Supplementary Materials for

On the origin of mongrels: Evolutionary history of free-breeding dogs in Eurasia

Małgorzata Pilot, Tadeusz Malewski, Andre E. Moura, Tomasz Grzybowski, Kamil Oleński, Anna Ruść, Stanisław Kamiński, Fernanda Fadel, Daniel Mills, Abdulaziz N. Alagaili, Osama B. Mohammed, Grzegorz Kłys, Innokentiy M. Okhlopkov, Ewa Suchecka & Wiesław Bogdanowicz

This file includes:

- Supplementary Text
- Supplementary Figures 1-20
- Supplementary Tables 1-5

Screening the dataset for related individuals

We used the software KINGROUP2 (Konovalov *et al*. 2004) to identify groups of kin (e.g. full-siblings and half-siblings) and CERVUS 3.0 (Marshall *et al*. 1998) to identify parent-offspring pairs. These analyses were carried out independently for two datasets consisting of the same individuals, but differing in the set of loci used. The number of loci for this analysis had to be reduced to less than 1,000 due to the limits set for KINGROUP input files. The first dataset was constructed by selecting highly variable loci (having minor allele frequencies between 0.496 and 0.500) without missing data, which resulted in 903 SNPs. The second dataset was constructed by selecting loci without missing data, pruning for linkage disequilibrium with a criterion of r^2=0, and then selecting highly variable loci (with minor allele frequency between 0.492 and 0.500; the lower frequency threshold being established empirically in order to achieve a required number of SNPs). This resulted in 993 SNPs.

The results of the CERVUS analysis were identical between the two datasets, and identified 23 groups of closely related individuals (parent-offspring and/or sibling groups). The results of the KINGROUP analysis were also consistent, identifying 21of the kin groups from CERVUS as related at a full-sibling level, and two groups as related at a half-sibling level (with one discrepancy between the results from the two datasets). The closely related individuals were removed from the dataset, with only one individual per kin group retained. This way we obtained a dataset of 200 unrelated individuals, which was used in all the subsequent analyses.

Analysis of genetic differentiation in Eurasian FBDs

Methods

We analysed population genetic structure using the LD-pruned FBD dataset. We used the Bayesian clustering methods with no prior population information as implemented in

ADMIXTURE (Alexander *et al*. 2009) and STRUCTURE (Pritchard *et al*. 2000). The maximum number of genetic clusters (K) assumed in these analyses was set to 14, equal to the number of the sampling sites. In ADMIXTURE, we used the termination criterion that stops the run when the log-likelihood increases by less than $\varepsilon = 10^{-4}$ between iterations. To identify the optimal value of K, we applied a cross-validation method, where 10 datasets were created by removing 10% of the genotypes at random, analysed as described above, and variation in the results was compared between different K values. The K value that exhibited the lowest cross-validation error was considered as the optimal K.

STRUCTURE was run assuming K from 1 to 14, with 100,000 MCMC iterations preceded by 20,000 burn-in iterations. We used the admixture model with correlated allele frequencies and ran three replicates for each K value. Selection of optimal K was carried out taking into account likelihood and Evanno *et al*. (2005) ΔK values.

In addition, we carried out a spatially explicit analysis of genetic structure using the software GENELAND (Guillot *et al*. 2008). The spatially explicit analyses are more computationally demanding, and we thus reduced the number of SNPs by carrying out a more stringent pruning for LD using the criterion of $r^2<0.1$, which resulted in a set of 43,024 SNPs. The analysis was run with four independent replicates for 10,000 MCMC iterations with 1,000 burn-in iterations, for K values between 1 and 14. The model assuming correlated allele frequencies among populations was used.

We assessed the presence of an isolation-by-distance pattern in the FBDs across Eurasia by testing for a correlation between $F_{ST}/(1-F_{ST})$ and the natural logarithm of geographic distance using a simple (univariate) Mantel test implemented in GENALEX 6.5 (Peakall & Smouse 2012). We also used GENALEX to carry out a spatial autocorrelation analysis based on pair-wise F_{ST} values between 14 sampling sites.

Results

ADMIXTURE indicated K=2 as the most likely number of clusters, but most individuals were admixed between these two clusters without clear biological interpretation. Most European FBDs had higher assignment probabilities to one cluster, while most Middle Eastern and East Asian FBDs had higher assignment to the second cluster. Most Central/West Asian individuals had nearly equal assignment probabilities to each cluster (Supplementary Fig. 1A). Individual assignment K=3 provided a clearer biological scenario, with one cluster defining the populations from the Middle East (Saudi Arabia and Iraq), the second cluster defining the populations from East Asia (China and Thailand), while most individuals from Europe, Central/West Asia and East Russia had an admixed ancestry between all three genetic clusters. At K=4 a clearer distinction between European and Central/West Asian populations was obtained. Larger values of K did not reveal further geographically-defined clusters (Supplementary Fig. 1A).

Genetic clustering patterns inferred in STRUCTURE were consistent with those described for ADMIXTURE. Both likelihood and Evanno's ΔK values indicated K=5 as the optimum number of clusters, but the fifth genetic cluster did not have any clear biological interpretation and its addition caused little change in the clustering patterns relative to K=4 (Supplementary Fig. 1B).

GENELAND analysis resulted in considerable inconsistencies between replicates, with K value indicated as optimal varying between 8 and 12. The geographic distribution of the inferred genetic clusters was largely inconsistent between different runs, and did not show any clear spatial pattern, except for populations from Saudi Arabia and Iraq being consistently grouped together.

The Principal Component Analysis (PCA) based on individual pair-wise distances showed little clustering, reflecting a large degree of similarity among FBD populations. The

PC1 differentiated East Asian and Middle Eastern dogs from European and Central/West Asian dogs (Supplementary Fig. 2A), while PC2 did not correspond to any clear geographic division. PC3 differentiated East Asia from the Middle East, with other regions having intermediate positions (Supplementary Fig. 2B). FBDs from Europe and Central/West Asia were not clearly differentiated. Similarly, PCA based on population-level pair-wise F_{ST} between sampling sites grouped together Europe, Central/West Asia and East Russia into a single cluster. This cluster was differentiated from the Middle Eastern populations at PC1 and from East Asian population at PC2 (Supplementary Fig. 2C).

Weak population differentiation was also supported by a tree of inter-individual identity-by-state (IBS) distances (Supplementary Fig. 3), where dogs from different sampling locations and broader regions of Eurasia did not form geographically distinct clades. Individuals from the Middle East were an exception, as they grouped into a single sub-clade, although within a larger clade containing dogs from multiple geographic locations. This pattern suggests that the Middle Eastern dogs have been relatively isolated from other dog populations in West Asia, but only recently.

Finally, we found a weak, but significant correlation between genetic differentiation ($F_{ST}/(1-F_{ST})$ coefficient) among the 14 sampling sites and their geographic distance (R^2=0.056, *P*=0.041; Supplementary Fig. 4A). The spatial autocorrelation analysis showed a trend of declining autocorrelation for distance classes between 1,000 and 4,000 km, and no dependence of spatial autocorrelation on geographic distance for larger distance classes (5,000-9,000 km; Supplementary Fig. 4B). The heterogeneity test was significant at the whole correlogram level (*P*=0.002).

Estimation of differences in past effective population size between local FBD populations

Methods

We estimated effective population sizes (N_E) based on the extent of LD, measured as r^2 coefficient, using the method described in Tenesa *et al*. (2007). We estimated N_E using the equation $E(r^2) = 1/(\alpha+4N_E c) + 1/n$, where c is the genetic distance between loci in Morgans, and 1/n is the adjustment for small sample size (Hayes *et al*. 2003). The parameter $\alpha = 1$ is assumed in the absence of mutations, and $\alpha = 2$ if mutations are accounted for. Here the assumption of $\alpha = 2$ is more accurate, but because our goal was not to accurately reconstruct the demographic history, but to assess whether any population has higher historical N_E values compared to other populations, we used both $\alpha = 1$ and $\alpha = 2$ to check whether this may affect our conclusions.

Average r^2 values were calculated within 20 distance classes between 2.5 kb and 1 Mb. Assuming that 1Mb = 1 cMorgan, they correspond to 0.0025–1 cMorgan. Each genetic distance (c) and its associated average r^2 value corresponds to a N_E estimate in a particular time in the past (measured in generations), estimated as $t \approx 1/(2c)$ (Hayes *et al*. 2003). The distance classes considered here provide N_E estimates in a period between 50 and 20,000 generations ago, or between 150 and 60,000 years ago, if we assume a 3-year generation time, estimated for grey wolves (Mech & Seal 1987). Although the generation time of modern dogs may be shorter (Freedman *et al*. 2014), it is unclear when the generation turnover became accelerated. However, the generation time assumed does not affect the relative N_E comparisons between different regions. Because the timing of the demographic changes estimated using this approach is not precise (as it assumes a linear relation between the recombination distance and time, which does not hold for all demographic scenarios; Hayes *et al*. 2003), we did not use this approach to estimate the absolute timing of demographic events

in the dog evolutionary history, but only to identify differences in the demographic history of different regional populations.

Results

We inferred temporal variation in effective population sizes (N_E) in FBD populations from LD patterns. This method of inferring past population dynamics is less precise than methods based on genome re-sequencing data (e.g. see Freedman *et al.* 2014), therefore we did not use it to obtain accurate reconstruction of population dynamics. We instead used this analysis to identify differences in the demographic history of different regional populations. LD-based N_E estimates are expected to be higher in source populations relative to derived populations in time periods preceding their split, as seen in African vs non-African human populations (McEvoy et *al.* 2011). We found that, independently of the value assumed for the model parameter alpha, Chinese FBDs had higher N_E estimates than any other FBD population throughout all the time periods assessed (Supplementary Fig. 7). The populations from Thailand and Mongolia had intermediate N_E estimates between the Chinese population and all the remaining populations until about 2,500 years ago, and in the following time periods had comparable N_E estimates to West Eurasian populations.

Measures of fit for the TREEMIX trees

Residuals from the TREEMIX tree without migration events indicated populations that did not fit well the tree model, with SE of the residual covariance exceeding 60 (Supplementary Fig. 16A). Such high residual covariance was due to two FBD populations from Saudi Arabia showing signs of admixture with grey wolves. Indeed, the addition of 10 migration edges revealed gene flow from wolves to these FBD populations, and improved the overall fit of the model, with SE of the residual covariance below 19 (Supplementary Fig. 16B). In consistence with this, P-values of all these 10 inferred migration edges obtained from the jackknife

analysis were below 0.0004, indicating that the inclusion of these gene flow events significantly improves the fit of this phylogenetic model to the data. The TREEMIX tree with 15 migration edges did not result in considerably better residual fit as compared with 10 migration edges, but additional gene flow events suggested by the residual plot were either between pairs of pure-breeds or between adjacent FBD populations (Supplementary Fig. 16C).

Assessment of the population divergence chronology in the evolutionary history of extant dogs

We used the program KIMTREE (Gautier & Vitalis 2013) to identify the tree topology that constitutes the most likely representation of the evolutionary history of extant dogs. KIMTREE uses Kimura's time-dependent solution of the diffusion equation describing genetic drift to estimate population divergence times. These time estimates are conditional on a history of population splits, which are represented as either bifurcations or polytomies. The estimates of branch lengths (reflecting the divergence times) can be obtained for different topologies, which can be subsequently compared using the deviance information criterion (Spiegelhalter *et al*. 2002).

In order to achieve relatively large sample sizes, we applied this analysis to five dog groups, identified based on the genetic structure and phylogenetic analyses described above. This included (a) FBDs from Western Eurasia (Europe and Central/West Asia), (b) FBDs from the Middle East, (c) FBDs from East Asia, (d) modern European breeds, (e) East Asian and Arctic breeds. We excluded dog breeds with ambiguous positions in the dog phylogeny, including European spitz-type breeds, Eurasier, and Tibetan terrier. We pruned the dataset from loci with MAF<0.01 and more than 10% of missing data, as well as loci in strong linkage disequilibrium (with the threshold of $r^2 < 0.5$). This resulted in 104,769 SNPs. Due to

limitations associated with computer memory, we could not complete the KIMTREE analysis using this entire dataset. Therefore, we divided the data into two datasets with almost equal number of SNPs (52,384 and 52,385). The analysis of these two datasets provided us with two pseudo-replicates, which gave very consistent results, supporting their validity.

Assessment of an indigenous status of East Asian FBDs

To address the question of whether East Asian FBDs are indigenous, i.e. whether there is evidence for the long-term presence of these lineages in East Asia, we analysed an additional SNP dataset obtained from published whole-genome sequences of dogs and wolves (Wang *et al*. 2013, Freedman *et al*. 2014, Gou *et al*. 2014), available in the DogSD database (Bai *et al*. 2015; http://dogsd.big.ac.cn). We analysed 40 individuals, including: 7 Chinese indigenous dogs from Diqing, 3 Chinese indigenous dogs from Lijlang, 3 Chinese indigenous dogs from Yingjiang, 4 Tibetan mastiffs, 1 Australian dingo, 1 Basenji, 4 Kunming dogs (a modern Chinese breed originating from an admixture of German shepherds with local cross-breed dogs), 4 German shepherds, 1 Malinois, and 8 grey wolves from China, Russia, Israel and Croatia. We used PLINK to merge SNP datasets downloaded from this database into one dataset, and prune it from loci with MAF<0.01 and loci in linkage disequilibrium (r^2<0.5). We constructed a neighbour-joining tree of inter-individuals IBS distances, a maximum-likelihood tree of population divergence, and a PCA plot, using the methods described above.

The analysis of this SNP dataset showed that Chinese FBDs are distinct from modern dog breeds (where modern breeds included the Kunming dog – a modern Chinese breed of European origin), but cluster with the Australian dingo (Supplementary Fig. 20). The dingo is a wild domestic dog whose ancient East Asian origin and isolation from other dogs for at least 3,500 years is well documented by genetic and archaeological studies (Larson *et al*. 2012, Oskarsson *et al*. 2012, Sacks *et al*. 2013). Clustering of Chinese FBDs with the dingo

indicates that they belong to the East Asian lineage from which the dingo originated at least 3,500 years ago, supporting their indigenous origin. These results also demonstrate a clear genetic distinctiveness of Chinese FBDs from modern pure-breed dogs, with the exception of one individual that might be admixed with Kunming dogs.

References

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–64. (doi:10.1101/gr.094052.109)

Bai B, Zhao W-M, Tang B-X, Wang Y-Q, Wang L, Zhang Z, Yang H-C, Liu Y-H, Zhu J-W, Irwin DM, *et al.* 2015. DoGSD: the dog and wolf genome SNP database. *Nucl Acids Res*. **43**, D777–83. (doi:10.1093/nar/gku1174)

Brown SK, Darwent CM, Sacks BN. 2013. Ancient DNA evidence for genetic continuity in Arctic dogs. *J Archaeol Sci*. **40**, 1279–88. (doi:10.1016/j.jas.2012.09.010)

Brown SK, Darwent CM, Wictum EJ, Sacks BN. 2015. Using multiple markers to elucidate the ancient, historical and modern relationships among North American Arctic dog breeds. *Heredity* 24 June 2015. (doi:10.1038/hdy.2015.49)

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol.* **14**, 2611–20. (doi:10.1111/j.1365-294X.2005.02553.x)

Freedman AH, Gronau I, Schweizer RM, Ortega-Del Vecchyo D, Han E, Silva PM, Galaverni M, Fan Z, Marx P, Lorente-Galdos B, *et al.* 2014. Genome sequencing highlights the dynamic early history of dogs. *PLoS Genet*. **10**, e1004016. (doi:10.1371/journal.pgen.1004016)

Gautier M, Vitalis R. 2013. Inferring population histories using genome-wide allele frequency data. *Mol Biol Evol*. **30**, 654–68. (doi:10.1093/molbev/mss257)

Gou X, Wang Z, Li N, Qiu F, Xu Z, Yan D, Yang S, Jia J, Kong X, Wei Z, *et al.* 2014. Whole genome sequencing of six dog breeds from continuous altitudes reveals adaption to high-altitude hypoxia. *Genome Res*. **24**, 1308–1315. (doi:10.1101/gr.171876.113)

Guillot G, Santos F, Estoup A. 2008. Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics* **24**, 1406–1407. (doi:10.1093/bioinformatics/btn136).

Hayes BJ, Visscher PM, McPartlan HC, Goddard ME. 2003. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Res*. **13**, 635–43. (doi:10.1101/gr.387103)

Konovalov DA, Manning C, Henshaw MT. 2004. KINGROUP: a program for pedigree relationship reconstruction and kin group assignments using genetic markers. *Mol Ecol Notes.* **4**, 779–82. (doi:10.1111/j.1471-8286.2004.00796.x)

Larson G, Karlsson EK, Perri A, Webster MT, Ho SYW, Peters J, Stahl PW, Piper PJ, Lingaas F, Fredholm M, *et al.* 2012. Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc Natl Acad Sci USA*. **109**, 8878–8883. (doi:10.1073/pnas.1203005109)

Marshall TC, Slate J, Kruuk LEB, Pemberton JM. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol.* **7**, 639–55. (doi:10.1038/nature04338)

McEvoy BP, Powell JE, Goddard ME, Visscher PM. 2011. Human population dispersal "Out of Africa" estimated from linkage disequilibrium and allele frequencies of SNPs. *Genome Res*. **21**, 821–829. (doi:10.1101/gr.119636.110)

Mech LD, Seal US. 1987. Premature reproductive activity in wild wolves. *J Mammal.* 68, 871–3. (doi:10.2307/1381570)

Oskarsson MCR, Klütsch CFC, Boonyaprakob U, Wilton A, Tanabe Y, Savolainen P. 2012. Mitochondrial DNA data indicate an introduction through Mainland Southeast Asia for Australian dingoes and Polynesian domestic dogs. *Proc R Soc B London*. **279**, 967–974. (doi:10.1098/rspb.2011.1395)

Parker HG, Kim LV, Sutter NB, Carlson S, Lorentzen TD, Malek TB, Johnson GS, DeFrance HB, Ostrander EA, Kruglyak L. 2004. Genetic structure of the purebred domestic dog. *Science.* **304**, 1160–1164. (doi:10.1126/science.1097406)

Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. *Bioinformatics.* **28**, 2537–2539. (doi:10.1093/bioinformatics/bts460)

Pickrell JK, Pritchard JK. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet*. **8**(11), e1002967. (doi:10.1371/journal.pgen.1002967)

Pritchard J, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics.* **155**, 945–59.

Sacks BN, Brown SK, Stephens D, Pedersen NC, Wu J-T, Berry O. 2013. 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, 1103–18. (doi:10.1093/molbev/mst027)

Spiegelhalter DJ, Best NG, Carlin BP, Linde AVD. 2002. Bayesian measures of model complexity and fit. *J Roy Stat Soc B.* **64**, 583–639.

Stronen, AV, Jędrzejewska B, Pertoldi C, Demontis D, Randi E, Niedziałkowska M, Pilot M, Sidorovich VE, Dykyy I, Kusak J, *et al.* 2013. North-south differentiation and a region of high diversity in European wolves (*Canis lupus*). *PLoS ONE.* **8**(10), e76454. (doi:10.1371/journal.pone.0076454)

Tenesa A, Navarro P, Hayes BJ, Duffy DL, Clarke GM, Goddard ME, Visscher PM. 2007. Recent human effective population size estimated from linkage disequilibrium. *Genome Res*. **17**, 520–26. (doi:10.1101/gr.6023607)

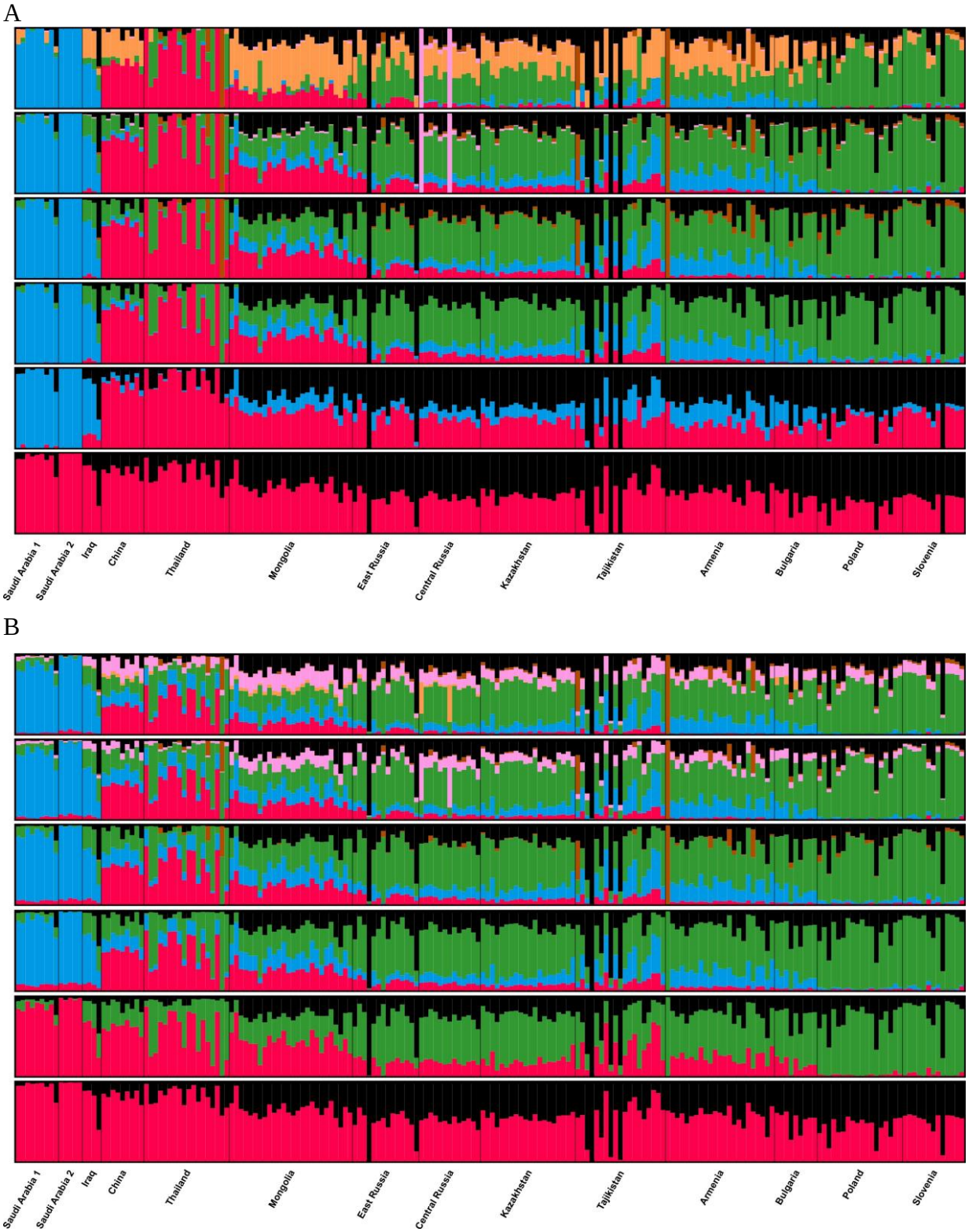
van Asch B, Zhang A-B, Oskarsson MCR, Klütsch CFC, Amorim A, Savolainen P. 2013. Pre-Columbian origins of Native American dog breeds, with only limited replacement by European dogs, confirmed by mtDNA analysis. *Proc R Soc B London*. **280**, 2013.1142. (doi:10.1098/rspb.2013.1142)

Vaysse A, Ratnakumar A, Derrien T, Axelsson E, Rosengren Pielberg G, Sigurdsson S, Fall T, Seppälä EH, Hansen MST, Lawley CT, *et al.* 2011. Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genet*. **7**(10), e1002316. (doi:10.1371/journal.pgen.1002316)

vonHoldt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, Degenhardt JD, Boyko AR, Earl DA, Auton A, *et al.* 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* **464**, 898–902. (doi:10.1038/nature08837)

Wang G-D, Zhai W, Yang H-C, Fan R-X, Cao X, Zhong L, Wang L, Liu F, Wu H, Cheng L-G, *et al.* 2013. The genomics of selection in dogs and the parallel evolution between dogs and humans. *Nature Commun*. **4**, 1860. (doi:10.1038/ncomms2814)

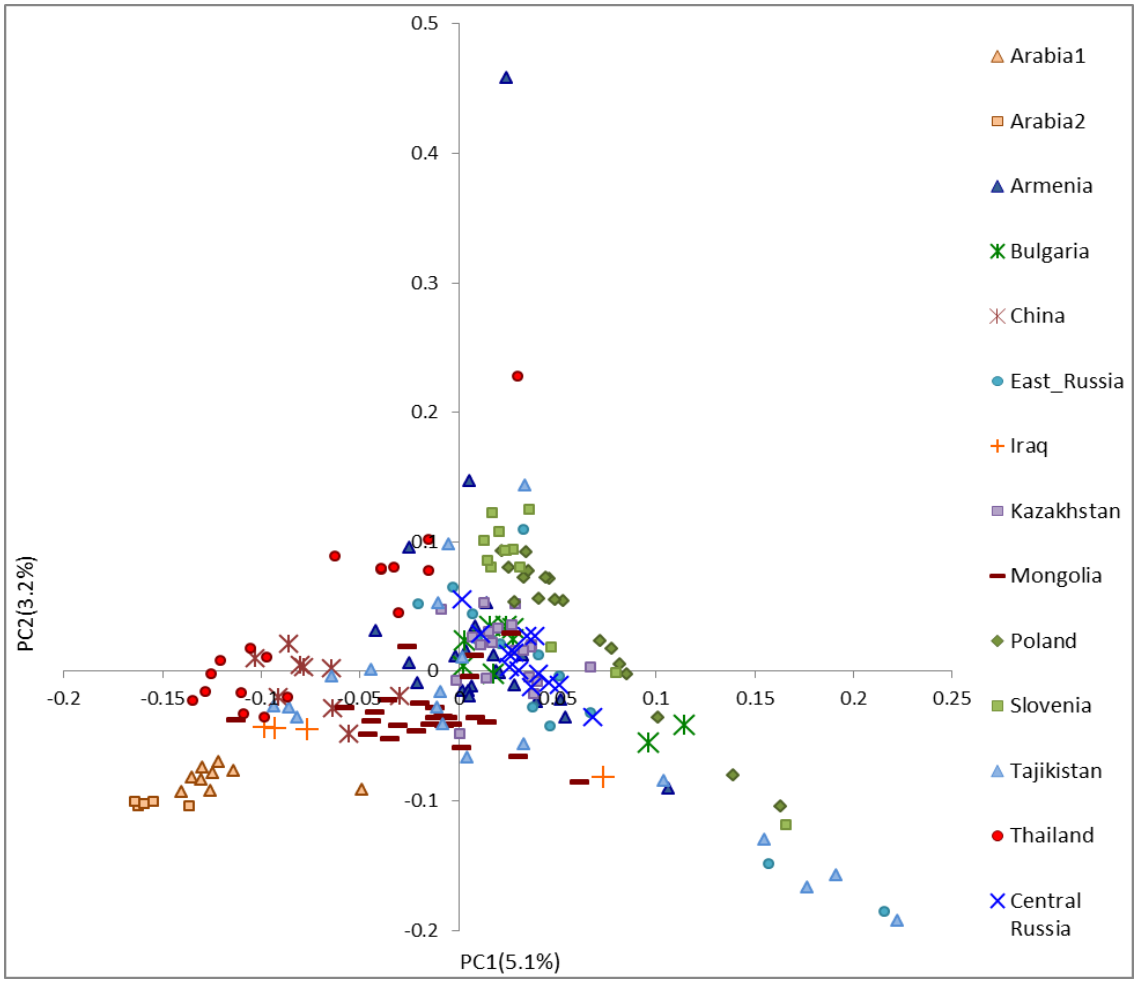
Supplementary Figure 1



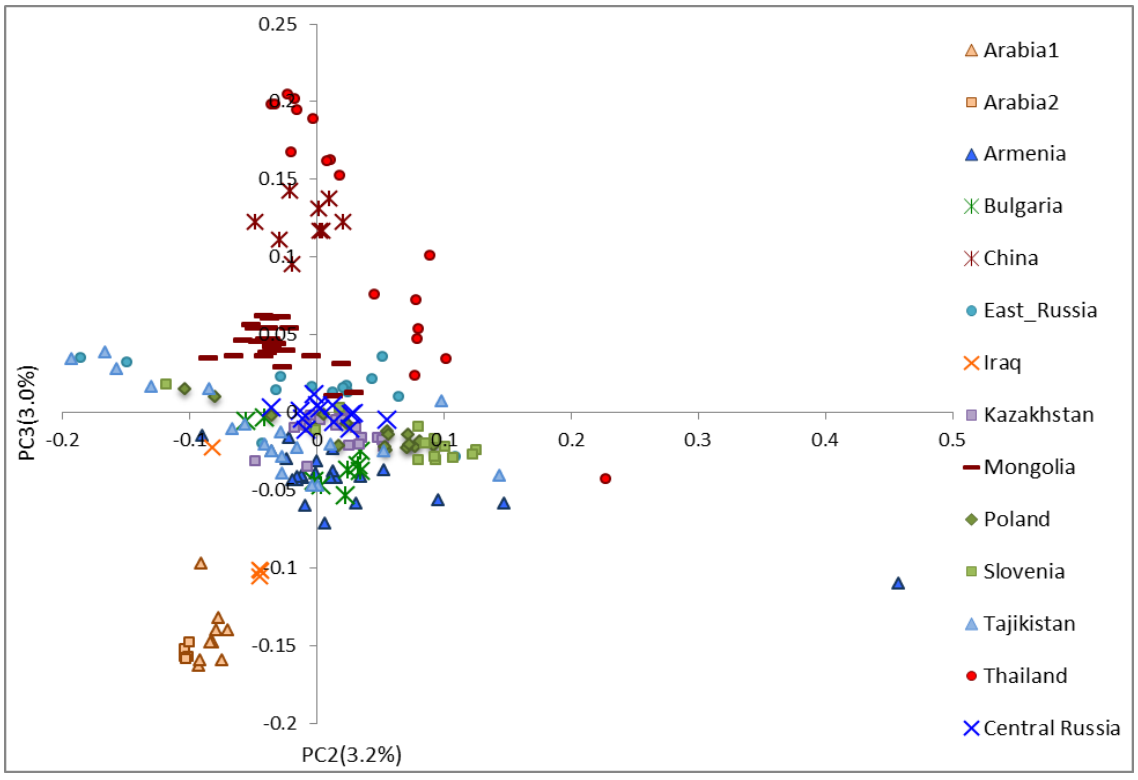
Supplementary Figure 1. Population genetic structure in Eurasian FBDs. Population genetic structure inferred using (A) ADMIXTURE and (B) STRUCTURE, for K between 2 (bottom) and 7 (top). Membership in different genetic populations is represented by different colours.

Supplementary Figure 2

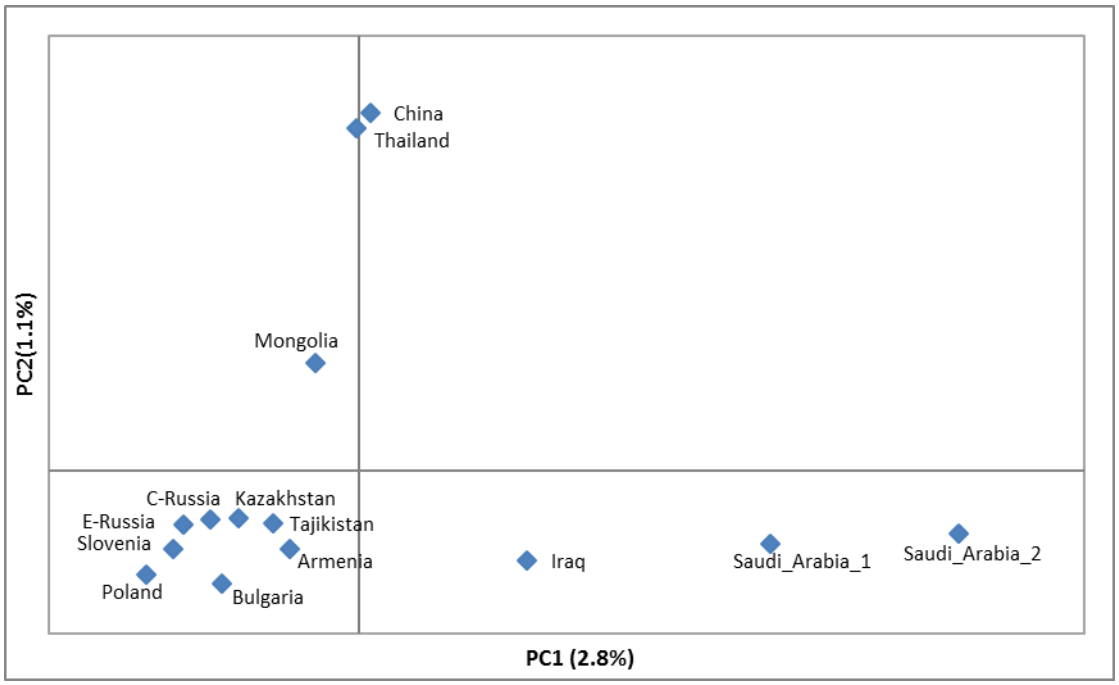
A



B



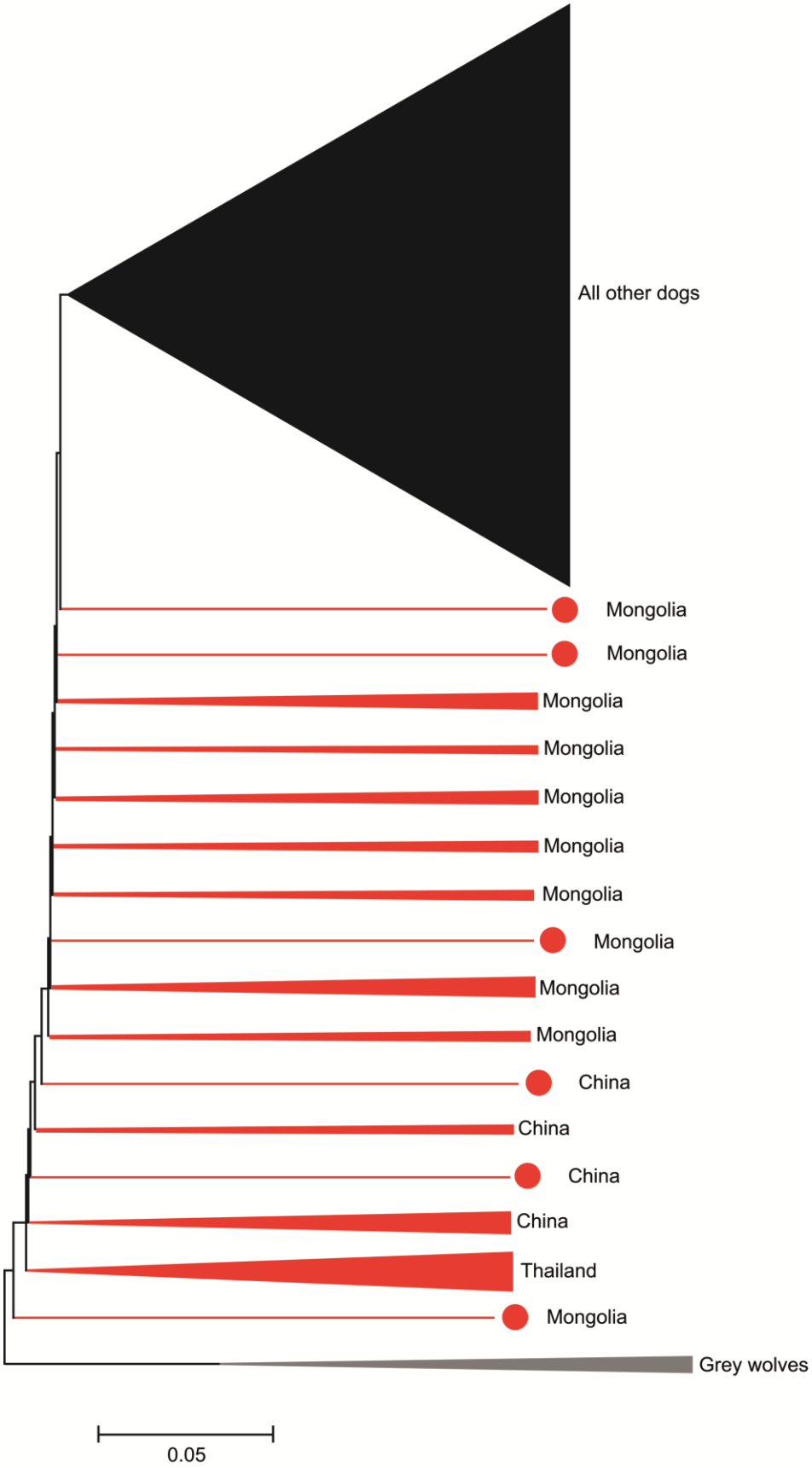
C



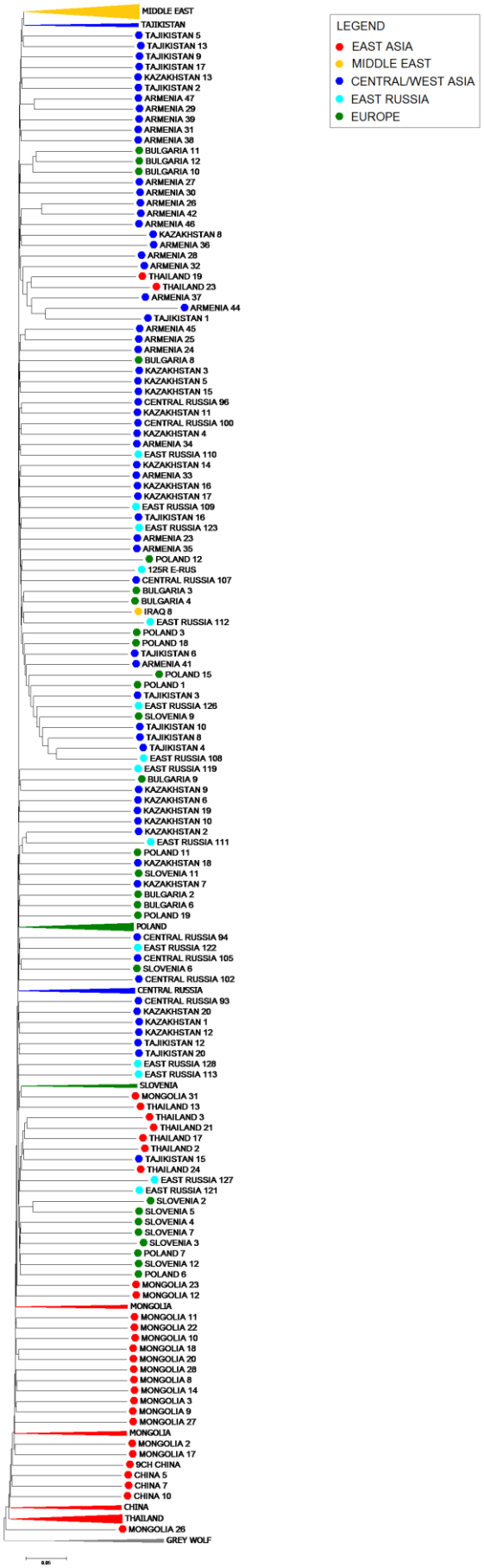
Supplementary Figure 2. Principal Component Analysis of genetic differentiation among FBDs from different parts of Eurasia. (A) Representing individual-based genetic distances: PC1 and PC2 (an outlier from Armenia groups with pure-breed dogs: see Supplementary Figure 8), (B) Representing individual-based genetic distances: PC2 and PC3. (C) Representing pair-wise F_{ST} between local populations of FBDs.

Supplementary Figure 3

A



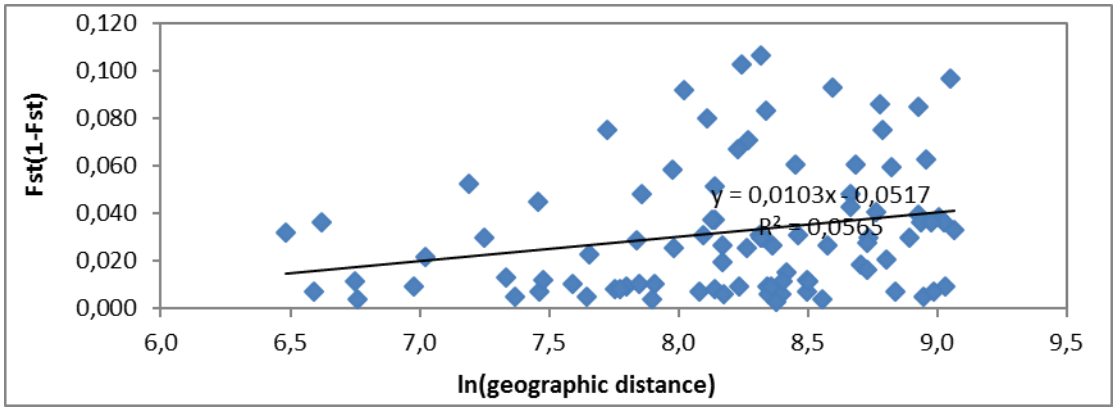
B


LEGEND
EAST ASIA
MIDDLE EAST
CENTRAL/WEST ASIA
EAST RUSSIA
EUROPE
MIDDLE EAST
TAJIKISTAN
TAJIKISTAN 5
TAJIKISTAN 13
TAJIKISTAN 9
TAJIKISTAN 17
KAZAKHSTAN 13
TAJIKISTAN 2
ARMENIA 47
ARMENIA 29
ARMENIA 39
ARMENIA 31
ARMENIA 38
BULGARIA 11
BULGARIA 12
BULGARIA 10
ARMENIA 27
ARMENIA 30
ARMENIA 26
ARMENIA 42
ARMENIA 46
KAZAKHSTAN 8
ARMENIA 36
ARMENIA 28
ARMENIA 32
THAILAND 19
THAILAND 23
ARMENIA 37
ARMENIA 44
TAJIKISTAN 1
ARMENIA 45
ARMENIA 25
ARMENIA 24
BULGARIA 8
KAZAKHSTAN 3
KAZAKHSTAN 5
KAZAKHSTAN 15
CENTRAL RUSSIA 96
KAZAKHSTAN 11
CENTRAL RUSSIA 100
KAZAKHSTAN 4
ARMENIA 34
EAST RUSSIA 110
KAZAKHSTAN 14
ARMENIA 33
KAZAKHSTAN 16
KAZAKHSTAN 17
EAST RUSSIA 109
TAJIKISTAN 16
EAST RUSSIA 123
ARMENIA 23
ARMENIA 35
POLAND 12
125R E-RUS
CENTRAL RUSSIA 107
BULGARIA 3
BULGARIA 4
IRAQ 8
EAST RUSSIA 112
POLAND 3
POLAND 18
TAJIKISTAN 6
ARMENIA 41
POLAND 15
POLAND 1
TAJIKISTAN 3
EAST RUSSIA 126
SLOVENIA 9
TAJIKISTAN 10
TAJIKISTAN 8
TAJIKISTAN 4
EAST RUSSIA 108
EAST RUSSIA 119
BULGARIA 9
KAZAKHSTAN 9
KAZAKHSTAN 6
KAZAKHSTAN 19
KAZAKHSTAN 10
KAZAKHSTAN 2
EAST RUSSIA 111
POLAND 11
KAZAKHSTAN 18
SLOVENIA 11
KAZAKHSTAN 7
BULGARIA 2
BULGARIA 6
POLAND 19
POLAND
CENTRAL RUSSIA 94
EAST RUSSIA 122
CENTRAL RUSSIA 105
SLOVENIA 6
CENTRAL RUSSIA 102
CENTRAL RUSSIA
CENTRAL RUSSIA 93
KAZAKHSTAN 20
KAZAKHSTAN 1
KAZAKHSTAN 12
TAJIKISTAN 12
TAJIKISTAN 20
EAST RUSSIA 128
EAST RUSSIA 113
SLOVENIA
MONGOLIA 31
THAILAND 13
THAILAND 3
THAILAND 21
THAILAND 17
THAILAND 2
TAJIKISTAN 15
THAILAND 24
EAST RUSSIA 127
EAST RUSSIA 121
SLOVENIA 2
SLOVENIA 5
SLOVENIA 4
SLOVENIA 7
SLOVENIA 3
POLAND 7
SLOVENIA 12
POLAND 6
MONGOLIA 23
MONGOLIA 12
MONGOLIA
MONGOLIA 11
MONGOLIA 22
MONGOLIA 10
MONGOLIA 18
MONGOLIA 20
MONGOLIA 28
MONGOLIA 8
MONGOLIA 14
MONGOLIA 3
MONGOLIA 9
MONGOLIA 27
MONGOLIA
MONGOLIA 2
MONGOLIA 17
9CH CHINA
CHINA 5
CHINA 7
CHINA 10
CHINA
THAILAND
MONGOLIA 26
GREY WOLF
0.05

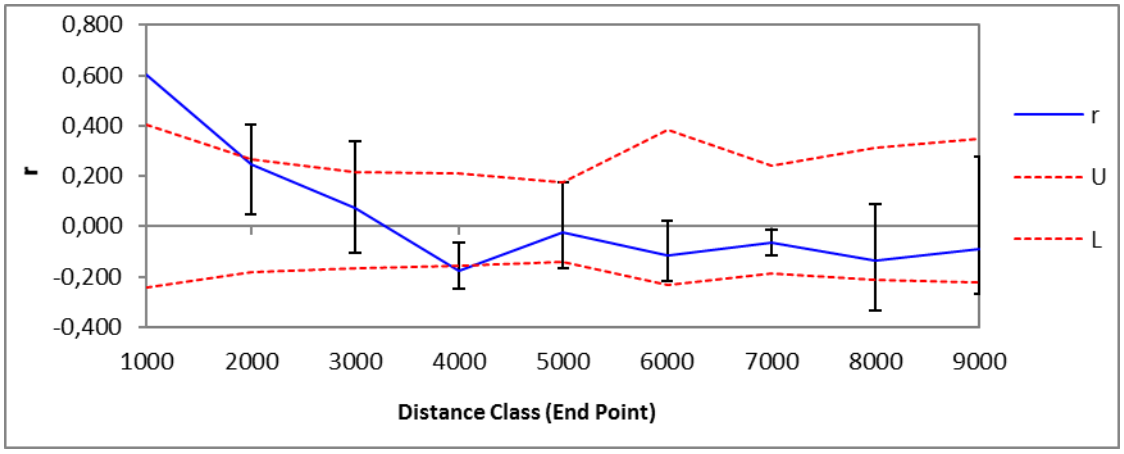
Supplementary Figure 3. Tree of inter-individual IBS distances between FBDs. (A) Simplified tree with only individuals branching from basal nodes distinguished, while all the others are presented as one clade. (B) Detailed tree, with individuals from different regions are distinguished by colours. Individuals from the same geographic location or the same breed clustering together are represented as triangles.

Supplementary Figure 4

A

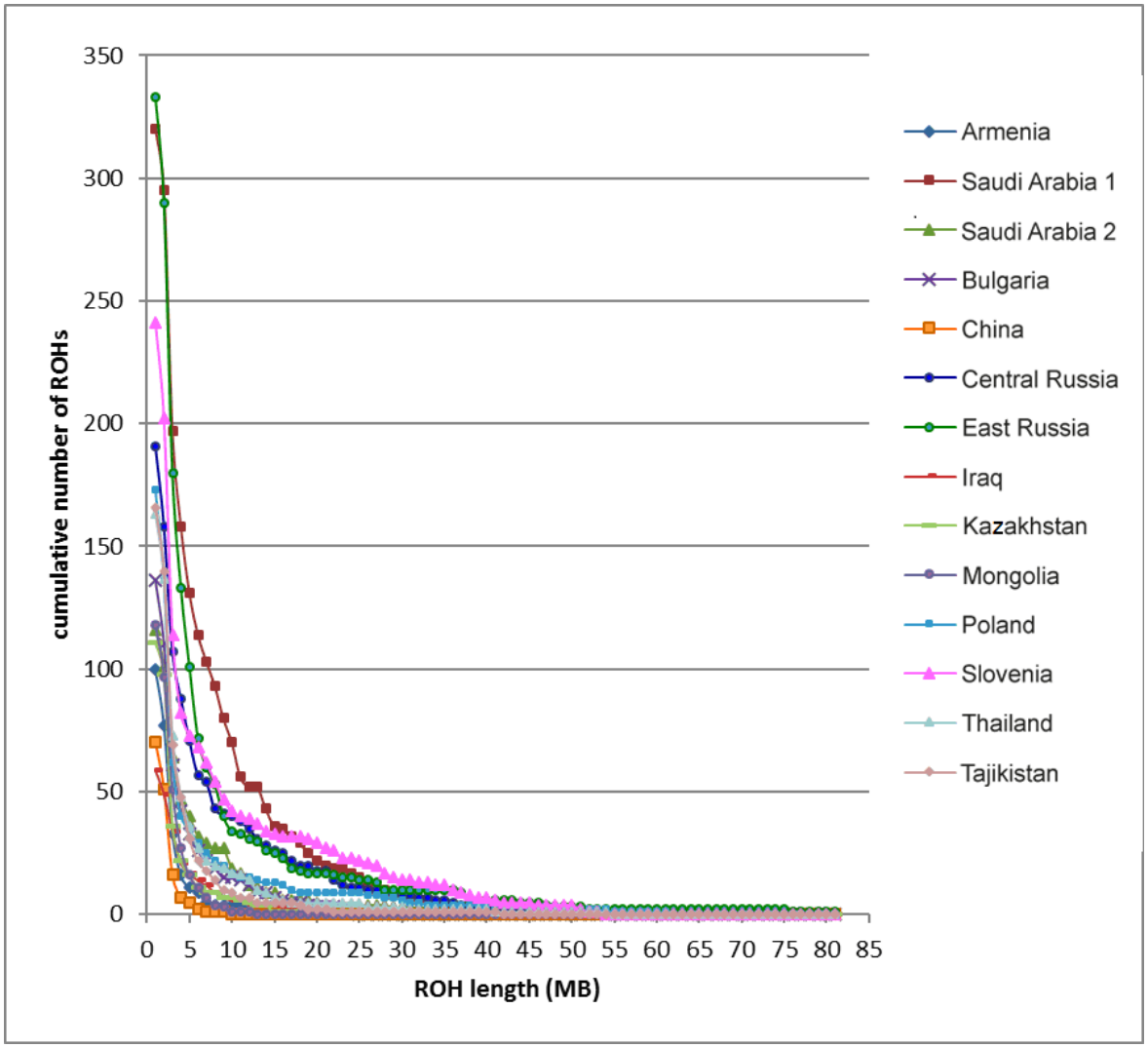


B



Supplementary Figure 4. Genetic differentiation between local populations of FBDs in relation to their geographic distance. (A) Relationship between transformed pair-wise F_{ST} values ($F_{ST}/(1-F_{ST})$) and natural logarithm of geographic distance, illustrating isolation by distance pattern; (B) Spatial autocorrelation of genetic differentiation between populations, represented as pair-wise F_{ST}. "r" denotes the autocorrelation coefficient, and U and L upper and lower limits of the 95% confidence interval for the null hypothesis of no autocorrelation.

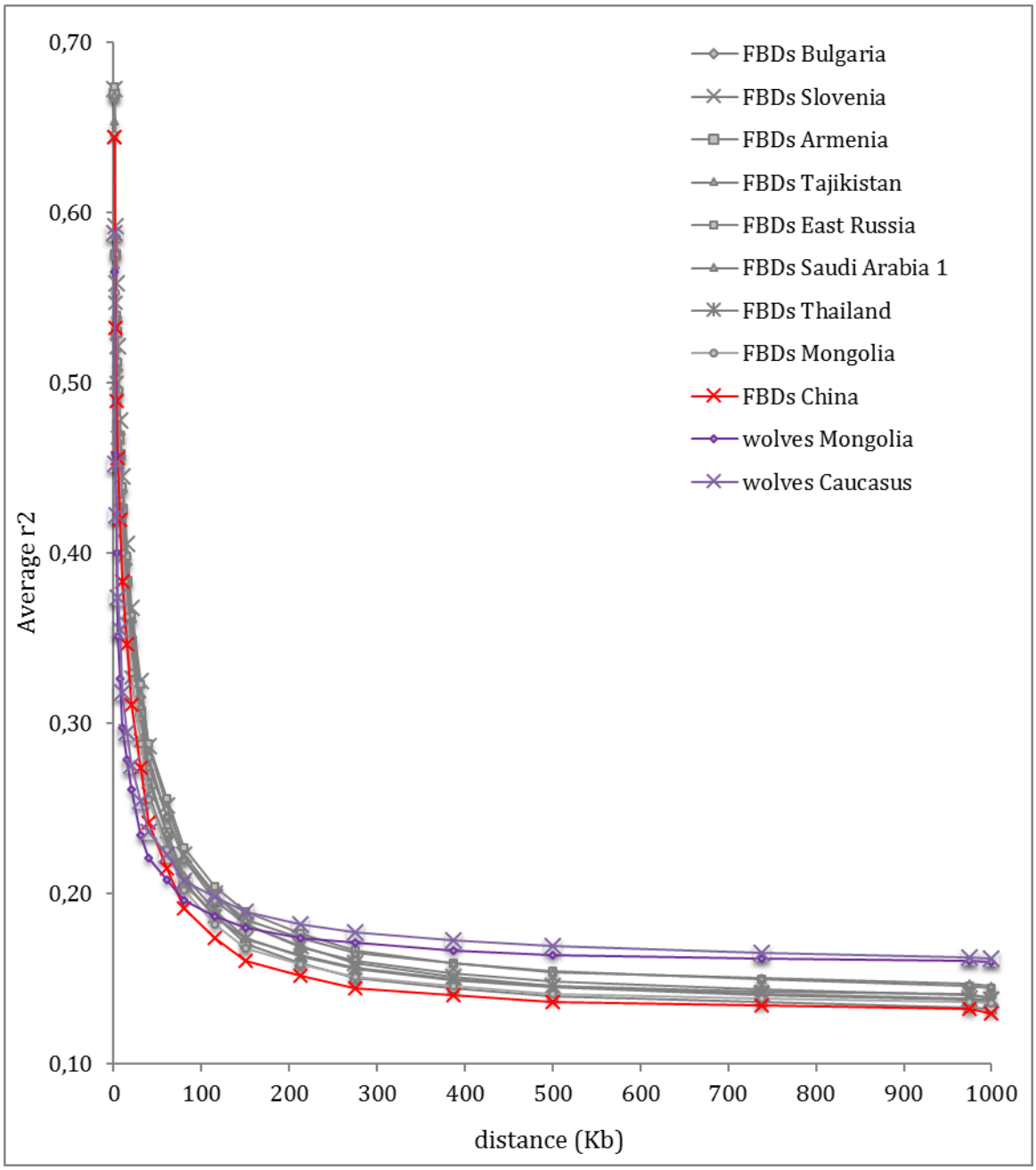
Supplementary Figure 5



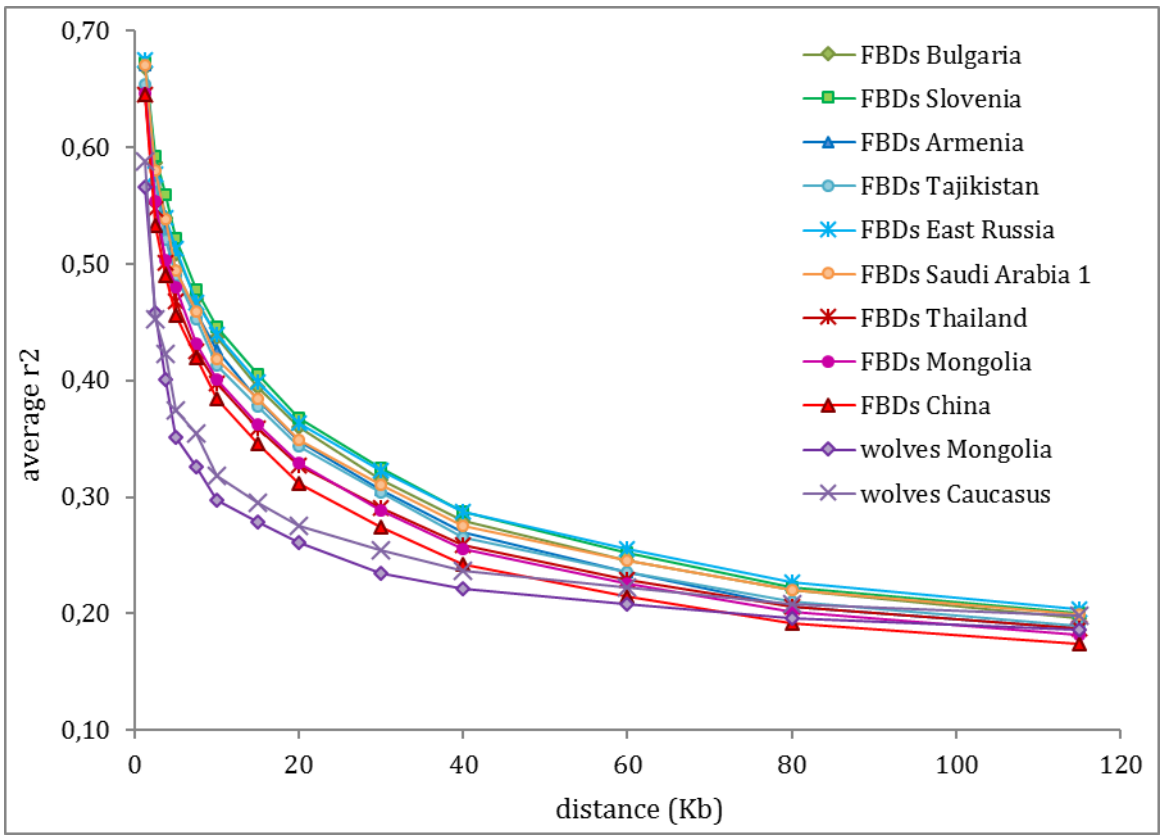
Supplementary Figure 5. Frequency distribution of runs of homozygosity (ROHs), illustrating autozygosity levels in local populations of FBDs. The vertical axis represents the cumulative number of ROHs of length equal to or greater than the length represented on the horizontal axis.

Supplementary Figure 6

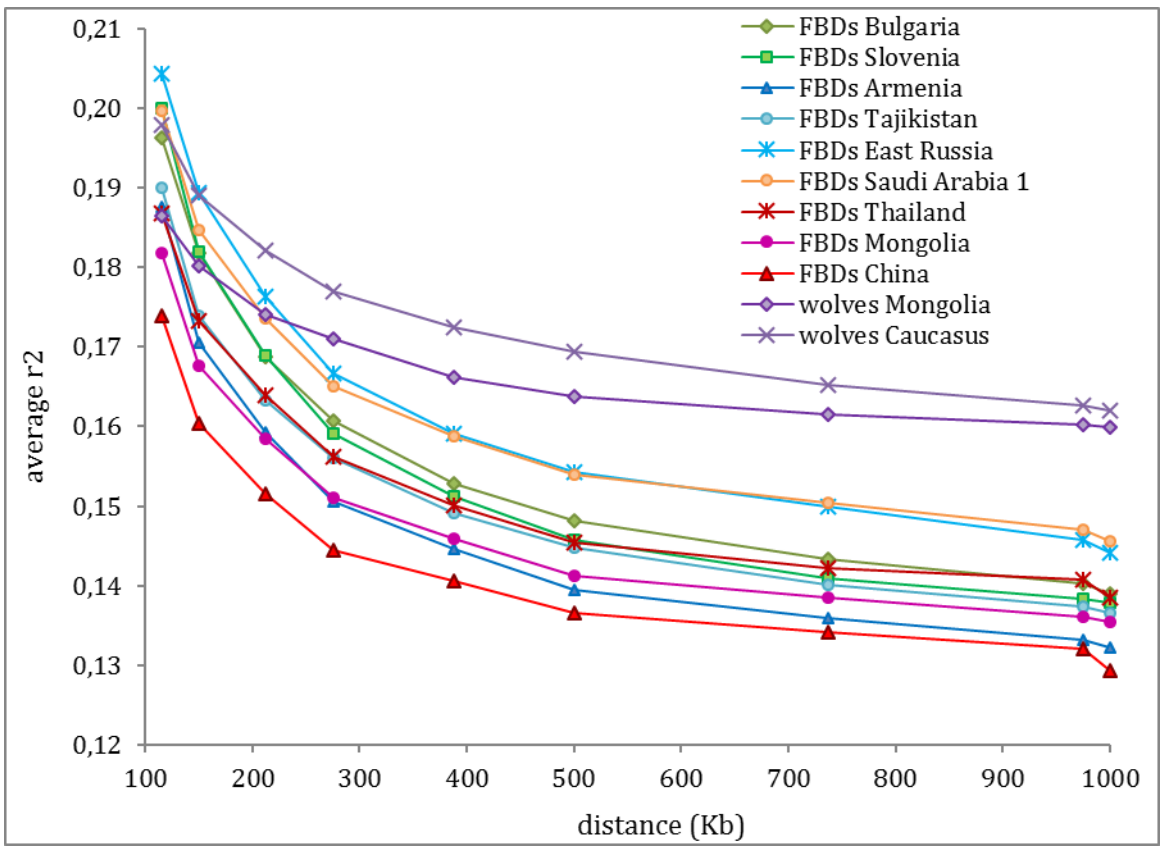
A



B


0,70
0,60
0,50
0,40
0,30
0,20
0,10
average r2
0
20
40
60
80
100
120
distance (Kb)
FBDs Bulgaria
FBDs Slovenia
FBDs Armenia
FBDs Tajikistan
FBDs East Russia
FBDs Saudi Arabia 1
FBDs Thailand
FBDs Mongolia
FBDs China
wolves Mongolia
wolves Caucasus

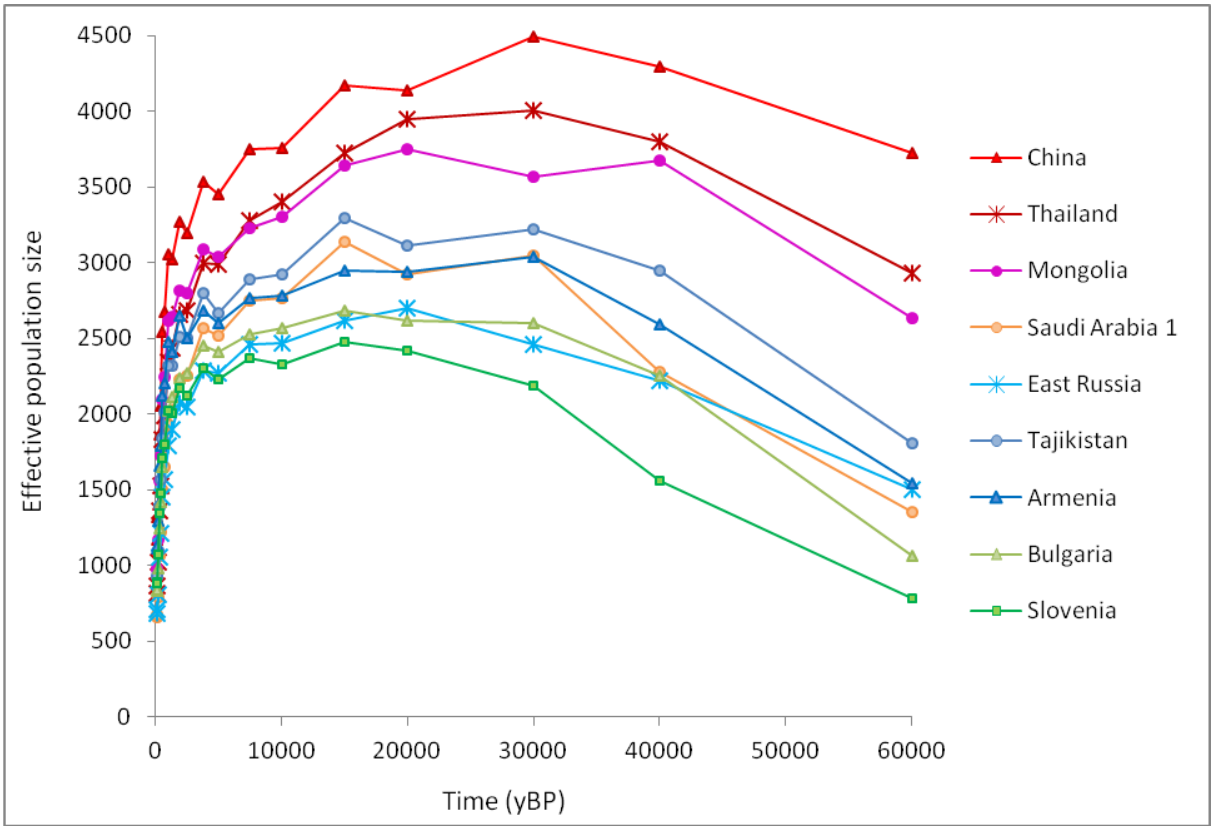
C


0,21
0,20
0,19
0,18
0,17
0,16
0,15
0,14
0,13
0,12
average r2
100
200
300
400
500
600
700
800
900
1000
distance (Kb)
FBDs Bulgaria
FBDs Slovenia
FBDs Armenia
FBDs Tajikistan
FBDs East Russia
FBDs Saudi Arabia 1
FBDs Thailand
FBDs Mongolia
FBDs China
wolves Mongolia
wolves Caucasus

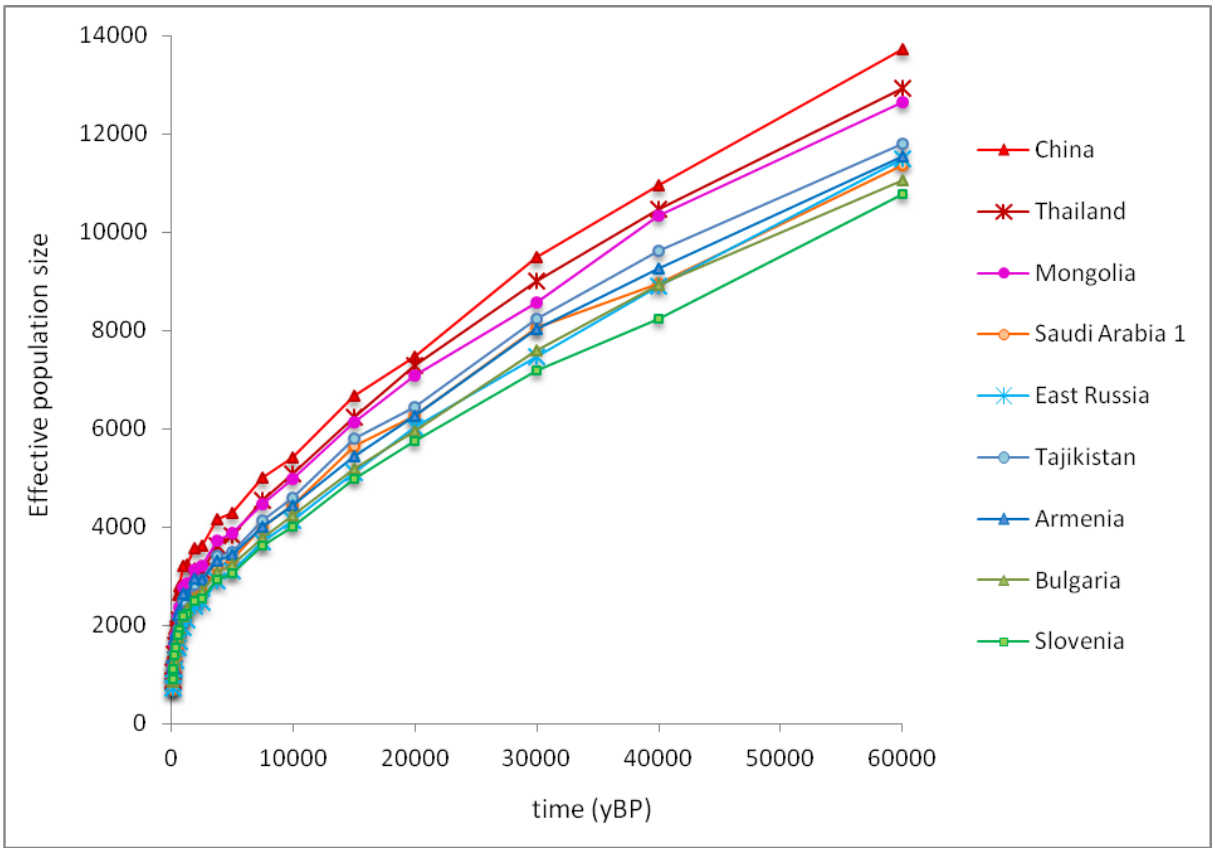
Supplementary Figure 6. Linkage disequilibrium in selected populations of FBDs in comparison with grey wolves. Linkage disequilibrium is represented by average genotypic association coefficient r^2 and is plotted as a function of inter-SNP distance. The sample size for each local dog and wolf population was 9 individuals. (A) LD values for the entire range of genetic distances, with only Chinese FBD and wolf populations distinguished, (B) LD values for the genetic distances between 1.25 and 115 kb, (C) LD values for the genetic distances between 115 and 1000 kb. It is important to stress that this analysis did not include pure-breed dogs but only FBDs.

Supplementary Figure 7

A

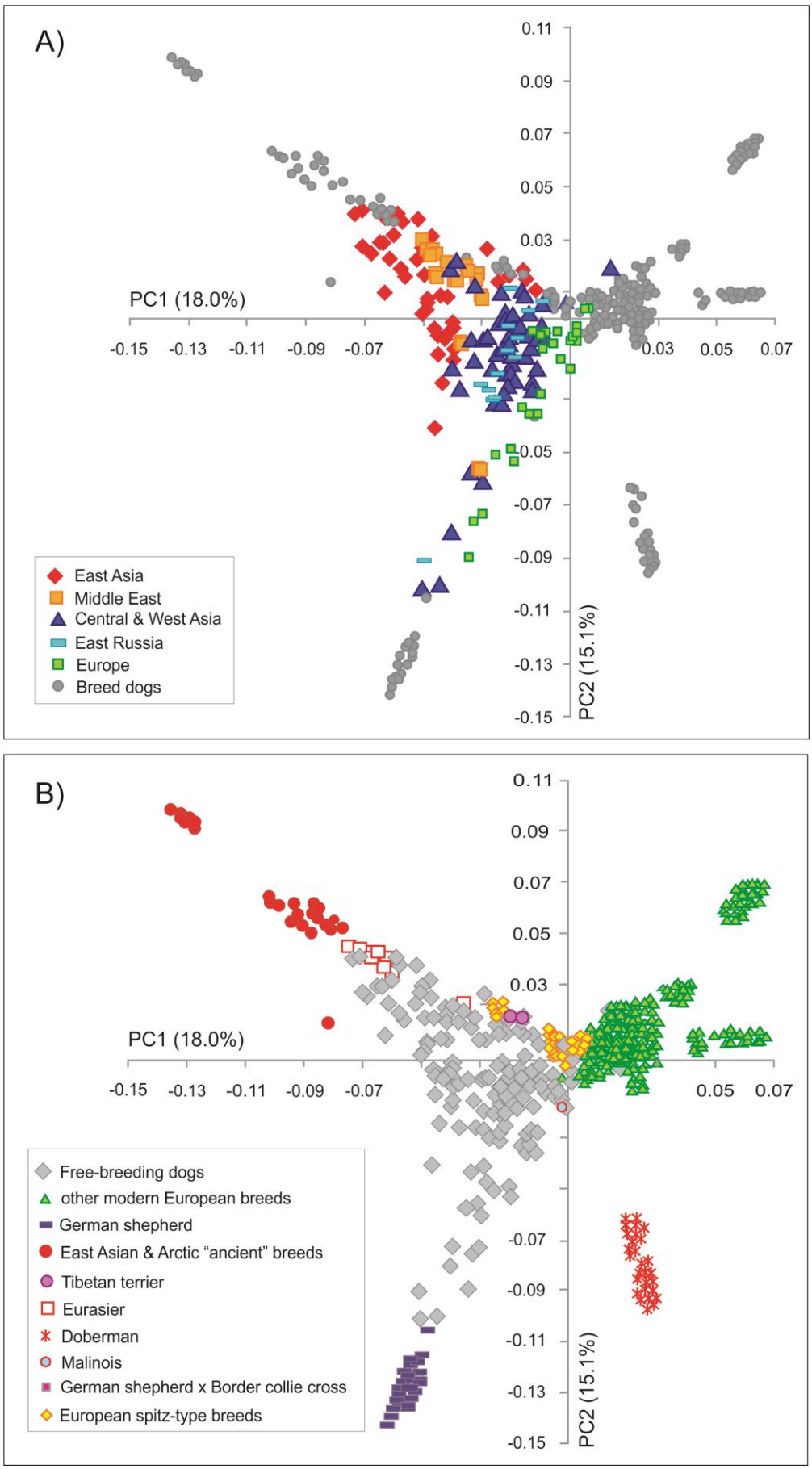


B

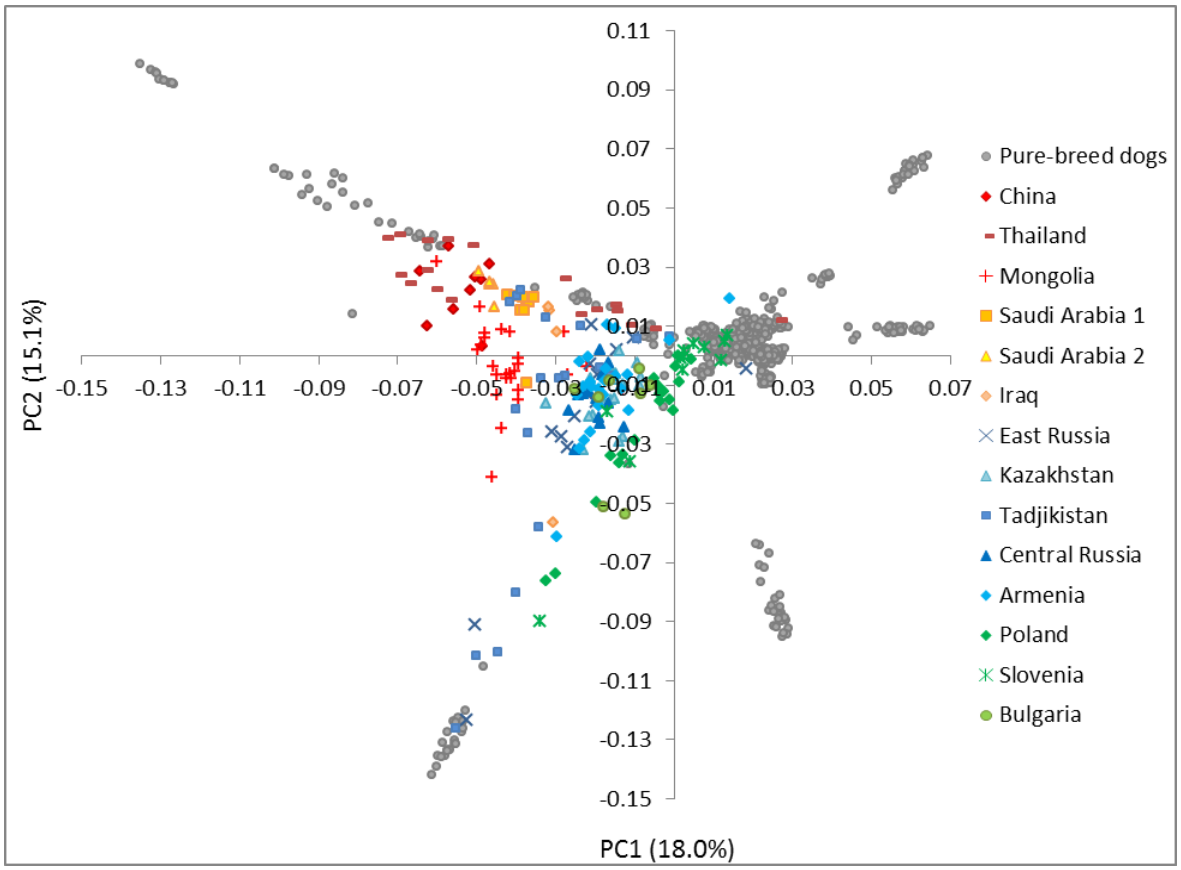


Supplementary Figure 7. Temporal changes in effective population size of Eurasian FBD populations inferred from the linkage disequilibrium (A) for alpha = 2, (B) for alpha = 1. Although none of these plots provides a correct representation of the dog population dynamics, they consistently show higher N_E estimates in China, Thailand and Mongolia compared to other regions, consistent with East Asian FBDs being ancestral to FBDs from other regions. It is important to stress that convergence of all the dog populations into a single N_E curve at times pre-dating domestication is not expected for LD-based estimates, but instead the ancestral populations are expected to have higher N_E estimates (see McEvoy *et al*. 2011).

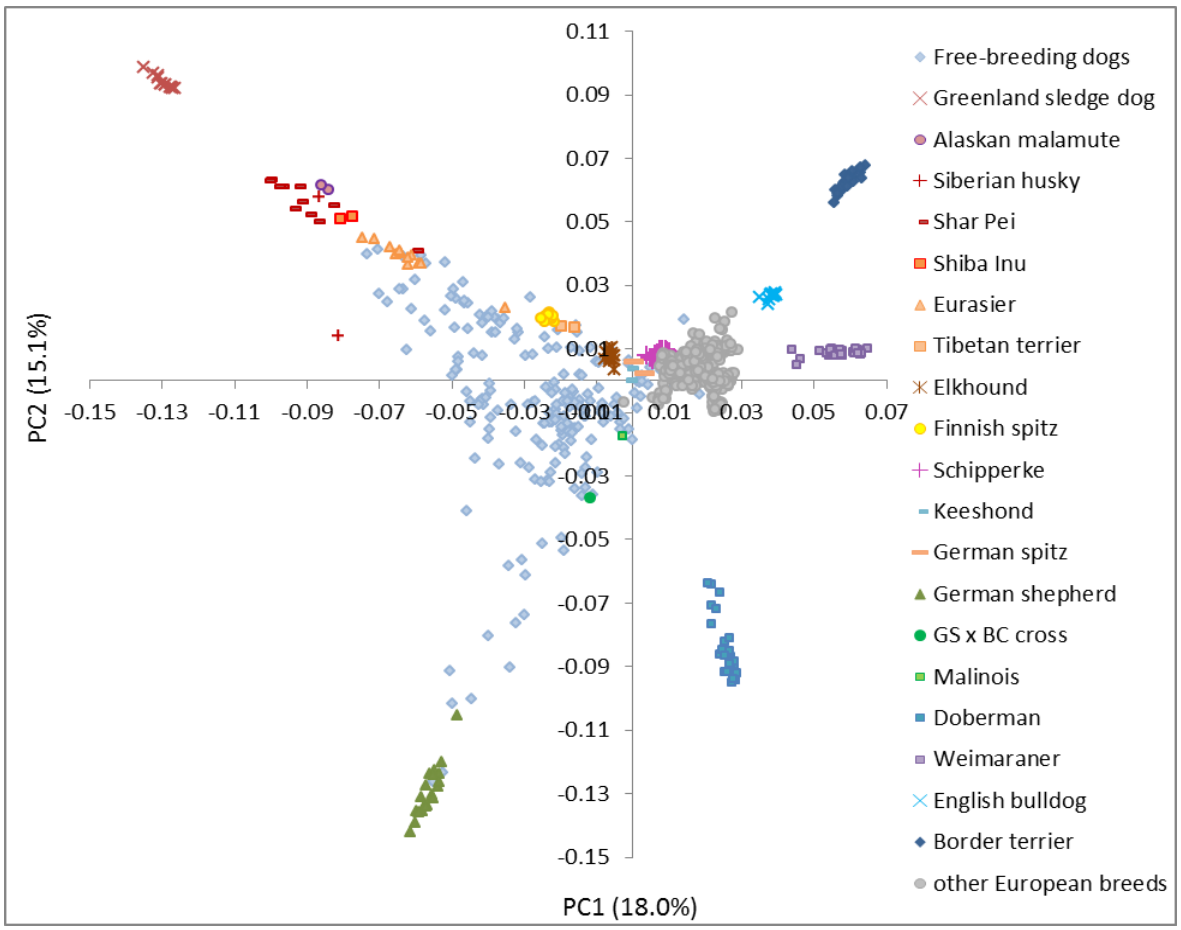
Supplementary Figure 8



C


0.11
0.09
0.07
0.05
0.03
0.01
-0.01
-0.03
-0.05
-0.07
-0.09
-0.11
-0.13
-0.15
-0.15 -0.13 -0.11 -0.09 -0.07 -0.05 -0.03 -0.01 0.01 0.03 0.05 0.07
PC2 (15.1%)
PC1 (18.0%)
Pure-breed dogs
China
Thailand
Mongolia
Saudi Arabia 1
Saudi Arabia 2
Iraq
East Russia
Kazakhstan
Tadjikistan
Central Russia
Armenia
Poland
Slovenia
Bulgaria

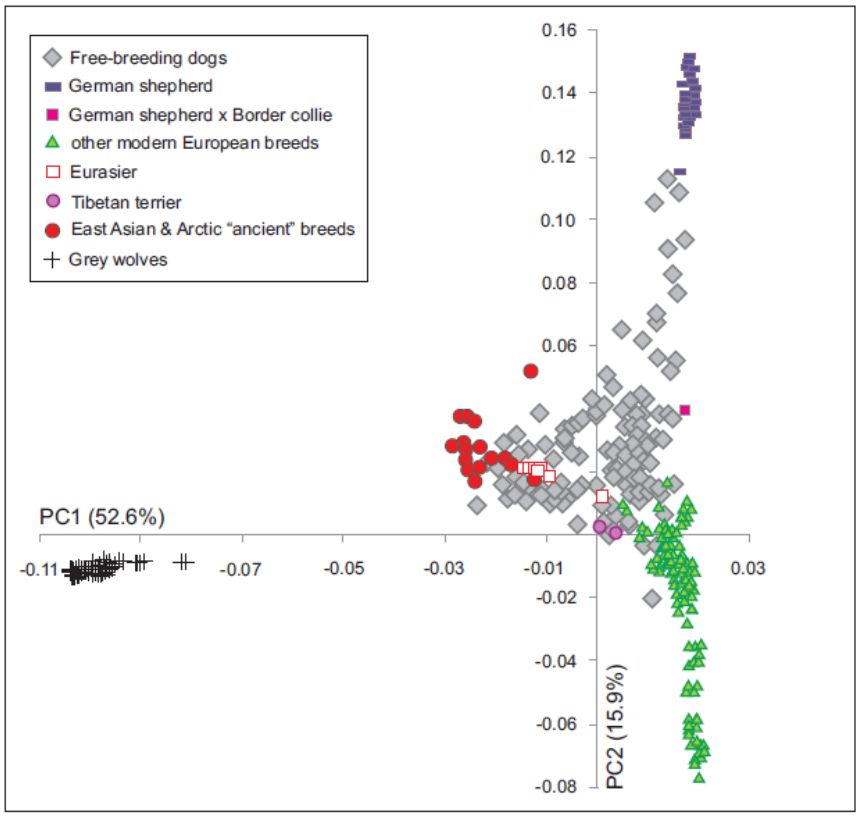
D


0.11
0.09
0.07
0.05
0.03
0.01
-0.01
-0.03
-0.05
-0.07
-0.09
-0.11
-0.13
-0.15
-0.15 -0.13 -0.11 -0.09 -0.07 -0.05 -0.03 -0.01 0.01 0.03 0.05 0.07
PC2 (15.1%)
PC1 (18.0%)
Free-breeding dogs
Greenland sledge dog
Alaskan malamute
Siberian husky
Shar Pei
Shiba Inu
Eurasier
Tibetan terrier
Elkhound
Finnish spitz
Schipperke
Keeshond
German spitz
German shepherd
GS x BC cross
Malinois
Doberman
Weimaraner
English bulldog
Border terrier
other European breeds

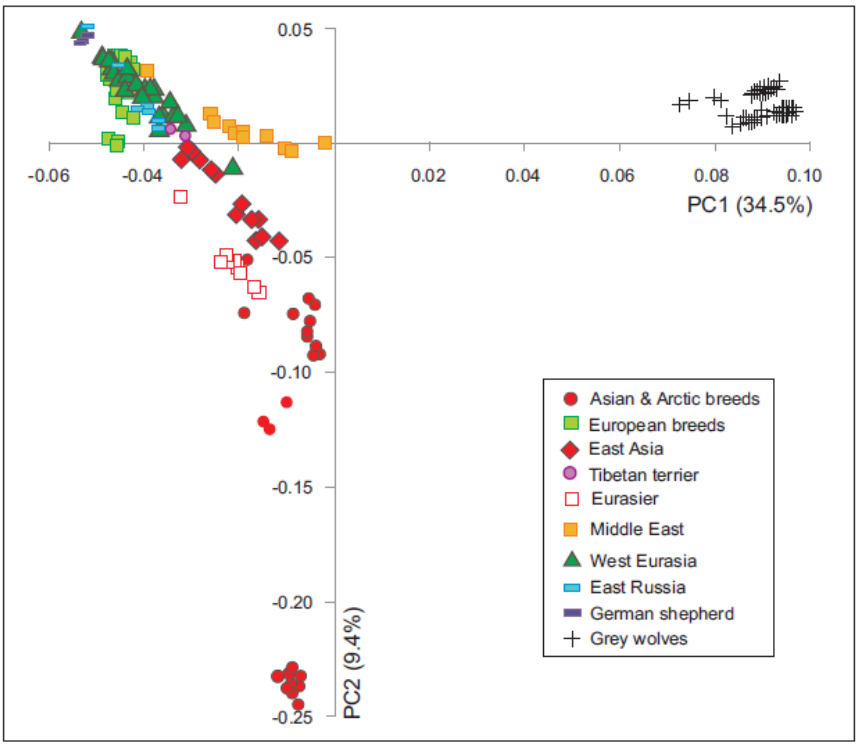
Supplementary Figure 8. Principal component analysis of genetic differentiation among FBDs and pure-breed dogs. (A) with FBDs from different regions distinguished by different symbols, (B) with groups of dog breeds distinguished by different symbols, (C) with local FBD populations distinguished by different symbols, (D) with individual dog breeds distinguished by different symbols. GS x BC cross – a F1 cross between German shepherd and Border collie.

Supplementary Figure 9

A

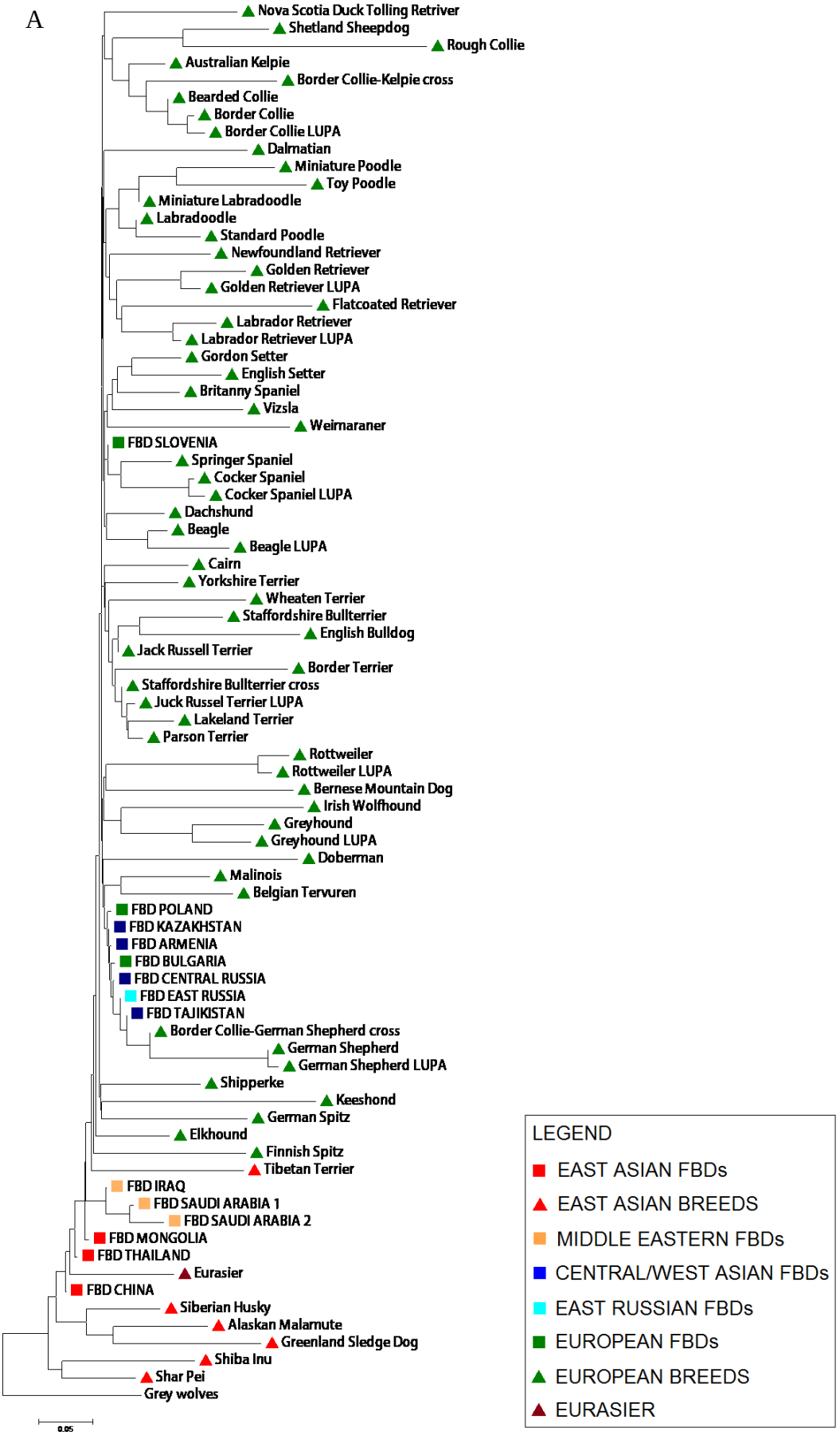


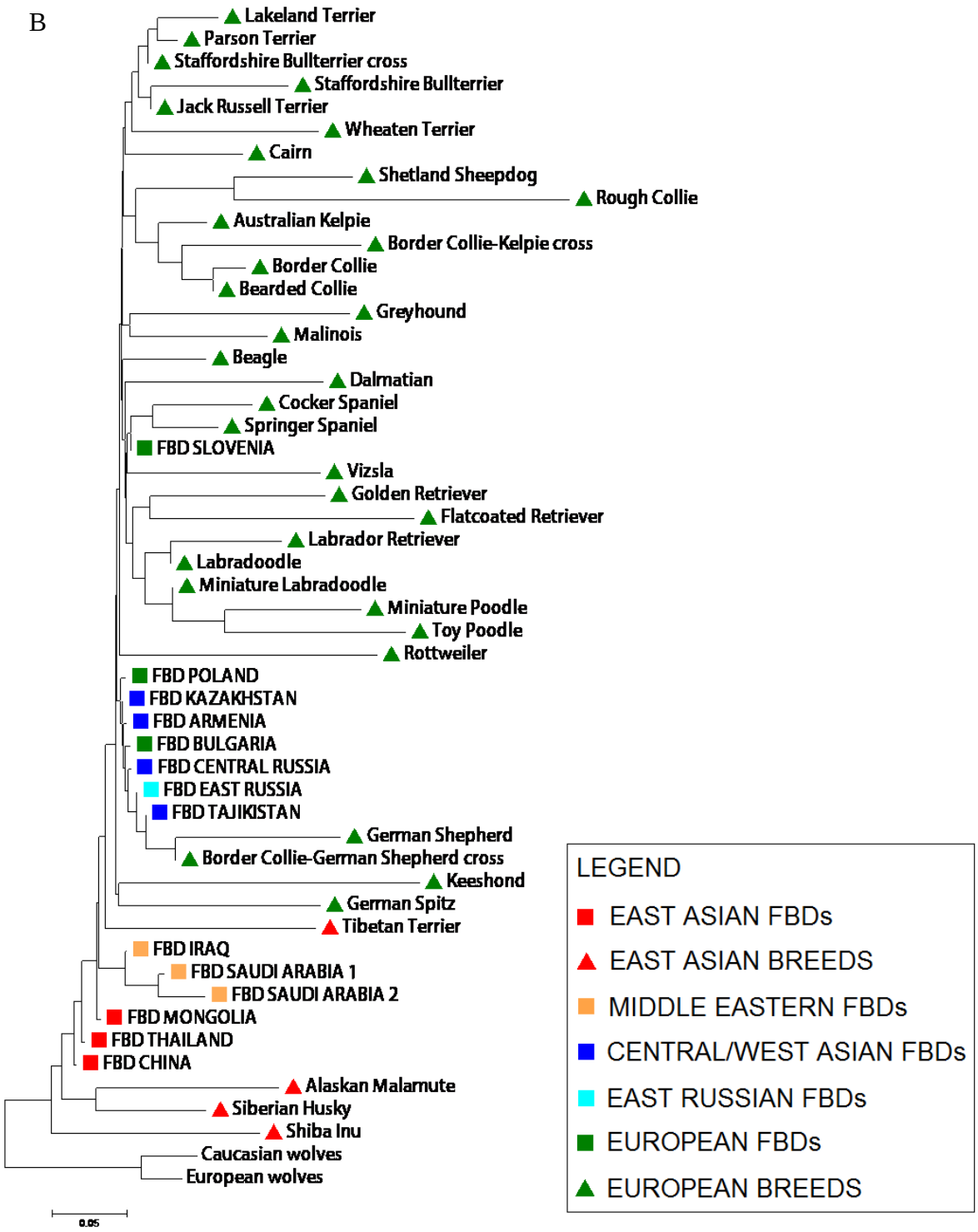
B



Supplementary Figure 9. Principal component analysis of genetic differentiation among dogs and grey wolves. (A) Including all non-related individuals from FBD, LUPA and UK databases and grey wolves (B) Based on a dataset reduced to even out the number of FBDs, "ancient" and modern dog breeds, and grey wolves. FBDs are labelled with region name.

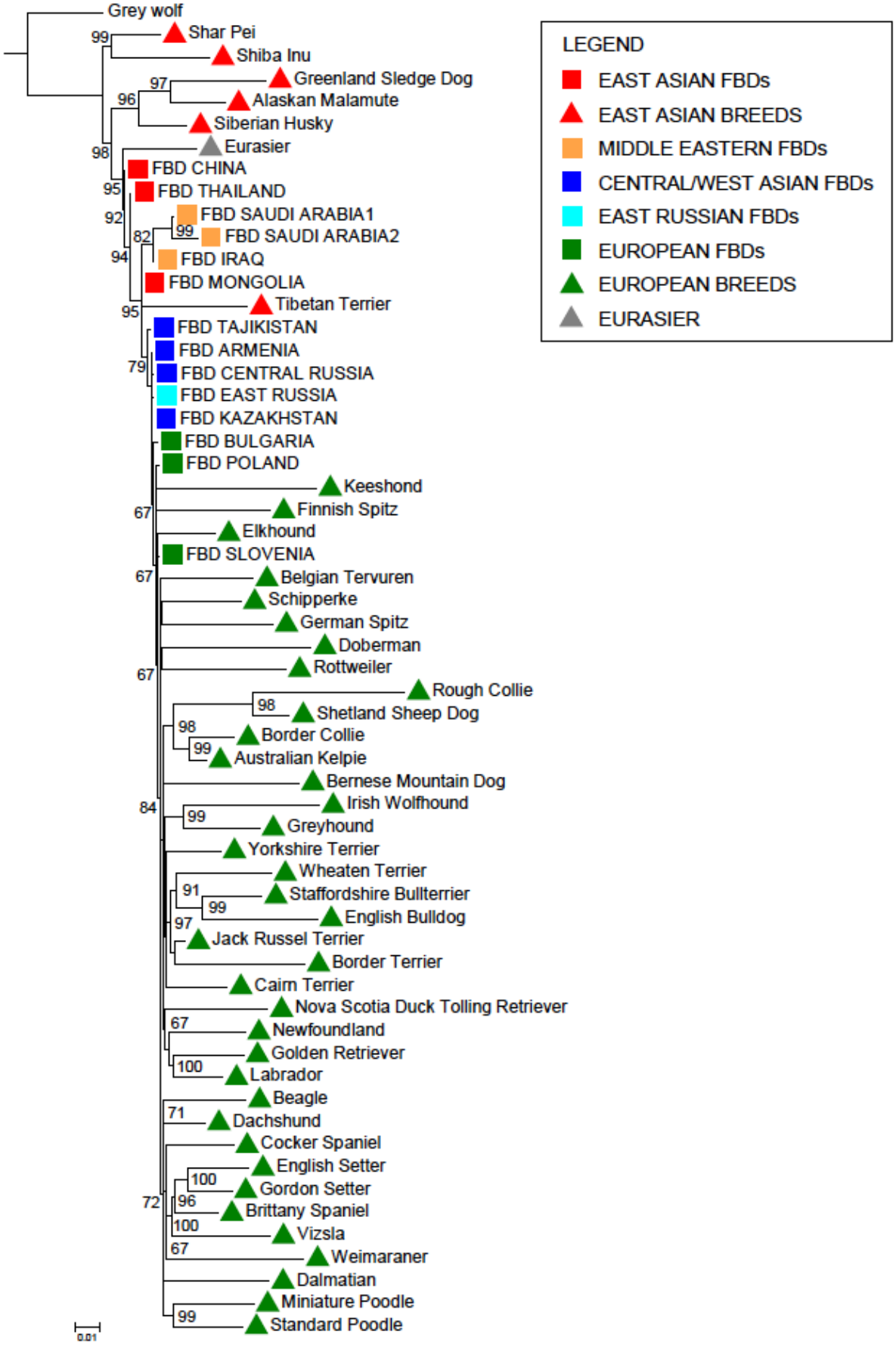
Supplementary Figure 10





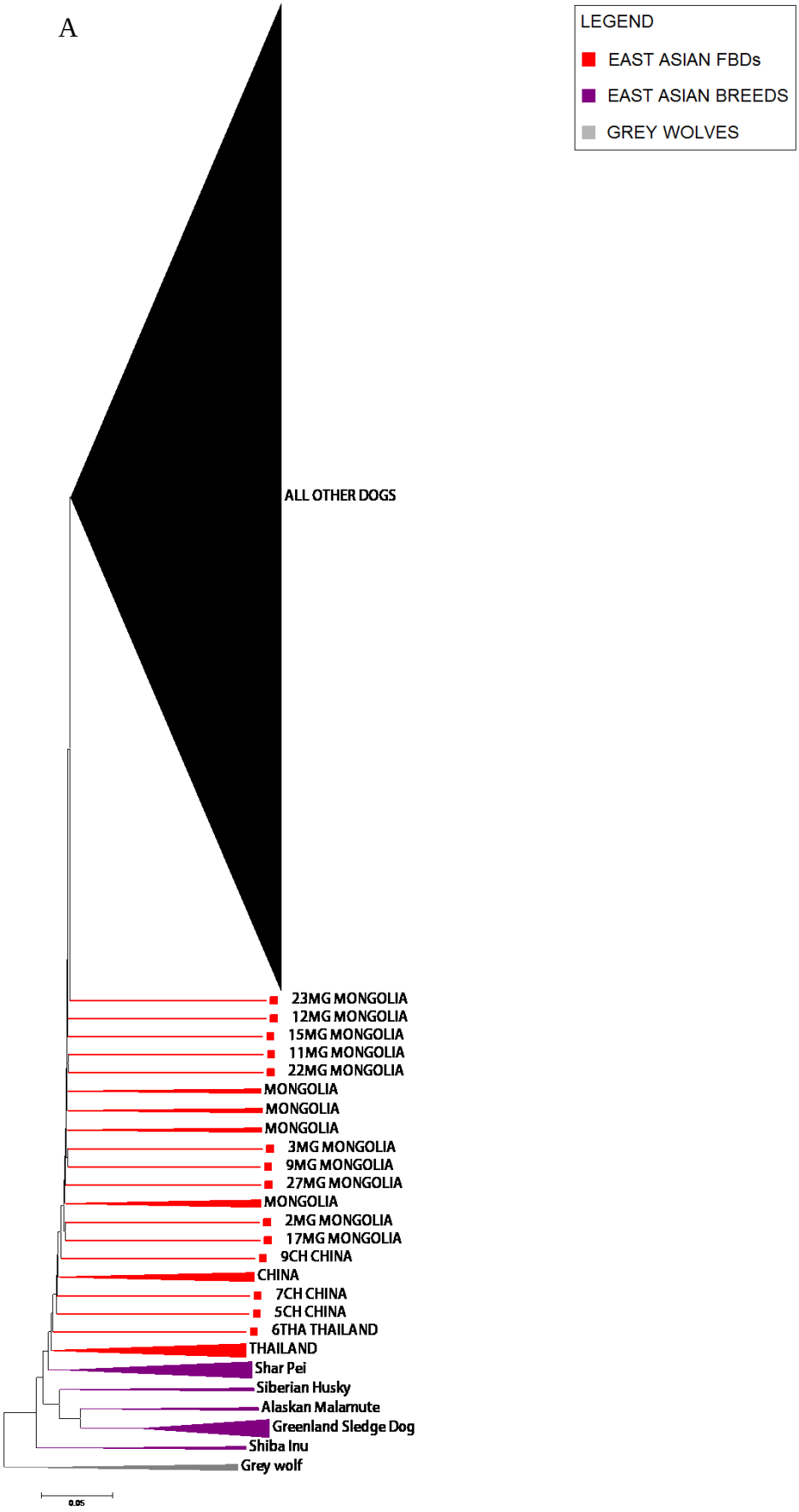
Supplementary Figure 10. Neighbour-joining tree of F_{ST} distances: (A) among FBD populations (marked with green circles) and pure breed dogs, with Caucasian wolves are used as an outgroup and (B) among FBD populations (marked with green circles) and pure breed dogs from the UK dataset. Both Caucasian and Eastern European wolves are used as an outgroup. Eastern European wolf genotypes are from Stronen *et al*. (2013). Pure breed dogs from the LUPA dataset (Vaysse *et al*. 2011) are not included here because genotype calling this dataset is incompatible with the Eastern European wolf dataset. Comparison between (A) and (B) shows that using a larger set of wolves as an outgroup does not change the tree topology.

Supplementary Figure 11

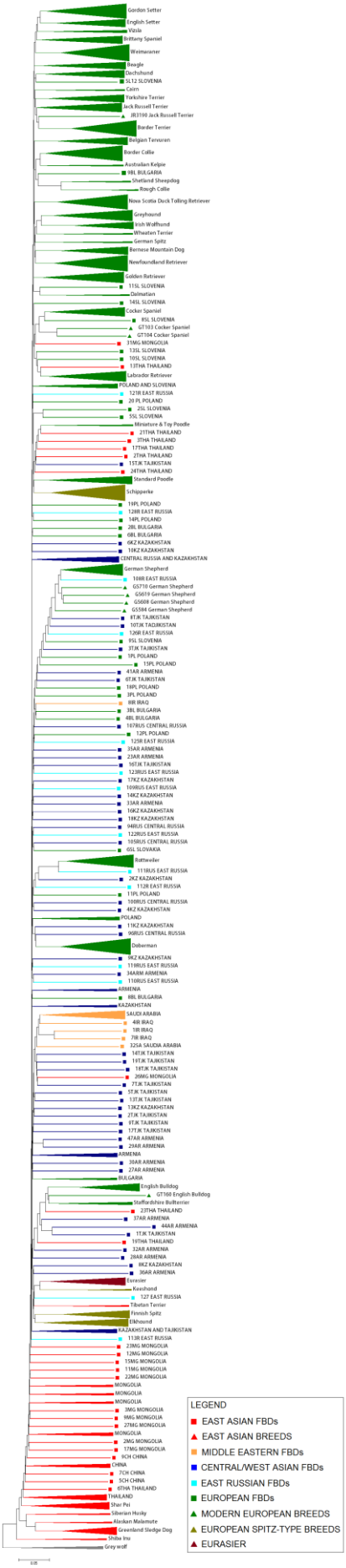


Supplementary Figure 11. Maximum likelihood tree of genetic differentiation among FBDs and breed dogs, constructed in TREEMIX without introducing the migration parameter, drawn using MEGA6 software. German shepherd was excluded from this analysis. This is the same tree as in Figure 1B in the main text, but showing the details of bootstrap support instead of presenting results of the RASP analysis.

Supplementary Figure 12



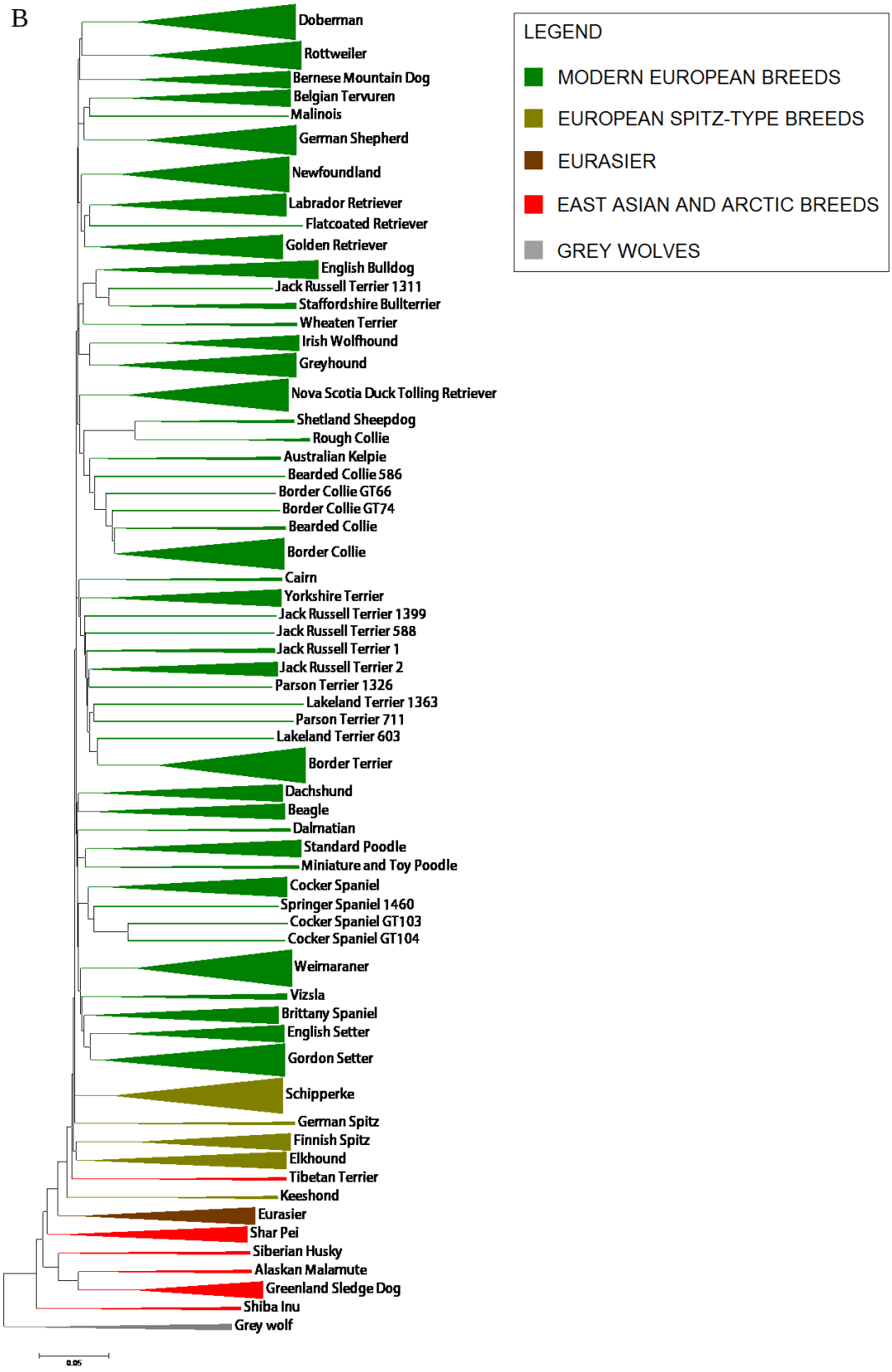
B


Gordon Setter
English Setter
Vizsla
Brittany Spaniel
Weimaraner
Beagle
Dachshund
SL12 SLOVENIA
Cairn
Yorkshire Terrier
Jack Russell Terrier
JR3190 Jack Russell Terrier
Border Terrier
Belgian Tervuren
Border Collie
Australian Kelpie
9BL BULGARIA
Shetland Sheepdog
Rough Collie
Nova Scotia Duck Tolling Retriever
Greyhound
Irish Wolfhound
Wheaten Terrier
German Spitz
Bernese Mountain Dog
Newfoundland Retriever
Golden Retriever
11SL SLOVENIA
Dalmatian
14SL SLOVENIA
Cocker Spaniel
8SL SLOVENIA
GT103 Cocker Spaniel
GT104 Cocker Spaniel
31MG MONGOLIA
13SL SLOVENIA
10SL SLOVENIA
13THA THAILAND
Labrador Retriever
POLAND AND SLOVENIA
121R EAST RUSSIA
20 PL POLAND
2SL SLOVENIA
5SL SLOVENIA
Miniature & Toy Poodle
21THA THAILAND
3THA THAILAND
17THA THAILAND
2THA THAILAND
15TJK TAJIKISTAN
24THA THAILAND
Standard Poodle
Schipperke
19PL POLAND
128R EAST RUSSIA
14PL POLAND
2BL BULGARIA
6BL BULGARIA
6KZ KAZAKHSTAN
10KZ KAZAKHSTAN
CENTRAL RUSSIA AND KAZAKHSTAN
German Shepherd
108R EAST RUSSIA
GS710 German Shepherd
GS619 German Shepherd
GS608 German Shepherd
GS584 German Shepherd
8TJK TAJIKISTAN
107JK TAOJIKISTAN
126R EAST RUSSIA
9SL SLOVENIA
3TJK TAJIKISTAN
1PL POLAND
13PL POLAND
41AR ARMENIA
6TJK TAJIKISTAN
19PL POLAND
3PL POLAND
8IR IRAQ
3BL BULGARIA
4BL BULGARIA
107RUS CENTRAL RUSSIA
12PL POLAND
125R EAST RUSSIA
3SAR ARMENIA
23AR ARMENIA
16TJK TAJIKISTAN
123RUS EAST RUSSIA
17KZ KAZAKHSTAN
109RUS EAST RUSSIA
14KZ KAZAKHSTAN
33AR ARMENIA
16KZ KAZAKHSTAN
18KZ KAZAKHSTAN
94RUS CENTRAL RUSSIA
122RUS EAST RUSSIA
105RUS CENTRAL RUSSIA
6SL SLOVAKIA
Rottweiler
111RUS EAST RUSSIA
2KZ KAZAKHSTAN
112R EAST RUSSIA
11PL POLAND
106RUS CENTRAL RUSSIA
4KZ KAZAKHSTAN
POLAND
11KZ KAZAKHSTAN
96RUS CENTRAL RUSSIA
Doberman
9KZ KAZAKHSTAN
111RUS EAST RUSSIA
34ARM ARMENIA
110RUS EAST RUSSIA
ARMENIA
8BL BULGARIA
KAZAKHSTAN
SAUDI ARABIA
4IR IRAQ
1IR IRAQ
7IR IRAQ
325A SAUDIA ARABIA
14TJK TAJIKISTAN
19TJK TAJIKISTAN
18TJK TAJIKISTAN
20MG MONGOLIA
7TJK TAJIKISTAN
5TJK TAJIKISTAN
13TJK TAJIKISTAN
13KZ KAZAKHSTAN
2TJK TAJIKISTAN
9TJK TAJIKISTAN
17TJK TAJIKISTAN
47AR ARMENIA
29AR ARMENIA
ARMENIA
36AR ARMENIA
27AR ARMENIA
BULGARIA
English Bulldog
GT168 English Bulldog
Staffordshire Bullterrier
23THA THAILAND
37AR ARMENIA
44AR ARMENIA
1TJK TAJIKISTAN
19THA THAILAND
32AR ARMENIA
29AR ARMENIA
8KZ KAZAKHSTAN
36AR ARMENIA
Eurasier
Keeshond
127 EAST RUSSIA
Tibetan Terrier
Finnish Spitz
Elkhound
KAZAKHSTAN AND TAJIKISTAN
113R EAST RUSSIA
23MG MONGOLIA
12MG MONGOLIA
15MG MONGOLIA
11MG MONGOLIA
22MG MONGOLIA
MONGOLIA
MONGOLIA
MONGOLIA
3MG MONGOLIA
9MG MONGOLIA
27MG MONGOLIA
MONGOLIA
2MG MONGOLIA
17MG MONGOLIA
8CH CHINA
CHINA
7CH CHINA
5CH CHINA
6THA THAILAND
THAILAND
Shar Pei
Siberian Husky
Alaskan Malamute
Greenland Sledge Dog
Shiba Inu
Grey wolf
LEGEND
EAST ASIAN FBDs
EAST ASIAN BREEDS
MIDDLE EASTERN FBDs
CENTRAL/WEST ASIAN FBDs
EAST RUSSIAN FBDs
EUROPEAN FBDs
MODERN EUROPEAN BREEDS
EUROPEAN SPITZ-TYPE BREEDS
EURASIER

Supplementary Figure 12. Tree of inter-individual IBS distances among free-breeding and pure-breed dogs. (A) Simplified tree with only individuals branching from basal nodes distinguished, while all the others are presented as one clade. Individuals from the same geographic location or the same breed clustering together are represented as triangles. (B) A tree showing details of this large clade, which is also presented in a separate file due to its large size. Breeds of different origin are distinguished with colours.

Supplementary Figure 13

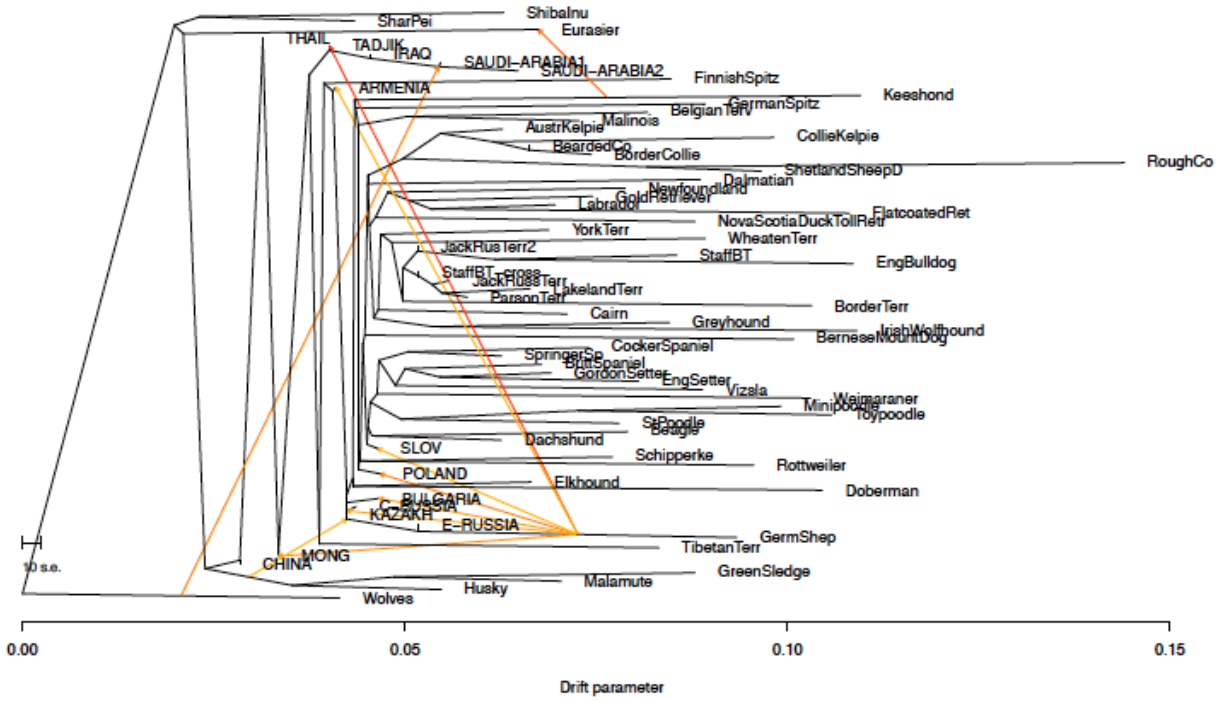
A



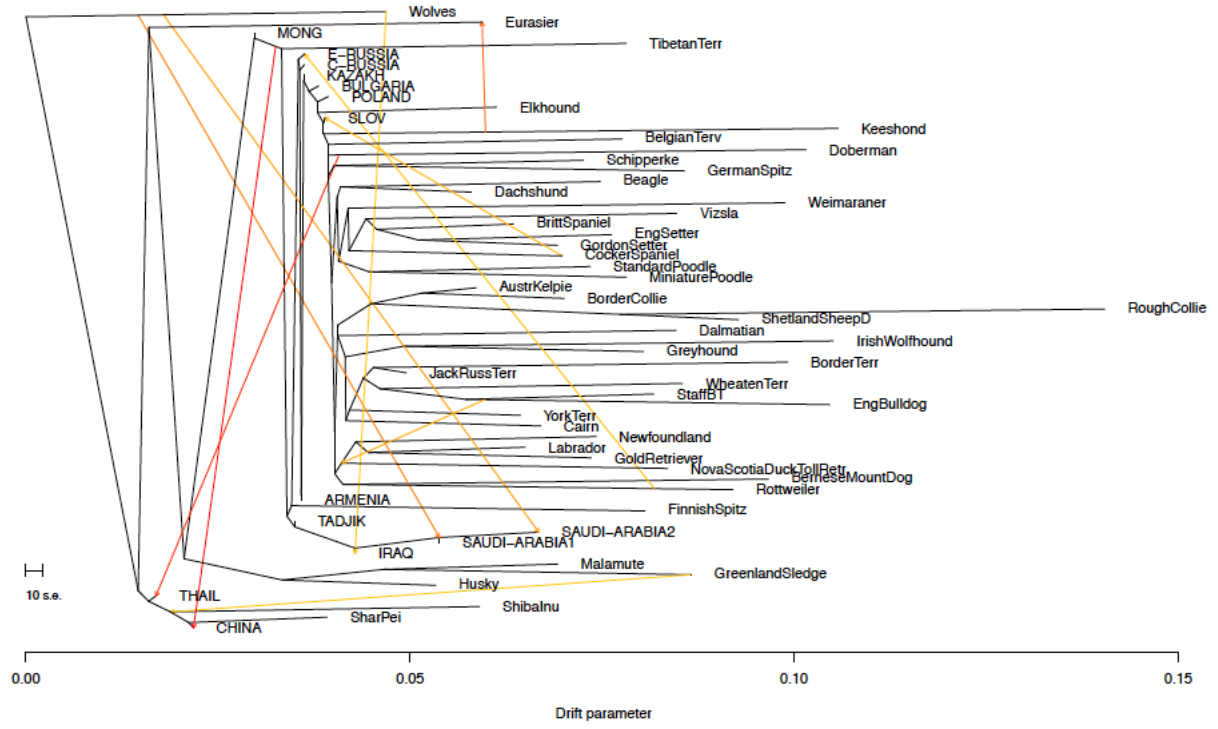
Supplementary Figure 13. Tree of IBS inter-individual distances among pure-breed dogs. (A) Simplified tree distinguishing only breeds branching from basal nodes in the tree, while all other breeds (all of which are of European origin) are grouped into one cluster. (B) A tree showing details of this large clade, with individuals from the same breed clustering together represented as triangles. Breeds of different origin are distinguished with colours.

Supplementary Figure 14

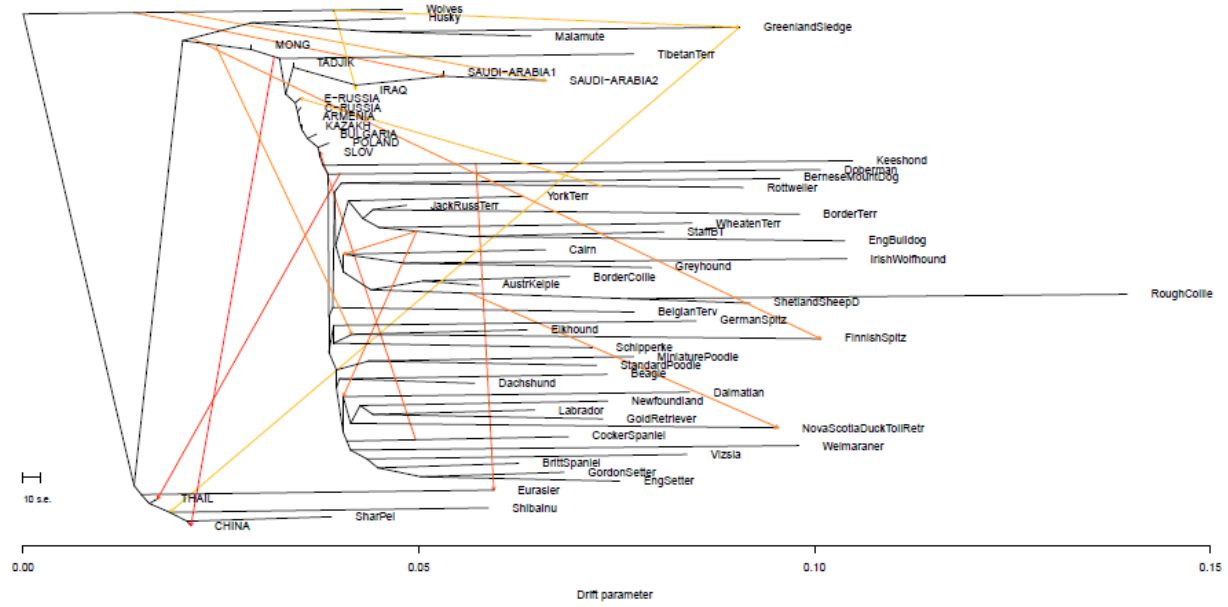
A



B

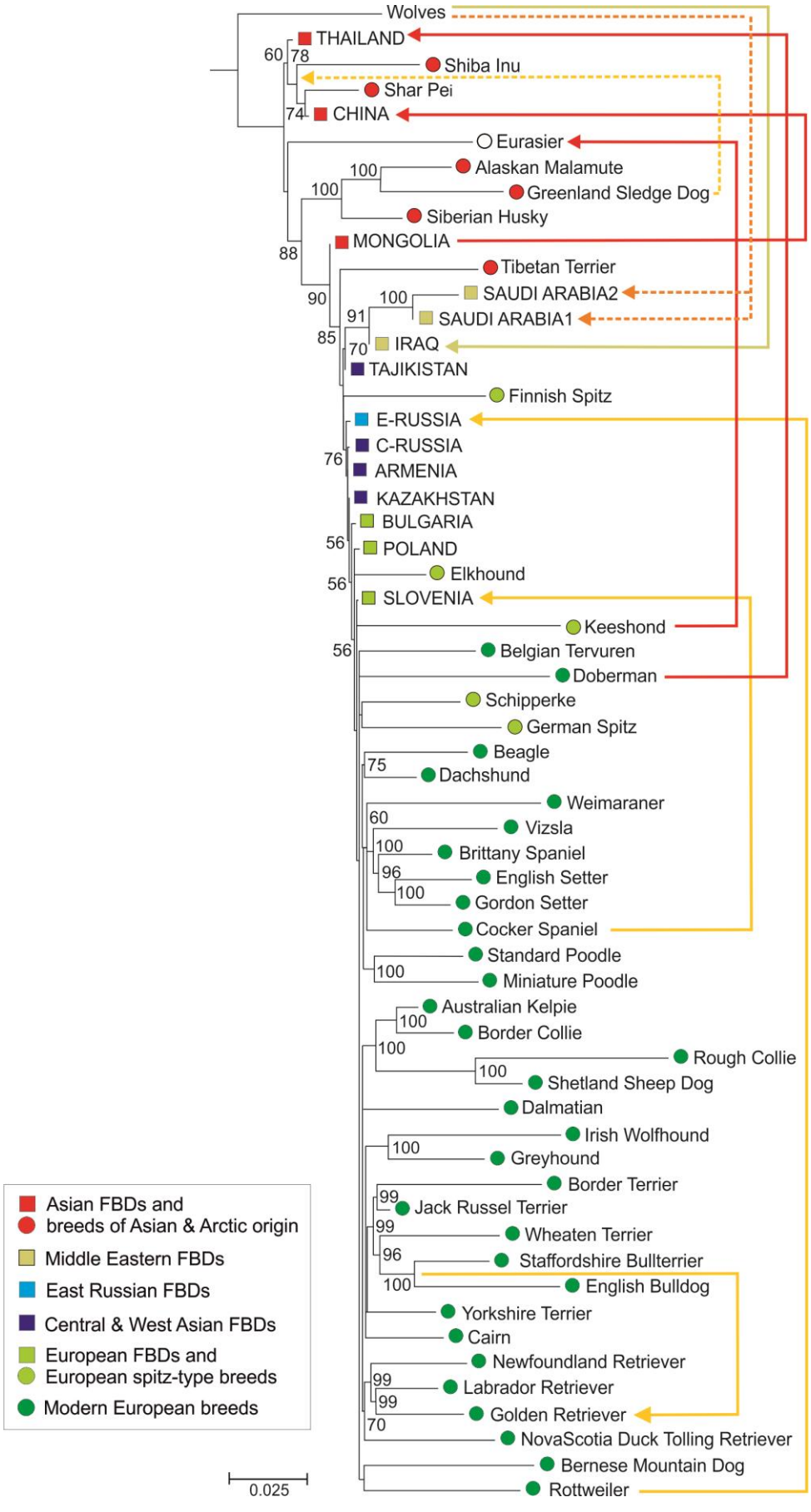


C



Supplementary Figure 14. Maximum likelihood tree of genetic differentiation among FBDs and breed dogs, constructed in TREEMIX (A) assuming 10 events of post-divergence gene flow and including German shepherd, (B) assuming 10 events of post-divergence gene flow and without German shepherd, (C) assuming 15 events of post-divergence gene flow and without German shepherd. The colour of arrows represents gene flow intensity, increasing from light yellow to red. The figures were created using an R script distributed with TREEMIX software.

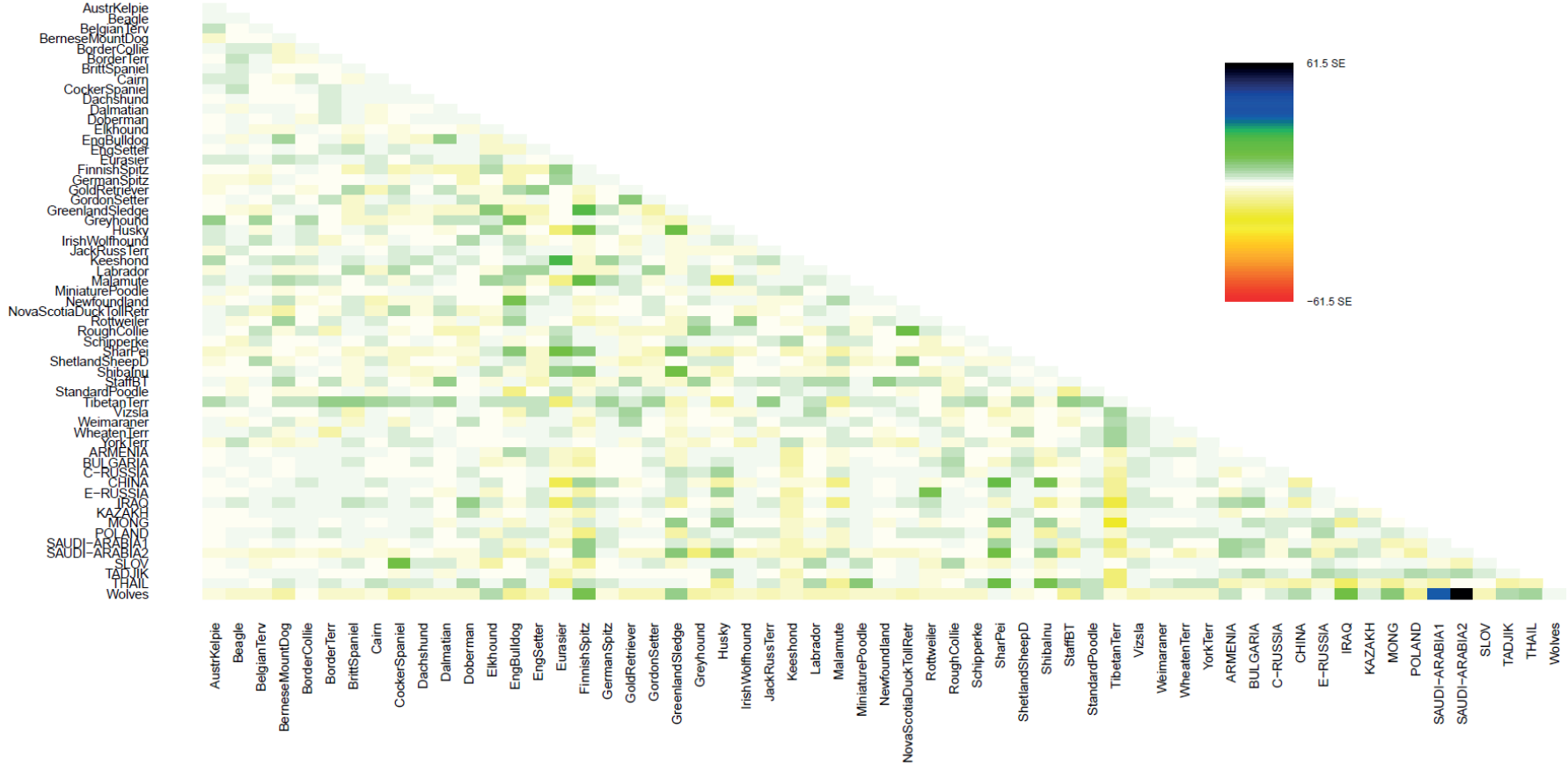
Supplementary Figure 15



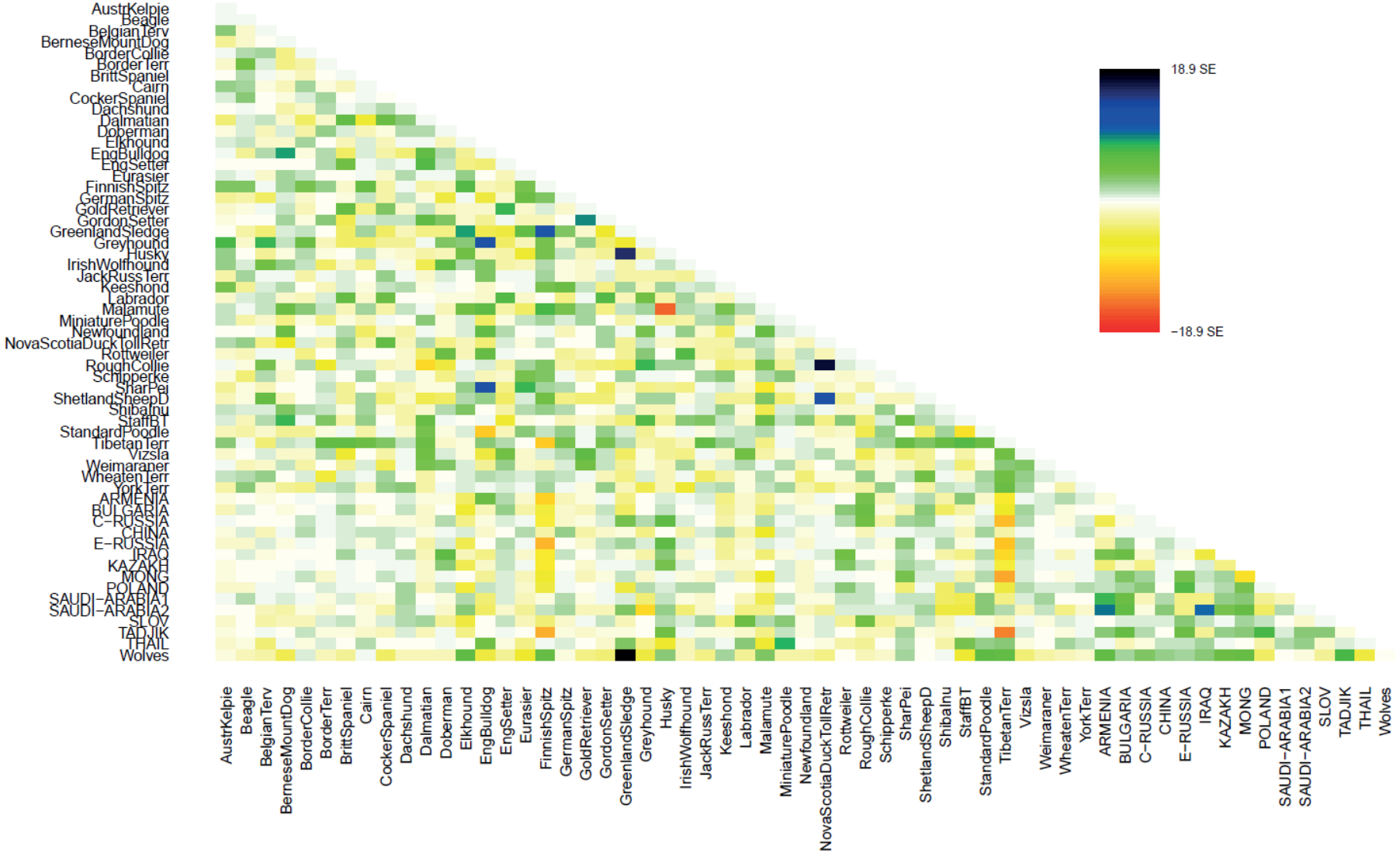
Supplementary Figure 15. Maximum likelihood tree of genetic differentiation among FBDs and pure-breed dogs, constructed in TREEMIX, assuming 10 gene flow events. The colours of arrows represent gene flow intensity, increasing from light yellow to red. P-values of all the inferred migration edges are below 0.0004. Bootstrap support for nodes is marked if above 50%. Bootstrap support for some groups of pure-breed dogs is not shown because of space constraints in the figure, but this information is non-essential for this study. Bootstrap support for the clade grouping Chinese and Thai FBDs, Shar Pei and Shiba Inu is relatively low, because in some replicates this clade also included Eurasier (a breed of mixed East Asian and European origin; trees representing bootstrap replicates are available upon request). Gene flow from wolves to Arctic dog breeds is only detected when 15 gene flow events are assumed (see Supplementary Fig. 14C).

Supplementary Figure 16

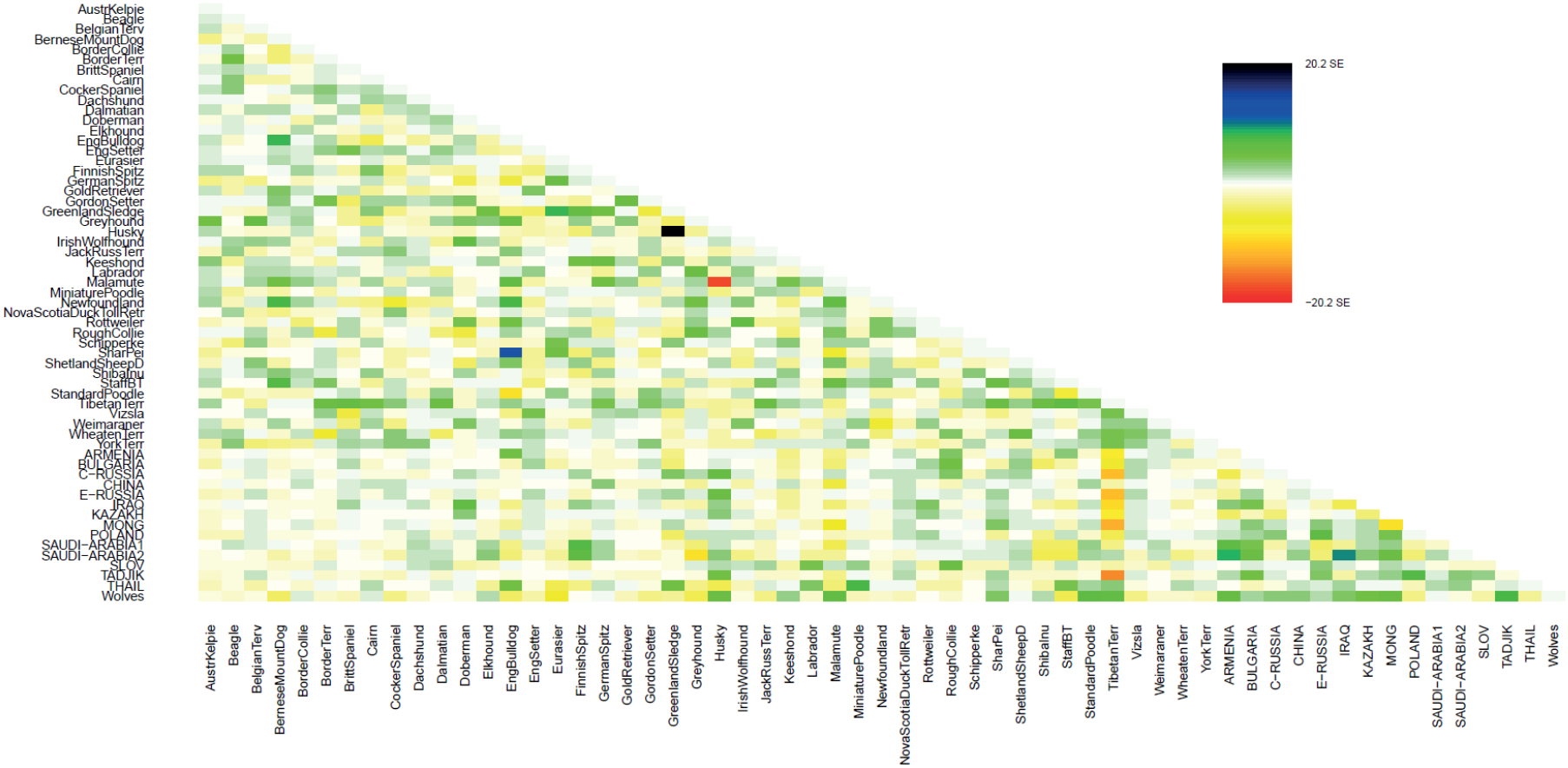
A



B


AustrKelpie
Beagle
BelgianTerv
BerneseMountDog
BorderCollie
BorderTerr
BrittSpaniel
Cairn
CockerSpaniel
Dachshund
Dalmatian
Doberman
Elkhound
EngBulldog
EngSetter
Eurasier
FinnishSpitz
GermanSpitz
GoldRetriever
GordonSetter
GreenlandSledge
Greyhound
Husky
IrishWolfhound
JackRussTerr
Keeshond
Labrador
Malamute
MiniaturePoodle
Newfoundland
NovaScotiaDuckTollRetr
Rottweiler
RoughCollie
Schipperke
SharPei
ShetlandSheepD
ShibaInu
StaffBT
StandardPoodle
TibetanTerr
Vizsla
Weimaraner
WheatenTerr
YorkTerr
ARMENIA
BULGARIA
C−RUSSIA
CHINA
E−RUSSIA
IRAQ
KAZAKH
MONG
POLAND
SAUDI−ARABIA1
SAUDI−ARABIA2
SLOV
TADJIK
THAIL
Wolves
18.9 SE
−18.9 SE

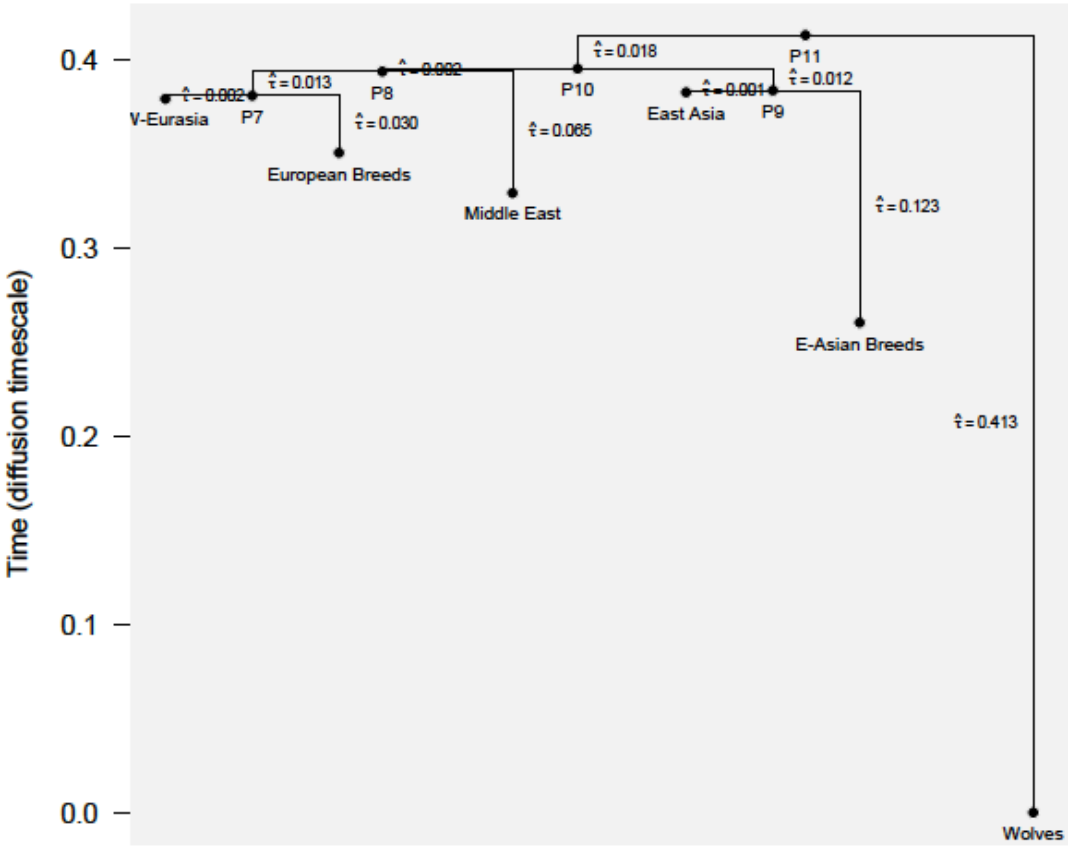
C


AustrKelpie
Beagle
BelgianTerv
BerneseMountDog
BorderCollie
BorderTerr
BrittSpaniel
Cairn
CockerSpaniel
Dachshund
Dalmatian
Doberman
Elkhound
EngBulldog
EngSetter
Eurasier
FinnishSpitz
GermanSpitz
GoldRetriever
GordonSetter
GreenlandSledge
Greyhound
Husky
IrishWolfhound
JackRussTerr
Keeshond
Labrador
Malamute
MiniaturePoodle
Newfoundland
NovaScotiaDuckTollRetr
Rottweiler
RoughCollie
Schipperke
SharPei
ShetlandSheepD
Shibalnu
StaffBT
StandardPoodle
TibetanTerr
Vizsla
Weimaraner
WheatenTerr
YorkTerr
ARMENIA
BULGARIA
C-RUSSIA
CHINA
E-RUSSIA
IRAQ
KAZAKH
MONG
POLAND
SAUDI-ARABIA1
SAUDI-ARABIA2
SLOV
TADJIK
THAIL
Wolves
20.2 SE
-20.2 SE

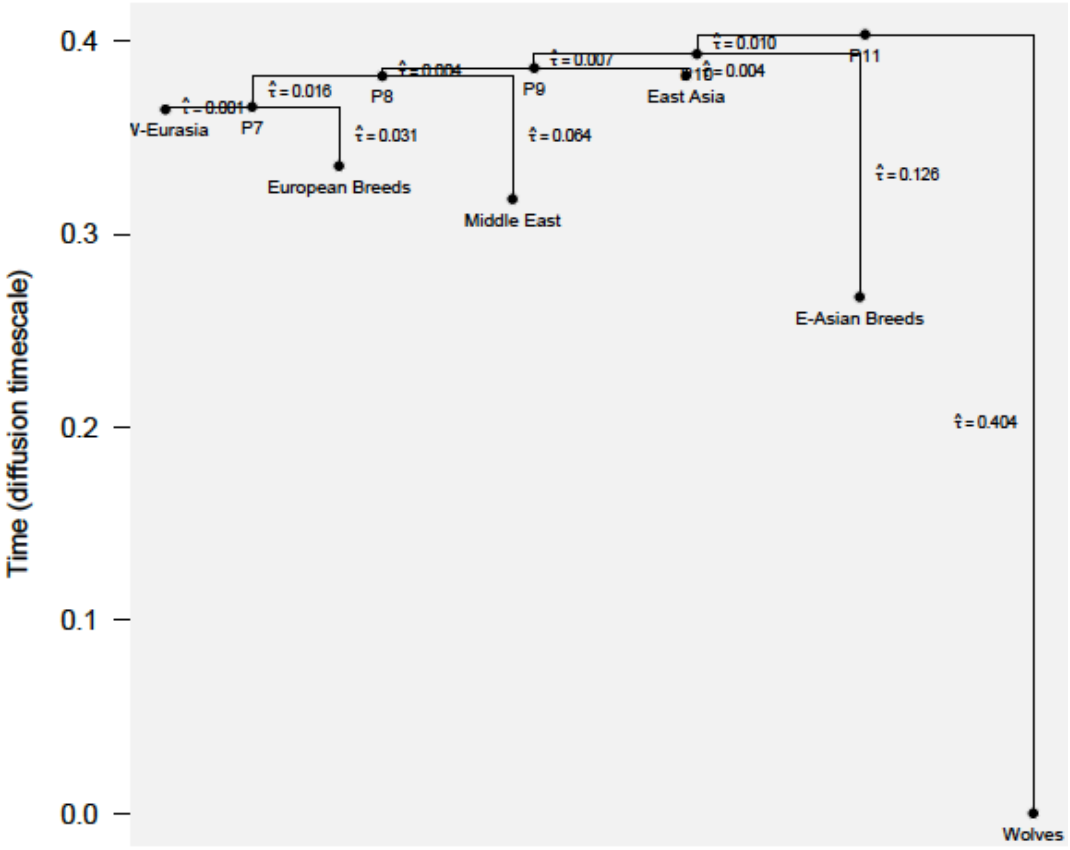
Supplementary Figure 16. Plots of residual fit from the TREEMIX maximum likelihood trees: (A) without gene flow events, (B) with 10 gene flow events, (C) with 15 gene flow events. These plots correspond to the TREEMIX trees from Supplementary Figure 11 and Supplementary Figure 14 B and C, respectively. Population pairs with residuals higher than zero are more closely related than it is represented in the tree, and therefore populations with high positive residuals are candidates for additional admixture events (Pickrell & Pritchard 2012).
The breeds and FBD populations included in the plot are as follows: Australian Kelpie, Beagle, Belgian Tervuren, Bernese Mountain Dog, Border Collie, Border Terrier, Brittany Spaniel, Cairn, Cocker Spaniel, Dachshund, Dalmatian, Doberman, Elkhound, English Bulldog, English Setter, Eurasier, Finnish Spitz, German Spitz, Golden Retriever, Gordon Setter, Greenland Sledge Dog, Greyhound, Siberian Husky, Irish Wolfhound, Jack Russel Terrier, Keeshond, Labrador Retriever, Alaskan Malamute, Miniature Poodle, Newfoundland Retriever, Nova Scotia Duck Tolling Retriever, Rottweiler, Rough Collie, Schipperke, Shar Pei, Shetland Sheep Dog, Shiba Inu, Staffordshire Bull Terrier, Standard Poodle, Tibetan Terrier, Vizsla, Weimaraner, Wheaten Terrier, Yorkshire Terrier, FBD-Armenia, FBD-Bulgaria, FBD-Central Russia, FBD-China, FBD-East Russia, FBD-Iraq, FBD-Kazakhstan, FBD-Mongolia, FBD-Poland, FBD-Saudi Arabia 1, FBD-Saudi Arabia2, FBD-Slovenia, FBD-Tajikistan, FBD-Thailand, and grey wolves.

Supplementary Figure 17

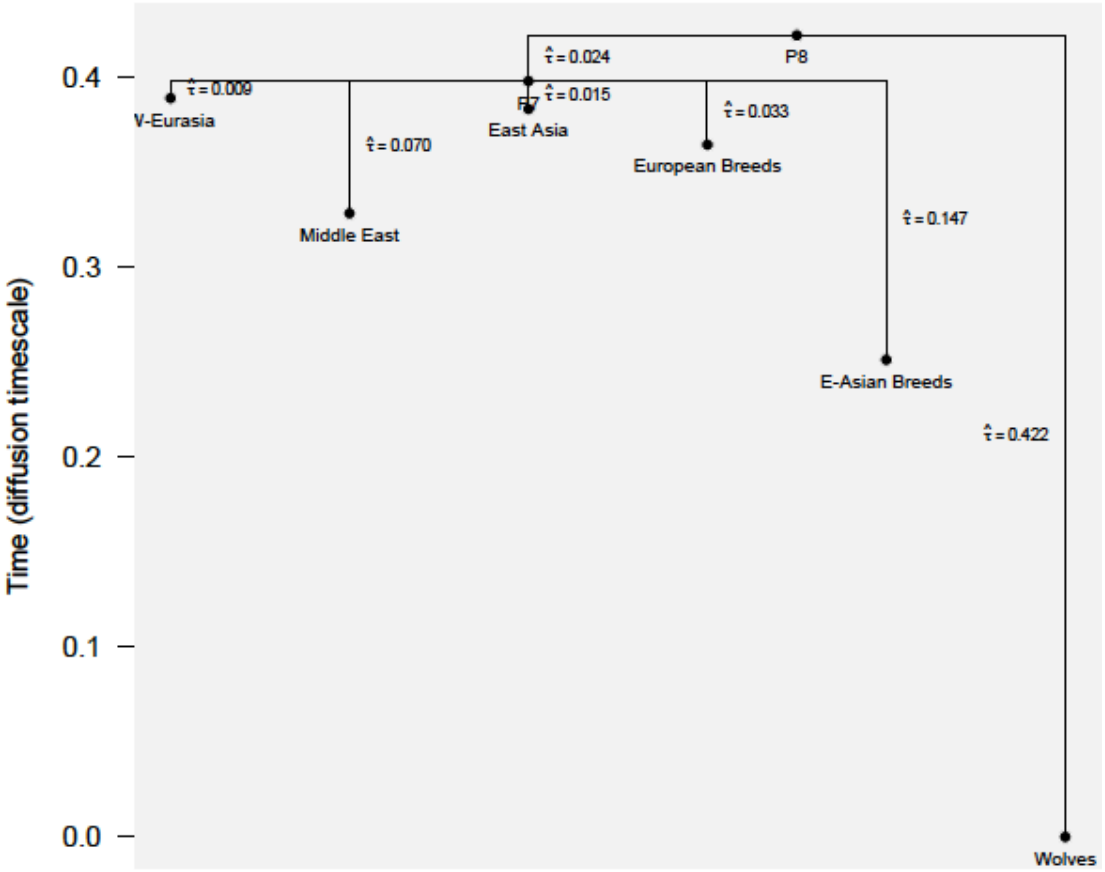
A1. DIC_{A1} = 1,817,414



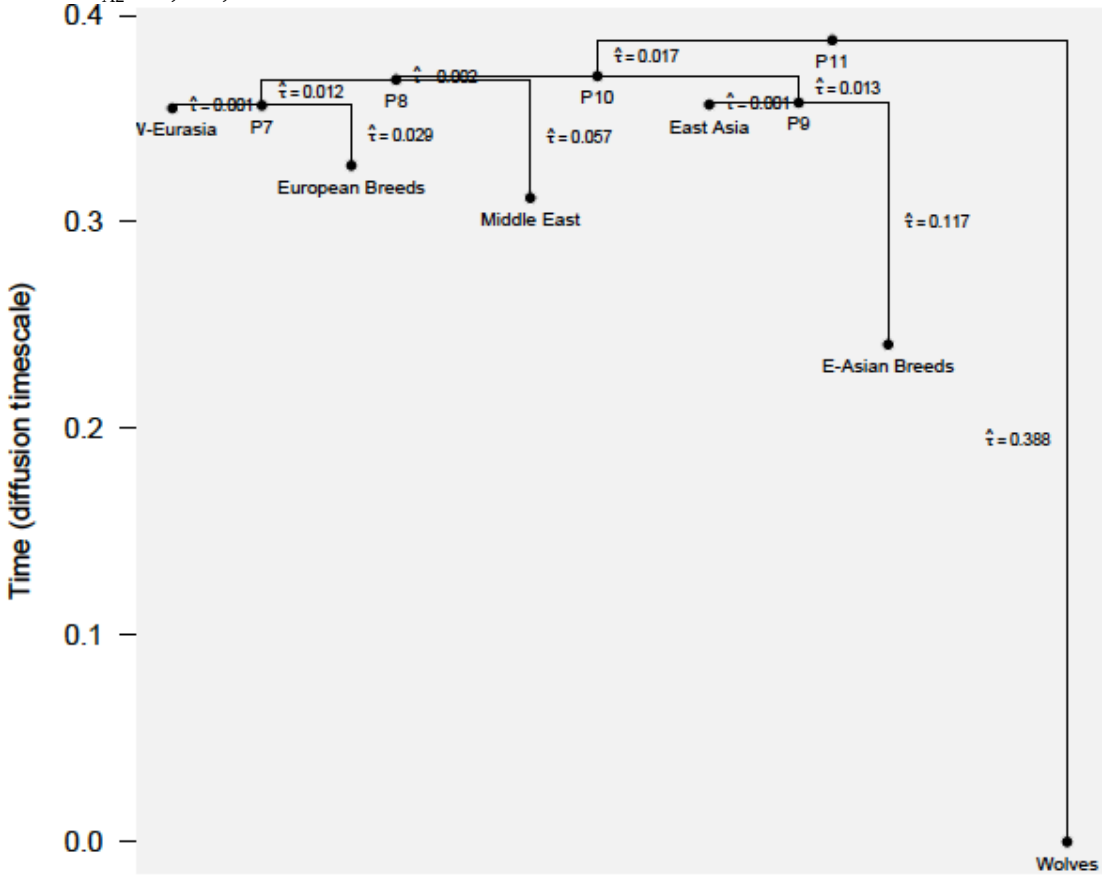
B1. DIC_{B1} =1,818,596; $DIC_{B1} - DIC_{A1}$ = 1,183



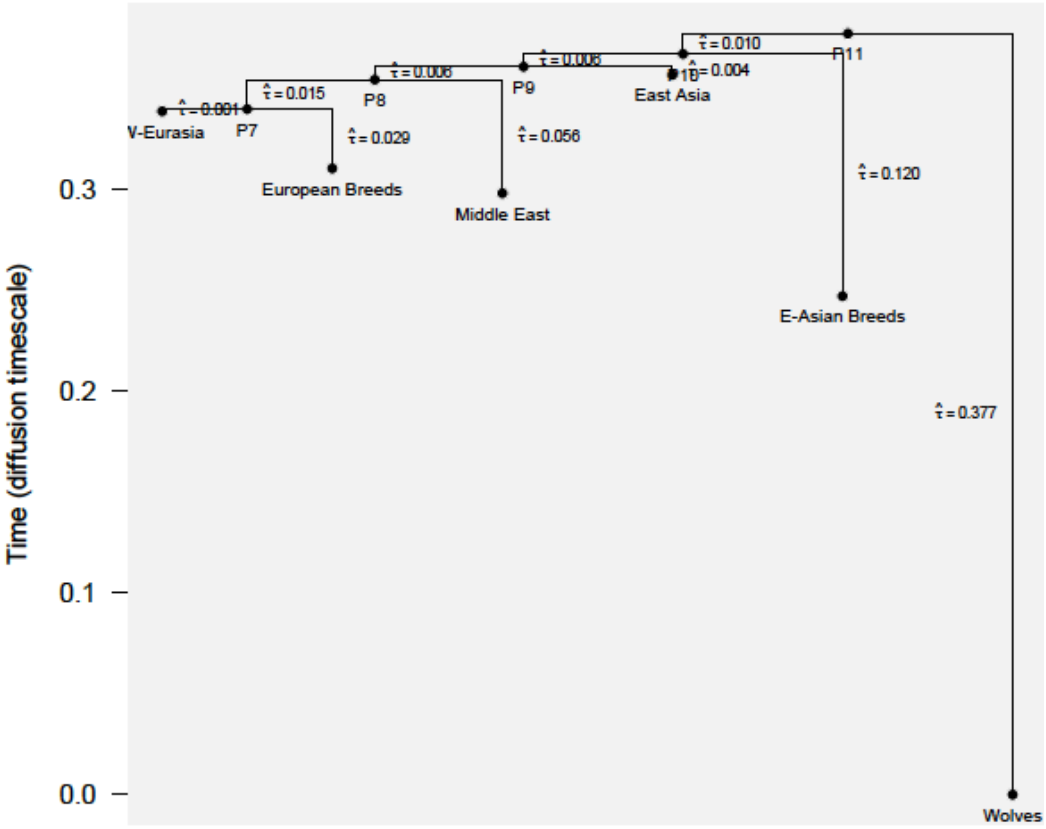
C1. DIC_{C1} = 1,822,605; $DIC_{C1} - DIC_{A1}$ = 5,191



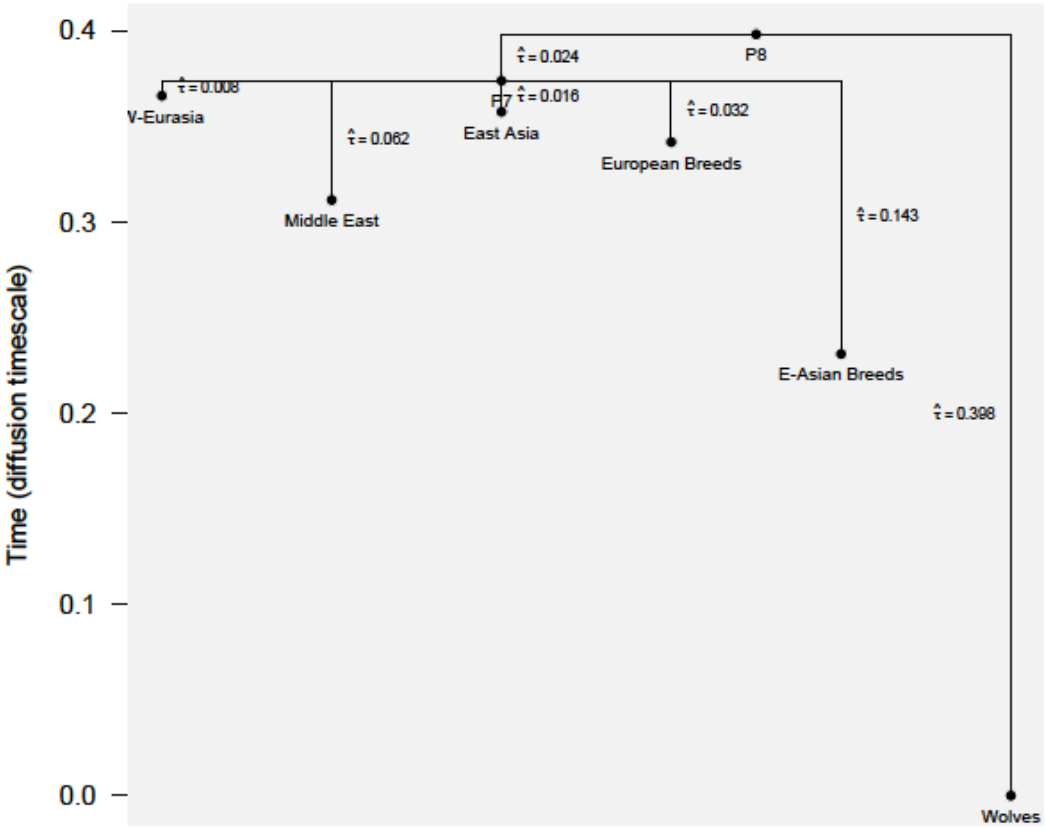
A2. DIC_{A2} = 1,824,530



B2. DIC_{B2} = 1,825,667; $DIC_{B2} - DIC_{A2}$ = 1,137



C1. DIC_{C2} = 1,830,211; $DIC_{C2} - DIC_{A2}$ = 5,681

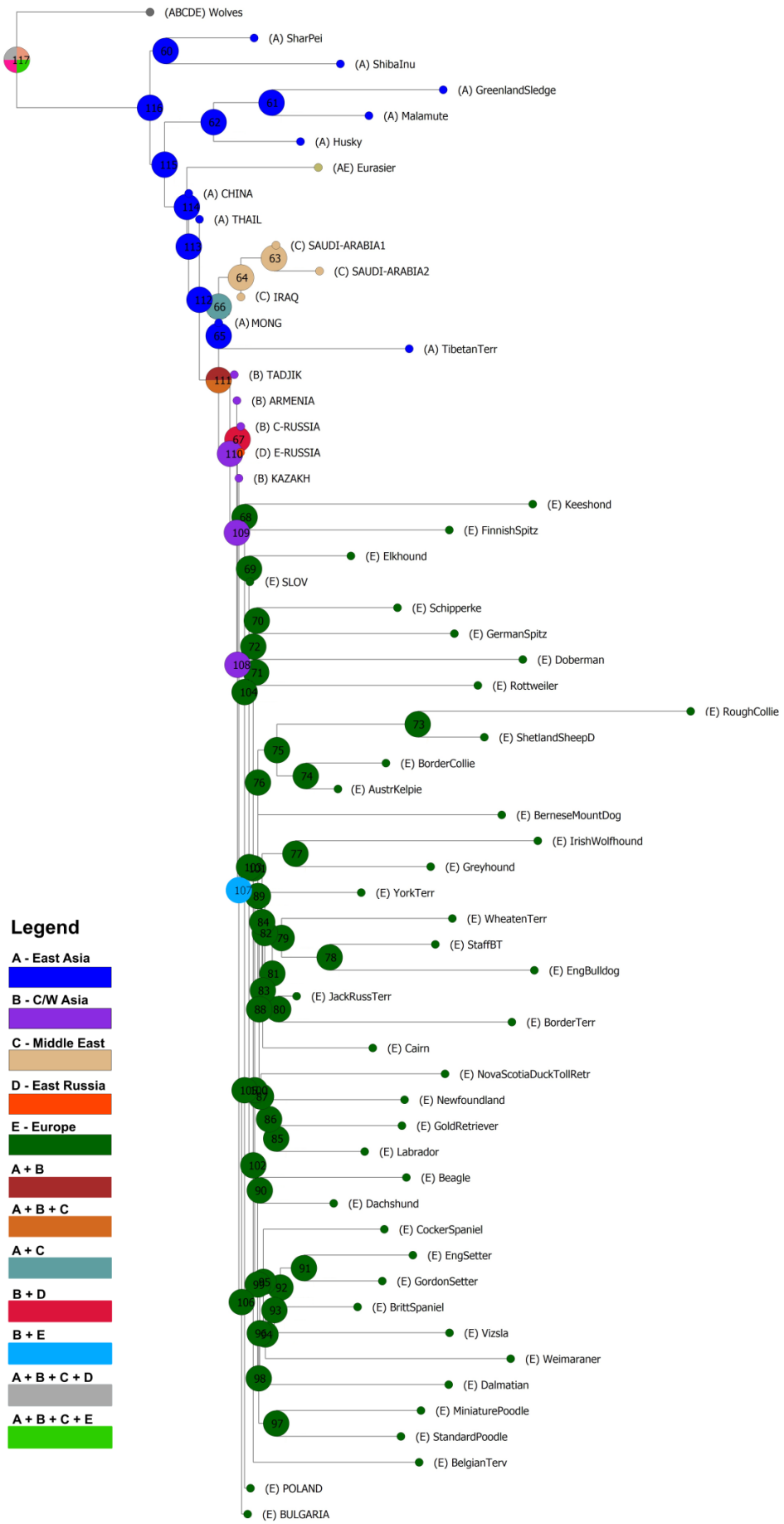


W-Eurasia: FBDs from Europe and West/Central Asia, Middle East: FBDs from the Middle East, East Asia: FBDs from East Asia, E-Asian Breeds: dog breeds originating from East Asia.

Supplementary Figure 17. Tree topologies inferred in KIMTREE, based on two sets of SNPs originating from the equal division of the original SNP set (made due to computational issues). Pairs of trees (A1 and A2), (B1 and B2), (C1 and C2) represent three different topologies analysed for two sets of SNPs. Trees (A1) and (A2) represent the topology having the strongest support based on the DIC values. The branch lengths in these trees are estimated as $\tau \approx T/(2N_E)$, where N_E is the effective size of the population the branch leads to, and T is time (in generations). The estimated τ values are provided for each branch. The most likely tree topology is identified using the deviance information criterion (DIC).

Supplementary Figure 18

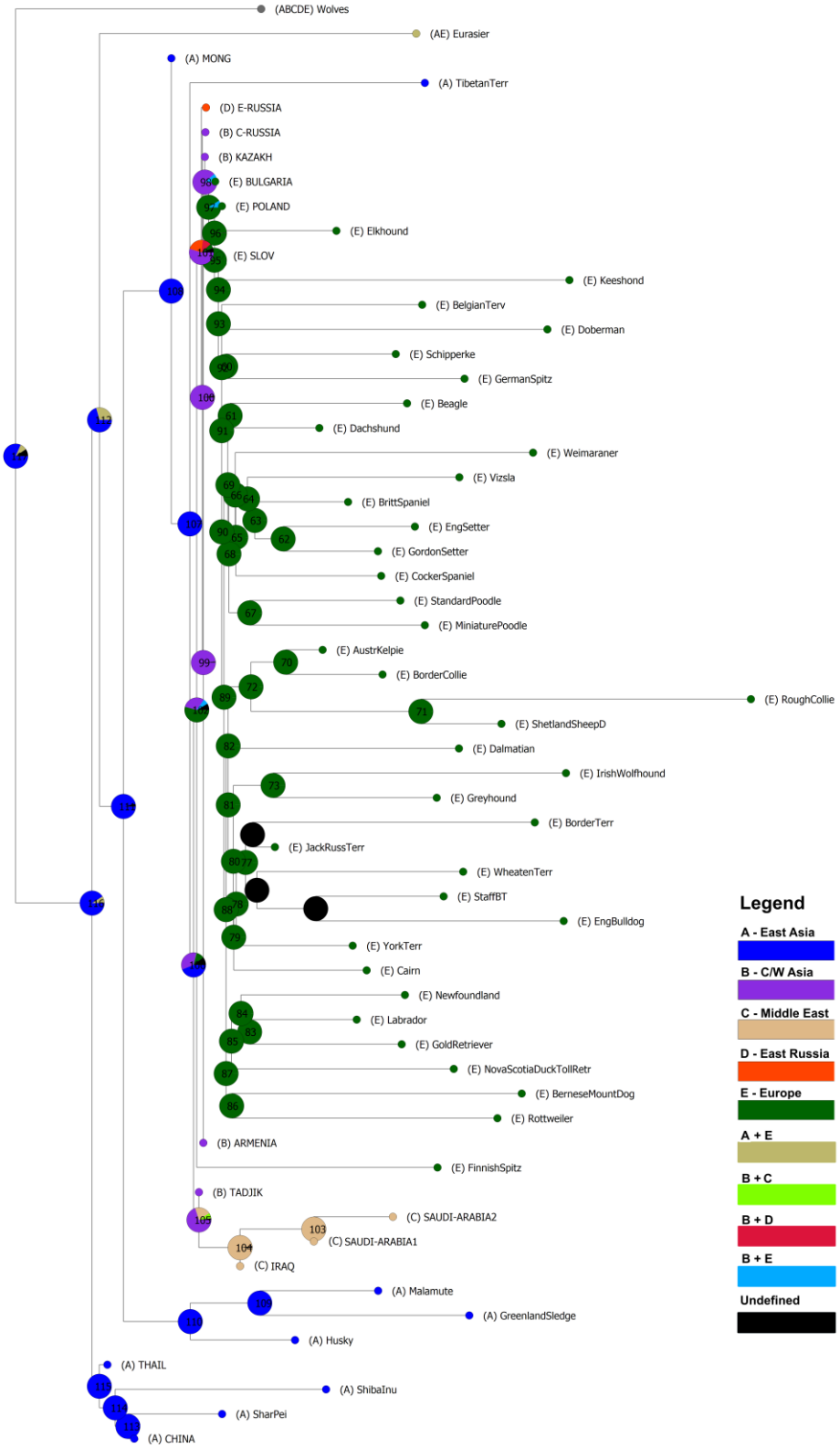
B


(ABCDE) Wolves
(A) SharPei
(A) ShibaInu
(A) GreenlandSledge
(A) Malamute
(A) Husky
(AE) Eurasier
(A) CHINA
(A) THAIL
(C) SAUDI-ARABIA1
(C) SAUDI-ARABIA2
(C) IRAQ
(A) MONG
(A) TibetanTerr
(B) TADJIK
(B) ARMENIA
(B) C-RUSSIA
(D) E-RUSSIA
(B) KAZAKH
(E) Keeshond
(E) FinnishSpitz
(E) Elkhound
(E) SLOV
(E) Schipperke
(E) GermanSpitz
(E) Doberman
(E) Rottweiler
(E) RoughCollie
(E) ShetlandSheepD
(E) BorderCollie
(E) AustrKelpie
(E) BerneseMountDog
(E) IrishWolfhound
(E) Greyhound
(E) YorkTerr
(E) WheatenTerr
(E) StaffBT
(E) EngBulldog
(E) JackRussTerr
(E) BorderTerr
(E) Cairn
(E) NovaScotiaDuckTollRetr
(E) Newfoundland
(E) GoldRetriever
(E) Labrador
(E) Beagle
(E) Dachshund
(E) CockerSpaniel
(E) EngSetter
(E) GordonSetter
(E) BrittSpaniel
(E) Vizsla
(E) Weimaraner
(E) Dalmatian
(E) MiniaturePoodle
(E) StandardPoodle
(E) BelgianTerv
(E) POLAND
(E) BULGARIA
Legend
A - East Asia
B - C/W Asia
C - Middle East
D - East Russia
E - Europe
A + B
A + B + C
A + C
B + D
B + E
A + B + C + D
A + B + C + E

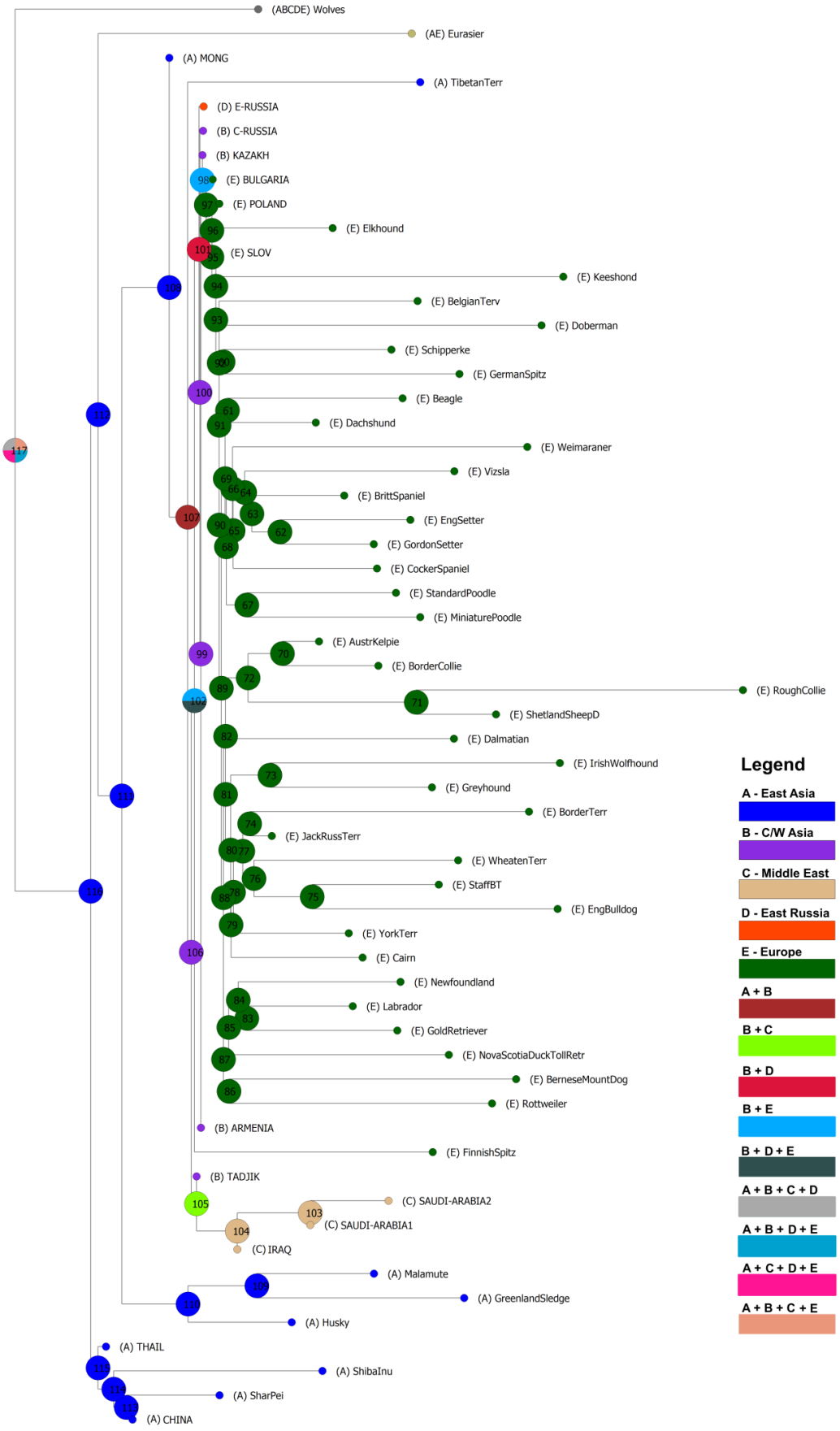
Supplementary Figure 18. Biogeographic reconstruction of the distribution of ancestral dog populations inferred using the RASP software, based on dog phylogeny created in TREEMIX assuming no post-divergence gene flow (A) with uncertainty assessed using Bayesian Binary MCMC method; this is the same tree as in Figure 1B, presented with more details; (B) with uncertainty assessed using S-Diva method. Distribution of ancestral dog populations is marked on nodes using colour-codes. Arctic breeds were assigned to East Asia, according to their primary origin (Brown *et al*. 2013, 2015, van Asch *et al*. 2013).

Supplementary Figure 19

A



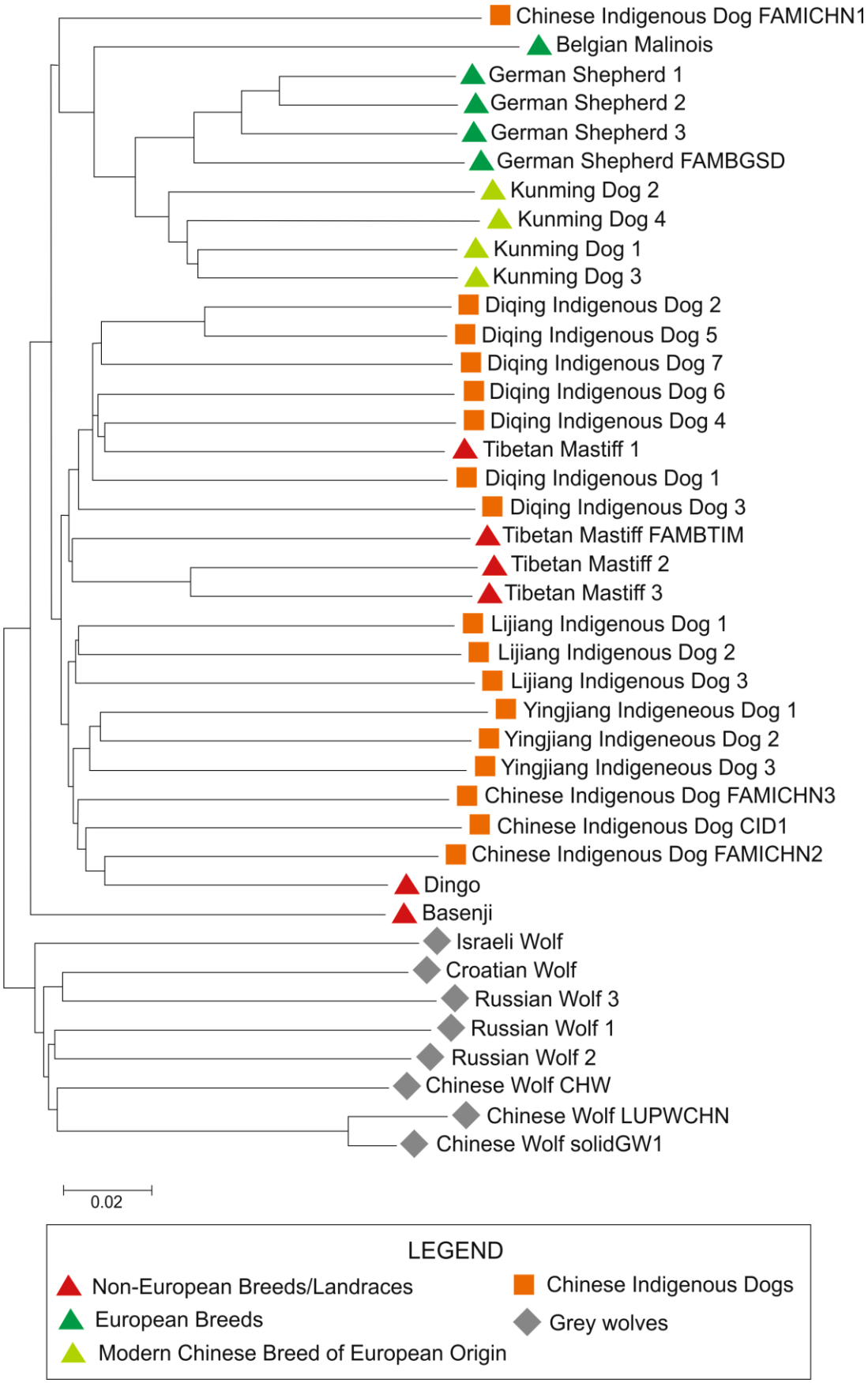
B


(ABCDE) Wolves
(AE) Eurasier
(A) MONG
(A) TibetanTerr
(D) E-RUSSIA
(B) C-RUSSIA
(B) KAZAKH
(E) BULGARIA
(E) POLAND
(E) Elkhound
(E) SLOV
(E) Keeshond
(E) BelgianTerv
(E) Doberman
(E) Schipperke
(E) GermanSpitz
(E) Beagle
(E) Dachshund
(E) Weimaraner
(E) Vizsla
(E) BrittSpaniel
(E) EngSetter
(E) GordonSetter
(E) CockerSpaniel
(E) StandardPoodle
(E) MiniaturePoodle
(E) AustrKelpie
(E) BorderCollie
(E) RoughCollie
(E) ShetlandSheepD
(E) Dalmatian
(E) IrishWolfhound
(E) Greyhound
(E) BorderTerr
(E) JackRussTerr
(E) WheatenTerr
(E) StaffBT
(E) EngBulldog
(E) YorkTerr
(E) Cairn
(E) Newfoundland
(E) Labrador
(E) GoldRetriever
(E) NovaScotiaDuckTollRetr
(E) BerneseMountDog
(E) Rottweiler
(B) ARMENIA
(E) FinnishSpitz
(B) TADJIK
(C) SAUDI-ARABIA2
(C) SAUDI-ARABIA1
(C) IRAQ
(A) Malamute
(A) GreenlandSledge
(A) Husky
(A) THAIL
(A) ShibaInu
(A) SharPei
(A) CHINA
Legend
A - East Asia
B - C/W Asia
C - Middle East
D - East Russia
E - Europe
A + B
B + C
B + D
B + E
B + D + E
A + B + C + D
A + B + D + E
A + C + D + E
A + B + C + E

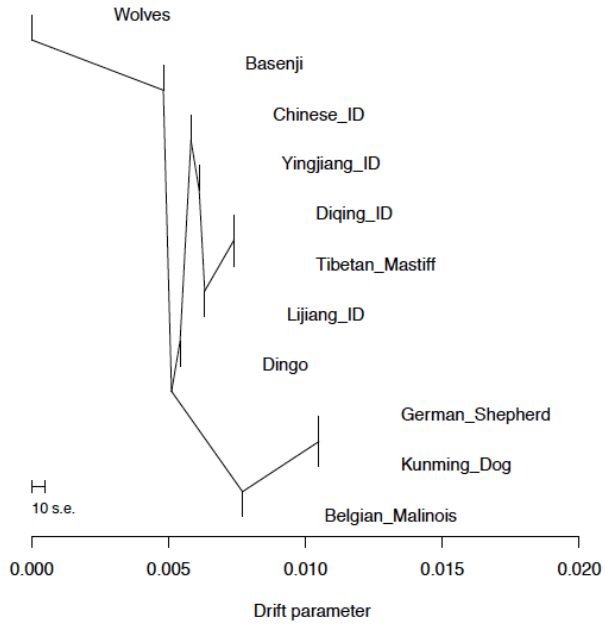
Supplementary Figure 19. Biogeographic reconstruction of the distribution of ancestral dog populations inferred using the RASP software, based on dog phylogeny created in TREEMIX assuming 10 events of gene flow (A) with uncertainty assessed using Bayesian Binary MCMC method, (B) with uncertainty assessed using S-Diva method. Distribution of ancestral dog populations is marked on nodes using colour-codes. Arctic breeds were assigned to East Asia, according to their primary origin (Brown *et al*. 2013, 2015, van Asch *et al*. 2013). The effect of the presence of a breed of admixed East Asian and European origin (Eurasier) is visible as an ambiguity of the geographical assignment of the respective nod to either East Asia or Europe.

Supplementary Figure 20

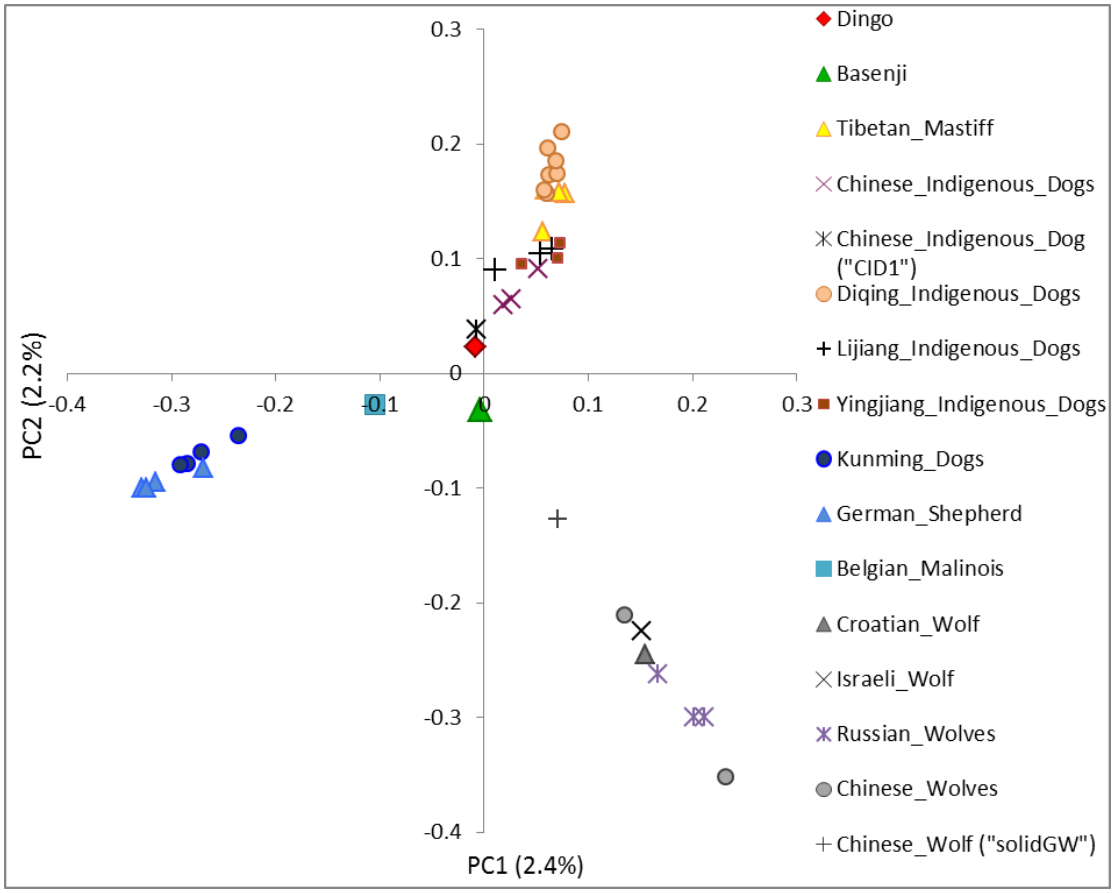
A



B



C



Supplementary Figure 20. Evolutionary relationships between indigenous Chinese dogs, pure-breed dogs of modern and ancient origin, and grey wolves from the DoGSD database (Bai *et al*. 2015). (A) Neighbour-joining tree of inter-individual IBS distances (B) Maximum likelihood TREEMIX tree constructed without assuming gene flow; (C) PCA plot.

Supplementary Table 1. List of dog breeds with basal position in the dog phylogeny according to four different studies (modified from Larson *et al*. 2012). Primary place of origin is provided if known from earlier studies, and different than present origin.

Breed	Parker *et al*. 2004	vonHoldt *et al*. 2010	Larson *et al*. 2012	Present study	Origin	Primary origin	Functional group
Chow Chow	y	y			EA		spitz
Shar-Pei	y	y	y	y	EA		spitz
Shiba Inu	y			y	EA		spitz
Akita	y	y	y	y	EA		spitz
Alaskan Malamute	y	y		y	AR	EA[1]	spitz
American Eskimo		amb			AR	EA[1]	spitz
Greenland Sledge Dog				y	AR	EA[1]	spitz
Siberian Husky	y	y		y	EA(AR)		spitz
Samoyed	amb	y			WA(AR)		spitz
Dingo		y			AU	EA[2]	wild dog
New Guinea Singing Dog		y			NG	EA[2]	wild dog
Afghan Hound	y	y			C-WA		sighthound
Saluki	y	y	y		ME		sighthound
Basenji	y	y	y		AF		sighthound
Tibetan Terrier	amb		amb	amb	EA		non-specified
Canaan		y			ME		spitz
Eurasier			y	amb	EU/EA		spitz
Finnish Spitz			y	amb	EU		spitz
Keeshond	n			amb	EU		spitz
Schipperke	n		amb	amb	EU		spitz
Elkhound	n			amb	EU		spitz
German Spitz				amb	EU		spitz

y – early-branching position, n – not early-branching position, amb – ambiguous

EA – East Asia, AR – Arctic, WA – West Asia, C-WA – Central/West Asia, ME – Middle East, EU – Europe, AF – Africa, AU – Australia, NG – New Guinea

[1] van Asch *et al*. 2013

[2] Oskarsson *et al*. 2012

Supplementary Table 2. A list of FBDs used in this study and their sampling sites.

ID	Sampling site	Sampling site name	Region
23A	Yerevan, Armenia	Armenia	Central/West Asia
24A	Yerevan, Armenia	Armenia	Central/West Asia
25A	Yerevan, Armenia	Armenia	Central/West Asia
26A	Yerevan, Armenia	Armenia	Central/West Asia
27A	Yerevan, Armenia	Armenia	Central/West Asia
28A	Yerevan, Armenia	Armenia	Central/West Asia
29A	Yerevan, Armenia	Armenia	Central/West Asia
30A	Yerevan, Armenia	Armenia	Central/West Asia
31A	Yerevan, Armenia	Armenia	Central/West Asia
32A	Yerevan, Armenia	Armenia	Central/West Asia
33A	Yerevan, Armenia	Armenia	Central/West Asia
34A	Yerevan, Armenia	Armenia	Central/West Asia
35A	Yerevan, Armenia	Armenia	Central/West Asia
36A	Yerevan, Armenia	Armenia	Central/West Asia
37A	Yerevan, Armenia	Armenia	Central/West Asia
38A	Yerevan, Armenia	Armenia	Central/West Asia
39A	Yerevan, Armenia	Armenia	Central/West Asia
40A	Yerevan, Armenia	Armenia	Central/West Asia
41A	Yerevan, Armenia	Armenia	Central/West Asia
42A	Yerevan, Armenia	Armenia	Central/West Asia
43A	Yerevan, Armenia	Armenia	Central/West Asia
44A	Yerevan, Armenia	Armenia	Central/West Asia
45A	Yerevan, Armenia	Armenia	Central/West Asia
46A	Yerevan, Armenia	Armenia	Central/West Asia
47A	Yerevan, Armenia	Armenia	Central/West Asia
89R	Tomsk, Russia	Central Russia	Central/West Asia
91R	Tomsk, Russia	Central Russia	Central/West Asia
92R	Tomsk, Russia	Central Russia	Central/West Asia
93R	Tomsk, Russia	Central Russia	Central/West Asia
94R	Tomsk, Russia	Central Russia	Central/West Asia
95R	Tomsk, Russia	Central Russia	Central/West Asia
96R	Tomsk, Russia	Central Russia	Central/West Asia

97R	Tomsk, Russia	Central Russia	Central/West Asia
98R	Tomsk, Russia	Central Russia	Central/West Asia
99R	Tomsk, Russia	Central Russia	Central/West Asia
100R	Tomsk, Russia	Central Russia	Central/West Asia
102R	Tomsk, Russia	Central Russia	Central/West Asia
104R	Tomsk, Russia	Central Russia	Central/West Asia
105R	Tomsk, Russia	Central Russia	Central/West Asia
106R	Tomsk, Russia	Central Russia	Central/West Asia
107R	Tomsk, Russia	Central Russia	Central/West Asia
1KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
2KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
3KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
4KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
5KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
6KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
7KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
8KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
9KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
10KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
11KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
12KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
13KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
14KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
15KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
16KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
17KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
18KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
19KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
20KZ	Almaty, Kazakhstan	Kazakhstan	Central/West Asia
1TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
2TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
3TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
4TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
5TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
6TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
7TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia

8TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
9TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
10TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
12TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
13TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
14TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
15TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
16TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
17TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
18TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
19TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
20TDZ	Dushanbe, Tajikistan	Tajikistan	Central/West Asia
2CH	Zibo, Shandong Province, China	China	East Asia
3CH	Zibo, Shandong Province, China	China	East Asia
4CH	Zibo, Shandong Province, China	China	East Asia
5CH	Zibo, Shandong Province, China	China	East Asia
6CH	Zibo, Shandong Province, China	China	East Asia
7CH	Zibo, Shandong Province, China	China	East Asia
8CH	Zibo, Shandong Province, China	China	East Asia
9CH	Zibo, Shandong Province, China	China	East Asia
10CH	Zibo, Shandong Province, China	China	East Asia
1MG	Ulan Bator, Mongolia	Mongolia	East Asia
2MG	Ulan Bator, Mongolia	Mongolia	East Asia
3MG	Ulan Bator, Mongolia	Mongolia	East Asia
4MG	Ulan Bator, Mongolia	Mongolia	East Asia
6MG	Ulan Bator, Mongolia	Mongolia	East Asia
8MG	Ulan Bator, Mongolia	Mongolia	East Asia
9MG	Ulan Bator, Mongolia	Mongolia	East Asia
10MG	Ulan Bator, Mongolia	Mongolia	East Asia
11MG	Ulan Bator, Mongolia	Mongolia	East Asia
12MG	Ulan Bator, Mongolia	Mongolia	East Asia
13MG	Ulan Bator, Mongolia	Mongolia	East Asia
14MG	Ulan Bator, Mongolia	Mongolia	East Asia
15MG	Ulan Bator, Mongolia	Mongolia	East Asia
16MG	Ulan Bator, Mongolia	Mongolia	East Asia
17MG	Ulan Bator, Mongolia	Mongolia	East Asia

18MG	Ulan Bator, Mongolia	Mongolia	East Asia
20MG	Ulan Bator, Mongolia	Mongolia	East Asia
21MG	Ulan Bator, Mongolia	Mongolia	East Asia
22MG	Ulan Bator, Mongolia	Mongolia	East Asia
23MG	Ulan Bator, Mongolia	Mongolia	East Asia
25MG	Ulan Bator, Mongolia	Mongolia	East Asia
26MG	Ulan Bator, Mongolia	Mongolia	East Asia
27MG	Ulan Bator, Mongolia	Mongolia	East Asia
28MG	Ulan Bator, Mongolia	Mongolia	East Asia
29MG	Ulan Bator, Mongolia	Mongolia	East Asia
30MG	Ulan Bator, Mongolia	Mongolia	East Asia
31MG	Ulan Bator, Mongolia	Mongolia	East Asia
1TAJ	Mueang Khon Kaen District, Thailand	Thailand	East Asia
2TAJ	Mueang Khon Kaen District, Thailand	Thailand	East Asia
3TAJ	Mueang Khon Kaen District, Thailand	Thailand	East Asia
4TAJ	Mueang Khon Kaen District, Thailand	Thailand	East Asia
5TAJ	Mueang Khon Kaen District, Thailand	Thailand	East Asia
6TAJ	Mueang Khon Kaen District, Thailand	Thailand	East Asia
7TAJ	Mueang Khon Kaen District, Thailand	Thailand	East Asia
8TAJ	Mueang Khon Kaen District, Thailand	Thailand	East Asia
10TAJ	Mueang Khon Kaen District, Thailand	Thailand	East Asia
11TAJ	Mueang Khon Kaen District, Thailand	Thailand	East Asia
12TAJ	Mueang Khon Kaen District, Thailand	Thailand	East Asia
13TAJ	Mueang Maha Sarakham District, Thailand	Thailand	East Asia
14TAJ	Mueang Maha Sarakham District, Thailand	Thailand	East Asia
16TAJ	Mueang Maha Sarakham District, Thailand	Thailand	East Asia
17TAJ	Mueang Maha Sarakham District, Thailand	Thailand	East Asia
18TAJ	Mueang Maha Sarakham District, Thailand	Thailand	East Asia
19TAJ	Mueang Maha Sarakham District, Thailand	Thailand	East Asia
21TAJ	Mueang Maha Sarakham District, Thailand	Thailand	East Asia
22TAJ	Mueang Maha Sarakham District, Thailand	Thailand	East Asia
23TAJ	Mueang Maha Sarakham District, Thailand	Thailand	East Asia
24TAJ	Mueang Maha Sarakham District, Thailand	Thailand	East Asia
108R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia

109R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
110R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
111R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
112R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
113R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
114R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
116R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
118R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
119R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
120R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
121R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
122R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
123R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
124R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
126R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
125R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
127R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
128R	Ussuriysk, Primorsky Krai, Russia	East Russia	East Russia
2BL	Bulgaria	Bulgaria	Europe
3BL	Bulgaria	Bulgaria	Europe
4BL	Bulgaria	Bulgaria	Europe
6BL	Bulgaria	Bulgaria	Europe
8BL	Bulgaria	Bulgaria	Europe
9BL	Bulgaria	Bulgaria	Europe
10BL	Bulgaria	Bulgaria	Europe
11BL	Bulgaria	Bulgaria	Europe
12BL	Bulgaria	Bulgaria	Europe
1PL	Zduny, Poland	Poland	Europe
2PL	Zduny, Poland	Poland	Europe
3PL	Zduny, Poland	Poland	Europe
4PL	Zduny, Poland	Poland	Europe
5PL	Zduny, Poland	Poland	Europe
6PL	Zduny, Poland	Poland	Europe
7PL	Zduny, Poland	Poland	Europe
8PL	Zduny, Poland	Poland	Europe
9PL	Zduny, Poland	Poland	Europe

10PL	Zduny, Poland	Poland	Europe
13PL	Zduny, Poland	Poland	Europe
14PL	Zduny, Poland	Poland	Europe
15PL	Zduny, Poland	Poland	Europe
16PL	Zduny, Poland	Poland	Europe
17PL	Zduny, Poland	Poland	Europe
18PL	Zduny, Poland	Poland	Europe
19PL	Zduny, Poland	Poland	Europe
20PL	Zduny, Poland	Poland	Europe
21PL	Zduny, Poland	Poland	Europe
11PL	Zduny, Poland	Poland	Europe
12PL	Zduny, Poland	Poland	Europe
2SL	Skofije, Slovenia	Slovenia	Europe
3SL	Portoroz, Slovenia	Slovenia	Europe
4SL	Ankaran, Slovenia	Slovenia	Europe
5SL	Koper, Slovenia	Slovenia	Europe
6SL	Koper, Slovenia	Slovenia	Europe
7SL	Portoroz, Slovenia	Slovenia	Europe
8SL	Skofije, Slovenia	Slovenia	Europe
9SL	Skofije, Slovenia	Slovenia	Europe
10SL	Hrvaška, Slovenia	Slovenia	Europe
11SL	Koper, Slovenia	Slovenia	Europe
12SL	Koper, Slovenia	Slovenia	Europe
13SL	Ankaran, Slovenia	Slovenia	Europe
14SL	Piran, Slovenia	Slovenia	Europe
1p	Basrah, Iraq	Iraq	Middle East
2p	Basrah, Iraq	Iraq	Middle East
3p	Basrah, Iraq	Iraq	Middle East
4p	Basrah, Iraq	Iraq	Middle East
5p	Basrah, Iraq	Iraq	Middle East
6p	Basrah, Iraq	Iraq	Middle East
7p	Basrah, Iraq	Iraq	Middle East
8p	Basrah, Iraq	Iraq	Middle East
2AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
10AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
11AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East

13AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
14AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
15AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
16AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
17AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
18AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
19AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
20AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
21AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
24AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
25AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
27AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
28AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
29AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
30AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
31AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
32AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
22AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
26AS	Riyadh, Saudi Arabia	Saudi Arabia 1	Middle East
36AS	Al-Baha, Saudi Arabia	Saudi Arabia 2	Middle East
37AS	Al-Baha, Saudi Arabia	Saudi Arabia 2	Middle East
38AS	Al-Baha, Saudi Arabia	Saudi Arabia 2	Middle East
40AS	Al-Baha, Saudi Arabia	Saudi Arabia 2	Middle East
41AS	Al-Baha, Saudi Arabia	Saudi Arabia 2	Middle East

Supplementary Table 3. **A list of dog breeds used in this study, and their regions of origin.** "UK" dataset was produced in the present study, and consisted of dogs sampled in the United Kingdom. "LUPA" is a published, publicly available dataset deriving from Vaysse *et al*. (2011). Columns N(LUPA) and N(UK) provide the number of representatives of each breed present in these two datasets.

Breed	N(LUPA)	N(UK)	Breed origin
Australian Kelpie		2	Australia, ultimately Europe
Bearded Collie		3	Europe
Border Collie	16	8	Europe
Rough Collie		2	Europe
Shetland Sheepdog		2	Europe
Beagle	10	1	Europe
Dachshund	12		Europe
Belgian Tervuren	12		Europe
German Shepherd Dog	12	9	Europe
Doberman	25		Europe
Malinois		1	Europe
Bernese Mountain Dog	12		Europe
Dalmatian		2	Europe
Rottweiler	12	8	Europe
Weimaraner	26		Europe
Vizsla		4	Europe
Border Terrier	25		Europe
Cairn Terrier		2	Europe
Jack Russell Terrier	12	4	Europe
Lakeland Terrier		2	Europe
Parson Terrier		2	Europe
Wheaten Terrier		2	Europe
Yorkshire Terrier	12		Europe
English Bulldog	13		Europe
Staffordshire Bullterrier		4	Europe
Brittany Spaniel	12		Europe
Cocker Spaniel	14	2	Europe
Springer Spaniel		1	Europe
English Setter	12		Europe
Gordon Setter	25		Europe
Golden Retriever	14	4	Europe
Flatcoated Retriever		1	North America, ultimately Europe
Labrador Retriever	14	2	North America, ultimately Europe
Newfoundland Retriever	25		North America, ultimately Europe
Nova Scottia Duck-Tolling Retriever	23		North America, ultimately Europe
Toy Poodle		1	Europe
Miniature Poodle		1	Europe
Standard Poodle	12		Europe
Greyhound	11	6	Europe

Elkhound	12		Europe
Irish Wolfhound	11		Europe
Keeshond		2	Europe
Schipperke	25		Europe
German Spitz		2	Europe
Finnish Spitz	12		Europe
Eurasier	12		mixed Europe and East Asia
Tibetan Terrier		2	East Asia
Shar Pei	11		East Asia
Shiba Inu		2	East Asia
Siberian Husky		2	East Asia (east Siberia)
Alaskan Malamute		2	North America, ultimately East Asia
Greenland Sledge Dog	12		Greenland, ultimately East Asia

Supplementary Table 4. Genetic diversity at genome-wide autosomal SNPs in FBDs populations in Eurasia. H_O – observed heterozygosity, H_E – expected heterozygosity, N – number of individuals used to calculate the diversity indices. Intra-population divergence was calculated using EIGENSOFT software. The table is sorted according to H_O values.

Location	Region	N	H_O	H_E	Intra-population divergence
Armenia	C/W Asia	9	0.3538	0.3341	1.061
Kazakhstan	C/W Asia	9	0.3502	0.3312	0.994
Mongolia	E Asia	9	0.3484	0.3293	0.973
Bulgaria	Europe	9	0.3447	0.3268	0.964
Tajikistan	C/W Asia	9	0.3441	0.3288	1.005
China	E Asia	9	0.3427	0.3239	0.972
Poland	Europe	9	0.3415	0.3295	1.024
Iraq	Middle East	4	0.3383	0.2967	0.946
Central Russia	C/W Asia	9	0.3336	0.3287	1.023
Thailand	E Asia	9	0.3336	0.3264	1.058
East Russia	NE Asia	9	0.3266	0.3275	1.040
Slovenia	Europe	9	0.3254	0.3285	1.029
Saudi Arabia 1	Middle East	9	0.3021	0.3070	0.986
Saudi Arabia 2	Middle East	5	0.2973	0.2783	0.924

Supplementary Table 5. Autozygosity in FBD populations in Eurasia.

Autozygosity is measured as number of homozygous segments (ROHs) of different length identified in individuals from each population.

Location	Region	Total number of ROHs	Number of ROHs <1Mb	Freq. of ROHs <1Mb	Number of ROHs 1-2Mb	Freq. of ROHs 1-2Mb	Number of ROHs >2Mb	Freq. of ROHs >2Mb	Largest ROH length (Mb)
Iraq	ME	59	10	0.17	15	0.25	34	0.58	19
China	EA	70	19	0.27	35	0.50	16	0.23	9
Armenia	C-WA	100	23	0.23	44	0.44	33	0.33	31
Kazakhstan	C-WA	111	13	0.12	62	0.56	36	0.32	26
S. Arabia 2	ME	116	15	0.13	37	0.32	64	0.55	53
Mongolia	EA	118	21	0.18	46	0.39	51	0.43	12
Bulgaria	EU	136	26	0.19	49	0.36	61	0.45	24
Thailand	EA	163	26	0.16	64	0.39	73	0.45	54
Tajikistan	C-WA	166	26	0.16	71	0.43	69	0.42	41
Poland	EU	173	37	0.21	74	0.43	62	0.36	61
Central Russia	C-WA	191	33	0.17	51	0.27	107	0.56	73
Slovenia	EU	241	39	0.16	88	0.37	114	0.47	53
S. Arabia 1	ME	320	25	0.08	98	0.31	197	0.62	52
East Russia	EA	333	43	0.13	110	0.33	180	0.54	88

EA – East Asia, C-WA – Central/West Asia, ME – Middle East, EU – Europe