New Guinea highland wild dogs are the original New Guinea singing dogs

Suriani Surbakti[a,1], Heidi G. Parker[b,1], James K. McIntyre[c], Hendra K. Maury[a], Kylie M. Cairns[d], Meagan Selvig[e], Margaretha Pangau-Adam[a,e], Apolo Safonpo[a], Leonardo Numberi[a], Dirk Y. P. Runtuboi[a], Brian W. Davis[f,2], and Elaine A. Ostrander[b,2]

[a]Department of Biology, Universitas Cenderawasih, Jayapura, Papua 99224, Indonesia; [b]National Human Genome Research Institute, National Institutes of Health, Bethesda MD 20892; [c]New Guinea Highland Wild Dog Foundation, Fernandina Beach, FL 32034; [d]Centre for Ecosystem Science, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia; [e]Department of Conservation Biology, University of Göttingen, 37073 Göttingen Germany; and [f]Department of Veterinary Integrative Biosciences, Texas A&M University College of Veterinary Medicine, College Station, TX 77843

Contributed by Elaine A. Ostrander, July 17, 2020 (sent for review April 16, 2020; reviewed by Klaus-Peter Koepfli, Fernando Racimo, and Robert D. Schnabel)

New Guinea singing dogs (NGSD) are identifiable by their namesake vocalizations, which are unlike any other canid population. Their novel behaviors and potential singular origin during dog domestication make them an attractive, but elusive, subject for evolutionary and conservation study. Although once plentiful on the island of New Guinea (NG), they were presumed to currently exist only in captivity. This conclusion was based on the lack of sightings in the lowlands of the island and the concurrent expansion of European- and Asian-derived dogs. We have analyzed the first nuclear genomes from a canid population discovered during a recent expedition to the highlands of NG. The extreme altitude (>4,000 m) of the highland wild dogs' (HWD) observed range and confirmed vocalizations indicate their potential to be a wild NGSD population. Comparison of single-nucleotide polymorphism genotypes shows strong similarity between HWD and the homogeneous captive NGSD, with the HWD showing significantly higher genetic diversity. Admixture analyses and estimation of shared haplotypes with phylogenetically diverse populations also indicates the HWD is a novel population within the distinct evolutionary lineage of Oceanic canids. Taken together, these data indicate the HWD possesses a distinct potential to aid in the conservation of NGSD both in the wild and under human care.

canine | genetic | conservation | SNP | genome

The New Guinea singing dog (NGSD) is a rare canid living in the New Guinea highlands that, in the wild, is the largest land predator on the island of New Guinea. The dogs produce a characteristic harmonic vocalization (1), described as a "wolf howl with overtones of whale song" (2). For years conservation biologists have concluded that the NGSD may be extinct or nearing extinction in the wild due to loss of habitat and the encroachment of mainland breed dogs and village dogs (1, 3, 4). In this context, nuclear DNA studies of highland wild dogs (HWD) from New Guinea, so named based on their habitat combined with the initial observation of a "wild dog" on Mount Scratchley in 1897 (5), which share striking morphologic similarity to both the NGSD and dingo, are of interest. However, because of their secretive nature and propensity to live at high altitudes distant from villages, HWD are rarely observed. Indeed, prior to 2016 they were photographed only twice (1989 and 2012) (2, 6). Although there was a report of a population of HWD in the HeLa Province of Papua New Guinea (PNG) in 2009, it was based on several assumptions and no animals were ever sighted (7).

The dogs of Oceania, unique populations found in Australia, New Zealand, and the islands nations of Melanesia, Micronesia, and Polynesia, originated from East Asian dog populations, with archaeological evidence supporting their arrival at least 3.5 kya (8). However, the dispersal timing of NGSD to New Guinea remains uncertain due to lack of archaeological evidence found on the island. NGSD were first described following collection of a specimen at an altitude of about 2,100 m in Central Province, PNG, in 1897 (9, 10). Originally classified as a distinct species, *Canis hallstromi*, their taxonomy remains controversial in part due to the availability of only captive specimens for genetic analysis and debate regarding their origin (3, 10, 11). Though genetically similar to the dingo, the NGSD represents a distinct population, as evidenced by both morphology and behavior (12). Resulting from a very small founder population, no more than 200 to 300 captive NGSD remain alive today, largely bred for conservation purposes. Therefore, the population of free-roaming HWD may not only represent a significant evolutionary unit important for conservation and management but possibly an important link to understanding dog domestication (3).

Significance

New Guinea singing dogs (NGSD) are distinctive among the Canidae because of their unique and characteristic vocalization, isolated habitat, and status as a rare representative of wild dogs. Their scarcity, combined with the knowledge that none have been captured or exported since the late 1970s, supports the hypothesis that NGSD are extinct in the wild. We have analyzed the nuclear genome of the first dogs captured from the highlands of Papua in approximately 50 y. We provide DNA-based evidence for an ancestral relationship between highland wild dogs (HWD) and captive NGSD suggesting that the founding population of the NGSD is not, in fact, extinct and that HWD should be resourced for conservation efforts to rebuild this unique canid population.

Author contributions: S.S., H.G.P., J.K.M., B.W.D., and E.A.O. conceptualized the study; H.G.P., J.K.M., B.W.D., and E.A.O. designed research; S.S., H.G.P., J.K.M., and B.W.D. performed investigation; J.K.M., H.K.M., M.S., and M.P.-A. performed field work; S.S., J.K.M., H.K.M., L.N., B.W.D., and E.A.O. provided resources; H.G.P. curated data; H.G.P. analyzed data; S.S., J.K.M., A.S., D.Y.P.R., and E.A.O. supervised the project; S.S., A.S., D.Y.P.R., M.P.-A., and E.A.O. were responsible for project administration; H.G.P., J.K.M., K.M.C., and E.A.O. created visualization; and H.G.P., J.K.M., K.M.C., B.W.D., and E.A.O. wrote the paper with assistance from M.S.

Reviewers: K.K., Smithsonian Conservation Biology Institute; F.R., University of Copenhagen; and R.D.S., University of Missouri.

Competing interest statement: E.A.O. is a coauthor with the dog10K Consortium on a 2019 review article; R.D.S. is a member of the consortium. J.K.M. heads and directs the New Guinea Highland Wild Dog Foundation, which funded the field work for the project. He personally was funded by PT Freeport Indonesia.

[1]S.S. and H.G.P. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: bdavis@cvm.tamu.edu or eostrand@mail.nih.gov.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2007242117/-/DCSupplemental.

First published August 31, 2020.

The need to study this presumably archaic lineage of dog is threefold. First, it is critical to determine if the captive NGSD are from the same population as the HWD or, alternatively, if the NGSD is truly extinct in the wild. Second, the primary aim of the NGSD Conservation Society (https://ngsdconservation.org/home/) is to maintain as much genetic variation within the population as possible. This is extremely challenging given the limited gene pool of the captive NGSD which descend from only eight partially related founders (3, 13). If the existing HWD does indeed represent the predecessor population of the famed NGSD, it is imperative that it be protected. Breeding programs should be established to infuse what is expected to be comparatively robust genetics of the free-breeding HWD into the genetically compromised captive NGSD population, for which inbreeding coefficients have been estimated at >0.50 (14). Finally, truly wild populations of dogs are rare, including the dingo and potentially the HWD, highlighting the need for rapid and well-organized conservation plans.

In 2016, an expedition led by the New Guinea Highland Wild Dog Foundation (NGHWDF) in collaboration with the University of Papua reported the existence of 15 HWD on the western side of the island of New Guinea near the open-cut Grasberg Mine (15). Photographs and fecal samples were collected but were insufficient for nuclear genome analysis. A subsequent 2018 field study led to the collection of blood samples from three putative HWD in their natural environment, as well as demographic, morphologic, and behavioral data. We have utilized those samples to produce a detailed analysis of the HWD nuclear genome, enabling us to determine the relationship between the captive NGSD and modern HWD and to answer the question of whether or not the NGSD is extinct in the wild.

Results

This research study was initiated as an affiliation and collaboration between the NGHWDF and the University of Cenderawasih, in Papua Province, Indonesia. Samples were collected during a 2018 expedition to the base of Puncak Jaya, within the Tembagapura district in the Mimika Regency of Papua, Indonesia. Wild dogs were conditioned to approach and enter cage traps over a period of 2 wk, during which time 18 dogs were observed, 10 of which were new to the study and 8 of which were observed in 2016 (15). Two captured dogs were fitted with Global Positioning System collars and released. All dogs fit the general description of an NGSD (Fig. 1*A*) and the gross body measurements were within or very close to the expected range based on the small number of phenotyped, captive NGSD described previously (*SI Appendix*, Table S1) (10). All three samples were used for the genomic studies described below.

Combining new and publicly available data, we created a dataset of 1,346 dogs from 161 breeds as well as 9 nondog canids, 16 captive NGSD, and 25 wild dingoes (14, 16, 17). Phylogenetic inference of this genotypic dataset places HWD as a monophyletic clade with 100% bootstrap support (Fig. 1*B*). This clade is positioned adjacent to the captive NGSD and dingoes on a single branch comprising all of the dog populations we sampled that originate in Oceania. The branch is within the Asian dog clade, which also includes purebred dogs of East Asian and Arctic ancestry. Ninety-six percent of the branches (25/26) in this clade were positioned with 100% confidence. Dingoes representing multiple wild populations in Australia are not monophyletic.

To determine the genetic composition of HWD compared to all available dogs we assessed possible hybridization using identical-by-descent (IBD) haplotype sharing. Haplotypes were phased across the genome and those inferred to be IBD with a logarithm of the odds (LOD) score >3.0 were summed for every pair of dogs. The distribution of those sums was plotted for each breed dog or wild canid-to-query pair where the query is either HWD, NGSD, or DING (Fig. 1*C*). HWDs display significant haplotype sharing with captive NGSDs and dingoes, but not with any modern breed dogs. However, the background level of sharing with non-Asian breeds is significantly higher in the HWD than in captive NGSD or dingoes (Wilcoxon P value < 2.2e-16) (*SI Appendix*, Fig. S1).

We used the phased haplotypes to assign regions of the HWD genome to the most similar group representing an ancestral population using RFmix (Fig. 2). By assigning each allele to exactly one of these representative populations based on haplotype, we estimate that 72% of the HWD genome is most like captive NGSD or dingoes, the representative groups for Oceanic populations (Fig. 3*A*). This is in stark contrast to the village dogs of New Guinea which share 87% of their genome with breed dogs and only 13% with the Oceanic dogs. An even weaker Oceanic signature is evident in village dogs from Vietnam (11%) and Namibia (<1%). The Australian cattle dog, a historically hybrid breed population, shows 1.5% Oceanic dog heritage using the same metrics. The HWD is predicted to share only 28% of its genome with breed dogs. The excess of European-specific haplotype signatures in all of the dogs from PNG suggests admixture as the cause. Alternatively, it is possible that any ancient variation found in the HWD could masquerade as a signature of recent admixture, as the dog populations representing the Oceanic branch have diverged in ways that would alter their allele frequencies from that of the HWD. The dingoes are as many as 1,200 generations removed from the original dogs of New Guinea, providing adequate time for drift and bottlenecks to allow for loss of diversity. In addition, the captive NGSD encapsulate a severely limited amount of the genomic variation of the original wild population. Thus, the allelic imbalances may, in part, be lost variation in the modern representatives of the ancestral line.

Observed heterozygosity in the HWD is not significantly different from that of breed dogs from any region but is significantly higher than captive NGSD (p_{wilcox} = 0.0021), or a mixed population of dingoes (p_{wilcox} = 0.0006) (*SI Appendix*, Fig. S2). It is also significantly lower than that of village dogs collected in New Guinea (p_{wilcox} = 0.0023). Because the entire captive population of NGSDs is descended from only eight individuals, low genetic variation is expected. However, the dingoes derive from a varied, free-breeding population and would be expected to show much more heterozygosity. The low levels detected are likely an indication of ascertainment bias in the Illumina CanineHD single-nucleotide polymorphism (SNP) set, as the markers were originally chosen because they are polymorphic in European breed dogs and were not widely tested outside of that population. Similarly, the small number of wolves from different parts of the world analyzed with these SNPs show reduced heterozygosity (P = 9e10^{-6} compared to village dogs) and increased inbreeding over what is predicted in a free-breeding population but is not significantly different from that found in the HWD (*SI Appendix*, Table S2).

To perform principal components analyses (PCA) we trimmed the dataset by removing SNPs in high linkage disequilibrium with other SNPs. To help balance the comparisons, we reduced the breed dogs to four groups of 46 to 48 dogs each, representing major geographic regions of breed development; East Asia and the Arctic (ASIA), northern Europe (NORD), western and central Europe (EURO), and the Mediterranean region (MEDI), including parts of southern Europe, northern Africa, and central Asia. Breeds displaying recent admixture across geographic regions were excluded (16). PCA shows distinct clustering of the Oceanic dog groups apart from the breed dogs and wild canines (Fig. 3*B*) along the first PC, which accounts for 32% of the total variation in the dataset ($p_{T\text{-}W}$ = 2.03e-297). The village dogs from New Guinea cluster closer to the breed dogs than the pure Oceanic populations. Analysis of the Oceanic group alone reveals distinct, nonoverlapping groupings of captive NGSD, dingoes, HWD, and village dogs (Fig. 3*C*). The first PC separates the village dogs from the NGSD, HWD, and dingoes (77% of variation,

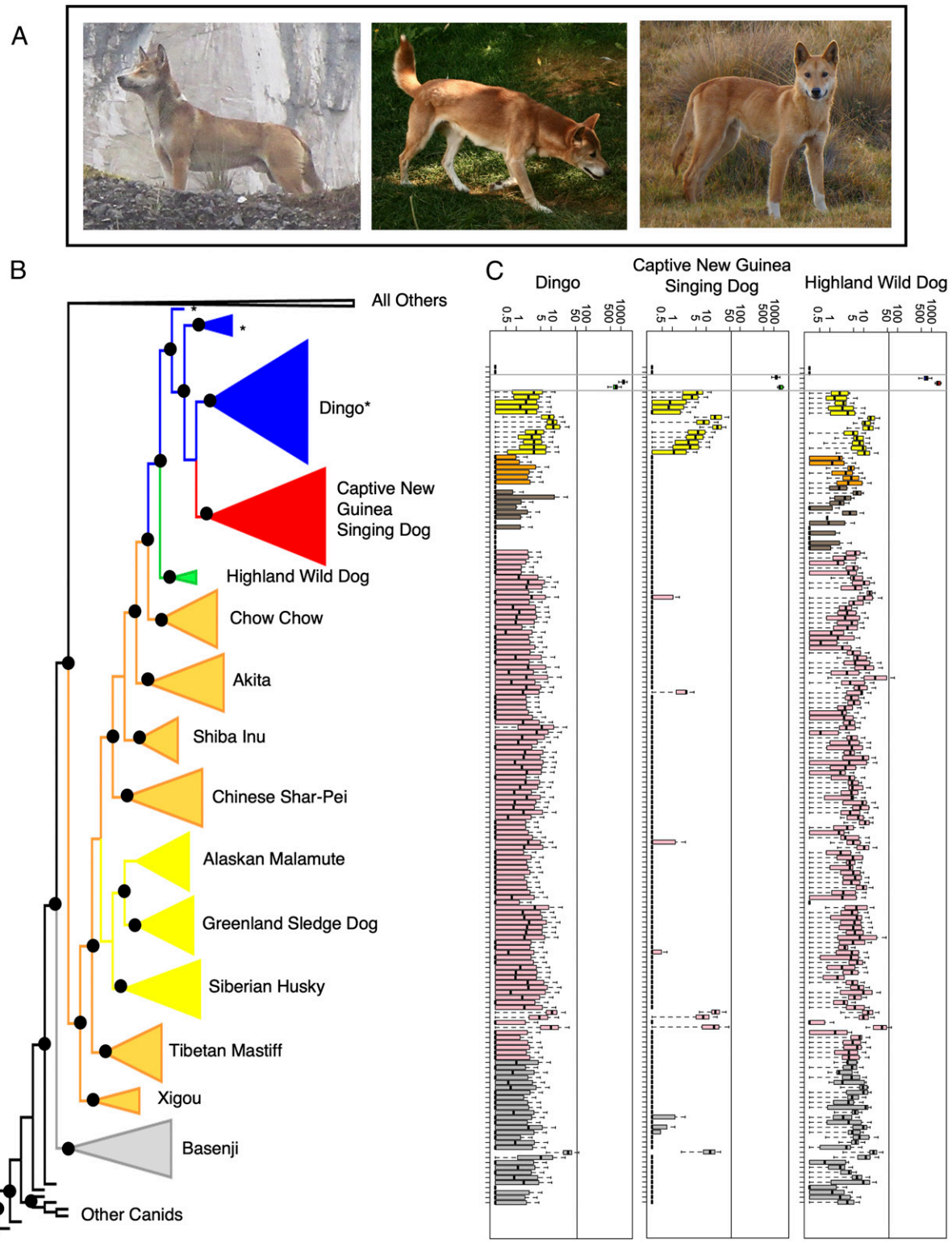

Fig. 1. Oceanic dog populations compared to breed dogs and wild canids. (*A*) Images from left to right: HWD seen during the 2018 expedition, NGSD, dingo. (*B*) Neighbor-joining dendrogram of dogs from 161 breeds places the HWD (green) within the clade of East Asian (gold) and Arctic (yellow) breed dogs on a monophyletic clade with the other Oceanic dog populations (NGSD in red, dingoes in blue on three branches indicated with an asterisk). Branches with 100% bootstrap values are marked with a black dot. (*C*) Box plots indicating the distribution of total haplotype sharing between all pairs of dogs from different populations. Each graph represents one population, indicated above the graph, sharing haplotypes with all others. Each box represents a breed. Haplotype sharing with individuals from the same breed is not included. The three Oceanic breeds are between the two gray lines in order from top to bottom dingo, NGSD, and HWD. The rest of the populations are grouped and colored as follows: black, nondog canids; yellow, Asian/Arctic origin, orange, Nordic origin; brown, Mediterranean origin; pink, Western European origin; gray, mixed or unclassified. The full list of breeds in order can be found in *SI Appendix*, Table S4.

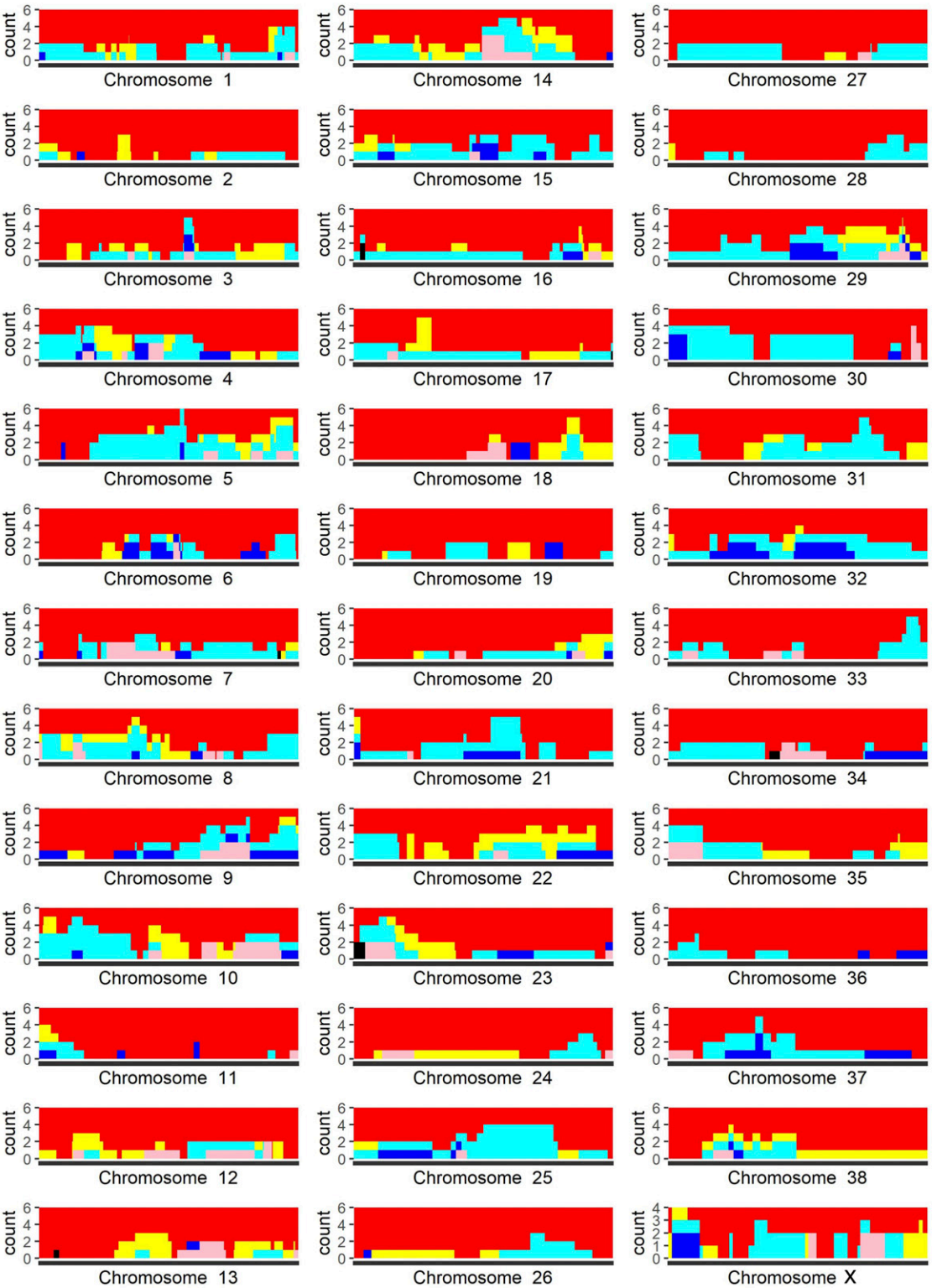

Fig. 2. Assignment of HWD chromosomes to their ancestral geographic origins. Each HWD was phased and each marker was assigned an origin from the modern populations provided based on haplotype sharing with individuals from that population group: red, Oceanic; yellow, Asian; light blue, Western European; dark blue, Mediterranean; pink, Nordic; black, wolf. The graph displays the number of alleles among the six available that were assigned to each representative ancestral population at each site and is ordered by representative ancestral population, not by individual HWD. Only four chromosomes are represented on the X, as two of the HWD are male.

$p_{T\text{-}W}$ = 7.70e-15). The second PC separates the NGSD from the dingoes (23% of variation, $p_{T\text{-}W}$ = 5.83e-15), with the HWD trending toward the NGSD.

Using allele frequencies from the independent genotypes, we built maximum likelihood trees of the Oceanic populations and geographic breed groups to assess the likelihood of admixture

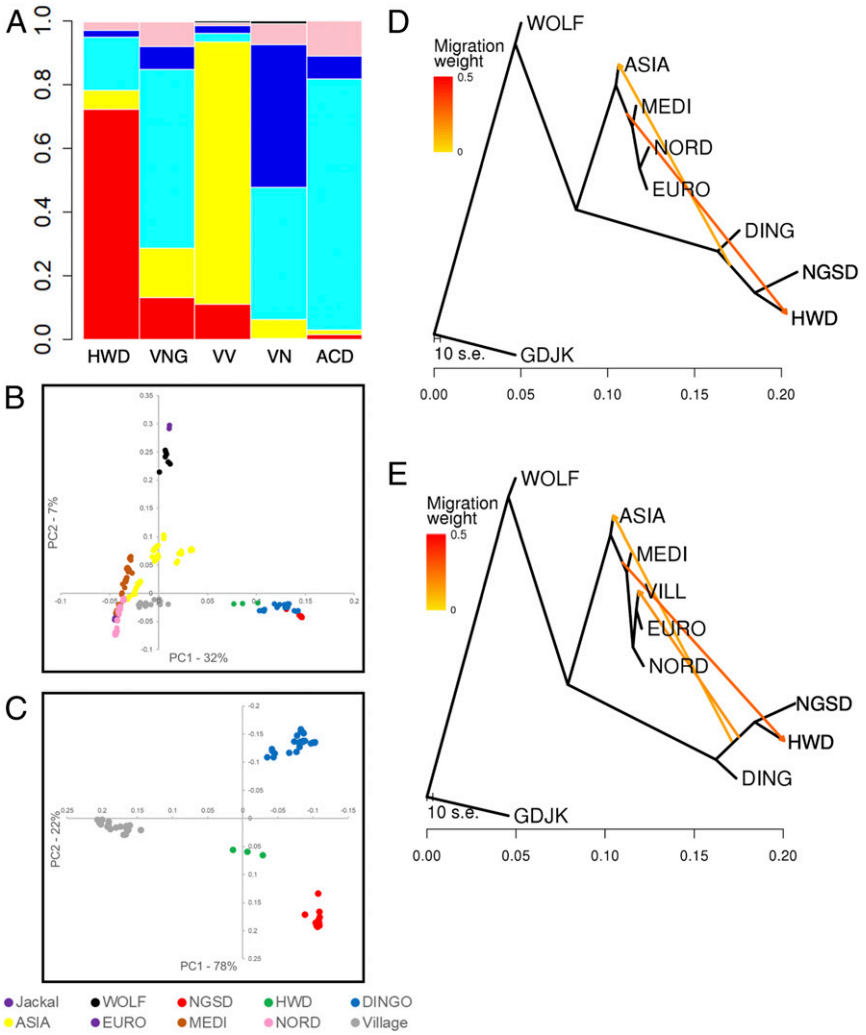

Fig. 3. Assessment of admixture in HWD. (*A*) Bar plot of admixture in HWD compared to village dogs and a putatively admixed breed. Admixture is based on the assignment of SNPs to putative ancestral chromosomes by haplotype sharing with modern representatives and averaged across all individuals of the population: 3 HWD and 10 each of village dogs from New Guinea (VNG), Vietnam (VV), and Namibia (VN) and Australian Cattle dogs (ACD). Red, Oceanic; yellow, Asian; light blue, European; dark blue, Mediterranean; pink, Nordic; black, wolf. (*B*) PCA of 251 dogs and nine wild canids shows division between Oceanic and other dogs on PC1 and wild canids and domestic dogs on PC2. Geographic origin of the dogs is indicated by color. (*C*) PCA analysis of only Oceanic dogs reveals separation of pure Oceanic and hybrid village dogs along PC1 and separation of New Guinea dogs and Australian dingoes along PC2. (*D*) Maximum likelihood trees of geographic-based breed populations with migration. Migration is indicated by the arrows shaded from yellow to red according to their weight. Golden jackal was used to root the tree. (*E*) Maximum likelihood tree with New Guinea village dogs added and migration allowed.

using the program Treemix (18). This method predicted the migration of an ancestor of all European breed dogs into the HWD, as well as compensatory migration from the New Guinea dogs to breed dogs (Fig. 3*D*). These same migration events are predicted if New Guinea village dogs are added to the tree, as well as a third migration of New Guinea naturalized dogs to the village dogs (Fig. 3*E*).

We tested these predictions further by calculating f3 which, when negative, provides evidence of admixture from two source populations into a third. We observe such evidence in HWD when the source populations include captive NGSD and any of the four breed groups (Table 1).

To further confirm the source and degree of possible admixture into the HWD we calculated D statistics and F4 ratios from four-branch tree arrangements including the HWD using either golden jackal or wolf as the outgroup. We find significant evidence for admixture ($Z > 3.0$) between all breed groups and the HWD (*SI Appendix*, Table S3). F4 ratios of tree arrangements swapping any of the breed groups for the HWDs predicted admixture ranging from 15 to 50%; however, the z-scores were insignificant for all combinations. We next used qpWave to examine the number of likely populations contributing to the HWD and determined that a two-source model best fits the data provided. That number did not change, regardless of which breed groups or Oceanic populations were included in the analysis. We ran qpAdm and obtained a prediction of 29.6% breed dog contribution to the HWD. This is very similar to the 28.8% migration from the breed clade predicted by Treemix and the 28% non-oceanic contribution predicted from haplotype sharing.

Discussion

The 2018 expedition to Indonesian New Guinea was motivated by analysis of specimens collected in 2016 (15). Specifically, analysis of mitochondrial DNA isolated from 9 of 24 presumed HWD scat samples revealed the A29 haplotype, rather than the A79 haplotype. The A79 haplotype is unique to NGSD, while the A29 haplotype is observed in dingoes, NGSD, some Asian and Arctic breed dogs (19, 20), and village dogs. Thus, the results are

Table 1. Detection of possible admixture between Oceanic populations and breed dogs in the HWD using the F3 statistic

Source 1: Breed	Source 2: Oceanic	Target	Z-score
ASIA	NGSD	HWD	−4.708
EURO	NGSD	HWD	−15.289
MEDI	NGSD	HWD	−13.639
NORD	NGSD	HWD	−13.871
ASIA	Dingo	HWD	n.s.
EURO	Dingo	HWD	n.s.
MEDI	Dingo	HWD	n.s.
NORD	Dingo	HWD	n.s.
All	All	Dingo	n.s.
All	All	NGSD	n.s.

"All" indicates all possible combinations of populations that fit the column descriptions. n.s = not significant (Z > 0).

not specific to the NGSD and can be generalized to dogs from East Asia and the Oceanic region (13, 21–23).

Three contemporary HWDs provided a look into the nuclear genome of this elusive, potential originator population that shares striking morphological similarity to both the NGSD and dingo. Distance measures place the HWD split basal to the dingo and NGSD division in a monophyletic clade with 100% bootstrap support. In addition, maximum likelihood trees, allowing for a small amount of migration, place the HWD, NGSD, and dingo on a branch excluding all breed dogs and diverging prior to the formation of breed dogs from four diverse geographic origins. These results suggest a distinct evolutionary lineage for the Oceanic populations. This is confirmed by the extensive haplotype sharing within dogs originating in Oceania to the exclusion of the remainder of canids. This shared genomic ancestry, along with estimates of genetic diversity, highlight the lack of an ancestral contribution from any lineage but Oceania, and specifically HWD, to the current captive NGSD. The lack of diversity renders it difficult to estimate the specific genetic origin of NGSD; however, the extant population of HWD is most similar of all populations thus far studied.

Using F3 statistics, admixture is detected in the HWD only when NGSD are part of the analysis. This suggests that the captive NGSD are derived solely from the HWD and not from another Oceanic population, that the HWD are not a new introduction of dingo to the islands, and that the presumptive admixture may represent early variation missing from the captive NGSD. The pattern of allele sharing is not found when HWD are analyzed with dingoes and breed dogs, suggesting that the derived alleles shared with breed dogs are similarly present in the dingo, and are missing only in the very restricted NGSD genome, likely as a result of reduced heterogeneity rather than recent migration.

As our knowledge of the genetic history of the dingo advances (14), it becomes increasingly necessary to understand the underlying contribution of phenotypically similar and closely related canids to the Oceanic dogs of the Canidae. Phylogenetic and haplotype analyses provide compelling evidence that the HWD studied here represent extant animals arising from the same historic stock as the captive NGSD. Analysis of admixture suggests that there is evidence for minimal introgression in the HWD, likely from a population ancestral to modern breeds. The fact that the New Guinea population of HWD possesses much higher genetic diversity than their largely inbred, captive counterparts argues that conservation efforts to revitalize the NGSD would benefit from the utilization of HWD in combination with the captive NGSD. Indeed, the inclusion of even a small number of individuals from a natural population can have a significant effect on the genetic diversity of a captive breeding population (24).

The size and extent of the HWD population, however, remains to be documented. Analyses in this work encompassed only three individuals. Sampling of dogs from more remote regions along the central range such as Puncak Mandala, Puncak Trikora, and equally remote areas on the PNG side of the Island of New Guinea would provide insight as to both the numbers of HWD in their natural habitat and their genetic diversity. Conservation efforts will benefit most from inclusion of the greatest number of specimens that best represent the original dogs, with the least amount of influence from outside sources, making it imperative that these studies be continued. In addition, studies of small pockets of wild living dogs will enable the search for the genetic variation that is required for free living, but irrelevant in a captive environment, and vice versa. There is urgency in this exercise, as each new generation of captive NGSD increases the risk for fixation of deleterious alleles, thus compromising subsequent breeding programs (25). This study provides insight into the genetic makeup of an isolated canid population and thus identifies a readily available genetic reservoir for the NGSD, previously thought to be extinct in the wild.

Materials and Methods

Sample Ascertainment. This research study was sponsored by PT Freeport Indonesia and resulted from an established collaboration between the NGHWDF, under the sponsorship of PT Freeport Indonesia, and the University of Cenderawasih, in Papua Province, Indonesia. The Foundation has participated in two research trips, the first in 2016 (15) and the second in 2018, both to the base of Puncak Jaya, which is within the Tembagapura district in the Mimika Regency of Papua, Indonesia. Samples utilized for this study were collected in 2018 from locations at the periphery of the Grasberg Mine using a novel capture approach for HWD. Wild dogs were conditioned to approach and enter cage traps over a period of 2 wk following enticement with crushed coyote gland lure and skunk essence. Traps were baited inside and out of cages. Auditory North American coyote calls were played to attract wild dogs and elicit return howls. Trail cameras reviewed the progress. After 2 wk the traps were set and the capture campaign was initiated. Two wild dogs were captured, immobilized, and underwent extensive examinations including measurements (*SI Appendix*, Table S1) and collection of biologic samples. The dataset included DNA and information from those dogs as well as a third recently deceased dog that appeared to be the victim of a car collision.

DNA Preparation. Three HWD samples were provided for processing. Two blood draws from live dogs were stored in acid-citrate-dextrose anticoagulant. DNA was isolated using described methods (26). The third sample, skin from the deceased dog, was stored in RNAlater (Qiagen) for transport. DNA was isolated using the DNeasy Blood and Tissue micro kit (Qiagen) and standard protocols following homogenization. All protocols completed at NIH and Texas A&M University followed Institutional Animal Care and Use–approved protocols.

SNP Genotypes and Phylogeny. DNA samples from the HWD as well as 10 captive NGSD were genotyped on the Illumina canine HD SNP chip (Illumina) using standard protocols. NGSD blood samples were collected from sanctuaries and private owners who had rescued them from zoos and exotic animal breeders in seven US states and Canada. They were chosen to include as much diversity from the captive population as possible. Genotype calls were made using Genome Studio v2.0.4 with genotyping module v2.0.4 (Illumina). Data from captive NGSD and HWD (Gene Expression Omnibus [GEO] accession no. GSE143824) (27) were analyzed with 1,346 domestic dogs representing 161 breeds, 7 wolves, and 2 golden jackals genotyped at 150,112 SNPs spanning all autosomes and the X chromosome (GEO accession nos. GSE90441, GSE83160, GSE70454, and GSE96736) (16, 28–31). In addition, publicly available SNP chip data from 24 dingoes (Dryad submission DOI:10.5061/dryad.sq8d0) (14, 32) was incorporated into the analysis, as was data from an additional five NGSD and 60 village dogs collected in New Guinea, Vietnam, and Namibia (Dryad submission DOI:10.5061/dryad.v9t5h) (17, 33).

Linkage disequilibrium pruning was carried out with an r^2 threshold > 0.5, as assessed by PLINK v1.9 using the –indep-pairwise option (34, 35). This reduced the dataset to 107,125 variants for subsequent principal component and admixture calculations.

To create a balanced dataset for PC and admixture analyses, purebred dogs were grouped according to ancestral geographic region of origin based on published phylogenetic results (16). Four groups were identified: Nordic, 46 dogs from breeds originating in Scandinavian countries; Mediterranean, 47 dogs from breeds originating in southern Europe, central Asia, and northern Africa; European, 46 dogs from breeds originating in Western European countries and the United Kingdom,; and Asian, 48 from breeds originating in East Asia and Arctic regions. Breeds that showed recent haplotype sharing across geographic groups were excluded. Each group was then trimmed to simulate the size of the smallest group, retaining as many different breeds as possible while randomly choosing representatives of the breeds. These were combined with the Oceanic populations, 25 dingoes, three HWD, and 16 NGSD, for the final balanced dataset.

Phylogenetic Analysis. A pairwise identity-by-state distance matrix was computed using PLINK v1.9 and the –distance 1-ibs command (34, 35). Bootstrapped distance matrices were created by randomly resampling 150,112 markers with replacement 100 times and input into PHYLIP using neighbor and consense to construct neighbor-joining dendrograms (36). Dendrograms were visualized using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

PCA. PCs were calculated for the balanced dataset along with the wolves and jackals and village dogs from New Guinea. A second PCA was performed including only dogs from the Oceanic region: HWD, captive NGSD, Dingo, and New Guinea village dogs. The eigensoft package was used to calculate PCs with smartpca (37, 38). Twstats was used to determine the significant components using the Tracy–Widom distribution. Proportion of variance was calculated from only the significant PCs ($P_{T\text{-}W} < 0.05$).

Haplotype Sharing. Data were phased and haplotypes determined using the program Beagle v4.1 (39) with sliding windows of 1,000 SNPs and a 50-SNP overlap. Haplotypes that are identical-by-state and that are predicted to be IBD, based on excess length and sharing (40–42) with a LOD score greater than 3.0, were calculated between every pair of dogs using the ibd option of Beagle (40). The inferred IBD haplotypes for each pair were summed across the genome to assess the recent breed ancestry of the HWD, captive NGSD, and dingoes.

Chromosomal regions were mapped to the most likely geographic ancestor population by RFmix (43). Phased genotypes from the balanced dataset were used after collapsing haplotypes that were separated by only one SNP and reassigning phase to the original Beagle output as described in Browning et al. (44). Total estimated ancestry was calculated by averaging the ancestral assignments of all 150,000 markers.

Admixture and Migration. The balanced dataset described above was used to assess the presence and extent of admixture in the HWD using Treemix (18) and AdmixTools (45). Wild canids were included in the dataset with the golden jackal serving as the outgroup for the maximum likelihood trees. Allele frequencies for each population were calculated in PLINK. Treemix was run with the bootstrap option and k = 1,000, allowing 0 to 10 migration events and repeated five times at each value. The optimal number of migrations was reached when the average variance came within one SD of 99.8% as calculated in the program OptM (https://rdrr.io/cran/OptM/) and by the Evanno method (46), the latter of which is calculated from the change in likelihood (Δm) with each additional migration event (*SI Appendix*, Figs. S3 and S4). This analysis was repeated including village dogs collected in New Guinea (17). Maximum likelihood trees with predicted migration events and residuals were viewed using plotting_functions.R, provided with the Treemix package.

F3, D statistics, and F4 ratios were calculated from three and four branch tree arrangements to evaluate the ancestral components of HWD and geographical groups of breed dogs. F3 was calculated using threepop from Treemix (18). D-statistics and F4 ratios were calculated using Admixtools with golden jackal as the outgroup. Because all of the breed groups can be assumed to have some shared ancestry with one another, wolf was included as the unrelated branch in the F4 ratio calculations (45).

Also using the R package admixr (47) to run AdmixTools we calculated the most likely number of source populations contributing to the HWD and the extent of estimated hybridization. These were calculated using the qpWave and qpAdm commands. Because no one breed group was determined to be the source of admixture, all breeds were combined to obtain the percent hybridization. Therefore, when running qpAdm the target was set as HWD, the sources were Oceanic and breed dogs, and the outgroups were the wolves and jackals.

Data Availability. Raw data files for SNP genotype arrays are deposited in the NCBI Gene Expression Omnibus under accession no. GSE143824 (27). Previously published data used in this paper can be found under GEO accession nos. GSE90441 (28), GSE83160 (29), GSE70454 (30), and GSE96736 (31). Dryad submissions: DOI:10.5061/dryad.sq8d0 (32) and DOI:10.5061/dryad.v9t5h (33).

ACKNOWLEDGMENTS. We gratefully acknowledge the University of Cenderawasih for their collaboration and logistical support in Indonesia and Andrew Hogan and Alex Harris at the National Human Genome Research Institute for sample preparation and genotyping. We thank Gayle Person for putting together a diverse collection of captive NGSD for our analysis. We acknowledge Michelle Brown for providing the photograph of the dingo. Expedition and sample acquisition were supported by NGHWDF. E.A.O. and H.G.P. were supported by the Intramural Program of the National Human Genome Research Institute of the National Institutes of Health. J.K.M. and the New Guinea Highland Wild Dog Project are funded and supported by PT Freeport Indonesia. K.M.C. is supported by a grant from the Australian Dingo Foundation. B.W.D. is supported by the Veterinary Biosciences Program at Texas A&M College of Veterinary Medicine.

1. I. L. Brisbin Jr., R. P. Coppinger, M. H. Feinstein, S. N. Austad, J. J. Mayer, The new Guinea singing dog: Taxonomy, captive studies and conservation priorities. *Sci. New Guinea* **20**, 27–38 (1994).
2. B. Crew, First photo of rare, wild New Guinea singing dog in 23 years. *Scientific American Blog*. https://blogs.scientificamerican.com/running-ponies/first-photo-of-rare-wild-new-guinea-singing-dog-in-23-years/. Accessed 10 December 2012.
3. J. Koler-Matznick, B. C. Yates, S. Bulmer, I. L. Brisbin Jr., The New Guinea singing dog: Its status and scientific importance. *Aust. Mammal.* **29**, 47–56 (2007).
4. ; IUCN Species survival commission and Canid specialist group, "Canids: Foxes, wolves, jackals and dogs" in *Status Survey and Conservation Action Plan Series*, C. Sillero-Zubiri, M. Hoffman, D. W. MacDonald, Eds. (IUCN, Gland, Switzerland and Cambridge, UK, 2004).
5. C. W. De Vis, "A wild dog from British New Guinea" in *Annals of the Queensland Museum*, (Government Printer 1891-1911, Brisbane, Australia, 1911), Vol. vol. 10, pp. 19–20.
6. T. Flannery, *Mammals of New Guinea*, (Cornell University Press, Ithaca, NY, 1995).
7. K. P. Alpin, E. Kale, "The non-volant mammal fauna of the Muller Range, Papua New Guinea" in *Rapid Biological Assessments of the Nakanai Mountains and the Upper Strickland Basin: Surveying the Biodiversity of Papua New Guinea's Sublime Karst Environments*, S. R. Richards, B. G. Gamui, Eds. (Conservation International, Arlington, VA, 2013), pp. 211–221.
8. P. Milham, P. Thompson, Relative antiquity of human occupation and extinct fauna at Madura Cave, Southeastern Western Australia. *Mankind* **10**, 175–180 (1976).
9. E. Troughton, The early history and relationships of the New Guinea Highland dog (*Canis hallstromi*). *Proc. Linn. Soc. N. S. W.* **96**, 93–98 (1971).
10. J. Koler-Matznick, I. L. Brisbin Jr., M. Feinstein, S. Bulmer, An updated description of the New Guinea singing dog (*Canis hallstromi*, Troughton 1957). *J. Zool.* **261**, 109–118 (2003).
11. P. D. Dwyer, M. Minnegal, Wild dogs and village dogs in new Guinea: Were they different? *Aust. Mammal.* **38**, 1–11 (2016).
12. K. Greig, R. Walter, E. A. Matisoo-Smith, "Dogs and people in Southeast Asia and the Pacific" in *The Routledge Handbook of Bioarchaeology in Southeast Asia and the Pacific Islands*, M. Oxenham, H. R. Buckley, Eds. (Routledge, London, 2015), chap. 21, pp. 462–482.
13. B. N. Sacks *et al*., Y chromosome analysis of dingoes and southeast asian village dogs suggests a neolithic continental expansion from Southeast Asia followed by multiple Austronesian dispersals. *Mol. Biol. Evol.* **30**, 1103–1118 (2013).
14. K. M. Cairns, L. M. Shannon, J. Koler-Matznick, J. W. O. Ballard, A. R. Boyko, Elucidating biogeographical patterns in Australian native canids using genome wide SNPs. *PLoS One* **13**, e0198754 (2018).
15. J. K. McIntyre, L. Wolf, B. N. Sacks, J. Koibur, I. L. Brisbin Jr., A population of free-living highland wild dogs in Indonesian Papua. *Aust. Mammal.* **42**, 160–166 (2019).
16. H. G. Parker *et al*., Genomic analyses reveal the influence of geographic origin, migration, and hybridization on modern dog breed development. *Cell Rep.* **19**, 697–708 (2017).
17. L. M. Shannon *et al*., Genetic structure in village dogs reveals a Central Asian domestication origin. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13639–13644 (2015).
18. J. K. Pickrell, J. K. Pritchard, Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **8**, e1002967 (2012).
19. P. Savolainen, Y. P. Zhang, J. Luo, J. Lundeberg, T. Leitner, Genetic evidence for an East Asian origin of domestic dogs. *Science* **298**, 1610–1613 (2002).
20. J. F. Pang *et al*., mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Mol. Biol. Evol.* **26**, 2849–2864 (2009).
21. P. Savolainen, T. Leitner, A. N. Wilton, E. Matisoo-Smith, J. Lundeberg, A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12387–12390 (2004).

22. M. C. Oskarsson *et al*., Mitochondrial DNA data indicate an introduction through Mainland Southeast Asia for Australian dingoes and Polynesian domestic dogs. *Proc. Biol. Sci.* **279**, 967–974 (2012).
23. K. M. Cairns, S. K. Brown, B. N. Sacks, J. W. O. Ballard, Conservation implications for dingoes from the maternal and paternal genome: Multiple populations, dog introgression, and demography. *Ecol. Evol.* **7**, 9787–9807 (2017).
24. W. E. Johnson *et al*., Genetic restoration of the Florida panther. *Science* **329**, 1641–1645 (2010).
25. J. R. Brandt *et al*., Genetic structure and diversity among historic and modern populations of the Sumatran rhinoceros (*Dicerorhinus sumatrensis*). *J. Hered.* **109**, 553–565 (2018).
26. G. I. Bell, J. H. Karam, W. J. Rutter, Polymorphic DNA region adjacent to the 5′ end of the human insulin gene. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5759–5763 (1981).
27. S. Surbakti *et al*, SNP genotyping of Highland Wild dogs, New Guinea Singing dogs, and a dingo. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE143824. Deposited 16 January 2020.
28. D. L. Dreger *et al*, Whole genome sequence, SNP chips and pedigree structure. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE90441. Deposited 22 November 2016.
29. D. L. Dreger *et al*, Studies of canine breed development on the island of Sardinia recapitulate genomic features of human population isolates. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE83160. Deposited 9 June 2016.
30. B. Decker *et al*, Comparison against 186 canid whole genome sequences reveals survival strategies of an ancient clonally transmissible canine tumor. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE70454. Deposited 1 July 2015.
31. H. G. Parker *et al*, Genomic analyses reveal the influence of geographic origin, migration and hybridization on modern dog breed development. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE96736. Deposited 16 March 2017.
32. K. M. Cairns, L. M. Shannon, J. Koler-Matznick, J. W. O. Ballard, A. R. Boyko, Elucidating biogeographical patterns in Australian native canids using genome wide SNPs. Dryad. https://doi.org/10.5061/dryad.sq8d0. Accessed 31 January 2019.
33. L. M. Shannon *et al*, Genetic structure in village dogs reveals a Central Asian domestication origin. Dryad. https://doi.org/10.5061/dryad.v9t5h. Accessed 16 January 2019.
34. C. C. Chang *et al*., Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
35. 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).
36. J. Felsenstein, PHYLIP–Phylogeny inference package (Version 3.2). *Cladistics* **5**, 164–166 (1989).
37. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
38. A. L. Price *et al*., Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
39. B. L. Browning, S. R. Browning, Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics* **194**, 459–471 (2013).
40. S. R. Browning, B. L. Browning, High-resolution detection of identity by descent in unrelated individuals. *Am. J. Hum. Genet.* **86**, 526–539 (2010).
41. S. R. Browning, B. L. Browning, Identity by descent between distant relatives: Detection and applications. *Annu. Rev. Genet.* **46**, 617–633 (2012).
42. A. Gusev *et al*., Whole population, genome-wide mapping of hidden relatedness. *Genome Res.* **19**, 318–326 (2009).
43. 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).
44. S. R. Browning *et al*., Ancestry-specific recent effective population size in the Americas. *PLoS Genet.* **14**, e1007385 (2018).
45. N. Patterson *et al*., Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
46. G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).
47. M. Petr, B. Vernot, J. Kelso, admixr-R package for reproducible analyses using ADMIXTOOLS. *Bioinformatics* **35**, 3194–3195 (2019).