Science Advances

AAAS

advances.sciencemag.org/cgi/content/full/2/8/e1600633/DC1

Supplementary Materials for

Genomic analysis reveals hidden biodiversity within colugos, the sister group to primates

Victor C. Mason, Gang Li, Patrick Minx, Jürgen Schmitz, Gennady Churakov, Liliya Doronina, Amanda D. Melin, Nathaniel J. Dominy, Norman T.-L. Lim, Mark S. Springer, Richard K. Wilson, Wesley C. Warren, Kristofer M. Helgen, William J. Murphy

Published 10 August 2016, *Sci. Adv.* **2**, e1600633 (2016)
DOI: 10.1126/sciadv.1600633

The PDF file includes:

- fig. S15. Biplot of PC3 and PC4 from PCA of colugo X chromosome variants.
- fig. S16. Biplot of PC4 and PC5 from PCA of colugo X chromosome variants.
- fig. S17. Biplot of PC5 and PC6 from PCA of colugo X chromosome variants.
- fig. S18. Biplot of morphometric PCA with males and females, and no dwarf specimens removed.
- fig. S19. Biplot of colugo morphometric PCA, with male, female, and dwarf specimens, where all measurements are normalized by body size.
- table S1. Mammals used in phylogenetic analyses.
- table S2. Presence/absence of phylogenetically informative retroposon markers and species distribution.
- table S3. Location of diagnostic retroposon markers in the human genome.
- table S4. χ^2 calculation for phylogenetically informative indels.
- table S5. Colugo and select mammal *V1R* and *OR* gene annotation summaries.
- table S6. Nuclear capture efficiency.
- table S7. Adapter blocking oligos.
- table S8. Off-target mtDNA assembly statistics.
- table S9. External fossil calibrations used to calculate divergence time between the two colugo genera.
- table S10. Principal component loadings for morphometric PCA using males and females with no dwarf individuals.
- table S11. Principal component loadings for morphometric PCA using males, females, and dwarf individuals, and all measurements were normalized for body size.
- table S12. mtDNA maximum composite likelihood + Γ distance matrix calculated between seven Sunda colugo groups and five Philippine groups.
- table S13. X chromosome maximum composite likelihood + Γ distance matrix calculated between seven Sunda colugo groups and three Philippine groups.
- table S14. Between-group mtDNA composite likelihood + Γ distance matrix for Sunda colugos from Thailand, Peninsular Malaysia, and Sumatra, and their satellite islands.
- table S15. Within-group mtDNA maximum composite likelihood + Γ distance matrix.
- table S16. Between-group mtDNA composite likelihood + Γ distance matrix for 10 Sunda colugo groups and 5 Philippine groups.
- table S17. BPP species estimation results for Sunda colugos.
- table S18. BPP species estimation results for Philippine colugos.
- table S19. Principal component loadings for X chromosome genetic variants.
- Legends for database S1 to S7
- References (*75–80*)

Other Supplementary Material for this manuscript includes the following:
(available at advances.sciencemag.org/cgi/content/full/2/8/e1600633/DC1)

- database S1 (Microsoft Excel format). Dermopteran positively selected genes.
- database S2 (Microsoft Excel format). Dermopteran WebGestalt results.

- database S3 (Microsoft Excel format). Ancestral primate positively selected genes.
- database S4 (Microsoft Excel format). Ancestral primate WebGestalt results.
- database S5 (Microsoft Excel format). Museum specimen information.
- database S6 (Microsoft Excel format). Primers used to amplify nuclear DNA capture probes.
- database S7 (Microsoft Excel format). Craniodental morphometric measurements for 125 Sunda colugos.

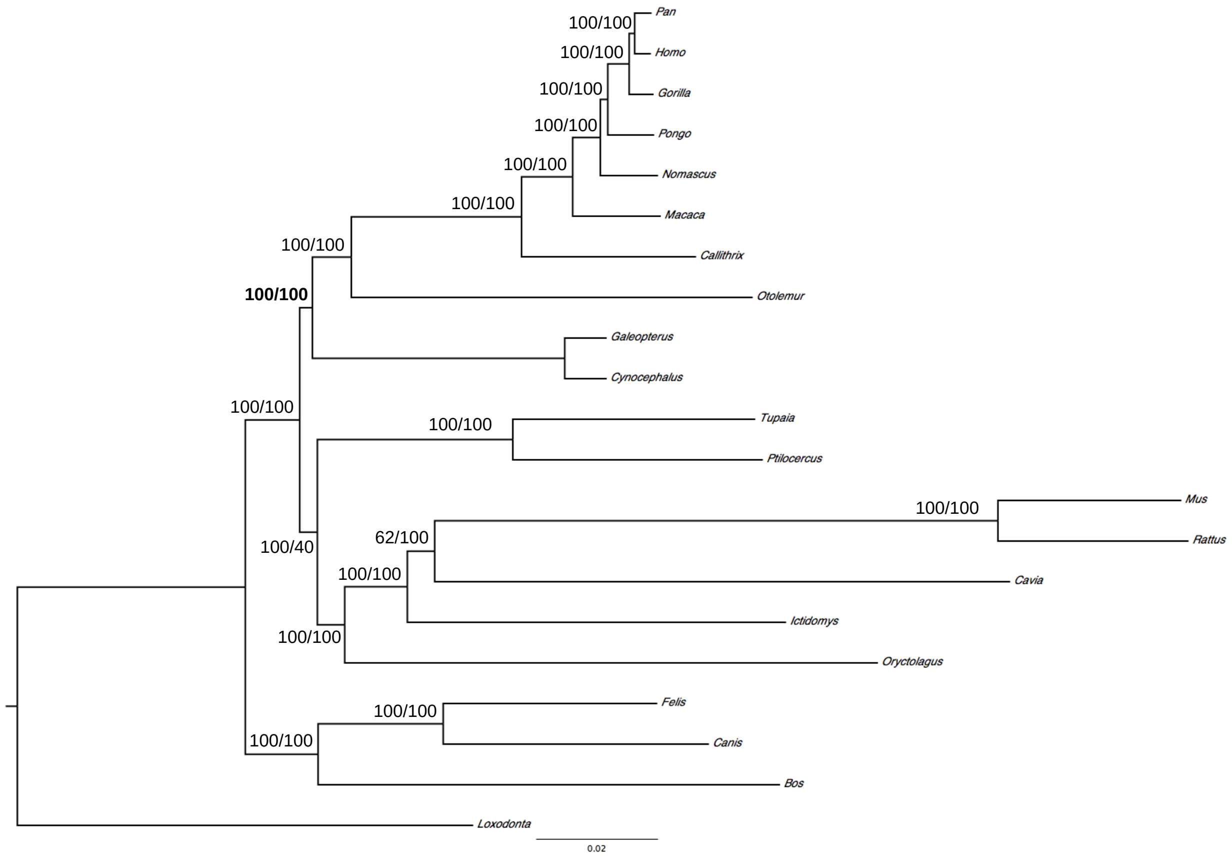

fig. S1. Maximum likelihood tree based on a concatenated alignment of 634 1:1 orthologous CDS, using all codon positions. ML bootstrap values (before the slash) are based on 1000 pseudoreplicates. Identical results were obtained when the analysis was restricted to first, second, or first+second codon positions, or amino acid sequences. SVD bootstrap values follow the slash, and are based on 500 pseudoreplicates. The SVD tree favored placement of treeshrews basal to other Euarchontoglires taxa (55%) and only weakly supported Euarchonta (5%).

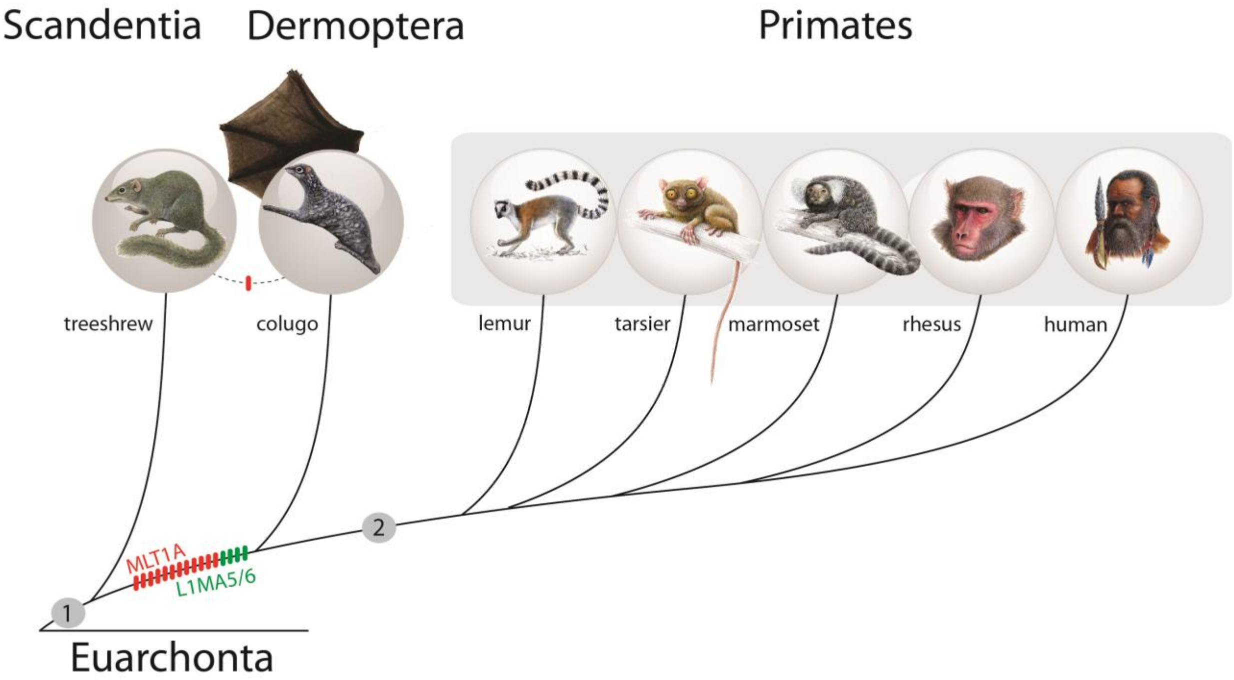

fig. S2. Phylogenetic tree of Euarchonta. 1) The monophyly of Euarchonta (Scandentia, Dermoptera, and Primates) is supported by one MLT1A and four MSTD elements (*11*), and 2) the monophyly of primates is confirmed by synapomorphic retrotransposon insertions (*6*). The close relationship of primates and colugo is supported by 12 MLT1A (red) and 4 L1MA5/6 (green) elements. One MLT1A insertion supports a sister-group relationship of colugo and tree shrew (at the dotted line between the two). This conflicting pattern is likely the result of incomplete lineage sorting.

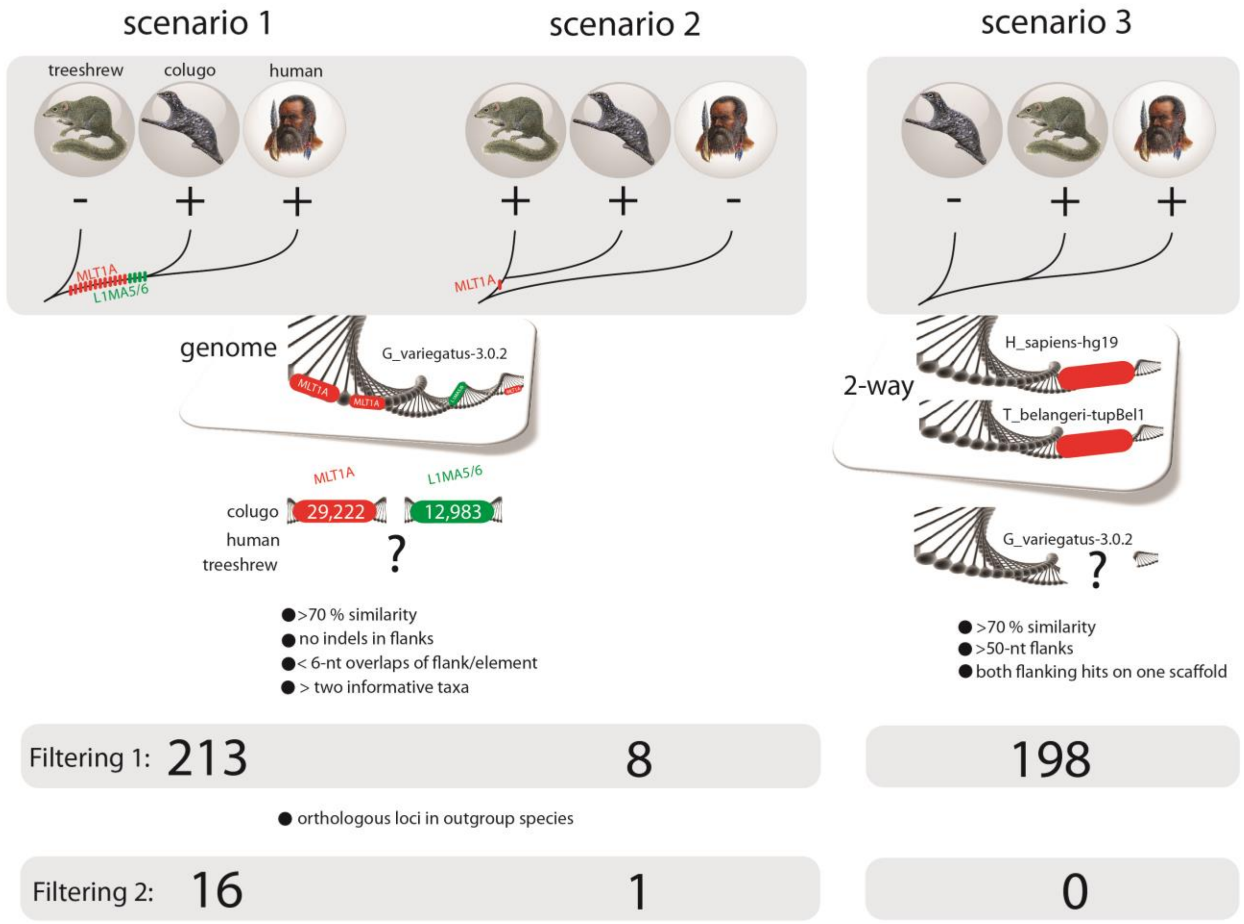

fig. S3. Description of the retroposon insertion screening strategy. The three possible phylogenetic scenarios for the relatedness of colugos are presented at the top. Scenario 1 describes a close relationship between the colugo and primates, the second places the treeshrew and colugo on one branch, and scenario 3 proposes a close relationship between the treeshrew and human. MLT1A (red), L1MA5/6 (green).

fig. S4. Phylogenetically informative indels supporting (P1-P15) Primatomorpha, (S1-S5) Sundatheria, and (PS1-PS5) Primates + Scandentia. (*Following pages*).

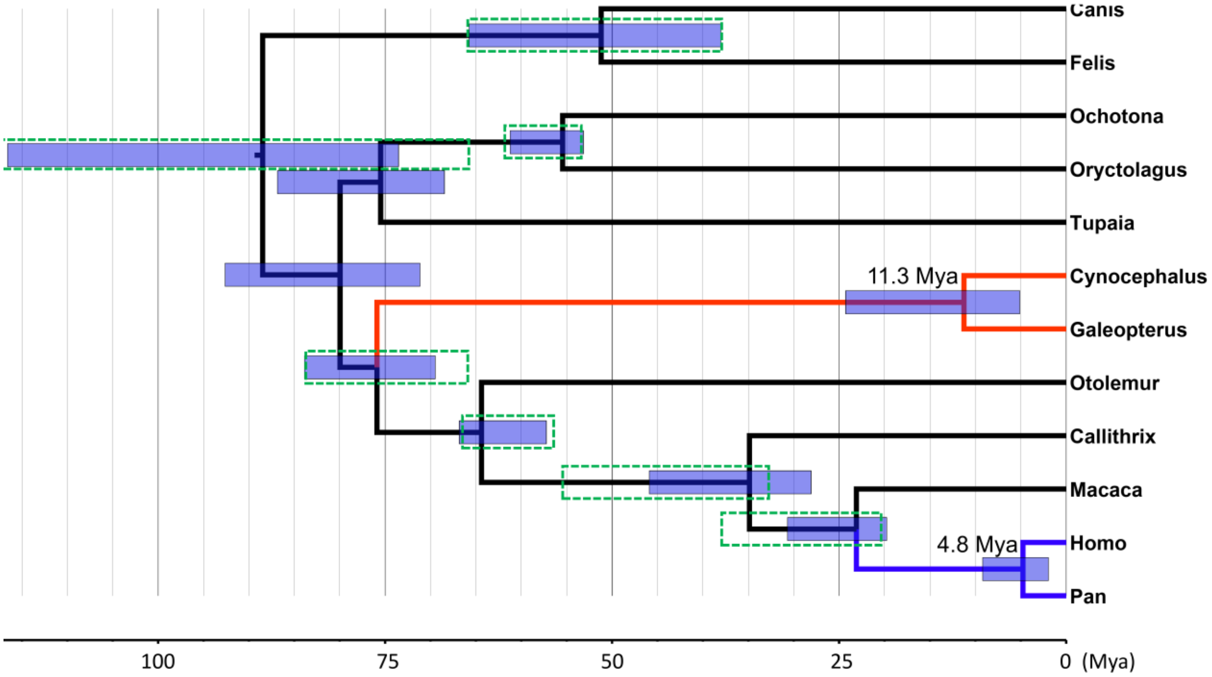

fig. S5. Time-tree showing seven external fossil calibration ranges as green boxes (table S9). Topology was based on the maximum likelihood 'best-tree'. The human/chimpanzee bifurcation (blue) node was left uncalibrated to serve as a control for terminal node age estimates that lack calibrations. Blue boxes are calculated 95% confidence intervals.

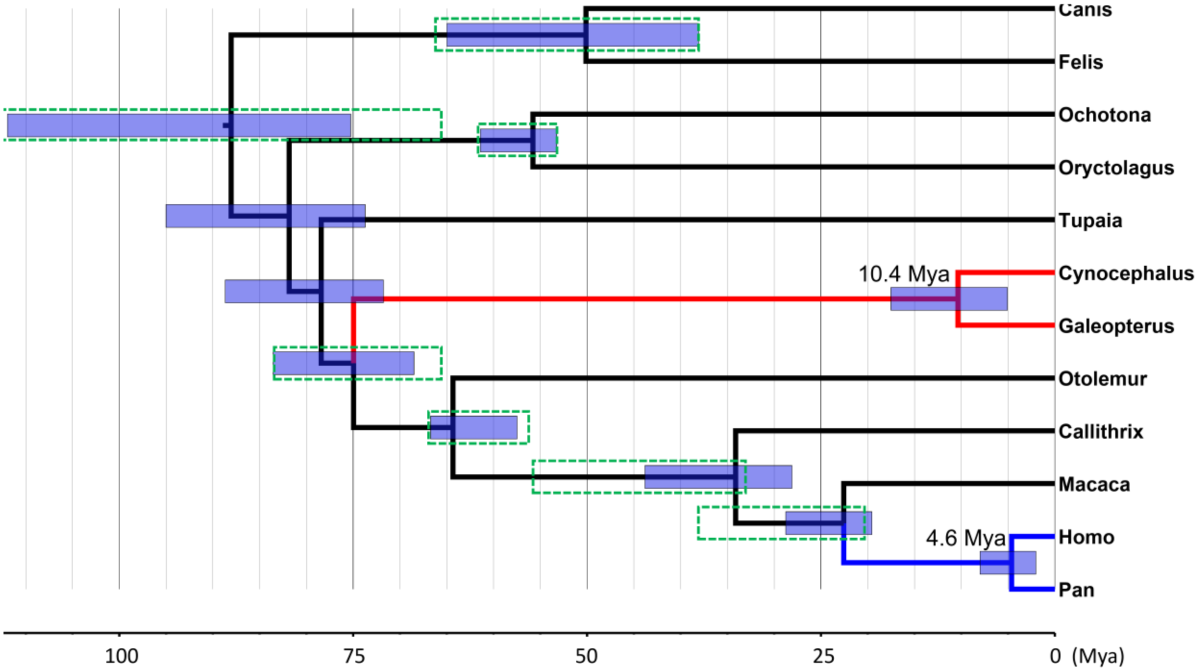

fig. S6. Time-tree with showing seven external fossil calibrations as green boxes (table S9). Euarchontan monophyly was enforced on the maximum likelihood 'best-tree' topology. The human/chimpanzee node (blue) was left to serve as a control for terminal node age estimates that lack calibrations. Blue boxes are calculated 95% confidence intervals.

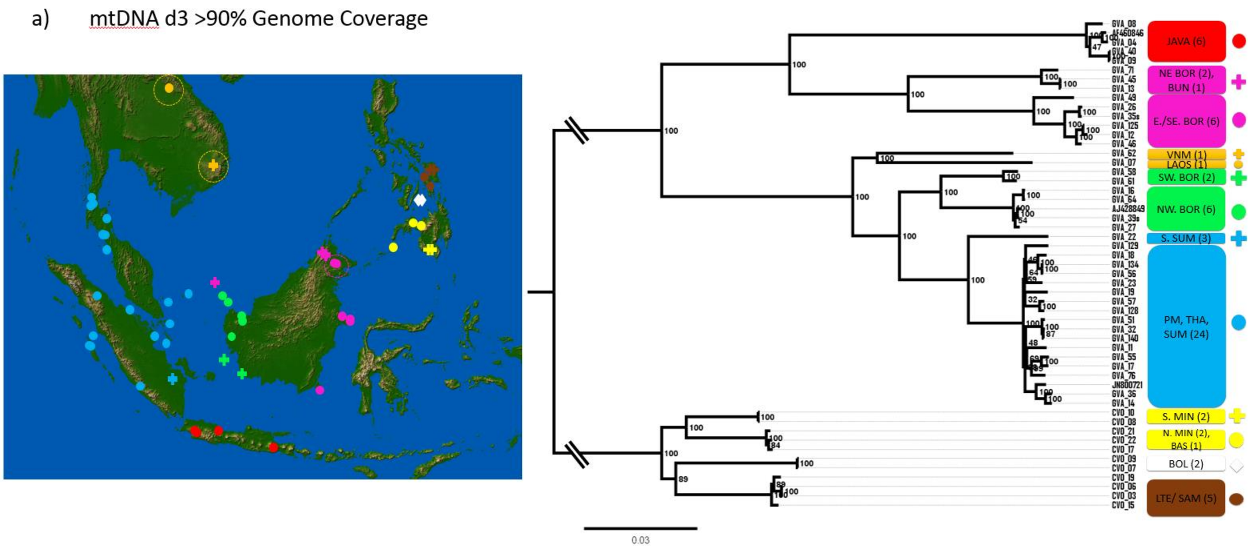
a) mtDNA d3 >90% Genome Coverage

fig. S7. Maximum likelihood mtDNA tree with 1000 bootstrap iterations, 53 taxa, and greater than 90% mitogenome coverage. Labels for geographic area are next to sample numbers. Abbreviations: Vietnam=VNM, Thailand=THA, Peninsular Malaysia=PM, Sumatra=SUM, Borneo=BOR, Bunguran=BUN (GVA_71), Basilan=BAS (CVO_21), Mindanao=MIN, Bohol=BOL, Leyte=LTE, Samar=SAM, Dinagat=DIN.

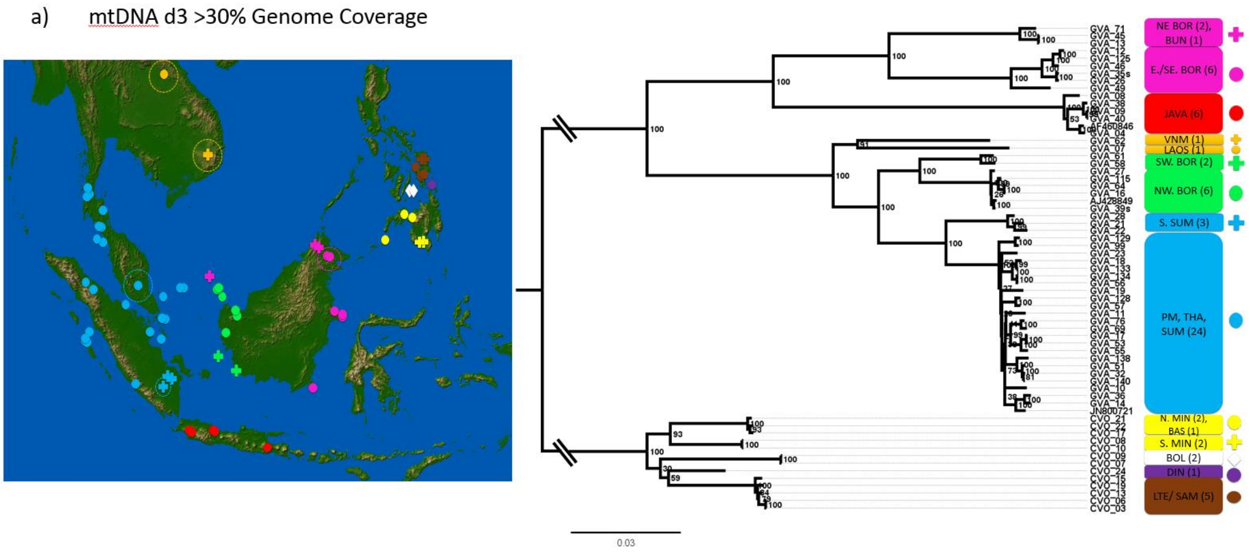

fig. S8. Maximum likelihood mtDNA tree with 1000 bootstrap iterations 65 taxa, and greater than 30% mitogenome coverage. Labels for geographic area are next to sample numbers. Abbreviations: Vietnam=VNM, Thailand=THA, Peninsular Malaysia=PM, Sumatra=SUM, Borneo=BOR, Bunguran=BUN (GVA_71), Basilan=BAS (CVO_21), Mindanao=MIN, Bohol=BOL, Leyte=LTE, Samar=SAM, Dinagat=DIN.

a) All Nuclear d3 > 40% Target Probe Coverage

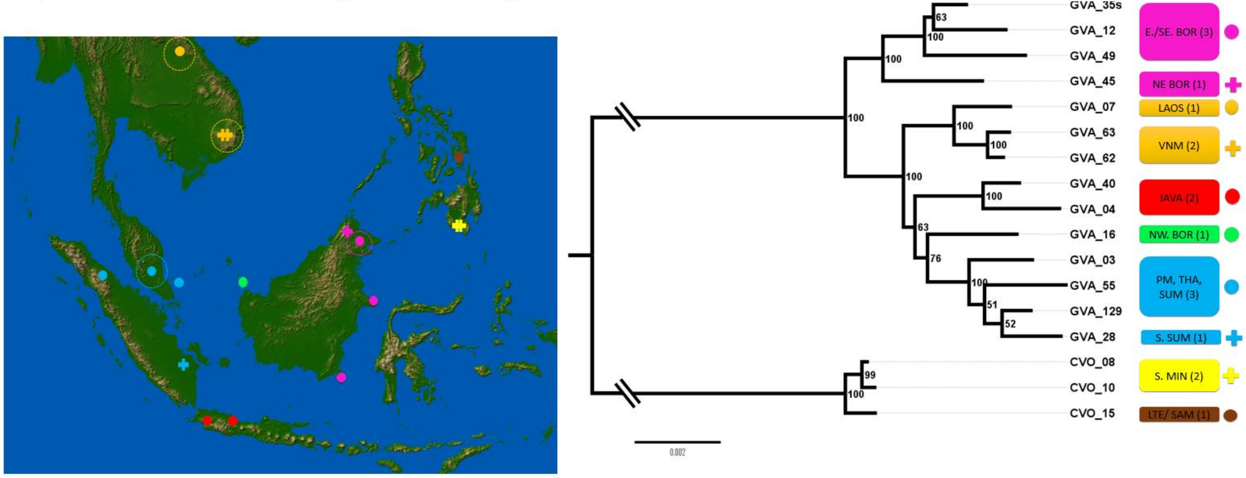

fig. S9. Maximum likelihood all-nuclear (biparental+Chr Y) tree with 1000 bootstrap iterations with sequence depth ≥3. Labels for geographic area are next to sample numbers. Abbreviations: Vietnam=VNM, Thailand=THA, Peninsular Malaysia=PM, Sumatra=SUM, Borneo=BOR, Mindanao=MIN, Leyte=LTE, Samar=SAM, Dinagat=DIN.

a) Biparental d3 > 40% Target Probe Coverage

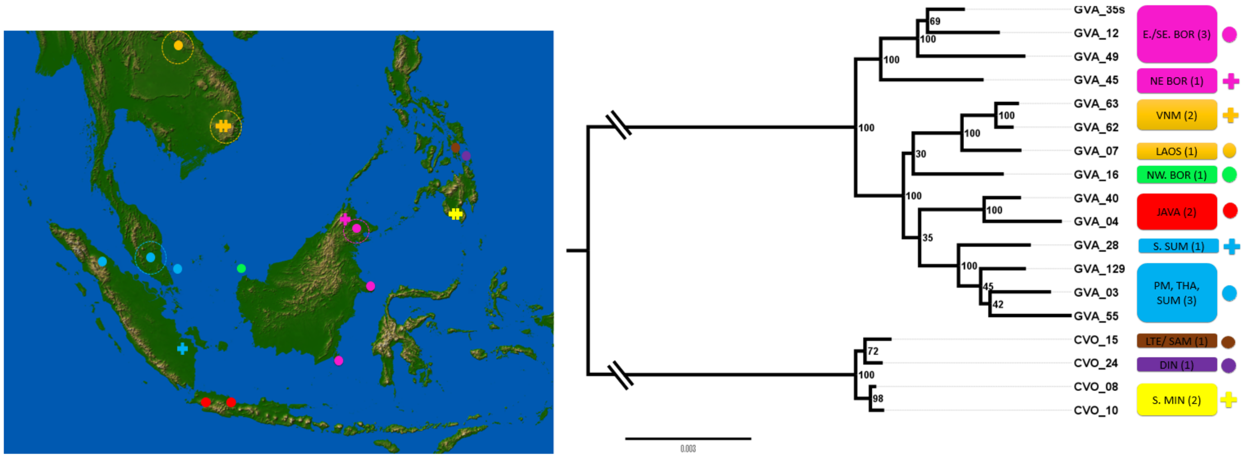

fig. S10. Maximum likelihood biparental locus tree with 1000 bootstrap iterations with minimum sequence depth $\geq$2. Abbreviations: Vietnam=VNM, Thailand=THA, Peninsular Malaysia=PM, Sumatra=SUM, Borneo=BOR, Mindanao=MIN, Leyte=LTE, Samar=SAM, Dinagat=DIN.

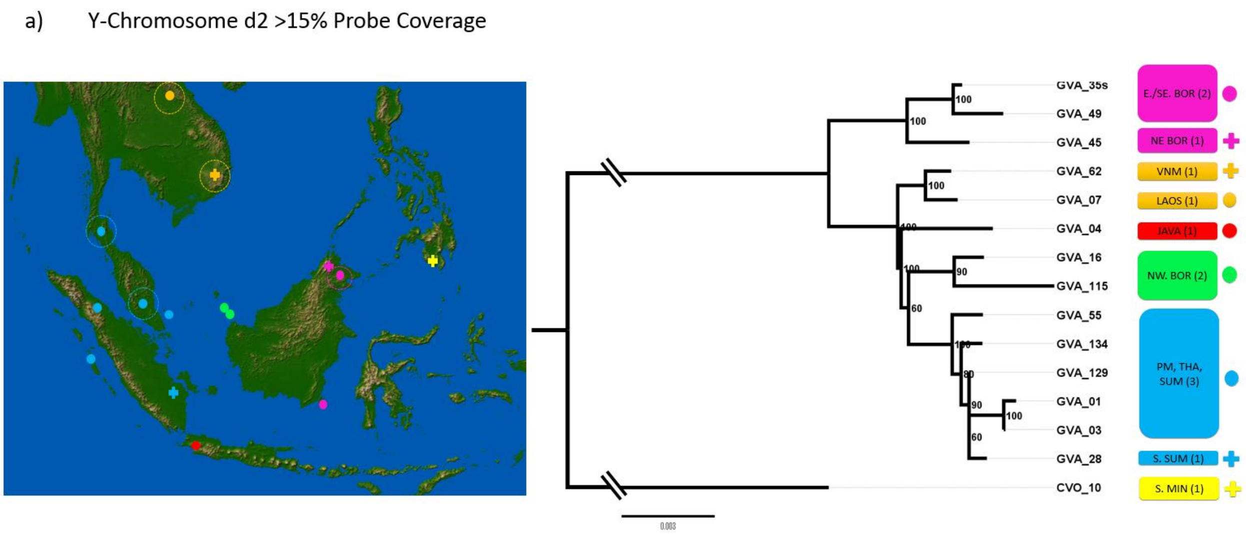

fig. S11. Maximum likelihood Y chromosome tree (minimum depth ≥2) with 1000 bootstrap iterations. Abbreviations: Vietnam=VNM, Thailand=THA, Peninsular Malaysia=PM, Sumatra=SUM, Borneo=BOR, Bunguran=BUN, Basilan=BAS, Mindanao=MIN, Bohol=BOL, Leyte=LTE, Samar=SAM, Dinagat=DIN.

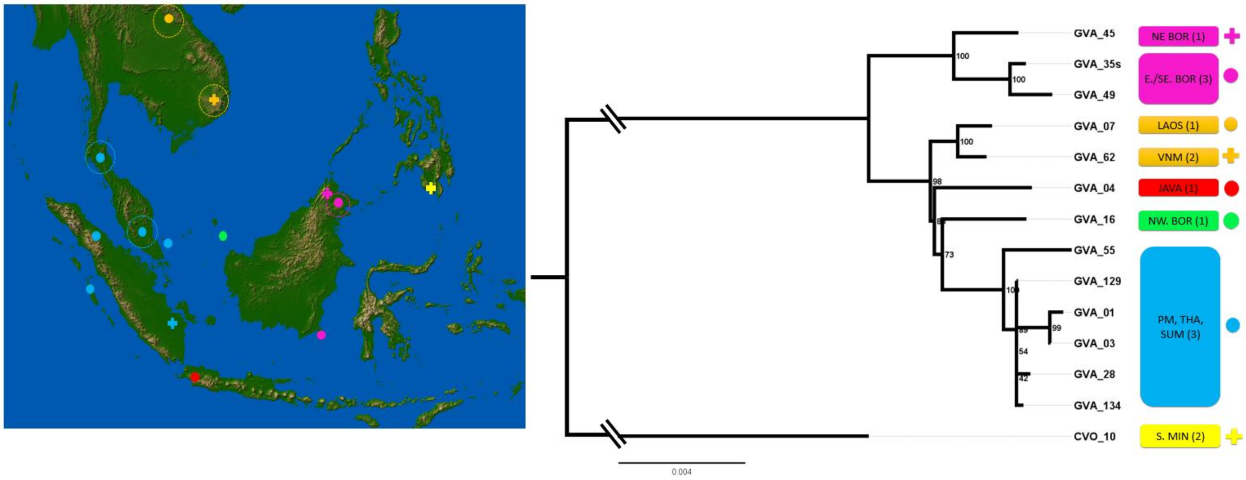

fig. S12. Maximum likelihood Y chromosome tree (minimum depth ≥ 3) with 1000 bootstrap iterations. Abbreviations: Vietnam=VNM, Thailand=THA, Peninsular Malaysia=PM, Sumatra=SUM, Borneo=BOR, Mindanao=MIN.

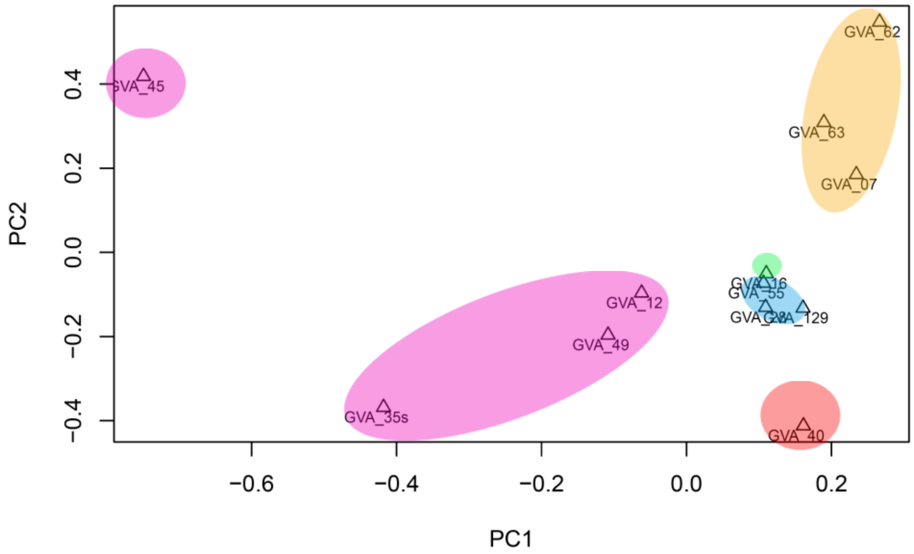

fig. S13. Principal component analysis (PC1 versus PC2) of X chromosome variation in Sunda colugos. Colors follow geographic clusters in Fig. 3.

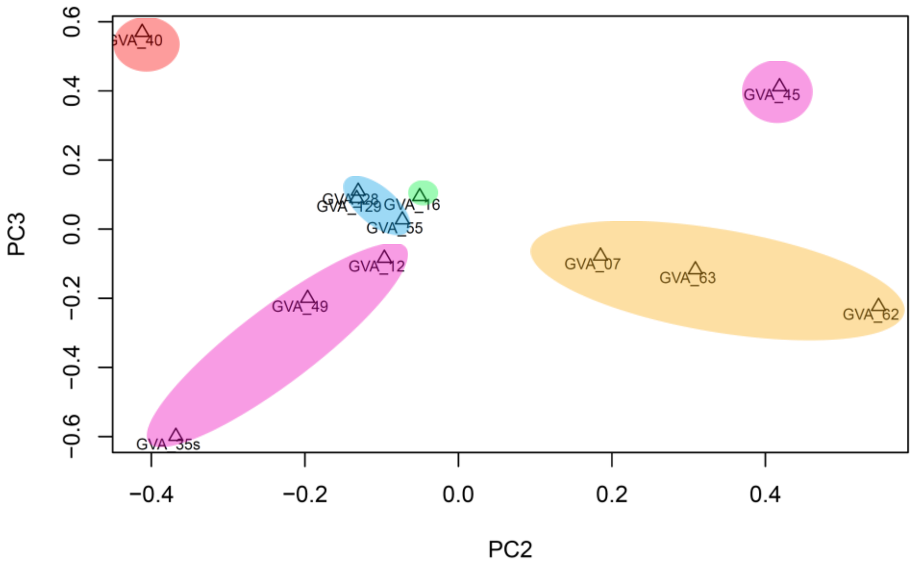

fig. S14. Principal component analysis (PC2 versus PC3) of X chromosome variation in Sunda colugos. Colors follow geographic clusters in Fig. 3.

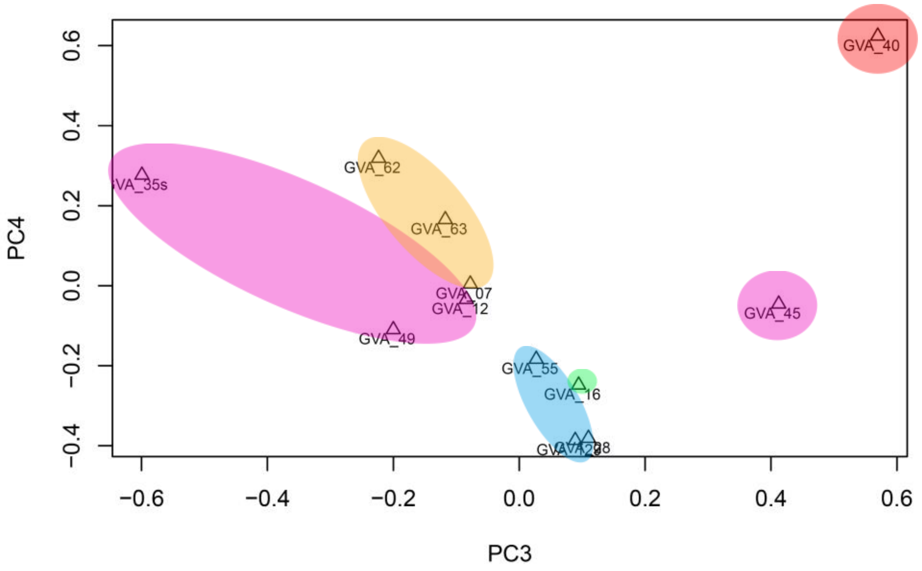

fig. S15. Principal component analysis (PC3 versus PC4) of X chromosome variation in Sunda colugos. Colors follow geographic clusters in Fig. 3.

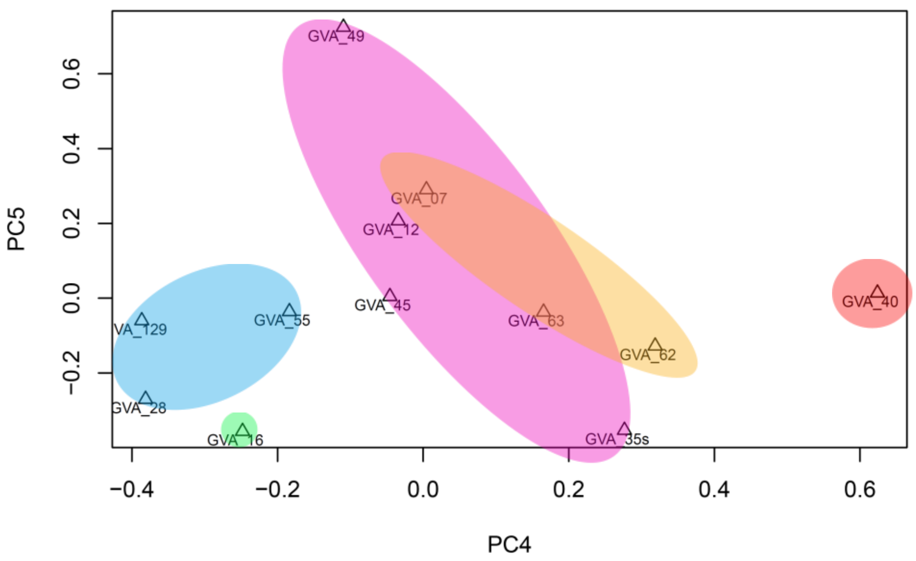

fig. S16. Principal component analysis (PC4 versus PC5) of X chromosome variation in Sunda colugos. Colors follow geographic clusters in Fig. 3.

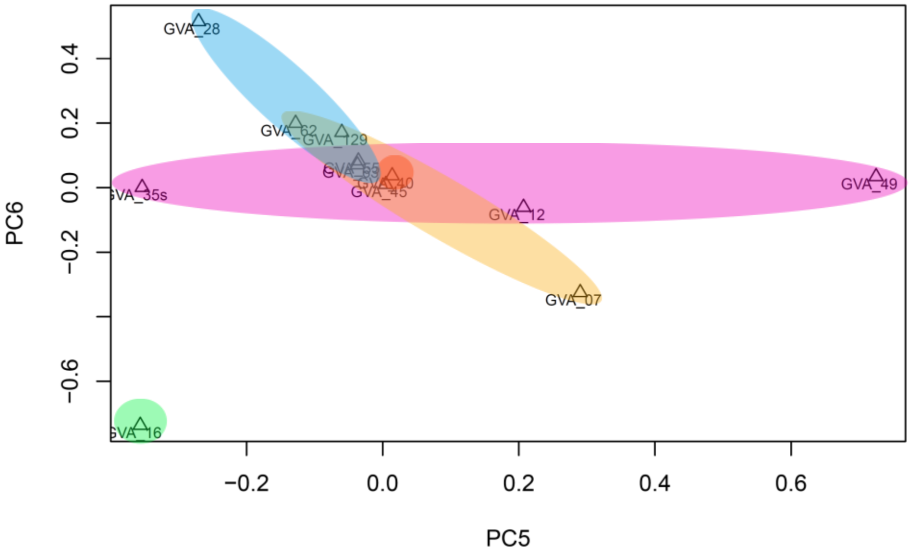

fig. S17. Principal component analysis (PC5 versus PC6) of X chromosome variation in Sunda colugos. Colors follow geographic clusters in Fig. 3. West Borneo (green) is well separated in this comparison.

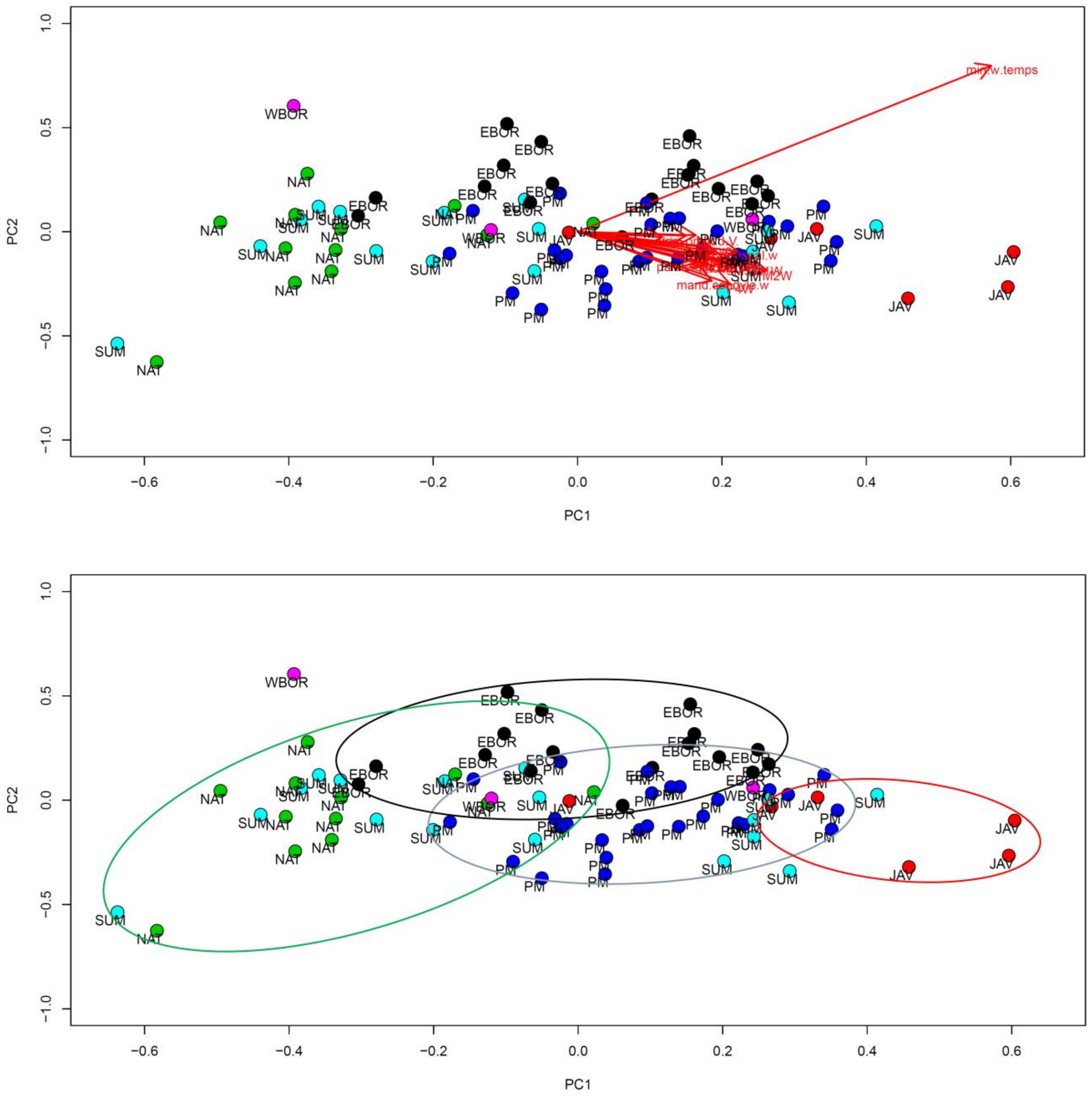

fig. S18. PCA analysis (singular value decomposition) of 19 craniodental measurements for male and female Sunda colugos. Dwarf individuals were not included. Top: Red arrows are the 19 variable vectors. Bottom: same plot as above except the red arrows have been removed for clarity. Ellipses encircle colugos from different regions. Each point represents one individual. 45% of the variation is explained by PC1 and 29% by PC2.

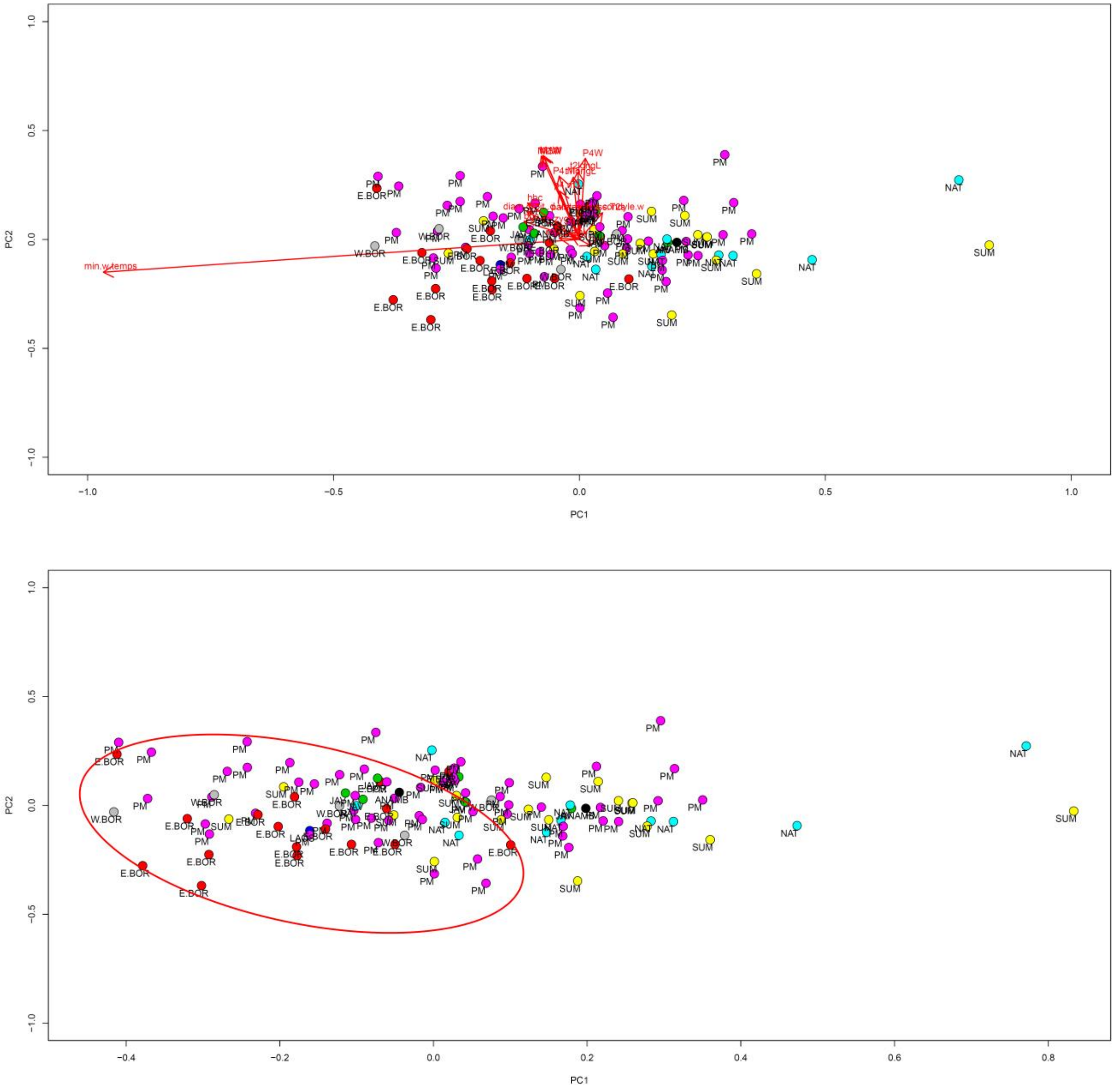

fig. S19. PCA analysis (singular value decomposition) of 19 craniodental measurements for male and female Sunda colugos, with dwarf individuals included for comparison to fig. S13. Top: Red arrows are variable vectors. Bottom: same plot as above except the red arrows have been removed and the image enlarged for clarity. Each point is an individual. Red ellipse encircles East Bornean individuals. All of the craniodental measurements are normalized for body size. 44% of the variation is explained by PC1 and 20% by PC2.

table S1. Eutherian mammals used in phylogenetic comparisons.

Common Name	**Species Name**	**Sequence Origin**
Human	*Homo sapiens*	Ensembl v.79
Chimpanzee	*Pan troglodytes*	Ensembl v.79
Gorilla	*Gorilla gorilla*	Ensembl v.79
Orangutan	*Pongo abelii*	Ensembl v.79
Macaque	*Macaca fascicularis*	Ensembl v.79
Marmoset	*Callithrix jacchus*	Ensembl v.79
Bushbaby	*Otolemur garnetti*	Ensembl v.79
Sunda colugo	*Galeopterus variegatus*	G_variegatus-3.0.2
Philippine colugo	*Cynocephalus volans*	14x ref. assembly-this study
Chinese treeshrew	*Tupaia belangeri chinensis*	(*53*)
Pen-tailed treeshrew	*Ptilocercus lowii*	5x ref. assembly-this study
Rabbit	*Oryctolagus cuniculus*	Ensembl v.79
13-lined ground squirrel	*Ictidomys tridecemlineatus*	Ensembl v.79
Guinea Pig	*Cavia porcellus*	Ensembl v.79
Rat	*Rattus rattus*	Ensembl v.79
Mouse	*Mus musculus*	Ensembl v.79
Cow	*Bos taurus*	Ensembl v.79
Cat	*Felis catus*	Ensembl v.79
Dog	*Canis familiaris*	Ensembl v.79
African elephant	*Loxodonta africana*	Ensembl v.79

table S2. Presence/absence of phylogenetically informative retroposon markers and species distribution.

	1	2	3	4	56	61	74	280	288	346	348	398	408	450	576	582	252
	L1MA				**MLT1A/MSTD**												
	A6	**A6**	**A6**	**A6**	**D**	**1A0**	**D**	**D**	**1A0**	**D**	**D**	**D**	**D**	**D**	**D**	**1A0**	**D**
Human	+	+	+	+	+	+	+	+	+dup	+	+	+	+	+	+	+	-
Chimp	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-
Baboon	+	+	+	+	+	+	+	+	+	+	+	+	+	n/s	+1/2	+	-
Rhesus	+	+	+	+	+	+	+	+	+	+	n/s	+	+	n/a	+	+	-
Marmoset	+	+	+	+	+	+	+	+	+	+I	n/a	+I	+	n/a	+	n/a	-
Tarsier	n/s	+	+	+I	+	+	+1/2	+	+1/2	+1/2	+	+1/2	+	+	+	+	n/a
Bushbaby	n/s	+	+	n/a	+	n/a	+1/2	+	n/a	+1/2	+	n/a	+	+	+	+	-
Mouse lemur	+	+	+	+	n/a	+	+	+	+1/2	+	+	+	+1/2	+1/2	+	+1/2	-
Colugo	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Treeshrew	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+dup
Pika	n/a	d	-	-	-	-	-	n/a	-	-	n/a	n/a	n/a	-	n/a	n/a	n/a
Rabbit	-	-	d	-	-	-	-	n/a	-	-	-	n/a	-	n/a	-	n/a	n/a
Mouse	n/a	n/a	n/a	n/a	n/a	n/a	-	-	n/a	-	-	n/a	n/a	-	n/a	n/a	n/a
Kangaroo rat	-	n/s	n/a	n/a	n/a	n/a	-	-	n/a	-	n/a	n/a	n/s	n/a	n/a	n/a	n/a
Guinea pig	n/a	-	n/a	n/a	-	d	-	n/a	-	-	-	-	-	-	d	n/a	-
Squirrel	n/s	n/s	n/a	n/a	-	-	-	-	n/s	-	-	n/s	n/a	n/s	n/s	n/a	n/a
Dog	n/a	-	-	n/a	-	-	-	-	-	-	-	n/a	-	-	-	-	-
Cat	-	n/s	n/a	n/a	n/a	n/s	-	-	-	n/a	n/s	n/a	-	-	-	n/s	n/a
Megabat	-	-	-	-	-	-	-	-	n/s	-	-	n/a	-	n/a	-	n/a	-
Horse	n/a	-	-	-	-	-	-	-	-	-	-	n/a	-	-	-	-	-
Elephant	n/a	n/a	-	-	n/a	n/a	-	-	-	-	-	n/a	n/a	-	n/a	n/a	n/a
Rock hyrax	-	n/s	n/s	n/s	n/a	n/a	n/a	n/a	n/s	-	-	n/a	n/a	-	n/a	n/s	-
Sloth	n/s	-	n/s	-	n/a	n/s	n/a	n/a	n/s	-	n/s	n/a	n/s	-	-	n/s	n/a
Armadillo	n/a	-	-	-	n/s	n/s	-	n/a	n/s	-	d	n/a	n/a	n/s	n/a	n/s	n/s
DRs	9 nt	11nt	11nt	-	7 nt	5 nt	4 nt	6 nt	5 nt	5 nt	-	-	5 nt	5 nt	5 nt	5 nt	5 nt

n/s=sequence information not available (gaps, scaffold boundary), n/a=alignment not available, d= deletion in the region of insertion with part of flank (only for absence), +I=additional insertion, +1/2=only one flank is supported (only for presence), dup=duplicated loci.

table S3. Location of diagnostic retroposon markers in the human genome (GRCh37/hg19).

Marker	Chr	Strand	From	To		Position	Gene	orthologous
1	**4**	+	169239940	169240432	**L1MA6**	Intergenic	Not	Gv+Hs
2	**2**	**-**	157141497	157141918	**L1MA6**	Intergenic	Not	Gv+Hs
3	**8**	+	121978322	121978810	**L1MA6**	Intergenic	Not	Gv+Hs
4	**3**	**-**	194397033	194396619	**L1MA6**	Intergenic	Not	Gv+Hs
56	**1**	+	57414697	57415317	**MSTD**	**Intron**	**C8B**	Gv+Hs
61	**7**	+	81125867	81126402	**MLT1A0**	Intergenic	Not	Gv+Hs
74	**8**	+	78898876	78899419	**MSTD**	Intergenic	Not	Gv+Hs
252c	**5**	+	17318796	17318992	**MSTD**	Intergenic	Not	Gv+Tb
280	**3**	+	85314313	85314946	**MSTD**	**Intron**	**CADM2**	Gv+Hs
288dup	**2**	+	112157798	112158410	**MLT1A0**	**Intron**	**Loc541471**	Gv+Hs
	2	**-**	87850105	87849493		**Intron**	**RMND5A**	
346	**7**	+	135186465	135187059	**MSTD**	**Intron**	**CNOT4**	Gv+Hs
348	**10**	+	20888530	20889206	**MSTD**	Intergenic	Not	Gv+Hs
398	**5**	+	32721370	32722324	**MSTD**	**Intron**	**NPR3**	Gv+Hs
408	**14**	+	46562841	46563514	**MSTD**	Intergenic	Not	Gv+Hs
450	**1**	+	176569974	176570619	**MSTD**	**Intron**	**PAPPA2**	Gv+Hs
576	**1**	+	215140155	215140814	**MSTD**	Intergenic	Not	Gv+Hs
582	**5**	+	27554295	27554924	**MLT1A0**	Intergenic	Not	Gv+Hs

table S4. Chi-square calculation for phylogenetically informative indels (*43*).

	Observed	**Expected**	**chi-square**	
Primatomorpha	19	9	11.1	
Sundatheria	4	9	2.8	
Primates+Tupaia	4	9	2.8	
Sum(values)	27	27	16.7	
***P*-value**			0.000045	df = 1

table S5. Size of Olfactory receptor (*OR*) and vomeronasal class 1 receptor (*V1R*) gene families found in colugo and 16 select mammalian genomes.

Node	*OR* gene families			*V1R* gene families		
	intact genes	**partial/ pseudo genes**	**Repertoire size**	**intact genes**	**Partial/ pseudo genes**	**Repertoire size**
human	396	425	821	5	115	120
orangutan	296	525	821	5	114	119
rhesus	319	321	640	0	51	51
marmoset	366	258	624	8	42	50
galago	356	585	941	60	49	109
mouse lemur	361	619	980	82	109	191
colugo	518	1480	1998	46	107	153
Chinese treeshrew	969	1396	2365	23	55	78
mouse	1127	202	1329	187	121	308
rat	1194	438	1632	106	66	172
rabbit	751	278	1029	160	132	292
cat	679	330	1009	21	47	68
dog	811	246	1057	8	29	37
horse	1063	1511	2574	36	47	83
pig	1254	782	2036	8	28	36
cow	1055	926	1981	40	24	64
opossum	1157	300	1457	100	26	126

table S6. Nuclear capture efficiency and comparison of individual versus multiplex capture results. Multiplex and Individual nuclear capture columns report the percentage of reads mapped to reference probe sequences after removal of single end and paired end duplicates and merging of overlapping read pairs. Indiv. / Multiplex column reports the relative fold-increase in efficiency when performing single individual capture compared to multiple individual capture.

Specimen	Multiplex Nuclear Capture	Individual Nuclear Capture	Indiv. / Multiplex
CVO_06	0.14	0.14	1.0
CVO_07	0.01	0.02	2.0
CVO_08	0.20	0.75	3.8
CVO_15	0.14	0.14	1.0
CVO_22	0.05	0.16	3.2
CVO_24	0.02	0.09	4.5
GVA_12	0.04	0.09	2.6
GVA_17	0.02	0.05	2.5
GVA_22	0.07	0.08	1.1
GVA_26	0.15	0.22	1.5
GVA_27	0.09	0.17	1.9
GVA_32	0.08	0.17	2.1
GVA_40	0.13	0.21	1.6
GVA_49	0.08	0.15	1.9
GVA_61	0.04	0.07	1.8
GVA_64	0.01	0.04	4.0
GVA_75	0.03	0.07	2.3
GVA_76	0.03	0.09	3.0
GVA_129	0.10	0.10	1.0
Average:	0.08	0.15	2.2

table S7. Adapter blocking oligos, modified slightly from (*48*).

Adapter Blocked	Blocking Oligo Sequence
#1 Blocks P5:	5' TGTAGATCTCGGTGGTCGCCGTATCATT-P 3'
#2 Blocks Rd1 SP:	5' AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-P 3'
#3 Blocks Rd2 SP:	5' GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-P 3'
#4 Blocks P7:	5' CAAGCAGAAGACGGCATACGAGAT-P 3'
#5 Blocks P5 (Complement)	5' AATGATACGGCGACCACCGAGATCTACAC-P 3'
#6 Blocks Rd1 SP (Complement)	5' ACACTCTTTCCCTACACGACGCTCTTCCGATCT-P 3'
#7 Blocks Rd2 SP (Complement)	5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-P 3'
#8 Blocks P7 (Complement)	5' ATCTCGTATGCCGTCTTCTGCTTG-P 3'

table S8. mtDNA reference assembly statistics from off-target nuclear capture reads.

Individual	Reads Mapped	Total Reads	Average Depth	% of Reference Bases Covered
CVO_06	1683	3315220	11.81	95.15
CVO_07	5683	9655652	36.07	96.64
CVO_08	3758	6397053	30.13	95.93
CVO_10	4643	10610200	23.89	95.81
CVO_13	146	1740932	0.75	43.55
CVO_15	4587	8669257	24.08	96.85
CVO_21	2096	4369074	8.24	95.44
CVO_22	12524	10240636	64.23	99.40
CVO_24	647	9673217	4.67	89.58
GVA_07	5292	6220128	39.63	98.85
GVA_09	1059	29808549	4.79	66.96
GVA_12	12411	14150951	81.37	99.95
GVA_16	6509	12224335	23.31	99.73
GVA_17	8125	6930959	54.78	99.97
GVA_18	6171	14485997	30.26	99.85
GVA_22	8263	3584858	21.87	99.70
GVA_27	364	1493420	1.41	73.37
GVA_28	650	7673989	2.12	82.41
GVA_32	2863	5050725	13.53	98.73
GVA_35s	2669	11384773	11.71	99.18
GVA_39s	13891	6333908	63.6	99.98
GVA_40	7273	8461912	41.68	99.55
GVA_45	16923	13081551	94.3	100.00
GVA_49	11396	5959804	91.28	100.00
GVA_52	103	6106934	0.52	37.50
GVA_55	4984	15265857	26.69	99.75
GVA_61	1392	1551730	9.12	97.92
GVA_62	4794	16809740	25.36	96.46
GVA_63	273	7911650	0.92	48.95
GVA_64	17615	11341089	79.11	99.99
GVA_69	2137	17168303	6.77	96.57
GVA_71	142	3077004	0.35	25.05
GVA_75	337	2478753	1.11	60.87
GVA_76	2314	9911491	7.18	99.05
GVA_78	36	5498695	0.12	10.72
GVA_106	64	4035807	0.25	18.01
GVA_115	1002	5719343	3.58	92.03
GVA_121	20	3982512	0.06	6.06
GVA_129	20842	12079124	162.86	99.99
GVA_133	570	5284258	2.36	76.68
GVA_134	3530	13585155	22.13	99.92
Mean:	4873	8617672	28.89	84.77
Median:	2863	7673989	21.87	96.85

table S9. External fossil calibrations used to calculate the divergence time between colugo genera. In some cases original calibrations have been updated to correspond to revisions in the Stratigraphic Code.

Pair of Taxa	Lower Bound (mya)	Upper Bound (mya)	Reference
Dog-Cat (*Canis-Felis)*	38	66	*75*
Rabbit-Pika (*Oryctolagus-Ochotona*)	53.7	61.6	Minimum based on Ypresian tarsals of crown lagomorph (*76*) with age of 53.7 Ma (*77*); maximum following (*75*)
Macaque-Marmoset (*Macaca-Callithrix*)	28.1	56	*75*
Macaque-Human (*Macaca-Homo*)	20.55	38	*75*
Primates	56	66	*78*
Primatomorpha	65.2	83.8	Minimum based on *Purgatorius* (oldest crown fossil, (*79*); maximum based on stratigraphic bounding as in Meredith et al. (*9*) with two stages
Euarchontoglires-Laurasiatheria	65.2	131.5	Minimum based on *Purgatorius* (*79*); maximum based on Benton et al., (*80*)

table S10. Morphometric principal component loadings for males and females with no dwarfs.

	PC1	**PC2**	**PC3**
GSL	0.18932	-0.09443	0.17312
CBL	0.18564	-0.10606	0.19291
ONL	0.20692	-0.14089	0.21679
zyg	0.18509	-0.07422	0.22832
min.w.temps	0.57311	0.797471	-0.06760
P4.M3	0.17849	-0.15021	-0.11203
bbc	0.16389	-0.01663	0.28649
hbc	0.11709	-0.02596	0.06355
P4W	0.21061	-0.24898	-0.13606
M1W	0.24827	-0.16391	-0.21752
M2W	0.26188	-0.18617	-0.20261
M3W	0.22163	-0.16353	-0.10616
max.pal.w	0.22208	-0.09415	0.32784
diam.orbit.d.V.	0.15018	-0.01950	0.11652
palate.across.T2s	0.16431	-0.14817	0.09456
p4.m3	0.18205	-0.13435	-0.01946
mand.condyle.w	0.18609	-0.23571	0.30417
t1longL	0.16841	-0.12435	-0.51270
t2longL	0.18179	-0.15644	-0.35573

table S11. Morphometric principal component loadings for males and females (dwarfs included) after normalizing by body size.

	PC1	PC2	PC3
GSL	-0.01663	0.017175	-0.00442
ONL	0.023339	-0.02825	-0.12767
zyg	-0.05148	0.062697	0.183351
min.w.temps	-0.96799	-0.15036	-0.07986
P4.M3	-0.04110	0.290691	0.060051
bbc	-0.11283	0.083246	0.14261
hbc	-0.10536	0.167393	0.146996
P4W	0.012156	0.372447	0.076299
M1W	-0.07394	0.385925	0.085834
M2W	-0.07768	0.382301	0.140197
M3W	-0.07258	0.381536	0.134362
max.pal.w	-0.04104	0.006042	0.125325
diam.orbit..d.V.	-0.10677	0.127659	0.206659
palate.across.T2s	0.002943	0.120638	-0.04047
p4.m3	-0.04006	0.221703	0.073767
mand.condyle.w	0.04722	0.122476	0.325222
t1longL	-0.01239	0.285843	-0.73167
t2longL	-0.00268	0.311116	-0.37502

table S12. Maximum mtDNA genetic divergence between seven Sunda colugo groups and five Philippine colugo groups, based on the composite likelihood + Γ genetic distance. Abbreviations: Vietnam=VNM, Thailand=THA, Peninsular Malaysia=PM, Sumatra=SUM, Borneo=BOR, Bunguran=BUN, Basilan=BAS, Mindanao=MIN, Bohol=BOL, Leyte=LTE, Samar=SAM, Dinagat=DIN.

	PM THA SUM	**W.BOR**	**JAVA**	**LTE/SAM**	**BOL**	**E.BOR**	**NE.BOR**	**E.MIN**	**W.MIN**	**LAOS**	**VNM**
THA/PM/SUM											
W.BOR	0.058										
JAVA	0.153	0.149									
LTE/SAM	0.216	0.210	0.220								
BOL	0.216	0.210	0.222	0.050							
E.BOR	0.150	0.146	0.121	0.216	0.216						
NE.BOR	0.144	0.141	0.121	0.211	0.212	0.070					
E.MIN	0.212	0.211	0.216	0.048	0.052	0.215	0.209				
W.MIN	0.211	0.209	0.216	0.050	0.053	0.214	0.207	0.035			
LAOS	0.080	0.076	0.151	0.210	0.209	0.145	0.141	0.210	0.207		
VNM	0.070	0.069	0.144	0.196	0.197	0.136	0.133	0.197	0.196	0.064	
DIN	0.209	0.201	0.215	0.032	0.041	0.212	0.208	0.041	0.040	0.202	0.190
					mean	stdev	min	max	median		
Average Sundaic Between Group Divergence					0.117238	0.035714	0.058	0.153	0.136		
Average Philippine Between Group Divergence					0.0442	0.007391	0.032	0.053	0.0445		

table S13. X chromosome sequence divergence between eight Sunda colugo groups and three Philippine colugo groups, based on maximum composite likelihood + Γ genetic distance. Includes all individuals with >40% probe base coverage. Abbreviations: Vietnam=VNM, Thailand=THA, Peninsular Malaysia=PM, Sumatra=SUM, Borneo=BOR, Bunguran=BUN, Basilan=BAS, Mindanao=MIN, Bohol=BOL, Leyte=LTE, Samar=SAM, Dinagat=DIN.

	E.BOR	**THA/PM/SUM**	**VNM**	**NW.BOR**	**S.SUM**	**LTE/SAM**	**NE.BOR**	**E.MIN**	**JAVA**	**DIN**	**SE.BOR**
E.BOR											
THA/PM/SUM	0.0070										
VNM	0.0057	0.0056									
NW.BOR	0.0051	0.0052	0.0039								
S.SUM	0.0051	0.0039	0.0043	0.0034							
LTE/SAM	0.0276	0.0291	0.0288	0.0289	0.0274						
NE.BOR	0.0041	0.0073	0.0059	0.0059	0.0057	0.0285					
E.MIN	0.0267	0.0285	0.0282	0.0280	0.0270	0.0014	0.0276				
JAVA	0.0062	0.0062	0.0052	0.0051	0.0047	0.0299	0.0069	0.0291			
DIN	0.0285	0.0301	0.0299	0.0302	0.0281	0.0010	0.0294	0.0012	0.0312		
SE.BOR	0.0038	0.0096	0.0080	0.0073	0.0071	0.0311	0.0053	0.0296	0.0089	0.0315	
LAOS	0.0060	0.0058	0.0027	0.0047	0.0047	0.0294	0.0066	0.0286	0.0056	0.0302	0.0086
					Mean	Stdev	Min	Max	Median		
Average Sundaic Between Group Divergence					0.0058	0.0016	0.0027	0.0096	0.0057		
Average Philippine Between Group Divergence					0.0012	0.0002	0.0010	0.0014	0.0012		

table S14. Pairwise sequence divergence between colugos from Peninsular Malaysia (PM), Sumatra (SUM) and Thailand (THA), and their satellite islands (blue).

	THA	P.Bakong	P.Tanahbala	P.Aur	P.Siantan	PM	P.Langkawi	P.Bintang	KoAdang	P.Pini	SUM	P. Penuba
THA												
P.Bakong	0.0109											
P.Tanahbala	0.0110	0.0102										
P.Aur	0.0106	0.0109	0.0104									
P.Siantan	0.0097	0.0101	0.0096	0.0078								
PM	0.0113	0.0109	0.0112	0.0107	0.0103							
P.Langkawi	0.0104	0.0106	0.0112	0.0101	0.0094	0.0102						
P.Bintang	0.0099	0.0102	0.0102	0.0032	0.0070	0.0097	0.0094					
KoAdang	0.0084	0.0103	0.0097	0.0114	0.0114	0.0099	0.0112	0.0108				
P.Pini	0.0104	0.0097	0.0030	0.0103	0.0094	0.0100	0.0104	0.0098	0.0105			
SUM	0.0111	0.0102	0.0101	0.0114	0.0106	0.0111	0.0112	0.0106	0.0109	0.0095		
P.Penuba	0.0110	0.0025	0.0103	0.0112	0.0103	0.0117	0.0107	0.0104	0.0105	0.0094	0.0103	
P.Rupat	0.0111	0.0097	0.0111	0.0111	0.0108	0.0121	0.0107	0.0115	0.0118	0.0095	0.0105	0.0100
					Mean	**Stdev**	**Min**	**Max**	**Median**			
THA/PM/SUM to island sequence divergence					0.0105	0.0007	0.0084	0.0121	0.0106			

table S15. Within-group mtDNA maximum composite likelihood + Γ genetic distance matrix for nine Sunda colugo groups and five Philippine colugo groups. Abbreviations: Vietnam=VNM, Thailand=THA, Peninsular Malaysia=PM, Sumatra=SUM, Borneo=BOR, Bunguran=BUN, Basilan=BAS, Mindanao=MIN, Bohol=BOL, Leyte=LTE, Samar=SAM, Dinagat=DIN, nc=not calculated (single individuals).

Thailand/Pen.Malaysia/Sumatra	0.010
SW Borneo	0.007
Java	0.007
NW Borneo	0.003
SAM/LTE	0.003
Bohol	0.000
E. Borneo	0.011
N.E. Borneo	0.006
E.MIN	0.000
W. MIN	0.002
S.SUM	0.006
LAOS	nc
VNM	nc
DIN	nc

table S16. mtDNA genetic distance between nine Sundaic colugo groups and five Philippine colugo groups, based on the maximum composite likelihood + Γ distance matrix. Abbreviations: Vietnam=VNM, Thailand=THA, Peninsular Malaysia=PM, Sumatra=SUM, Borneo=BOR, Bunguran=BUN, Basilan=BAS, Mindanao=MIN, Bohol=BOL, Leyte=LTE, Samar=SAM, Dinagat=DIN.

	PM THA SUM	**SW.BOR**	**JAVA**	**NW.BOR**	**LTE/SAM**	**BOL**	**E.BOR**	**NE.BOR**	**E. MIN**	**W.MIN**	**S.SUM**	**LAOS**	**VNM**
PM/THA/SUM													
SW.BOR	0.059												
JAVA	0.154	0.149											
NW.BOR	0.058	0.037	0.149										
LTE/SAM	0.217	0.208	0.22	0.210									
BOL	0.217	0.209	0.222	0.210	0.050								
E.BOR	0.151	0.145	0.121	0.146	0.216	0.216							
NE.BOR	0.146	0.140	0.121	0.141	0.211	0.212	0.07						
E.MIN	0.213	0.208	0.216	0.212	0.048	0.052	0.215	0.209					
W.MIN	0.213	0.206	0.216	0.210	0.050	0.053	0.214	0.207	0.035				
S.SUM	0.034	0.057	0.146	0.057	0.204	0.205	0.138	0.136	0.200	0.199			
LAOS	0.081	0.075	0.151	0.076	0.210	0.209	0.145	0.141	0.210	0.207	0.074		
VNM	0.070	0.067	0.144	0.070	0.196	0.197	0.136	0.133	0.197	0.196	0.070	0.064	
DIN	0.210	0.202	0.215	0.201	0.032	0.041	0.212	0.208	0.041	0.040	0.193	0.202	0.190
					mean	stdev	min	max	median				
Average Sundaic Between Group Divergence					0.107	0.040	0.034	0.154	0.127				
Average Philippine Between Group Divergence					0.044	0.007	0.032	0.053	0.045				

table S17. BPP species estimation (*70*) results for the *Galeopterus* biparental dataset (53 partitions, with 9,628 sites after removal of sites with missing data, clean data=1).

Model	**Nodes**[a]	**Posterior Probability**	**Posterior Probability**	**Posterior Probability**	**Posterior Probability**	**Posterior Probability**
1	000000	0.00000	0.00000	0.00000	0.00000	0.00000
2	100000	0.00000	0.00000	0.00000	0.00000	0.00000
3	101000	0.00000	0.00000	0.00000	0.00000	0.00000
4	101010	0.00000	0.00000	0.00000	0.00000	0.00000
5	101011	0.00028	0.00000	0.00000	0.00000	0.00000
6	101100	0.00000	0.00000	0.00000	0.00000	0.00000
7	101110	0.00000	0.00004	0.00000	0.00000	0.00211
8	101111	0.00007	0.00001	0.00006	0.00000	0.00145
9	110000	0.00000	0.00000	0.00000	0.00575	0.04417
10	111000	0.00000	0.00000	0.00000	0.00989	0.00496
11	111010	0.00595	0.00199	0.00099	0.15714	0.06777
12	111011	0.42029	0.17070	0.16314	0.82608	0.64262
13	111100	0.00000	0.00000	0.00000	0.00000	0.00092
14	111110	0.01321	0.00720	0.00590	0.00006	0.02099
15	111111	0.56021	0.82007	0.82993	0.00109	0.21502
Parameters		Parameters estimated from data	Sensitivity analysis	Sensitivity analysis	Sensitivity analysis	Sensitivity analysis
theta (α, β)		2, 408	2, 1000	2, 1000	2, 10	2, 10
tau (α, β)		2, 344	2, 10	2, 1000	2, 10	2, 1000

[a] The ancestral nodes (left to right) for *Galeopterus*: **1)** East Borneo (EB) to Northeast Borneo (NEB), **2)** EB+NEB, **3)** Node 4 to Node 5, **4)** Laos (L) +Vietnam (V), **5)** Java (J) to node 6, **6)** West Borneo (W) to Pen.Malay./Sumatra (PS).

Guide tree with PP for each node (i.e., support for distinct species) in the five analyses shown above (left to right):

((EB, NEB) #0.999650,((V, L) #0.573490, (J, (W, PS) #0.980845) #1.000000) #1.000000) #1.000000;
((EB, NEB) #0.999955,((V, L) #0.827310, (J, (W, PS) #0.990770) #1.000000) #1.000000) #1.000000;
((EB, NEB) #0.999945,((V, L) #0.835875, (J, (W, PS) #0.993115) #1.000000) #1.000000) #1.000000;
((EB, NEB) #1.000000,((V, L) #0.001145, (J, (W, PS) #0.827160) #0.984355) #0.994245) #1.000000;
((EB, NEB) #0.996440,((V, L) #0.240490, (J, (W, PS) #0.859100) #0.949955) #0.955830) #1.000000;

table S18. BPP species estimation (*70*) results for the *Cynocephalus* biparental dataset (101 partitions, with 44,072 sites after removal of sites with missing data, clean data=1).

Model	**Nodes**[a]	**Posterior Probability**	**Posterior Probability**	**Posterior Probability**	**Posterior Probability**
1	00	0.0000	0.0000	0.0000	0.0000
2	10	0.3913	0.2381	0.3365	0.3799
3	11	0.6087	0.7619	0.6635	0.6202
Parameters		Parameters estimated from data	Sensitivity analysis	Sensitivity analysis	Sensitivity analysis
theta (α, β)		2, 2040	2, 1000	2, 2040	2, 2040
tau (α, β)		2, 771	2, 771	2, 10	2, 1000

[a] The ancestral nodes (left to right) for *Cynocephalus*: **1)** Mindanao to Leyte+Dinagat, **2)** Leyte+Dinagat.

Guide tree with PP for each node (i.e., support for distinct species) in the four analyses shown above (left to right):

(Mindanao, (Leyte, Dinagat) #0.608665) #1.000000;
(Mindanao, (Leyte, Dinagat) #0.761880) #1.000000;
(Mindanao, (Leyte, Dinagat) #0.663515) #1.000000;
(Mindanao, (Leyte, Dinagat) #0.620150) #1.000000;

table S19. X chromosome genetic variation principal component loadings.

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
GVA_62	0.266233	0.547404	-0.22344	0.318666	-0.12787	0.199536	-0.22883	-0.17031	0.491531	-0.11319	-0.05209	0.288675
GVA_55	0.106196	-0.07303	0.027003	-0.1837	-0.03574	0.083016	0.033093	-0.13947	-0.01564	-0.09455	0.90977	0.288675
GVA_35s	-0.417930	-0.36815	-0.59937	0.276448	-0.35334	0.000792	0.174165	-0.02659	0.021265	0.119834	-0.0103	0.288675
GVA_40	0.161232	-0.41209	0.569543	0.623993	0.013976	0.037057	-0.00685	-0.01491	0.043918	0.046969	-0.03364	0.288675
GVA_49	-0.108230	-0.19643	-0.20001	-0.10953	0.7248	0.034303	-0.42293	-0.00116	0.130662	0.300926	-0.03684	0.288675
GVA_45	-0.749030	0.418413	0.412216	-0.0457	0.004591	0.010241	0.0647	-0.02743	0.01838	0.055598	0.006742	0.288675
GVA_07	0.234199	0.185082	-0.07773	0.00447	0.290208	-0.3249	0.710371	0.291528	0.131815	0.155633	-0.02253	0.288675
GVA_63	0.189492	0.308727	-0.11744	0.165388	-0.03705	0.070891	-0.14557	0.061808	-0.81878	0.202542	-0.03823	0.288675
GVA_16	0.109996	-0.05044	0.094384	-0.24809	-0.35616	-0.73628	-0.37578	0.054117	0.088315	0.068156	-0.07762	0.288675
GVA_12	-0.06196	-0.09647	-0.08403	-0.03419	0.206989	-0.0607	-0.01122	0.091066	-0.16144	-0.88744	-0.16348	0.288675
GVA_129	0.160713	-0.13246	0.088955	-0.38649	-0.06008	0.171362	0.252997	-0.70848	-0.04948	0.071862	-0.33086	0.288675
GVA_28	0.109087	-0.13056	0.109916	-0.38127	-0.27032	0.514682	-0.04415	0.589822	0.119453	0.07366	-0.15091	0.288675