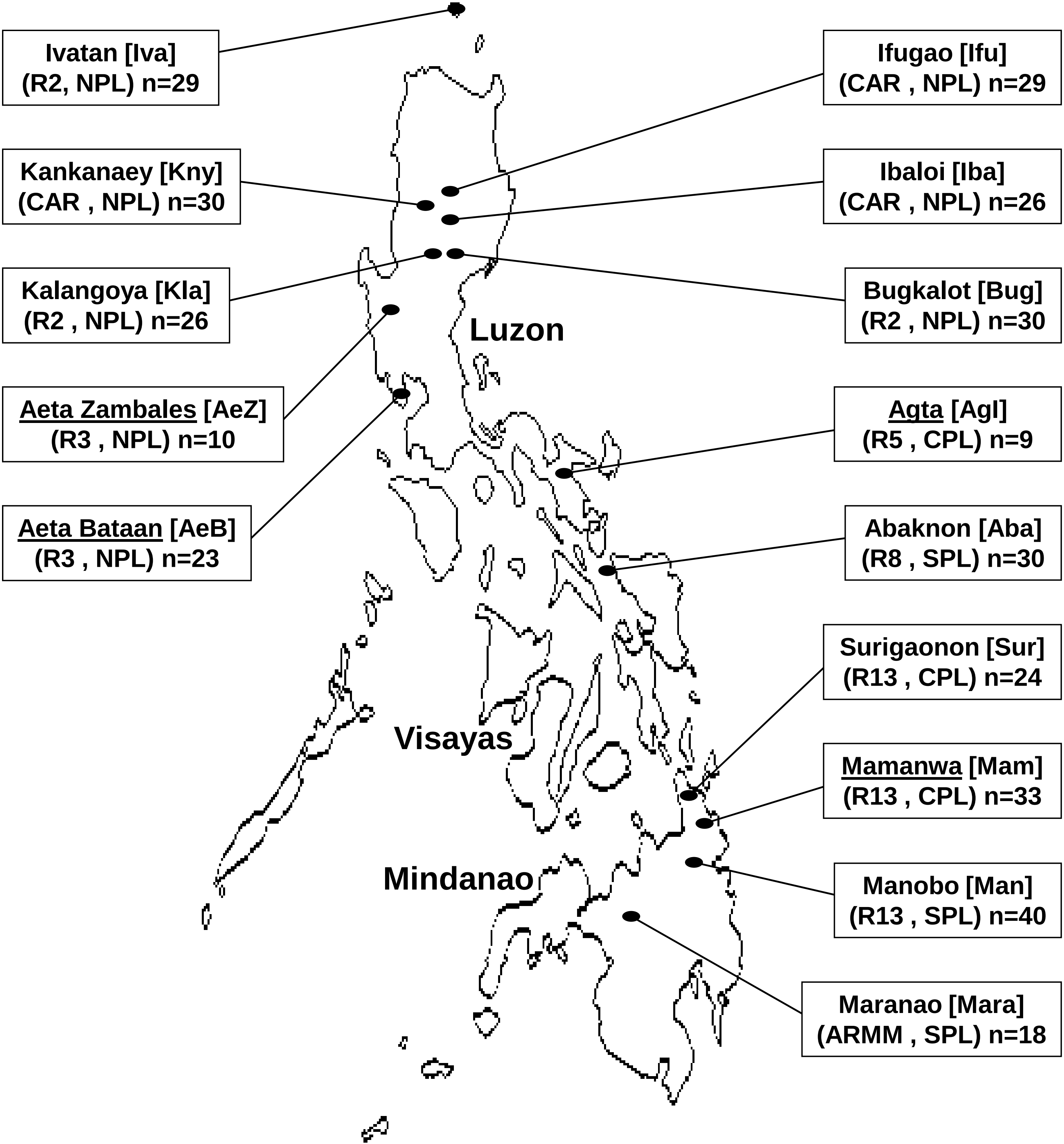


Supplementary Figure 1.

Supplementary Figure 1. Map of the Philippines showing geographical locations of FE groups.

The three main island groups of Luzon, Visayas and Mindanao are indicated. Underlined FE group names refer to FEN groups. Population codes are in square brackets. In parenthesis are geo-political regional groupings (ARMM: Autonomous Region of Muslim Mindanao; CAR: Cordillera Administrative Region; R2-R13: Regions 2-13) and general Philippine language classifications (NPL: Northern Philippine languages; CPL: Central Philippine languages; SPL: Southern Philippine languages) (Lewis 2009). Population sample sizes (‘n’) are indicated. For general geographical groupings: Ivatan to Aeta Bataan are Northern groups; Agta and Abaknon are central groups; Surigaonon to Maranao are southern groups. Other relevant details in Supplementary Table 1.

Supplementary Figure 2.

Supplementary Figure 2. Sequence coverage for the 260 FE group mtDNA genomes in this study.

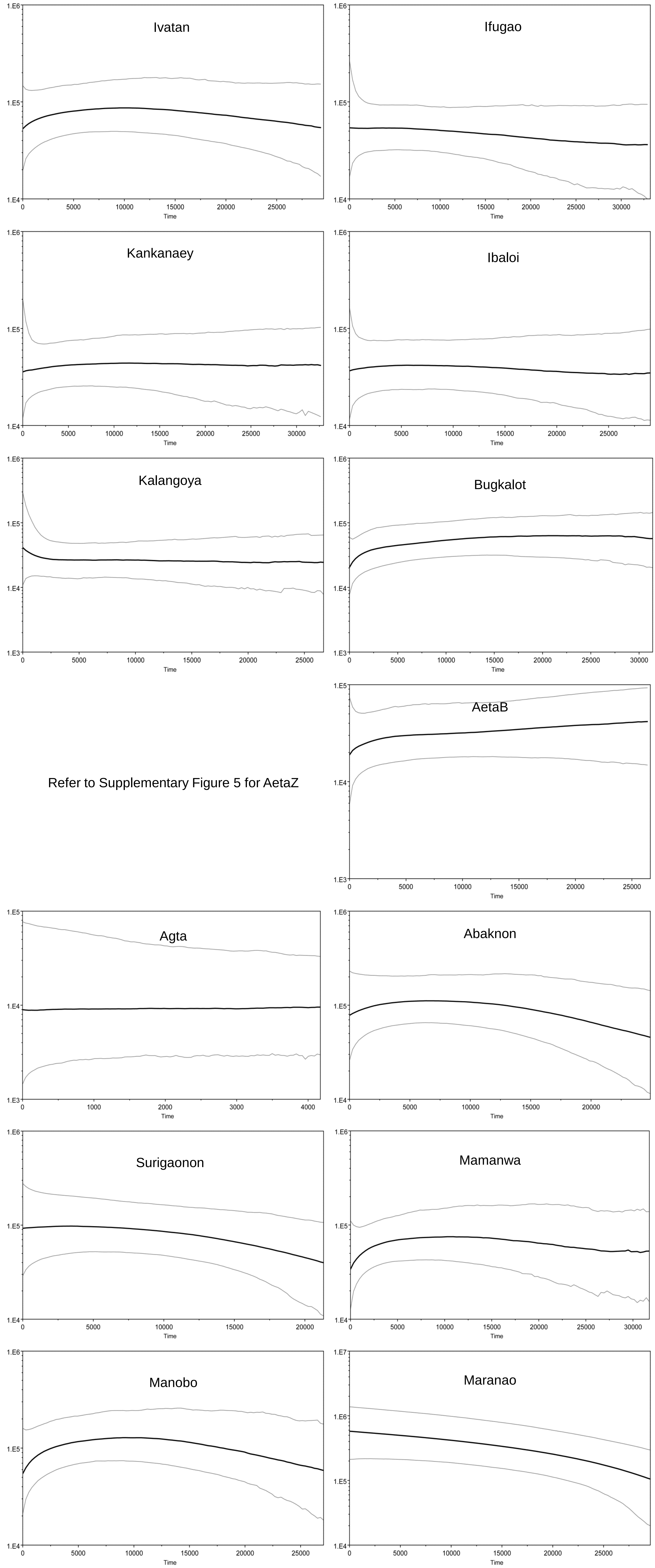
Black solid line indicates average coverage while gray dotted lines refer to minimum and maximum coverage. FE group samples are ordered by year of collection, indicated after group name. Other relevant details in Supplementary Table 2.

Ht#	Northern								Central		Southern			
	Ivatan	Ifugao	Kankanaey	Ibaloi	Kalangoya	Bugkalot	AetaZ	AetaB	AgtaI	Abaknon	Surigaonon	Mamanwa	Manobo	Maranao
1			1	1										
2		2	3	1										
3				1	3									
4				2	1									
5		3		1										
6		1		1	1									
7		1	4											
8		1	1											
9		2	1											
10											1	1		
11												1	1	
12											1	1		
13											3			1
14	1									3				
15				2	1	3				2				
16	2									2				
17			2	1	1					1	1	1		
18		2		1	2					1				
19	2					6					1			
20						1				5	1	5	2	1
21	1	4	4	5	4	7	10		6	1		1	5	
22								7	1	1				

Supplementary Figure 3.

Supplementary Figure 3. Haplotype sharing patterns among FE groups.

Haplotypes are based on complete mtDNA sequence data. Ht# refers to random arbitrary haplotype designations. FE group names are indicated and arranged from left to right as northern, central and southern geographical groups. Filled blocks indicate presence of the shared haplotype. Number within filled blocks indicates the number of individuals that possess the shared haplotype.

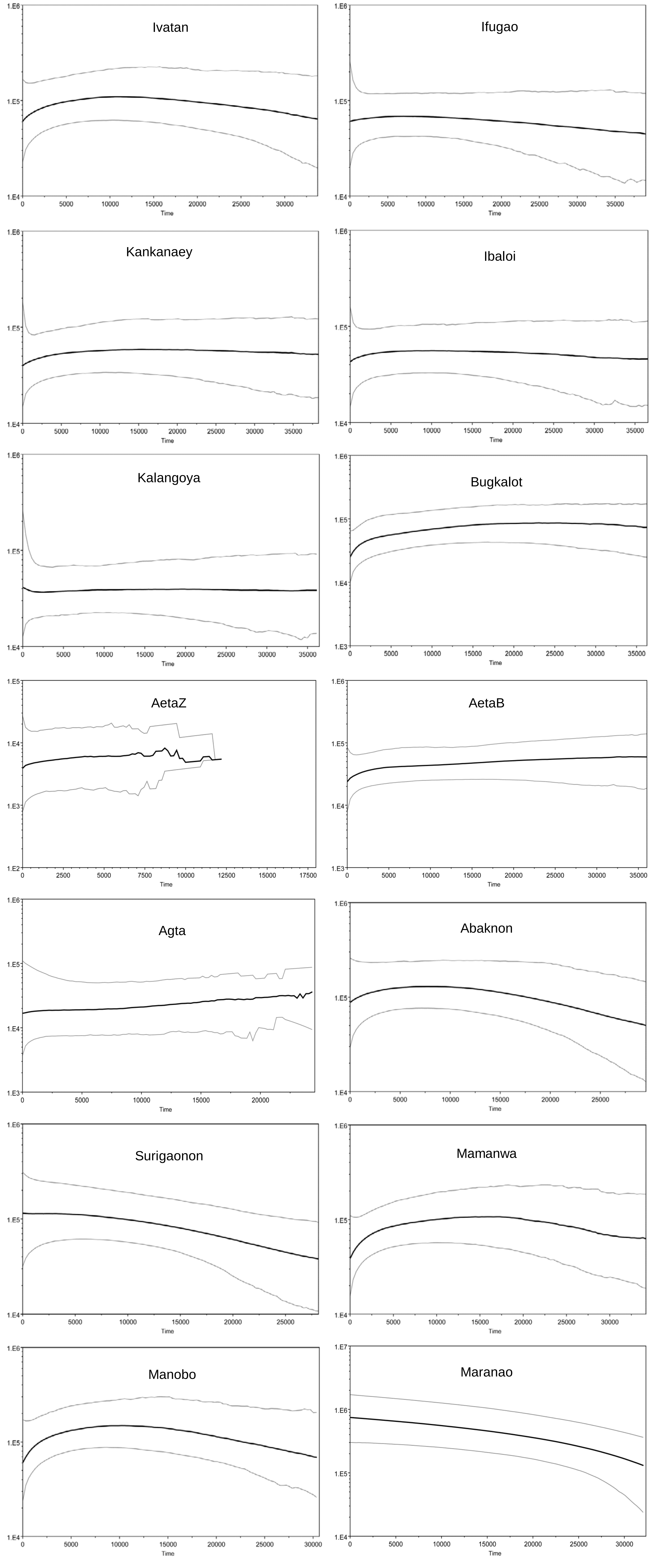


Refer to Supplementary Figure 5 for AetaZ

Supplementary Figure 4.

Supplementary Figure 4. Bayesian Skyline Plots (BSP) for FE groups.

These BSPs are the output of analyses that included the RSRS and are comparable to the information in ‘A’ columns in Supplementary Table 5. Each plot is labelled with the FE group name. The product of effective population size (N_e) and time is on the y-axis. In each plot, the light lines are the 95% highest posterior density (HPD) that bound the Bayesian mean estimate represented by a heavy black line.



Supplementary Figure 5.

Supplementary Figure 5. Bayesian Skyline Plots (BSP) for FE groups.

These BSPs are the output of analyses that included the RSRS and are comparable to the information in ‘B’ columns in Supplementary Table 5. Each plot is labelled with the FE group name. The product of effective population size (N_e) and time is on the y-axis. In each plot, the light lines are the 95% highest posterior density (HPD) that bound the Bayesian mean estimate represented by a heavy black line.

Refer to Figure 3 for biplot of Dimensions 1 and 2

Supplementary Figure 6.

Supplementary Figure 6. MDS plots of FE groups and reference data set.

MDS plot of Φ_{ST} genetic distances based on complete mtDNA data of FE groups (this study; Gunnarsdóttir *et al*. 2011a) and a Reference data set. Shown here are biplots of dimensions 2 to 5. The biplot of dimensions 1 and 2 is shown in Figure 3. Stress value is 0.12. FE group names are indicated and are adjacent to shapes. These shapes indicate northern (triangles), central (circles) and southern (squares) geographical locations. Open shapes indicate FEnN groups, while filled shapes indicate FEN groups. Asian and Pacific groups of the reference data set are represented by the following symbols: filled diamonds (South Asia: 28 Indian groups, including Andaman Islanders); asterisks (Southeast Asia: Vietnam, the Besemah and Semende groups of Sumatra, Indonesia and the Bidayuh, Jehai, Seletar and Temuan groups of Malaysia); open diamonds (North East Asia: Han Chinese, Japan and Taiwan); plus signs (Australia, Near Oceania and Remote Oceania). Other relevant details in Supplementary Table 3 and Supplementary Text.

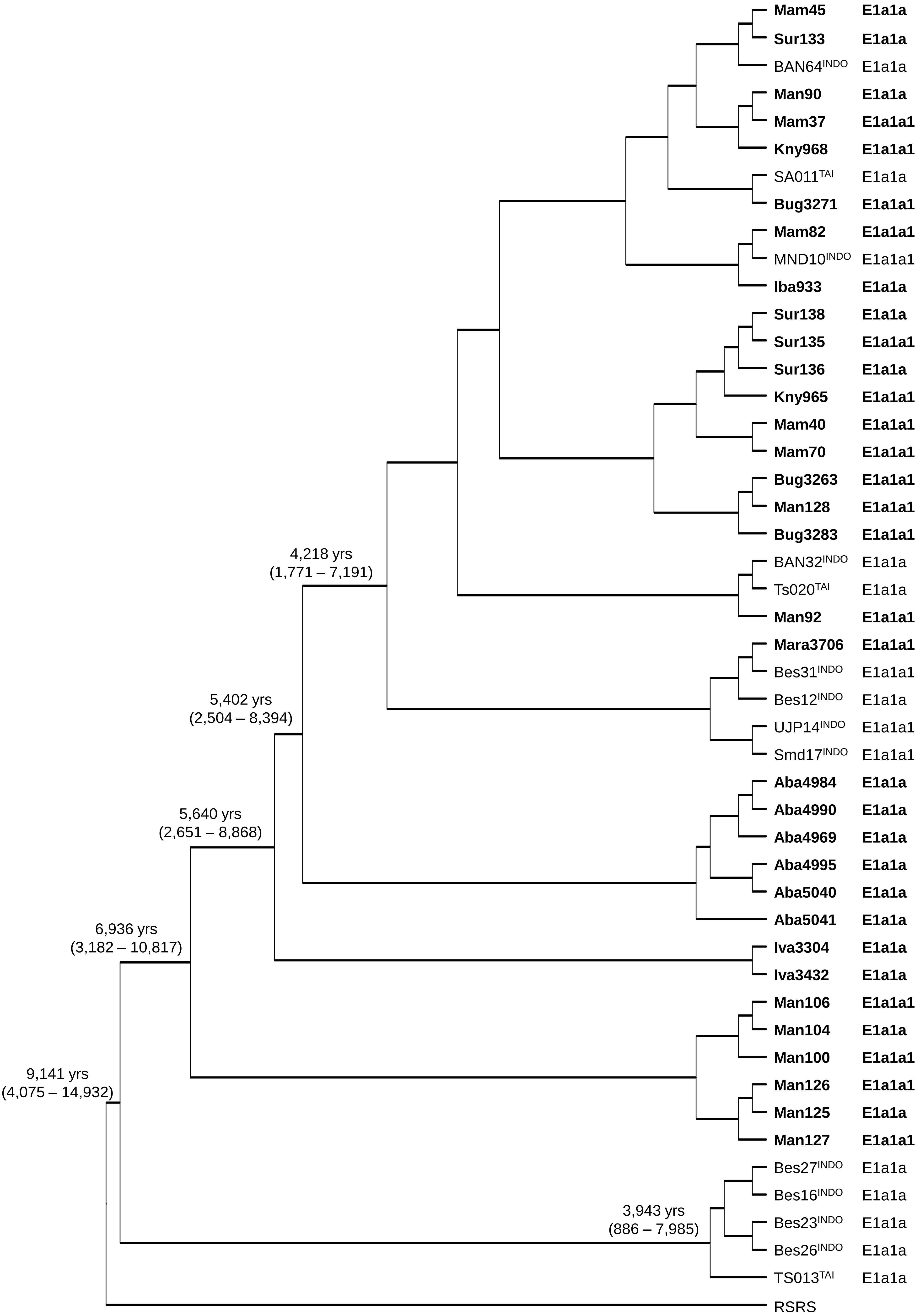
Supplementary Figure 7.

Supplementary Figure 7. CA plot of haplogroup counts in FE groups and reference data set. The CA results here are a part of a series of CA that started with the inclusion of 18 reference data set groups (Supplementary Table 7; Supplementary Figure 7A), with subsequent CA involving removal of outliers (Supplementary Figure 7B-I) and ending with a CA that involved five reference data set groups (Supplementary Table 8; Figure 4). FE group and reference data set group names are indicated in bold print. Haplogroup names are in *narrow italicized* font. Other relevant details in Supplementary Table 3 and Supplementary Text.

Supplementary Figure 8.

Supplementary Figure 8. Maximum clade credibility tree for haplogroup B4a1a.

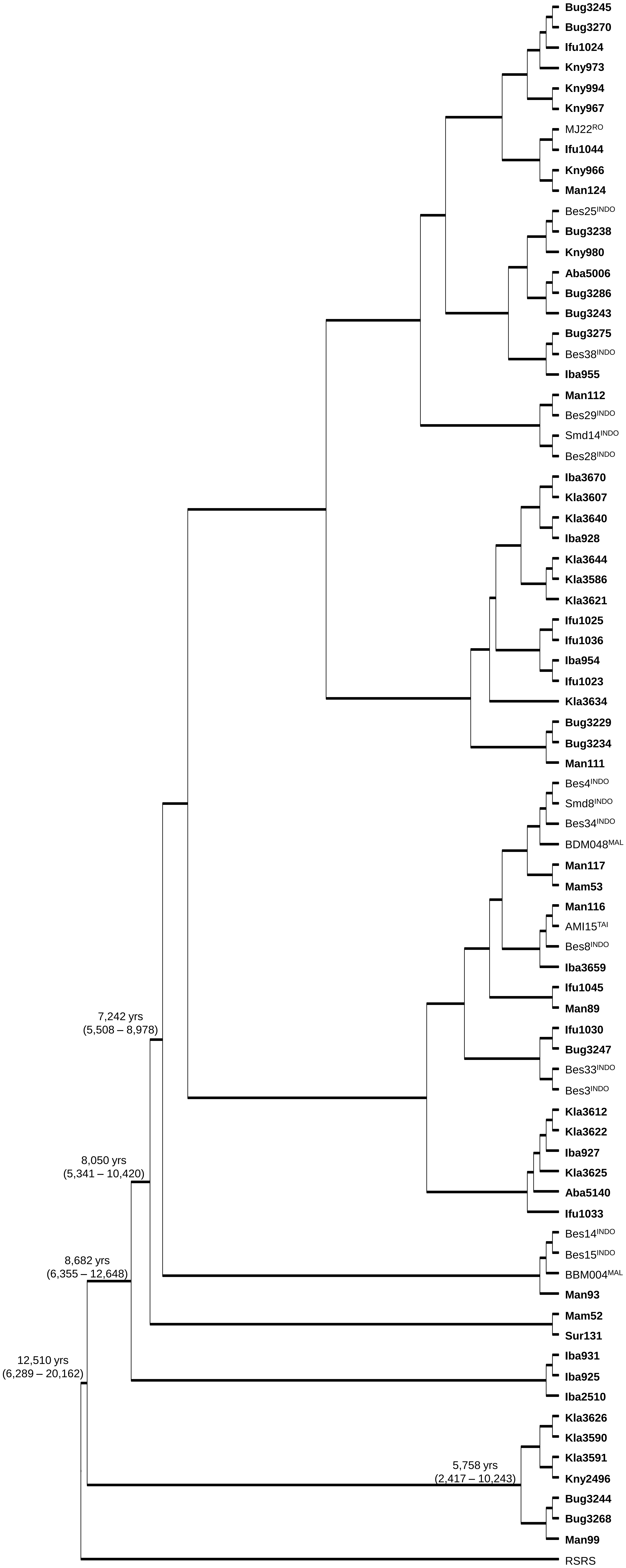
Tree topology and coalescent time estimates with 95% highest posterior density (HPD) are results of BEAST (Drummond and Rambaut 2007) analyses. Sample names (RSRS, Reference data set, FE group) and haplogroup labels are indicated at the tips of the tree. Reference data set group labels are superscripts ($^{\mathrm{TAI}}$: Taiwan; $^{\mathrm{INDO}}$: Indonesia including Besemah and Semende; $^{\mathrm{NO}}$: Near Oceania; $^{\mathrm{RO}}$: Remote Oceania). FE group samples and haplogroup labels are in bold text. Haplogroup labels are results of mtDNAble (Behar *et al*. 2012) haplogroup assignment.



Supplementary Figure 9.

Supplementary Figure 9. Maximum clade credibility tree for haplogroup E1a1a.

Tree topology and coalescent time estimates with 95% highest posterior density (HPD) are results of BEAST (Drummond and Rambaut 2007) analyses. Sample names (RSRS, Reference data set, FE group) and haplogroup labels are indicated at the tips of the tree. Reference data set group labels are superscripts (^{TAI}: Taiwan; ^{INDO}: Indonesia including Besemah and Semende). FE group samples and haplogroup labels are in bold text. Haplogroup labels are results of mtDNAble (Behar *et al*. 2012) haplogroup assignment.



Supplementary Figure 10.

Supplementary Figure 10. Maximum clade credibility tree for haplogroup M7c3c.

Tree topology and coalescent time estimates with 95% highest posterior density (HPD) are results of BEAST (Drummond and Rambaut 2007) analyses. Sample names (RSRS, Reference data set, FE group) and haplogroup labels are indicated at the tips of the tree. Reference data set group labels are superscripts (^{TAI}: Taiwan; ^{INDO}: Indonesia including Besemah and Semende; ^{MAL}: Malaysia; ^{RO}: Remote Oceania). FE group samples and haplogroup labels are in bold text. Haplogroup labels are results of mtDNAble (Behar *et al*. 2012) haplogroup assignment.

	RSRS	MN34[A] (M52b)	MN98[A] (M52b)	MN99[A] (M52b)	MN43[A] (M52b)	NI36[A] (M52b)	SH17[B] (M52a)	SH53[B] (M52)	SU18[B] (M52a)	SU57[B] (M52a)	SU54[B] (M52a)	SH107[B] (M52a)	SH103[B] (M52a)	SH96[B] (M52a)	SH07[B] (M52a)	**AeZ1790 (M52a)**	**AgI4396 (M52'58)**	**AeZ1787 (M52'58)**	**AeZ1785 (M52'58)**	**AeZ1769 (M52a)**	**AeZ1791 (M52'58)**	**AgI4397 (M52a)**	**AeZ1783 (M52a)**	**AeZ1789 (M52a)**	**AgI4386 (M52'58)**	**AgI4401 (M52'58)**	**AgI4385 (M52'58)**	**AeZ1795 (M52a)**	WA103[A] (M52'58)
M52																													
A16275G		A16275G	A16275G	A16275G	A16275G	A16275G	A16275A	A16275G	A16275G	A16275G	A16275G	A16275G	A16275G	A16275G	A16275G	A16275G	A16275G	A16275N	A16275G	A16275G	A16275G	A16275N	A16275G	A16275G	A16275G	A16275G	A16275G	A16275G	A16275A
M52a																													
C15349T		C15349C	C15349C	C15349C	C15349C	C15349C	C15349T	C15349C	C15349T	C15349T	C15349T	C15349T	C15349T	C15349T	C15349T	C15349T	C15349C	C15349N	C15349N	C15349T	C15349T	C15349T	C15349T	C15349T	C15349N	C15349N	C15349N	C15349T	C15349C
G16390A		G16390G	G16390G	G16390G	G16390G	G16390G	G16390A	G16390G	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390A	G16390G
M52a1																													
C573.XC		C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC	C573N	C573.XC	C573.XC	C573.XC	C573.XC	C573.XC
G6755A		G6755G	G6755G	G6755G	G6755G	G6755G	G6755A	G6755A	G6755A	G6755A	G6755A	G6755A	G6755A	G6755A	G6755A	G6755A	G6755N	G6755N	G6755A	G6755A	G6755A	G6755A	G6755N	G6755N	G6755N	G6755A	G6755A	G6755A	G6755G
C16327a		C16327C	C16327C	C16327C	C16327C	C16327C	C16327a	C16327C	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327a	C16327C
M58																													
T152C!		T152C!	T152C!	T152C!	T152C!	T152C!	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	C152T	T152C!
G13135A		G13135G	G13135G	G13135G	G13135G	G13135G	G13135G	G13135G	G13135G	G13135G	G13135G	G13135G	G13135G	G13135G	G13135G	G13135A	G13135A	G13135A	G13135A	G13135N	G13135A	G13135N	G13135A	G13135G	G13135A	G13135A	G13135A	G13135A	G13135A
A16230G!		G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230N	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	G16230A	A16230G!
C16270T		C16270C	C16270C	C16270C	C16270C	C16270C	C16270C	C16270C	C16270C	C16270C	C16270C	C16270C	C16270C	C16270C	C16270C	C16270T	C16270T	C16270N	C16270T	C16270T	C16270T	C16270N	C16270T	C16270T	C16270T	C16270T	C16270T	C16270T	C16270T

Supplementary Figure 11.

Supplementary Figure 11. Maximum clade credibility tree for haplogroup M52’58.

Tree topology and coalescent time estimates with 95% highest posterior density (HPD) are results of BEAST (Drummond and Rambaut 2007) analyses. Sample names (RSRS, Reference data set, FE group) and haplogroup labels (in parenthesis) are indicated at the tips of the tree. South Asian Reference data set group markers are superscripts (^{A}: non-Muslim Indians from Chandrasekar *et al*. 2009; ^{B}: Indian Muslims from Eaaswarkhanth *et al*. 2009). FE group samples and haplogroup labels are in bold text. Haplogroup labels are results of mtDNAble (Behar *et al*. 2012) haplogroup assignment. Haplogroups (underlined) and diagnostic nucleotides are indicated.

Supplementary Figure 12.

Supplementary Figure 12. Maximum clade credibility tree for haplogroup N.

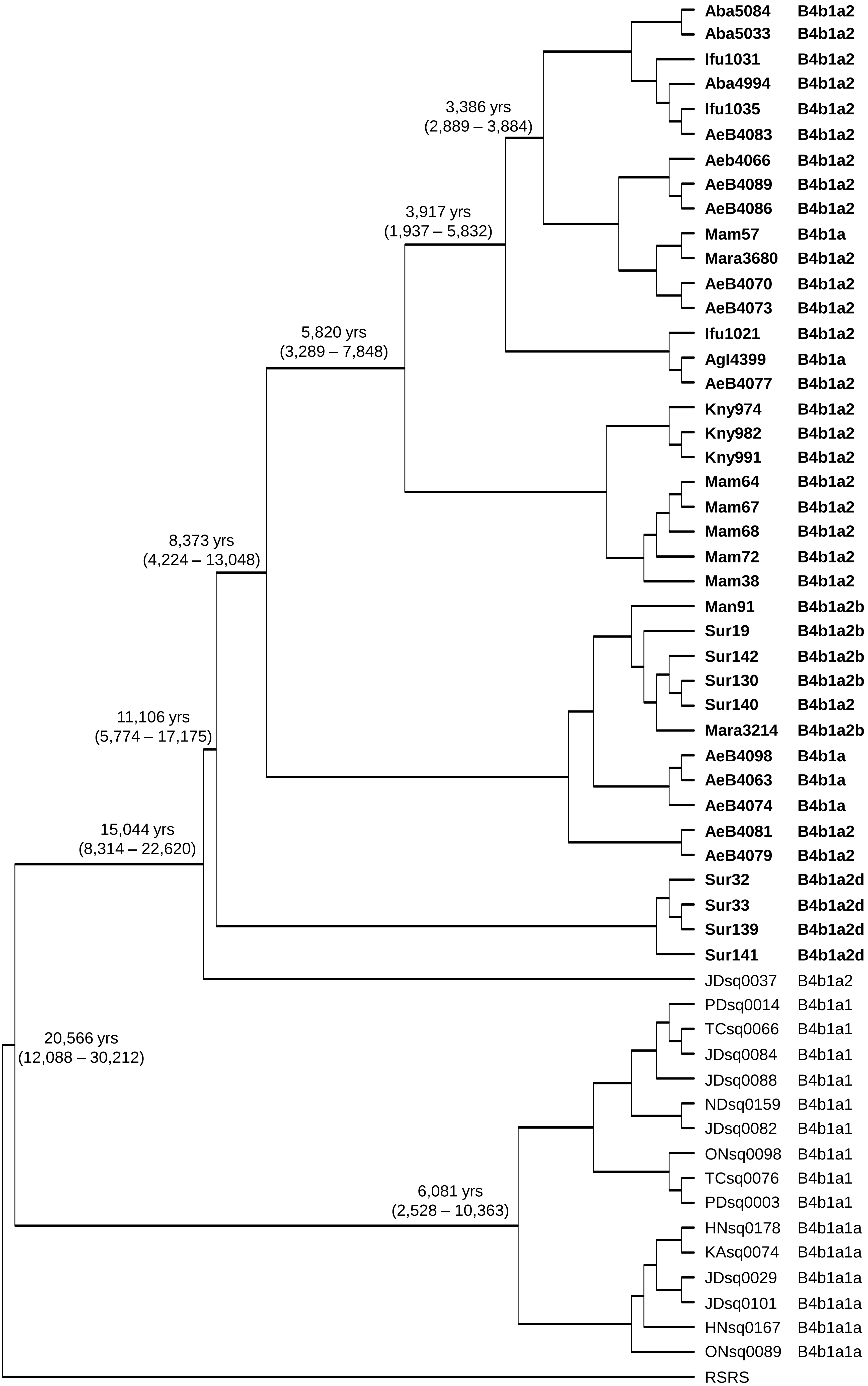
Tree topology and coalescent time estimates with 95% highest posterior density (HPD) are results of BEAST (Drummond and Rambaut 2007) analyses. Sample names (RSRS, Reference data set, FE group) and haplogroup labels are indicated at the tips of the tree. Reference data set group labels are superscripts ($^{\text{INDO}}$: Indonesia including Besemah and Semende; $^{\text{MAL}}$: Malaysia; $^{\text{IND}}$: India; $^{\text{AUS}}$: Australia). Japan Reference data samples are labelled with the following code pattern: –sq0--- (i.e. JDsq0037). Mamanwa FE group samples and haplogroup labels are in bold text. Haplogroup labels are results of mtDNAble (Behar *et al*. 2012) haplogroup assignment.

	RSRS	AUR17^{AUS} (P4b1)	Bundi246LL^{NO} (P4)	UV1145^{NO} (P4)	B009-099NO (P4a1)	BlueNG230^{NO} (P4a1)	**Mara3675** **(P2’10)**	**Bug3269** **(P2’10)**	**Bug3272** **(P2’10)**	B015-040NO (P2)	TRO122^{NO} (P2)	TRO133^{NO} (P2)	AUD32^{AUS} (P8)	**Mara3677** **(P1d1)**	AL25^{NO} (P3b1)	Mon515^{NO} (P3b1)	**AeB4088** **(P)**	**AeB4080** **(P)**	**AeB4082** **(P)**	**AeB4069** **(P)**	**AeB4087** **(P)**	**AeB4094** **(P)**	**AeB4090** **(P)**	**AeB4068** **(P)**	**AeB4097** **(P)**
P3																									
573.XC		573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC	573.XC
P5																									
T152C!		C152T	T152C!	T152C!	T152C!	T152C!	C152T	T152C!	T152C!	C152T	T152C!	T152C!	T152C!	C152T	C152T	T152C!	C152T	C152T	C152T	T152C!	T152C!	T152C!	T152C!	T152C!	T152C!
T16311C!		C16311T	C16311T	C16311T	T16311C!	C16311T	T16311C!	T16311C!	T16311C!	T16311C!	C16311T	C16311T	C16311T	C16311T	C16311T	T16311C!	C16311T	C16311T	C16311T	T16311C!	T16311C!	T16311C!	T16311C!	T16311C!	T16311C!
P6																									
T6719C		T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T	T6719T
T16311C!		C16311T	C16311T	C16311T	C16311T	C16311T	T16311C!	T16311C!	T16311C!	T16311C!	C16311T	C16311T	C16311T	C16311T	C16311T	T16311C!	C16311T	C16311T	C16311T	T16311C!	T16311C!	T16311C!	T16311C!	T16311C!	T16311C!
T16362C		T16362C	T16362T	T16362T	T16362T	T16362T	T16362T	T16362T	T16362T	T16362C	T16362T	T16362T	T16362C	T16362T	T16362T	T16362T	T16362T	T16362T	T16362T	T16362T	T16362T	T16362T	T16362T	T16362T	T16362T
P7																									
T152C!		C152T	T152C!	T152C!	T152C!	T152C!	C152T	T152C!	T152C!	C152T	T152C!	T152C!	T152C!	C152T	T152C!	T152C!	C152T	C152T	C152T	T152C!	T152C!	T152C!	T152C!	T152C!	T152C!
P9																									
C3172a		C3172C	C3172C	C3172C	C3172C	C3172C	C3172C	C3172C	C3172C	C3172C	C3172C	C3172C	C3172C	C3172C	C3172C	C3172C	C3172a	C3172a	C3172a	C3172a	C3172a	C3172a	C3172a	C3172a	C3172a
C13452a		C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452C	C13452a	C13452a	C13452a	C13452a	C13452a	C13452a
T14788C		T14788T	T14788T	T14788T	T14788T	T14788T	T14788T	T14788T	T14788T	T14788T	T14788T	T14788T	T14788T	T14788T	T14788T	T14788T	T14788C	T14788C	T14788C	T14788C	T14788C	T14788C	T14788C	T14788C	T14788C
C16278T!		T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C	T16278C
Coding Region																	G1888A C4354T A4547G C6053T G8860A A9067G T9656C T10101C C10718T G11440A C13934T C15885T	G1888A C4354T A4547G C6053T G8860A A9067G T9656C T10101C C10718T G11440A C13934T C15885T	G1888A C4354T A4547G C6053T G8860A A9067G T9656C T10101C C10718T G11440A C13934T C15885T	G2831A A4985G A5414G A6125G T8610C C9159g G12372A T12757C A14878G A14879G	A4985G A5414G A6125G T8610C C9159g G12372A T12757C A14878G A14879G	G2831A A5414G A6125G T8610C C9159g G12372A T12757C A14878G A14879G	G2831A A5414G A6125G T8610C C9159g G12372A T12757C A14878G A14879G	G2831A A5414G A6125G T8610C C9159g G12372A T12757C A14878G A14879G	G2831A A5414G A6125G T8610C C9159g G12372A T12757C A14878G A14879G

Supplementary Figure 13.

Supplementary Figure 13. Maximum clade credibility tree for haplogroup P.

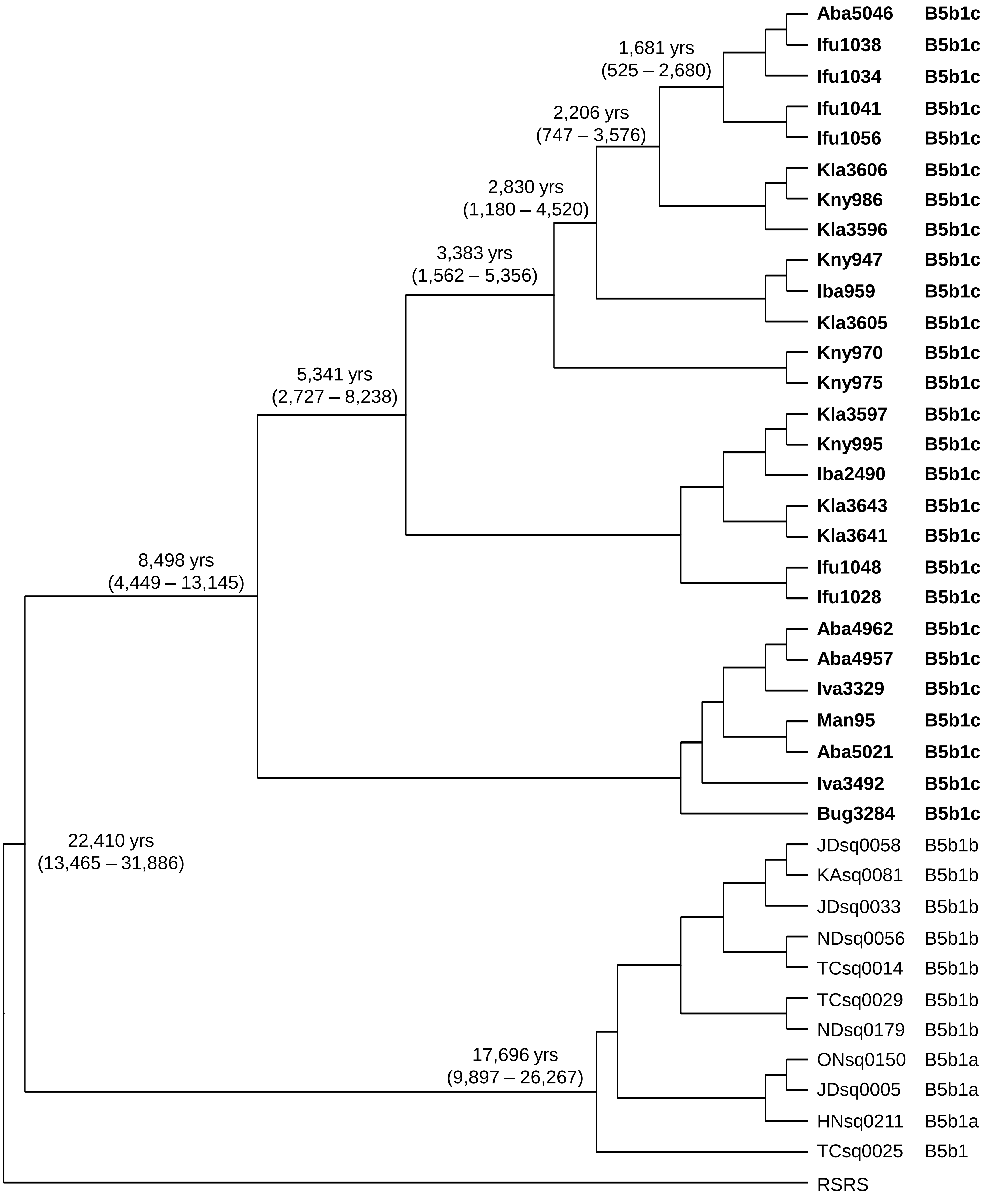
Tree topology and coalescent time estimates with 95% highest posterior density (HPD) are results of BEAST (Drummond and Rambaut 2007) analyses. Sample names (RSRS, Reference data set, FE group) and haplogroup labels (in parenthesis) are indicated at the tips of the tree. Reference data set group labels are superscripts (^{AUS}: Australia; ^{NO}: Near Oceania) FE group samples and haplogroup labels are in bold text. Haplogroup labels are results of mtDNAble (Behar *et al*. 2012) haplogroup assignment. Haplogroups (underlined), diagnostic nucleotides and FE group coding region polymorphisms are indicated.



Supplementary Figure 14.

Supplementary Figure 14. Maximum clade credibility tree for haplogroup B4b1a.

Tree topology and coalescent time estimates with 95% highest posterior density (HPD) are results of BEAST (Drummond and Rambaut 2007) analyses. Sample names (RSRS, Reference data set, FE group) and haplogroup labels are indicated at the tips of the tree. Japan Reference data samples are labelled with the following code pattern: –sq0--- (i.e. JDsq0037). FE group samples and haplogroup labels are in bold text. Haplogroup labels are results of mtDNAble (Behar *et al*. 2012) haplogroup assignment.



Supplementary Figure 15.

Supplementary Figure 15. Maximum clade credibility tree for haplogroup B5b1c.

Tree topology and coalescent time estimates with 95% highest posterior density (HPD) are results of BEAST (Drummond and Rambaut 2007) analyses. Sample names (RSRS, Reference data set, FE group) and haplogroup labels are indicated at the tips of the tree. Japan Reference data samples are labelled with the following code pattern: –sq0--- (i.e. JDsq0037). FE group samples and haplogroup labels are in bold text. Haplogroup labels are results of mtDNAble (Behar *et al*. 2012) haplogroup assignment.