Supplementary information, Data S10 Serial founder models can generate a gradient in distances from the outgroup observed in the PCA plot and are unlikely to be caused by wolf/dog migration

In the PCA analysis (Figure 1G and 1H), we observed a serial gradient in genetic distances from the wolves to the domestic dogs. The question is then, can a serial founder model create patterns like this? When we simulated a simple serial founder model (Supplementary information, Figure S10) and performed a PCA analysis, we easily recapitulated the pattern observed in Figure 1G. The basal lineage is the one closest to the outgroup with recently derived populations being much further away from the outgroup. This suggests that genetic drift in the basal groups is significantly less and they tend to stay genetically more similar to the outgroup. When we added migrations between the outgroup and all dog populations, the pattern stays qualitatively very similar.

Pickrell and Reich outlined an alternative scenario to the serial founder model for human evolutionary history[20]. The basic idea is that, if there are high migration rates between a source population (outgroup) and a geographically nearby population (e.g., Figure 1B and 1C of Pickrell and Reich, 2014), the resulting genetic pattern mimics a serial founder model.

There are multiple lines of evidence against this possibility for the dataset in our study. First, the levels of migrations between wolves and dogs are all very low (migration rates mostly less than 1 and a maximum migration rate of $2Nm = 5$ was observed in our dadi analysis, Supplementary information, Data S8). In order to estimate the amount of admixture between multiple dog/wolf population pairs, we performed F4 tests to study migrations between multiple dog and wolf pairs. For example, using the Israeli wolf and Croatian wolf downloaded from the study by Freedman *et al.*[9], we estimated the gene flow between Israeli wolves and African dogs to be 1-F4(GSD, Outgroup; AF, ISW)/F4(GSD, outgroup; SAM, ISW) = 0.161, where the outgroup is the sequenced dhole (Supplementary information, Table S1). Likewise, estimating the admixture proportions between Croatian wolves and European dogs using the SNP array dataset, the amount of admixture is mostly around 4-5%. (The analysis is only possible with the SNP array data. For example, using 1-

F4(TIM, outgroup; GREY, the Croatian wolf)/ F4(TIM, outgroup; LAH, the Croatian wolf), we reached a value of 4%. Here GREY is the Greyhound and LAH is the Lapponian Herder.)

In summary, we find comparable levels of gene flow between wolves and dogs in Europe and Asia. The wolf/dog migration is slightly higher in the Middle East. In general, the proportion is between a few percent to as high as 12-16%. Thus, there is no evidence that gene flow in East Asia is much higher than in other regions and admixture is unlikely to create the distance gradient observed in the PCA plot.