Supplementary Table 1: Molecular diversity indices for populations in West Timor based on Y chromosome STRs (*n* = 14)

Population	**N**	**h**	**S**	**Diversity Indices**					
				Haplotype diversity	**SD**	**MNPD**	**SD**	**Nucleotide diversity**	**SD**
Umaklaran	41	29	13	0.97	0.013	8.38	3.96	0.60	0.31
Fatuketi	35	26	13	0.98	0.011	6.72	3.25	0.48	0.26
Tialai	24	18	13	0.96	0.025	8.35	4.01	0.60	0.32
Raimanawe	50	32	13	0.96	0.017	7.53	3.58	0.54	0.28
Kamanasa	62	52	14	0.99	0.005	7.99	3.76	0.57	0.30
Kletek Rainan	17	15	13	0.99	0.025	8.80	4.27	0.63	0.34
Kletek Suai	15	11	13	0.95	0.040	7.70	3.80	0.55	0.30
Kletek Wefatuk	15	14	13	0.99	0.028	7.60	3.76	0.54	0.30
Kateri	50	29	13	0.96	0.015	6.66	3.20	0.48	0.25
Laran	50	45	13	1.00	0.005	7.45	3.54	0.53	0.28
Kakaniuk	47	26	13	0.95	0.017	7.57	3.60	0.54	0.29
Umanen Lawalu	49	41	14	0.99	0.007	8.29	3.91	0.59	0.31
Besikama	42	32	13	0.98	0.012	8.52	4.02	0.61	0.32

Abbreviations: N, number of sequences; h, number of haplotypes; S, number of polymorphic sites; MNPD, mean number of pairwise differences; SD, standard deviation.

Supplementary Table 2: Y chromosome haplogroup frequencies (percentages) in West Timor

	TIM_BS	TIM_FT	TIM_KM	TIM_KT	TIM_KK	TIM_KR	TIM_KS	TIM_KW	TIM_LR	TIM_RM	TIM_TI	TIM_UM	TIM_UL	TIMOR
	n=42	*n=35*	*n=62*	*n=50*	*n=47*	*n=17*	*n=15*	*n=15*	*n=50*	*n=50*	*n=24*	*n=41*	*n=49*	*n=497*
C-RPS4Y	2.4		3.2		8.5	11.8		6.7	6.0	4.0		12.2	4.1	4.4
C-M38	9.5	37.1	22.6	46.0	25.5	23.5	20.0	53.3	32.0	24.0	4.2	17.1	30.6	26.6
C-P355	14.3	28.6	11.3	12.0	2.1	5.9	6.7		8.0	24.0	20.8	4.9	14.3	12.5
C-M208			1.6								4.2			0.4
C-P343		2.9			2.1									0.4
D-M116												2.4		0.2
E-P1									2.0					0.2
F-P14*				2.0	2.1						20.8	2.4	2.0	1.8
J-M172			1.6					6.7						0.4
K-P397	9.5	2.9	19.4		4.3	5.9		6.7	14.0	8.0	16.7	19.5	10.2	9.9
K-P79	2.4													0.2
K-P336												2.4		0.2
M-P34	9.5		12.9	2.0		17.6	13.3	6.7	2.0	2.0		17.1	6.1	6.2
O-P201	4.8		14.5	2.0	4.3				2.0	2.0			2.0	3.4
O-JST002611	2.4					5.9								0.4
O-M134		5.7			2.1				2.0	2.0				1.0
O-M119	11.9	2.9		2.0	4.3				2.0		4.2		2.0	2.4
O-P203		2.9							4.0			7.3		1.2
O-M110	9.5	8.6	1.6	18.0	12.8		6.7	6.7	10.0	6.0	8.3	4.9	6.1	8.0
O-M95				2.0			6.7		2.0	2.0				0.8
P-P295	4.8	8.6	4.8	14.0	25.5	5.9	6.7	6.7	10.0	18.0	16.7	2.4	10.2	10.9
Q-P36													2.0	0.2
S-M254	14.3		4.8			11.8	40.0	6.7	2.0	8.0	4.2	7.3	6.1	6.0
S-P377	4.8		1.6		6.4	11.8			2.0				4.1	2.2

Populations: TIM_BS: Besikama; TIM_FT: Fatuketi; TIM_KM: Kamanasa; TIM_KT: Kateri; TIM_KK: Kakaniuk; TIM_KR: Kletek Rainan; TIM_KS: Kletek Suai; TIM_KW: Kletek Wefatuk; TIM_LR: Laran; TIM_RM: Raimanawe; TIM_TI: Tialai; TIM_UM: Umaklaran; TIM_UL: Umanen Lawalu.

Supplementary Table 3: Regional Y chromosome haplogroup frequencies (percentages) in West Timor and surrounding islands

	TIM	SUM	FLO	LEM	SLW	PAN	ALO	MAL	BLI	TAB	PHI	PNG	MEL	AUS
	n=497	*n=649*	*n=394*	*n=92*	*n=177*	*n=26*	*n=22*	*n=30*	*n=502*	*n=48*	*n=47*	*n=48*	*n=70*	*n=40*
C-RPS4Y	4.4	1.2	25.4	3.3	7.3		4.5		1.8			2.1		
C-M38	26.6	49.5	5.8	4.3	4.0	7.7	18.2	16.7				6.3	12.9	
C-P355	12.5	1.5	9.1	17.4	0.6	3.8	36.4	16.7						
C-M208	0.4											2.1	1.4	
C-P343	0.4		4.3	17.4		3.8								
C-M347														55.0
C-M217										2.1	2.1			
D-M116	0.2													
E-P1	0.2													
F-P14*	1.8	0.2	1.5	5.4	2.3									
H-M69									3.4					
H-Apt					1.1									
J-M304									0.2					
J-M172	0.4				0.6									
J-M12									0.4					
K-M526*					0.6									
K-P261									1.0					
K-P397	9.9	4.5	6.1	3.3	5.1	3.8	4.5	16.7				6.3	1.4	
K-P405		0.3												
K-P60														45.0
K-P79	0.2											4.2	8.6	
K-P315												2.1	8.6	
K-P336	0.2						22.7							
K-P378											31.9			
L-M20					1.1				2.4					

M-P256													2.9
M-M5			2.8			3.8						6.3	
M-P34	6.2	4.5	18.5	12.0	1.1	15.4	9.1	16.7	0.8			25.0	2.9
M-P87													10.0
M-P22													22.9
M-M353													2.9
M-P118													11.4
O-M122		2.2							0.6				
O-P197											2.1		1.4
O-P201	3.4	1.4	4.8		22.0	38.5		10.0	6.0	6.3	23.4	4.2	2.9
O-M7					0.6				1.0		2.1		
O-JST002611	0.4										2.1		
O-M134	1.0				0.6				0.2		4.3		
O-M119	2.4	9.6		2.2	10.7	7.7			8.4				1.4
O-P203	1.2	0.5	9.1	1.1	6.2			3.3	10.4	70.8	2.1	2.1	
O-M110	8.0	4.5			14.7			3.3	0.8	18.8	8.5		1.4
O-M95	0.8	1.1	4.8		9.0				58.4		2.1		
O-M111										2.1	2.1		
P-P295	10.9	3.2			0.6						14.9		
Q-P36	0.2												
Q-M346									0.4				
R-M124					1.7				2.2				
R-P25									0.4				
R-M269									0.4				
R-M17					5.1				0.6				
S-M230		0.5	0.5	2.2									2.9
S-M254	6.0	15.1	6.9	19.6	5.1			16.7	0.4		2.1	37.5	4.3
S-P377	2.2	0.5	0.3	12.0		15.4	4.5						
S-M226												2.1	

Populations: TIM: Timor; SUM: Sumba; FLO: Flores; LEM: Lembata; ALR: Alor; PTR: Pantar; SLW: Sulawesi; MAL: Maluku; BLI: Bali; TAB: Taiwanese Aboriginals; PHIL: Philippines; PNG: Papua New Guinea; MEL: Melanesia; AUS: Australia. All data, except Timor, from Karafet and colleagues[14].

Supplementary Table 4: mtDNA haplogroup frequencies (percentages) in West Timor

	TIM_BS	TIM_FT	TIM_KM	TIM_KT	TIM_KK	TIM_KR	TIM_KS	TIM_KW	TIM_LR	TIM_RM	TIM_TI	TIM_UM	TIM_UL	TIMOR
	n=42	*n=35*	*n=67*	*n=50*	*n=49*	*n=31*	*n=20*	*n=20*	*n=50*	*n=50*	*n=24*	*n=41*	*n=50*	*n=529*
B4a	7.1	17.1	7.5	4.0	2.0	45.2	15.0	15.0	6.0	8.0	20.8		4.0	9.6
B4a1a1a	2.4		6.0	36.0	2.0		5.0		18.0	6.0			4.0	7.4
B4b1		2.9	3.0	2.0	2.0		25.0	15.0	6.0			2.4	4.0	3.6
B4c1b3	2.4		3.0						2.0				2.0	0.9
B4c2	7.1													0.6
B5a						3.2		15.0						0.8
B5b1			6.0						2.0	2.0		4.9		1.5
D			3.0	4.0										0.8
D4		5.7												0.4
D5			1.5											0.2
D5b1c	2.4		3.0	22.0										2.6
E		2.9	7.5				5.0		4.0	8.0	8.3		2.0	3.0
E1a1a	16.7	17.1	1.5	6.0	4.1	9.7		5.0	6.0	2.0			4.0	5.5
E1b	9.5	11.4	7.5			12.9		5.0	8.0	4.0	20.8	12.2		6.4
E2	2.4						5.0							0.4
F1a			1.5					5.0	2.0				2.0	0.8
F1a3			1.5						18.0	2.0	4.2		6.0	2.8
F1a4	16.7	2.9	13.4	8.0	53.1	6.5	30.0	30.0		18.0	16.7	9.8	48.0	19.3
F3								5.0	2.0				4.0	0.8
M*	4.8	2.9	3.0	2.0					2.0	8.0	4.2	12.2		3.2
M7												2.4		0.2
M7c3	2.4			8.0		3.2				4.0		4.9		1.9
M7c3c		2.9	10.4		14.3		5.0	5.0	4.0	14.0		9.8	2.0	5.9
N*	2.4	5.7								2.0				0.8
N9a	7.1		1.5	2.0		6.5							4.0	1.7

P*	2.4		4.5		2.0									0.9
Q	2.4					3.2	5.0			2.0		2.4		0.9
Q1	9.5	25.7	14.9	4.0	18.4	6.5			14.0	16.0	25.0	29.3	10.0	14.0
R*				2.0		3.2								0.4
R9c1	2.4	2.9							2.0	4.0		9.8	4.0	2.1
Y2					2.0		5.0		4.0					0.8

* Mitochondrial paragroups: Individuals were not assigned to any derived haplogroup.

Populations: TIM_BS: Besikama; TIM_FT: Fatuketi; TIM_KM: Kamanasa; TIM_KT: Kateri; TIM_KK: Kakaniuk; TIM_KR: Kletek Rainan; TIM_KS: Kletek Suai; TIM_KW: Kletek Wefatuk; TIM_LR: Laran; TIM_RM: Raimanawe; TIM_TI: Tialai; TIM_UM: Umaklaran; TIM_UL: Umanen Lawalu.

Supplementary Table 5: Regional mtDNA haplogroup frequencies (percentages) in West Timor and surrounding islands

	TIM Present Study *n*=529	SUM Tumonggor et al, 2013[6] *n*=639	FLO Tumonggor et al, 2013[6] *n*=453	LEM Tumonggor et al, 2013[6] *n*=92	ALR Tumonggor et al, 2013[6] *n*=23	PTR Tumonggor et al, 2013[6] *n*=27	SLW Tumonggor et al, 2013[6] *n*=188	MAL Hill et al, 2007[58] *n*=43	BLI Tumonggor et al, 2013[6] *n*=457	TAB Tabbada et al, 2010[41] *n*=640	PHIL Tabbada et al, 2010[41] *n*=423	PNG Friedlaender et al, 2007[42] *n*=231	MEL Friedlaender et al, 2007[42] *n*=607	AUS Hudjasov et al, 2007[45] *n*=38
B		0.3	0.2						0.4					
B4a	9.6	5.8	4.9	13	13		8	9.3	2.4	8.7	11.8	19.0	16.0	
B4a1a1a	7.4	0.6			4.3			14	0.4					
B4b1	3.6	5.6	2.9	3.3	4.3		4.3	4.7		6	7.6		0.3	
B4c1b3	0.9	4	0.3			14.8	10.6		11.8	4.9	5			
B4c2	0.6	0.2	1.2						7.4					
B5a	0.8	3.1	4	1.1				7	18.4	5.9	0.7			
B5b1	1.5	5.6	2.6	1.1	4.3		1.1		3.7		7.8			
D	0.8	2	3.8				0.5	4.7						
D4	0.4													
D5	0.2	1.4	1.8	1.1			1.6	2.3		4.8				
D5b1c	2.6	4.2	1.5	3.3										
E	3.0	1.4	1.8				2.7		1.5	2.3				
E1a1a	5.5	3	2.4	8.7	8.7	3.7	18.1	4.7	2.4	6.9	11.1	0.4	1.2	
E1b	6.4	4.7	1.1	1.1	4.3	3.7	6.9		0.7		1.4	0.4	9.1	
E2	0.4	3.3	3.1				5.9	4.7	1.1	2.8	3.3	0.4	0.2	
F1a	0.8	3.8	0.2				0.5		1.8	2.2				
F1a1		1.3	2.4	2.2			0.5	2.3	1.8	3.9	0.2			
F1a3	2.8	5.5	5.3	2.2			2.1	7	2.6		3.1			
F1a4	19.3	4.7	3.8	14.1	13		2.7	4.7	1.5		4.3			
F1ac			2				1.1		0.2					
F3	0.8	0.6									1.9			
M*	3.2	8.3	14.7	2.2	8.7	14.8	6.9	2.3	10.7					
M7	0.2							2.3	1.1					
M7b		1.4	0.7					2.3			0.2			

M7b1			0.4	1.1				2.3	1.5	0.7	1.2			
M7b3		5.2	1.6				2.1			8.2	3.3			
M7c3	1.9		0.2	2.2					1.5					
M7c3c	5.9	11	4.6	8.7	4.3	7.4	7.4	2.3	3.3	7.9	11.3			
M9			0.2											
M13b1			0.2	1.1							0.7			
M17a		0.8	0.9		4.3	3.7	2.7	2.3	5.5					
M73		1.7	4.4			7.4	8	2.3	3.9		0.47			
M47		0.3												
N*	0.8	0.8	6.6	2.2		3.7								
N21									2					
N22									0.7		0.2			
N9a	1.7		0.2				0.5		1.1	1.2				
P*	0.9	1.4	5.7	13	4.3	3.7	0.5		0.2			42.9	4.9	15.7
Q	0.9													
Q1	14.0	2.8	6	7.6	21.7	33.3	1.6	11.6				29.9	12.2	
Q2				3.3								4.8	9.9	3.1
Q3									0.7			0.9	0.8	
R*	0.4	1.6	3.3	1.1	4.3	3.7	1.1		0.2					
R9			2.2				1.6							
R9b1		0.2		1.1										
R9c1	2.1	2.5	0.4							2.9	5			
R22		0.6	1.8	4.3					4.6					
R23		0.2							1.5					
Y2	0.8		0.4	1.1			1.1		3.3	1.4	4.7			
M27													2.3	
M28												0.4	36.7	
M29													6.1	
M13														21.9
M14														3.1
N13														21.9
S2														28.1

S5		6.3

* Mitochondrial paragroups: Individuals were not assigned to any derived haplogroup.

Populations: TIM: Timor; SUM: Sumba; FLO: Flores; LEM: Lembata; ALR: Alor; PTR: Pantar; SLW: Sulawesi; MAL: Maluku; BLI: Bali; TAB: Taiwanese Aboriginals; PHIL: Philippines; PNG: Papua New Guinea; MEL: Melanesia; AUS: Australia.

Supplementary Table 6: Standardized genetic distances in West Timor (G'_{ST})

	Number of Subpopulations	G'_{ST}			
		mtDNA HVS I	mtDNA SNP	Y STR	Y SNP
All Indonesian populations	70	0.86	0.57	0.97	0.79
Timor	13	0.69	0.42	0.72	0.24

Supplementary Table 7: Mantel tests for genetic, linguistic and geographic associations (Based on seven populations in Timor that are primarily monolingual)

Correlation	*r* (correlation coefficient)	*P*
Full correlation		
mtDNA HVS1 Slatkin and geography	0.12	0.27
mtDNA HVS1 Slatkin and language	0.13	0.37
mtDNA SNP Slatkin and geography	0.12	0.31
mtDNA SNP Slatkin and language	0.15	0.31
YSTR Slatkin and geography	0.36	0.097
YSTR Slatkin and language	0.24	0.19
YSNP Slatkin and geography	0.43	0.08
YSNP Slatkin and language	-0.31	0.90
Geography and language	0.02	0.43
Partial correlation		
mtDNA HVS1 Slatkin and geography (language*)	0.12	0.29
mtDNA HVS1 Slatkin and language (geography *)	0.13	0.37
mtDNA SNP Slatkin and geography (language*)	0.12	0.31
mtDNA SNP Slatkin and language (geography*)	0.15	0.34
YSTR Slatkin and geography (language*)	0.36	0.12
YSTR Slatkin and language (geography*)	0.25	0.19
YSNP Slatkin and geography (language*)	0.46	0.057
YSNP Slatkin and language (geography*)	-0.35	0.93
Geography and language (HVS! Slatkin*)	0.0051	0.46
Geography and language (mtDNA SNP Slatkin*)	0.0027	0.46
Geography and language (YSTR Slatkin*)	-0.071	0.57
Geography and language (YSNP Slatkin*)	0.18	0.26

* Indicates constrained parameter.

Supplementary Table 8: Asian admixture rates in Indonesia and surrounding islands (Cox and colleagues[9]; Timor data is from the present study)

			Asian Admixture			
Region	Location	Population	All	A	X	Difference
ISEA	Taiwan	Aboriginal	1.00	1.00	1.00	0
	Philippines	Filipino	1.00	1.00	1.00	0
Indonesia	Sumatra	Toba	1.00	1.00	1.00	0
	Nias	Nias	1.00	1.00	1.00	0
	Mentawai	Mentawai	1.00	1.00	1.00	0
	Java	Dieng	0.98	0.96	1.00	0.04
	Java	Javanese	1.00	1.00	1.00	0
	Borneo	Borneo	1.00	0.99	1.00	0.01
	Bali	Bali	0.99	0.95	1.00	0.04
	Sulawesi	Sulawesi	0.97	0.91	0.99	0.08
	Flores/Lembata	Flores/Lembata	0.66	0.62	0.69	0.08
	Sumba	Sumba	0.81	0.74	0.86	0.12
	Alor	Alor	0.51	0.49	0.54	0.04
	Timor	**Timor**	**0.64**	**0.58**	**0.70**	**0.12**
Melanesia	PNG	Coastal	0.14	0.13	0.16	0.03
	Bismarck Archipelago	New Britain	0	0	0	0
	Bougainville	Nasioi	0	0	0	0
	Vanuatu	Maewo	0.06	0.02	0.14	0.12

Supplementary Table 9: Asian admixture rates in West Timor populations

			Asian Admixture			
Region	Location	Population	All	A	X	Difference
Indonesia	Timor	Besikama	0.65	0.59	0.71	0.12
		Fatuketi	0.64	0.59	0.69	0.10
		Kakaniuk	0.67	0.58	0.74	0.16
		Kamanasa	0.62	0.54	0.69	0.15
		Kateri	0.61	0.56	0.66	0.10
		Kletek Rainan	0.63	0.59	0.67	0.08
		Kletek Suai	0.63	0.58	0.67	0.09
		Kletek Wefatuk	0.72	0.66	0.77	0.11
		Laran	0.62	0.57	0.67	0.10
		Reinawe	0.63	0.57	0.68	0.11
		Tialai	0.62	0.59	0.65	0.06
		Umaklaran	0.63	0.57	0.68	0.11
		Umanen Lawalu	0.68	0.58	0.78	0.20
Mean			*0.64*	*0.58*	*0.70*	*0.11*

Supplementary Table 10: Tentative estimates of the Time of the Most Recent Common Ancestor (TMRCA) of selected mitochondrial DNA haplogroups inferred from the ρ statistic within Indonesia and major surrounding islands

Locations	mtDNA Haplogroup E1a1a		mtDNA Haplogroup F1a4		mtDNA Haplogroup P		mtDNA Haplogroup Q	
	Years BP ± SD	ρ ± SD	Years BP ± SD	ρ ± SD	Years BP ± SD	ρ ± SD	Years BP ± SD	ρ ± SD
Indonesia (Lesser Sunda Islands)	6,700 ± 3,100	0.351 ± 0.16	5,150 ± 1,600	0.268 ± 0.08				
Timor	4,050 ± 2,900	0.211 ± 0.15	3,050 ± 1,200	0.158 ± 0.06	59,600 ±16,050	3.109 ± 0.84	26,000 ±9,250	1.356 ± 0.48
Philippines	3,850 ± 2,500	0.200 ± 0.13	950 ± 600	0.048 ± 0.03				
Taiwan Aboriginals	4,000 ± 1,800	0.210 ± 0.09						
New Guinea					53,700 ± 11,850	2.8 ± 0.62	38,900 ± 8,350	2.03 ±0.43

Note: Dates before present (BP) are rounded to the nearest 50 years.

Supplementary Table 11: Time to the Most Recent Common Ancestor (TMRCA) of selected Y chromosome haplogroups in individuals from Indonesia, Philippines (Aeta), Oceania and Australia using Batwing†

Y Chromosome Haplogroup	TMRCA (95% Confidence Interval) Years BP
C-RPS4Y	41,650 (17,850 – 92,100)
C-M38	32,950 (14,350 – 70,850)
C-P343	16,000 (6,500 – 36,550)
C-P355	13,600 (5,600 – 30,200)
C-M347	15,900 (5,950 – 39,550)
M-P256	30,600 (13,100 – 66,600)
S-M230	22,750 (9,750 – 49,300)

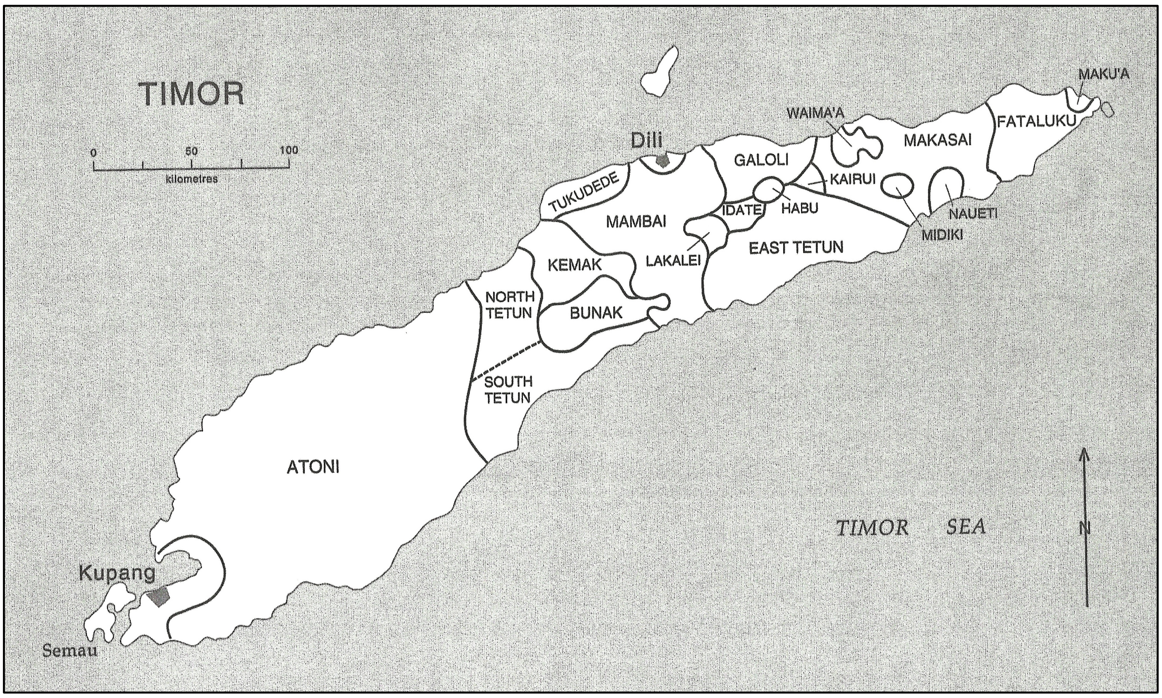
Note: Dates before present (BP) are rounded to the nearest 50 years.

† http://www.mas.ncl.ac.uk/~nijw/#batwing

Supplementary Table 12: A list of new Y chromosome markers, position, primer information and reference SNP ID

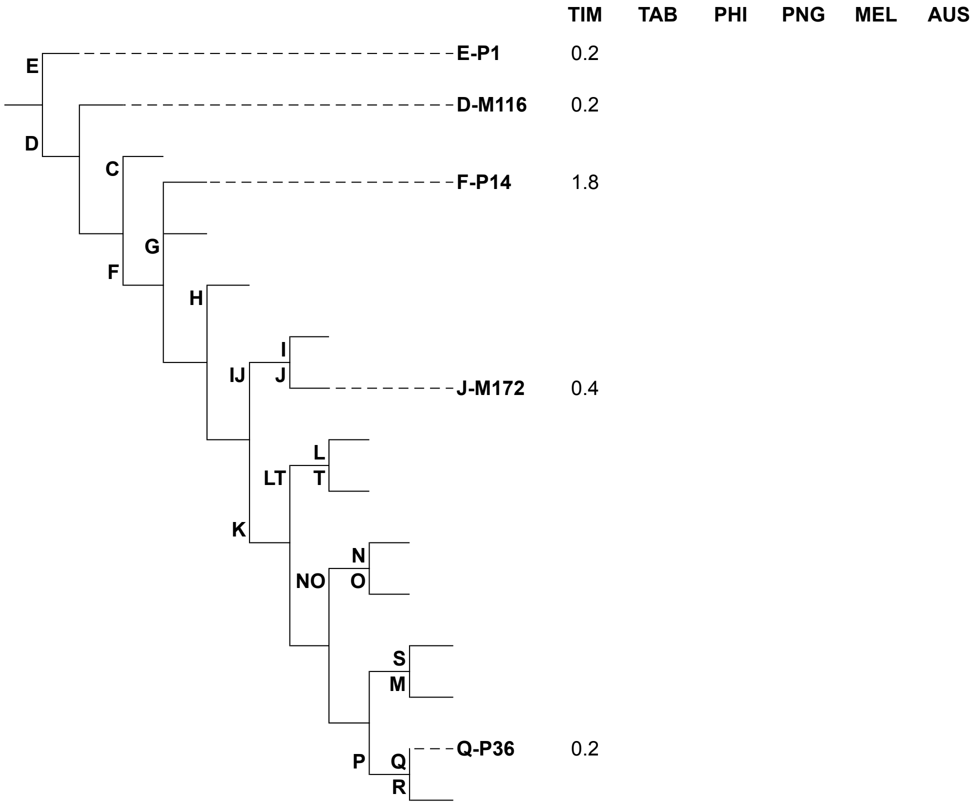
SNP	Haplogroup	Y-position (GRCh37/ hg19)	Forward Primer	Reverse Primer	PCR Size (bp)	Mutation	Site (bp)	Strand
P343	C	2,815,407	ATGAACAGAATGAATTAAATGGT	TATGTTGTAAGATGAAAAATGT	360	Cdel	85	Positive
P355	C	22,077,527	AATCTCCTGCCCTGAATGACTGTCC	TATCCTGTAATCCTTTGTGAGCCTC	399	G→A	250	Negative
P377	S	5,847,373	CAAAAAGTGGGCAATCAGATGACT	ATTTTATTCATTTTTATGGCTGAG	513	A→G	258	Positive

Supplementary Figure 1: Language groups spoken on the island of Timor (adopted from Therik, 2004)



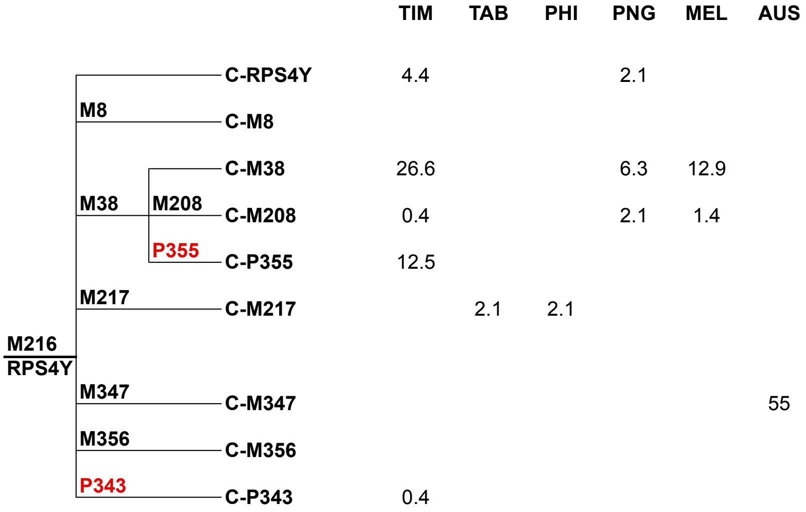
Supplementary Figure 2: Maximum Parsimony tree of Y chromosome haplogroups in West Timor

a) Simplified tree showing lineages from the D and E trunk onward.



	TIM	TAB	PHI	PNG	MEL	AUS
E-P1	0.2					
D-M116	0.2					
F-P14	1.8					
J-M172	0.4					
Q-P36	0.2					

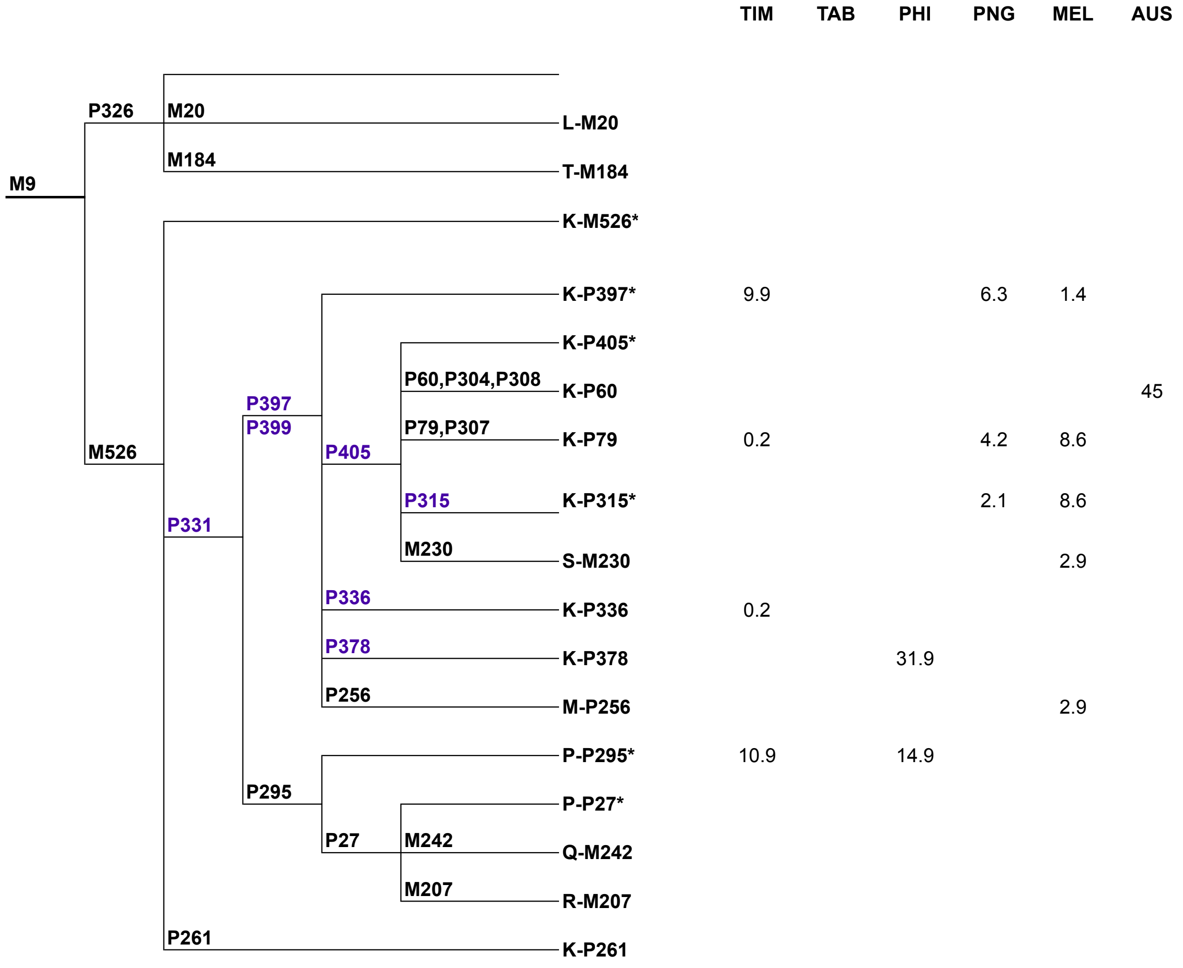
b) Detailed tree of C lineages (including newly reported haplogroups in red).



	TIM	TAB	PHI	PNG	MEL	AUS
C-RPS4Y	4.4			2.1		
C-M8						
C-M38	26.6			6.3	12.9	
C-M208	0.4			2.1	1.4	
C-P355	12.5					
C-M217		2.1	2.1			
C-M347						55
C-M356						
C-P343	0.4					

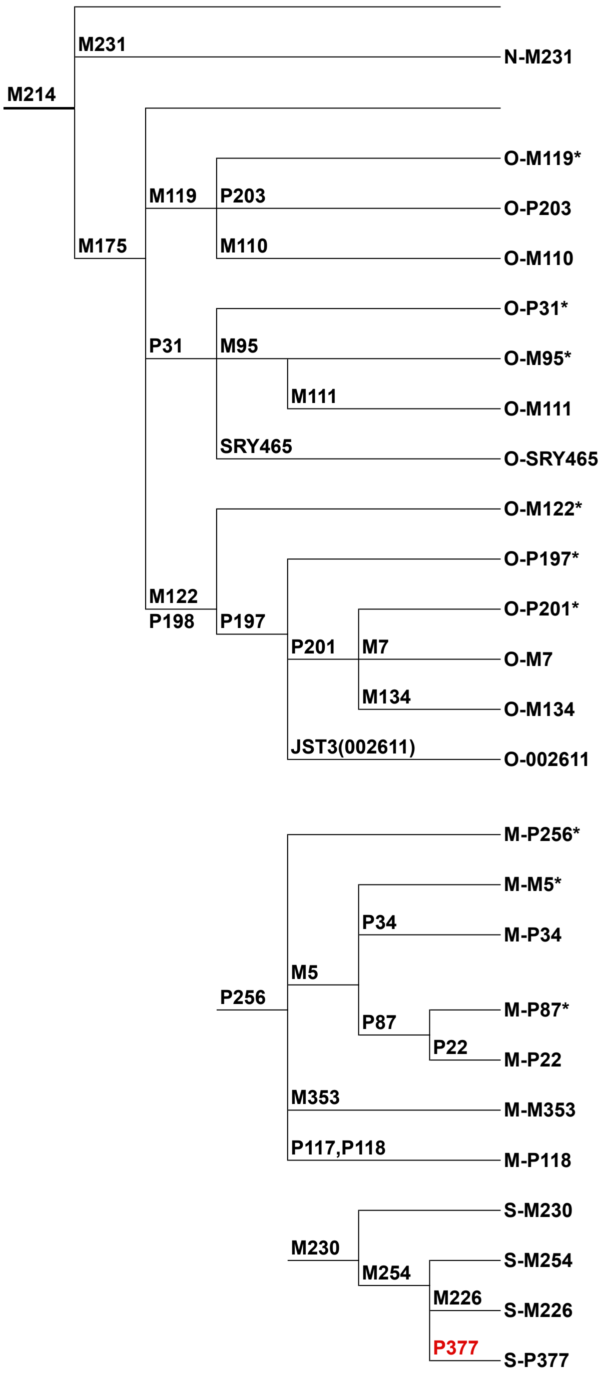
Supplementary Figure 2 *continued*

c) Lineages from the K trunk (including recently reported haplogroups by Karafet and colleagues, *in press,* in purple)



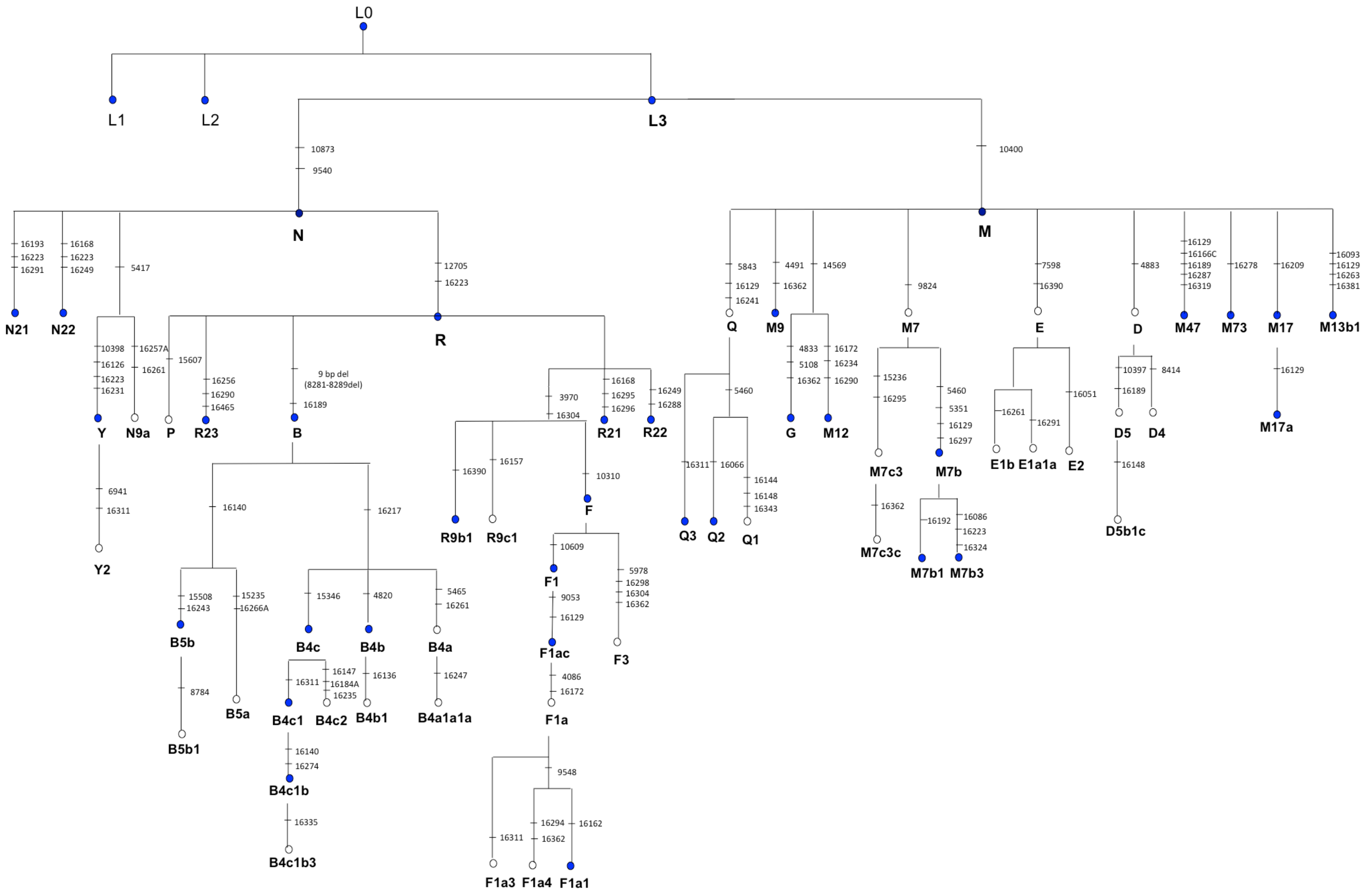
Supplementary Figure 2 ***continued***

d) Lineages O, M and S (including a newly reported haplogroup in red)

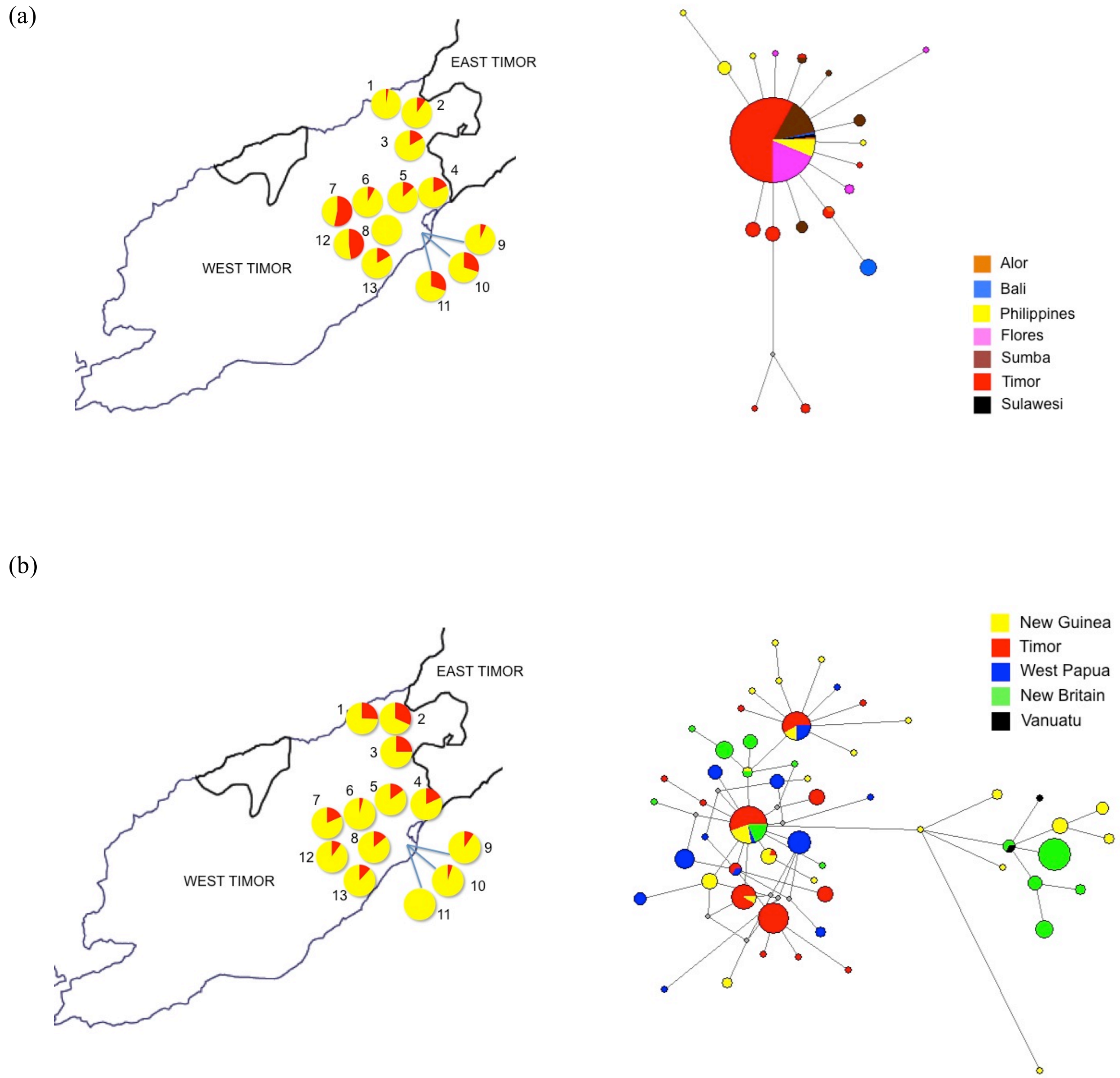


	TIM	TAB	PHI	PNG	MEL	AUS
N-M231						
O-M119*	2.4				1.4	
O-P203	1.2	70.8	2.1	2.1		
O-M110	8	18.8	8.5		1.4	
O-P31*						
O-M95*	0.8		2.1			
O-M111		2.1	2.1			
O-SRY465						
O-M122*						
O-P197*			2.1		1.4	
O-P201*	3.4	6.3	23.4	4.2	2.9	
O-M7			2.1			
O-M134	1		4.3			
O-002611	0.4		2.1			
M-P256*					2.9	
M-M5*				6.3		
M-P34	6.2			25	2.9	
M-P87*					10	
M-P22					22.9	
M-M353					2.9	
M-P118					11.4	
S-M230					2.9	
S-M254	6		2.1	37.5	4.3	
S-M226				2.1		
S-P377	2.2					

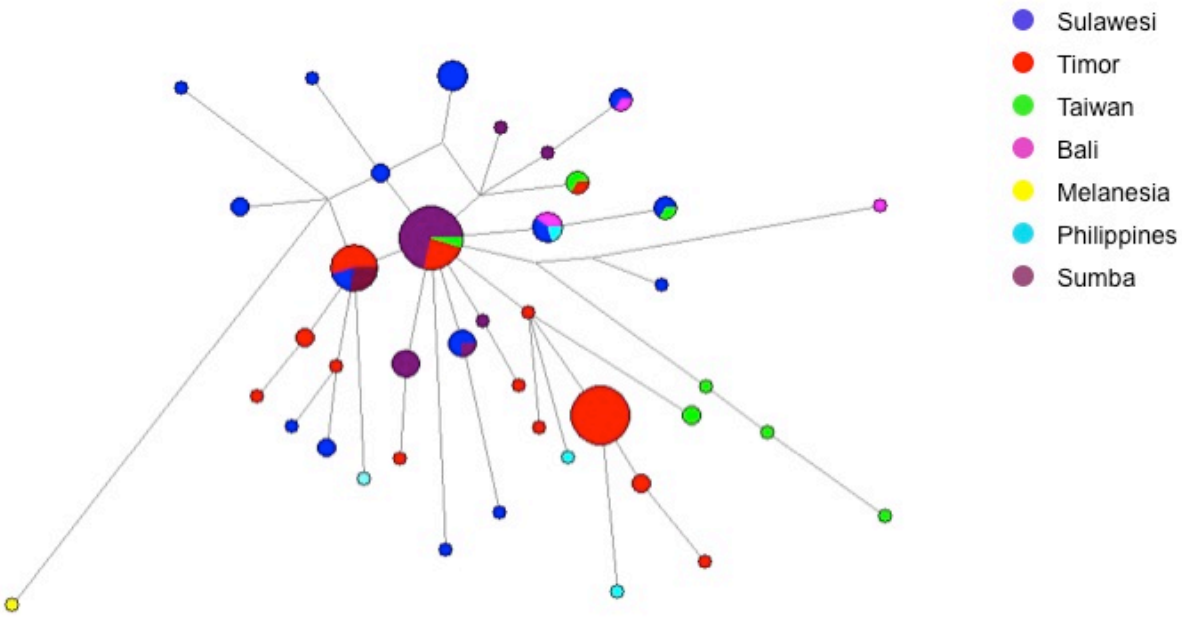
Supplementary Figure 3: Phylogeny of mitochondrial DNA haplogroups observed in Indonesian populations (adapted from Tumonggor et al., 2013). Open circles: mtDNA lineages found in West Timor.



Supplementary Figure 4: Distribution and network of mtDNA haplogroups (a) F1a4 and (b) Q in West Timor. Populations: (1) Fatuketi, (2) Umaklaran, (3) Tialai, (4) Raimanawe, (5) Kamanasa, (6) Kateri, (7) Kakaniuk, (8) Laran, (9) Kletek Rainan, (10) Kletek Suai, (11) Kletek Wefatuk, (12) Umanen Lawalu, and (13) Besikama.



Supplementary Figure 5: Network of Y chromosome haplogroup O-M110



Supplementary Text 1: DNA Extraction and Genetic Screening

DNA was extracted from peripheral blood samples using the salting-out procedure of Miller, Dykes and Polesky[68]. DNA from buccal swabs was extracted using standard phenol-chloroform protocols.

The first hypervariable segment (HVS I) of mitochondrial DNA was amplified using primers L15926 (5′-TCAAAGCTTACACCAGTCTTGTAAACC-3′) and H639 (5′-GGGTGATGTGAGCCCGTCA-3′). PCR amplicons were sequenced in both forward and reverse directions using primers L15965 (5′-CAAGGACAAATCAGAGAA-3′) and H11 (5′-GTGGTTAATAGGGTGATAG-3′). Traditional Sanger sequences were aligned and edited with Sequencher® v. 5.0 (Gene Codes Corporation, Ann Arbor, MI, USA; http://www.genecodes.com). Polymorphisms were scored relative to the revised Cambridge reference sequence (rCRS)[69]. Haplogroups were initially predicted from HVS I sequences using known reference genomes[30, 41, 59, 70-73]. Assignments were subsequently confirmed using Taqman and Restriction Fragment Length Polymorphism (RFLP) assays (as described in Tumonggor and colleagues[6]). For mtDNA comparative analysis, we used previously published data from the following geographic regions: surrounding Indonesian islands[6, 59], the Philippines[41], indigenous Taiwanese[74], New Guinea and Melanesian populations[42], and Australian aboriginals[45].

Polymorphic sites from the non-recombining portion of the human Y chromosome (NRY) included a set of 85 binary markers published previously[14, 40] and three newly discovered mutations: C-P355, C-P343 and S-P377. Information about these new mutations is provided in **Supplementary Table 12** and their phylogenetic positions are shown in **Supplementary Figure 2**. We use the mutation-based naming system recommended by the Y Chromosome Consortium[75, 76]. We analyzed 14 Y chromosome short tandem repeats (STRs) (DYS19, DYS385a, DYS385b, DYS388, DYS389I, DYS389CD, DYS390, DYS391, DYS392, DYS393, DYS426, DYS438, DYS457, and TAGA or DYS439+4) using methods described by Redd and colleagues[77]. Additional populations were included for comparative analysis of the Y chromosome data: indigenous Taiwanese (n = 48), Filipinos (n = 47), Papua New Guineans (n = 48), Melanesians (n = 70) and Australians (n = 40).

ADDITIONAL REFERENCES

68 Miller, S. A., Dykes, D. D., Polesky, H. F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl Acids Res* 16: 1215 (1988).

69 Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., Howell, N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23: 147 (1999).

70 Tanaka, M., Cabrera, V. M., Gonzalez, A. M., Larruga, J. M., Takeyasu, T., Fuku, N., Guo, L. J., Hirose, R., Fujita, Y., Kurata, M., Shinoda, K., Umetsu, K., Yamada, Y., Oshida, Y., Sato, Y., Hattori, N., Mizuno, Y., Arai, Y., Hirose, N., Ohta, S., Ogawa, O., Tanaka, Y., Kawamori, R., Shamoto-Nagai, M., Maruyama, W., Shimokata, H., Suzuki, R., and Shimodaira, H. Mitochondrial genome variation in eastern Asia and the peopling of Japan. *Genome Res* 14: 1832-1850 (2004).

71 Kong, Q. P., Yao, Y. G., Sun, C., Bandelt, H. J., Zhu, C. L., Zhang, Y. P. Phylogeny of east Asian mitochondrial DNA lineages inferred from complete sequences. *Am J Hum Genet* 73: 671-676 (2003).

72 Friedlaender, J., Gentz, F., Friedlaender, F., Kaestle, F., Schurr, T., Koki, G., Schanfield, M., McDonough, J., Smith, L., Cerchio, S., Mgone, C., and Merriwether, D. A. Mitochondrial genetic diversity and its determinants in Island Melanesia. In *Papuan Pasts: Cultural, Linguistic and Biological Histories of Papuan-Speaking Peoples*, A. Pawley, R. Attenborough, J. Golson, and R. Hide, eds. Canberra, Pacific Linguistics, pp 693-716 (2005).

73 van Oven, M. Revision of the mtDNA tree and corresponding haplogroup nomenclature. *Proc Natl Acad Sci USA* 107: E38-E39 (2010).

74 Trejaut, J. A., Kivisild, T., Loo, J. H., Lee, C. L., He, C. L., Hsu, C. J., and Li, Z. Y. Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. *PLoS Biol* 3: e247 (2005).

75 Y Chromosome Consortium. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 12: 339-348 (2002).

76 Karafet, T.M., Mendez, F.L., Meilerman, M.B., Underhill, P.A., Zegura, S.L., Hammer, M.F. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 18: 830-838 (2008).

77 Redd, A.J., Agellon, A.B., Kearney, V.A., Contreras, V.A., Karafet, T., Park, H., de Knijff, P., Butler, J.M., Hammer, M.F. Forensic value of 14 novel STRs on the human Y chromosome. *Forensic Sci Int* 130: 97–111 (2002).