PNAS
www.pnas.org

Supplementary Information for

Multiple migrations to the Philippines during the last 50,000 years

Maximilian Larena[a,1], Federico Sanchez-Quinto[a,b], Per Sjödin[a], James McKenna[a], Carlo Ebeo[c,d], Rebecca Reyes [e,f], Ophelia Casel[g], Jin-Yuan Huang[h], Kim Pullupul Hagada[f,I], Dennis Guilay[j], Jennelyn Reyes[k], Fatima Pir Allian[l,m], Virgilio Mori[m], Lahaina Sue Azarcon[n], Alma Manera[o], Celito Terando[p,q], Lucio Jamero Jr[e], Gauden Sireg[r,s], Renefe Manginsay-Tremedal[s], Maria Shiela Labos[t,u], Richard Dian Vilar[v,w], Acram Latiph[x], Rodelio Linsahay Saway[y], Erwin Marte[z], Pablito Magbanua[f,aa], Amor Morales[bb], Ismael Java[cc,dd], Rudy Reveche[dd,ee], Becky Barrios[ff,gg], Erlinda, Burton[hh,2], Jesus Christopher Salon[hh,ii], Ma. Junaliah Tuazon Kels[a], Adrian Albano[jj,kk], Rose Beatrix Cruz-Angeles[ll], Edison Molanida[ll,mm], Lena Granehäll[a], Mário Vicente[a], Hanna Edlund[a], Jun-Hun Loo[h], Jean Trejaut[h], Simon Y. W. Ho[nn], Lawrence Reid[d,oo], Helena Malmström[a,pp], Carina Schlebusch[a,pp,qq], Kurt Lambeck[rr,1], Phillip Endicott[ss], Mattias Jakobsson[a,pp,qq,1]

correspondence to: maximilian.larena@ebc.uu.se (M.L.);
Kurt.Lambeck@anu.edu.au (K.L.)
mattias.jakobsson@ebc.uu.se (M.J.)

This PDF file includes:

Supplementary text
Figures S1 to S10
SI References

Supplementary Information Text

List of Supplementary Figures

List of Supplementary Tables

1 Extended Materials and Methods

1.1 Sampling. Saliva samples were collected with the use of Oragene Saliva Collection Kit (DNA Genotek Inc, Ottawa, ON, Canada). Participants were requested to have nothing by mouth for 30 minutes, before sample collection, and those who were eating, drinking, smoking or chewing betel nuts were requested to gargle with water, and rest for 30 minutes with nothing by mouth before sample collection. All participants were at least 18 years old. In each ethnolinguistic group, samples were collected from unrelated individuals and included only one individual from sets of individuals that self-reported to be (up to 2nd degree) relatives. Individuals were also asked to report on their 4 grandparents' ethnicity, and only individuals who had all 4 grandparents from the same ethnic group were included in the study. The list of indigenous cultural communities included in the study is summarized in Table S1.

1.2 Ethical Considerations. This research project, with the aim of establishing baseline scientific data for the genetic diversity, interrelatedness, and migration history of Philippine indigenous cultural communities, is recognized by and implemented in partnership with the National Commission for Culture and the Arts (NCCA) of the Philippines, in accordance with the provisions of Philippine Republic Act 7356, or the Law Creating the NCCA. All research participants were at least 18 years old, and were able to autonomously and voluntarily provide an informed consent. Consent was secured from each individual, and whenever necessary, from each respective Indigenous Cultural Community Council. The consent process, sampling, and/or subsequent validation were performed in coordination with the NCCA and, in some regional areas, with local partners or agencies including non-governmental, cultural organizations, local educational institutions, Indigenous Cultural Community Councils, Local Government Units, and/or regional offices of National Commission on Indigenous Peoples (Supplementary Section 10). The processing of samples and analysis of data was approved by the Regional Ethical Review Board in Uppsala, Sweden (Dnr 2016/103).

1.3 DNA extraction and SNP genotyping. The saliva samples were processed for DNA isolation at the Mattias Jakobsson Laboratory, Human Evolution, Department of Organismal Biology, using the prepIT DNA isolation kit (DNA Genotek Inc., Ottawa, ON, Canada). DNA extracts were quantified using the NanoDrop ND-1000 Spectrophotometer (Coleman Technologies Inc., Orlando, FL, LabVIEW@). The purified 1,094 DNA samples were sent to SNP&SEQ Technology Platform at Uppsala University for genotyping of SNP markers with the Illumina Infinium assay (San Diego, CA) using InfiniumOmni2-5Exome-8v1-3 Bead Chip (2,612,357 SNPs). Initial analysis of newly genotyped samples with GenomeStudio 2011.1 (Illumina Inc.) demonstrated an average call rate of 99.95% (2,611,039 / 2,612,357) and reproducibility of 99.99% (329 conflicts in 5,204,499 duplicate tests),

1.4 Quality control and filtering of genotyped data. Using PLINK v1.9 software (1), the newly genotyped data were filtered for SNP missingness (10%) and SNPs that were not in Hardy-Weinberg equilibrium. Indels, duplicates, non-autosomal and unmapped SNPs were also removed. Four individuals with missing data of more than 15% were removed (Umayamnon-05, Sambal-06, Gaddang-02, and Gaddang-06). Furthermore, family relationships were inferred using KING software, with estimated kinship coefficients of at least 0.0844, corresponding to at least 3rd-degree relationship, set as a cut-off value for identifying related individuals. A total of 62 individuals were removed from pairs of related individuals, producing a final Phil_2.35M dataset with 1028 individuals and 2,359,167 SNPs.

1.5 Merging of Datasets. For comparative analysis, the Phil_2.35M dataset was merged with a worldwide panel of populations from 1000 Genomes Project (2) and Simon's Genome Diversity Project (3), generating a Phil_1KGP_SGDP_1.69M dataset with 3,331 individuals and 1,690,499 SNPs. To cover more populations in the Asia-Pacific region for comparative analysis, the Phil_1KGP_SGDP_1.69M dataset was subsequently merged with published Illumina-based Indonesian (4-7), Malaysian (8), and mainland Asian (9) datasets to generate a Phil_AsiaPacific_315K dataset with 5132 individuals and 315,692 SNPs. For an alternative dataset with distinct set of populations, the Phil_2.35M dataset was merged with Affymetrix Human Origins-based worldwide (10, 11), MSEA (12), Indian (13), and Oceanian (14, 15) datasets to produce Phil_HO_201K dataset with 5402 individuals and 201,387 SNPs.

1.6 Population Genetic Analysis. Measures of genetic diversity including heterozygosity, runs of homozygosity, and inbreeding coefficient were computed using PLINK v1.9. Principal component analysis and calculation of between population F_{ST} was performed using EIGENSOFT v7.1 (16, 17), and a F_{ST}-based neighbor-joining tree was plotted using MEGA7(18). Mantel tests was performed to determine statistical significance on correlations between genetic and linguistic and between genetic and geographic distances.

ADMIXTURE v1.3 (19) and CLUMPP (20) were used to analyze population structure, which was subsequently visualized using Pong v1.4. Outgroup f3 statistics and formal tests of admixture were performed using qp3Pop and qpDstat of AdmixTools v5.0 package(11). We used qpF4Ratio and qpAdmin of AdmixTools v5.0 package for estimating admixture proportions in populations, and qpGraph for fitting populations in an admixture graph with baseline framework based on earlier publications (21). For estimating dates of admixture, we utilized a weighted linkage disequilibrium (LD) statistic-based method, MALDER (22), which also allows detection of multiple admixture events. To infer local ancestries and subsequent masking in admixed populations, we utilized RFMix (23), which employs a conditional random field parameterized by random forests trained on reference panels of least admixed populations.

Detection of identity-by-descent (IBD) segments was implemented using Beagle version 4.1 software (24), and IBDne was used for the estimation of recent effective population size based on IBD (25). Using a coalescent model and a maximum likelihood framework(26-29), divergence time was estimated with the formula $\textit{Divergence Time} = T \times 2N_e \times g$, where T is the drift parameter, N_e as the effective population size and g as the generation time (30 years). The effective population size for Cordillerans was assumed to be 2500 (7, 30), and 3500 for Papuans (31); while the drift parameter was calculated following the method in an earlier study (26).

Our divergence time estimation method is limited by the assumptions on the effective population size and potential issues of ascertainment bias. It is notable however that the estimates we generated here are in line with the estimates generated using other methods and using sequenced genomes. The divergence time estimate we had for

Papuans vs Australians (25 kya) is at the lower boundary of the estimate in Malaspinas *et al*., 2016 (~ 37 kya; 95% CI 25–40 kya). Likewise our divergence time estimate for Cordillerans vs Amis/Atayal (8.4 and 8.9 kya, respectively) is also at the lower limit of the estimates by Choin *et al*., 2020 (~15 kya; 95% CI: 9.2–18 kya).

We further estimate the divergence time between Cordillerans vs Amis/Atayal using the 'TTo' method based on computing sample configurations in a population divergence model and using genome sequence data(29, 32). This approach estimates the number of generations since a population divergence for a pair of individuals (or populations), and the method produces direct estimates in generations that are unaffected by the effective population size of the population of each of the individuals in the comparison. Hence, the 'TTo' approach alleviates both issues with ascertainment bias and (potentially fluctuating) effective population size. The (mean) estimated divergence time between Kankanaey and Ami/Atayal individuals is ~17 kya (95% CI: 9.5 – 25 kya).

1.7 Laboratory processing of ancient samples. Two phalanges, with lab ID tai001 (LDDW-I, TPI-Mo1) and tai002 (LSSW-I, Mo2), were used for DNA extraction in a dedicated ancient DNA laboratory (see e.g.(29)). The surface of the phalanges was decontaminated through UV irradiation (6J/cm2 at 254 nm) and the outer layer was removed using a Meisinger ISO 021 dental burr or a Dremel drill with either cutting discs or dental burs. To investigate if the samples contained endogenous ancient DNA, ~50 mg bone powder was retrieved from each phalanx using the above-mentioned drill. DNA was extracted using a silica-based method(33) with modifications as in (34). Blunt-end Illumina DNA libraries was prepared using P5 and P7 adapter and indices(35), omitting the shearing step, followed by amplification for 16-20 cycles using IS4 and index primers from (35), cleaning of libraries and DNA quality control as described in (36).

The libraries were screened using Illumina HiSeqX with 150bp paired-end length and v2.5 chemistry at the SNP & SEQ Technology Platforms at Uppsala University. Upon successful retrieval of endogenous ancient DNA, five UDG-treated libraries were built (37) to reduce the effect of post-mortem DNA damage, for the best quality sample, tai002, following (36). We then processed the phalanges using a Starbeater mill (VWR) to provide additional material. Three additional DNA extractions were performed for tai001 and nine additional extractions for tai002, using ~ 50 mg bone powder per extraction and a silica-based method specifically suited for heavily degraded DNA (38).

For tai001, one blunt-end Illumina library (described above) was produced per extract. One library was further subjected to whole genome capture using MyBaits Human Whole Genome Capture kit (Mycroarray, Ann Arbor, MI) following the manufacturer's instructions (MyBaits manual version 2.3.1). For tai002, one or two UDG-treated libraries (see above) were prepared from each extract. The libraries were shotgun sequenced as described above. Negative controls followed the ancient samples throughout the entire laboratory process and they did not yield any DNA and were therefore not sequenced.

1.8 Processing of aDNA data. For processing of aDNA data, paired-end reads were merged and their adapters trimmed and subsequently mapped to the human reference genome using BWA (39) following (40, 41). PCR duplicates with identical start and end coordinates were collapsed into consensus sequences (42). Both tai001 and tai002 sequences demonstrated the deamination pattern towards the read fragment-ends (43), characteristic of ancient DNA. The level of contamination was estimated using the Green (44), Contamix (45), and Schumtzi (46) methods for mitochondrial DNA and VerifyBamID for autosomal DNA. In order to investigate if the mitochondrial contamination estimates could be extrapolated as a proxy for nuclear contamination in Liangdao2, we estimated the mitochondrial to nuclear ratio as described previously (47). The ratios lying between 33-50 and are below the threshold of 200 suggested as the highest ratio at which mitochondrial contamination estimates could be interpreted as a nuclear contamination proxy (47).

1.9 Bayesian phylogenetic analysis of mitochondrial genomes. We used BEAST 1.10(48) to perform Bayesian phylogenetic analyses of two separate data sets, comprising samples from hg E (n=207) and hg R9 (n=137). The data sets were partitioned into five subsets: the D-loop, RNA genes, and the first, second, and third codon positions of the 13 protein-coding genes. For each subset, the best-fitting model of nucleotide substitution was chosen using the Bayesian information criterion. For each data set, we compared two models of among-lineage rate variation (strict clock and relaxed clock) and two tree priors (constant-size coalescent and skyride coalescent) for a total of four model combinations. These model combinations were compared using marginal likelihoods, estimated using path sampling with 25 power posteriors.

The dating analyses were calibrated using the ages of the ancient samples and using an informative prior for the substitution rate. We specified a normal distribution for the prior on the substitution rate, with a mean of 2.67×10^{-8} substitutions/site/year and a standard deviation of 2.551×10^{-9} (45). The posterior distribution was estimated using Markov chain Monte Carlo analysis, with samples drawn every 5000 steps over a total of 10^6 steps. To check for convergence to the stationary distribution, we ran each analysis in duplicate. We checked for sufficient sampling by inspecting the effective sample sizes of all parameters in Tracer (49).

2 Geology of Island Southeast Asia and Taiwan-Southern China

During the glacial cycles of the Quaternary Period, the dominant global-scale process that affects the planet's morphology is the changing sea level and concomitant shoreline migration as ice sheets grew and decayed, lowering and raising sea level by typically 100-150 m and exposing or flooding shallow continental shelves and facilitating or inhibiting human (and other flora and fauna species) migration, particularly from the last interglacial period ~120,000 years ago to the end of the last major deglaciation at ~10,000 years ago. On regional and local scales, tectonic processes can also play an important part in these modifications, most notably at active plate margins. Global internal processes (mantle convection) play only secondary roles unless the time scales extend out to millions of years. In SE Asia, both these global and regional processes have modified the geography of the region on these glacial time scales.

The landscape of island SE Asia has been shaped over a period of ~ 100 million years by the convergence of the three major tectonic plates - the Eurasian, the Indo-Australian and Pacific - the recent expression of which is characterized in the west and southwest by the Sunda Trough and fault, and in the eastern half by a complex zone of subduction and faulting defining the margins of small crustal blocks (50, 51). In between lies a relatively stable area, the Sunda Block, of continental crustal structure and shallow seas (50). To the north, the block extends to the Malay-Thai peninsula.

The current plate margins are clearly defined by seismic activity and volcanism whereas the Sunda Block itself is largely aseismic. In the east, this block includes the shallow Sulu and Celebes Seas, extending to the western islands of the Philippines archipelago defined primarily by transform faults. The dominant land body today is Borneo with small island groups extending to the Philippines: northern Sabah-Palawan-Mindoro-Luzon in the north and eastern Sabah to Mindanao in the south. In times of low sea level much of the block was exposed, forming an exposed landmass giving rise to its name 'Sundaland'.

The Sunda Block itself has not been stable on the million-year time scale with uplift recorded in the geology of Borneo and with geophysical signatures of high upper-mantle temperatures and low seismic velocities (52) that are indicative of tectonic deformation with uplift rates of the order of 0.10 mm/year or less (or < 1m in10^4 years); well below both the detection level of the available geodetic data and the accuracy of any palaeographic reconstructions. At the margins of the block, past vertical rates may have been greater with uplift of several km over the last 3-5 million years (e.g in the Banda Arc and Mollucas) (50, 51).

For the Sunda Shelf with its shallow seas, the more important contribution to changing shorelines on the multi-thousand year time scales is from the growth and decay of the large high latitude ice sheets and their concomitant fluctuation in ocean volume: with an eustatic change (the globally averaged change in sea level, equal to the change in grounded ice volume distributed uniformly over the world's oceans) of the order 100-150 m during a full glacial cycle (Figs. S3A,B) and the shallow shelves for much of the present seas, considerable parts of the block have been exposed at glacial times opening up, if not land bridges, then much reduced water crossings between southern Asia and the outlying islands.

The response of the crust and oceans to changes in ice volume produces a complex pattern of temporal and spatial change in land elevations and shoreline locations and departures from the eustatic approximation can be significant due to the changing shape and gravity of the solid earth, oceans and ices sheets as the planet deforms under the changing ice-water surface loads (53). The combined gravitational and deformational response is usually referred to glacio- or glacio-hydro- isostasy, the latter in recognition that changes in water-loading are also important.

In low latitude regions, far from the former ice sheets, the major departure from eustatic sea level results from the earth-ocean response to the changes in water load during times of growth and decay of the ice sheets and may exceed 25% of the eustatic change, notably in SE Asia(53, 54) and become important in palaeo-shoreline reconstructions, particularly during glacial and late-glacial times.

Physical-mathematical models for this combined ocean-earth response are well developed and provide a predictive capability whose reliability will depend on the modellers' knowledge of the ice history and earth's rheology, knowledge of which is inferred from field evidence, of ice margins and retreat histories, and from a calibration of the models against observations of sea level change in both high- and low-latitude regions. Figs. S3C and S3D compare observed and predicted rsl for SE Asia. The former are for an area bounded by longitudes 95° and 130° E and latitudes 27° N to 15° S and are from (55) as are the predicted values. Agreement between the two is within the uncertainties of both predictions and observations such that predictions across the region as a whole should give a realistic representation of the regional sea level change back to about 30 kya. Then, if present ocean depths are known, palaeo-depths can also be predicted as can the location of shorelines throughout this time interval.

Figure S3E gives the shoreline reconstructions for SE Asia at selected epochs, from the time of the maximum glaciation at ~21 kya BP to ~6 kya by which time sea levels reached values close to present-sea level. Included are the onset of the Bølling–Allerød 'warming' at ~15 kya after which sea levels rose rapidly by ~40-45 m in ~500 years, and the Younger Dryas 'cold' periods at ~12.5 kya when sea levels remained nearly constant for ~300 years. Higher resolutions reconstructions for selected periods are shown in Figure S3F for the area between North Borneo and the Philippines. These reconstructions are based on (i) earth rheology parameters that are appropriate for continental and ocean island sites far from the former ice sheets, (ii) on ice models extending back to the penultimate glacial maximum ~140 kya BP that have been inferred from analyses of sea level change across areas of former glaciation (55), and (iii) sea-floor bathymetry from used. The GEBCO 30″ global gridded data. (https://www.gebco.net/data_and_products/).

Throughout the glacial period much of the Sunda Block is exposed and access to the Philippines at the time of maximum glaciation would involve only narrow water crossings (< 3 km), either via Palawan and Mindoro, or from eastern Sabah to Zamboanga, and beyond to Luzon Island. Once across these channels the entire archipelago is accessible without further water crossings. This scenario is maintained until about 15-14 kya with the more tractable route from Borneo being the southern one, the Sibutu-Basilan Ridge. Similar configurations would have occurred in the ~

70-60 kya interval and later at ~40 kya. To the north, the entire shelf separating Taiwan from mainland China is dry throughout the glacial interval with submergence initiated before ~13 kya (Fig. S3G). By 8 kya predicted sea levels are approaching their present levels and the reconstructions on these scales are indistinguishable from present maps.

Predictions for earlier periods become increasingly uncertain because of the paucity of sea-level data that is also of increasingly greater uncertainty, and because constraints on the ice sheets becomes increasingly uncertain. For the low latitude locations, the isostatic effects are dominated by the water-load changes and the source of the melting ice, the global ice-volume function, becomes of secondary importance provided that the mass is conserved in the ocean-ice system. To infer this latter, we use sea-level constraints (Fig. S3A) derived from the uplifted terrace sequences of Papua New Guinea(53, 56): relative highstands are determined from coral reef elevations corresponding to times when rates of uplift were about equal to rates of sea-level rise; low stands are determined from the elevation of erosional features that formed in these reefs during subsequent periods of sea-level fall. To reduce this local sea-level curve to an ice-volume function we assume that the relative contributions from the individual ice sheets to the total ocean volume function is the same for the pre-LGM period as for the post-LGM period. At 65 kya, for example, corresponding to the interstadial MIS 4, the local sea level estimate is about the same as at ~14 kya and we assume that the ice distribution at the former epoch is the same as that at 14 kya. This allows us to build up a global ice scenario that is consistent with the estimate of the global ice volume fluctuation and to predict the palaeo topography for the earlier periods.

3 General population structure of the Philippines

3.1 Philippine ethnic groups cluster with Asia-Pacific populations

To determine the relationship of Philippine ethnic groups to a worldwide set of populations, PCA was performed using EIGENSOFT (16,17) on Phil_1KGP_SGDP_1.69M, Phil_AsiaPacific_315K, and Phil_HO_201K datasets with the following set of parameters: r2 threshold of 0.2, sample size limit of 20, and no outlier removal (Fig. 1B & Figs. S1B,C). Consistently among 3 datasets, PC1 is defined by East Asian vs African axis, while PC2 is defined by European vs African axis. All Philippine ethnic groups cluster together with Asia-Pacific populations, where Philippine Negritos (maroon circular marker) form a distinctive cline that aligns between Oceanians (i.e. Australian & Papuan) and East Asians, indicating an ancestry that is a mixture of Australasian-related & East Asian-related. Cordillerans, interestingly, lie at the extreme edge that defines the East Asian cluster of PC1, even more extreme than indigenous Taiwanese, Amis or Atayal, or any other East Asian population (i.e. Han, Dai, Japanese, etc.) or MSEA population (i.e. Kinh, Cambodian, Thai, etc.).

3.2 Asia-Pacific-restricted PCA of merged datasets

PCA restricted to Asia-Pacific populations was performed on Phil_AsiaPacific_315K and Phil_HO_201K datasets, with similar parameters indicated above (Figs. S1D,E). For Phil_AsiaPacific_315K dataset, the distribution of populations corresponds well with geography, where PC1 is defined geographically by east (East Asians) vs west (indigenous populations of India) and PC2 by north (North Asians) vs south (Oceanian ethnic groups). Again, in all datasets, Philippine Negritos lie between Oceania and East Asia, while Cordillerans lie at the extreme edge of PC1, defining the East Asian cluster. All other non-Negrito and non-Cordilleran Philippine ethnic groups (orange circular marker) cluster together between Cordillerans, Philippine Negritos, and MSEA-based populations (i.e. MahMeri & Seletar Orang Asli Malays, Htin, Mlabri, Vietnamese & Cambodians), indicating a variable combination of Negrito-related, Cordilleran-related and MSEA-related ancestries for these groups.

Implementing an East and Southeast Asian-restricted PCA on Phil_AsiaPacific_315K and Phil_HO_201K datasets further revealed detailed genetic relationships between ISEA ethnic groups (Figs. S1F,G). Both datasets are characterized by Cordilleran vs continental East Asian clusters in PC1, and Cordilleran vs MSEA-based (MahMeri & Seletar Orang Asli Malays for Phil_AsiaPacific_315K) or MSEA-related (Sama ethnic groups for Phil_HO_201K) clusters in PC2. In Phil_AsiaPacific_315K dataset, Manobo, Sama, western Indonesian, and other non-Cordilleran populations lie between AN-speaking Cordillerans and AA-speaking Orang Asli Malays, supporting a dual Cordilleran and MSEA-related ancestral sources for the non-Negrito component of these groups. The genetic relationships between Manobo, Sama, and Mlabri/Htin groups of MSEA were more evident in Phil_HO_201K dataset: all three clusters lie opposite to and away from Cordillerans, where Sama, Malaysians, and western Indonesians lie closer to Mlabri/Htin groups relative to Manobo. In both datasets, Mangyan consistently form their own cluster between Cordillerans and Manobos. Morever, Amis and Atayal also form their own cluster, lying in the middle of Cordilleran, MSEA, and East Asian ethnic groups, indicating admixture and complex demographic history for these indigenous Taiwanese populations.

3.3 PCA of Phil_2.35M dataset

When PCA was applied to Phil_2.35M dataset (Fig. 1C), PC1 is defined by Negrito (maroon circular marker) vs Cordilleran (green circular marker) clusters, and PC2 by Sama (dark blue circular marker) vs Mangyan/Cordilleran clusters. The Negrito axis is best defined by the least admixed Negritos, Ayta Magubukon and Ayta Ambala, the Cordilleran axis by the least admixed Cordillerans: Balangao, Bontoc, Ayangan Ifugao, Tuwali Ifugao, Kalanguya, Kankanaey and Ibaloi, and the Sama axis by Sama Dilaut sea nomads: Sama Dilaut Bongao, Sama Dilaut Mampang and Sama Dilaut Taluksangay. Among the Mangyan ethnic groups, which forms their own cluster (purple circular marker), Mangyan Iraya distinctively appears to attract more towards the Negrito axis, indicating a greater affiliation to Negritos compared to other Mangyan or any other non-Negrito ethnic groups. All Manobo (orange circular marker) and other non-Negrito ethnic groups (pink circular marker) are found in the midst of the Cordilleran, Sama, and Negrito axis, indicating a complex demographic history that is composed of at least three ancestral sources.

3.4 Population stratification of Philippine ethnic groups by ADMIXTURE analysis

To investigate the general population structure of the Philippines, based on maximum number of SNP genotypes (2.35 million SNPs), we utilized an unsupervised clustering algorithm implemented in ADMIXTURE (19). Using default parameters, we ran 50 iterations for each K (number of clusters). Common modes of replicates, based on LargeKGreedy algorithm with 1000 repeats, were identified using CLUMPP (20). A cutoff symmetric coefficient G’>=0.9 between pairs of replicates were considered to classify the replicates as belonging to the same nodes. After selection of the most frequent common mode, a second run of CLUMPP was implemented for all K values with the most frequent common mode based on LargeKGreedy algorithm of 10,000 repeats. The result was then processed for analysis and visualization of latent clusters using Pong v1.4 (57), where the major mode for each K was plotted (Fig. S2B and Table S2).

Assuming two clusters (*K2*), Philippine ethnic groups are split into maroon component, found predominantly among Negrito groups covering Ayta, Agta, Arta, Atta, Ati, Batak, & Mamanwa, and a pink component, found predominantly among non-Negrito ethnic groups covering Batanic, Cordilleran, Cagayan, Sambalic, Palawanic, Mangyan, Tagalog, Bicolano, Visayan, Danao, Bilic, Manaskan, Manobo, Sama, and Sangir ethnic groups. The least admixed Negrito group are best represented by Ayta Magbukon and Ayta Ambala, while the least admixed non-Negrito is best represented by Cordillerans including Kankanaey, Bontoc, Balangao, Ayangan Ifugao, Tuwali Ifugao, Kalanguya, and Ibaloi. Among self-identified non-Negritos, Mangyan Iraya have higher than average levels of Negrito ancestry, which even exceed the levels in some Negrito groups, such as Ati Panay, Ati Negros, Agta Bulusan and Agta Matnog. This is consistent with the earlier findings in the PC analysis where Mangyan Iraya appears to have more Negrito ancestry (Fig. 1C), relative to other Mangyan groups of Mindoro or any other non-Negrito groups of the northern Philippines.

At K3, non-Negritos split into green and pink components, representing Cordilleran-related or non-Cordilleran-related ancestry, respectively. The latter is best represented by Sama Dilaut sea nomads of Tawi-Tawi and Zamboanga peninsula, and is associated with MSEA-based ancestry which will be emphasized later in an Asia-Pacific-wide ADMIXTURE analysis. At K4, Mangyan ethnic groups acquires their own purple component cluster, with Mangyan Buhid and Mangyan Bangon as the best surrogates. At K5, Sama ethnic groups acquires their own dark blue cluster, separating them from other non-Negrito ethnic groups of southern Philippines (pink component), which are mainly defined by Manobo ethnic groups.

At K6, Negrito groups split into Ayta (brown component, Central Luzon Negritos covering Ayta Ambala, Ayta Magbukon, Ayta Mag-antsi, Ayta Mag-indi & Ayta Sambal) vs all other Negrito ethnic groups (maroon component), and at K7, Manobo ethnic groups acquire their own orange cluster, which is best represented by Ata Manobo, Bukidnon Matigsalug, Davao Matigsalug, and Bukidnon Tigwahanon. Further stratification of non-Ayta Negritos can be observed from K8 to K10, where Southeast Luzon Negritos (Agta Manide, Agta Lopez, Agta Isarog, Agta Iraya & Agta Iriga) acquires their own light blue cluster at K8, and subsequently Northeast Luzon Negritos (Agta Dupaningan, Agta, Labin, Atta Rizal, Agta Maddela, Agta Casiguran, Arta) and Southern Negritos (Mamanwa) acquire their own clusters at K9 and K10, blue and black components, respectively.

At K11, Manobos are further split into central, northern and western Manobos (orange component) and southern and eastern Manobos & Mansakans (dark orange component). The former are best represented by Ata Manobo, Bukidnon Matigsalug, Davao Matigsalug, and Bukidnon Tigwahanon, while the latter are best represented by Manobo Sarangani & Manguangan Manobo as well as Mansakan ethnic groups including Kalagan, Mansaka, and Tagakaulo. At K12, further stratification is observed within Mangyan ethnic groups, where Mangyan Iraya acquires its own sky blue cluster . At K13, Agta Dumagat and Agta Remontado of Central Luzon acquire their own light purple cluster, and lastly at K14, a pink component arises among ethnic groups of southern Philippines, and is best represented by Ati Negros and Ati Panay.

The optimal K where the highest proportion of the most frequent common mode was observed is K7, 45 out of 50 independent runs, demonstrating a complex population structure within the Philippines represented by Cordilleran, Mangyan, Manobo & Sama non-Negritos, and Ayta & Agta negritos. However, we cannot discount the observations at K8 and K9, which also displayed a relatively high proportion of the most frequent common mode (both at 42 out 50 of independent runs), and where the Negritos further cluster into Northeast Luzon Negritos, Southeast Luzon Negritos, Central Luzon Negritos, and Southern Negritos.

3.5 ADMIXTURE Analysis on a subset of Phil_AsiaPacific_315K dataset

Additionally, we ran ADMIXTURE on a subset of Phil_AsiaPacific_315K dataset (Fig. S2C and Table S3), covering all Asia-Pacific populations with a minimum sample size of 3 individuals, plus 4 African groups (Ju/'hoansi, Luhya, Mbuti, Yoruba) & 3 European populations (CEU, IBS & TSI). Each population is limited to a maximum of 15 individuals, and each K was ran for 50 iterations. After processing with CLUMPP, the major mode for each K was plotted using Pong v1.4. From the start at K2, the populations are split into African + West Eurasian + Australasian (maroon component) vs East Asian clusters (pink component, best represented by Cordillerans). At K3, Australasians (maroon component, best represented by Papuans, and found in high proportions among Andaman Island Negritos, Malay Negritos, Philippine Negritos, and South Asians) separate from Africans and West Eurasians (brown component). Then Africans acquire their own (gray) cluster at K4. Subsequently, Andaman Islanders (Onge & Jarawa) acquire their own (purple) cluster, which is also found in high proportions among Malay Negritos (Bateq, Jehai, Kintaq & Mendriq), consistent with previous findings of shared ancestry between the Andaman Island & Malay Negritos (58).

At K6, a component (blue) emerges, which represents northern East Asian ancestry. This is found, in highest proportion, among Evenki, Koryak, and Yakutian ethnic groups. Papuan ancestry (black component) splits from Philippine Negrito ancestry (maroon component) at K7, and is found in high proportion among eastern Indonesians, relative to western Indonesians, and among non-Negrito ethnic groups of middle to southern Philippines, relative to non-Negrito ethnic groups of the north (where it is largely absent). The east-to-west and south-to-north gradient of Papuan ancestry in Indonesia and Philippines, respectively, is consistent with the likely westward gene flow of Papuan-related ancestry into eastern Indonesia and subsequently into southern Philippines.

The split between Andaman Islanders (light gray component) and Malay Negritos (light purple component) is observed at K8. At K9, a green Cordilleran cluster appears and separates from the rest of non-Negrito groups of ISEA (pink component), and subsequently at K10, a (purple) Mangyan cluster appears (found predominantly in Mangyan Buhid, Mangyan Bangon, Mangyan Iraya & Mangyan Hanunuo). Europeans, represented by CEU, IBS & TSI, acquire its own (light teal) cluster at K11.

From K12 to K13, further stratification of non-Negrito ethnic groups is observed, with the appearance of a (dark blue) Sama cluster at K12 and of teal Orang Asli Malay non-Negrito cluster at K13. The blue Sama component is found in highest proportion among Sama sea nomads (Sama Dilaut Bongao, Sama Dilaut Mampang, Sama Dilaut Taluksangay), and in minor proportion among shoreline-dwelling Sama (Sama Kabingaan & Sama Banguigi), urbanized inland Sama (Sama Deya Bongao), Yakan of Basilan, and Tausug of Sulu. Outside of the Philippines, the Sama component is also found in high proportion among the Bajo groups of Indonesia, including Bajo Derawan, Bajo Kotabaru & Bajo Kendari. The Malay non-Negrito component is best represented by MahMeri and Seletar of Malaysia, and is contrasted from the pink Manobo cluster of Mindanao (best represented by Ata Manobo, Manobo Matigsalug, Bukidnon Matigsalug). Malay non-Negrito component is also found among most non-Negrito ethnic groups of ISEA; East Asians (Dai, Japanese, Han); Kinh of Vietnam; and Austroasiatic-speaking (Gond, Ho, Santal, Korwa, Shor) & Tibeto-Burman-speaking (Brahmin, Tharu, Tripuri, Jamatia) populations of India.

At K14, Ayta groups of Central Luzon (brown component) split from all other Philippine Negrito groups (maroon component), and finally at K15, the Manobo cluster (orange component) splits from the other non-Negrito-related cluster (pink component) which is found in high proportions among non-Negrito ethnic groups of the northern Philippines.

The optimal K with the highest proportion of the most common mode is K12 (50/50), although K13 (42/50) and K14 (41/50) also display high frequency of the most common mode, and thus should not be discounted. The clusters found in K12 to K14 portray a complex population structure within the Philippines, composed of Ayta & Agta Negritos and Cordilleran, Mangyan, Manobo & Sama non-Negrito clusters, as well as MahMeri-related component that is shared with Malay non-Negritos.

3.6 ADMIXTURE Analysis on a subset of Phil_HO_201K dataset

ADMIXTURE analysis was also implemented on a subset of Phil_HO_201K dataset with a particular focus on Asia-Pacific populations (Fig. S2D and Table S4). Starting with African + West Eurasian + Australasian (light green component) vs East Asian (pink component) at K2, it is followed sequentially by the appearance of an Australasian (black) cluster at K3, (gray) African cluster at K4, (blue) northern East Asian (nEA) cluster at K5, separation of (brown) Negrito cluster from (black) Papuan cluster at K6, and delineation of (dark brown) Philippine Negrito cluster from (light brown) Andaman Island Negrito + Ancestral South Indian cluster at K7.

Distinct to this panel is the appearance of Htin/Mlabri-related cluster at K8 (teal component), which is mainly found in AA-speaking Mlabri and Htin ethnic groups of MSEA (as the least admixed), as well as Sama Dilaut sea nomads of southwestern Philippines, Lebbo and Bajo ethnic groups of western Indonesia, all East Asian & MSEA groups, some North Asians, and AA-speaking and Tibeto-Burman-speaking populations of India. The widespread distribution of AA-related ancestry demonstrates the shared genetic origins of Mlabri & Htin ethnic groups of MSEA and Barito-speaking populations of Borneo(12) and, as presented in this paper, Sama-related ethnic groups of southwest Philippines. Furthermore, in the context of Mlabri/Htin-related genetic component, our findings also demonstrate the shared AA-related genetic ancestries between Sino-Tibetan speaking populations of East Asia & India and AA-speaking ethnic groups of MSEA & India.

Further stratification of Philippine non-Negrito populations is observed from K9 to K11, as demonstrated by purple Mangyan cluster at K9, green Cordilleran cluster at K10 as contrasted to pink Manobo cluster, and dark blue Sama cluster at K11. At K12, Andaman Island Negrito forms its own cluster (gray component) that is differentiated from Ancestral South India (brown component) found among ethnic groups of mainland India. From K13 to K19, further stratification among Philippine Negritos and appearance of Manobo, Mlabri-distinct, and South African clusters are observed with the following sequence: Southeast Luzon Negritos at K13 (light blue), Manobo at K14 (orange), Mlabri at K15 (light purple), Northeast Luzon Negritos at K16 & K17 (sky blue), Southern Africa at K18 (dark purple component), and Agta & Atta Negritos of Cagayan province at K19 (light green component).

It is at K10 where the optimal K with the highest proportion of the most common mode is observed (44/50), although K12 should not be discounted given its display of relatively high frequency of the most common mode (38/50). The substructure that is observed in K10 and K12 include Philippine Negrito & Papuan-related clusters, and Cordilleran, Mangyan, Manobo & Sama non-Negrito clusters, as well as AA-related component that is shared with Mlabri/Htin ethnic groups of MSEA.

3.7 Neighbor-joining tree based on pairwise Population F_{ST} reveals a dichotomy between Negrito and non-Negrito groups

A matrix of pairwise F_{ST} of Philippine ethnic groups was determined using the *pyhlipoutname:* parameter flag of EIGENSOFT, and was used as an input for constructing a phylogenetic tree in Mega7 software (18). The plot of the neighbor-joining tree based on pairwise F_{ST} dissects Philippine ethnic groups into Negrito vs non-Negrito clusters (Fig. S2A). However, four self-identified Negrito groups, such as Panay Ati, Negros Ati, Agta Matnog and Agta Bulusan, cluster with non-Negritos, indicating high levels of admixture with non-Negritos among these groups consistent with what was observed in Admixture analysis (Figs. S2B-D and Tables S2-4). Negritos display a nested population structure, where southern Negritos, such as Mamanwa and Batak, are basal to the rest. Likewise, non-Negritos also display a nested population structure, where all non-Negrito ethnic groups south of Luzon are more basal in relation to ethnic groups of the north, which would reflect the idea that Cordilleran-related ancestry is a more recent arrival than MSEA-related ancestry to the Philippines.

3.8 Assessing genetic drift based on Runs of homozygozity

Some ancestry clusters observed in Admixture may be partly explained by genetic drift caused by recent bottlenecks, founder events, and/or inbreeding. Using the --homozyg and --het flags of PLINK v1.9 software (1), we calculate the total runs of homozygosity (ROH) and inbreeding coefficients, respectively, for all individuals in the Phil_2.35M dataset, and plotted a bar graph indicating the average and SD values for each population. The patterns observed in plotting total ROH and inbreeding coefficient per population generated similar results, where Mangyan ethnic groups and Sama Dilaut sea nomads displayed the highest levels of total ROH and inbreeding coefficients (Figs. S1H,I). Plotting concurrently the number and length of ROH in populations can also be informative in inferring demographic history, where long tracts of ROH are characteristic of bottlenecks or endogamous populations ((59); Figs. S1J). All Philippine ethnic groups, similar to all other non-African populations, display high numbers of short length ROH category, consistent with the ancient bottleneck likely brought about by a shared Out-of-Africa event. On the other hand, longer tracts of ROH are uniquely evident amongst Mangyans and Sama Dilaut sea nomads, probably reflecting high levels of endogamy practiced by these groups.

4 Genetic diversity & Archaic ancestry of Philippine Negritos

The Philippine Negritos included in the study were asked with regards to the acceptability of the term Negrito. All Negrito participants self-identify as Negritos and do not object to the use of this exonym.

4.1 Population structure of Philippine Negritos

To investigate the genetic relationships between Philippine Negritos, a subset PCA was performed using EIGENSOFT on Phil_2.35M dataset that is restricted only to Negrito ethnic groups (Figs. S4A,J). Evidently, Negritos are structured into Northern Negritos vs Southern Negritos (PC1), and within the northern island of Luzon, into Ayta vs Agta/Arta/Atta/Dumagat Negritos (PC2). The Southern Negritos include all Negrito populations that reside in islands south of Luzon, including Batak of Palawan, Ati of Negros and Panay islands, and Mamanwa of Mindanao. Further investigation into Northern Negritos with subset PCA reveals a deep population structure composed of 3 distinct clusters; Central Luzon Negritos (Ayta Ambala & Ayta Magbukon of Bataan, Ayta Mag-antsi of Tarlac, Ayta Mag-indi of Pampanga, and Ayta Sambal of Zambales), Southeastern Luzon Negritos (Agta Iraya, Agta Iriga & Agta Isarog of Camarines Sur, Agta Manide of Camarines Norte, and Agta Lopez of Quezon), and Northeast Luzon Negritos (Agta Dupaningan, Agta Labin & Atta Rizal of Cagayan; Agta Casiguran, Agta Maddela & Arta of Quirino and Aurora provinces) (Fig. S4B).

4.2 Negritos of northern Philippines

To determine the presence of a distinct Northern Negrito ancestry, we utilized the test *D(Mbuti,AytaMagbukon,Balangao,X)*, where we examine whether any X Negrito population share alleles with the least admixed Negrito of Luzon, Ayta Magbukon, relative to the least admixed East Asian, Cordilleran Balangao (Fig. S4C and Table S5A). We find that most Negrito groups of Luzon displayed detectible presence of Ayta-related Northern Negrito ancestry. All Negrito groups south of Luzon, on the other hand, such as Mamanwa, Batak, Ati Panay & Ati Negros, as well as the highly admixed Negrito populations of southeast Luzon (Agta Matnog and Agta Bulusan of Sorsogon) did not present with any signal of Ayta-related Northern Negrito ancestry. We then investigated further whether Negrito groups of the Philippines share ancestry with AustraloPapuans using the test *D(Mbuti,Australian/Papuan,Balangao,X)*. We find that Northern Negritos exhibit uniform low-level AustraloPapuan genetic signal, which reflects shared ancestry between Negritos and AustraloPapuans (Fig. 4D). The shared AustraloPapuan genetic signal detected in Northern Negritos is however noticeably lower than what is found in Mamanwa of Mindanao, indicating that Negritos of southern Philippines are genetically closer to the AustraloPapuan clade (discussed in Section 4.3).

To examine whether any Northern Negrito population exhibit greater genetic affinity with either Australians or Papuans, we utilized the test *D(Mbuti;X,Australian,Papuan)* (Fig. S4E). We observe that all Northern Negrito groups are equidistant to Australians and Papuans (indicating a lack of directional geneflow from either Australians or Papuans into any Negrito population of Luzon), and that Northern Negritos are evidently an outgroup to the AustraloPapuan clade. The (NorthernNegrito,(Papuan,Australian)) topology was supported by the tests *f3(Mbuti;Papuan,X), f3(Mbuti;Australian,X)* or *f3(Mbuti;AytaMagbukon,X),* where we show that Australians and Papuans share more drift with each other than with any other Northern Negrito population (Figs. S4G-I). Hence, we infer that the ancestors of present-day Northern Negritos who entered the island of Luzon likely diverged from a common ancestral Australasian population (Ancestral Sunda) in the old continental landmass of Sundaland, earlier than the divergence between Australians and Papuans.

We then estimated the divergence time between Northern Negritos and AustraloPapuans using a previously established method (26). First, we observe a divergence time of ~25 kya between Australians and Papuans (Fig. S4J), which is in the lower limit of the range of a previous estimation using a different method (60). Northern Negritos were estimated to diverge from a common ancestral Australasian population approximately 46 kya, which is significantly older than the divergence between Papuans and Australians (Fig. S4J). This was prior to the LGM which lasted from 31 until 16 kya and peaked around 26 kya. Thus, this provides a window of opportunity for ancestors of Northern Negritos to have easier access to reach Philippines via land migration. During this period, there is an open corridor for populations to cross from MSEA to Borneo and then to Palawan, given the exposed land bridges brought about by lower sea levels (Figs. 3E,F). However, it will still require a sea-based journey to reach Luzon, traversing via Mindoro strait between Palawan and Mindoro island, and via Verde island passage between Mindoro and Luzon. Interestingly, traces of Northern Negrito ancestry can be detected among Mangyan populations of Mindoro, which is particularly high in Mangyan Iraya (Fig. 4N and Tables S2-4 & S5B). This suggests that Mangyans share a genetic ancestry with an ancient transitory Negrito population that crossed from Palawan to Luzon via Mindoro (which eventually diverged in Luzon to became the present-day Ayta and Agta Negritos). This would additionally imply that, aside from ancestral AustraloPapuans, the ancestors of Northern Negritos are likely one of the earliest ocean navigators in human prehistory.

We constructed a topology to depict the genetic relationships between Northern Negritos using qpGraph (Table S5C). The model that fits reveals a trifurcation between Central Luzon, Northeast Luzon & Southeast Luzon Negritos (worst fitting f statistic of $Z = 2.2$). Upon entry into Luzon, Northern Negritos initially diverged around 36 kya into what became the present-day Ayta Negritos that settled in Central Luzon, Southeast Luzon Negritos of Bicol region and Quezon province, and Northeast Luzon Negritos of Cagayan valley region (Fig. 4J). The latter subsequently diverged into Negritos of Cagayan province (Agta Dupaningan, Agta Labin & Atta Rizal) and Negritos of Quirino and Aurora provinces (Arta, Agta Maddela & Agta Casiguran).

4.3 Southern Philippine Negrito of Mindanao Island

Southern Negritos are best represented by Mamanwa ethnic group of Surigao del Norte and Agusan del Norte provinces of Mindanao. Mamanwa possess higher AustraloPapuan-related genetic signal relative to Northern Negritos of Luzon, which indicates that Mamanwa are genetically closer to Australians and Papuans (Figs. 4J,K,L). Based on *f3(Mbuti;Papuan,X)*, *f3(Mbuti;Australian,X)* or *f3(Mbuti;Mamanwa,X)*. Where we examine levels of shared drift between Mamanwa, Australians, and Papuans, we find that Mamanwa also appears to be an outgroup to the AustraloPapuan clade (Fig. 4M).

In addition, we performed the test D(Mbuti,Papuan/Australian,Mamanwa, NorthernNegrito), which indicates that Australo-Papuans form a clade with Mamanwa relative to Northern Negritos (Table S5D). In addition to Southern Negrito ancestry, Mamanwa show slightly higher genetic affinity to Papuans relative to Australians, which could reflect an additional ancestry derived from an expansion of Papuan-related populations into southeastern Philippines (i.e. Sangil, Lambangian, Tboli, and Blaan), and eastern Indonesia, that is post Australo-Papuan divergence (discussed further in Section 4.5).

We constructed seven admixture graph models for Mamanwa, where all models portray Mamanwa as an admixture between Australasian and Cordilleran-related ancestries (Figures S5A-G). Models where the Australasian-related ancestry in Mamanwa are exclusively derived from either via divergence within Philippine Negritos or divergence from Papuans post Australo-Papuan split are both rejected (Models 1 & 2, Z=-8.8 and Z=-6.7, respectively). A model where the Australasian-related ancestry in Mamanwa is exclusively derived from the ancestor of Australo-Papuans prior to the split between Australians and Papuans has a better fit, but nevertheless rejected (Model 3, Z=-4.0). One of the best fit models is when, apart from the presence of East Asian-related ancestry, the Australasian-related ancestry in Mamanwa is derived mainly from an ancestral population that diverged from ancestral Australo-Papuans and minimally, at ~5%, from a Papuan-related population post Australo-Papuan divergence (Model 4, Z=-1.7). A model with Manmawa forming a clade with Ayta plus an additional Papuan ancestry, without accounting for East Asian admixture, is rejected (Model 5). An alternative model where an ancestral Negrito group admixed with East Asians prior to divergence into Northern and Southern Negritos is also rejected (Model 6). A model where after the divergence of ancestral Negritos, East Asian admixture is accounted for separately into Northern and Southern Negritos, and Southern Negritos subsequently received additional Papuan-related ancestry is not rejected (Model 7). Hence, we can not distinguish between the Model 4 and Model 7, where Mamanwa forms a clade with AustraloPapuans plus additional Papuan ancestry. Although the latter has better Z score of 1.7 relative to the Z score of 2.02 of Model 7.

Our findings indicate that the ancestors of Mamanwa came about as an offshoot population of Basal Oceania that diverged around 37 kya, and entered Mindanao prior to divergence of Australians and Papuans (Fig. 2A, Figs. S4J,S5D). Their entry into Mindanao is likely through the Sulu archipelago, during a period when the islands were interconnected with each other, and passable during the LGM, except for the narrow Sibutu passage (Figs. S2E,F). Moreover, we also observe some minimal gene flow from Papuans that is post Australian-Papuan divergence (Figs. S4K,S5D and Table S5E), which points to transoceanic migrations of Papuans or gene flow of Papuan-related ancestry from Papua New Guinea into Eastern Indonesia and then into southern Philippines (see Section 4.5).

The seprate pathways for migration of Northern and Southern Negritos are largely based on geography. The archipelagic nature of the Philippines allows two portals of entry for migrating groups from Borneo Island (or Sundaland landmass during the LGM). One is via Palawan to get to the north reaching Mindoro Island and then subsequently Luzon Island. Palawan, Mindoro, and Luzon Islands are inhabited today by Negrito or Negrito-like ethnic groups (Ayta, Agta, Atta, Batak, Mangyan Iraya). The second pathway to get to the south is via Sulu archipelago to reach Mindanao Island, where Southern Negritos (Mamanwa) or ethnic groups with Southern Negrito ancestry (Ata Manobo, Matigsalug, Tigwahanon, Teduray, etc) reside.

The alternative scenario would be for ancestral Negritos to enter only a single portal of entry. In this case, if they entered Mindanao Island via Sulu archipelago, they have to cross Samar and Leyte Islands in the east coast to reach Luzon, and then cross the long stretch of Bicol peninsula to reach southern Luzon, and go all the way to head back southward by traversing Verde Island passage to reach Mindoro Island and Mindoro passage to reach Palawan Island. It will be of the reverse order if the ancestral Negritos only entered the Palawan port of entry. Although these models are less parsimonious, we can not exclude this possibility.

4.4 Negrito ancestry among non-Negritos

Negrito ancestry is uniformly detectable in almost all non-Negrito ethnic groups, depicting historical admixture between indigenous Negritos and more recent migrants. We utilize the test *D(Mbuti;AytaMagbukon,Balangao,X)* to examine the presence of Northern Negrito ancestry, in contrast to Cordilleran ancestry, among non-Negritos, using Ayta Magbukon as the surrogate for Northern Negritos (Fig. S4N and Table S5B). As expected, we observe that Northern Negrito ancestry is only found among non-Negrito ethnic groups of Luzon. Mangyan ethnic groups, particularly Mangyan Iraya, a self-identified non-Negrito group of Mindoro island, appear to have stronger than usual genetic affinity with Ayta, relative to any other non-Negrito groups of northern Philippines. This may explain their documented phenotypic features observed in Mangyan Iraya that are commonly ascribed to Negritos, as well as the linguistic evidence which indicates a close relationship between Mangyan Iraya and Central Luzon Ayta languages (61). The high Northern Negrito ancestry in Manygan Iraya was consistently observed when we estimated the relative proportion of Australasian-related vs Cordilleran-related ancestries among non-Negritos of northern Philippines using qpAdmix (Table S5M).

Among non-Negritos of Luzon, the Bugkalots, also known as Ilongot, displayed the highest detectable level of Northern Negrito ancestry (Fig. S4N and Tables S5B,M). Bugkalots are found across the provinces of Nueva Ecija, Nueva Viscaya, Qurino, and Aurora of Luzon. The Bugkalot participants in this study are from Nagtipunan municipality of Quirino, which lie close to the cultural communities of Arta, Agta Maddela and Agta Casiguran. Hence, due to geographical proximity and consequent admixture, this may explain the high levels of Northern Negrito ancestry detected in this group. Next to Bugkalot, high levels of Northern Negrito ancestry were also detected among Ga'dang, Apayao, Itneg, Kalinga, & Malaweg which are ethnic groups found in contiguous provinces of Abra, Apayao, Cagayan & Kalinga. Atta Negritos are known to reside in these area, including Atta Rizal (included in our study) & Atta Pamplona of Cagayan, Atta Pudtol of Apayao, and Atta Villa Vicsiosa of Abra (which is known to speak a language that was established to be already extinct). Again, geographical proximity and consequent admixture between Negritos and non-Negritos may explain the high Northern Negrito ancestry detected among these aforementioned groups.

Interestingly, some non-Negrito groups of northern Luzon, such as the Cordillerans of central mountain range, appear to be largely unadmixed or presented with undetectable levels of Negrito ancestry (Figs. 2B-D & S4D,N,O). This is in spite of the fact that Cordillerans reside in a confined area of mountainous topography, surrounded by Negrito groups (Ayta in the south and Agta/Atta/Arta/Dumagat in the east and the north), and that Cordillerans had documented historical trade relations with Negritos (62-64). Therefore, these groups of the Cordillers represent an example in human demographic history, where longstanding interaction between two genetically distinct populations resulted in no dectable levels of population admixture.

The Negrito ancestry detected among non-Negrito groups of southern Philippines is mainly attributed to Southern Negrito or Mamanwa-related ancestry. The ancestry component is detected by the test *D(Mbuti,Papuan,Amis,X)*, where we examine whether any X non-Negrito population share more alleles with Australians and Papuans relative to Amis (using Papuans as surrogate for Basal Oceanian ancestry found in Mamanwa) (Tables S5G), and the test *D(Mbuti;X,Australian,Papuan)*, to examine whether such a signal in X populations is equidistant to Australian and Papuan, which is expected for Southern Negrito branch of Basal Oceania (Fig. S4O and Tables S5E,F). Consistent with the findings in section 4.2, we find that the Southern Negrito ancestry found in these groups primarily forms an outgroup to the AustraloPapuan clade, with some groups receiving minimal gene flow from Papuan-related ancestry that is post Australian-Papuan divergence (see section 4.5 below). Moreover, the highest signal of Southern Negrito ancestry among non-Negrito groups of southern Philippines is found among Manobo ethnic groups, particularly among Ata Manobo, Davao Matigsalug, and Bukidnon Matigsalug, and Bilic-speaking groups (Figs. S2B-D and Tables S2-4).

4.5 Papuan-related signal in southern Philippines

With the test *D(Mbuti;X,Australian,Papuan)*, we investigated whether any population X share more alleles with either Papuans or Australians, or whether any population X possesses a Papuan-related ancestry that is post-divergence between Australians and Papuans. We find that some ethnic groups of southeastern Philippines possess significant Papuan-related ancestry (Fig. S4O and Tables S5E,F). The Papuan-related gene flow is highest among Sangil ethnic group, an Islamized and mostly agriculturalist/fisherfolk communities that not reside not only in the southern tip of Mindanao, such as Glan Municipality and Sarangani Islands, but also, in greater numbers, in Northern Sulawesi and Gorotalo provinces of eastern Indonesia. With Phil_AsiaPacific_315K dataset, we also find a significant Papuan-related ancestry in ethnic groups of eastern Indonesia covering Lesser Sunda Islands, North Maluku, Sulawesi, and East Kalimantan of Borneo (Tables S5F,H). With Phil_HO_201K dataset, higher Papuan-related than Australian-related ancestry is found not only in Lebbo ethnic group of East Kalimantan, but also in all ethnic groups of Bougainville, New Britain, New Ireland, Manus, Vanuatu, and the Solomon islands (Tables S5F). This would imply that the Australasian ancestry detected in all ethnic groups of Oceania, eastern Indonesia, and some ethnic groups of southeast Philippines is largely an effect of a recent transoceanic migration of Papuan-related populations, after the divergence between Australians and Papuans (Tables S5E,F,I).

What drove the long-distance migrations of Papuan-related populations still remains unclear. Although a key event that happened in Papua New Guinea after the Papuan-Australian divergence 25 kya is the independent development of agriculture dated around 10 kya (31). Population expansion after this period, together with technological innovations in seafaring (either independently developed or acquired through interaction with a nearby non-Papuan ocean navigators), and increasing inter-island trade, may altogether have played a role in the movement of Papuan-related groups into other islands.

4.6 Australasian ancestry in Mainland Asia and Americas

Based on various analyses using Admixture and combination of D tests, we find that Philippine Negrito ancestry is geographically confined within the Philippines. Using the test (*Mbuti; X, Onge, Papuan)*, where we examine whether any population X has more Onge-related or AustraloPapuan-related ancestry, we find that the Australasian ancestry present in ethnic groups of western Indonesia, East Asia, MSEA, and South Asia is mainly attributed to Onge/Jarawa-related or Hoabinhian-related ancestry (Table S5J). Interestingly, all D test combinations of *D(Mbuti;Papuan,Cordilleran,Amis)* results in positive value, with a Z score of at least 2 when either Ifugao Tuwali or Kalanguya was used as a surrogate for Cordilleran (Table S5K). If this result is confirmed in a larger sample size of Amis or a more comprehensive coverage of indigenous Taiwanese ethnic groups, this would indicate an Australasian-related gene flow with the direction from ethnic groups of the Philippines into indigenous Taiwanese groups.

Earlier publications reported a presence of Australasian-related ancestry among indigenous Native American communities of Brazil, Surui and Karitiana(9). This was suggested to be either a two wave founding of the Americas or a more recent gene flow of Australasian-related ancestry into Native Americans. The Australasian signal was detected in a single ancient individual from Lagoa Santa, Brazil in a recent study, which suggests an ancient population structure, whereby a subset population among the first Native Americans carry some Australasian genetic ancestry(65). Given these findings, we utilized the test *D(Mbuti;X,Mixe/Piapoco/Pima,Y)*, to examine whether any Australasian X population (Papuan, Australian, Philippine Negrito, Malay Negrito, or Andaman Island Negrito) shares more alleles with any Native American Y population relative surrogates for least admixed Native American populations (Mixe, Piapoco & Pima). We find that Onge & Jarawa, AustraloPapuans, Northeast Luzon Negritos, and Southern Negritos are the best surrogates for the Australasian signal detected among Surui and Karitiana (Table S5L).

5 Post-glacial northward migrations of non-Negritos into the Philippines

5.1 Detection of distinct Manobo ancestry in southern Philippines

We used the test *D(Mbuti,AtaManobo,Balangao,X)* to determine whether any population X share more alleles with the least admixed Manobo, Ata Manobo, or with a representative Cordilleran, Balangao (Fig. S6A). Manobo ancestry is detectable among inland ethnic groups of southeastern Mindanao covering the provinces of Bukidnon and Davao del Norte and Paquibato district of Davao City (Davao Matigsalug, Bukidnon Matigsalug, and Bukidnon Tigwahanon). This is consistent with our findings in the Admixture analysis, where Ata Manobo, Davao Matigsalug, Bukidnon Matigsalug, and BukidnonTigwahanon represent the least admixed populations that define the Manobo cluster (Figs. S2B-D & Tables S2-4).

We investigated further the geographic spread of Manobo ancestry in southern Philippines by masking away the non-Manobo alleles of Mindanao ethnic groups. We implemented RFMix (23) on shapeit-phased (66, 67) subset of Phil_1KGP_SGDP_1.69M dataset using the flag -fb 1 -e 1 -n 10 -u 1, and discriminate the different ancestries within the Philippines using the following reference populations: 15 Ifugao Cordillerans for the least admixed Cordilleran, 15 combination of AtaManobo, Davao Matigsalug & Bukidnon Matigasalug for the least admixed Manobo, 15 combination of Sama Dilaut populations for the least admixed Sama, 15 combination of Ayta Magbukon & Ayta Ambala for the least admixed Negrito, and 15 Papuans for the least admixed Papuan. Using the test *D(Mbuti,AtaManobo,Balangao,*X^{rf}*)*, where X^{rf} is any population where non-Manobo alleles were masked away, we detect significant levels of Manobo ancestry in all Manobo ethnic groups as well as Danao (Meranao, Maguindanao, Iranun), Bilic (Tboli, Obo, Bagobo Klata, Blaan Koronadal, Blaan Sarangani), Mansakan (Kalagan, Mandaya, Mansaka, Tagakaulo) and Sangil ethnic groups, indicating an ancient island-wide geographic distribution of Manobo ancestry in Mindanao (Fig. S6B).

5.2 Manobo ancestry is an outgroup to Cordilleran-Taiwanese clade

We investigated the relationship between indigenous Taiwanese, Cordillerans, and Manobos with the test; *D(Mbuti,Balangao,X,Ami/Atayal)*, where we examine whether Cordillerans forms a clade with indigenous Taiwanese populations, Amis or Atayal, relative to any other Manobo ethnic group. As expected, when X is another Cordilleran, D score is significantly negative, which shows that Cordillerans form a clade together relative to Amis or Atayal (Figs. S6C,D and Table S6A). On the other hand, when X is a Manobo ethnic group, D score is significantly positive, indicating that Cordillerans forms a clade with Amis or Atayal relative to any other Manobo ethnic group, or that Manobo ancestry is an outgroup to Cordilleran-Taiwanese divergence. This is consistent when we implment the following tests: *f3(Mbuti;Balangao,Amis), f3(Mbuti;Balangao,Manobo)* or *f3(Mbuti;Manobo,Amis)*, where we show that Amis and Balangao share more drift with each other relative to any Manobo ethnic group (Table S6B).

We then probed further into the relationships between Manobos, indigenous Taiwanese or Cordillerans, and mainland Asian or Native American populations, with the test *D(Mbuti,Balangao/Amis,X,AtaManobo)*, where we examine whether any X mainland Asian or Native American population shares more alleles or forms a clade with Balangao or Amis compared to the least admixed Manobo, Ata Manobo (Figs. S6E,F and Tables S6C,D). All Native American and Siberian populations were consistently shown to be as an outgroup to Manobo-Cordilleran/Ami clade, supporting the evidence that the east Asian component in Native Americans and Siberians is a split from Basal East Asian that is older than the subsequent split of Ancestral Manobo branch of Basal Austric (Fig. 2B). In line with the Austric hypothesis (68-72), we define 'Basal Austric' in this study as the common ancestor of what became the present-day AN-speaking, Kra-Dai-speaking & AA-speaking populations.

The *D(Mbuti,Amis,X,AtaManobo)* test additionally demonstrated that mainland Asian populations, such as Lahu, Japanese, Korean, Northern Han, Southern Han, Miao, Tujia, Dai, She and Kinh Vietnamese share more alleles with Amis than with Ata Manobo (Figs. S6E,F and Tables S6C,D). This indicates that that the East Asian ancestry component of the aforementioned mainland Asian ethnic groups forms a clade with Amis relative to Ata Manobo, and that Ancestral Manobo likely diverged from Basal Austric earlier than the split between Amis and East Asian ancestral components of Han, Dai, Kinh and other groups. This inferred topology is consistent when the test *D(Mbuti,Amis,X,AtaManobo)* was performed on both Phil_1KGP_SGDP_1.69M and Phil_HO_201K datasets (Figs. S6E,F and Tables S6C,D).

5.3 Manobos are admixed with Cordilleran-related and Southern Negrito ancestry

Taking advantage of the topology framework we used earlier, we tried to fit the topology of Manobos in relation to other Asian ethnic groups using qpGraph (Table S6E). Ancestral Manobo is best modelled as a branch of Basal Austric of East Asia that diverged earlier than the Cordilleran-Amis split and younger than the separation of the East Asian component that contributed to the ancestry of Native Americans (represented by Mixe). Provided that Cordilleran-Amis divergence is estimated to be at least ~8k years old (see section 6.4), and that the founding population of all Native Americans and ancient Beringians diverged from Basal East Asian around 36 kya, with continuous gene flow maintained from other East Asians up until ~25 kya(65) , the divergence between Ancestral Manobo and the common ancestor of Cordilleran-Taiwanese clade must have occurred sometime in between 8-25 kya. This expectation is in line with our calculated divergence time between Manobos and Cordillerans (Fig. S7E), which is around 15 kya, falling well within the postulated 8-25 kya date of divergence. The calculated Manobo-Cordilleran divergence time also falls within the Last Glacial Maximum period, right around the time when there is accessible route for people to migrate from MSEA into southern Philippines, when the land mass of Mindanao Island was only separated from continental Sundaland by the narrow Sibutu strait (Figs. S3E,F).

The most parsimonious explanation for the arrival of Manobos into southern Philippines is likely through land-based migration from southern China into Mindanao via MSEA. This migration route may have been through an ancient extended coastline of MSEA and continental Sundaland that became submerged by rising sea levels of Postglacial period, commencing from 12 kya and stabilizing until 8 kya (Figs. S3E,F).

The loss of extended coastline likely resulted in the loss of archeological evidence, which is supported by the fact that one can hardly find coastal archeological sites in Asia that is older than 8000 years (73). Alternatively, a direct 1,800-km sea route from southern China into the southern Philippines is also possible. However, there is no documented evidence for long-distance ocean navigation skills among Manobos, nor there is any, as of yet, archeological evidence for navigational technologies. Moreover, direct ocean-based migration from southern China into the southern Philippines would inevitably require crossing between Luzon, Mindoro, Panay, Negros and/or Palawan Islands of northern to middle Philippines, where our current results show that these aforementioned islands are inhabited by populations with low to undetectable levels of Manobo ancestry.

Following their entry into southern Philippines, at least two major admixture events have occurred among Manobo-like populations. The first is accounted for by admixture with the earliest populations in the southern Philippines, the Southern Negritos (Table S6F). This is supported by the test, *D(Mbuti,Papuan,Balangao,X)* & *D(Mbuti,Australian,Balangao,X)*, where we find significant AutraloPapuan-like ancestry among all Manobo ethnic groups similar to the Southern Negrito ethnic group, Mamanwa. The second major admixture event is brought about by the expansion of Cordilleran-like populations. This can be demonstrated by spatial interpolation of *D(Mbuti,Cordilleran,X,Dai)*, where we discriminate between Manobo-related vs Cordilleran-related ancestry in any X Manobo ethnic group (Fig. S7G). A positive D test would indicate more allele sharing between X Manobo population and Cordillerans due to substantial admixture with Cordilleran-like ancestry, while a negative D test would indicate that X Manobo population possesses higher ancestral Manobo ancestry or ancestry that is older than the Cordilleran-Dai divergence. Remarkably, the spatial interpolation of *D(Mbuti,Amis,X,Dai), D(Mbuti,Cordilleran,X,Dai)* or *D(Mbuti,Cordilleran,AtaManobo,X)* shows that the greatest impact of Cordilleran-like admixture is found among the coastal populations of Mindanao (Figs. S7F-H). This scenario is plausible due to the fact that the expansion of Cordilleran-like ancestry is likely brought about by the migration of expert ocean navigators from South China-Taiwan area. Thus it is not unexpected to see the greatest impact of Cordilleran-related gene flow in populations lying across the coastal areas of Mindanao.

Another ancestry component, although minor, found in some Manobo and other related ethnic groups of the southern Philippines is the Papuan-related one (see Section 4.4). We infer this as likely to be a gene flow or movement of Papuan-related populations from Papua New Guinea into southern Philippines which occurred post Australian-Papuan divergence, given that *D(Mbuti,X,Australian,Papuan)* in these X Mindanao populations (Lambangian Manobo, Sangil and Bilic-speaking Blaan populations) are significantly positive (Fig. S4O). This indicates that these populations share more alleles with present-day Papuans than with Australians, similarly observed among all populations of Oceania covering Bougainville, Samoa, and Vanuatu islands, as well as ethnic groups in various islands of eastern Indonesia. The test *D(Mbuti,X,Australian,Papuan)* demonstrates that the epicenter of Papuan-related genetic signal is expectedly found closer to the Papuan mainland, with a gradual dilution of the signal westwards into Indonesia and southern Philippines, and with a variable signal intensity detected eastwards into Oceania (Tables S5E,F). The greatest impact of this signal in the southern Philippines is among the southern and eastern coastal populations, which supports a northwest directionality of gene flow from Papua New Guinea into eastern Indonesia and southeastern Philippines.

5.4 Sama ancestry in southwestern Philippines

Sama Dilaut sea nomads are one of the most culturally distinct ethnic groups of the Philippines. Aside from their uncommon ocean-based lifestyle, they speak a distinct language that is not classified under Philippine languages. Oral traditions among Sama Dilaut sea nomads of southwestern Philippines would highlight peninsular Malaysia of MSEA, specifically Johor, as the point of origin of their ancestors (74, 75). In our analyses presented below, we show that Sama-related populations are genetically affiliated to AA-speaking ethnic groups of MSEA, supporting the narrative of their oral traditions, and providing additional incontrovertible evidence for a south-to-north migration of some ancient populations from MSEA into Indonesia and the Philippines(12, 76, 77), which is separate and distinct from the north-to-south migration of Cordilleran-related populations from Taiwan/South China area into the Philippines and the rest of ISEA.

Our admixture analysis using different panels demonstrated that Sama and Sama-related ethnic groups consistently form their own genetic cluster (Figs S2B-D and Tables S2-4). We formally investigated the presence of Sama ancestry using the test *D(Mbuti,SamaDilautBongao,Cordillleran,X)*, to address whether any population X possess a ubiquitous Sama-related and non-Cordilleran ancestry, utilizing Sama Dilaut of Bongao, Tawi-Tawi island as the reference population for the least admixed Sama (Fig. S6G). We found that this can indeed be found among coastal-dwelling Sama (Kabingaan & Banguigi), inland Sama (Sama Deya of Bongao) and SamaDilaut cultural communities of Zamboanga peninisula (Taluksangay and Mampang), Sama-related ancestry is also detectable among seafaring population of Borneo, Bajo Derawan (Fig. S6G). To distinguish between Sama vs Manobo-related ancestries in southern Philippines, we plotted an interpolation map of the test *D(Mbuti,X,AtaManobo,SamaDilautBongao)*, which examines whether any population X contains alleles that are predominantly Manobo-related or Sama-related (Figs. S7I,J). We observe a greater Manobo-related ancestry in most ethnic groups of mainland Mindanao, including Manobo, Danao, Bilic, Sangil & Mansakan-speaking populations, and a higher Sama-related ancestry among Sama & Subanon-speaking ethnic groups of the Sulu archipelago and Zamboanga peninsula of Mindanao.

Using the RFMix-treated subset of Phil_1KGP_SGDP_1.69M panel, we ran *D(Mbuti,SamaDilautBongao,Balangao,*X^{rf}*)*, where X^{rf} is any population which had their non-Sama alleles being masked away and thus retain only the alleles that are distinct for Sama-related populations. We observe a significant detectable levels of Sama ancestry among ethnic groups of the southwestern Philippines as well as Palawan Island. Sama ancestry is detected in Yakan of Basilan Island, Tausug of Sulu Island, Subanon and Subanen groups of Zamboanga peninsula, and Pa'lawan, Molbog & Tagbanwa of Palawan Island (Fig. S6H). This indicates that aside from entry of Sama-related populations into islands of the Sulu archipelago, Sama-like groups may also have independently entered Palawan Island from Borneo, and contributed to the ancestry of present-day Pa'lawan, Molbog & Tagbanwa ethnic groups.

5.5 Sama ancestry is affiliated to Htin/Mlabri-related ancestry of MSEA

Our Admixture analysis using the Phil_HO_201K dataset indicate that both Sama and Manobo-related populations share ancestry that cluster together with Htin & Mlabri AA-speaking ethnic groups of MSEA (Fig. S2D). We then utilze the test *D(Mbuti,X,AtaManobo,Htin/Mlabri)* to examine whether any X least admixed Sama-related ethnic groups exhibits greater genetic affinity with either the least admixed Manobo (Ata Manobo) or the least admixed AA-related populations (Htin or Mlabri); and the test *D(Mbuti;X,AtaManobo,SamaDilautBongao)*, to examine whether any X AA-speaking population shares more alleles with AtaManobo or with Sama Dilaut. We find that populations with high Sama ancestry share more alleles with Mlabri and Htin populations relative to Manobo Ata (Figs. S6K,S7J,K and Tables S6G,K,L). This AA-related genetic signal is extensively found in southwestern Philippines, including all Sama Dilaut and inland Sama groups, Palawanic populations such as Tagbanwa, Molbog, and Pa'lawan and Zamboanga peninsula populations such as Subanon and Subanen (Fig. S6K and Table S6G).

Using qpgraph, a simple model where Htin and Sama forms a clade together is not rejected, as opposed to other alternatives where Manobo and Sama or Manobo and Htin forms a clade together (Figures S7A-C). When we construct a more complex admixture graph, Sama is modelled to derived marjority of its ancestry (80%) from a Htin-related ancestral group (Figure S7D).

The non-Manobo AA-related genetic signal was also detected amongst Lebbo, Indonesian Bajo, Malay, Nicobarese, ethnic groups of MSEA & mainland East Asia, and AA-speaking ethnic groups of India (such as Juang and Munda) (Table S6H). Our observations are in line with previous findings that an AA-related genetic signal extends into western Indonesia covering the islands of Borneo, Sumatra, and Java (12, 76). Thus, our findings further extend the geographic region of Mlabri/Htin-affiliated genetic ancestry from western Indonesia into southwestern Philippines, covering the ethnic groups of Palawan, Sulu archipelago, and Zamboanga peninsula.

5.6 Topology of Sama, Htin/Mlabri, and Manobo-related populations of Southeast Asia

Using *D(Mbuti;Balangao/Amis/Atayal,X,SamaDilautBongao)*, we investigated whether any population X forms a clade with Cordillerans/indigenous Taiwanese or whether population X is an outgroup to Sama Dilaut + Cordilleran/indigenous Taiwanese clade (Fig. S6L and Table S6I). We consistently find that Native Americans and Siberians are an outgroup to a clade that includes mainland East Asian and non-Negrito MSEA & ISEA populations. The latter has a common ancestor that we labelled as Ancestral Austric, in line with the Austric hypothesis that AN-speaking, AA-speaking, Kra-Dai-speaking & Hmong-Mien-speaking populations share a common ancestor (70). We find that the shared Sama/Manobo/Htin/Malbri ancestry is a branch of Basal Austric that diverged from Basal East Asia ~15 kya, which subsequently diverged into Sama/Htin/Mlabri vs Manobo branches ~12 kya (Fig. 2B & Fig. S7E). Given this topology, Ancestral Sama is older than the divergence between Cordilleran/Amis/Atayal and East Asian component of Han, Dai, Japanese and Kinh.

The ancestral AA-related genetic signal most likely originated from southern China, traversed through MSEA, likely via the Mekong river. The latter expansion includes the spread of Htin/Mlabri-related genetic signal, via Sundaland, into western Indonesia and southwestern Philippines and contribution of Htin/Mlabri-related ancestry into AA-speaking ethnic groups of east India (to become the present-day Munda ethnic groups) (78, 79). The spread of Htin/Mlabri-related genetic signal into southwestern Philippines most likely happened post-entry of Manobo-like populations into mainland Mindanao, given the disparate geographical distributions of Manobo and Sama ancestries in southern Philippines (Fig. S7I), and given that Manobo diverged earlier from a common ancestral Sama/Manobo/Htin/Malbri population (Fig. S7E), and that Ancestral Sama still has to diverge from a common Sama/Htin/Mlabri population in MSEA before entering western Indonesia and southwestern Philippines (Figs. 2B & 3B,C).

5.7 Detection of South Asian-genetic signal in Sama Dilaut

Various cultural communities of ISEA had long-distance historical trade with the Indian subcontinent via the Indian Ocean Trading Network since the 1st millennium BCE (80) until the commencement of colonial period. Along this period are various Hindu-Buddhist Kingdoms or empires including, among others, Srivijaya (650-1377), Medang (732-1006), Kediri (1245-1221), Singhasari (1222-1292) and Majapahit (1293-1527). Both Srivijaya and Majapahit kingdoms had ruled over a wide geographical area covering coastal MSEA, western Indonesia, & Malaysia, and may have exerted influences as far as Sulu archipelago of the Philippines. It is then not surprising to find South Asian genetic signal among lowlander Malays, Javanese, Balinese, Sumatran and Bajo populations of Indonesia, which provide evidence for the historical long-distance interactions between ISEA and the Indian subcontinent.

Admixture analysis of Phil_AsiaPacific_315K & Phil_HO_201K datasets in section (Figs. S2C,D) revealed a low amount of West Eurasian-like ancestry among Sama Dilaut ethnic groups, which can be attributed to gene flow from South Asian populations with high West Eurasian ancestry (5, 81-83). We then formally tested this by using the test *D(Mbuti;X,AtaManobo,Sama)*, where we investigate whether any X South Asian or control population share more alleles with Sama-related ethnic groups relative to Ata Manobo (Table S6J). We find that Sama Dilaut sea nomads of Sulu archipelago and Sama coastal dwellers of Zamboanga peninsula exhibited evidence of gene flow from South Asians. The best surrogate for South Asian genetic signal are populations with high West Eurasian ancestry or populations labelled as ‘Ancestral North Indian’ (9). The admixture date calculated via LD-based method, Malder, revealed a South Asian gene flow in Sama populations around 750 +/- 150 years ago, which is before the period of Spanish colonization, and well within the period when ISEA was involved in an active trading network with the Indian subcontinent (Table S6M).

6 Origins and genetic legacy of Cordillerans

6.1 Cordillerans are the least admixed descendants of Basal East Asian

Consistent with previous findings (7, 15) and with all ADMIXTURE analyses we performed (Figs. S2B-D and Tables S2-4), Kankanaey of Mountain Province display no evidence of admixture with Negrito or Papuan-related ancestry. Aside from Kankanaey, we additionally have shown for the first time that other Cordillerans: Bontoc and Balangao of Mountain Province, Tuwali, Ayangan, and Kalanguya of Ifugao, and Ibaloi of Benguet, also displayed with no detectable signal of admixture with Australasians, Negritos or Papuans. We formally investigated this with the test *D(Mbuti;Papuan,Han,X)*, where we examine whether any Philippine population X had exchanged gene flow with Papuans, relative to Han, a population which is known to have no historical admixture with Papuans or Negritos (Fig. S8A and Table S7A). Papuans were used as a surrogate for the least admixed Australasian ancestry, given that all Philippine Negritos are admixed with Cordilleran-related ancestry. For the above-listed Cordilleran populations, our results show a *D(Mbuti;Papuan,Han,X)* score of close to zero, further providing statistical evidence that Kankanaey, Bontoc, Balangao, Tuwali, Ayangan, Kalanguya & Ibaloi are unadmixed with Australasian ancestry. In contrast, other populations that are linguistically classified as Cordillerans, and that reside outside of Mountain Province, Ifugao, and Benguet provinces of Luzon (Apayao of Apayao province, Bugkalot of Quirino province, Itneg of Abra province, Kalinga of Kalinga province, and Pangasinan of Pangasinan province), displayed admixture with Australasians, both detectable in ADMIXTURE analysis as well as with *D(Mbuti;Papuan,Han,X)* test (Figs. S2B-D & S8A and Tables S2-4 & S7A).

We further confirmed the lack of admixture between central Cordillerans and Negritos, using the test *D(Mbuti;RfAytaMagbukon,Han,X)*, where we examine whether any population X, relative to Han, shares more alleles with RFMix-processed Negrito, RfAytaMagbukon (Fig. S8C). The processing with RFMix for Ayta Magbukon was implemented to remove all East Asian and Cordilleran-related ancestry, and to retain alleles that are classified as Negrito ancestry only. Again, Kankanaey, Bontoc, Balangao, Tuwali, Ayangan, Kalanguya & Ibaloi displayed no evidence of admixture with Negritos. All other non-Negrito ethnic groups of northern Philippines, including other Cordillerans (Apayao, Bugkalot, Itneg, Kalinga & Pangasinan), did exhibit significant admixture with Negrito-related ancestry.

We also performed f3 Admixture tests *f3(Papuan,Ami,X), f3(AytaMagbukon,Amis,X),* & *f3(AgtaManide,Amis,X)* to examine whether any target Cordilleran population X is an admixture between East Asians, represented by Amis, and Australasians, represented by Papuan or least admixed Ayta and Agta Negritos, Ayta Magbukon and Agta Manide, respectively. Consistently we observe no evidence of admixture for Kankanaey, Bontoc, Balangao, Tuwali, Ayangan, Kalanguya & Ibaloi Cordillerans with Papuan, Ayta, or Agta-related ancestry (Fig. S8B and Table S7B).

The finding that some non-Negrito groups of northern Luzon, such as Cordillerans of central mountain range, remain largely unadmixed or presented with no detectable evidence of Negrito ancestry is highly unusual in human demographic history. All other ethnic groups in Asia-Pacific region are either admixed with Hoabinhian-related (MSEA), Papuan-like (southern Philippines, eastern Indonesia, and Oceania), Northern Negrito (northern Philippines), Southern Negrito (southern Philippines), AA-related (western Indonesia, southwestern Philippines, and MSEA), or northern East Asian-like (Taiwan and mainland East Asia) ancestries. Hence, being the least admixed representative for East Asian, it is not surprising to find Cordilleran groups to consistently define the axis of PC1 in worldwide PCA, in polar opposite to the African Khoe-San ethnic groups at the other extreme (Fig. 1B & Figs. S1B,C).

Using the test *D(Chimp;X,BalitoBay,Y)*, we examine the level of shared ancestry between any worldwide population X and Asia-Pacific population Y (Balangao, Amis, Han or Papuan, plus control CEU population) relative to the least admixed African, Balito Bay (Table S7C). The highest D scores are found among the comparisons between Balangao and worldwide populations, followed by Amis and then Han, indicating that Cordillerans possess the least admixed signal of Basal East Asian ancestry that is shared by a wide range of populations. We extended this investigation by using the test *D(Mbuti;Balangao,Papuan,X),* where we examine whether any X population possess a Cordilleran-related ancestry that is found in the least admixed Cordilleran, Balangao, relative to Papuans (Table S7D). We demonstrate that Cordillerans, being the least admixed, share the common Basal East Asian ancestry with populations from a wide geographic area covering Asia-Pacific and Americas. These findings were consistent when we substituted X in the test *D(Mbuti;Balangao,Papuan,X)* with ancient individuals or populations (Table S7E).

6.2 Cordilleran ancestry is a surrogate genetic signal for AN-speaking populations

All Admixture analyses we performed reveal Cordillerans, and not Amis or Atayal, as the population that defines the unadmixed genetic cluster for AN-speaking populations (green cluster in Figs. S2B-D). This cluster is found in varying proportions, admixed with other ancestries, amongst AN-speaking populations of Malaysia, Indonesia, Oceania, and Taiwan. For instance, in the Admixture plot with Phil_HO_315K panel, both Ami and Atayal not only possess 61-65% of the Cordilleran cluster, but also with 28-31% of AA-related (defined by Htin & Mlabri ethnic groups) and 1.5-3.4 % of nEA-related cluster (defined by Chukchi, Eskimo, and Koryak ethnic groups). We formally tested the presence of AA-related and nEA-related genetic ancestries in Amis/Atayal with *D(Mbuti;X,Balangao,Amis/Atayal),* where we investigate whether any X North Asian, East Asian or MSEA population share more alleles with Amis/Atayal relative to Cordillerans. We find that populations that have high ancestral nEA or AA-related ancestries share more alleles with Amis or Atayal than with Cordillerans (Figs. S8D,F-H and Tables S7F,G).

Additionally, present day Cordillerans, and not Ami or Atayal, serve as the best surrogate for the AN-related genetic signal of all AN-speaking ethnic groups of the Philippines. With the test *D(Mbuti;X,Amis/Atayal,Cordilleran)*, we evaluate whether any X Philippine ethnic group is more genetically affiliated with Ami/Atayal or with any Cordilleran group (Table S7H). We find that all Philippine populations may they be Negrito, Sama-like, Manobo-like or any other ethnic group, share more alleles with Cordillerans than with Ami or Atayal. This is likewise confirmed with the test *D(Mbuti,X,Papuan,Y*), where we examine whether any Philippine population X, accounting for Australasian admixture in them (represented by Papuans), share more alleles with any Y Cordilleran, Ami or Atayal ethnic group (Table S7I). Again, all

Philippine ethnic groups display greater genetic affiliation with Cordillerans than with Ami or Atayal.

We also extended our investigation to examine whether Cordillerans are the best surrogate for AN-related ancestry among AN-speaking populations of Malaysia, Indonesia and Oceania. With the test *f3(Mbuti,X,Y)*, where we examine which population X shares the most drift with Malaysian or Indonesian populations Y (Table S7J). As expected, Dusun, Malay, Murut, Lebbo, and Indonesian Bajo shares the most drift with the least admixed Cordillerans. For admixed Oceanian populations, we used the test *D(Mbuti;X;Papuan,Y)* to examine which populations X share the most alleles with Oceanian population Y while accounting for Papuan-related admixture in them (Tables S7K,L). All Oceanian populations in Vanuatu as well as Manus, New Ireland, and New Britain provinces of Islands region, Papua New Guinea, share the most alleles with Cordillerans.

We also looked into whether ancient individuals from northern Philippines, peninsular Malaysia and Oceania share the most drift with Cordillerans, using the test *f3(Mbuti,X,Y)*, where X is an ancient individual/population and Y as Cordillerans or any AsiaPacific population (Figs. S8I,J and Table S7M). As expected, the 1880-year-old Philippine individual in Nagsabaran site of Northern Luzon is found most genetically affiliated with Cordillerans. Moreover, historical Malays from Supu Hujung4 and Kinabatagan sites of Sabah, Malaysia as well as 2900-year-old Lapita individuals from Vanuatu also share the most drift with Cordillerans relative to any other present day populations of AsiaPacific region.

6.3 Relationship of Cordillerans to indigenous Taiwanese & Batanic Islanders

Batanes is an archipelago that lies at the northernmost boundary of the Philippines, across the Luzon strait in between Taiwan and Luzon island. It has been argued that the Batanic islands are one of the probable initial settlement sites for AN-speakers from Taiwan prior a subsequent southward migration into Luzon (84). This model of demographic movement would require a nested population structure whereby Bashiic-speaking populations (Ibatan, Ivatan, and Itabayaten) should cluster together with Cordillerans, relative to indigenous Taiwanese. However, such a model is rejected using qpGraph. The model that fits requires the following: i.) Cordilleran ancestry in Ivatan, Balangao, and Amis/Atayal form a trifurcation rooted from a shared common ancestral source, ii.) Amis & Atayal receive additional input from an admixed population with nEA ancestry, and iii.) Ivatan receive additional gene flow from both Australasian and nEA sources (Table S7N).

Hence, a parsimonious interpretation of these findings is that an ancestral Cordilleran population from south China area gave rise to genetically equidistant populations: one group that went directly to mainland Taiwan and diverged to become Amis & Atayal, another to Batanic islands to become Itbayaten, Ivatan & Ibatan, and another to mainland Luzon to become the present-day Cordillerans. Both Amis and Atayal subsequently admixed with a population with high nEA ancestry, consistent with the evidence of detectable nEA ancestry among these groups (Figs. S2B-D & S8D,F-H, and Tables S2-4). An alternative scenario for the shared Cordilleran ancestry observed among Amis and Atayal is the initial presence of an ancestral population with nEA ancestry (consistent with nEA ancestry observed in 8,000-year-old Liangdao individuals from neighboring islands of Matsu archipelago, see Section 7), and a second major migration of Cordilleran-related population from South China (or from Luzon) into Taiwan which subsequently admixed with the initial settlers.

The ancestral Bashiic group settled in one of the islands of Batanes, who subsequently diverged into Itabayaten and Ivatan/Ibatan ancestral populations (Table S7N). The former settled in Itbayat Island, while the latter subsequently diverged into Ivatan which settled mainly in Batan & Sabtang Islands, and Ibatan who settled mainly in Babuyan Island. Provided the intervisibility between the islands and documented evidence for inter-island exchanges, it is not surprising to find Bashiic groups to have received gene flow from populations with nEA ancestry (likely from an admixed indigenous Taiwanese population or directly from mainland China ethnic groups) and Australasian ancestry (from Luzon ethnic group with Northern Negrito ancestry). An alternative scenario is that the ancestral population that arrived in Batanes from South China/Taiwan area may have been an admixed population with Cordilleran-related and nEA ancestries.

6.4 Into the Philippines in multiple pulses

An recent comprehensive review on the origins and dispersal of AN languages has presented an argument that the expansion of MP languages likely happened in various pulses at different points in time (85). In this study, we provide genetic evidence for a non-monolithic expansion of Cordilleran-related populations from South China-Taiwan area into the Philippines. We used the tests *D(Mbuti,Balangao,X,Amis),* where we examine whether the Cordilleran-related ancestry in any X population forms a clade with Cordillerans or is an outgroup to Amis and Cordilleran, and *D(Mbuti;Balangao,ManoboAta,X),* where we examine whether any population X forms a clade with central Cordillerans relative to Manobo Ata, or that Cordilleran-related ancestry in these populations is younger than the Manobo-Cordilleran divergence (Tables S7O,P). We find a non-uniform distribution of genetic relationship between non-Negrito ethnic groups of middle to northern Philippine and non-Negrito, non-Manobo, non-Sama ethnic groups of southern Philippines relative to Balangao. Some groups (Bashiic and Cagayanic) appear to form a clade with Cordillerans or (Mansaka and Meranao of southern Philippines) are equidistant to Amis and Balangao, while most groups (Mangyan, Tagalog, Ilocano, Sambalic, Bicolano, Visayan, Bilic, Sangil) form an outgroup to an Ami-Cordilleran clade. In addition, the Cordilleran-like ancestry in these populations is younger than the divergence between Manobo and central Cordillerans. This is suggestive that the ancestral source population of most Cordilleran-related populations in the Philippines came about post Manobo-Cordilleran divergence ~15 kya, and is coming from outside of Taiwan, probably from an ancestral population in South China area after ~10 kya. The latter is supported by the fact that the divergence between Amis/Atayal and various Cordilleran-related groups started at least from ~10 kya (Figs. S7E). Moreover, the divergence time estimate between central Cordillerans and various Cordilleran-related groups (Bashiic, Tagalog, Ilocano, Cagayan, Sambalic and Mangyan) fall between ~7 to 10 kya, setting the boundary for the arrival of these groups in the Philippines at ~7 – 8 kya or earlier (Figs. S8K-P).

The arrival of Cordilleran-related populations at various locations and at various points in time resulted in admixture with earlier Negrito migrants or with admixed indigenous populations with Negrito ancestry. When we estimate the admixture date between Negrito and non-Negrito ancestry using Malder, we observe a random distribution of oldest dates across the archipelago, such as 4,877 +/- 1,250 years BP for Agta Labin in Luzon, 5,246 +/- 1,704 years BP for Kinaray-a in Visayas, and 6,504 +/- 1,204 years BP for Bukidnon Manobo in Mindanao, 5,381 +/- 1,037 years BP for Sama Dialut Taluksangay in Zamboanga peninsula, and 4,689 +/- 1,098 years BP for Palawano in Palawan Island (Fig. S8U and Table S7Q). The absence of north-to-south gradient of admixture dates as can be expected in a stepwise unidirectional movement of Cordilleran-related populations from the Batanic Islands of the north to the Mindanao Island of the south discounts the possibility of a simplistic monolithic expansion of populations espoused by the Out-of-Taiwan hypothesis. Instead, our current multiple lines of evidence support a complex non-uniform migration of Cordilleran-related populations from South China and/or Taiwan area into several locations across the various islands of the Philippine archipelago.

6.5 Diversification of Ethnic Groups in the Visayas

The migration and diversification of Visayan-speaking populations is in itself distinctive due to the geographic nature of middle Philippines: multiple islands that are largely intervisible with each other. Without accounting for admixture, the F_{ST}-based NJ tree characterizes a Visayan clade that can be dissected between east vs west divided by a line drawn along the political boundary that separates Negros Oriental and Negros Occidental (Fig. S2A). The eastern clade includes ethnic groups covering the islands of Cebu, Bohol, Camiguin, Leyte, Samar, and eastern side of Negros (Cebuano, Boholano, Cinamiguin, and Waray), while the western clade includes ethnic groups residing in the islands of Agutaya, Cagayancillo, Cuyo, Panay and western side of Negros Negros (Agutaynen, Kagayanen, Cuyonon, Kinaray-a, Sulodnon, Ati Panay, Ati Negros, Hiligaynon and Bukidnon Negros). All are classified as Visayan-speaking populations except for Agutaynen, who speak a language belonging to the Kalamian group, and Cinamiguin & Kagayanen, who speak a language that are classified under the Manobo language family.

The probable reason why non-Visayan-speaking ethnic groups cluster together with Visayan-speaking populations may be due to the historical inter-island exchange or unidirectional movement of populations that are geographically proximate to each other, such as the migration of Boholanos and Cebuanos into nearby Caminguin Island, interaction between ethnic groups of Cuyo and Agutaya Islands, and interaction between Kagayanen of Cagyancillo and indigenous groups of Negros and Panay Islands. Ati Panay and Ati Negros self-identify as Negritos, and speak languages (Inati/Inata) that are classified under the Visayan branch based on Ethnologue (86). However, Reid (64) and Glottolog (87) classify Inati as an isolate among Philippine languages, and hence place them as a first order subgroup of the MP language family who later experienced heavy borrowing from neighboring Visayan speakers. Given this, Reid posits that Ati likely acquired their language from an early MP-speaking population, which was later confined to relative isolation following the arrival of subsequent populations that spoke the ancestral form of Visayan language.

We constructed a simple phylogenetic framework that fits the topology of Visayan ethnic groups using qpGraph (Table S7R). In this model (*worst-fitting f test of Z=2.2*), we utilized Balangao and Papuan as representative for populations with least admixed Cordilleran or Australasian-related ancestry, respectively. We find that all populations possess a high proportion of Cordilleran-related ancestry (more than 82%), and, at varying degrees, some minimal proportion of Papuan-related ancestry and Basal East Asian ancestry that is older than the divergence of Liangdao and Cordilleran Balangao. Kinaray-a and Hiligaynon cluster together with a Coridilleran-like ancestral root population labelled as *Visayas East*, while Cebuano, Boholano, and Waray cluster together with a shared Cordilleran-like ancestral populations labelled as *Visayas West*.

6.6 Expansion of Cordilleran-related populations into Mindanao

Following Southern Negrito and Manobo-like population migrations, we showed in Section 5.3 that the genetic ancestry of coastal ethnic groups of Mindanao was largely affected by expansion of Cordilleran-related populations. Interestingly, these coastal ethnic groups speak languages that form distinct clusters under the Philippine language macrofamily, including Bilic for Tboli, Obo, Bagobo Klata, Teduray, Blaan Sarangani, and Blaan Koronadal; Danao for Meranao, Iranun, and Maguindanao; and Mansakan for Mansaka, Mandaya, Kalagan, and Tagakaulo. This again likely reflects the layer of complexity in the pulsatile migration of Cordilleran-related populations into Mindanao, which reached different destinations at various points in time, forming linguistically distinct ancestral populations that subsequently settled at various coastal sites: western coast of Maguindanao & Lanao del Sur (or alternatively northern coast of Lanao del Norte) for Danao, eastern coast of Davao Oriental for Mansaka, and southern coast of Sarangani & South Cotabato for Bilic groups. The foundation of these distinct Philippine language clusters may have been brought about by the relative isolation of Cordilleran-related groups in polarized coastal areas of Mindanao, or amalgamation of a dominant ancestral MP language with a language spoken by an earlier indigenous population, or a combination of both.

We modelled the expansion of Cordilleran-related populations into Mindanao in a phylogenetic framework using qpGraph, with Manobo Ata as a surrogate for Manobo ancestry, Papuan for Australasian ancestry, and Balangao for Cordilleran-related ancestry (Table S7S). As expected all groups demonstrate an admixed background ancestry that is Cordilleran-related and Manobo-like, indicating the earlier sequential migration of Ancestral Southern Negrito and ancestral Manobo populations. In addition, all populations received Cordilleran-like ancestry at varying proportions; with the highest amounts detected in Mansaka within Mansakan-speaking groups, Meranao within Danao-speaking groups, and Tboli within Bilic-speaking groups (Figs. S2B-D and Tables S2-4). This is consistent with the findings based on the tests *f3(Mbuti,Balangao,X)* or *D(Mbuti,X,Papuan,Balangao)*, where the order of populations in terms of shared alleles with Balangao Cordilleran is Mansaka > Danao > Bilic (Tables S7T,U).

The ancestral Bilic-speaking group subsequently spread out into isolated communities that later formed the Bagobo Klata of Mt Apo surrounding terrain of Davao area, Teduray of Upi, Maguindanao, Tboli & Obo of Tboli and Lake Sebu, South Cotabato, and Blaan of South Cotabato and Sarangani provinces (Table S7V). The ancestral Danao-speaking group diverged into i) closely related Maranao & Iranun ethnic groups of Lanao provinces and ii) Maguindanao ethnic group of mainly Maguindanao province (Table S7W). Lastly, the ancestral Mansakan-speaking group also diverged into Kamayo of Bislig, Surigao del Sur; Mandaya & Mansaka of Davao Oriental, Davao del Norte & Compostela provinces; and farther into the west, Kalagan and Tagakaulo of Davao del Sur (Table S7X).

6.7 Detection of West Eurasian gene flow dated to the Spanish colonial period

The Philippines was a colony of Spain from 1565 until 1898, and subsequently under American rule from 1901 until 1946. In contrast to the Philippines, the genetic legacy of the Colonial Period in the Americas is readily apparent today through the detectable high West Eurasian ancestry among the majority of lowland and/or urbanized populations of Latin America (88, 89). In the Philippine context, however, admixture between Spanish and local indigenous populations is largely limited (Figs. S2B-D & Tables S2-4), and can only be detected at a population level, using the test *D(Mbuti,CEU,Balangao,X)*, among Bicolano and Creole-speaking Chavacano ethnic groups (Table S7Y). The signal is likely driven by 4 out of 10 individuals tested among Chavacanos, and 1-2 individuals with high levels of West Eurasian ancestry out of 10 tested among Bicolanos. If the threshold of significance is set at $Z > 3$, the presence of West Eurasian ancestry was also detected in random individuals ($n = 4$) among Yogad, Ibaloi, Kapampangan, and Pangasinan populations. If the threshold of significance is stretched to $Z > 2$, the signal can also be detected in some indviduals from Bolinao, Cebuano, Ibaloi, Itabayaten, Ilocano, Ivatan, Kapampangan, Pangasinan, and Yogad populations. All of these aforementioned ethnic groups predominantly reside in lowland and/or urbanized areas (Table S7Y). No West Eurasian ancestry was detected among Negrito, Manobo, and almost all highland ethnic groups. The West Eurasian genetic signal detected among Sama-related ethnic groups can be attributed to South Asian (Ancestral North Indian with high West Eurasian ancestry) gene flow into these populations. Using Malder (Table S7Z), we were able to detect a single admixture event characterized as West Eurasian plus Cordilleran-related, 239 +/- 54 years BP in Bicolano, 156 +/- 36 years BP in Chavacano, 424 +/- 90 years BP in Cuyonon, 161 +/- 42 years BP in Itawis, 429 +/- 109 years BP in Tagalog, and 178 +/- 28 in Yogad, all of which fall within the period of Spanish Colonization (Table S7Z). The mean estimates for Kapampangan (548 +/- 153 years BP) and Hiligaynon (507 +/- 150 years BP) were older than the initial arrival of Spanish colonialists into the Philippines, but have wide confidence intervals where the lower limits still fall within the Spanish Colonial Period (392 and 357 years BP, respectively).

7 Genetic affiliation of Liangdao individuals

7.1 Archeological context

Liang island is situated on the higher part of the continental shelf of East Asia, and only became an island following the end of the Younger Dryas cold period ~12 kya. The Liangdao individuals were excavated from the lower layer (Liangdao-1), dated to 8,380-8204 BP, and upper layer (Liangdao-2), dated to 7512-7374 BP, of a shell midden at Daowei-I, Liang Island, in the Matzu archipelago, situated off the coast of East Asia (90). During the time when these two individuals were alive, the rise in post-glacial sea levels were approaching the maximum, and Taiwan had been separated from the mainland following the ~42m rise in sea level accompanying the Bølling-Allerød warming period that occurred ~15 kya (Fig. S2G). The separation of Liang Island from coastal East Asia, however, does not appear to happen until the early Holocene.

Archaeological remains from both the upper and lower layers are indicative of people who were not practicing settled agriculture and had a reliance on maritime resources. This is consistent with a general pattern of resource procurement in coastal East Asia at this time, which is designated as ‘Complex Hunter-Gatherer’ (CHG) (73, 91). It is defined as a flexible mode of existence involving the use of wild mammals, fish, shellfish and arboriculture, and the production of pottery. The layer above Liangdao-1 contains shell midden and pottery. However, it is not clear whether occupation of Liang Island at the time these two ancient individuals were alive was seasonal or permanent.

According to archaeological evidence, the CHG subsistence strategy was able to support relatively high population densities in coastal East Asia during the Early to Mid-Holocene, but was replaced by sedentary agriculture, including the cultivation of rice and millet, 5,000-3,000 BP (73). Hung (2019) argues that this change in subsistence is accompanied by a ‘second layer’ of material culture, which indicates the arrival of speakers of Austronesian languages. This transition is also marked by changes in both burial practice and biological traits in cranial and dental morphology (92). This new archaeological horizon is attributed to the replacement of people of an ‘Australo-Melanesian’ physical affinity, by Austronesian speaking rice and millet farmers, whose physical attributes are predicted to originate from Siberia (73, 92, 93).

The situation at Liangdao, however, runs contrary to expectations of this model. The analysis of the cranium of Liangdao-1 places him within the cluster attributed to Mesolithic hunter-gatherers (‘Australo-Melanesian’), whilst that of Liangdao-2 indicates affinity to the nEA ancestry component (92). Yet, the age of the individual predates the arrival of the Neolithic layers by more than two millennia and her burial context is clearly CHG. This chronological conundrum was previously treated as an outlier, indicating mobility of people prior to the large-scale arrival of people from the north in the hypothesized population replacement associated with settled agriculture and the putative expansion of the Austronesian language family (73, 92). The recovery of genomic DNA from both Liangdao-1 and Liangdao-2, however, provides a means to test this hypothesis directly.

7.2 Authentication of aDNA & Contamination estimation

In order to assess if the genetic data obtained from the Liangdao individuals is endogenous, levels of contamination were estimated for both mitochondrial and autosomal DNA using Contamix (94) and VerifyBam, respectively (45). Table S8A shows that the contamination estimate prediction for tai001 (Liangdao-1) is low, but the confidence interval is quite wide, probably due to an extremely low level of coverage for the mitochondrial genome. Several extracts were screened from tai002 (Liangdao-2) with quite disparate results for preservation of nuclear DNA. The contamination estimate for the best extract from tai002 (e1) presents slightly higher contamination estimates for mtDNA (given in Table S8A). However, the estimate for autosomal contamination at 1.46% is well below the 5% limit above which extracts are usually treated as unreliable for demographic inference. The demographic analyses presented below were performed with either all data for tai001, or from tai002 (e1) only.

Previous studies have reported that highly fragmented ancient DNA molecules have cytosine deaminations present at high frequencies towards the ends of the sequence. Such patterns of deanimation (DNA damage) are observed from blunt end libraries as an excess of C> T and G>A transitions at the 5' and 3' ends of sequencing reads, respectively. Figure S9A show the deamination patterns for tai001 (Liangdao-1) and tai002 (Liangdao-2) produced with the mapDamage2 software (95). Despite the low coverage of tai001 an endogenous DNA deamination pattern is observed, the same as that observed for tai002. These patterns are consistent with a substantial portion of the genetic material obtained from Liangdao-1 and Liangdao-2 originating from the archeological remains used for DNA extraction.

7.3 Liangdao Man - mitochondrial haplogroup E

The sequencing results from Liangdao-1 Man show that his mitochondrial DNA lineage belonged to haplogroup (hg) E, and has an ancestral position within the phylogeny. This confirms, and provides independent replication for, the results of a previous study of the same individual (96). Our Bayesian analysis, however, was not able to provide convincing statistical support for the precise location of his mtDNA genome within hg E (Fig. S9O and Tables S8B,C). The date of the MRCA of Liangdao Man should, therefore, lie between the root of hg E (8.2-10.1 kya) and the oldest nodes of hg E2 (7.3-9.4 kya) and hg E1 (7.5-9.6 kya) (Table S8C). This distribution (95% HPDs) of dates (7.3-10.1 kya) is concordant with the radiocarbon age of the sample centred on 8.2-8.4 kya.

We found that the age of the root of hg E is very similar to the estimate of (96) but younger than other estimates made with global data sets using various methods (97-99). We also found that the age of the root was very sensitive to the choice of demographic model but that the skyride model was very strongly favored over a model with constant population size (Fig. S9K). Moreover, the demographic history that is reconstructed in the skyride plot (Fig. S9M) is consistent with a strong population expansion within hg E, which stems from the base of the phylogeny. This supports the archaeological evidence that the CHG subsistence strategies supported high population densities in this region prior to the advent of settled agriculture (73).

It is also noteworthy that the E1a2a clade is geographically restricted to Near Oceania with a MRCA 4.7-7.5 kya. This raises the possbility that its translocation predated the expansion of people speaking Malayo-Polynesian languages, if this was concomitant with the Lapita archaeological horizon dating to less than 3.5 kya (100, 101). The E1a1a1 clade is currently restricted to Madagascar with a MRCA (2.7-6.7) that predates most estimates for the time of first settlement(4, 102, 103). Clades of hg E that are specific to Taiwan have MRCAs that span the time of settled agriculture and the end of the CHG period, i.e. 1.5-7.5 kya (Table S8C).

7.4 Liangdao Woman – mitochondrial haplogroup R9

The second Liangdao individual is, by molecular sexing, a woman, who has a mitochondrial lineage belonging to haplogroup R9; a hg that lies within macro haplogroup N (prevalent across Eurasia). Our data, therefore, confirms that which was reported in(90) by way of independent replication. Haplogroup R9 divides into a major clade F, which occurs at substantial frequencies across Island Southeast Asia, resulting from a recent star-like expansion, and two minor ones R9b and R9c (104).

As with hg E, our Bayesian analysis of mitochondrial DNA hg R9 favors a relaxed molecular clock and a skyride demographic model (Fig. S9L). The age of hg R9 is older (16.4-28.3 kya) than that of hg E (Tables S8D,E), but a marked increase in the slope of the skyride plot is evident around 8 kya, i.e. at the same time as in hg E (Fig. S9N). However, the slope is slightly less steep and the reduction ~5 kya is much less marked than that of hg E. It should be noted that the estimated *Ne* for R9 is an order of magnitude greater than that for hg E, which would lead to less sensitivity to local variation, due to the assumption of panmixia in demographic analyses of this type.

There is strong support for the internal clades of R9b1 (14.4-25.9 kya), R9b2 (6.3-14.5 kya) and R9c1 (12.4-22.7 kya) (Table S8E) and the lineage of Liangdao-2 is placed ancestral to R9c1 with 89% support and with an age to the Most Recent Common Ancestor (MRCA) of 14.4-25.7 kya. Within R9b1, there is strong support for R9b1b (5.7-14.8 kya) and R9b1a3 (7.6-16.5 kya), which are both restricted in distribution to MSEA, and R9b1a (12.4-22.2 kya), whose distribution includes ISEA and mainland East Asia. Within R9c1 there is strong support for the clades of R9c1a, R9c1b1 and R9c1b2, which all have MRCAs within the time period of 5.6-13.9 kya, similar to R9b1b, R9b2, and R9b1a3. These and other less well supported clades display internal structures consistent with the demographic expansion indicated by the skyride plot (Fig. S9N), which precedes the transition to rice and millet agriculture occurring 5-3 kya (73).

7.5 Inferences from the ages and distribution of mitochondrial DNA haplogroups E and R9

Previous research into mtDNA haplogroup E found a distinctive geographic distribution centred on ISEA (e.g., (77, 105)) with a general absence in MSEA and mainland East Asia. The presence of hg E on Liang island around 8,000 years ago was interpreted by an earlier study as part of a southward movement from mainland East Asia to Taiwan, with hg E subsequently being lost to genetic drift outside of ISEA (96); this putative migration being associated with the concomitant expansion of the Austronesian language family. An alternative hypothesis is that hg E originated in situ within ISEA and was carried to Taiwan from the south so that its general absence today in East Asia is not due to genetic drift but its original sphere of influence being ISEA (105).

Our Bayesian phylogenetic analysis, using an updated data set comprising complete hg E mtDNA genomes (Tables S8B,C), reaffirms these geographical distributions. It also confirms that the two main regions outside of ISEA where these mtDNA lineages occur, Madagascar and Near Oceania, are both associated with the expansion of the Austronesian language family (Fig. S9O and Table S8C). In contrast, hg R9b is present in both MSEA and ISEA but the two principal sub-groups (R9b and R9c) exhibit very different geographical skewing relative to each other (Tables S8D,E). Table S8E and Fig. S9P also show that R9b is much more common in MSEA than it is in ISEA and that it is particularly well represented in Thailand, where it is found among both Tai-Kadai and Austroasiatic speaking groups. Hg R9b is also found in mainland East Asia, but is virtually absent in ISEA except for a minor distribution in western Indonesia and a single individual each in Taiwan and the southern Philippines. As Taiwan mtDNA has been systematically researched (96), this lack of individuals is unlikely to be a consequence of sampling bias.

Haplogroup R9c has a wider geographic distribution that includes Taiwan (including four mountain tribes), the Philippines (including Ivatan), eastern Indonesia, western Indonesia, Thailand, Vietnam, Laos, and China (Table S8E & Fig. S9P). Of particular interest is the clade R9c1a2, which so far is Taiwan specific with an MRCA of 2.7-9.0 kya, and is part of R9c1a, which is dominated by ISEA populations and has an MRCA of 7.0-13.9 kya. This distribution focus appears to be real because MSEA has been sampled quite densely for complete mtDNA genomes (106, 107). In fact, there is only one clade, R9c1b1, that has a prevalence in MSEA. All three well-supported clades of R9c have MRCAs between 5.6 and 13.9 kya, matching the demographic expansion during the Early to Mid Holocene, indicated by the Bayesian skyride plot (Table S8E).

These current phylogeographic representations of hg R9 and hg E, together with the Bayesian reconstruction of their demographic history, suggest that both haplogroups expanded significantly during the early Holocene. This is consistent with the CHG scenario in coastal East Asia during the early post-glacial. The ultimate origin of hg E likely requires more ancient DNA to resolve with confidence, but the haplogroup is likely to have been present on Taiwan and in Near Oceania by the Mid-Holocene, consistent with the idea of movements of people along long-established Ocean corridors well before the currently accepted dates for the expansion of the Austronesian language family. Moreover, a tentative case can be argued that R9c may have entered Taiwan from the coastal East Asia region, where it was part of a much larger network of people connected to MSEA, where R9b is prevalent.

7.6 Autosomal DNA analysis

For autosomal analysis, Liangdao-1 and Liangdao-2 genomes were merged with Phil_1KGP_SGDP_1.69M, Phil_AsiaPacific_315K, and Phil_HO_201K datasets, and with published ancient DNA data from Asia, to produce the Phil_1KGP_SGDP_Ancient_1.6M, Phil_AsiaPacific_Ancient_315K, and Phil_HO_Ancient_201K datasets. Following haploidization and filtering to keep trasversion sites only, three additional datasets were generated:

Phil_1KGP_SGDP_Ancient_Transv_317K, Phil_AsiaPacific_Ancient_Transv_58K, and Phil_HO_Ancient_Transv_32K with 317,220 SNPs, 57,828 SNPs, and 32,018 SNPs, respectively.

An intial assessment of the regional genetic affinities of the autosomal DNA recovered from Liangdao-1 and Liangdao-2 was performed by PCA, using comparative data from East Asia, Taiwan, Philippines, South East Asia and Polynesia (Extended Data Figs. S9B,C). The results for Liangdao2 generated in this study and in a recently published work (108) are very similar, placing them somewhere between the genetic variation of present-day Philippines, MSEA, and mainland East Asia.

We extended our investigation using outgroup f3 tests: f3(Mbuti,Liangdao1,X) and f3(Mbuti,Liangdao2,X); where we investigate which among the populations X share the most drift with Liangdao-1 or Liangdao-2 (Figure S9F-I and Tables S8F-I). With Phil_1KGP_SGDP_Ancient_Transv_317K panel, we find that both Liangdao-1 and Liangdao-2 share the most drift with ancient Lapita individuals, Cordillerans, and indigenous Taiwanese groups, Amis and Atayal. The results were consistent when we use the Phil_HO_Ancient_Transv_32K panel: all present day Cordillerans and indigenous Taiwanese, as well as ancient individuals from Lapita, northern Philippines, Taiwan, Vietnam, and coastal southern East Asia, share the most drift with Liangdao individuals. This indicates that Liangdao individuals serve as best surrogate for the genetic signal that anchors Cordilleran-related ancestry into mainland East Asia.

To determine whether any control population X in comparison to Liangdao individuals share alleles with Australasian-related ethnic groups, as represented by Papuans, we used the test D(Mbuti;Papuan,Balangao,X). We find that, similar to the least admixed Cordillerans, both Liangdao-1 and Liangdao-2 do not possess any signal of admixture with populations containing Australasian-related ancestry, in contrast to a control population, Sangil, who were earlier established to have Papuan-related ancestry (Table S8S).

Hence, we do not find evidence for strong genetic affiliation of Liangdao-1 to ‘Australo-Melanesian’ or Negrito-related populations. This is in contrast to an earlier study based on cranio-morphometric analysis, which characterized Liangdao-1 as more Australo-Melanesian-like relative to Liangdao-2. Using various combination of D tests, with Phil_1KGP_SGDP_Ancient_Transv_317K and Phil_HO_Ancient_Transv_32K datasets (Tables S8J & S8K, respectively), we find that Liangdao-1 shares more alleles with East Asians relative to any other Australasian-related populations.

Interestingly, admixture analysis using either Phil_1KGP_SGDP_Ancient_Transv_317K or Phil_HO_Ancient_Transv_32K panels demonstrate that both Liangdao-1 and Liangdao-2 individuals not only possess Cordilleran-related (i.e. Balangao and Kankanaey), but also consistently possess nEA-related ancestry (blue component) that is found highest among North Asians (Chukchi, Koryak, and Yakutian) and, in large proportion, among East Asians such as Han, Dai, and Japanese (Figs. S9D,E and Tables S9 & S10). This is consistent with our earlier PCA findings that places Liangdao-1 and Liangdao-2 within the genetic variation of Cordilleran-related, MSEA-based, and mainland East Asian (possessing nEA-related ancestry) populations. Moreover, we noted that the admixture profile of Liangdao-1 and Liangdao-2 were best represented by (among the modern day populations) Amis and Atayal, similarly presenting with combinatory Cordilleran-related, MSEA-based and nEA-related ancestries.

We then performed a formal test to confirm the presence of nEA-related ancestry in Liangdao-1 and Liangdao-2. We utilized the test D(Mbuti;X,AtaManobo, Tujia/Japanese) to investigate whether any populations X share alleles with populations known to have high nEA-related ancestry, like Tujia or Japanese, relative to a population that does not have nEA-related ancestry, Ata Manobo (Tables S8L,M). As expected, ancient Shamanka and Devil's Gate individuals, as well as Northern and East Asian populations presented to have detectable signal of nEA ancestry. In addition, both Amis and Atayal also presented to have nEA-related ancestry, in contrast to least admixed Cordillerans. Liangdao individuals, especially more consistently with Liangdao-2, displayed positive signal for the presence of nEA-related ancestry. More importantly, we can model Amis, Atayal, and ancient Taiwanese individuals as an admixture between nEA and Cordilleran-related ancestries using qpAdm (Fig. S9J and Table S8W). On the other hand, modelling the least admixed central Cordillerans (Kankanaey, Balangao, and Bontoc) as an admixture between nEA and Ami, Atayal, Taiwanese Hanben, Taiwanese Gongguan, or Liangdao-2 were all rejected (Tables S8V-X).

7.7 Asia-Pacific demographic history in light of Liangdao DNA evidence

The cranial and dental affinities of Liangdao Man and Woman, as well as their burial positions (flexed and extended, respectively), classify them unambiguously with the two biological and cultural 'layers' proposed to represent coastal East Asian CHG ancestry and intrusive Siberian related ancestry (73, 92). From a chronological perspective, Liangdao-2 is much too early to fit with this hypothesis, and, moreover, she predates settled agriculture as a mode of existence too (by a considerable margin).

The results from the autosomal analysis cloud the waters still further, because a nEA-related ancestry component (blue component in Figs. S9D,E) is present in Liangdao-1 and Liangdao-2 (Figs. S9D,E,J). This suggests that the observed changes in cranial and dental morphology between the two Liangdao individuals are not possible to predict from the underlying genetic diversity attributable to nEA-related ancestry using ADMIXTURE and formal D tests. Based on the results presented here, there appears to be considerable genetic continuity between Liangdao-1 and Liangdao-2, despite the physical and cultural differences.

Note that the Liangdao genomes also include the Green AN-related 'Cordilleran' component, which is found in a relatively unadmixed state among Cordillerans of northern Luzon. This is confirmed by outgroup f3 statistics, where both Liangdao-1 and Liangdao-2 individuals share the most drift with Cordillerans, Amis, and Atayal, relative to any other modern day populations. These findings suggest that the Liangdao individuals might serve as an important link of Cordilleran-related ancestry from mainland East Asia into ISEA and Taiwan.

Additionally, the findings of nEA-related ancestry in Liangdao-1 and Liangdao-2, but not in present-day Cordillerans, pushes the boundary for the arrival of early Cordillerans into the Philippines to around 7-8 kya or earlier, which is prior to the dates of established agriculture. This demonstrates that the first Cordilleran-related populations in the Philippines were likely hunter-gatherers, similar to Liangdao-1 and Liangdao-2 and the CHG societies in coastal mainland Asia (73). The question whether the ancestral Cordilleran populations, as well Liagndao-1 and Liangdao-2, spoke an ancestral form of AN language still remains unclear and is difficult to ascertain. One possibility is that the ancestral Cordillerans spoke an early form of MP-language of AN, and that the divergence between proto-MP and other primary branches of AN (Formosan languages) were essentially earlier than previous estimates; i.e. that the initial diversification and spread of An languages is not linked to farming. The previous estimates for constructing phylogenetic trees for AN languages were calibrated to the dates of farming archeological sites (109, 110), and not to other older sites resembling CHG societies. Alternatively, another possible scenario is that the ancestral Cordillerans spoke a pre-AN form of language, and then shifted to AN, along with other groups in the Philippines when agriculture was introduced into these communities.

8 Classification of Philippine Languages

The Austronesian (AN) languages are one of the major language families that is spoken by ~ 380 million people or almost 5% of world's population. AN includes approximately 1250 distinct languages spoken in a wide geographical area of Asia-Pacific region, extending as far as Madagascar in the west and to Easter Island in the east. The word "Austronesian", originally coined by Wilhelm Schmidt, is derived from the German word "austronesisch" which stems from Latin "auster" or "south wind" and "νῆσος" or "island". There are 10 primary branches of the AN language family, all of which except one are spoken in mainland Taiwan. The extra-Formosan primary branch of AN is the Malayo-Polynesian (MP) branch, which includes all AN languages outside of Taiwan, thus including the Philippine languages. All Austronesian Philippine languages are classified under the AN language family.

8.1 Origin and Directionality of Spread of Austronesian Languages.

The early works of Blust demonstrated higher lexical diversity of Aboriginal Taiwanese languages relative to all other Austronesian languages beyond Taiwan combined (111). This positions Taiwan as the putative ancestral source of Austronesian languages. The north-south directional movement of Austronesian language from Taiwan into the Philippines is supported by current available linguistic evidence including phonological, pronominal, and morphosyntactic innovations. As Ross (2005) says "If a set of innovations is shared by the languages of a group, it is inferred that they are shared because they have been inherited from a single interstage language. This is far more probable than the alternative assumption—that the innovations have occurred independently in each language which reflects them (112)."

There are two phonological mergers that support the north-south directionality of Austronesian spread from Taiwan into the Philippines (113, 114). Proto-AN *t and *C are merged as Proto-MP *t, as in the following examples: Proto-AN *Cau > Proto-MP *tau 'person', Proto-AN *kuCu > Proto-MP *kutu 'head louse', Proto-AN *batu > Proto-MP *batu 'stone', Proto-AN *telu > Proto-MP *telu 'three'. Proto-AN *L and *n are merged as Proto-MP *n, as in the following examples: Proto-AN *bulaL > Proto-MP *bulan 'moon', Proto-AN *tiaL > Proto-MP *tian 'belly', Proto-AN *zalan > Proto-MP *zalan 'road', Proto-AN *nipen > Proto-MP *nipen 'tooth'. In addition to phonological mergers, the phonological shift from Proto-AN *S to Proto-MP *h is also clear evidence for directionality, given that a shift from a sibilant to /h/, and not the opposite, is a standard natural sound change widely represented in other languages throughout the world. This is exemplified by the following: Proto-AN *Sajek > Proto-MP *hajek 'smell'; Proto-AN *Suab > Proto-MP *huab 'yawn'; Proto-AN *taSiq > Proto-MP *tahiq 'sew'; Proto-AN *SulaR > PMP *hulaR 'snake'.

The pronominal evidence of Proto-AN *=mu 'GEN.2PL' shift to Proto-MP *=mu 'GEN.2SG' is also evidence of a north-south directional movement of Austronesian language spread. Amongst Aboriginal Taiwanese, the pronoun is always a genitive second person plural pronoun, while the Austronesian languages outside of Taiwan mainland, the pronoun is regularly a second person singular pronoun. This is a common change known as a "politeness shift" similar to what is observed in French, "tu" replaced "vous", or in English, "you" replaced "thou" (115). While is not uncommon in some languages to have a shift from second person plural to second person singular, the opposite is rare. Another pronominal evidence of north-south directionality is the reconstruction of a dual pronoun in Proto-MP. A lot of the western MP languages, such as Ilocano in the Philippines, differentiate a first person dual pronoun (i.e., ‘we two’) from a first person inclusive plural pronoun (i.e., ‘we all’). In Ilocano, when one says “Mapan ta idiay bantay”, it translates to ‘Let’s go to the mountain (the two of us)’; while when one says “Mapan tayo idiay bantay”, it translates to ‘Let’s all go to the mountain’.

Some morphological evidence also characterize directionality, differentiating MP languages from their likely origins in mainland Taiwan (116). One example of morphological innovation is the verbal prefix *maN- which does not occur systematically in Aboriginal Taiwanese languages, but occurs widely in MP languages. Additionally, a set of morphological and syntactic differences are found between Aboriginal Taiwanese and MP languages, such as the “recent perfective” construction.

8.2 Classification of Malayo-Polynesian Languages.

The MP languages share features that are not found in AN languages spoken in Taiwan (111, 113-115, 117). The only primary branch of AN language found outside of Taiwan mainland is further classified into second-order groupings. Blust earlier proposed two major branches of MP: Western Malayo-Polynesian (WMP) and Central Eastern Malayo-Polynesian (CEMP), and further categorizes WMP into higher-level subgroupings: i) Philippine, which include all Philippine languages except the Sama-Bajaw languages of southwestern Philippines, ii) North Sarawak, which include languages spoken in northern Sarawak of Malaysian Borneo, iii) Barito, which includes languages spoken in southern Borneo and Malagasy of Madagascar, iv) Malayo-Chamic, which include the Malay language and Chamic languages of MSEA, v) Celebic, which includes all languages in central and southeast Sulawesi. The CEMP, on the other hand, includes all AN languages of eastern Indonesia and Oceania, and is further subdivided into two subcategories: Central Malayo-Polynesian (CMP) and Eastern Malayo-Polynesian (EMP). EMP covers approximately 120 languages spoken in Lesser Sunda Islands and southern and central Moluccas of eastern Indonesia, while CMP includes South Halmahera-West New Guinea languages, spoken in northern Moluccas and Cenderawasih Bay of Papua and West Papua, and Oceanic languages, spoken in Polynesia, Melanesia, and Micronesia.

However, the unifying branch of WMP has been questioned due to lack of unity, and thus was not regarded as a valid primary division of MP, but as an umbrella term for all MP languages that do not belong to CEMP. Blench argued that the lack of internal hierarchy within the AN phylum and the unstructured characteristics of WMP must have been brought about by almost simultaneous and rapid dispersal of AN languages widely into many parts of ISEA and eastwards into Micronesia. For Blench, the spread of AN languages in ISEA must have occurred through assimilation and not by colonization or population replacement. The assimilation process was efficient to a point of complete language replacement, leading to limited evidence of non-AN substrates originally spoken by the first settlers of ISEA (118).

A recent review summarized the linguistic, arceheological, and genetic evidence on the diversity and dispersal history of MP languages (85). It was argued that the lexical and grammatical diversity of MP languages cannot be explained solely by a single block dispersal, but rather likely via multiple dispersal events at different directions and at various points in time. Given the complex migrational history of populations in ISEA, currently available linguistic evidence cannot provide substantial data to infer the higher order temporal and spatial relations of various MP languages. In addition, Klamer (85) argued that, given the complex mobility of populations, the evolution of MP languages likely involved long term interactions within and between multilingual communities, where new migrants and indigenous inhabitants thrive together over a long period of time.

8.3 Branches of Philippine Languages.

Blust supports a unified Philippine subgroup of MP languages, proposing a hypothetical Proto-Philippine language source. He then summarizes the research of several linguists into Philippine languages dividing them into 15 subgroups: 1) Bashiic (Batanic islands of northern Philippines and Orchid island of Taiwan), 2) Cordilleran, 3) Central Luzon, 4) Inati, 5) Kalamian, 6) Bilic, 7) South Mangyan, 8) Palawanic, 9) Central Philippine, 10) Manobo, 11) Danaw, 12) Subanen/Subanun, 13) Sangiric, 14) Minahasan (five languages in the northern peninsula of Sulawesi), 15) Gorontalic (nine languages in the northern peninsula of Sulawesi).

However the unity of Philippine languages, together with the proposition of a Proto-Philippine language, is questioned by Ross, Pawley, and Reid (114). Aside from the absence of commonly shared phonological and grammatical innovations among Philippine languages (112), the reconstruction of Proto-Philippine phonology and morphology is almost identical to that reconstructed for Proto-MP (119, 120). The rapid spread of Proto-MP language into Philippines must have been a chain of dialects that have developed into multiple branches of Philippine languages today. Hence for Reid, there is no Proto-Philippine language, and rather that the subcategories of Philippine languages are first-order branches of MP (114), and include the following: i) Bashiic, ii) Northern Luzon, iii) Central Luzon, iv) Inati, v) Kalamian, vi) Greater Central Philippine, vii) Bilic, viii) Sangiric, and ix) Minahasan.

8.4 Languages spoken by Philippine Negritos

There are approximately 36 distinct ethnic groups of Negritos that are found throughout the Philippines. They are regarded as the first inhabitants of the Philippine archipelago. They must have entered the Philippines via Sundaland way before the AN expansion. All Negritos of the Philippines speak a language that is included under the MP branch of AN, and are hence described as Austronesian-speakers. Negrito groups speak an AN language that is usually clustered with the language group spoken by non-Negrito groups surrounding them. The linguistic positions of the Negrito groups in relation to their neighboring non-negritos is substantial, where Negrito groups are usually first-order groups within their language subfamily or in some cases are isolates. The former is exemplified by Inati, the language spoken by the Ati of Panay and Negros, and Manide, the language spoken by Negritos of Camariñes Norte, where neither seems to be related genetically to any other Philippine language, but have some words that are borrowed heavily from the non-Negrito groups surrounding them.

It is held that structure and position of Negrito languages in relation to other Philippine languages is instructive of their probable history (64). Although there is no record of the original language that was earlier spoken by Negritos, there are some items in their vocabulary that may reflect some retention of a non-Austronesian substratum (62). Alternative explanations of these unique lexical elements among Negritos is that these forms may be retentions of a proto-MP etymology that was lost in all groups other than Negritos, or that these are post-AN innovations in one of the Negrito groups that have diffused via contact into other languages.

8.5 Position of Sama Languages

The language spoken by Sama Dilaut sea nomads and related ethnic groups is not classified under Philippine languages, but as a first-order branch of MP language family. The linguistic position of Sama within the MP language family has not been clearly established, but is recently categorized within the Barito Branch of MP. Blust would classify the vocabulary of Sama into three distinct strata: i) Philippine loanwords, ii) Malay loanwords, and iii) native material, and has inferred, based on lexical evidence of the native material, that Sama languages originated from Borneo (117). In addition, Blust has proposed that the homeland of Sama people, like the Malagasy of Madagascar, is most likely in the basin area of Barito river and its tributaries. Sama languages probably have loanwords from early Southeast Barito languages, suggesting that early speakers of Sama have a connection with the populations inland of Borneo, upstream of the Barito river basin. There are also common loanwords in Samal and Southeast Barito languages that point to trade contacts with Srivijayan Malays, who are ancestral to the modern Banjar communities, since 800 AD.

8.6 Alternatives to the Tree Model

There are series of events that occurred in Philippine history and prehistory that should be accounted for in considering language diversification and evolution. Most of these events can potentially contribute to the extensive lexical shifting and borrowing across Philippine languages, such as arrival of early Austronesian-speaking groups, trading through long-distance maritime networks with Hindu-Buddhist polities of Malaysia and Indonesia, extensive exchange between Philippine coastal cultural communities and Chinese traders, Islamization of southwestern Philippines, Spanish colonization for 333 years, and American rule from 1901 until 1946 (121). All of these events have substantially affected and modified Philippine languages which is impossible to be modelled by simple tree or network diagrams.

Simple tree models are limited in capturing the linguistic transformations (122) that likely have occurred in Negritos, from the loss of their original language to the adoption of the language of their non-Negrito neighbors (123). In addition, trees cannot provide information on the merging of language subgroups brought about by dialect and language chaining. Moreover, tree models are limited in providing insights on the current findings of this study, such as the presence of distinct Manobo and Htin/Mlabri-related Sama genetic ancestries in the southern Philippines. All Manobo and Sama ethnic groups speak languages that are classified under the Malayo-Polynesian branch of Austronesian. So far, current available linguistic evidence does not provide any support for prehistoric contact between AN-related languages and any other Austroasiatic languages of MSEA, nor does it show any non-Austronesian lexical items in the linguistic substratum of Manobo and Sama languages. Given our genetic findings and given the limitations of tree models, it is not clear whether the ancestral populations of Manobo and Sama groups underwent complete linguistic replacement just as the change of language by Negritos.

8.7 Correlations between language, geography, and genetics

Labelling the linguistic affiliation of each Philippine ethnic group in the plotted neighbor-joining tree based on pairwise F_{ST} displays some concordance between language and genetics within the Negrito or non-Negrito-wide clusters (Fig. S2A). The clustering between non-Negritos relatively reflects the clustering of linguistic families within the Philippines, with some exceptions, for instance among others: Cagayanen which speaks a Manobo language clusters with Visayan groups, Tausug which speaks a Visayan-related language clusters with Sama groups, and Blaan Sarangani which speaks a Bilic language clusters with Manobo Sarangani (Fig. S2A). This inconsistency reflects a recent admixture between the two groups that have different linguistic affinities. but are geographically adjacent or accessible to each other.

As expected, provided the complex demographic history of the Philippines with at least 5 major waves of human migrations, there is poor correlation between genetic relatedness and linguistic distance as well as between genetic relatedness and geography (Figs. S10B,C). A better correlation is observed when the comparisons were restricted to Negrito groups only, or when the comparisons were made among non-Negritos that reside within the same region of the same island (Figs. S10B,C).

9 Ethnolinguistic Groups of the Philippines

A key thrust of the NCCA is to develop and promote the Filipino national culture and arts, and to preserve Filipino cultural heritage. In addition, part of the mandate of NCCA is to ensure the widest dissemination of artistic and cultural products among the greatest number across the country and overseas for their appreciation and enjoyment. For that reason, and to highlight the rich diversity in culture and traditions of Philippine ethnolinguistic groups, we have listed below brief ethnographic descriptions of what are known as the indigenous cultural communities of the Philippines.

Most Philippine groups were animists traditionally, but many have become Roman Catholic in religion following the influence of Spanish colonisation. Islam has established itself in some provinces and these will be noted in the following descriptions. Where each group is located and the dialect names given will be mentioned, along with unique facts about the group. Each language is given its subgrouping according to Ethnologue (and or Glottolog). All (Austronesian) Philippine languages are Malayo-Polynesian, so these are not mentioned in the subgrouping statements.

Agutaynen. Agutaynen (also known as Agutayno or Agutaynon) is an ethnic group in Palawan that resides mostly in Agutaya Island and its surrounding islands (Diit, Maracañao, Matarawis, Algeciras, Concepcion, and Quiniluban Islands). Some Agutaynen have migrated to mainland Palawan and established cultural communities in the coastal sites of Roxas, San Vicente, and Brooke's Point, Linapacan, and Taytay municipalities. The Agutaynen language is classified as Kalamian. Most Agutaynen are also fluent in Cuyonon, English, and Tagalog.

Approximately 50% of Agutaynen live in Agutaya municipality situated in Cuyo archipelago. The municipality of Agutaya has a population of approximately 12,500, and is composed of 10 barangays, five of which are in Agutaya Island and the remaining five are island barangays. The other half of Agutaynen population has settled in mainland Palawan, with more 100 families per communities in Roxas, San Vincente, and Brooke's Point and around 20-50 families in Linapacan and Taytay. Most Agutaynen, just like other coastal Palawan ethnic groups, subsist on farming and fishing (124).

Apayao. Apayao (also known as Isneg, Isnag, or Maragat) is the main ethnic group of the northernmost province of Cordillera Administrative Region, the province of Apayao(125). Apayao people speak the Isnag (or Isneg) language which is classified as Northern Luzon, Northern Cordilleran, Cagayan Valley, Isnag. Isnag language has the following dialects: Bayag, Calanasan, Dibagat-Kabugao, Daragawan or Karagawan, and Talifugu-Ripang or Tawini. Most Apayaos used to live along the waterways of major rivers. Recently, increasing number of Apayaos are settling in the more urbanized municipal centers of their province.

Agta of Cagayan province (Agta Labin, Agta Dupaningan and Atta Rizal). Three distinct Negrito groups (but linguistically related) reside in the province of Cagayan: Agta Labin, Agta Dupaningan, and Atta Rizal. The Agta Dupaningan language (also known as Eastern Cagayan Agta, and Dupaninan) is classified as Northern Luzon, Northern Cordilleran, Northeastern Luzon. Dialects of Agta Dupaningan listed in Ethnologue include Barongagunay, Bolos Point, Camonayan, Palaui Island, Peñablanca, Roso, Santa Ana-Gonzaga, Santa Margarita, Tanglagan, Valley Cove, and Yaga. Agta Labin (also known in the literature as Central Cagayan Agta). Atta Rizal is also known as Atta Faire, Rizal Agta, or Southern Atta). Both the Central Cagayan and Atta Faire languages are classified as Northern Luzon, Northern Cordilleran, Cagayan Valley, Ibanagic.

Agta Dupaningan indigenous cultural communities are found along the eastern coastal municipalities of Cagayan and northeast Isabela provinces (from Sta. Ana in the north to Divilican & Maconacon in the south), as well as in Palaui Island (126). Agta Dupaningan are skilled hunter-gatherers using various kinds of hunting technologies (127). Both men and women equally engage in hunting expeditions (128). The Dupanigan Agta included in this the study are the Agta indigenous cultural community of Palaui Island, locally known as Palaui Agta. The Palaui Agta people are integrated with the resident Ilocano population of Palaui Island.

Agta Labin indigenous cultural communities reside in the municipalities of Gattaran, Lal-lo, Peñablanca, and Baggao, Cagayan. The Agta Labin participants included in the study are from Lal-lo, Cagayan. Agta

Labin traditionally have a seminomadic lifestyle, and through marriage alliances, enable them to travel extensively to other places where they have kinship relations (129).

Atta Negritos reside in the western section of Cagayan province, with groups extending towards some areas of Apayao province. Atta Negritos are geographically classified into Atta Pamplona (Atta in Sanchez-Mira, Pamplona, and Abulug municipalities), Atta Faire or Atta Rizal (Atta in Rizal municipality), and Atta Pudtol (Atta in Pudtol municipality of Apayao) (63). The Atta indigenous cultural community included in this study are from Rizal, Cagayan.

Agta Dumagat & Agta Remontado. The Negrito groups found in Rizal province and the neighboring northern section of Quezon province are the Agta Dumagat and Agta Remontado. Agta Dumagat indigenous cultural communities (also known as Umiray Agta, Umirey Dumagat, or Dumaget) reside in Umiray, General Nakar & Infanta, Quezon, as well a Polilio Island, Quezon, and even in southern section of Baler, Aurora. Agta Remontado indigenous cultural communities (also known as Hatang-Kayey, Remontado Agta, Remontado, Sinauna, Sinauna Tagalog), on the other hand, are found in Tanay, Rizal and General Nakar, Quezon. Agta Remontado speak the Remontado language which is classified under Austronesian, Malayo-Polynesian, Central Luzon; while Agta Dumagat speak the Umiray Dumagat language which is classified under Austronesian, Malayo-Polynesian, Greater Central Philippine.

The Agta Remontado indigenous cultural community included in this study is from Barangay Barangay Daraitan, Tanay, Rizal. Remontado is a Spanish term which means 'people who have returned back to the mountains'. The group self-identify as the Remontados, and they refer to their language as Hatang-Kayi 'this language'. Most Remontados have intermarried with Tagalogs, Ilocanos, and Agta Dumagats, and their admixed children hardly speak the Remontado language. It was reported that only 325 speakers of Remontado are remaining, placing it as one of the moribund languages of the Philippines (130).

The Agta Dumagat is known locally as Dumaget or Dumagat, or as Umiray Dumagat in most published literature. The Agta Dumagat indigenous cultural community included in the study is from General Nakar, Quezon. The early Agta Dumagats were traditionally semi-nomadic, who subsist in a combination of hunting-gathering, swidden farming, and fishing practices. Most Agta Dumagat have intermarried with other ethnic groups in the region, which is evidenced by substantial diffusion of lexical items from various neighboring languages into their own language(131), as well as our genetic findings that Agta Dumagat possess high levels of non-Negrito ancestry relative to other Negritos. The term Dumagat is popularly defined as "people from the ocean". This is based on a Tagalog folk etymology, *dagat* 'ocean', with the common Philippine infix <*um*>. However, the term Dumagat has been shown to derive from an old locative prefix *du* (found also in Dupaningan) and Magat, the name of one of the major tributaries of the Cagayan River (64).

Agta of Quirino & Aurora provinces (Agta Maddela, Agta Casiguran & Arta). The ethnic groups of Quirino and Aurora provinces that self-identify as Negritos include Agta Casiguran (also known as Nagtipunan Agta), Agta Maddela (also known locally as Agta Ilagen), and Arta. All of whom speak distinct languages that are classified as Northern Luzon, Northern Cordilleran, Northeastern Luzon. The term Dumagat is only used as andonym by the Umiray Dumaget people. It is not used as endonym by any of the other Negrito groups of the Philippines, but is commonly applied as an exonym to all Negrito groups. There is another Negrito group in Aurora, the Alta people that will be included in a future study.

The Agta Casiguran participants included in our study are from from Barangay Dissimungal, Nagtipunan, Quirino. Our respondents maintain that they sometime travel and cross the Sierra Madre mountains to meet with their relatives in Casiguran, Aurora. In the early 1960's the lifestyle of Agta Casiguran was predominantly based on hunting and gathering (132). However nowadays, due to the increasing loss of the natural tropical forest habitat, incursions of non-Negrito farming communities, and interventions provided by the government and non-governmental organizations, Agta Casiguran people are becoming more sedentary and settled in small communities.

Arta speak a language that is considered to be one of the most endangered languages of the Philippines (133). There are currently at least 10 remaining fluent speakers of the Arta language, with approximately 35-45 individuals who are secondary speakers (134). Some unique innovations, that are characteristic of the Arta language, are found within their numerical system, such as *sipang* ‘one’ and *tallip* ‘two’. Arta mostly live together with the Agta Casiguran indigenous cultural communities of Quirino province. All Arta are multilingual, and are fluent in Agta Casiguran and Ilocano.

Agta Maddela still retain their hunting and gathering way of living. It is one of the least documented ethnic groups of the Philippines. They were estimated to be around 300 individuals according to early field notes by Thomas Headland (132). The local NCIP staff, together with our Agta Casiguran and Arta respondents, claim that Agta Maddela speak a language that is different from Arta or Agta Casiguran languages. In addition, our Arta and Agta Casiguran respondents assert that they find the language of Agta Maddela mostly incomprehensible to them.

Agta Iriga, Agta Isarog, & Agta Iraya. The Negrito groups of Camarines Sur province of the Bicol region that are named after the mountains they live nearby are the Agta Iriga, Agta Isarog, and Agta Iraya. The Agta Isarog and Agta Mt Iraya (also known as Inagta of Mt. Iraya, Iraya, Iraya Agta, Itbeg Rugnot, or Lake Buhi East) languages are both are classified as Greater Central Philippine, Central Philippine, Bikol, Coastal. The Agta Mt. Iriga language (also known as Iriga, Iriga Agta, Lake Buhi West, Mt. Iriga Negrito, San Ramon Inagta) which is classified as Greater Central Philippine, Central Philippine, Bikol, Inland.

Most Agta of Camarines Sur have intermarried with and assimilated to the culture of the non-Negrito Bicolano lowlanders. They largely subsist on farming as their means of livelihood. Some are highly educated and work as professionals in government or private industries. In addition, most Agta Iriga, Agta Isarog, and Agta Iraya are fluent in the regional language of Bicolano, as well as the national language, which is largely based in Tagalog.

Agta Lopez & Agta Manide. Agta Lopez is the Negrito ethnic group who resides in Barangay Villa Espina, Lopez, Quezon. Our Agta Lopez respondents claim that they are relatives of the Manide, the main Negrito ethnic group of Camarines Norte province of Bicol region. Both Manide and Agta Lopez speak the Manide language which is classified as Greater Central Philippine, Umiray Dumaget, although this classification has been challenged in the literature, which claims that Manide is a Philippine isolate. Agta Lopez is sometimes referred to as Ayta by local government agencies and in publications, and is likely an exonym by Tagalog settlers in the region, who commonly label Negritos as *Ayta*.

Manide are a times referred to as the *Abiyan, Bihug, or Kabihug* (135). There are around more than two dozen indigenous cultural communities of Manide in the province of Camarines Norte, with a population of approximately 3700. The Manide participants in this study are from Jose Panganiban municipality, Camarines Norte. Manide are multilingual communities. They are fluent in Bicolano and Tagalog.

Agta Matnog & Agta Bulusan. Agta Matnog and Agta Bulusan are ethnic groups of Sorsogon province of Bicol region that self-identify as Negritos. Both Agta Matnog and Agta Bulusan indigenous cultural communities speak the Southern Sorsoganon language, which is classified as Greater Central Philippine, Central Philippine, Bisayan, Central, Warayan, Gubat. They are also fluent in Central Bikol and Tagalog languages. Ethnologue lists Ayta Sorsogon as their main language, which is apparently already extinct. All the inhabitants self-identify as an admixed Agta, instead of Ayta. The Ayta is likely an exonym, patterned after the label for Negritos given by Tagalog people.

The labels, Agta Matnog and Agta Bulusan, were coined after the names of the municipalities where they reside. Matnog is the site of the Matnog port, serving as one of the busiest ports of the Philippines, which connects Luzon to Samar Island. Bulusan, on the other hand, is the site of the highest peak of Sorsogon, Mount Bulusan (1,565 m) and the scenic crater lake, Bulusan lake. Most members of the Agta indigenous cultural communities in Sorsogon are farmers. Phenotypically, they resemble the non-Negrito inhabitants of Sorsogon. This is evident in our genetic findings, where they consistently cluster with non-Negrito Bicolanos.

Ata Manobo. Ata is a Manobo ethnic group commonly referred to as Ata of Davao or Ata Manobo. They mostly reside in the highlands of Davao region, including the municipalities of Kapalong, Asuncion and Talaingod, Davao del Norte and Paquibato district, Davao City. Their indigenous cultural communities also stretch into southwest Bukidnon as well as the northwest border of Compostela Valley.

They are different from and shouldn't be confused with Ata of Mabinay, Negros Oriental or Ati of Panay, Guimaras, and Negros islands, which are self-identified Negrito groups of the central Philippines.

The Ata Manobo and Ata of Davao language is classified as Greater Central Philippine, Manobo, Central, South, Ata-Tigwa. Aside from primarily speaking the Ata Manobo language in household settings or within the community, Ata Manobo people also use Cebuano for interaction or trade with lowlanders. Ata Manobos practice swidden farming, and supplement their agricultural practices with hunting, fishing, or food gathering.

Ati Panay & Ati Negros. Ati is the name for the Negrito ethnic groups of the Visayas region. They are found in Boracay, Panay, Guimaras, and Negros Islands. The Ati ethnic group represented in this study are the Ati indigenous cultural communities of Nagpana, Barotac Viejo, Iloilo and Marikudo, Isabela, Negros Occidental. The Ati or Inati language is classified as Greater Central Philippine, Central Philippine, Bisayan, Central, Peripheral based on Ethnologue. However, Reid (64) argues that Ati is a primary branch of Malayo-Polynesian which has received many borrowings from neighboring Visayan languages. Documented dialects of Ati include Baratoc Viejo Nagpana Ati in Iloilo and Malay Ati in Aklan. Most of the younger generation are fluent in the language of their province, Hiligaynon or Kinaraya-a, and can speak Tagalog and English as the medium of instruction in schools.

Ayta Magbukon & Ayta Ambala. Ayta Magbukon and Ayta Ambala are the Negrito ethnic groups of Bataan province, Central Luzon. Ayta Ambala indigenous cultural communities are not only found in Bataan (Dinalupihan), but also in Zambales province: in the towns of Castillejos, San Marcelino, Subic, & Olongapo. Ayta Magbukon cultural communities, on the other hand, are found exclusively in Bataan province, in the towns of Abucay, Bagac, Balanga, Limay, Maraviles, Morong, Orani, Orion, and Samal. Ayta Magbukon are also known as Bataan Ayta, Bataan Sambal, or Mariveles Ayta; while Ayta Ambala are also known as Ambala Agta and Ambala Sambal. The Magbukon and Ambala languages (136), are classified as Central Luzon, Sambalic. Most Ayta Magbukon and Ayta Ambala can speak Sambal and Tagalog.

Both Ayta Magbukon and Ayta Ambala were historically semi-nomadic hunter-gatherers. They move from one place to another. Presently, most Ayta Magbukon and Ayta Ambala rely on agriculture for their subsistence, with root crops, fruits and vegetables as their usual crops (137). Some of the Ayta Magbukon and Ayta Ambala still retain the classical physical description of Ayta Negritos: short, dark-skinned, and kinky haired; however nowadays, increasing numbers are losing what was phenotypically described as Ayta, due to intermarriage with neighboring lowlanders.

Preserving the culture and tradition among younger generations has been a challenge among Ayta indigenous cultural communities. One documented approach that was implemented in Ayta Magbukon is the establishment of NCCA-funded School of Living Traditions or SLT. Through SLT, cultural masters and elders were provided an avenue to impart their indigenous knowledge and skills to younger Aytas, may it be traditional way of cooking using indigenous plant resources and recipes, singing songs in their own, learning medicinal plants, or engaging in traditional dance.

Ayta Sambal, Ayta Mag-antsi & Ayta Mag-indi. Outside of Bataan, the Ayta Negrito ethnic groups of Zambales, Pampanga, and Tarlac provinces include Ayta Sambal, Ayta Mag-antsi, and Ayta Mag-indi. Other names for these groups include Ayta Hambali, Botolan Zambal, or Sambal Botolan for Ayta Sambal; Ayta Mag-Anchi or Mag-Anchi Sambal for Ayta Mag-antsi; and Indi Ayta or Mag-Indi Sambal for Ayta Mag-indi. Ayta Sambal indigenous cultural communities are found in Botolan and Cabagan municipalities of Zambales, while both Ayta Mag-antsi and Ayta Mag-indi cultural communities are scattered across two or more provinces. The indigenous cultural communities of Ayta Mag-indi are found in Floridablanca and Porac, Pampanga and San Marcelino, Zambales. Ayta Mag-antsi, on the other hand, reside in Bamban & Capas, Tarlac; Botolan, San Marcelino & Castillejos, Zambales; Mabalacat, Pampanga; and Sapang Bato, Angeles City. These Ayta groups speak distinct but closely related languages that are classified as Central Luzon, Sambalic. Ayta Mag-indi language is most closely related to Ayta Mag-antsi language, while Ayta Sambal language is most closely related to Sambal or Sambali, a language spoken by a non-Negrito Sambal ethnic group.

Historically, Aytas of Central Luzon were skilled hunter-gatherers (138). They usually hunted in bands, and at times used traps to catch wandering animals. Nowadays, the subsistence of Ayta Negritos is mainly supported by swidden agriculture (139), accompanied by the maintenance of small-scale livestock and poultry. Some work as casual laborers in the lowlands, and increasing numbers are becoming educated up to university level and later work as professionals.

The eruption of Mount Pinatubo in 1991, which is the second largest volcanic eruption of the 20th century, has impacted the life and ways of Ayta Sambal, Ayta Mag-antsi, and Ayta Mag-indi ethnic groups (140-142). Mount Pinatubo is situated in the Zambales mountains, at the boundary of Pampanga, Tarlac and Zambales provinces. by The pyroclastic flow and *lahar* deposits caused by Mount Pinatubo eruption devastated the old villages of Ayta cultural communities. Aytas were displaced from their cultural roots and their traditional source of livelihood, and were relocated in government-arranged resettlement areas. Long after the eruption, and when vegetation gradually came back into habitable zones, some Aytas returned close to the slopes of Mount Pinatubo to settle again (143).

Bagobo Klata. Bagobo Klata are also known as the Clata, Atto, Eto, Guanga, Giangan, Gulanga, Guiangan, or Jangan ethnic groups. They are the only Bilic-related ethnic groups of Davao region. Their indigenous cultural communities are found in Catalunan, Tugbok, Calinan, and Baguio districts of Davao City. They are linguistically different from the Manobo-related group, Tagabawa, who are also referred to as Bagobo or Bagobo Tagabawa. Bagobo Klata indigenous cultural communities are geographically separated from Tagabawa by the Lipadas river. Bagobo Klata speak the Giangan language which is classified as Bilic. Their language is more related to Tboli and Blaan of South Cotabato and to Teduray of Sultan Kudarat, than to the Manobo-related languages of their neighboring cultural communities.

Bagobo Klata are typically swidden farmers, with upland rice, root crops, and vegetables as main crops. An indigenous cultural community of a Bagobo Klata is ruled by a tribal council and headed by a chieftain. Disputes are resolved through amicable settlements.

Bagobo Klata traditionally have weavers or *Talanaw'wo*, whom they believe make designs that were guided by the spirits in their dreams (144, 145). The colors for these garments are sourced from their surroundings. Nowadays, there are limited number of garments with authentic designs, given that the tradition of weaving is vanishing. There are attempts by the communities to revitalize the tradition by identifying the skilled weavers, in the aspiration that they will be able to train a new generation of weavers.

Balangao. Most Balangao reside in the eastern section of Mountain province, mainly in the municipality of Natonin, and some in the neighboring Paracelis and Barlig municipalities. In addition to low hill areas, their environment is largely mountainous topography with moderate to steep slopes and woodland forests. The Balangao language (also called Balangaw, Farangao, or Balangao Bontoc) is classified as Northern Luzon, Meso-Cordilleran, South-Central Cordilleran, Central Cordilleran, North Central Cordilleran, Nuclear Cordilleran. Most Balangao are multilingual. They speak English and Tagalog that are mainly learned from school, as well as Ilocano which is the *lingua franca* of the region.

The majority of Balangao are agriculturists. Arable lands are mostly found at the slope of mountains, where some are terraced and irrigated. The majority of the crops produced in Natonin and Paracaelis are corn and rice, with camote, peanuts, vegetables, and banana serving as secondary crops.

Bangon. Bangon is one of the Mangyan ethnic groups of Mindoro (146). Their indigenous cultural communities are found in Bansud, Bongabong, and Gloria municipalities of Oriental Mindoro. Bangon Mangyan are also known as Barangan, Batangan, Binatangan, Fanawbuid, Suri, Tabuid, Eastern Taubuid, or Taubuid. The Bangon Mangyan or Eastern Taubuid language is classified as Greater Central Philippine, South Mangyan, Buhid-Taubuid.

Most Bangon Mangyan still practice animism, although an increasing number have converted to Christianity with the entry of Christian missionaries into the area. Bangon subsist on a combination of hunting and gathering and swidden agriculture. Their main crops include rice, yams, coconuts, sweet potatoes, jackfruit, taro, banana, and vegetables. The excess produce that Bangon communities have are sold in the lowland markets. Most of the farm produce are harvested in the uplands. Their most efficient way of transporting products is through the use of improvised wooden floaters, which carry their farm goods along the river.

Batak. Batak is the Negrito ethnic group of Palawan Island. They are ethnolinguistically different from the Batak group of Indonesia. The Batak or Binatak language is classified as Greater Central Philippine. The Batak language is heavily influenced by Palawano and Tagbanwa languages of neighboring communities, which is reflected as evidence of admixture between Batak Negritos and non-Negritos (30, 147). Just like Ayta Negritos of central Luzon, Batak are phenotypically described as dark-skinned, with curly hair, and short stature. There are approximately less than 500 individuals that self-identify as Batak, and an increasing number of younger generation intermarry with individuals from other ethnic groups, and hence become absorbed into a larger non-Negrito ethnic communities.

Bataks were traditionally nomadic hunter-gatherers. With increasing economic development of the Palawan region, and government intervention, Bataks recently changed their subsistence systems and became settled into small indigenous cultural communities (148). Nevertheless, some still visit the forests and intermittently engage in hunting and gathering, including hunting of wild pigs and collection of honey. Batak people also engage in low-impact shifting cultivation such as *kaingin* farming.

Bicolano. Bicolano refers to the people living in the Bicol region, which is also known as Bicolandia. The region is composed of Bicol Peninsula and the neighboring islands of southeast Luzon including Albay, Camarines Sur, Camarines Norte, Catanduanes, Sorsogon and Masbate. Some Bicolanos also live in the province of Quezon.

The Bicolano language is related to other languages in the central Philippines. The people of Bicol region have four major groups of languages based on the location: Coastal Bikol (with four sub-languages), Inland Bikol (with six sub-languages), Pandan Bikol (lone language) and Bisakol (with three sub-languages). The majority of Bicolanos understand the coastal Bikol specifically the central Bikol language, since this language has been widely used in the literature and mass media and in the major cities of Legaspi and Naga. Bisakol (a combination of Bisaya and Bikolano), a Visayan language that was heavily influenced by the Bicol language, is mainly used in Masbate and Sorsogon.

Nowadays, Bicolanos in the major cities of the region live in a modern urban lifestyle, while the majority of people in the rural areas largely depend on agriculture. Their main products include rice, coconuts, abaka, fruit trees and corn. They also produce tobacco. Mining industry was later introduced in the region with copper, zinc, gold and limestone as the major minerals. Commercial fishing is also being utilized. Some engage in logging and handicrafts from abaca plant.

Blaan Koronadal and Blaan Sarangani. Blaan people and their language is classified as the Bilic group of the Malayo-Polynesian languages. Blaan are mostly multilingual, and are fluent in the languages of their neighboring communities, such as Hiligaynon, Cebuano, Tagalog, or Ilocano. According to oral history, the Blaan and the Tboli people are the only groups that live in the plains and hills of southern Mindanao. The Blaans call the Tboli people as *To Bali* or people who live on the other side of the mountain. Living side by side with each other, Blaans and Tbolis regard each other as equals that live on parallel streams.

Blaan are clustered into three groups depending on their geographical area of residence: such as Blaan from Davao del Sur and Davao Occidental, Blaan from South Cotabato, Sultan Kudarat, North Cotabato and Maguindanao and the Blaan from Sarangani province. Linguistically, Blaan people are grouped into Blaan Sarangani and Blaan Koronadal. Blaans are alternatively grouped into *To Lagad* or those who live up on the high plains and *To Baba* or those who live in the valleys and plains. The Koronadal Blaan people are also referred as *To Kalon* or the *Blaan of the Cogon* grass, likely attributed to the early natural environment of Koronadal which then was densely covered with cogon grass.

Boholano. Boholano or *Bol-anon* is a term for the people residing in the island province of Bohol. A large number of Boholanos have migrated to southern Leyte and the northeastern portion of Mindanao. Boholanos speak the Boholano dialect of Cebuano, which is classified as Greater Central Philippine, Central Philippine, Bisayan, Cebuan. Agriculture is the major source of income for Boholanos, which includes farming of rice, corn, coconut and bananas. Some coastal communities engaged in fishing. Tourism has also become a significant impact on the island's economy. The island of Bohol is a popular tourist destination. The Philippine tarsier which is one of the world's smallest primates is found in Bohol Island.

Bontoc. Most Bontocs are found in the municipality of the same name, Bontoc of Mountain province. Bontok is the name of the language while Bontoc is the name of the municipality. Bontoc people speak a macrolanguage that is classified as Northern Luzon, Meso-Cordilleran, South-Central Cordilleran, Central Cordilleran, North Central Cordilleran, Nuclear Cordilleran, Bontok-Kankanay. The macrolanguage is grouped into Northern Bontok (spoken in Sadanga, Belwang, Betwagan and Anabel), Southern Bontok (spoken in Talubin, Bayyo, and Can-eo), Southwestern Bontok (spoken in Alab, Balili, Gonogon), Eastern Bontok (spoken in Barlig, Kadaklan, Lias), Central Bontok (spoken in Bontoc Ili, Caluttit, Dalican, Guina-ang, Ma-init, Maligcong, Samoki, and Tocucan towns of Bontoc). Each of these groups is a different language, and the towns which compose them speak different dialects. Central Cordilleran languages have recently been described as a network of languages.

The main source of income and livelihood of the Bontoc people is agriculture, with rice as the predominant crop, supplemented by other crops such as camote, millet, corn, sugarcane, legumes, and vegetables. Rice cropping is practiced twice a year, with the first cropping from March until August, and subsequent cropping from September until January. Bontocs are also known for their stoned-walled terraced farming with intricate irrigating canals, for instance, the farming technology that was used in the Maligcong rice terraces. The use of stone walls in terraces create a stable structure that prevents erosion of soil.

Buhid. Buhid is one of the Mangyan ethnic groups of Mindoro (146). The word *buhid* is derived from *bukid* 'mountain'. Buhid people reside in the municipalities of Mansalay, Roxas, Bongabong, and Bansud, Oriental Mindoro, and the municipalities of San Jose and Rizal, Occidental Mindoro. The Buhid language is classified as Greater Central Philippine, South Mangyan, Buhid-Taubuid. Buhid indigenous cultural communities are known for their written script (of Indic origins) that is closely related to the Hanunuo script.

Most Buhid live along mountain slopes and beside rivers. Swidden farming is commonly practiced by the Buhid. Other complementary livelihood includes animal husbandry, hunting, fishing handicraft making and farming labor services. They are also known as pot makers. They usually plant rice, corn, garlic, onions and fruits.

Bugkalot. The Bugkalot ethnic group (also known as Ilongot) are mostly found in Nueva Viscaya and Nueva Ecija provinces along the borders of the Caraballo and southern Sierra Mountain ranges. Bugkalots speak the Ilongot language that is classified as Northern Luzon, Meso-Cordilleran, South-Central Cordilleran, Southern Cordilleran, Ilongot; with the following documented dialects: Abaka, Egongot, Ibalao, Italon, and Iyongut. Bugkalot people also speak Ilocano as the main language of the region, and Tagalog & English as their educational and institutional languages.

Bugkalots are mostly swidden agriculturists, with some engaging in hunting and fishing to supplement their needs (149). Rice and root crops are their staple food. Other crops they cultivate include corn, bananas, manioc, tobacco, coffee, sugar, legumes, and various vegetables. Betel nut chewing is a common practice.

Binukidnon or Bukidnon Negros. The Bukidnon ethnic group of Negros Island are also known as Magahat, Karolano, Mangahat, or Buquitnon *(150)*. They are linguistically different from the Manobo-related ethnic group of Bukidnon province in Mindanao, who are also called as Bukidnon, Binukid or Binukidnon.

Bukidnon Negros speak the Binukidnon language which is classified as Greater Central Philippine, Central Philippine based on Ethnologue, and is classified under Central Philippine, Bisayan, Negrosanon based on Glottolog. Bukidnon Negros are grouped into Northern Binukdinon or Karolano, who are based in Kabankalan municipality of Negros Occidental, and Southern Binukidnon or Magahat, who are based in different barangays of Tanjay, Santa Catalina, Bayawan, and Siaton municipalities. The ethnic group included in this study are the Karolanos of Kabankalan, a municipality in the southern section of Negros Occidental. Most Bukidnon Negros are fluent in Cebuano or Hiligaynon, and in Tagalog and English which serve as the medium of instruction in schools.

The majority of Bukidnon Negros are agriculturists, with rice, corn, banana, legumes and vegetables as regular crops. Livestock and poultry are also maintained in small scale, including chicken, goat, swine, carabao, horse, or duck. Presently, the indigenous cultural community of Kabankalan is ruled by an Indigenous Peoples Council, and headed by the Indigenous Peoples Council chairperson.

Bukidnon Binukid. Bukidnon is one of the indigenous ethnic groups in the province of Bukidnon, a landlocked area in the north-central part of Mindanao. Bukidnon people speak the Binukid language which is classified as Greater Central Philippine, Manobo, North. Their language is grouped together with Kagayanen, Higaonon, and Cinamiguin Manobo languages.

With the influx of migrants and other foreign traders into Bukidnon province in the past centuries, intermarriages occurred between the early Bukidnon communities and settled migrants, having descendants that became the Cebuano-speaking communities of the lowland areas of the province. Bukidnon indigenous cultural communities can be found at the foot of Mt. Sumilao and in Malaybalay town. They are also in the lowland villages of Central Bukidnon.

Bukidnon indigenous cultural communities are headed by a datu. The datu is not elected by all members of the community, instead he is selected by the majority of the clan member officials. Historically, Bukidnon clans live in a communal house with various quarters where a group of at as many as fifty related families could live. Houses are usually made of bamboo and thatch, and many are raised above the ground with floors of bamboo strips.

Bukidnon Higaonon. Higaonon is one the seven ethnic groups of the province of Bukidnon. They were also thought to historically settle in the coastal side of north-central coast of Mindanao. The Higaonon language is classified as Greater Central Philippine, Manobo, North. Most Higaonon are fluent in Cebuano, and also speak Tagalog and some English.

The political system of the Higaonon revolves around a datu. There is usually a principal datu who rules over an entire group composed of several units that are each headed by minor datus. These minor datus form a counselling body for the whole community. The datu assumes multiple roles in the community.

Bukidnon Manobo. Bukidnon Manobo people reside in northcentral Mindanao and form part of one of the seven ethnic groups of Bukidnon province. The traditional religion of Bukidnon Manobo people is based on animism. Manobo practiced slash and burn before they cultivate the area for farming. From clearing to planting to growing and harvesting of commodities. Rice is the main crop. Some have coconuts and other crops like corn and fruit trees. Some Bukidnon Manobos are skilled hunters. They also collect honey from the beehives of the forest. When trees start to bloom, the hunter waits for the coming of the bees that will lead him to their beehives.

Bukidnon Matigsalug and Davao Matigsalug. Matigsalug is one of the indigenous peoples of Mindanao that reside in San Fernando & Kitataotao, Bukidnon; Arakan, North Cotabato; and Marilog, Calinan, Paquibato, Davao City. Most indigenous cultural communities of Matigsalug can be found along the riverbanks of tributaries of the Davao river (Salug river).

The Matigsalug language is classified as Greater Central Philippine, Manobo, Central, South, Ata-Tigwa. The Matigsalug language is most related to the Tigwahanon dialect and to the Ata language. Increasing number of Matigsalug are also fluent in Cebuano. The beliefs, culture and traditions of the Matigsalug are deeply rooted in their surroundings. Trees, mountains, rivers, and valleys surround the Matigsalug indigenous cultural communities, providing them a regular source for daily sustenance.

Matigsalugs practice shifting cultivation with rice, root crops, vegetables and fruits as usual crops. They also supplement their food source by hunting wild animals including wild boar, monkey, deer and lizard, by fishing in the river, and by small scale livestock and poultry.

Bukdinon Talaandig. Talaandig people are known as dwellers of steep places. As slopes provide a dwelling place for Talaandig indigenous cultural communities, they regard themselves as the guardians of the mountains. Talaandig speak the Binukid language which is classified as Greater Central Philippine, Manobo, North. Talaandig people have extended kinship relationships, which is expanded through intermarriage. They place a high priority in preserving the value of family relationships.

Talaandig people have three kinds of houses, a multilayered house where they do their cooking; tree houses as lookouts; and a shack near their cultivated fields where they stay during planting and harvesting season. Most Talaanding are agriculturists. Their crops include rice, root crops, fruit trees, and various vegetables. They also supplement their needs through hunting and trapping animals.

Bukidnon Tigwahanon. Tigwahanon is one of the seven ethnic groups of Bukidnon. Their indigenous cultural communities are found in the municipality of San Fernando close to the border of Davao del Norte. The Tigwahanon variant of Matigsalug language is classified as Greater Central Philippine, Manobo, Central, South, Ata-Tigwa. Most Tigwahanon are agriculturists. Their crops include rice, corn, sweet potato and cassava. Aside from farming, Tigwahanon also engage in hunting and fishing. They also gather honey from the forests.

Bukidnon Umayamnon. Umayamnon people are one of the ethnic groups of Bukidnon. They are known as the forest experts and warriors of the mountains of the headwaters of Pulangi River in Bukidnon Province and Umayam River in Agusan del Sur. The Umayam variant of Manobo Agusan is classified as Greater Central Philippine, Manobo, Central, East.

Umayamnon are mostly dependent on natural resources for their daily food. They till lands to farm. They catch fish from a nearby river. Historically, most Umayamnon seldom left their ancestral land for better opportunities in the lowlands. Accordingly, some Umayamnons retained their traditional belief system. Umayamon people value kinship relationships. Family relations among Umayamnon indigenous cultural communties are traced on both sides. They consider both paternal and maternal sides of the family as important members of the clan. They are expected to help each other to toil the land and farm, when celebrating special occasions such as weddings, burial, farming, and when the need arises.

Cebuano. Cebuano (also known as Sebuano or Sugbuanon or Sugbuhanon) is a name for people whose primary mode of communication is the Cebuano language. From Cebu, a number of Cebuano have spread out to other places such as Siquijor, Bohol, Negros Oriental, southwestern Leyte, western Samar, Masbate, and parts of Mindanao, specifically on the northern and western coasts.

Cebuano language is classified as Greater Central Philippine, Central Philippine, Bisayan, Cebuan. Known dialects of Cebuano include Boholano, Cebu, Leyte, and Mindanao Visayan. The language was greatly influenced by the Spanish as thousands of loanwords were borrowed into the language. In Cebuano, vowels were expanded to five from three, the number of vowels in the parent language. It is closely related to other languages in the Visayas regions such as the languages of Hiligaynon (Ilongo) and Waray-Waray.

Chavacano. Chavacano (also known as Zamboangueño, Chabacano, Chabakano) is an ethnic group who mostly reside in Zamboanga City, and speaks a creolized language of Spanish. The grammar of Chavacano is mainly based on Philippine languages: Tagalog, Hiligaynon, or Cebuano; while the Spanish version is largely of Mexican Spanish. Linguists have identified six varieties of Chavacano: Chavacano Zambonagueño, Caviteño, Ternateño, Ermiteño, Chabacano Davaoeño, and Chabacano de Cotabato. The Spanish creole languages of Davao, Cotabato, and Ermita are considered to be extinct. The Chavacano speakers in this study are from Zamoboanga City.

The language of Zamboanga Chavacano likely commenced in 1635, when the Spanish Jesuits started to establish settlements in Zamoboanga City. The settlement was expanded following the construction of a local military defense fortress, that was ordered by the Spanish Colonial government. The admixture between Iberian/Hispanic-American migrants, Visayan and Tagalog migrants, and local indigenous people of Zamboanga (Subanon and Sama-related groups) likely contributed to the make-up of present-day Chavacanos.

Cinamiguin Manobo. Cinamiguin Manobo is the only Manobo ethnic group of Camiguin Island, an island located in the northcentral section of Mindanao. Cinamiguin Manobo indigenous cultural communities are found in Sagay and Guinsilaban municipalities of Camiguin Island. The Kinamigin language (also known as Kamigin, Cinamiguin, Quinamiguin) is classified as Greater Central Philippine, Manobo, North. Most Cinamiguin Manobo are also fluent in Cebuano, which is the main language spoken by most inhabitants of Camiguin Island.

Most Cinamiguin Manobos have intermarried with Cebunao and Boholano migrants, so the younger generation of Cinamiguin Manobos only use Cebuano as their main language, and can hardly understand any of their original language.

Cuyonon. Cuyonon (also known as Cuyo, Cuyono, Cuyunon, Kuyonon or Kuyunon) is an ethnic group that mainly resides in Cuyo Islands, situated between the islands of Palawan and Panay. Some Cuyonon migrated to the Palawan mainland and established clusters of communities at coastal sites. They speak the Cuyonon language that is classified as Greater Central Philippine, Central Philippine, Bisayan, West, Kuyan. The Cuyonon language is reportedly similar to Ratagnon, a language spoken by a Mangyan ethnic group of Mindoro Occidental.

Cuyonon of Cuyo archipelago had historical trade with Chinese and Islamized ethnic groups of southern Philippines, Malaysia, and Brunei. Since the colonial period, the Spaniards had a strong presence in Cuyo Island, as evidenced by a constructed defensive fort. It is then not surprising to find Spanish influence in the culture of Cuyonons.

Most Cuyonon that settled in Palawan are agriculturists, cultivating rice, corn, and root crops. Some, especially those living along the coastlines and in Cuyo archipelago, still retain the traditional livelihood: fishing. Living close to the sea provides them with a broad seafood resource including various fish species, crabs, shrimps, squids, octopus, sea urchins, and seaweeds.

Gaddang & Ga'dang. Ga'dang indigenous cultural communities are mostly found in the upland municipalities of Paracelis, Mountain Province and Potia, Ifugao; while the closely related Gaddang indigenous cultural communities are mostly found in the lowland municipalities of Bagabag, Bayombong, and Solano, Nueva Vizcaya, as well as various areas of Isabela and Quirino provinces. The languages of Gaddang and Ga'dang are classified as Northern Luzon, Northern Cordilleran, Cagayan Valley, Ibanagic, Gaddangic. Most Gaddang and Ga'dang also speak Ilocano, which is the *lingua franca* of the region.

Both upland Ga'dang and lowland Gaddang populations are predominantly farmers, although they differ in the form of farming (151). Upland Ga'dang rely on swidden agriculture, while the lowland Gaddang utilize a wet farming technology of plowed and irrigated fields. Rice serves as the main crop, and is supplemented by legumes, millet, bananas, vegetables, and even tobacco as cash crop. The upland Ga'dang also supplement their food resource by fishing in the streams and hunting wild animals in the remaining forests. Domesticated pigs and chickens are raised for local consumption or for use as sacrificial animals in rituals.

Hanunuo. Hanunuo is one of the Mangyan ethnic groups of Mindoro. They live in the municipalities of Mansalay, Bulalacao, and some parts of Bongabong along the periphery of southeastern Mindoro. The Hanunuo language is classified as Greater Central, South Mangyan, Hanunoo. Dialects of Hununuo include Binli, Bulalakawnon, Gubatnon, Kagankan, Waigan, and Wawan.

Hanunuo is known for their script, which is one of the Indic scripts of the Philippines. Hanunuo are traditionally animist. The indigenous cultural communities of Hanunuo are usually found in the valleys of southeastern Mindoro, often situated near streams or rivers. The villages are small with around fifty people. The houses are made of wood, bamboo and nipa roofs. The majority of Hanunuo grow their food through swidden farming or slash and burn agriculture, wherein a section of forest is cut down, following by burning of plant debris, and planting of crops after clearing.

Hiligaynon. Hiligaynon is also referred to as Panayan or Ilonggo, although the latter is the more specific term for people living in Iloilo, Guimaras and Panay. The term Hiligaynon defines the language and culture of the Ilonggo people. They mainly reside in Panay island (Aklan, Antique, Capiz and Iloilo). They are also scattered in Western Negros, Southern Mindoro, Tablas, Romblon, Sibuyan, Guimaras and Northwestern Masbate. A minority of the population migrated to Central Mindanao particularly in the SOCCKSARGEN provinces (South Cotabato, Cotabato, Sultan Kudarat, Sarangani and General Santos).

The Hiligaynon language is classified as Greater Central Philippine, Central Philippine, Bisayan, Central. The coastal cities of Iloilo in Panay island and Bacolod in Negros are the economic and administrative centers for the region. The majority of Hiligaynon people live in rural areas, subsisting on farming with rice, corn, sugarcane and coconuts as the major products. Some coastal residents are involved in fishing. A minority of families in Negros Occidental have large sugarcane plantations where locals are employed and work. Tourism has also become a significant source of income since Boracay Island is part of the region.

Iraya. Iraya is one of the Mangyan ethnic groups of Mindoro. The Iraya Mangyans live in the municipalities of Puerto Galera, San Teodoro and Baco in Oriental Mindoro, as well as in Occidental Mindoro, in the municipalities of Abra de Ilog, Paluan, Mamburao and Santa Cruz. The Iraya language is classified as North Mangyan. Most Iraya occupy permanent or semi-permanent villages or clusters of houses around and beyond the fringes of the farmlands. Because of the proximity of their communities to barangay centers, and their intermittent participation in wage labor of big farms, most Iraya also speak Tagalog.

Irayas mainly rely on farming for their subsistence. They plant rice, banana, sweet potato, beans, papaya, corn squash and other root crops. They are also skilled in nito-weaving. Handicrafts such as jars, trays, plates and cups of different sizes and design are being marketed to the lowlanders.

Ibaloi. Ibaloi (also known as Ibadoy, Ibaloy, Inibaloi, Nabaloi, Benguet-Igorot, iBenguet) is the predominant ethnic group of Benguet province of the Cordillera Administrative Region (152), in the municipalities of Bokod, Itogon, Kabayan, La Trinidad, Sablan, Tublay, Tuba, and the southern sections of Kapangan and Atok. There are also Ibaloi people in the western side of Nueva Viscaya province.

The Ibaloi or Inibaloi language is classified as Northern Luzon, Meso-Cordilleran, South-Central Cordilleran, Southern Cordilleran, West Southern Cordilleran, Nuclear Southern Cordilleran, Ibaloy. Documented dialects include Bokod, Daklan, and Kabayan Ibaloi. Traditionally, the Ibaloi people engaged in swidden agriculture, wet-rice agriculture, mining, fishing and hunting. Majors crops being produced recently include rice, root crops, legumes, and vegetables.

The ancient Ibaloi tradition of mummification is a complex process that requires cleaning the body, covering it with salt and herbs, and smoking it in seated position for months until complete dehydration occurs (153). All body parts are preserved in the mummification process, including the internal organs

Ibanag. Ibanag is a major ethnic group of Cagayan and Isabela provinces of northeastern Luzon. The Ibanag language is classified as Northern Luzon, Northern Cordilleran, Cagayan Valley. Ibanag is considered as the de facto language and provincial identity of Cagayan and Isabela provinces. Two major dialects of Ibanag include Northern Ibanag (spoken in the northern municipalities of Cagayan province: Abulug, Aparri, Camalaniugan, Pamplona and Lal-lo) and Southern Ibanag (spoken in Tuguegarao, Cagayan and Isabela provinces). Most Ibanag people are also fluent in Ilocano, which is the main language of the region.

Ibatan. Ibatan is the indigenous ethnic group of Babuyan Claro Island. The entire Babuyan Claro Island and the surrounding 5-km ocean was awarded to the Ibatan indigenous cultural community in 2007 as their ancestral domain. The Ibatan language is classified as Bashiic. Their language is most closely related to Ivatan of the neighboring islands in Batanes. Ibatan language has 72% lexical similarity with Itbayaten Ivatan and 74% with Basco Ivatan. Most Ibatan are fluent in Ilocano, and use Tagalog and English as mediums of instruction in schools and governmental transactions.

Ilocano. Ilocano is also known as Ilokano or Samtoy. The Ilocano language is classified as Northern Luzon., and is related to the languages of the Cordilleran ethnic groups. Ilocanos are the third largest ethnolinguistic group in the Philippines. Four major Ilocano provinces include La Union, Ilocos Sur, Ilocos Norte, situated in the northwest of Luzon, and Abra. The homeland of the Ilokanos stretch from Cape Bojeador on the northwestern tip of Luzon down to the Gulf of Lingayen. Most of the indigenous cultural communities lie in the narrow coastal plain. A larger number of Ilocanos have migrated into the neighboring regions of Luzon (Cordillera, Central Luzon, and Cagayan Valley), the southern Philippine island of Mindanao (Cotabato, Sultan Kudarat, Davao and Zamboanga), and overseas into Hawai'i and California.

Iranun. Along the eastern shore of Ilana bay or mostly coastal municipalities of Maguindanao province (Parang, Matanog, SultanKudarat, Sultan Mastura, and Barira) are the Iranun ethnic group (also known as Iranon, Ilianon, Ilanum, or Ilanos). Some Iranun also reside in the hills lying between the western coastlines of Maguindanao and the southern edge of Lanao plateau. Some have migrated into Sulu or even further into the west coast of Sabah, Malaysia, including some villages in Kota Belud and Lahad Datu districts. The Iranun language is closely associated with the Meranao language. Both Iranun and Meranao languages are classified as Greater Central Philippine, Danao, Maranao-Iranon.

Itawis & Malaweg. Itawis and Malaweg are minority indigenous ethnic groups of Cagayan province in northeastern Luzon. The indigenous cultural communities of Malaweg are found in the municipality of Rizal, Cagayan; while the indigenous cultural communities of Itawis are mainly in the municipalities of Enrile, Iguig, Peñablanca, Tuao, Piat, Tuguegarao, Amulung, Sto. Niño, Solana, Baggao and Alcala, Cagayan. There are some documented Itawis communities in Echague municipality of Isabela. Itawis and Malaweg speak the Itawit and Malaweg languages, respectively, and are classified asNorthern Luzon, Northern Cordilleran, Cagayan Valley, Ibanagic. Glottolog lists Malaweg and Itawit as distinct languages under the Gaddangic subgroup, while Ethnologue lists Malaweg and Itawis as dialects under Itawit language. Most Itawis and Malaweg also speak Ibanag, which is a major language in Cagayan province, and Ilocano, which is the dominant language of the Cagayan Valley region.

Itneg. Itneg (also known as Tingguian, Tingyan, or Tinguian) is the main ethnic group of Abra province, Cordillera Administrative Region. The Itneg or Tinguian language is classified as Northern Luzon, Meso-Cordilleran, South-Central Cordilleran, Central Cordilleran, North Central Cordilleran, Kalinga-Itneg. There are at least six documented Itneg languages that are distinct from each other: Binongan Itneg (spoken in Licuan-Baay municipality); Inlaod Itneg (spoken in Danglas, Lagangilang, Langiden, and Peñarubia municipalities); Maeng Itneg (spoken in Luba, Tubo, and Villaviciosa municipalities); Masadiit Itneg (Boliney, Bucloc, and Sallapadan municipalities); Moyadan Itneg (Manabo municipality); Banao Itneg (spoken in western Abra with Banao Pikekj, Gubang Itneg, Malibcong Banao dialects, more similar to Kalinga languages than to other Itneg languages). Most Itneg also speak Ilocano, as the dominant language in the region.

Ivatan & Itbayaten. At the northermost part of the Philippine archipelago are the Batanic group of islands, where Ivatan and Itbayaten indigenous cultural communities reside. The Ivatan languages are classified as Bashiic. Ivatan language is the de facto provincial linguistic identity of the Batanic group of islands. Dialects of Ivatan language include Basco Ivatan and Southern Ivatan.

The Batanic group of islands is a common passageway for typhoons. In order to protect them from the strong winds of a typhoon, Ivatan built their houses with meter-thick coral or limestone walls and a thick layer of cogon grass roof (154). Ivatan subsist mainly on farming and seasonal fishing.

Kagayanen. Kagayanen, also known as Cagayano, Kagay-anen, Cagyanen, Kinagayanen, is an ethnic group of Palawan that mainly reside in Cagayancillo, a remote island situated between Negros and Palawan Islands. Some Kagayanen indigenous cultural communities are also found along the coastal sites of Palawan including Quezon, Rizal, Coron, and Balabac Islands. The Kagayanen language is classified as Greater Central Philippine, Manobo. It is the only ethnic group in Palawan that speaks a Manobo-related language. In addition, it is the northernmost Manobo-speaking ethnic group, and is one of the ethnic groups that speak a Manobo language outside of mainland Mindanao (Cinamiguin being the other). The majority of Kagayanens are also fluent in Cuyonon, Tagalog, English, Cebuano and/or Hiligaynon languages.

The Kagayayen people in Cagayancillo have a vast maritime resource. The island of Cagayancillo lies close to Tubbataha Reef National Park, a 97-thousand-hectare marine reserve that was declared by UNESCO as a world heritage site. Cagayancillo is 330 km east of Puerto Princesa the capital of Palawan province and 133 km southwest of Antique, and is composed of 31 islands categorized into 12 barangays.

Kankanaey. Kankanaey is also known as Kankanai, Lepanto Igorot, Kankanay, Applai, Katangnan, Sagada Igorot, Kataugnan, or Western Bontoc. The Kankanaey language is classified as Northern Luzon, Meso-Cordilleran, South-Central Cordilleran, Central Cordilleran, North Central Cordilleran, Nuclear Cordilleran, Bontok-Kankanay. The Kankanaey language is categorized into Central Kankanaey and Northern Kankanaey languages. The former is spoken in north Benguet, southwest Mountain province, southeast Ilocos Sur, and northeast La Union provinces while the latter is spoken in the southeast sections of Ilocos Sur and the western section of Mountain province, including Sagada, Besao, Bauko, Tadian, and Sabangan municipalities.

Kankanaeys are predominantly agriculturists with most households having a piece of land to raise crops. Rice serves as the main crop, which is supplemented by corn, root crops, legumes, banana, coffee, fruits, and vegetables. Small-scale raising of pigs is also practiced for domestic consumption or for utilization in the performance of rituals. There is also a vibrant weaving industry that is centered in Sagada, which is becoming more prevalent given the expansion of tourism industry in the past decade. Thus, tourism industry is increasingly contributing to the local economy of Kankanaey indigenous cultural communities.

Kalanguya. The Kalanguya ethnic group is also known as Kalangoya, Kallahan, Ikalahan, Kalanggutan, or Keley-i. They reside across multiple provinces surrounding Mount Pulag: Ifugao (Hungduan, Asipulo, Kiangan, and Tinoc), Nueva Viscaya (Kayapa, Santa Fe, and Ambaguio), Benguet (Bokod, Buguias, and Kabayan), and Pangasinan (San Nicolas). The majority of the members of the First Kalanguya Tribal Congress in 1993 (and affirmed in subsequent congresses) decided to have Kalanguya as the appropriate name for the ethnic group, while a small group from Imugan, Santa Fe, Nueva Viscaya maintained Ikalahan as the name of their cultural community.

The Kalanguya language is classified as Northern Luzon, Meso-Cordilleran, South-Central Cordilleran, Southern Cordilleran, West Southern Cordilleran, Nuclear Southern Cordilleran, Kallahan. Dialects of Kalanguya include Northern Kalanguya (Ambaguio municipality of Neuva Viscaya and Tinoc municipality of Ifugao), Southern Kalanguya (Santa Fe municipality of Nueva Viscaya), Western Kalanguya (Bokod municipality of Benguet), and Central Kalanguya (Kayapa municipality of Nueva Viscaya). Ethnologue lists Keley-i Kalanguya or Kallahan Keley-i as a distinct language, spoken in Kiangan and Tinoc municipalities of Ifugao, with documented dialects of Bayninan and Yatuka (155).

Kamayo. Kamayo is a Mansakan-related ethnic group at the east coast of Mindanao. Kamayo people mostly reside in southern section of Surigao del Sur province, between Marihatag and Lingig municipalities, with a large concentration of communities in the city of Bislig. The Kamayo language (also known as Kinamayo, Kadi, or Kinadi) is classified as Greater Central Philippine, Central Philippine, Mansakan. Ethnolgue identifies two dialect versions of Kamayo: North Kamayo and South Kamayo. The Kamayo language is most closely related to the language of their neighboring ethnic group, the Mandaya. Moreover, most Kamayo are fluent in Cebuano, which serves as their major secondary language.

Kalinga. Kalinga is the main ethnic group of the landlocked province of Kalinga, Cordillera Administrative Region. The Kalinga languages are classified as Northern Luzon, Meso-Cordilleran, South-Central Cordilleran, Central Cordilleran, North Central Cordilleran, Kalinga-Itneg. There are seven documented Kalinga languages: Butbut Kalinga (spoken in Tinglayan municipality and Tabuk City); Limos Kalinga (spoken in Pinukpuk and Rizal muncipalities, and Tabuk City); Lubuagan Kalinga (spoken in Lubuagan municipality and Tabuk City); Mabaka Kalinga (spoken in western Abra; northern Kalinga, and Conner, Apayao); Majukayang Kalinga (spoken in Tabuk City); Southern Kalinga (spoken in Tinglayan municipality); and Tanudan Kalinga (spoken in Tanudan municipality).

Kapampangan. The early settlers resided along the river banks of the Rio Grande de la Pampanga in Luzon that covers a wide land area bordering from the Gulf of Lingayen on the North, Zambales Mountains on the west, Sierra Madre on the east and Manila Bay on the south. Kapampangans are the fifth largest ethnolinguistic group of the Philippines. They are also known as Pampangan, Pampanguenos or Pampangos as they live mainly in the Province of Pampanga. They also occupy parts of Bataan, Tarlac, Bulacan, Nueva Ecija and Zambales.

The Pampangan or Kapampangan language is classified as Central Luzon, Pampangan. In addition, Kapampangans have their own indigenous Indic writing system representing their language. Their main industries include farming and fishing. Rice serve as the main crop, and they reside in a region that is referred to as the rice granary of the Philippines. Other crops include sugarcane, corn, vegetables, fruit trees and root crops.

Kinaray-a. Other alternative native names include the following: Hiniraya, Hinaraya, Binisaya na Karay-a and Bisaya na Kinaray-a. Kinaray-a people mainly reside in Antique province, as well as in various areas of Iloilo province, the southern part of Guimaras Island, Southern Aklan, Occidental Mindoro (particularly in Ilin Island) and in the western parts of Capiz. A small part of the population also migrated to some parts of Mindanao specifically in the SOCCSKSARGEN provinces.

The Kinaray-a language is classified as Greater Central Philippine, Central Philippine, Bisayan, West. Most Kinaray-a living inland depend on agriculture. Their main products include rice, corn, coconut, legumes, fruits and vegetables. Livestock and poultry are also maintained for local consumption. Those living in the coastal areas, on the other hand, depend on fishing as their major source of livelihood. Fishing is all year round, which is also supplemented by farming and production of seaweeds. Forest products such as bamboo, rattan, buri, abaca, vines and wild flowers were used in cottage industry and as raw materials for furniture and handicrafts. Several towns had different specialty in their productions ranging from native hats, toys, gifts, bags, bamboo crafts, mats and loom-woven barrrel skirts. The province has rich mineral resources that are being exported.

Maguindanao. Maguindanao (also spelled or called as Magindanaw, Maguindanaw, Maguindanaon, Magindanaoan) is an Islamized ethnic group in the southwestern region of Mindanao mainland, with the largest concentration of communities in Dinaig, Datu Piang, Maganoy, and Buluan municipalities of Maguindanao province, which stretch into Cotabato, South Cotabato, Sultan Kudarat, and Zamboanga del Sur provinces.

The Maguindanao language is classified as Greater Central Philippine, Danao, Maguindanao. Documented dialects of Maguindanao include Biwangan, Ilud, Laya, Sibugay, Tagakawanan. The secondary languages for Maguindanao are Cebuano, Tagalog, English and Arabic. Arabic is mainly used for the practice of Islam, which serves as their liturgical language.

Mamanwa. The only self-identified Negrito population of the southern Philippines is the Mamanwa ethnic group. Mamanwa indigenous cultural communities are found in the provinces of Surigao del Norte, Surigao del Sur, and Agusan del Norte. The Mamanwa language is classified as Greater Central Philippine, Central Philippine, Mamanwa. Most Mamanwa are fluent in Cebuano or Surigaonon, which are the major languages of their surrounding region. Some Mamanwa are also fluent in Manobo, especially the Agusan Manobo language of the Manobo indigenous cultural communities of Agusan del Sur.

Mandaya. Mandaya indigenous cultural communities are found in various municipalities of Davao del Norte; Manay, Caraga, Baganga, & Cateel municipalities of Davao Oriental; Togo municipality of Surigao del Sur; and southern section of Agusan del Sur. The Mandaya language is classified as Greater Central Philippine, Central Philippine, Mansakan, Eastern. Dialects of Mandaya include Caraga Mandaya, Cateel Mandaya, Manay Mandaya, Mangaragan Mandaya, and Sangab Mandaya. Mandaya language has 89% lexical similarity with Mansaka and 83% lexical similarity with Kalagan. Most Mandaya can speak Cebuano as their secondary language, while some can speak Mansakan.

Farming is the main source of livelihood for most Mandayas. Their crops include rice, root crops, bananas, vegetables, legumes, and abaca. To supplement farming, some also engage is fishing or hunting wild pigs, deer, birds, monkeys, or monitor lizards in their forest. Aside from spears, they also use traps to catch wild animals. The loss of forests due to logging activities and increasing urbanization have reduced the hunting and fishing activities of Mandaya people.

Manguangan and Dibabaon. The Manobo-related indigenous cultural communities of Davao del Norte provinces are the Manguangan (also known as Manguwangan or Manguagan) and Dibabaon (also known as Debabawon, Dibabaon Mandaya, Mandaya). Both Manguangan and Dibabaon speak a related language labelled as Dibabawon Manobo which is classified as Greater Central Philippine, Manobo, Central, East. The Dibabawon Manobo language is spoken as a secondary language for Manobos of Rajah Kabunsuwan, Lingig, Surigao del Sur.

Most Dibabawon and Manguangan are swidden farmers. Rice is the predominant crop, with corn, legumes, root crops, banana, and vegetables serving as secondary crops. Some also maintain abaca as their cash crop. Some Dibabawon and Manguangan still engage in fishing along rivers and streams and hunting or trapping wild animals in the forests. In addition, some families also maintain domesticated chickens and pigs in their backyards for their domestic consumption.

Manobo Agusan. The Manobo indigenous cultural communities of Agusan del Sur, Agusan del Norte, and Surigao del Norte provinces are collectively called Agusan Manobo (156). Agusan Manobo is named after the Agusan river, which is the third largest river basin of the Philippines. The Agusan Manobo language is classified as Greater Central Philippine, Manobo, Central, East. Dialects of Agusan Manobo listed in Ethnologue include Umayamnon, Surigao, and Agdawan.

The presence of Agusan Marsh is significant to the life Agusan Manobos. It is one of the most ecologically diverse and most important wetlands in the Philippines. It is declared a protected area, and designated as the *Agusan Marsh Wildlife Sanctuary*.

Manobo Dulangan & Lambangian Manobo. The Manobo ethnic groups of western Mindanao are the Manobo Dulangan and Lambangian Manobo. Manobo Dulangan indigenous cultural communities are found in Esperanza, Kalamansig, Lebak, Sen Ninoy Aquino, and Palimbang municipalities of Sultan Kudarat and in the western sections of South Cotabato province. Most Lambagian Manobo indigenous cultural communites are found in Upi and South Upi municipalities of Maguindanao. The Manobo Dulangan or Cotabato Manobo language and the Lambangian (also known as Lambanguian or Lambagian) language, are both classified as Greater Central Philippine, Manobo, South. Ethnologue lists Blit Manobo and Tasaday Manobo (groups that will be included in a later paper) as dialects of Dulangan Manobo. Most Dulangan and Lambanguian Manobos are fluent in Cebuano, and some can speak some Tagalog or Hiligaynon. Some are also fluent in the language of their neighbouring ethnic group, the Teduray.

Manobo Ilianen. Manobo Ilianen is the Manobo-related ethnic group of the the northern section of Cotabato province of Mindanao Island. Some Manobo Ilianen indigenous cultural communities are also found in Kandingilan, Kibawe, and Darnulong municipalities of Bukidnon and Northern Kambutalan and Datu Montawal municipalities of Maguindanao. Manobo The Manobo Ilianen language is classified as Greater Central Philippine, Manobo, Central, West. Dialects documented in Ethnologue for Manobo Ilianen include Arakan, Livunganen, and Pulangiyan.

Manobo Rajah Kabunsuwan. Manobo Rajah Kabunsuwan is the Manobo ethnic group that resides in Barangay Rajah Cabungsuan, Lingig, Surigao del Sur, and the adjacent areas extending to the northeast border of Davao Oriental and the southeast border of Agusan del Sur. The Manobo Rajah Kabunsuwan language is classified as Greater Central Philippine, Manobo, Central, East. The Manobo Rajah Kabunsuwan language has 82% lexical similarity with Dibabawon Manobo, and 76% lexical similarity with Sagunto dialect of Agusan Manobo.

Manobo Sarangani. Manobo Sarangani is the only Manobo-related ethnic group that resides in the provinces of Sarangani and Davao Occidental. Their indigenous cultural communities are found in Glan, Sarangani and the neighboring municipality of Jose Abad Santos, Davao Occidental. Opposite the western section of Davao Gulf, communities of Manobo Sarangani are also found in Governor Generoso, Davao Oriental. The Manobo Sarangani language is classified under Austronesian, Malayo-Polynesian, Greater Central Philippine, Manobo. Most Manobo Sarangani are multilingual. In addition to an ability to speak Cebuano, some are also fluent in the indigenous languages of their neighbors, Blaan or Tagakaulo.

Mansaka. Mansakans are one of the main ethnic groups of the Davao region, especially in the provinces of Davao del Norte and Compostela Valley. Mansakan indigenous cultural communities are found in Tagum City and in the municipalities of Mabini, Maco, Maragusan, Mawab, Nabunturan and Pantukan. The Mansaka language is classified as Greater Central Philippine, Central Philippine, Mansakan, Eastern. Just like any other ethnic group in Davao region, most Mansakan people are also fluent in Cebuano. The Mansakan language is most closely related to the language of their neighboring ethnic groups, the Mandaya and Kalagan, 89% and 74% lexical similarity, respectively.

Meranao. Meranao, also spelled as Maranaw, Ranao, or Maranao, is the ethnic identity of Lanao del Norte and Lanao del Sur provinces. Most Meranao are settled around Lake Lanao, the second largest lake of the Philippines, and the largest lake in the island of Mindanao. Meranao is one of the ethnic groups that predominantly profess Islam, and hence are grouped together as one of the 13 Moro ethnic groups of the Philippines.

The Maranao language is classified as Greater Central Philippine, Danao, Maranao-Iranon. Most Meranao are also fluent in Cebuano, due to interaction with Cebuano-speaking populations of northern Mindanao and the Visayas, as well as English and Tagalog which are mainly learned via formal education as mediums of instruction used in schools. The minority of Meranao can speak Arabic, which is the liturgical language used in the practice of their Islamic faith.

Molbog. Molbog (also known as Malebugan or Malebuganon) is the only Islamized ethnic group of Palawan. Molbog communities are found in Balabac and Bataraza municipalities at the southern tip of Palawan, mainly in Balabac and Ramos Islands. Some clusters of Molbog communities also reside in Banggi and Balambangam Islands of Sabah, Borneo, Malaysia. Molbogs of Sabah are alternatively called Balabak, which may pertain to their origins in Balabac Island of Palawan.

The Molbog language is classified as Greater Central Philippine, Palawanic. An alternative view is grouping Molbog with the Bonggi language of Sabah, as the Molbog-Bonggi subgroup. The latter may be due to assimilation of some Bonggi words into Molbog, due to geographic proximity and consequent exchange. Molbog is categorized into three dialect clusters: Balabac Molbog, Banggi Molbog, and Southern Palawan Molbog.

Ovu Manuvu. Ovu Manuvu are also known as Obo Manobo, Ubo, Bagobo, Kidapawan Manobo, or Obo Bagobo. Ovu Manuvu indigenous cultural communities are found in Kidapawan City, Arakan and Magpet municipalities of Cotabato, and Marilog district of Davao City. The Obo Manobo or Ovu Manuvu language is classified as Greater Central Philippine, Manobo, Central, South, Obo. Dialects documented in Ethnologue include Arakan Manobo, Kidapawan Manobo, Magpet Manobo, and Marilog. Most Ovu Manuvu are fluent in Cebuano, which is the regional *lingua franca* of the area, or Tagabawa, which is the language of their neighboring ethnic group. The Uvo Manuvu included in this study are from Marilog district, Davao City.

Palawan. Palawan (also called Palawano, Pala'wan, Palawanun, Palawanen or Palaweño) is one of the ethnic groups in Palawan Island of western Philippines. The Palawano language is classified as Greater Central Philippine, Palawanic. Palawanic languages are categorized into three subgroups: Brooke's Point Palawano (the southeast section of Palawan Island from south of Abu Abu to Bataraza), Central Palawano (the southwest section of Palawan Island from north of Quezon to north of Rizal municipalities), and Southwest Palawano (the southwest section of Palawan Island from north of Rizal to the southern tip and eastern Bataraza). In addition, Brooke's Point Palawano has a documented dialect in Bugsuk Island known as South Palawano or Bugsuk Palawano.

Most Palawan settle in the highlands, with their stilt houses along the hillsides close to a stream or river system. The Palawan group were historically known to be nomadic hunter-gatherers, but have shifted to agriculture with the influx of agrarian settlers. Palawan practice swidden agriculture in small patches of lands in the forest, with variable crops including upland rice, root crops, legumes, and some vegetables. Some Palawan also hunt wild pigs and collect honey. They also collect and sell rattan and resin.

A small community of Palawano seasonally reside in the southern side of the crater of Mount Mantalingaan, an extinct volcano in southwest Palawan. They are referred to as the *Tau't Bato*.

Panay Bukidnon. The Panay Bukidnon or Sulodnon ethnic group are also known as the Sulod, Monteses, Halawodnon, Panaynon, Tumandok or Mondo *(157)*. They are linguistically different from Bukidnon of Negros Island or the Manobo-related Bukidnon groups of Mindanao. All Panay Bukidnon indigenous cultural communities are located in the interior municipalities of Panay Island: Tapaz, Capiz; Lambunao, Iloilo; Valderrama, Antique. The Panay Bukidnon or the Sulod language is classified as Greater Central Philippine, Central Philippine based on Etnologue, and is classified as Central Philippine, Bisayan, West Bisayan, Kinarayan based on Glottolog. Their language is most similar to Kinaray-a language that is spoken mainly in the province of Antique.

Pangasinan. Pangasinan was a term used to refer to the coastal villages that later became the name of the province. Pangasinan is a term referring to the province, people and language spoken in the province of Pangasinan. There has been considerable intermarriage with neighboring Ilocano people from northern Luzon, with whom they share many traditions. The Pangasinan language is classified as Northern Luzon, Meso-Cordilleran, South-Central Cordilleran, Southern Cordilleran, West Southern Cordilleran.

Sama Deya. The Sama Deya ethnic group (also known as Southern Sama or Sama Tawi-Tawi) is the main ethnic group of Tawi-Tawi and neighboring islands (Simunul Island, Sibutu Island, etc). Sama Deya of Tawi-Tawi speak the Southern Sama language that is classified as Greater Barito, Sama-Bajaw, Sulu-Borneo, Inner Sulu. Dialects of Southern Sama listed in Ethnologue and Glottolog include Bajau Banaran, Bajau Darat, Bajau Laut, Bajau Semporma, Balimbing, Bongao, Languyan, Obian, Sama Sibutu, Sapa-Sapa, Sibutu, Simunul, Sitangkai, and Tandubas. The Sama Deya indigenous cultural community included in this study is from Bongao, Tawi-Tawi.

Tawi-Tawi was historically a center of Sama or Bajau culture until the arrival of Muslim missionaries in the 14^{th} century. Most Sama in Tawi-Tawi have converted to Islam, and this became the majority religion of the area. An important archeological site in Tawi-Tawi is the Balobok Rockshelter (158) which was declared as an Important Cultural Property by the National Museum in 2017. Three cultural layers were dated: 8760 +/ 100 B.P (Early Occupational Phase), 7290 +/- 120 B.P. (Middle Occupational Phase), and 5140 +/- 100 B.P. (Later Occupational Phase). It remains to be known whether the human remains discovered in the site have genetic continuity with the present day Sama Deya and Sama Dilaut population of Tawi-Tawi.

Sama Dilaut. The Sama Dilaut ethnic group is also known as the Sinama, Orang Laut, Dilaut, Samal, Central Sinama, Central Sama, Samal, Bajao, Bajau, Badjao, or Bajaw. The Central Sama language is classified as Greater Barito, Sama-Bajaw, Sulu-Borneo, Inner Sulu Sama. Ethnologue lists Sama Deya, Sama Dilaut, Sama Laminusa, Sama Siasi, Sama Tabawan as dialects of Southern Sama. The Sama Dilaut participants of this study include Sama Dialaut indigenous cultural communities of Bongao, Tawi-Tawi and Barangays Mampang & Takusangay of Zamboanga City.

Sama Dilaut are nomadic seafaring people who travel using their traditional handmade wooden boats *(75, 159)*. Some Sama Dilaut dwell in their boats, while most dwell in characteristic stilt houses that are located close to coastal sites. Sama Dilaut rely mainly on fishing for subsistence. They also engage in trade, selling their goods such as seafood products and pearls to land-based ethnic communities.

Sama Kabingaan & Sama Banguigi. The Sama ethnic groups that reside in the coastal barangays of Zamboanga peninsula, as well as in the islands of Sulu archipelago, include the Sama Kabingaan and Sama Banguigi. The Northern Sama or Sama Balangingih language (also known as Baangingi, Balanguingui, Bangingi, Bangingih Sama, Sama Bangingih, or Samal) is classified as Greater Barito, Sama-Bajaw, Sulu-Borneo, Inner Sulu Sama. Ethnologue lists the following dialects of Northern Sama or Sama Balangingih: Balangingi, Daongdung, Kabinga'an, Lutangan, Sibuco-Vitli or Sibuko, and Sibuguey.

Both Sama Kabingaan and Sama Banguigi predominantly rely on a sea-based economy, either via fishing, gathering of shells, or seaweed farming. In addition, Sama Kabingaan and Sama Banguigi are believers of the Islamic faith. From the 18^{th} to the 19^{th} century, Sama Banguigi were known for their sea raiding activities targeting coastal settlement areas of middle to northern Philippines and Borneo (160).

Sambal & Bolinao. Sambal and Bolinao are the Sambalic-related non-Negrito ethnic groups of Central Luzon. Bolinao indigenous cultural communities are found in Bolinao and Anda municipalities of Pangasinan, while Sambal are found all throughout the province of Zambales. The Bolinao language (also known Binubolinao, Binubulinao, Bolinao Sambal, Bolinao Zambal, Bino-Bolinao, Bulinaw, Sambal Bolinao) and the Sambal language (also known as Zambal or Sambali) languages are both classified as Central Luzon, Sambalic. Sambal is considered as the *de facto* provincial language and identity of Zambales. Documented dialects of Sambal in Ethnologue include Iba, Masinloc and Santa Cruz.

The town of Bolinao is known for the discovery of the Bolinao skulls (Balangasay archeological site). The characteristic feature of Bolinao skulls are the golden decorative elements of the teeth which resemble fish scales. The skulls were dated to around 14th to 15th century AD, and were discovered together with trade ware ceramics that date back to the Ming Dynasty of China.

Surigaonon. Surigaonon is a term for people living in Surigao. Some Surigaonon also migrated southwards into Agusan del Norte, Agusan del Sur and Davao Oriental provinces. The Surigaonon language is classified as Greater Central Philippine, Central Philippine, Bisayan, South. It has been profoundly influenced by the Cebuano language due to migration of many Cebuano people to the region.

Tagabawa Manobo. Tagabawa Manobo are also known as Bagobo, Tagabawa Bagobo, or Tagavawa. Tagabawa indigenous cultural communities are found at the foothills of Mount Apo, in Toril District of Davao City and Bansalan, Kapatagan, and Santa Cruz municipalities of Davao del Sur. The Tagabawa language is classified as Greater Central Philippine, Manobo. Most Tagabawa are bilingual, with Cebuano as their second language.

Tagakaulo and Kalagan. The Tagakaulo (also known as Tagakaulu, Tagakaulo, Kaolo, Kalagan) and Kalagan ethnic groups (also known as Kaagan, Kagan Kalagan, or Kinalagan) are the Mansakan-speaking populations mostly residing in the western sections of Davao region. Most Tagkaulo indigenous cultural communities are found in Davao del Sur and Sarangani provinces, while Kalagan indigenous cultural communities are found in Davao del Sur. The Kalagan particpants included in this study are from the Minuslim Kalagan or the Islamized Kalagan that reside across Davao Gulf, in Padada, Davao del Sur; Tagum, Madaum and Matiao in Davao del Norte; Pantukan, Compostela Valley; and Lupon, Davao Oriental. Both Tagkaulo and Kalagan peoples speak closely related languages that are classified as Greater Central Philippine, Central Philippine, Mansakan, Western.

Tagalog. The Tagalog people are the most dominant ethnolinguistic group in the Philippines. They are one of the largest groups with a highly urbanized society, given that they inhabit the capital region of the country, Metro Manila. Most Tagalogs live in Nueva Ecija, Bulacan, Tarlac, Zambales, Marinduque, Bataan, and Aurora provinces, and in the Calabarzon region, which includes 5 provinces: Batangas, Quezon, Cavite, Rizal and Laguna. A significant Tagalog population is also found in Mindoro and Palawan. The Tagalog language is classified as Greater Central Philippine, Central Philippine. The national language of the Philippines, Filipino, is largely based on Tagalog.

Early Tagalog settlements were commonly seen on the banks near the delta or mouth of the river. The earliest written record pertaining to Tagalog communities dates back to the 9th century. It is inscribed in a copperplate known as the Laguna Copperplate Inscription. This plate documents the existence of several Philippine polities, specifically the Pasig River Delta polity of Tondo that is believed to indicate trade, cultural and political ties. The inscription was written in Kawi script, and used a mixture of languages including Sanskrit, Old Javanese and Old Malay.

Tagbanwa. The Tagbanwa ethnic group is also known as Aborlan Tagbanwa, Apurawnon, or Tagbanua. Most Tagbanwa indigenous cultural communities are found in the central and northern sections of Palawan Island covering Quezon and Aborlan municipalities as well as Puerto Princesa City. The Tagbanwa language is classified as Greater Central Philippine, Palawanic. The Tagbanwa language is different from Calamian Tagbanwa which is spoken by indigenous cultural communities of northern Palawan and the Calamian group of islands.

Most Tagbanwa are agriculturists. They cultivate rice, root crops or corn in patches of land via swidden method. In addition, Tagbanwa also engage in collecting forest products like honey, resin, and rattan. Fishing is an important source of livelihood for Tagbanwa settled along the coasts. A supplemental source of income for Tagbanwa is selling their handicrafts including mats, baskets, and carved woodworks.

Tausug. The indigenous ethnic group of Sulu Island province are the Tausug people. Indigenous cultural communities of Tausug can also be found in Basilan and Tawi-Tawi Islands, as well as coastal sites of Zamboanga peninsula. The Tausug language is the *de facto* provincial linguistic identity of Sulu. It is classified as Greater Central Philippine, Central Philippine, Bisayan, South, Butuan-Tausug. The

Tausug language is the only Visayan-related language in the Sulu archipelago, which is mostly populated by ethnic groups speaking a Sama-related language such as Yakan, Central Sama, Southern Sama, and Northern Sama. Linguistically, Tausug is most closely-related to Butuanon and Suriganon languages of northern Mindanao. The majority of Tausug people are followers of Islam.

Teduray. Teduray people are the Bilic-related ethnic group of Sultan Kudarat, mainly in the municipalities of Upi and South Upi. The Teduray language is classified as Bilic. Most Teduray people are fluent in Cebuano, and some are fluent in the language of their neighboring ethnic groups: Lambanguian Manobo, Dulangan Manobo, or Maguindanao.

Tedurays have a social structure with strong kinship relationships, as reflected in their socioeconomic activities of working together in farming, hunting, fishing and basket weaving. The main livelihood is based on agriculture. Fishing and hunting are usually done during dry season when mountain, creeks and streams are shallow and when fishes are easy to cats. Hunting expedition also done during dry season. Bows and arrows, spears with barbs are used.

Tboli and Obo. Tboli (also known as T'boli or Tagabili) is one of the indigenous ethnic groups of South Cotabato province. Tboli indigenous cultural communities are found in Lake Sebu and Tboli municipalities. Obo (also known as Ubo), on the other hand, is another ethnic group that resides in the municipality of Lake Sebu. Obo are different from the Manobo-related ethnic group of Davao region, Uvo Manuvo. The Tboli language is classified as Bilic. Obo is a dialect of Tboli, and is a Bilic-related language.

Tuwali Ifugao & Ayangan Ifugao. Tuwali and Ayangan are the main ethnic groups of Ifugao province, Cordillera Administrative Region. Tuwali Ifugaos are mostly found in Hungduan and Kiangan municipalities, while Ayangan Ifugaos are mostly settled in the eastern sections of Ifugao province. The Ifugao languages are classified as Northern Luzon, Meso-Cordilleran, South-Central Cordilleran, Central Cordilleran, North Central Cordilleran, Nuclear Cordilleran, Ifugaw. Locally, the inhabitants of Ifugao province classify themselves as speaking Ayangan, Tuwali and Kalanguya languages. Ethnologue classifies Ifugao into four distinct languages: Mayoyao Ifugao, Amganad Ifugao, Batad Ifugao, and Tuwali Ifugao. Amganad is spoken in the Banaue barangay of the same name, and has dialects of Banaue and Burnay; Batad is spoken in Batad barangay of Banaue municiplaity and has a Ducligan dialect; Mayoyao is spoken in Mayoyao municipality; and Tuwali is spoken in Kiangan and Hungduan municipalities with Hapao, Hungduan, and Lagawe dialects. Ifugaos are also known for their rice terraces which were declared as UNESCO World Heritage Sites (Nagacadan, Hungduan, Mayoyao, Bangaan, and Batad), and for their UNESCO-declared Intangible Cultural Heritage.

Western Subanon and Southern Subanen. Subanen (also known as Subanon or Subanun) are the indigenous ethnic groups of Zamboanga peninsula. Some indigenous cultural communities of Subanen are also found in Misamis Occidental and Misamis Oriental provinces. The languages of Subanen are classified as Greater Central Philippine, Subanon. Ethnologue classifies Subanon/Subanen-related languages into six distinct clusters: Central Subanen or Sindangan Subanun; Eastern Subanen or Guinselugnen; Northern Subanen or Tuboy Subanon; Southern Subanon or Lapuyan Subanen; Kolibugan Subanon or Kalibugan; and Western Subanon or Siocon Subanon. In this study, the participants are from the Western Subanon indigenous cultural community of Zamboanga City and the Southern Subanen indigenous cultural community of Lakewood municipality, Zamboanga del Sur. Islamized Subanen are usually locally referred to as Kalibugan or Kolibugan.

Waray. Waray is a native word which means none or nothing, and is used to call the Visayan people who speak the Waray language, so this is an exonym. They are also referred to as the Lineyte-Samarnon. They mainly reside in the eastern part of the Visayas region which includes the islands of Samar (where they are called Samarenos/Samarnons), Northern Leyte (where they are called Leyteños), Biliran (where they are called Biliranon), and Masbate particularly Ticao Island (they are called Ticaonon) and Sorsogon. The Waray language (also called Waray-waray or Binisaya) is classified as Greater Central Philippine, Central Philippine, Bisayan, Central, Warayan, Samar-Waray.

Yakan. The indigenous ethnic group of Basilan Island province of the Sulu archipelago are the Yakan people. Some Yakan indigenous cultural communities can also be found in the surrounding islands of Basilan province and eastern coastal sites of Zamboanga peninsula. The Yakan language is the de facto provincial linguistic identity of the provincial island of Basilan. The Yakan language is classified as Greater Barito, Sama-Bajaw.

Yakans practice the Islamic faith. In addition, most Yakans subsist on agriculture, with rice, coconut and cassava as predominant crops. Some Yakan communities, especially those residing along the coasts, engage in fishing.

Yogad. Yogad indigenous cultural communities are found mainly in the municipality of Echague, and the nearby towns of Angadanan, Jones, and Santiago, Isabela. The Yogad language is classified as Northern Luzon, Northern Cordilleran, Cagayan Valley, Ibanagic. Most Yogad can speak Ilocano, Tagalog, and English, while some are fluent in language of a neighboring ethnic group, Ibanag.

10 Acknowledgements

For assistance in the consent process, sampling, validation and dissemination of results:

National Commission for Culture and the Arts: Felipe M. de Leon (NCCA chairperson 2010-2016), Orlando Magno, Rico Pableo Jr., Al Ryan Alejandre, Manny Pallatao, Shirley Maloles, Edwin Antonio, Elena Toquero, Rongie Moli, Vicente Handa, & Micahel Angelo Yambok

Cultural Center of the Philippines: Eva Marie Salvador, Chris Mallado, Ronnie Mirabuena & Nikko Zapata

Other Cultural or Non-Governmental organizations: Staff and volunteers of Dumendingan Arts Guild Collective, Kaliwat Performing Artists Collective Inc, Museo Dabawenyo, Museo de Oro, Nisa Ul Haqq, Negros Cultural Foundation, Inc., Musikahan sa Tagum Foundation, Inc; Panaghiusa Alang Sa Kaugalingnan Ug Kalingkawasan Inc., Surigao Heritage Center, Lam le few school of living traditions, Kafye Blaan Empowerment, Inc., South Cotabato Culture and Arts Foundation, Inc., and Tarbilang Foundation. Honey Villegas, Criz Tacumba Diuyan, Winnie Rose Galay, Ronel Nonan, Fernando Almeda Jr., Alma Uy, Mary Ann Orcullo.

Educational institutions: Samuel Benigno and Davie Ted Dolojan of Quirino State University; Maria Luz Badiola, Julie Lucille del Valle-Lopez, Alfredo Fabay, Fr. Roberto Exequiel Rivera S.J. of Ateneo de Naga University, Visitacion Simbulan and Aleth Mamauag of Isabela State University, Richard Daenos of City College of Angeles, Romeo Quilang of Cagayan State University, Eva Marie Codamon-Dugyon of Ifugao State University, Habib Macaayong of Mindanao State University, Dipunudun Marohom and Corazon Mangelen of Institute for Peace and Development of Mindanao of Mindanao State University – Marawi Campus, Mary Joyce Guinto-Sali and Jurma Tikmasan of Mindanao State University – Tawi-Tawi College of Technology & Oceanography), Loreta Sol Dinlayan of Bukidnon State University, Elnora Dudang of Pangasinan State University

National Commission for Indigenous Peoples: Leonor Oralde-Quintayo, Masli A. Quilaman, Atty. Ronaldo M. Daquioag, Evelyn L. Jacob, Woy Lim P. Wong, Juliet Akia, Jr., Vynn Laurilla, Atty Geroncio Aguio, Vivian Tanamor, Agnes Salvino, Atty. Rhodex Valenciano, Benito Batan, Imelda Pantaleon, Vernon Joy Repatacodo, Vicente Yaguel Jr., Sylvia Serrano, Catalina Rivera, Juan Gonzaga, Ernesto Banzales, Ike Encarnacion, Roberto Trestiza, Christopher Barcia, Wenceslao Dionesio, Richard Canuto, Roldan Parangue, Carina Fabregas, Jennifer Gerones, Atty. Josefina Rodrigues-Agusti, Ruben Bastero, Mary Jane Kinoc, Bernard Daytec, Dana Bunnol Ernesto Mendoza, Josephine Pattaguan, Victor Calingayan, Roger Dela Rosa, Felipe Lumiwes, Simplicia Hagada, Josie Pataueg, Nelson Castillo, Shirley Iguianon

Other Governmental organizations: Marie Rafael-Banaag, Assistant Secretary of Presidential Communications Operation Office and Teodoro 'Teddy' Baguilat Jr., Representative for Ifugao province and Chairperson of the Philippine Congress's Committee on National Cultural Communities (2010-2013) and Vice Chairperson of the Committee on Indigenous Peoples (2013-2016)

Local Government Units (LGUs): Staff and officials of the following city or municipal LGU's: Basco, Batanes; Bangued, Abra; Tabuk, Kalinga; Conner, Apayao; Lagawe, Ifugao; Paracaelis, Sagada, and Bontoc, Mountain Province; La Trinidad, Benguet; Bolinao, Pangasinan; Barotac Viejo, Iloilo; Kabankalan, Negros Occidental; Jose Abad Santos, Davao Occidental; Linging, Surigao del Sur; Koronadal and Tboli, South Cotabato, Lebak and Kalamanisg, Sultan Kudarat; Malungon and Glan, Sarangani; Kabacan, Cotabato, Davao City; Upi and South Upi, Maguindanao; Malaybalay, Bukidnon; Tagum City, Davao del Norte; Zamboanga City. Jason John Joyce, Nielo Tupas, Emelita Balingon, Dibu Tuan, Ann Viola, Dionesio Besana, Ronan Eugene Garcia, Camilo Lammawin, Ferdinand Tubban, Allen Jesse Mangaoang, Alexandre Claver, Osman Salisipan, Captain Victor Siao Sr., Efraim Englis, Neil John Yangco, Joel A. Quinanahan, Jhon Kevin O. Belec, Silverio Quibradero, Jose Restificar, Genaro Yap Aizon, Bentor Ganado, Clarita Prudencio, Djarma Rafael

Indigenous Cultural Communities: Indigenous Cultural Community Councils and/or Council of Elders of all indigenous cultural communities included in the study. Datu Cain Hukman, Datu Sanorio Abentong, Datu Simeon Piang, Datu Adolino Saway, Datu Jimboy Catawanan, Datu Ampuan Jeodoro Sulda, Bai Impandi Preciosa Sulda, Bai Vivian Sulda, Datu Roberto, Datu Rodrigo Kegod, Datu Angelito Omos, Datu Bernard Lumikid, Bai Jessica P. Ado, Bai Fedela G. Balquin, Datu Rodolfo Villegas, IPMR Eduarado Bualan, Timuay Nilda Mangilay, IPMR Sario Copas, Datu Roel Arthur Ali, Datu Carlito Guinto Sr, Bae Estrelita Uy, Bai Gregoria Uy, Datu Anthony Longakit Duyan, Bai Herminia Buisan-Ortiz, Chieftain Joseph Eling , Bai Inatlawan Adelina Tarinom Bai Tinangkil Herminia Saway, IPMR Pagbelngen Nestor Saavedra, Pagbelngen Danilo Bonales, Panglima Robert Cursod, Chieftain Verginia Hamora, Panglima Sanol Casim, Panglima Sapeon Diohane, Panglima Sarilan Puntas, Panglima Kevin Ayo, Chieftain Arlito Bacosa, Chieftain Agustin Bacosa, Apo Rogelio Bunag, Apo Roger Fresnillo, Maelam Floresto Zambrano, Maelam Rene Alvarez, Marilyn Banes, Teresita Sinceda Ricky Noblesala, Eufemia Enciso, Ariston Pranada, Corazon Sta. Ana, Lourdes Torcuator, Cleofa Dagaraga, Dodong Acuten, Althea Uy, Armando Anac, Estela Mediano, Helen Lumbos, Mispa Perong, Rodel Hilado, Maria Todi, Rosie Sula, Allen Sula, Kristel Peligrino, Chieftain Cruzaldo Rosales, Thomas Pumbaya

Other Individuals: Maria Noelyn Dano, Michelle Leonardo, Paul Adolfo, Jennelyn Aguinaldo, Alice Inovejas, Stephen Eric Doloriel, Leilani Duguran Ugali, Andre Narros Lluch, Martin Persson, Sebastian Bengtsson, Christine Arquiza Tulio, Edrick Curato, Lodar Esbobillo, Mae Paner, Moses Gabayno Villacrusis, Stephen Eric Doloriel, Esmeraldo Rotoni, Michael Rotoni, Gerard Lico, Farina Marohomsalic, Amirel Usman, Ruby Ferda Alcala, Jose Marcel Laud, Lorraine Hermosura-Faeldon, Jacqueline Frances Momville, Delia Lawian, Jim Sagario, Shirley Brillos, Flor Gaviola, Arthur Larena, Marissa Marasigan-Torentera, Retchie Sotomayor, Jennifer Versoza, Noel Yamon, J-Ann Buling, Fritzie Dael

For provision of legal services:

Atty. George Ahmed Paglinawan, Legal Consultant to the National Commission for the Culture and the Arts (2014-2016), Atty Eirene Jhone Aguila of Aguila Rances Law Offices, and Atty. Lemuel Lopez of Melbourne Law School, University of Melbourne.

For provision of archeological samples:

Suei-Sheng Yang, former Lienchiang County Governor and Lienchiang County Cultural Affairs Department. Marie Lin and Tseyi Wang of Mackay Memorial Hospital, Taipei City 10449, Taiwan

For laboratory technical assistance:

Luciana Simoes, Irene Urena and Berenice Villegas Ramos for processing and extraction of aDNA from Liangdao-1 and Liangdao-2 archeological samples.

For provision of published datasets:

Maude Phipps and Farhang Aghakhanian of Jeffrey Cheah School of Medicine and Health Sciences, Monash University, Selangor, Malaysia for sharing the Malaysian dataset. Partha Majumder and Analabha Basu of National Institute of Biomedical Genomics for sharing Indian dataset. David Reich, Mark Lipson, and Nathan Nakatsuka of Harvard Medical School and Pontus Skoglund of Francis Crick Institute for sharing the Human Origins dataset. Murray Cox of Massey University, Palmerston North, New Zealand and Francois Ricaut of French National Center for Scientific Research, Toulouse, France for sharing the Indonesian datasets. Toomas Kivisild of KU Leuven, Leuven, Belgium for sharing the ISEA dataset. Luca Pagani of University of Padova, Padova Italy and Mait Metspalu & J. Rodrigo Flores of Estonian Biocenter, Tartu, Estonia for sharing the ISEA, Malagasy, and Estonian Genome Variation Project dataset.

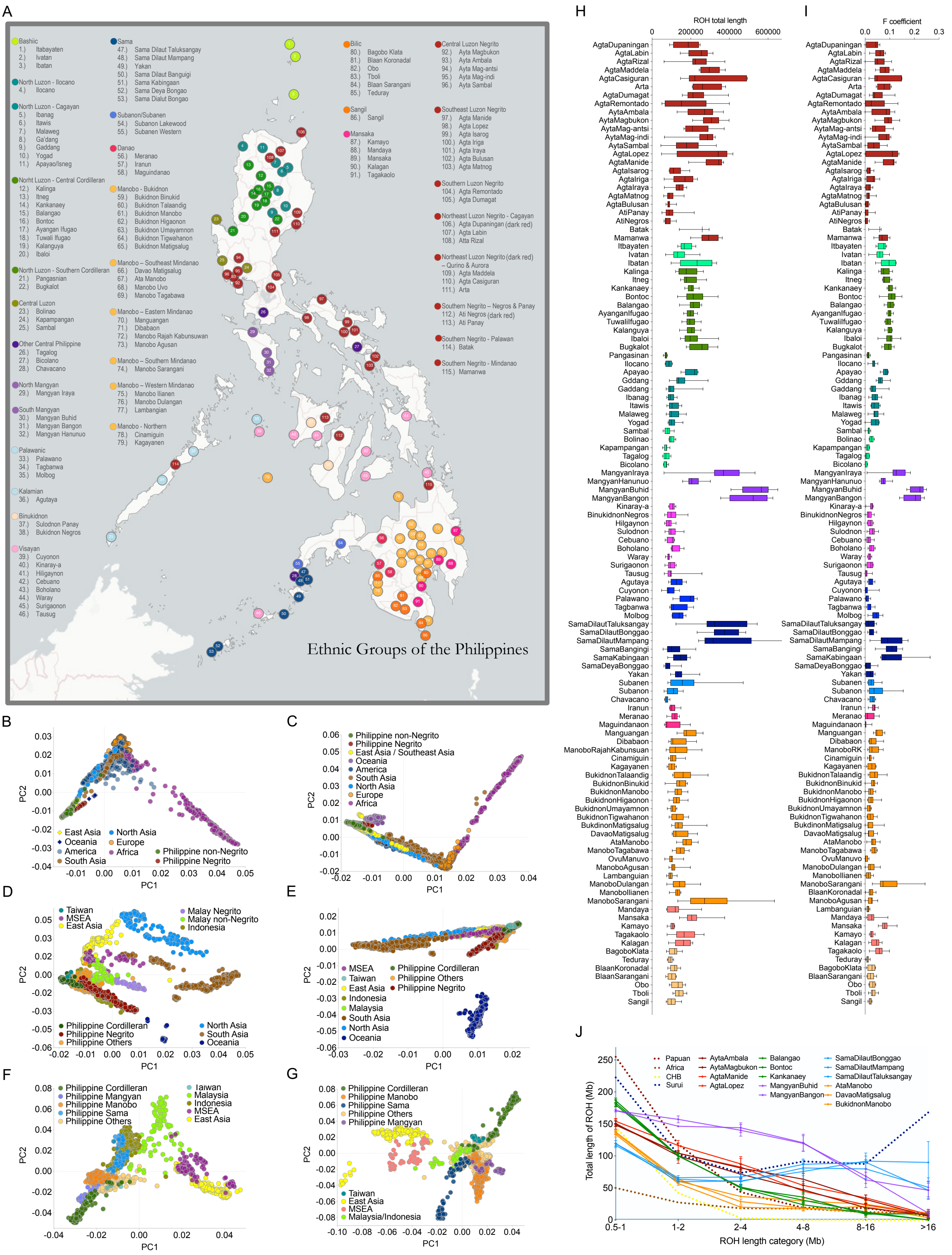

Figure S1. Ethnic groups of the Philippines and their genetic affiliation. (A) Location of 115 Philippine indigenous cultural communities labelled with a colour that corresponds to their ethnic group cluster. Principal Component analysis of worldwide populations using Phil_AsiaPacific_315K (B) or Phil_HO_201K (C) datasets; or of populations in the Asia-Pacific region using Phil_AsiaPacific_315K (D) or Phil_HO_201K (E) datasets; or of populations restricted to East Asia, Island Southeast Asia, and Mainland Southeast Asia using Phil_AsiaPacific_315K (F) or Phil_HO_201K (G) datasets. Runs of homozygosity (H) and F inbreeding coefficient (I) of Philippine ethnic groups. (J) Runs of homozygosity tracts of Philippine Negrito, Cordilleran, Mangyan, Manobo, and Sama ethnic groups, with Papuan, African, East Asian, and Native American reference populations.

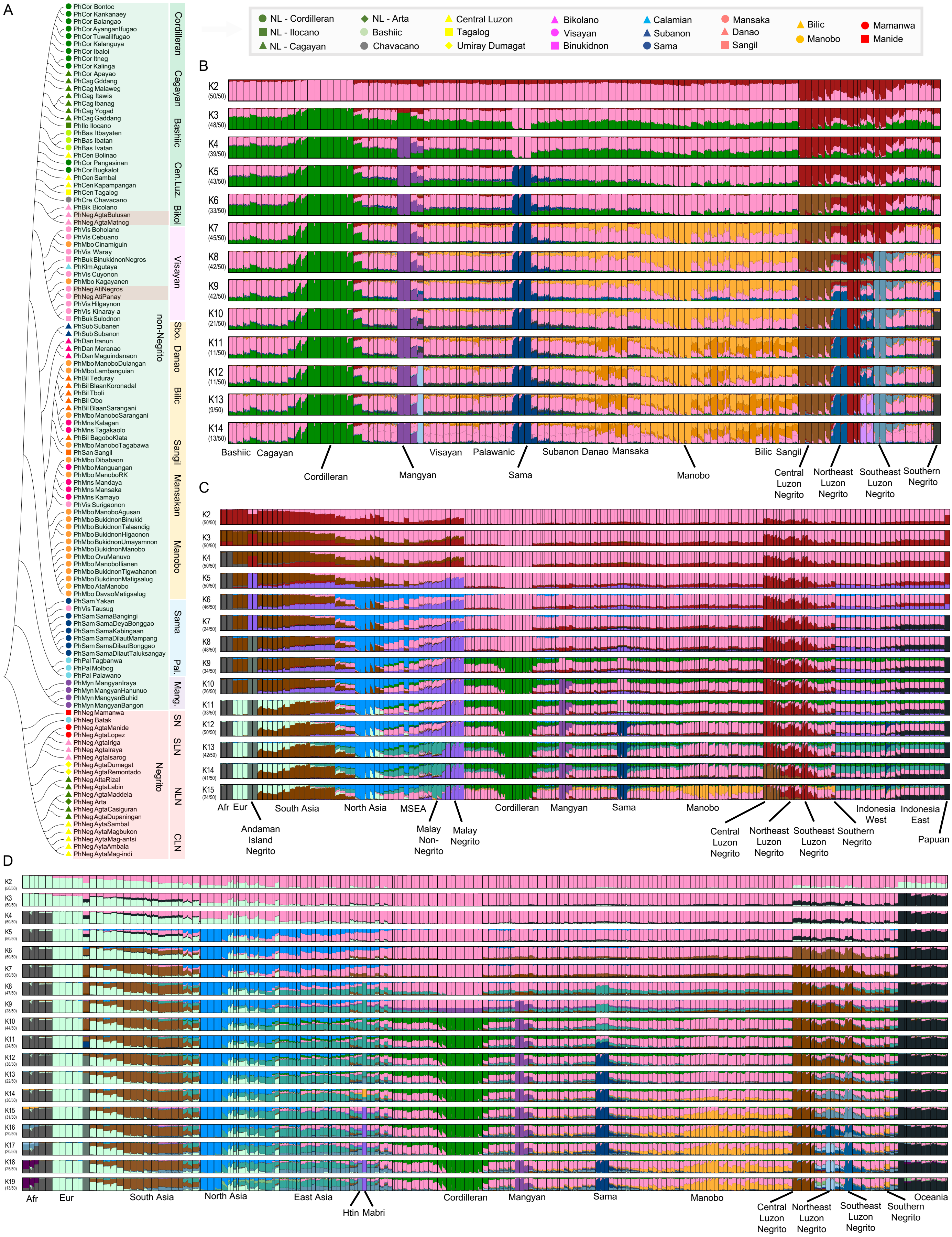

Figure S2. Population structure of the Philippines. (A) Neighbour-joining tree based on pairwise FST of Philippine ethnic groups. Neighbour-joining tree was plotted using MEGA7 software. Populations are shaded with light red colour to indicate self-identified Negritos or light green colour to indicate self-identified non-Negritos, and labelled with coloured markers to indicate ethnolinguistic classification. Ethnic group clusters are listed on the right side section of the main graph; with Sbo = Subanon, Pal = Palawanic, Mang = Mangyan, SN = Southern Negrito, SLN = Southeast Luzon Negrito, NLN = Northeast Luzon Negrito, and CLN = Central Luzon Negrito. Admixture analysis of Philippine populations using thePhil_2.35M dataset (B), of Asia-Pacific populations using Phil_AsiaPacific_315K dataset (C) or using Phil_HO_201K dataset (D).

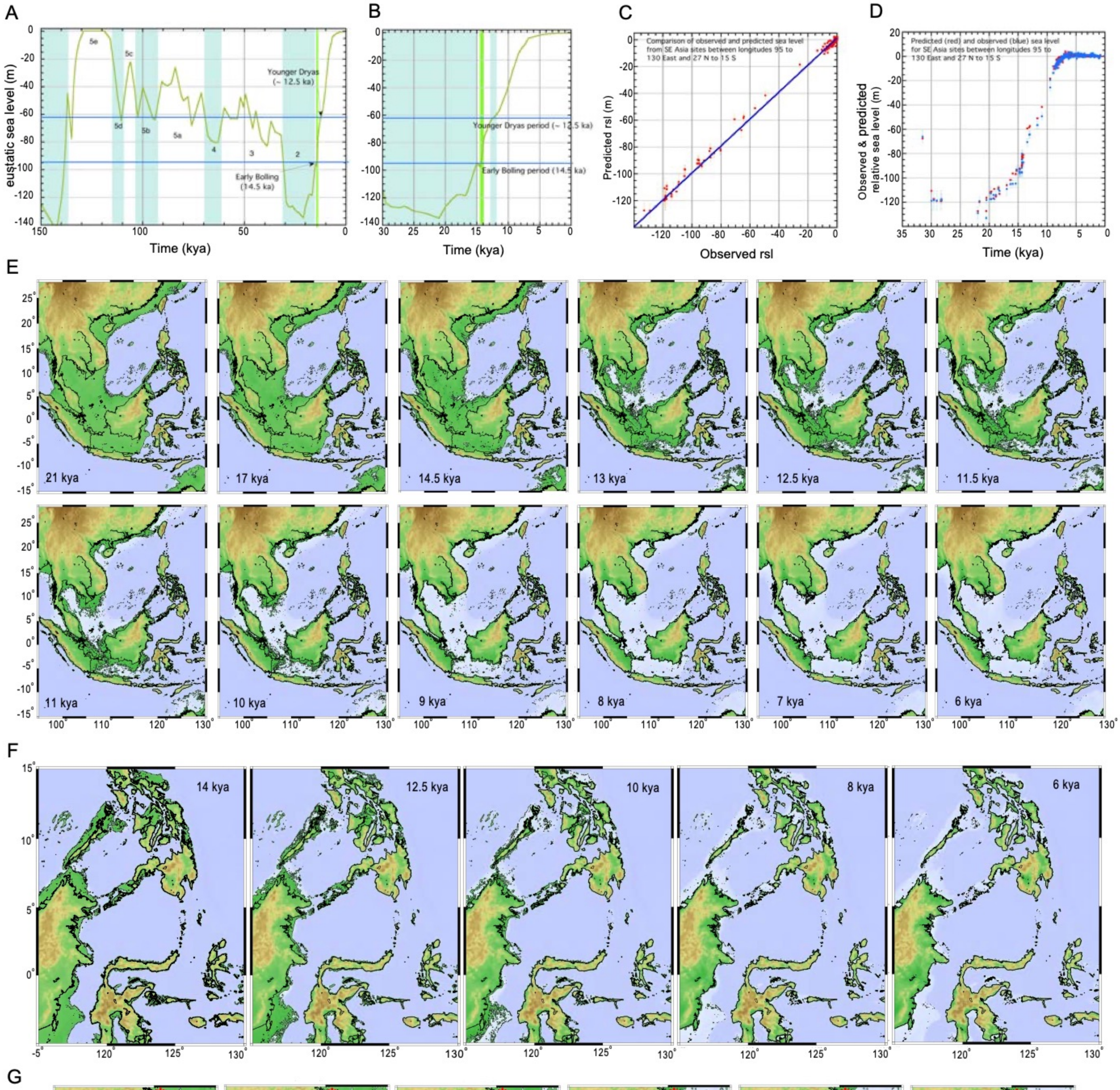

Figure S3 | Geologic features of ISEA. (A) Eustatic sea level (or equivalent-ice-volume function) for the last glacial cycle from 150 kya to present or from (B) the last 30 kya on an expanded scale. (C & D) Comparisons between observed and predicted model sea level from sites in Southeast Asia. (E) Reconstructions of the shoreline location for the entire Island Southeast Asia from 21 kya until 6 kya. (F) Higher resolution plots from 14 to 8 kya indicating possible land-based migration routes from Sabah to the Philippines. (G) Reconstructions of Taiwan-South China area from 14.5 to 6 kya, with red asterisk indicator for the location of Liangdao Island.

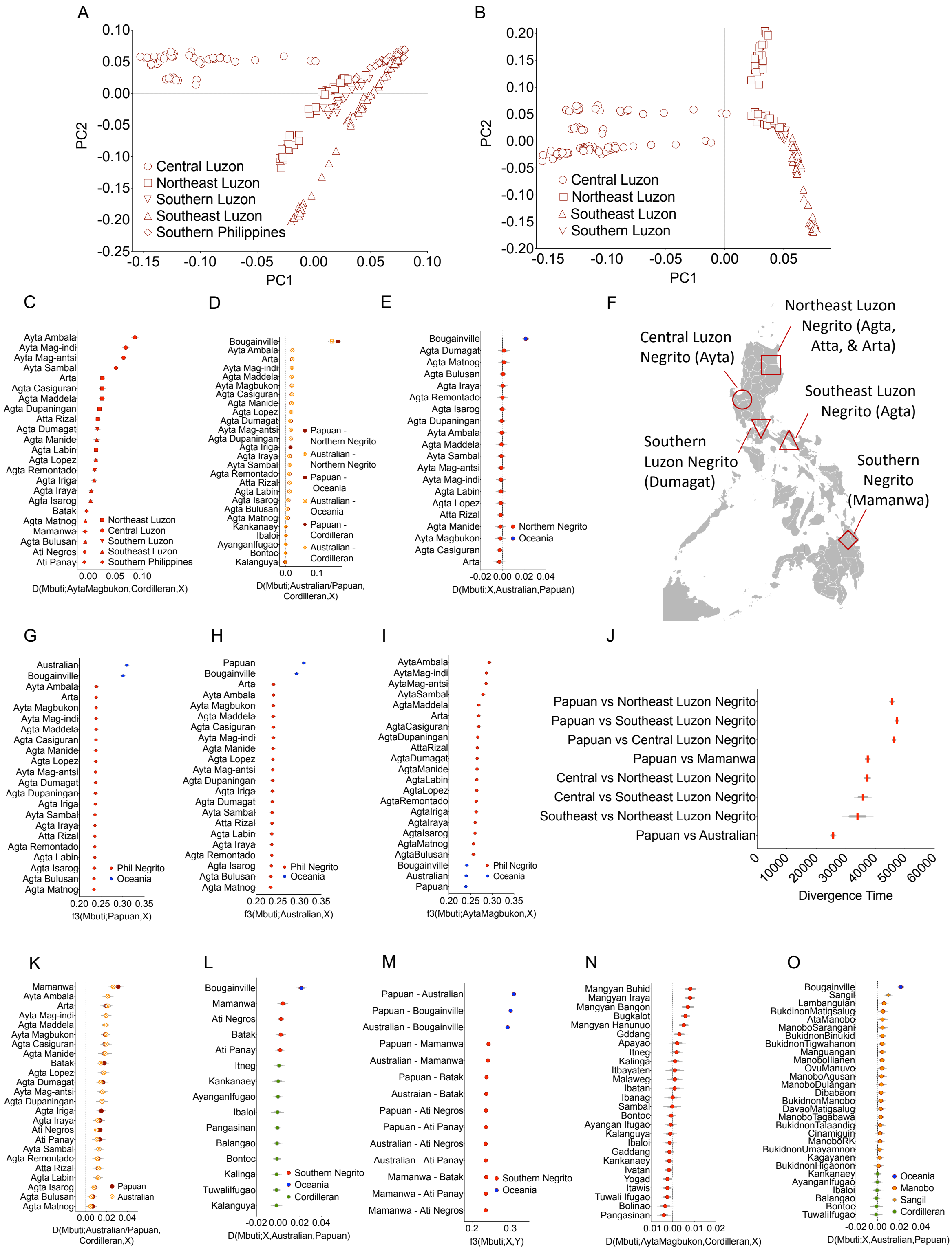

Figure S4. Deep-diverging populations of Philippine Negritos. (A) Principal Component analysis of all Philippine Negritos. (B) Principal Component analysis restricted to Northern Negritos of the Philippines. (C-E) Northern Negrito ancestry of the Philippines. (C) Levels of Northern Negrito ancestry using Ayta Magbukon as the reference population for least admixed Negrito. (D) Genetic relationships between Northern Negritos and Australians and Papuans. (E) Levels of non-Australian Papuan-related ancestry in Northern Negritos in comparison with the Bougainville Islander. (F) Map showing location of ethnic groups clusters of Northeast Luzon Negritos, Central Luzon Negritos, Southern Luzon Negritos, Southeast Luzon Negritos, and Southern Negritos of the Philippines. (G-I) Combination of outgroup f3 statistics showing Northern Negritos of the Philippines form as an outgroup to the Australian and Papuan clade. (J) Estimation of divergence time between Papuans and Northern Negritos, between Northern and Southern Negritos, between Papuan and Australian, and in between Northern Negritos. (K-M) Southern Negrito ancestry of the Philippines. (K) Genetic relationships between Southern Negritos (represented by Mamanwa) and Australians and Papuans. (L) Levels of non-Australian Papuan-related ancestry in Northern Negritos in comparison with Northern Negritos and Bougainville Islander. (M) Outgroup f3 statistics showing the relationships between Papuans, Australians, and Philippine Negritos, where both Northern and Southern Negritos form as an outgroup to the Australo-Papuan clade. (N) Level of Northern Negrito ancestry among non-Negrito populations of Luzon, using Ayta Magbukon as the reference population for the least admixed Negrito. (O) Levels of non-Australian Papuan-related ancestry in non-Negritos in comparison with the Bougainville Islander, likely as a result of more recent westward gene flow of Papuan-related ancestry into western Indonesia and southern Philippines. Thick and thin error bars represent 1 and 1.96 standard error of the estimate, respectively.

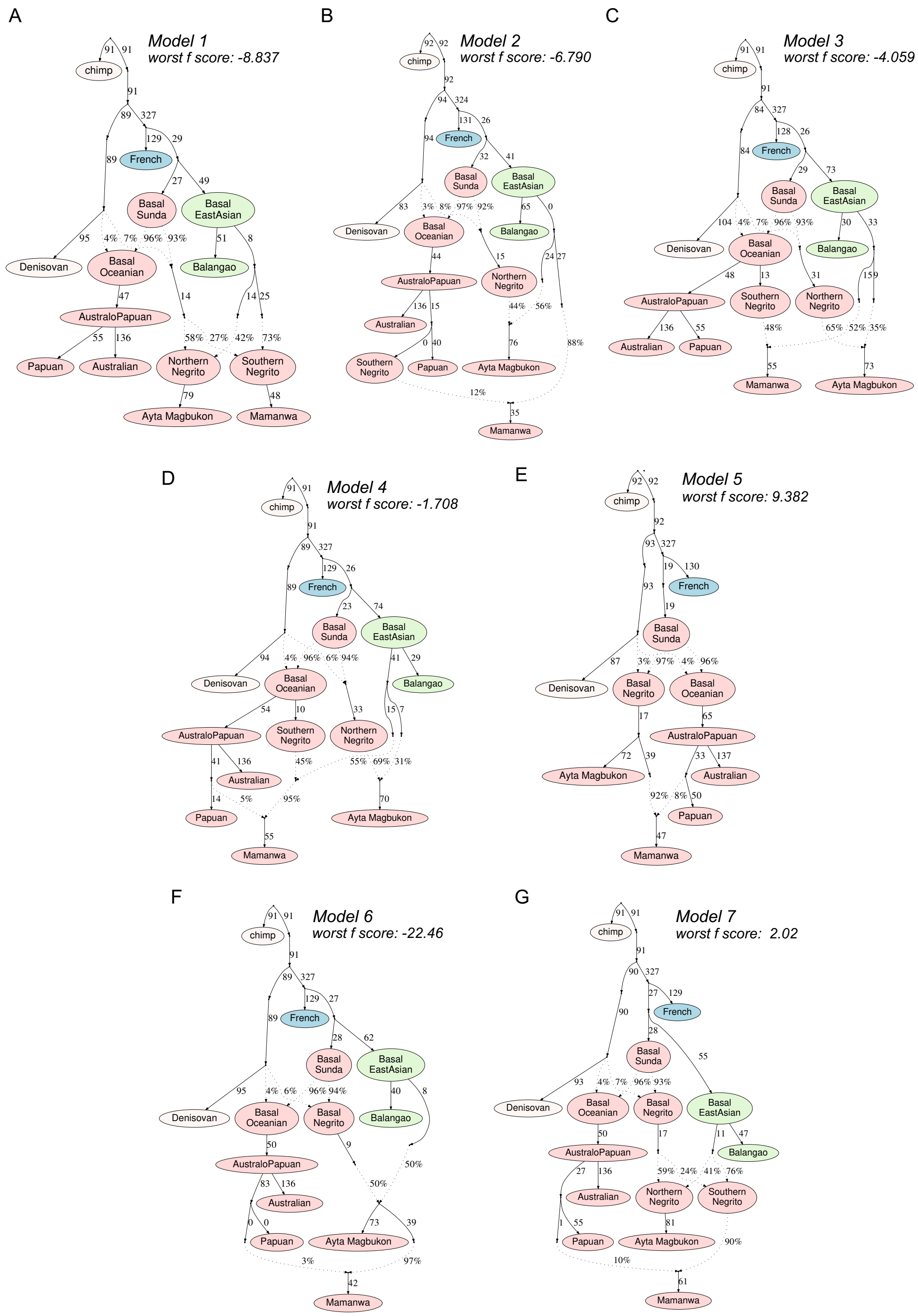

Figure S5. Genetic affiliation of Northern Negritos and Southern Negritos. Admixture graph depicting genetic relationships of Northern Negritos, Southern Negritos, AustraloPapuans, and East Asians, where Ayta Magbukon is modelled as an admixture between Basal East Asian and Northern Negrito ancestries (A-G). Admixture graph depicting models where the Australasian-related ancestry in Mamanwa are derived from divergence within Philippine Negritos (A), from Papuans following the AustraloPapuan split (B), from Basal Oceanians prior to the AustraloPapuan split (C), and from a combination of Basal Oceanian plus minimal contribution from Papuans post AustraloPapuan divergence (D), from Basal Negrito plus Papuan gene flow post AustraloPapuan divergence without East Asian admixture (E), from Basal Negrito with East Asian admixture then diverging into Northern Negritos and Southern Negritos plus Papuan gene flow post AustraloPapuan divergence (F), from Basal Negrito then diverging into Northern Negritos and Southern Negritos pluse East Asian admixture and Papuan gene flow post AustraloPapuan divergence (G).

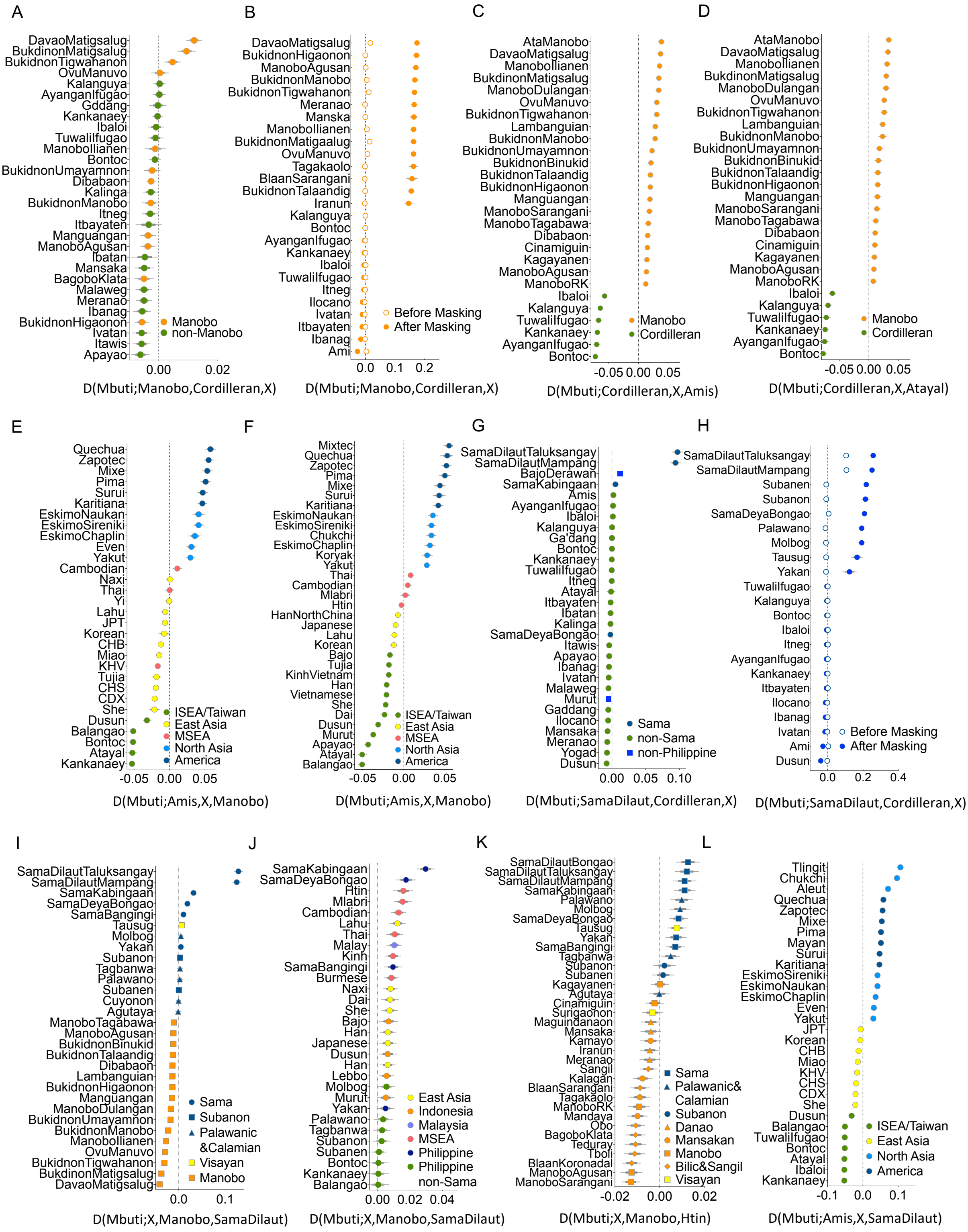

Figure S6. Manobo and Sama-related ancestries of southern Philippines. (A) Levels of Manobo-related ancestry using Ata Manobo as the reference least admixed Manobo population, and using Phil_1KGP_SGDP_1.69M dataset. (B) Levels of Manobo-related ancestry after masking non-Manobo-related alleles in populations. Genetic relationships between Manobo ethnic groups, Balangao Cordilleran, and Amis (C) or Atayal (D) using Phil_1KGP_SGDP_1.69M, showing Manobo is an outgroup to the Cordilleran-Aboriginal Taiwanese clade. Genetic relationships between Amis, Ata Manobo, and various ethnic groups of mainland Asia and the Americas using (E) Phil_1KGP_SGDP_1.69M or (F) Phil_HO_201K datasets, showing Manobo-related ancestry as an outgroup to the Amis-Han/Dai/Kinh clade. (G) Levels of Sama-related ancestry using Sama Dilaut Bongao as the reference population for the least admixed Sama, and using the Phil_1KGP_SGDP_1.69M dataset. (H) Levels of Sama-related ancestry after masking non-Sama-related alleles in populations. (I) Distinguishing between Manobo-related versus Sama-related ancestries among in southern Philippines using Ata Manobo as the reference last admixed Manobo and Sama Dilaut Bongao as the reference least admixed Sama. (J) Distinguishing levels of shared alleles between ISEA, MSEA, and continental Asian populations relative to Manobo or Sama Dilaut ethnic group. (K) Distinguishing between Manobo-related versus Htin-related ancestries among ethnic group of southern Philippines, indicating Sama-related and Palawanic ethnic groups of southwestern Philippines are more affiliated with Austroasiatic–speaking ethnic group of MSEA. (L) Genetic relationships between Amis, Sama Dilaut Bongao, and various ethnic groups of mainland Asia and the Americas using the Phil_1KGP_SGDP_1.69M dataset, showing Sama-related ancestry as an outgroup to the Amis- Han/Dai/Kinh clade. Thick and thin error bars represent 1 and 1.96 standard error of the estimate, respectively.

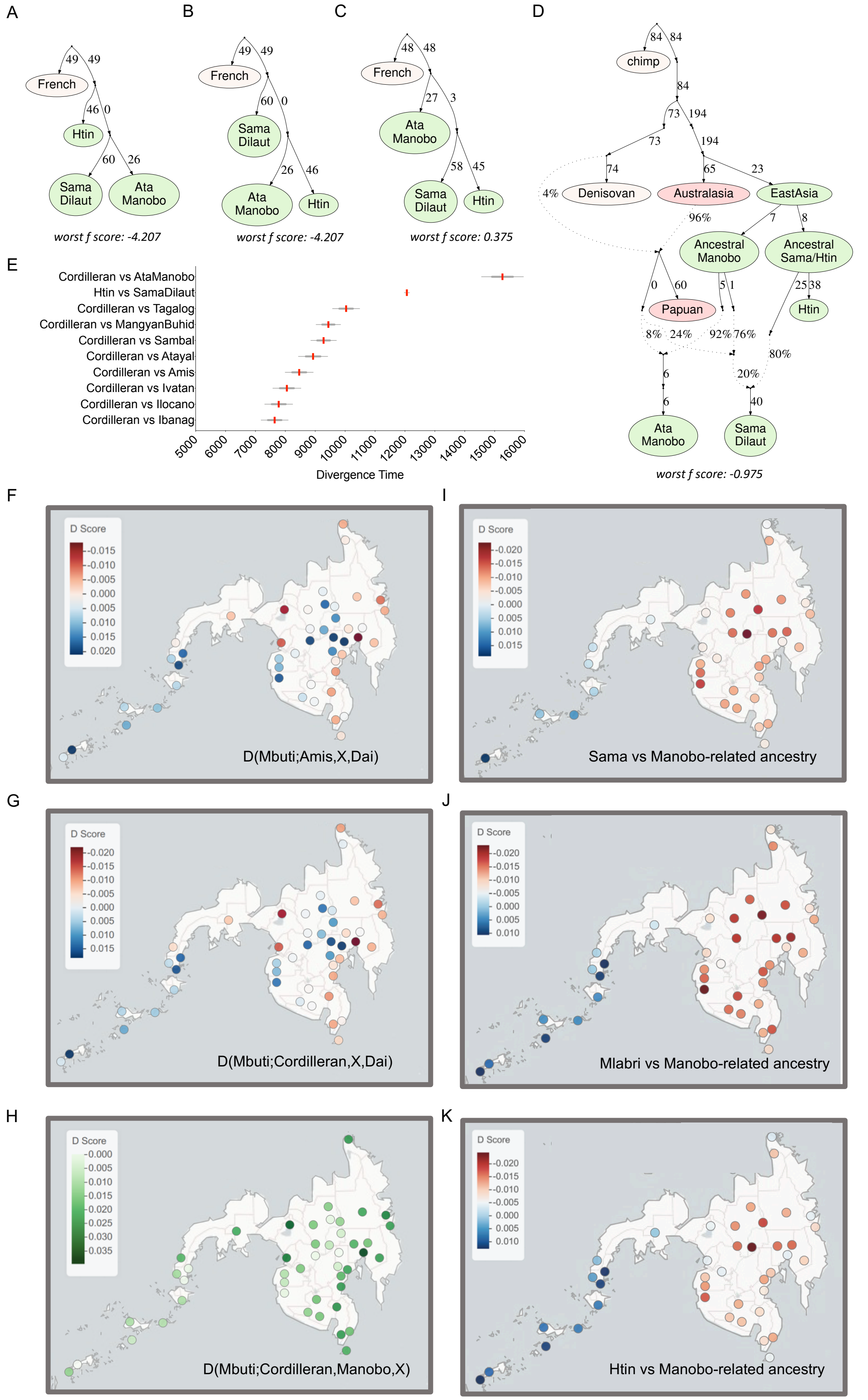

Figure S7. Genetic affiliations of Manobo and Sama-related ethnic groups of southern Philippines. (A-D) Admixture graph depicting genetic relationships of Manobo, Sama, and Htin-related populations. Sama Dilaut forms a clade with Htin relative to Ata Manobo (all f-statistics are within 0.37 standard error for C and within -0.97 standard error for D). (E) Estimation of divergence time between all Cordilleran ethnic groups and Amis, Atayal, Ata Manobo, Mangyan Buhid, or non-Negrito ethnic groups of northern Philippines, and between all Sama Dilaut ethnic groups and Htin ethnic group of MSEA. (F-H) Expansion of Cordilleran-related groups have the greatest impact on coastal populations of southern Philippines. Levels of Manobo-like (blue) versus Cordilleran-related (red) ancestry in Mindanao Island, using (F) Amis or (G) Cordilleran Balangao as the genetic signal for Cordilleran-related ancestry.(H) Levels of Cordilleran-related ancestry in Mindanao Island, relative to Ata Manobo. (I-K) Map showing distribution, levels, and genetic affiliation of Sama-related vs Manobo-related ancestries in southern Philippines. (I) Map of southern Philippines indicating levels of Sama-related (blue) versus Manobo-related (red) ancestries, based on the test D(Mbuti;X,SamaDilautBongao,AtaManobo). Map of southern Philippines indicating levels (J) Mlabri-related or (K) Htin-related (blue) versus Manobo-related (red) ancestries, based on the test D(Mbuti;X,Htin/Mlabri,AtaManobo), demonstrating that Sama-related groups have genetic affinity to Austroasiatic-speaking Htin and Mlabri ethnic groups of MSEA.

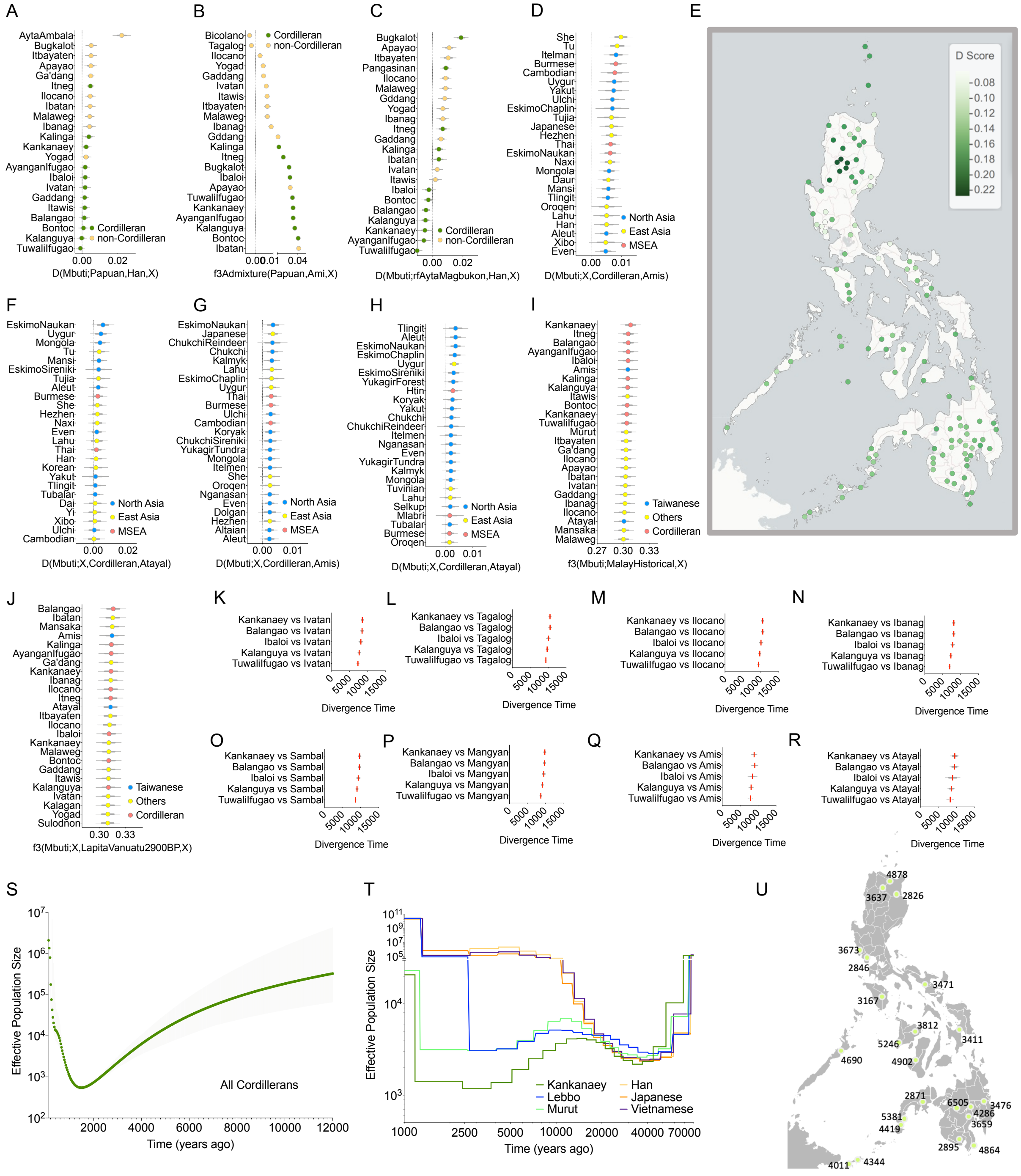

Figure S8. Cordillerans are the least admixed descendants of Basal East Asians. (A) Absence of Australasian ancestry in central Cordillerans, using Papuans as reference for least admixed Australasian. (B) f3 admixture tests showing absence of admixture signal in Cordillerans.. (C) Absence of Northern Negrito ancestry in Bontoc, Balangao, Kankanaey, Kalanguya, Tuwali Ifugao, Ayangan Ifugao, and Ibaloi Cordillerans, with the use of rfmix-treated Ayta Magbukon as a reference population with 'pure' Northern Negrito ancestry. Both (D,F) Amis and (G,H) Atayal share more alleles with ethnic groups with northern East Asian ancestry relative to Cordilleran Balangao. The analysis was done in both (D,G) Phil_1KGP_SGDP_1.69M and (F,H) Phil_HO_201K datasets. (E) Map showing levels of Cordilleran-related ancestry in present-day of the Philippines based on the test D(Mbuti;X,Papuan,Balangao). Cordillerans serve as one of the best surrogate genetic signal for the East Asian ancestry detected in historical Malay (I) and ancient Lapita (J) individuals. Estimation of divergence time between Cordillerans (Kankanaey, Balangao, Ibaloi, Kalanguya, Tuwali Ifugao) and Ivatan (K), Tagalog (L), Ilocano (M), Ibanag (N), Sambal (O), Mangyan (P), Amis (Q), & Atayal (R) ethnic groups, whre most divergence date estimates fall within ~ 8-10 kya. (S) Effective population size of Cordilleran ethnic groups based on shared IBD, and calculated with the software IBDne. (T) Effective population size inferred with MSMC using 4 genomes for each ISEA or mainland Asian populations (adapted from Pagani *et al*., 2016), showing a consistent bottleneck among populations of ISEA (Kankanaey Cordilleran, Lebbo, and Murut) since ~ 10 kya. (U) Map with geographic location of top 25 ethnic groups with the oldest estimated dates of admixture between Cordilleran-related and Negrito-related populations using Malder. Thick and thin error bars represent 1 and 1.96 standard error of the estimate, respectively.

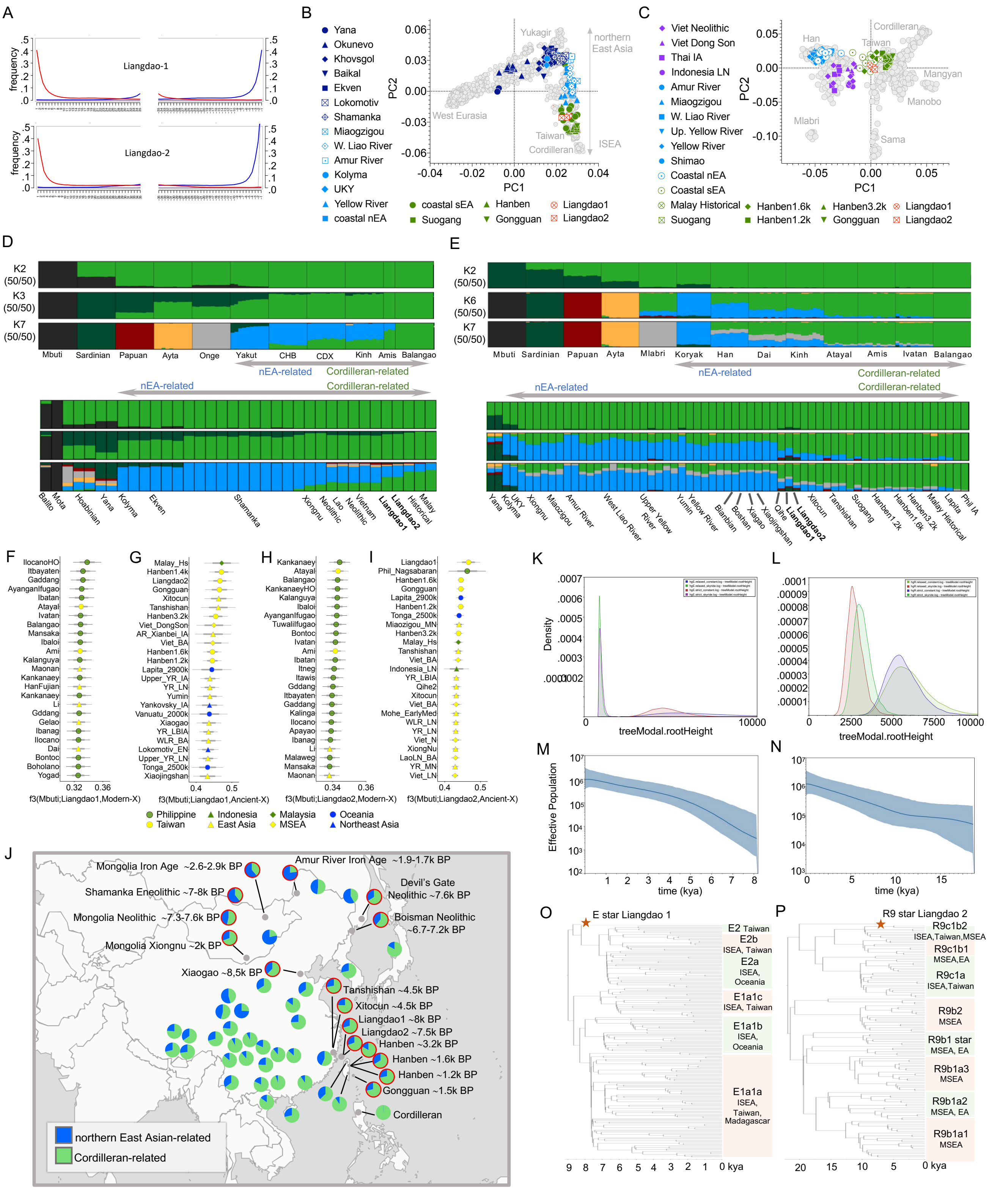

Figure S9. Genetic affiliation of Liangdao individuals. (A) Cytosine deamination patterns for Liangdao-1 and Liangdao-2 individuals. Principal Component analysis applying least square equations projection, and using the Phil_HO_Ancient_Transv_32K dataset. Liangdao individuals and other ancient individuals (genome coverage of at least 0.05x) are projected to present-day ethnic groups of Eurasia (B) or of ISEA, MSEA, and continental East Asia (C). Two Liagndao2 samples represent the same individual generated independently in this study and in Yang et al., (2020). Admixture analysis of present-day Asia-Pacific populations & ancient individuals using the Phil_1KGP_SGDP_Ancient_Transv_317K (D) or the Phil_HO_Ancient_Transv_32K (E) datasets, where Liangdao individuals demonstrated to have a combination of Cordilleran-related and northern East Asian (nEA)-related ancestries. Estimation of shared drift, using f3 statistics, between Liangdao-1 (F,G) & Liangdao-2 (H,I) and ancient & present-day Asia-Pacific populations using Phil_HO_Ancient_202K dataset. (J) Proportions of northern East-Asian and Cordilleran-related ancestries among populations estimated using qpAdm, assuming a 2-source admixture, with UKY/Kolyma as a proxy for northern East Asian ancestry and Kankanaey as a proxy for Cordilleran-related ancestry. Model comparison for the analyses of mitochondrial DNA haplogroups E (K) and R9 (L), using two different tree models. In both analyses the support for the skyride model does not justify the use of a model that assumes a constant population size. The choice clearly influences the root height. The results for strict and relaxed molecular clocks are shown in both cases, and the relaxed clock is strongly favored. Skyride plots from the Bayesian phylogenetic analyses of mitochondrial DNA haplogroups E (M) and hg R9 (N). The steep angle of the curve from the start of the plot in (O) is consistent with an expansion of this haplogroup commencing ~8 Kya, which started to tail off just before 5 Kya, the time traditionally associated with the spread of the Austronesian language family. For R9 (P) a slightly less marked expansion also occurs ~8 Kya but continues until the present. The difference may be due to R9 occurring over a much wider geographic area, reflected in the higher estimated effective population size. Phylogenetic trees inferred from the Bayesian analyses of complete mitochondrial genomes from haplogroup E (O) or R (P) using the software BEAST.

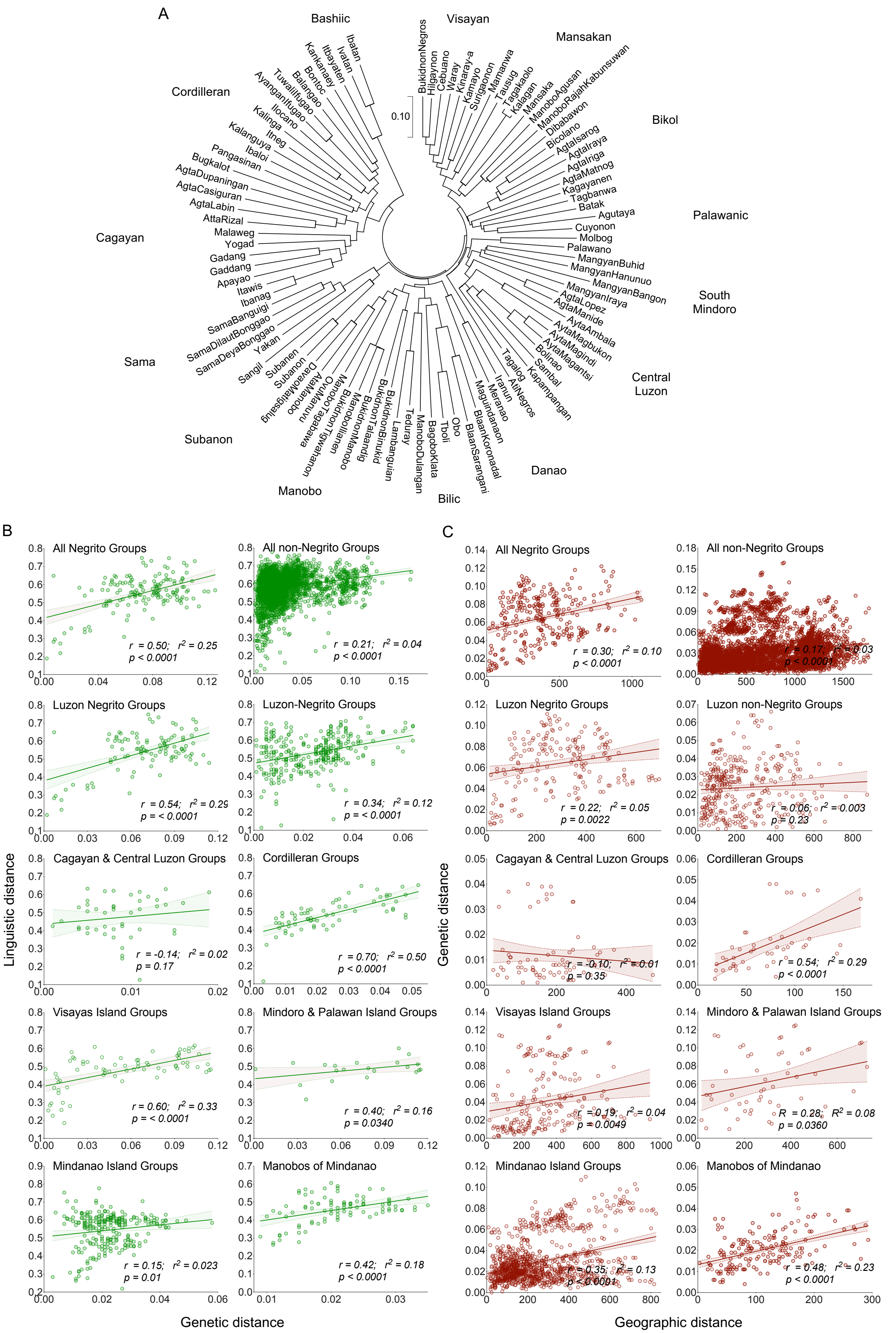

Figure S10. Correlation analysis between genetic, linguistic, or geographic distance. a, Neighbour-joining tree of Philippine languages based on linguistic distance. **b & c,** Correlation between genetic distance based on pairwise FST and linguistic distance based on proportion of cognate similarities or between genetic distance and geographic distance in km among all Negritos, all non-Negritos, Northern Negritos of Luzon, non-Negritos of Luzon, ethnic groups of Cagayan Valley and Central Luzon, ethnic groups of Cordillera region, ethnic groups of the Visayas region, ethnic groups of Mindoro and Palawan group of islands, ethnic groups of Mindanao Island, and Manobo ethnic groups of Mindanao Island.

References

1. S. Purcell *et al.*, PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575 (2007).
2. G. R. Abecasis *et al.*, An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65 (2012).
3. S. Mallick *et al.*, The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201-206 (2016).
4. P. Kusuma *et al.*, Contrasting linguistic and genetic origins of the Asian source populations of Malagasy. *Scientific reports* **6**, 26066 (2016).
5. P. Kusuma *et al.*, The last sea nomads of the Indonesian archipelago: genomic origins and dispersal. *European Journal of Human Genetics* **25**, 1004 (2017).
6. M. P. Cox *et al.*, Small traditional human communities sustain genomic diversity over microgeographic scales despite linguistic isolation. *Molecular biology and evolution* **33**, 2273-2284 (2016).
7. A. Mörseburg *et al.*, Multi-layered population structure in Island Southeast Asians. *European Journal of Human Genetics* **24**, 1605 (2016).
8. F. Aghakhanian *et al.*, Unravelling the genetic history of Negritos and indigenous populations of Southeast Asia. *Genome Biol Evol* **7**, 1206-1215 (2015).
9. A. Basu, N. Sarkar-Roy, P. P. Majumder, Genomic reconstruction of the history of extant populations of India reveals five distinct ancestral components and a complex structure. *Proc Natl Acad Sci U S A* **113**, 1594-1599 (2016).
10. I. Lazaridis *et al.*, Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419-424 (2016).
11. N. Patterson *et al.*, Ancient admixture in human history. *Genetics* **192**, 1065-1093 (2012).
12. M. Lipson *et al.*, Ancient genomes document multiple waves of migration in Southeast Asian prehistory. *Science*, eaat3188 (2018).
13. N. Nakatsuka *et al.*, The promise of discovering population-specific disease-associated genes in South Asia. *Nat Genet* **49**, 1403-1407 (2017).
14. M. Lipson *et al.*, Population Turnover in Remote Oceania Shortly after Initial Settlement. *Curr Biol* **28**, 1157-1165.e1157 (2018).
15. P. Skoglund *et al.*, Genomic insights into the peopling of the Southwest Pacific. *Nature* **538**, 510 (2016).
16. A. L. Price *et al.*, Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-909 (2006).
17. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. *PLoS Genet* **2**, e190 (2006).
18. S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**, 1870-1874 (2016).
19. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome research* **19**, 1655-1664 (2009).
20. M. Jakobsson, N. A. Rosenberg, CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801-1806 (2007).
21. M. Lipson, D. Reich, A Working Model of the Deep Relationships of Diverse Modern Human Genetic Lineages Outside of Africa. *Mol Biol Evol* **34**, 889-902 (2017).
22. J. K. Pickrell *et al.*, Ancient west Eurasian ancestry in southern and eastern Africa. *Proc Natl Acad Sci U S A* **111**, 2632-2637 (2014).

23. B. K. Maples, S. Gravel, E. E. Kenny, C. D. Bustamante, RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. *Am J Hum Genet* **93**, 278-288 (2013).
24. B. L. Browning, S. R. Browning, Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics* **194**, 459-471 (2013).
25. S. R. Browning, B. L. Browning, Accurate Non-parametric Estimation of Recent Effective Population Size from Segments of Identity by Descent. *Am J Hum Genet* **97**, 404-418 (2015).
26. C. M. Schlebusch *et al.*, Genomic variation in seven Khoe-San groups reveals adaptation and complex African history. *Science* **338**, 374-379 (2012).
27. J. Wakeley (2008) Coalescent Theory. Roberts & Co. (Greenwood Village, Colorado, USA.[Google Scholar]).
28. P. Skoglund, A. Götherström, M. Jakobsson, Estimation of population divergence times from non-overlapping genomic sequences: examples from dogs and wolves. *Mol Biol Evol* **28**, 1505-1517 (2011).
29. C. M. Schlebusch *et al.*, Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. *Science* **358**, 652-655 (2017).
30. L. Pagani *et al.*, Genomic analyses inform on migration events during the peopling of Eurasia. *Nature* **538**, 238-242 (2016).
31. A. Bergström *et al.*, A Neolithic expansion, but strong genetic structure, in the independent history of New Guinea. *Science* **357**, 1160-1163 (2017).
32. P. Sjödin (2020) Estimating divergence times from DNA sequences. In revision., ed J. McKenna.
33. D. Y. Yang, B. Eng, J. S. Waye, J. C. Dudar, S. R. Saunders, Technical note: improved DNA extraction from ancient bones using silica-based spin columns. *Am J Phys Anthropol* **105**, 539-543 (1998).
34. H. Malmström *et al.*, More on contamination: the use of asymmetric molecular behavior to identify authentic ancient human DNA. *Mol Biol Evol* **24**, 998-1004 (2007).
35. M. Meyer, M. Kircher, Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* **2010**, pdb.prot5448 (2010).
36. T. Günther *et al.*, Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation. *PLoS Biol* **16**, e2003703 (2018).
37. A. W. Briggs, P. Heyn, Preparation of next-generation sequencing libraries from damaged DNA. *Methods Mol Biol* **840**, 143-154 (2012).
38. J. Dabney *et al.*, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl Acad Sci U S A* **110**, 15758-15763 (2013).
39. H. Li, R. Durbin, Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595 (2010).
40. I. Lazaridis *et al.*, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409-413 (2014).
41. P. Skoglund *et al.*, Genomic diversity and admixture differs for Stone-Age Scandinavian foragers and farmers. *Science* **344**, 747-750 (2014).
42. M. Kircher, "Analysis of high-throughput ancient DNA sequencing data" in Ancient DNA. (Springer, 2012), pp. 197-228.
43. A. W. Briggs *et al.*, Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci U S A* **104**, 14616-14621 (2007).

44. R. E. Green *et al.*, A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* **134**, 416-426 (2008).
45. Q. Fu *et al.*, A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol* **23**, 553-559 (2013).
46. G. Renaud, V. Slon, A. T. Duggan, J. Kelso, Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol* **16**, 224 (2015).
47. A. Furtwängler *et al.*, Ratio of mitochondrial to nuclear DNA affects contamination estimates in ancient DNA analysis. *Sci Rep* **8**, 14075 (2018).
48. M. A. Suchard *et al.*, Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol* **4**, vey016 (2018).
49. A. Rambaut, A. J. Drummond, D. Xie, G. Baele, M. A. Suchard, Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst Biol* **67**, 901-904 (2018).
50. R. Hall, Southeast Asia's changing palaeogeography. *Blumea-Biodiversity, Evolution and Biogeography of Plants* **54**, 148-161 (2009).
51. R. Hall, The Eurasian SE Asian margin as a modern example of an accretionary orogen. *Geological Society, London, Special Publications* **318**, 351-372 (2009).
52. Z. Huang, D. Zhao, L. Wang, P wave tomography and anisotropy beneath Southeast Asia: Insight into mantle dynamics. *Journal of Geophysical Research: Solid Earth* **120**, 5154-5174 (2015).
53. K. Lambeck, J. Chappell, Sea level change through the last glacial cycle. *Science* **292**, 679-686 (2001).
54. K. Lambeck, Y. Yokoyama, T. Purcell, Into and out of the Last Glacial Maximum: sea-level change during Oxygen Isotope Stages 3 and 2. *Quaternary Science Reviews* **21**, 343-360 (2002).
55. K. Lambeck, H. Rouby, A. Purcell, Y. Sun, M. Sambridge, Sea level and global ice volumes from the Last Glacial Maximum to the Holocene. *Proc Natl Acad Sci U S A* **111**, 15296-15303 (2014).
56. J. Chappell, Geology of coral terraces, Huon Peninsula, New Guinea: a study of Quaternary tectonic movements and sea-level changes. *Geological Society of America Bulletin* **85**, 553-570 (1974).
57. A. A. Behr, K. Z. Liu, G. Liu-Fang, P. Nakka, S. Ramachandran, pong: Fast analysis and visualization of latent clusters in population genetic data. *Bioinformatics* **32**, 2817-2823 (2016).
58. G. Chaubey, P. Endicott, The Andaman Islanders in a regional genetic context: reexamining the evidence for an early peopling of the archipelago from South Asia. *Hum Biol* **85**, 153-172 (2013).
59. F. C. Ceballos, P. K. Joshi, D. W. Clark, M. Ramsay, J. F. Wilson, Runs of homozygosity: windows into population history and trait architecture. *Nat Rev Genet* **19**, 220-234 (2018).
60. A. S. Malaspinas *et al.*, A genomic history of Aboriginal Australia. *Nature* **538**, 207-214 (2016).
61. L. A. Reid, Re-evaluating the position of Iraya among Philippine languages. *JSEALS*, 23 (2017).
62. L. A. Reid, Possible non-Austronesian lexical elements in Philippine Negrito languages. *Oceanic Linguistics*, 37-72 (1994).
63. L. A. Reid, "Historical linguistics and Philippine hunter-gatherers". (2007).
64. L. A. Reid, Who are the Philippine negritos? Evidence from language. *Hum Biol* **85**, 329-358 (2013).

65. J. V. Moreno-Mayar *et al.*, Early human dispersals within the Americas. *Science* **362** (2018).
66. O. Delaneau, J. Marchini, J. F. Zagury, A linear complexity phasing method for thousands of genomes. *Nat Methods* **9**, 179-181 (2011).
67. O. Delaneau, J. Marchini, G. P. Consortium, Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. *Nat Commun* **5**, 3934 (2014).
68. L. A. Reid, Morphological evidence for Austric. *Oceanic linguistics*, 323-344 (1994).
69. L. A. Reid, A review and evaluation of the lexical and morphosyntactic evidence. *The peopling of East Asia: Putting together archaeology, linguistics and genetics*, 132 (2005).
70. G. Jäger, Support for linguistic macrofamilies from weighted sequence alignment. *Proc Natl Acad Sci U S A* **112**, 12752-12757 (2015).
71. W. Schmidt, *Die Mon-Khmer-völker: ein bindeglied zwischen völkern Zentralasiens und Austronesiens* (F. Vieweg und sohn, 1906).
72. H. H. La Vaughn, On the Track of Austric _ _ Part II. Consonant Mutation in Early Austroasiatic. *Mon-Khmer Studies* **27**, 13 (1997).
73. H.-c. Hung, Prosperity and complexity without farming: the South China Coast, c. 5000–3000 BC. *Antiquity*, 1-17 (2019).
74. P. Grang (2017) A collaboration between linguistics and genetics: tracing the Bajo diaspora history in Eastern Indonesia. in *2nd International Conference on Innovative Research Across Disciplines (ICIRAD 2017)* (Atlantis Press).
75. C. Sather, *The Bajau Laut: Adaptation, history, and fate in a maritime fishing society of South-eastern Sabah* (Oxford University Press, USA, 1997).
76. M. Lipson *et al.*, Reconstructing Austronesian population history in Island Southeast Asia. *Nat Commun* **5**, 4689 (2014).
77. P. A. Soares *et al.*, Resolving the ancestry of Austronesian-speaking populations. *Human genetics* **135**, 309-326 (2016).
78. K. Tätte *et al.*, The genetic legacy of continental scale admixture in Indian Austroasiatic speakers. *Sci Rep* **9**, 3818 (2019).
79. G. Chaubey *et al.*, Population genetic structure in Indian Austroasiatic speakers: the role of landscape barriers and sex-specific admixture. *Mol Biol Evol* **28**, 1013-1024 (2011).
80. B. Bellina, The archaeology of early contact with India and the Mediterranean world, from the fourth century BC to the fourth century AD. *Southeast Asia: from prehistory to history*, 68-88 (2004).
81. N. Brucato *et al.*, Genomic admixture tracks pulses of economic activity over 2,000 years in the Indian Ocean trading network. *Scientific Reports* **7**, 2919 (2017).
82. P. Kusuma *et al.*, Western Eurasian genetic influences in the Indonesian archipelago. *Quaternary International* **416**, 243-248 (2016).
83. G. Hudjashov *et al.*, Complex patterns of admixture across the Indonesian archipelago. *Molecular biology and evolution* **34**, 2439-2452 (2017).
84. P. Bellwood, E. Dizon, The Batanes archaeological project and the “Out of Taiwan” hypothesis for Austronesian dispersal. *Journal of Austronesian Studies* **1**, 1-33 (2005).
85. M. Klamer, The dispersal of Austronesian languages in Island South East Asia: Current findings and debates. *Language and Linguistics Compass* **13**, e12325 (2019).
86. G. F. Simons, C. D. Fennig, *Ethnologue: Languages of Asia* (sil International, 2017).
87. H. Hammarström, R. Forkel, M. Haspelmath, Glottolog 3.0. *Max Planck Institute for the Science of Human History* (2017).

88. S. Wang *et al.*, Geographic patterns of genome admixture in Latin American Mestizos. *PLoS Genet* **4**, e1000037 (2008).
89. A. Ruiz-Linares *et al.*, Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. *PLoS Genet* **10**, e1004572 (2014).
90. C. Chen (2003) *The Excavation of Daowei Site Group on Liangdao , Matsu Archipelago and the Reconstruction of Liangdao Ma* n. ed H. Chiu, Yu, K., Yin, Y. & Lin, F. (Liangjian Government, Liangjian).
91. Z. Chi, H.-c. Hung, Later hunter-gatherers in southern China, 18 000–3000 BC. *Antiquity* **86**, 11-29 (2012).
92. H. Matsumura *et al.*, Craniometrics Reveal "Two Layers" of Prehistoric Human Dispersal in Eastern Eurasia. *Scientific reports* **9**, 1451 (2019).
93. H. Matsumura, K. Shinoda, *Bio-anthropological Studies of Early Holocene Hunter-gatherer Sites at Huiyaotian and Liyupo in Guangxi, China* (National Museum of Nature and Science Tokyo, 2017).
94. G. Jun *et al.*, Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. *Am J Hum Genet* **91**, 839-848 (2012).
95. H. Jónsson, A. Ginolhac, M. Schubert, P. L. Johnson, L. Orlando, mapDamage2. 0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682-1684 (2013).
96. A. M. Ko *et al.*, Early Austronesians: into and out of Taiwan. *Am J Hum Genet* **94**, 426-436 (2014).
97. P. Soares *et al.*, Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet* **84**, 740-759 (2009).
98. D. M. Behar *et al.*, A "Copernican" reassessment of the human mitochondrial DNA tree from its root. *Am J Hum Genet* **90**, 675-684 (2012).
99. A. Rieux *et al.*, Improved calibration of the human mitochondrial clock using ancient genomes. *Mol Biol Evol* **31**, 2780-2792 (2014).
100. P. Bellwood, J. J. Fox, D. Tryon, *The Austronesians: historical and comparative perspectives* (ANU E Press, 2006).
101. P. J. Sheppard *et al.*, Lapita colonization across the Near/Remote Oceania boundary. *Current Anthropology* **52**, 000-000 (2011).
102. D. Pierron *et al.*, Genome-wide evidence of Austronesian-Bantu admixture and cultural reversion in a hunter-gatherer group of Madagascar. *Proc Natl Acad Sci U S A* **111**, 936-941 (2014).
103. D. Pierron *et al.*, Genomic landscape of human diversity across Madagascar. *Proc Natl Acad Sci U S A* **114**, E6498-E6506 (2017).
104. M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* **30**, E386-394 (2009).
105. P. Soares *et al.*, Climate change and postglacial human dispersals in southeast Asia. *Mol Biol Evol* **25**, 1209-1218 (2008).
106. W. Kutanan *et al.*, Complete mitochondrial genomes of Thai and Lao populations indicate an ancient origin of Austroasiatic groups and demic diffusion in the spread of Tai-Kadai languages. *Hum Genet* **136**, 85-98 (2017).
107. W. Kutanan *et al.*, New insights from Thailand into the maternal genetic history of Mainland Southeast Asia. *Eur J Hum Genet* **26**, 898-911 (2018).
108. M. A. Yang *et al.*, Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science* (2020).
109. R. D. Gray, A. J. Drummond, S. J. Greenhill, Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* **323**, 479-483 (2009).

110. S. J. Greenhill *et al.*, Evolutionary dynamics of language systems. *Proc Natl Acad Sci U S A* **114**, E8822-E8829 (2017).
111. R. Blust, *The austronesian languages* (Asia-Pacific Linguistics, School of Culture, History and Language, College …, 2013).
112. M. Ross, The Batanic languages in relation to the early history of the Malayo-Polynesian subgroup of Austronesian. *Journal of Austronesian Studies* **1**, 1-24 (2005).
113. R. Blust (1999) Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. in *Selected papers from the eighth international conference on Austronesian linguistics*, pp 31-94.
114. L. A. Reid, Revisiting the position of Philippine languages in the Austronesian family. *Br. Andrew Gonzalez (FSC) Distinguished Professorial Chair Lecture," organized by Linguistic Society of the Philippines and held at De La Salle University, Manila* **18** (2017).
115. R. A. Blust, The proto-Austronesian pronouns and Austronesian subgrouping. *Working Papers in Linguistics, University of Hawaii, Honolulu* **9**, 1-15 (1977).
116. L. A. Reid, Accounting for variability in Malayo-Polynesian pronouns. *Journal of Historical Linguistics* **6**, 130-164 (2016).
117. R. Blust, The linguistic macrohistory of the Philippines: Some speculations. *Current issues in Philippine linguistics and anthropology: Parangal kay Lawrence A. Reid*, 31-68 (2005).
118. R. Blench, Was there an Austroasiatic presence in Island Southeast Asia prior to the Austronesian expansion? *Bulletin of the Indo-Pacific Prehistory Association* **30**, 133-144 (2011).
119. A. Pawley (2006) Origins of the Filipinos: Linguistic and archaeological evidence. in *9th Philippine Linguistics Congress (9 PLC), University of the Philippines, Diliman, January*, pp 25-27.
120. A. Pawley (1999) Chasing rainbows: implications of the rapid dispersal of Austronesian languages for subgrouping and reconstruction. in *Selected papers from the eighth international conference on Austronesian linguistics* (Symposium Series of the Institute of Linguistics, Academia Sinica), pp 95-138.
121. L. A. Reid, 6. Modeling the Linguistic Situation in the Philippines. *Senri Ethnological Studies* **98**, 91-105 (2018).
122. S. Kalyan, A. François, H. Hammarström (2018) Understanding language genealogy: Alternatives to the tree model.
123. K. Ritsuko, L. A. Reid, Let's talk about trees: Genetic relationships of languages and their phylogenic representation. *Senri Ethnological Studies (Japan) eng no. 98* (2018).
124. J. F. Eder, Of fishers and farmers: Ethnicity and resource use in coastal Palawan. *Philippine quarterly of culture and society* **31**, 207-225 (2003).
125. F. M. Keesing, The Isneg: Shifting Cultivators of the Northern Philippines. *Southwestern Journal of Anthropology* **18**, 1-19 (1962).
126. L. C. Robinson, *Dupaningan Agta: Grammar, vocabulary, and texts* (2008).
127. P. B. Griffin, "Technology and variation in arrow design among the Agta of northeastern Luzon" in Projectile technology. (Springer, 1997), pp. 267-286.
128. M. J. Goodman, P. B. Griffin, A. A. Estioko-Griffin, J. S. Grove, The compatibility of hunting and mothering among the Agta hunter-gatherers of the Philippines. *Sex roles* **12**, 1199-1209 (1985).
129. R. Mayfield, *Central Cagayan Agta Texts* (Linguistic Soc. of the Philippines, Summer Inst. of Linguistics, 1987).

130. J. W. Lobel, O. V. Surbano, Notes from the Field: Remontado (Hatang-Kayi): A Moribund Language of the Philippines. (2019).
131. R. S. Himes, The relationship of Umiray Dumaget to other Philippine languages. *Oceanic Linguistics*, 275-294 (2002).
132. T. N. Headland, Thirty endangered languages in the Philippines. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session* **47**, 1 (2003).
133. L. A. Reid, Arta, another Philippine negrito language. *Oceanic Linguistics* **28**, 47-74 (1989).
134. Y. Kimoto, A Grammar of Arta: A Philippine Negrito Langage. (2017).
135. J. W. Lobel, Manide: an undescribed Philippine language. *Oceanic Linguistics*, 478-510 (2010).
136. R. S. Himes, The Central Luzon group of languages. *Oceanic Linguistics*, 490-537 (2012).
137. V. S. Balilla, J. Anwar-McHenry, M. P. McHenry, R. M. Parkinson, D. T. Banal, Aeta Magbukún of Mariveles: traditional Indigenous forest resource use practices and the sustainable economic development challenge in remote Philippine regions. *Journal of Sustainable Forestry* **31**, 687-709 (2012).
138. W. A. Reed, *Negritos of Zambales* (Bureau of public printing, 1904), vol. 2.
139. J. P. Brosius, The Zambales Negritos: Swidden agriculture and environmental change. *Philippine Quarterly of Culture and Society* **11**, 123-148 (1983).
140. C. B. Bautista (1996) The Mount Pinatubo disaster and the people of Central Luzon. (Seattle: University of Washington Press), pp 151-161.
141. J.-C. Gaillard, Was it a cultural disaster? Aeta resilience following the 1991 Mt Pinatubo eruption. *Philippine quarterly of culture and society* **34**, 376-399 (2006).
142. S. Seitz, Coping strategies in an ethnic minority group: the Aeta of Mount Pinatubo. *Disasters* **22**, 76-90 (1998).
143. E. M. Ragragio, C. N. Zayas, J. J. A. Obico, Useful plants of selected Ayta communities from Porac, Pampanga, twenty years after the eruption of Mt. Pinatubo. *Philipp J Sci* **142**, 169-182 (2013).
144. C. A. Quizon, Between the field and the museum: the Benedict Collection of Bagobo abaca ikat textiles. (1998).
145. C. A. Quizon, Costume, kóstyom, and dress: Formulations of Bagóbo ethnic identity in Southern Mindanao. *Ethnology*, 271-288 (2007).
146. C. E. Tweddell, The identity and distribution of the Mangyan tribes of Mindoro, Philippines. *Anthropological Linguistics*, 189-207 (1970).
147. C. Scholes *et al.*, Genetic diversity and evidence for population admixture in Batak Negritos from Palawan. *American journal of physical anthropology* **146**, 62-72 (2011).
148. J. F. Eder, *On the road to tribal extinction: depopulation, deculturation, and adaptive well-being among the Batak of the Philippines* (Univ of California Press, 1987).
149. R. Rosaldo, Ilongot hunting as story and experience. *The anthropology of experience*, 97-138 (1986).
150. T. S. Oracion (1952) An introduction to the culture of the Magahats of the upper Tayabanan River Valley, Tolong, Negros Oriental, Philippines. (University of Chicago).
151. B. J. Wallace, Plants, pigs, and people: Studying the food web in Pagan Gaddang. *Ethnology* **22**, 27-41 (1983).
152. A. Bagamaspad, Z. Hamada-Pawid, A people's history of Benguet. *Baguio City: Baguio Printing and Publishing Corporation* (1985).

153. T. D. Balangcod, A glimpse of the fire mummies of Kabayan, Benguet, Luzon, Philippines and the role of plants associated with the mummification process. (2018).
154. F. R. R. Blolong, The Ivatan Cultural Adaptation to Typhoons: A Porttrait of a Self-Reliant Community from the Indigenous Development Perspective. *Agham-Tao* **8** (1996).
155. A. Albano, E. van Dongen, S. Takeda, Legal Pluralism, Forest Conservation, and Indigenous Capitalists: The Case of the Kalanguya in Tinoc, the Philippines. *Nature and Culture* **10**, 103-127 (2015).
156. J. M. Garvan, A survey of the material and sociological culture of the Manobo of Eastern Mindanao. *American Anthropologist* **29**, 567-604 (1927).
157. F. L. Jocano, *Sulod Society: A study in the kinship system and social organization of a mountain people of Central Panay* (UP Press, 2008).
158. W. P. Ronquillo, The 1992 Archaeological Reexcavation of the Balobok Rockshelter, Sanga Sanga, Tawi Tawi Province, Philippines--A Preliminary Report (バロボク岩蔭遺跡発掘調査概報). *史料編集室紀要*, p218-179 (1993).
159. C. SATHER, World: The Sama Dilaut of. *Keeping the peace: Conflict resolution and peaceful societies around the world* **8**, 123 (2004).
160. J. F. Warren, *The Sulu Zone, 1768-1898: The dynamics of external trade, slavery, and ethnicity in the transformation of a Southeast Asian maritime state* (NUS Press, 2007).