Supplementary information, Data S8 Estimating the amount of gene flow between dogs and wolves

Since admixture can confound many population genetic analyses, we reviewed the migration intensities between dogs and wolves. We have three major sources of information.

a) In our dadi/MCMC analysis, the migration rate between dogs and wolves was found to be quite low (2Nm = 1 from dogs to wolves and 2Nm = 5 from wolves to dogs). An independent estimate using whole genome data was presented by Freedman *et al.*, 2014[9]. The highest gene flow between dogs and wolves was found between the Israeli wolf and the Basenji (e.g., Figure 5 panel A in the Freedman *et al.*'s study). The migration intensity (denoted as migration band) is given by the total migration probability, which equals the per generation migration rate/probability times the number of generations. In other words, this is the probability of being a migrant in a specific timespan (denote as P_{band}). For example, the P_{band} from the Israeli wolf to the Basenji was found to be 0.12-0.24. If we take 0.18 as a point estimate, this is equivalent to m = 0.18/(14,900/3) = 3.62E-5 per generation. Here 14,900 is the estimated split time between the two populations and we assume a generation time of 3 years. The true migration rate from Israeli wolves to Basenjis is 2Nm = 2*2600*3.63E-5 = 0.19 (here we are taking 2,600 as the population size for the Basenji, see Figure 5 panel A in Freedman *et al.*'s study). Similar calculations can be done for other migration bands; they give estimates less than the above estimated value (Figure 5, Panel A in Freedman *et al.*'s study). In other words, the demographic analysis estimates low levels of migrations between wolves and dogs.

b) From earlier mtDNA studies[17], researchers have found that, "Clades A, B, and C were represented in every population, representing 97.4% of all dogs and in most regions 100% of the dogs". If we assume that the rest of the haplotypes are due to admixture from wolves, we can conservatively estimate wolf admixture at about 2.6% using this mtDNA data.

c) An independent estimate of admixture that is different from the mtDNA and demographic inference is the four-population test[18]. By calculating the ratio between the two different F4 statistics[18] (Figure 4 in Patterson *et al.*'s 2012 study[18] and formula 5 therein), we can extract the admixture proportion of the source population to the target population. For example, a good combination of F4 test will be 1-F4(GSD, outgroup; SI, W)/F4(GSD, outgroup; TIM, W), where GSD is the German Shepherd, TIM is the Tibetan Mastiff and outgroup is the dhole (Supplementary information, Table S1). Using genome-wide SNP data, we estimated the admixture proportion from wolves to southern Chinese indigenous dogs to be 0.126. Many other combinations gave similar estimates (data not shown). Further analyses of local gene flow between dogs and wolves across Eurasia, using F4 and D tests, can be found in Supplementary information, Data S10 and S14. Combining the above evidence, we conclude that the amount of wolf/dog admixture is very limited, and the level of admixture is similar among multiple dog/wolf pairs across Eurasia.