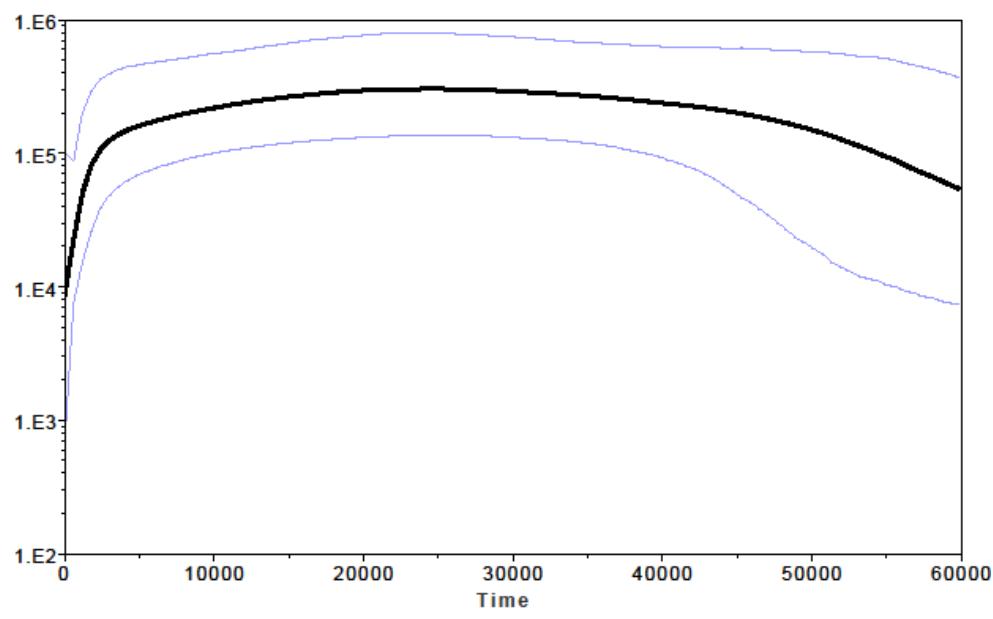


Mon1 (MO1)

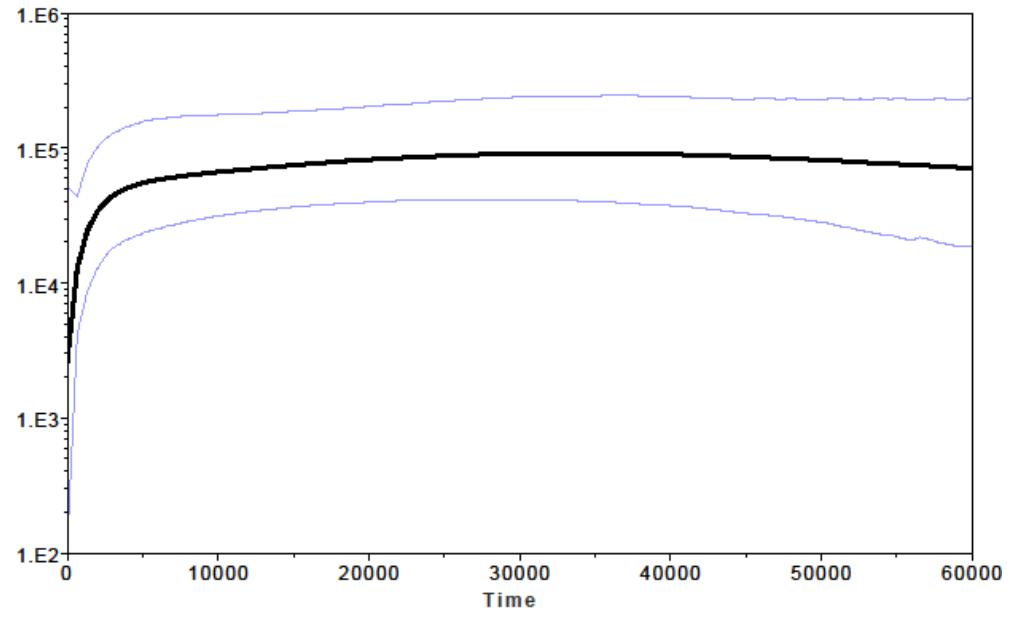


Mon2 (MO2)

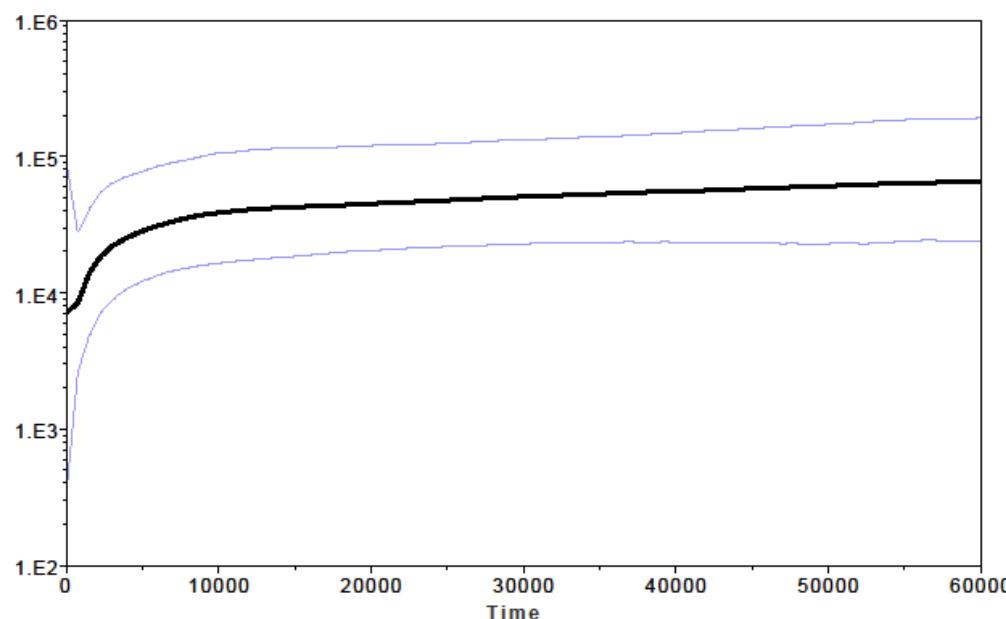




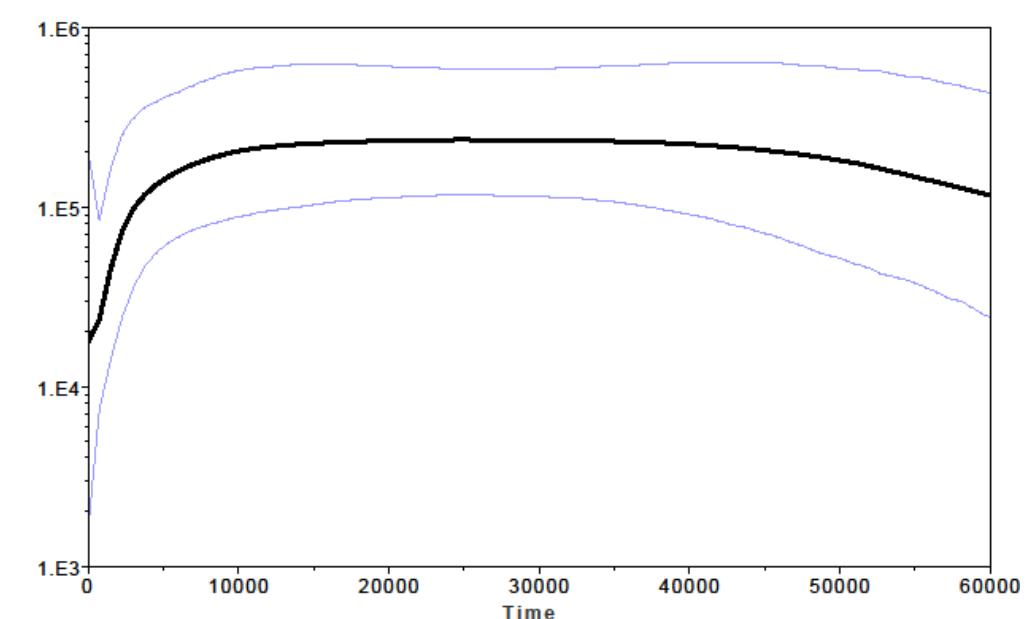
Bru (BU)



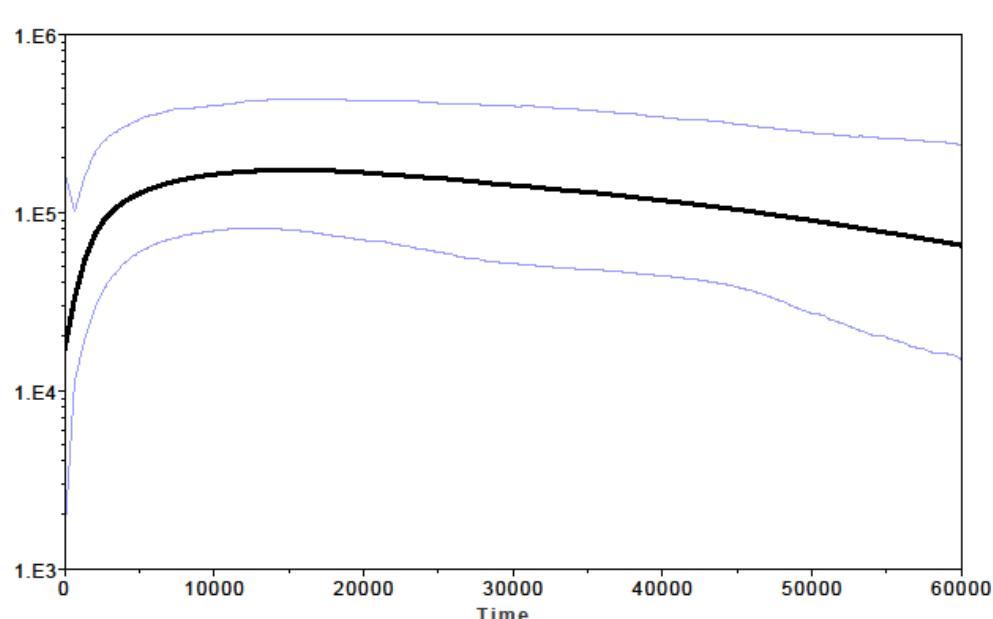
H'tin1 (TN1)



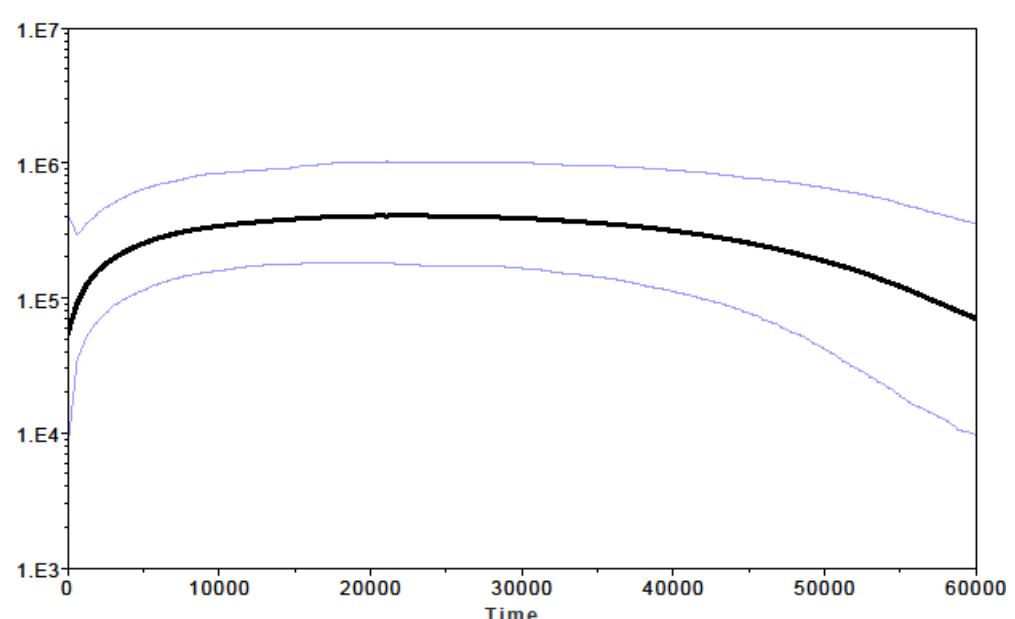
H'tin2 (TN2)



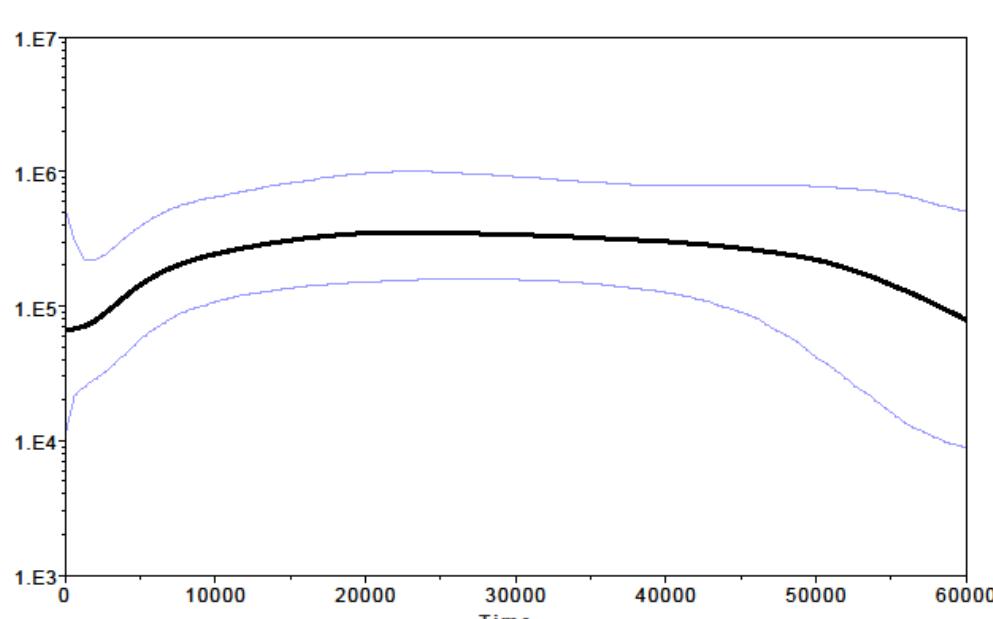
H'tin3 (TN3)



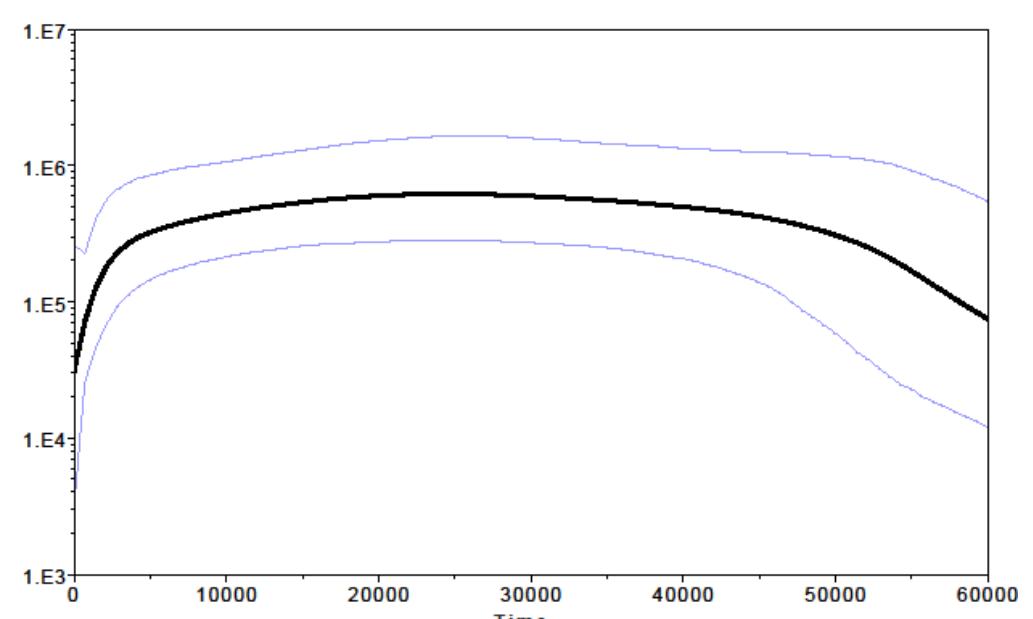
Khamu (KA)



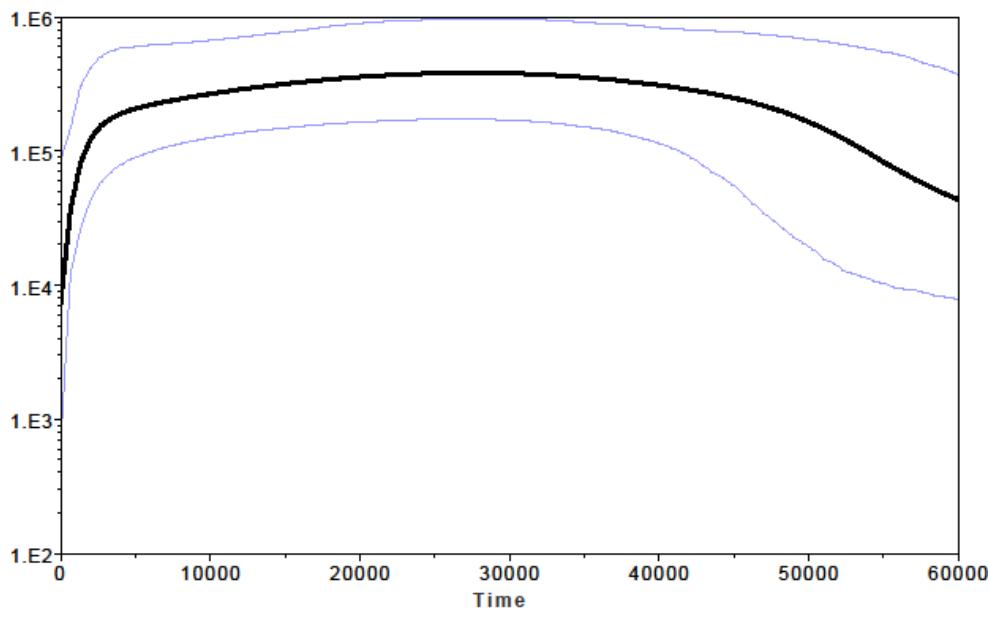
Blang1 (BL1)



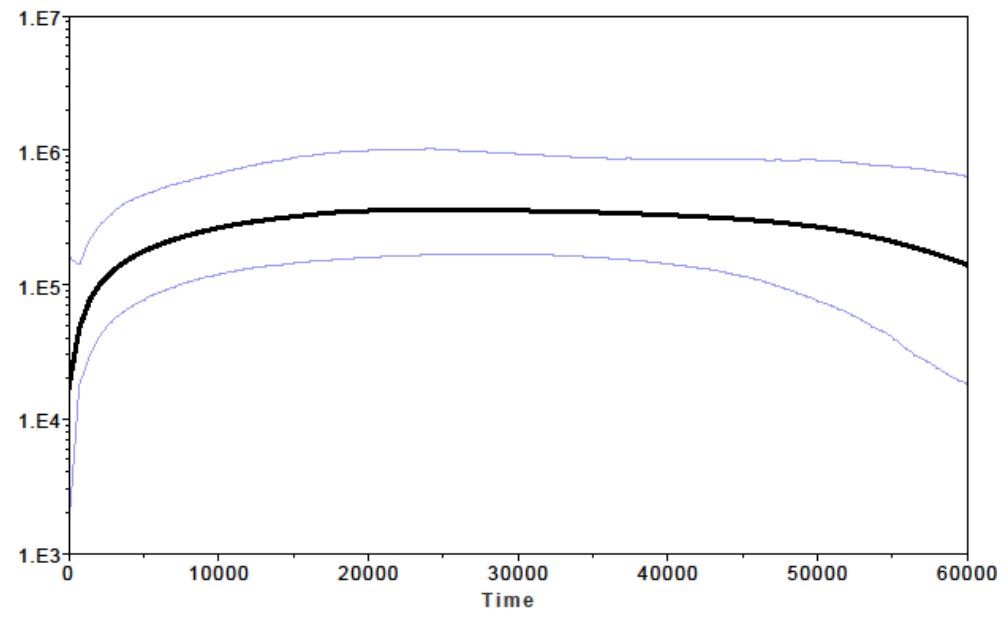
Blang2 (BL2)



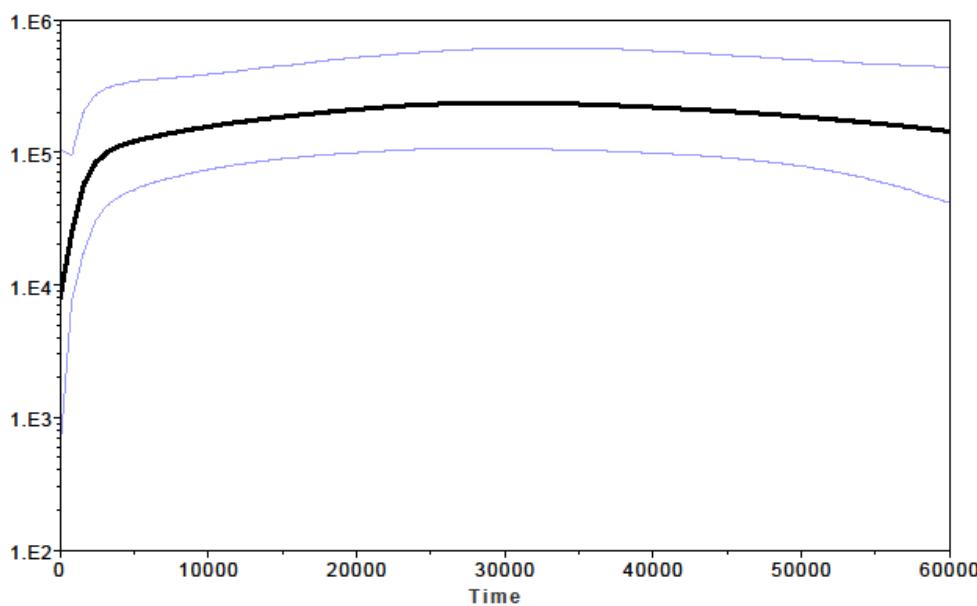
Paluang (PL)



Lawa1 (LW1)



Lawa2 (LW2)

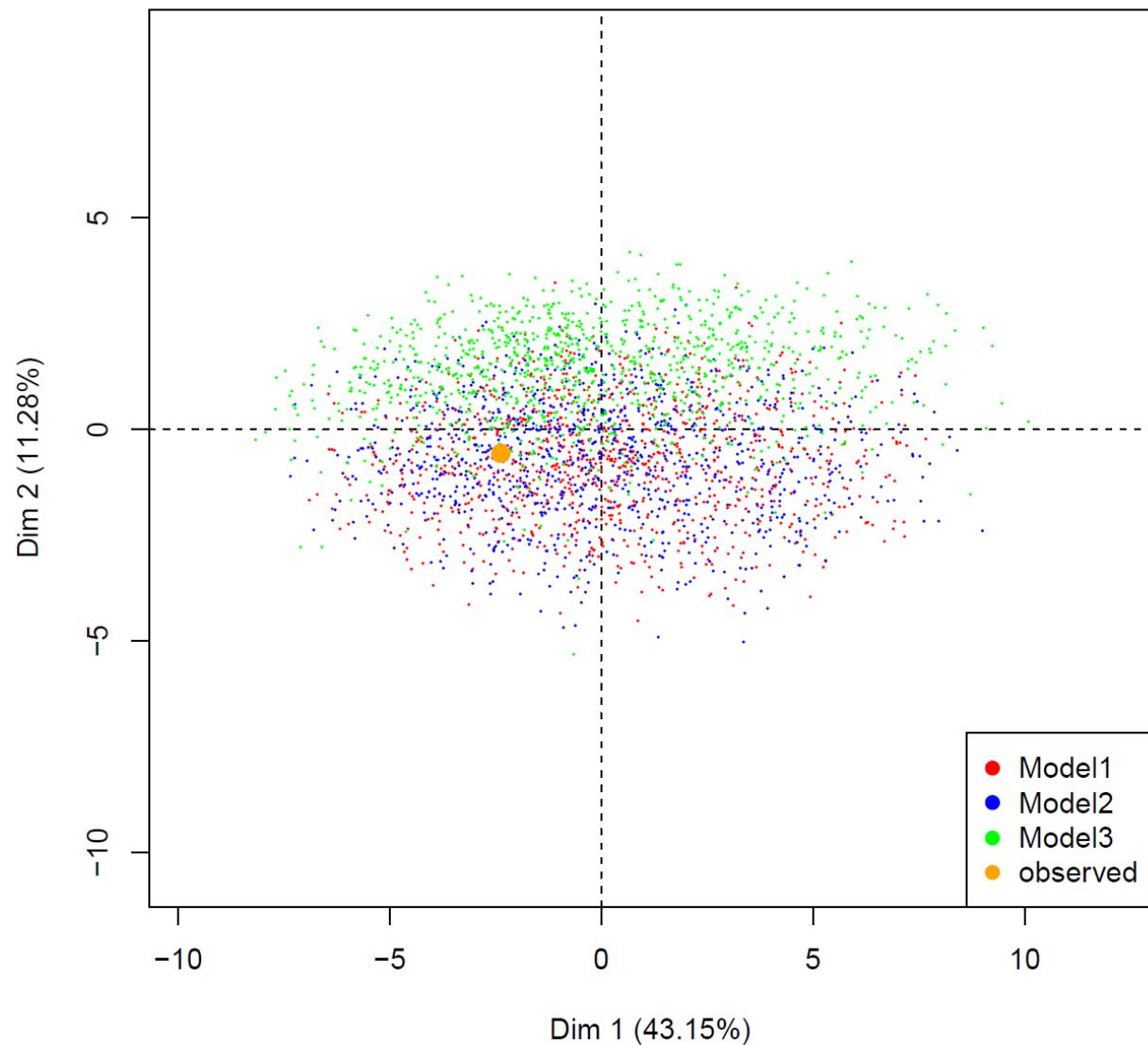


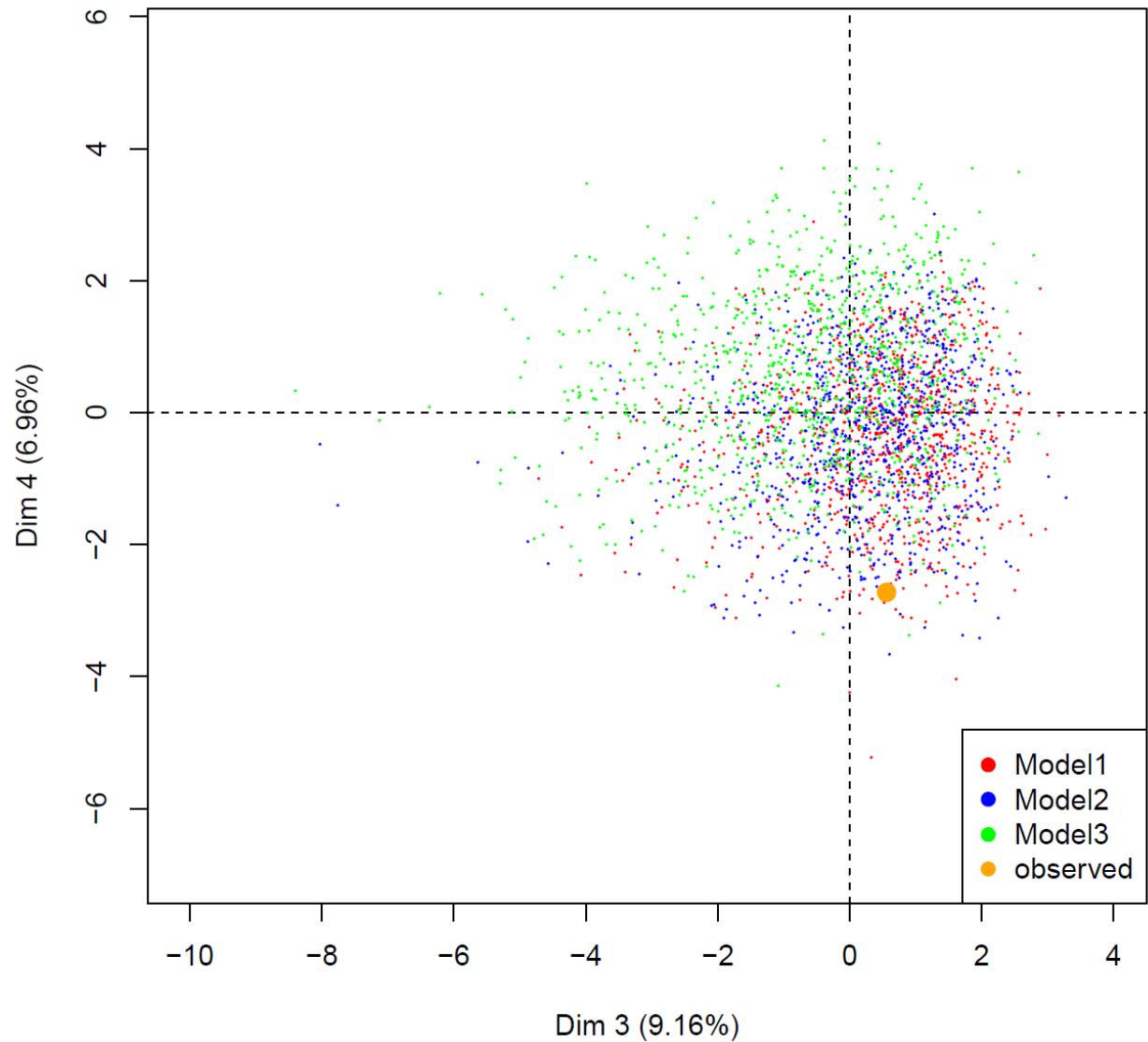
Lawa3 (LW3)

**Fig. S6** Bayesian skyline plots show the effective population size change over time in each studied population. Details concerning the settings are provided in the Methods.

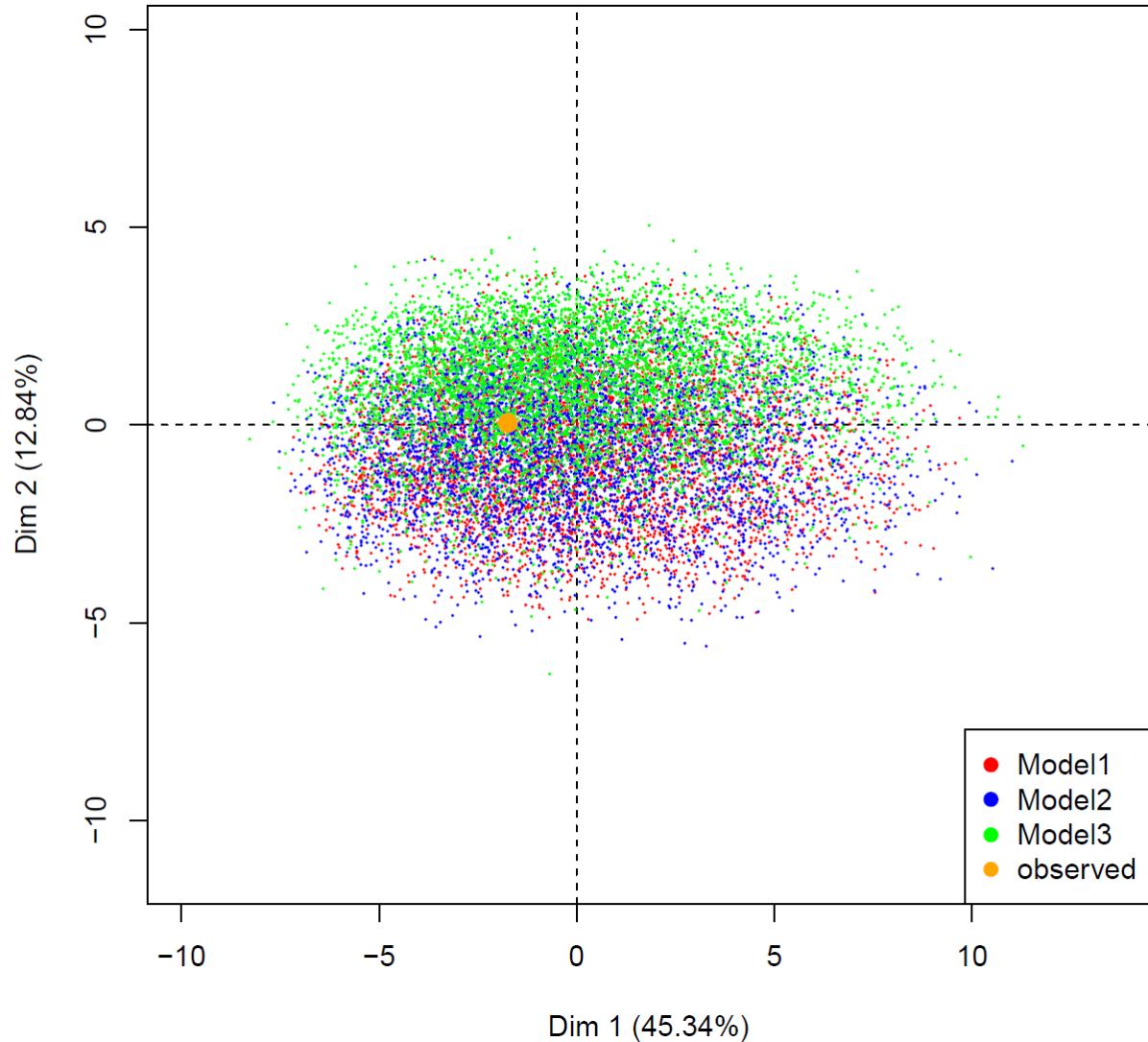
### Test1 (Northern Thai people)

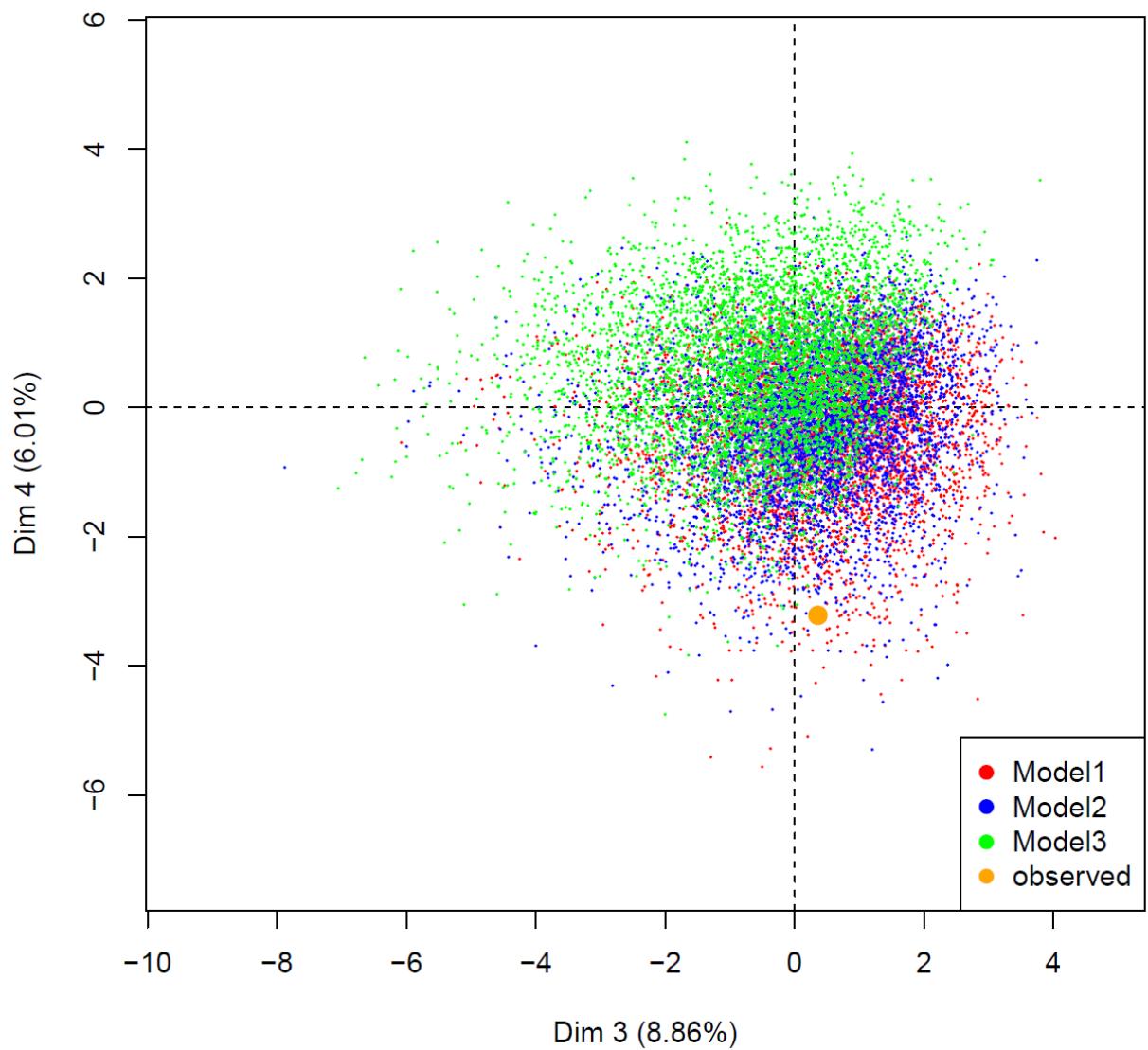
1,000 best simulations





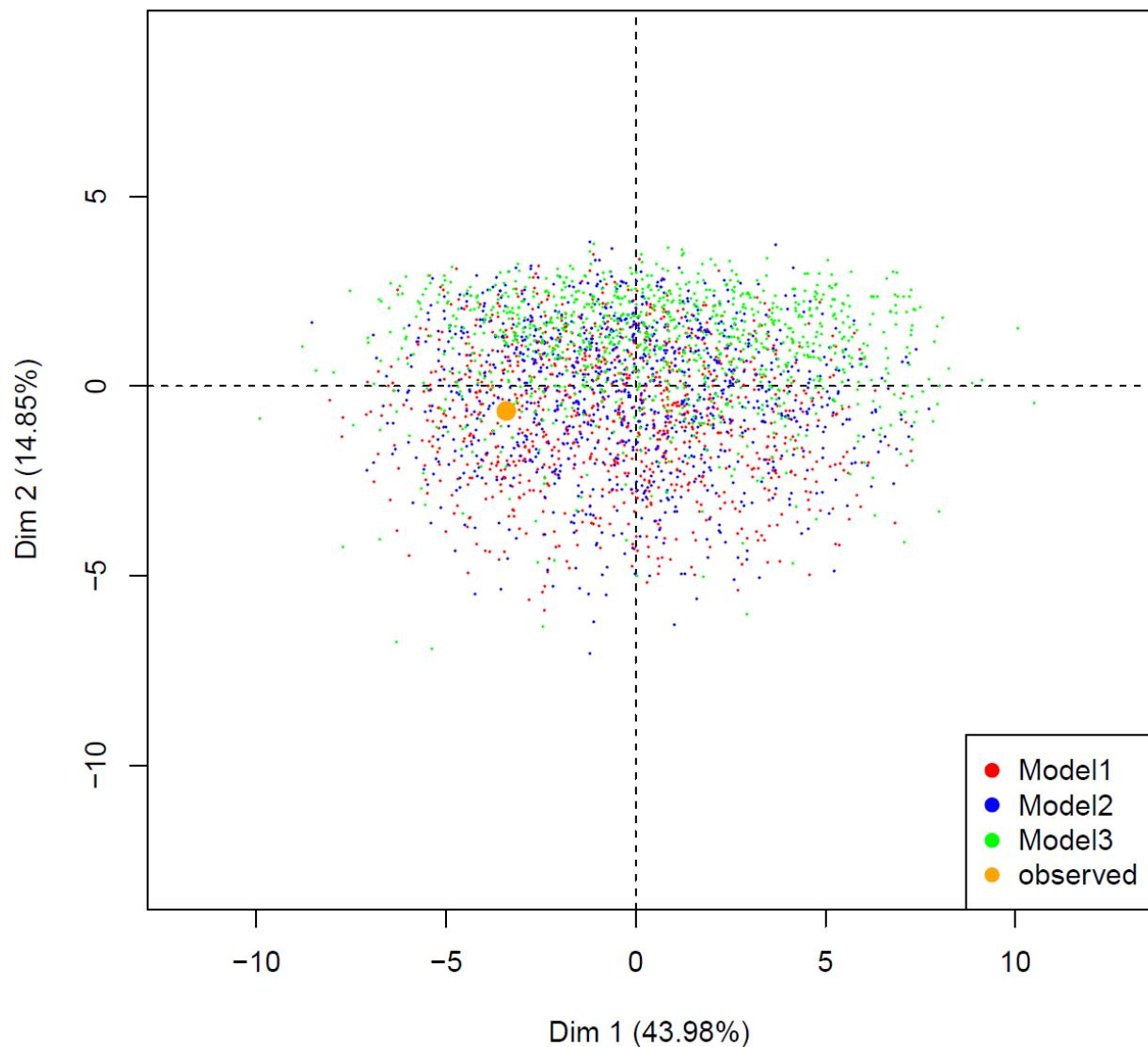
5,000 best simulations

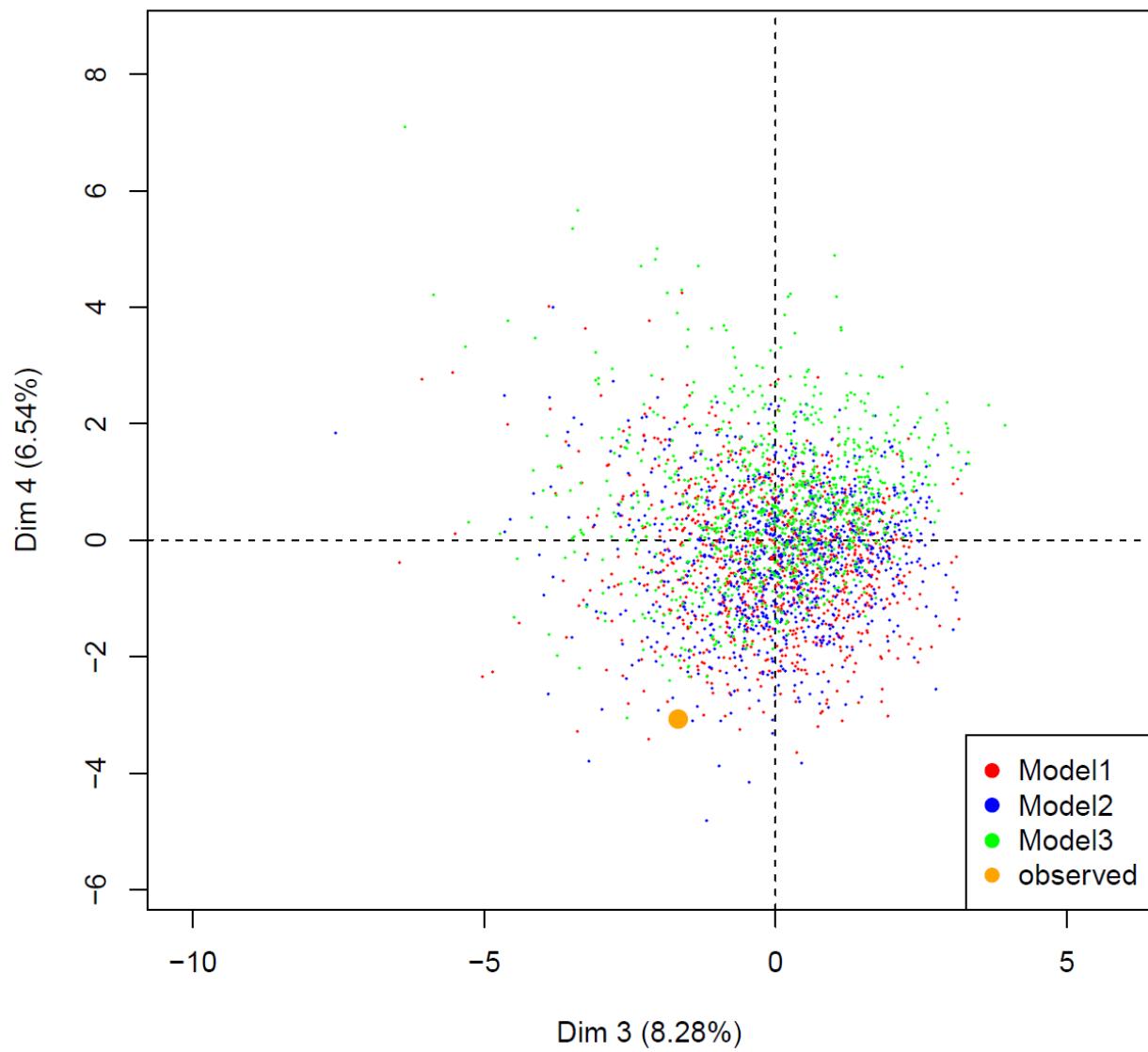




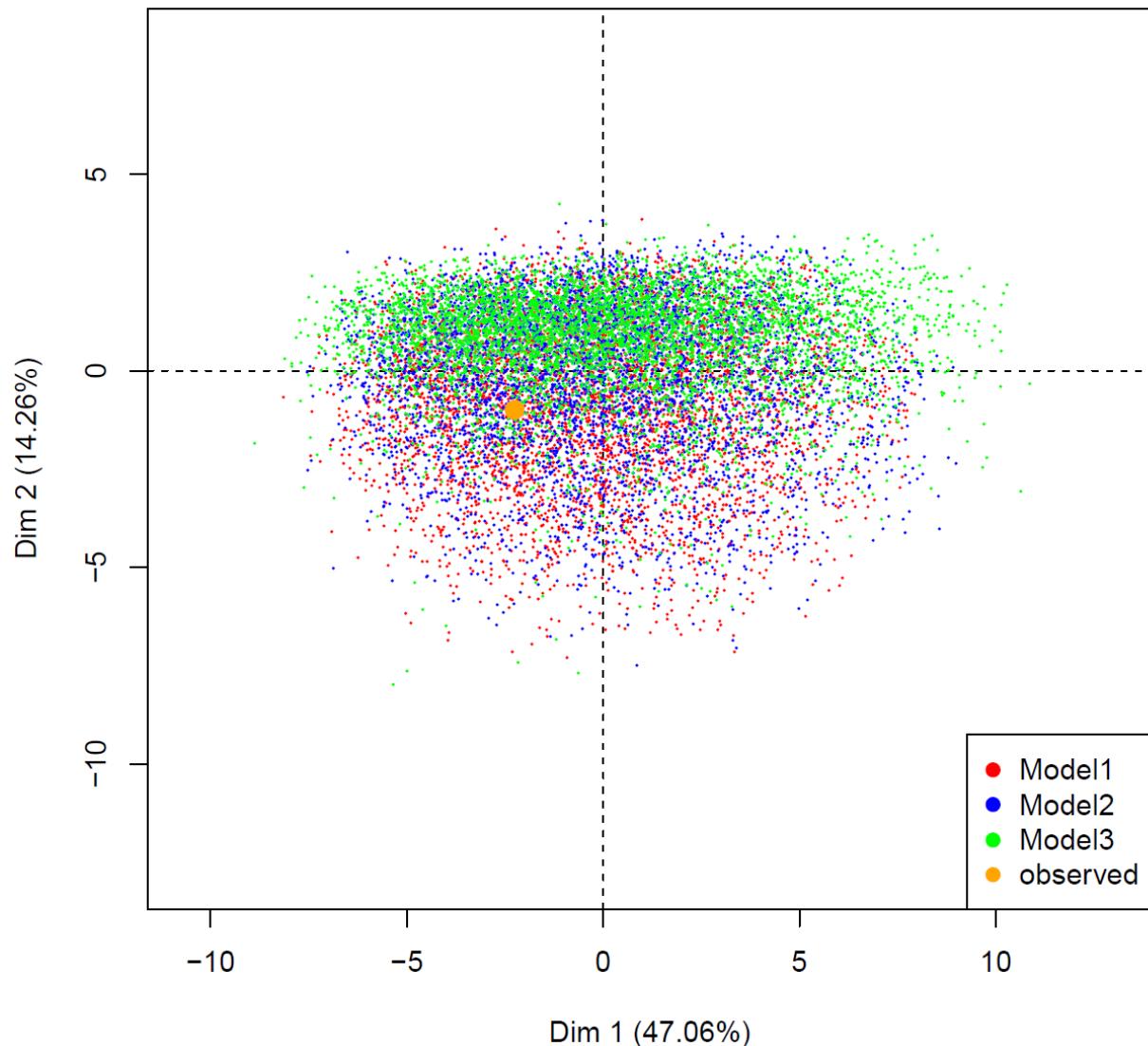
## Test 2 (Northeastern Thai and Lao people)

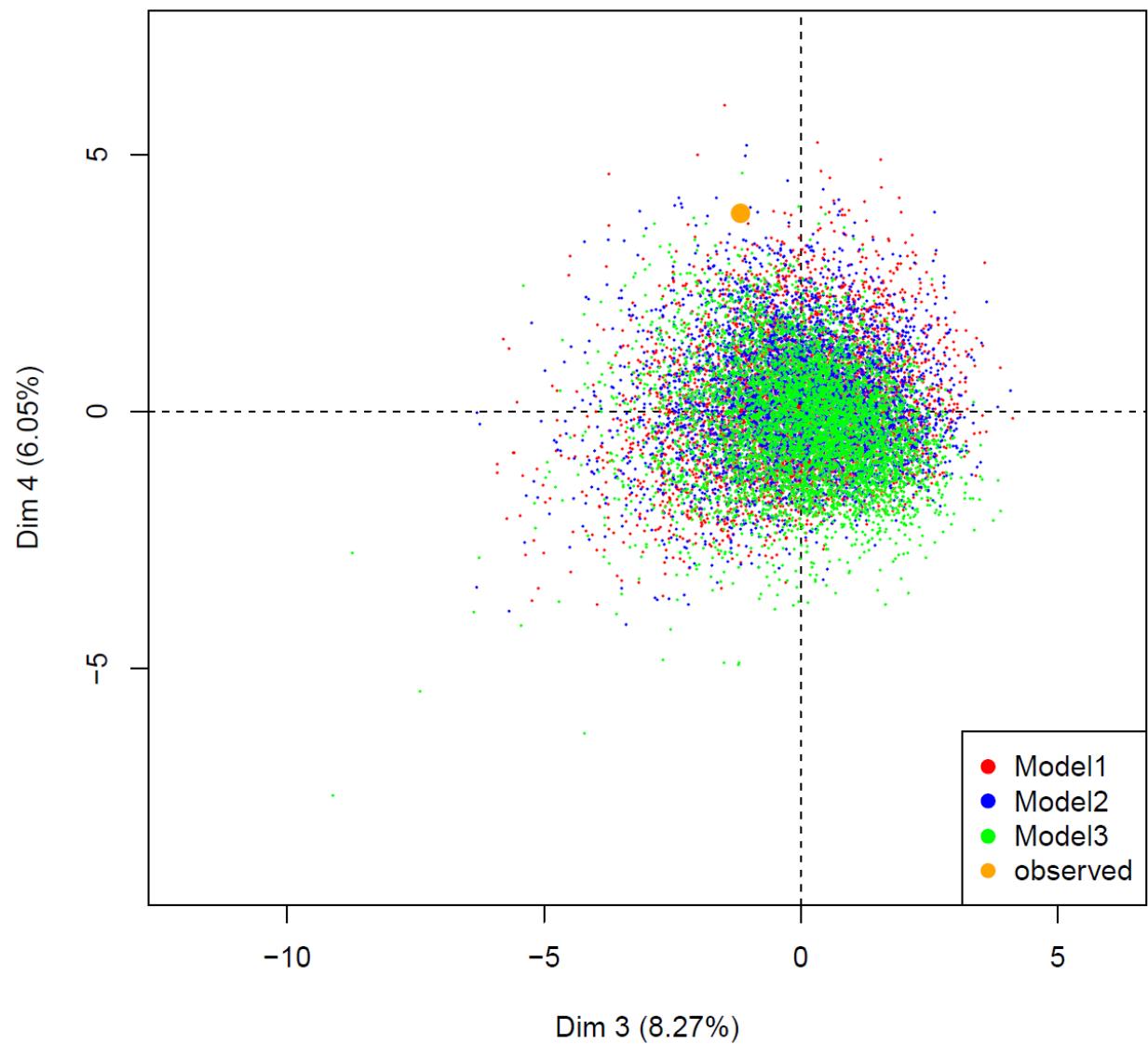
1,000 best simulations





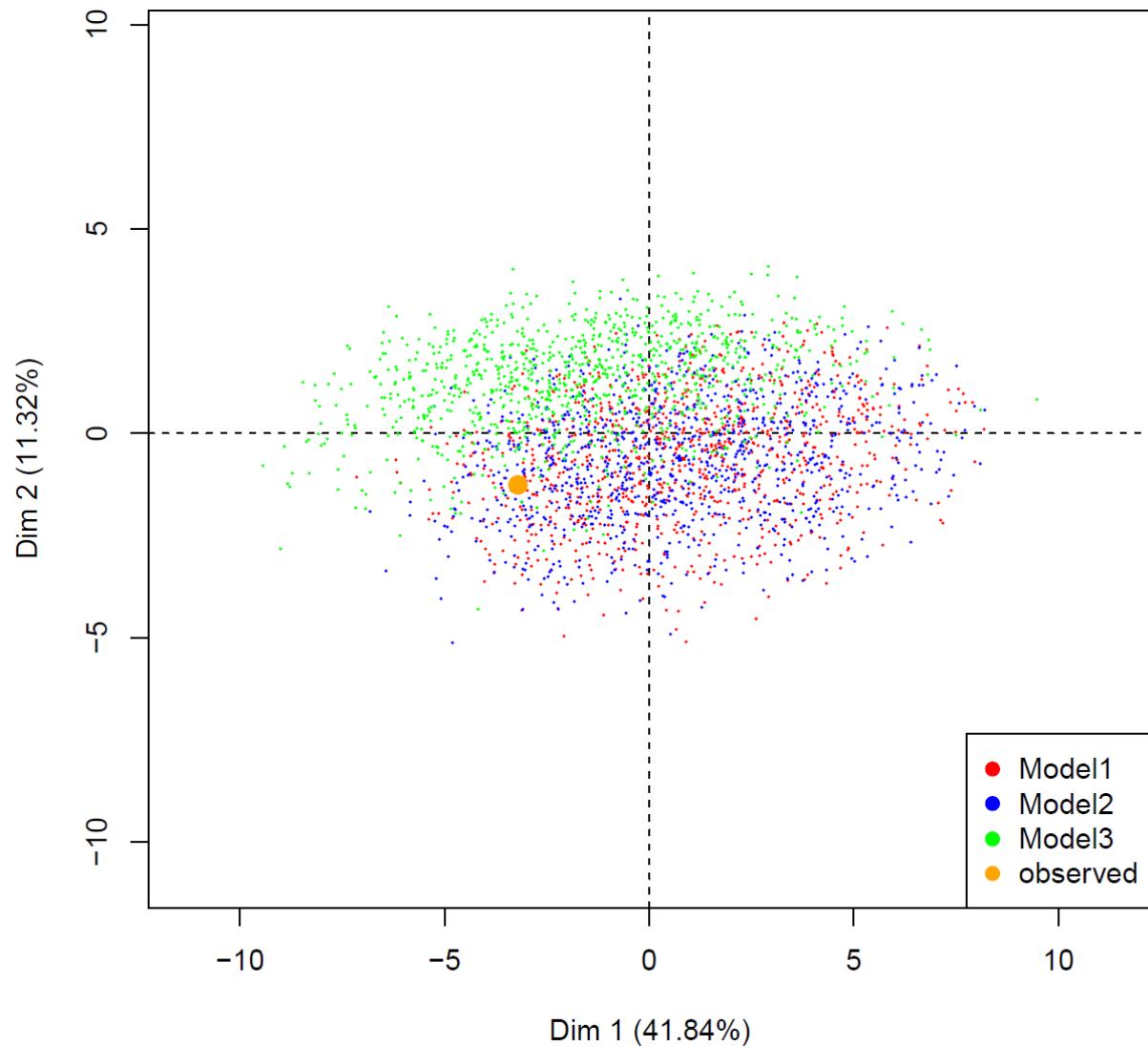
5,000 best simulations

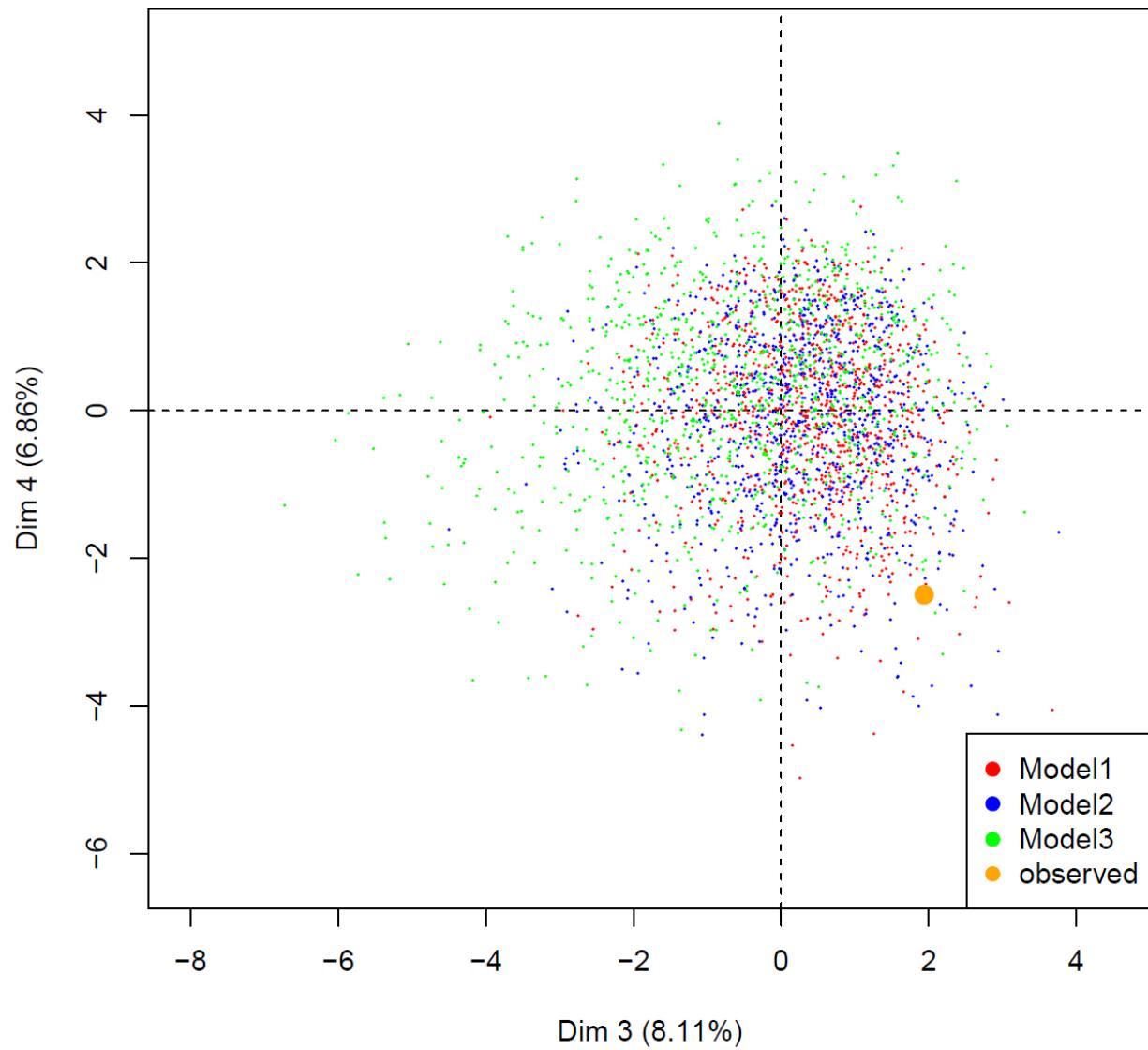




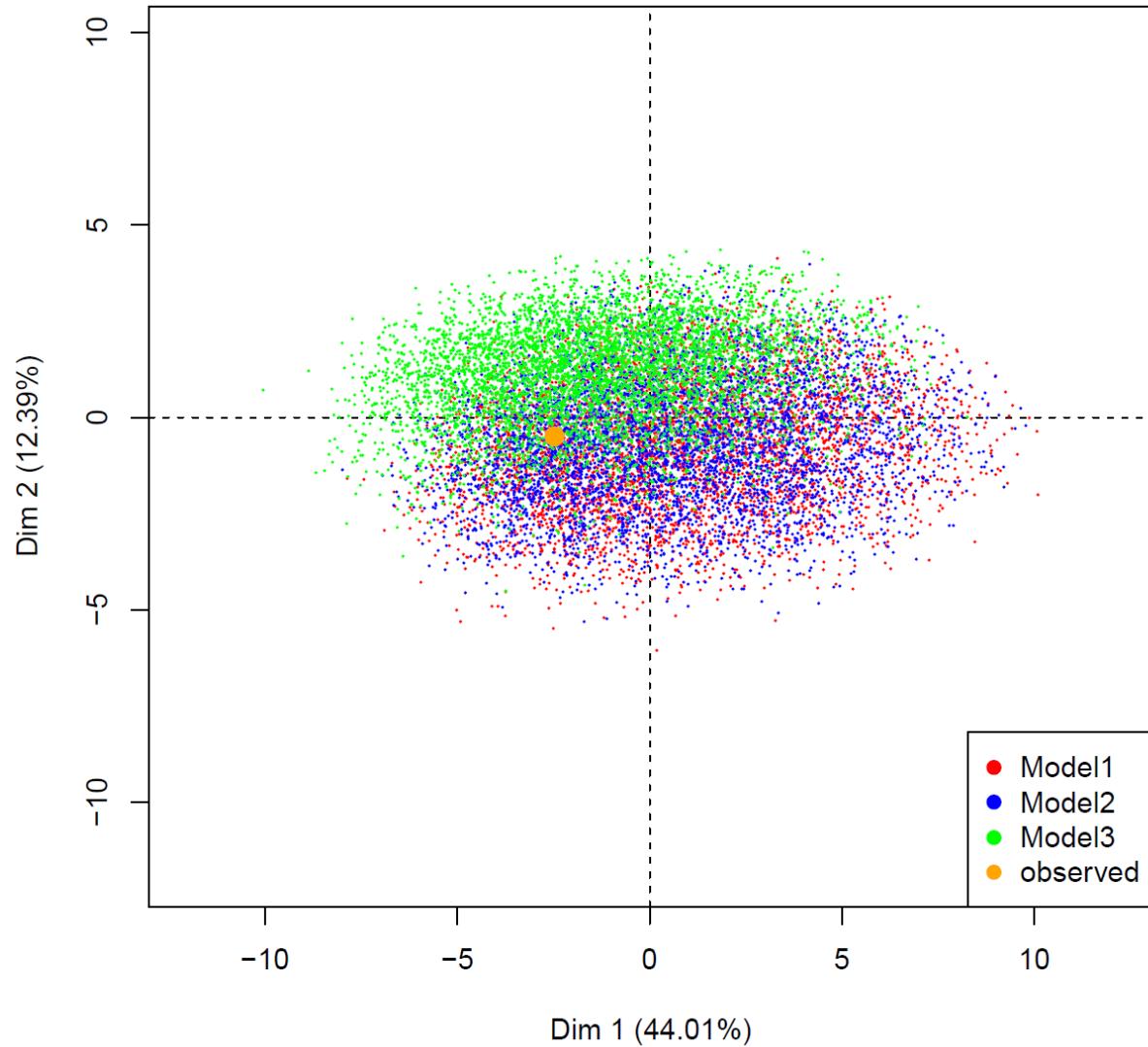
### Test 3 (Northeastern Thai people)

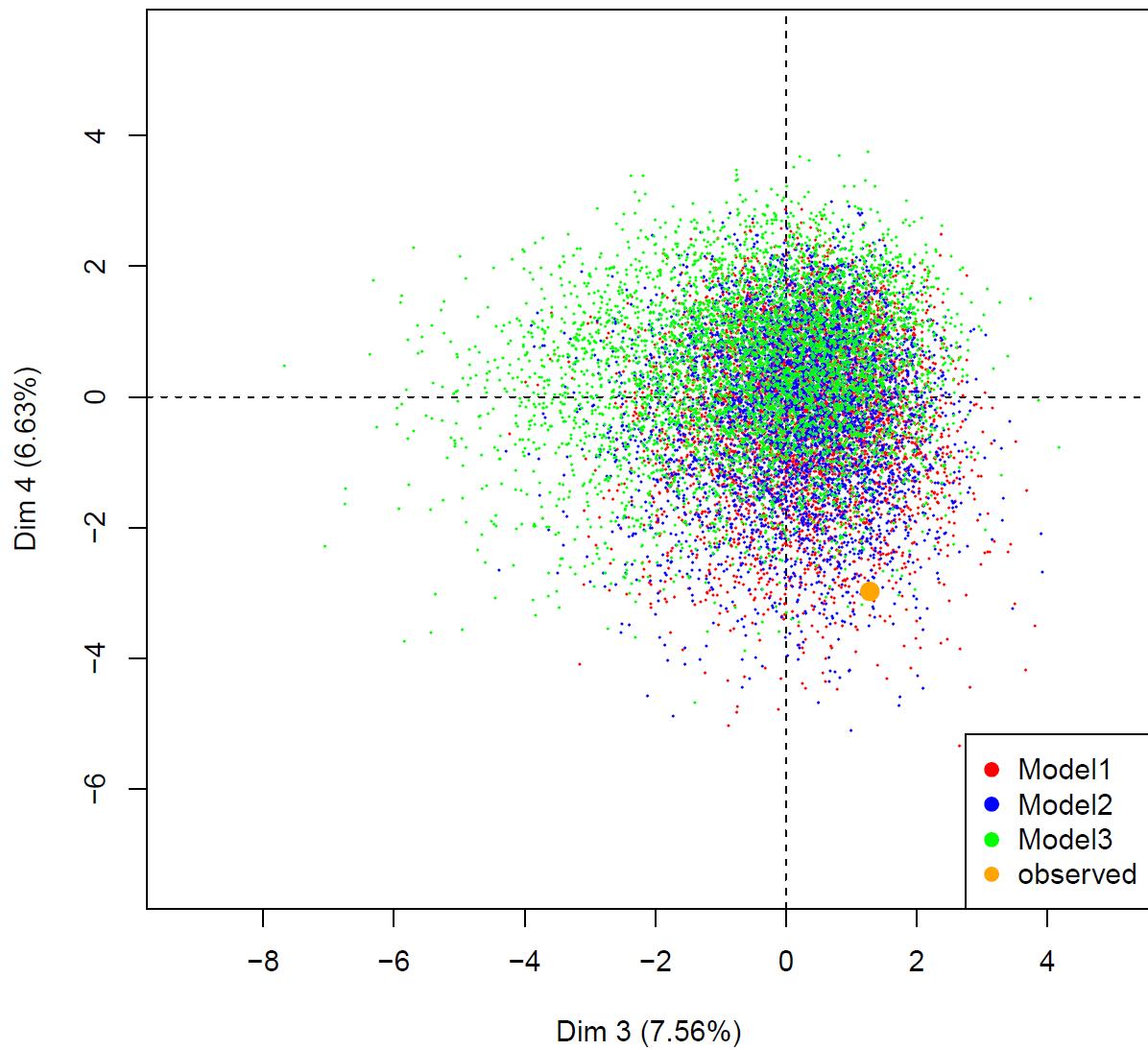
1,000 best simulations





5,000 best simulations





**Fig. S7** Principal Component Analysis (PCA) plots reveal the fit between the simulated and observed data generated by the best 1,000 (or 5,000) simulations for each model. Model 1, Model 2, and Model 3 correspond to models of demic diffusion, admixture and cultural diffusion, respectively.