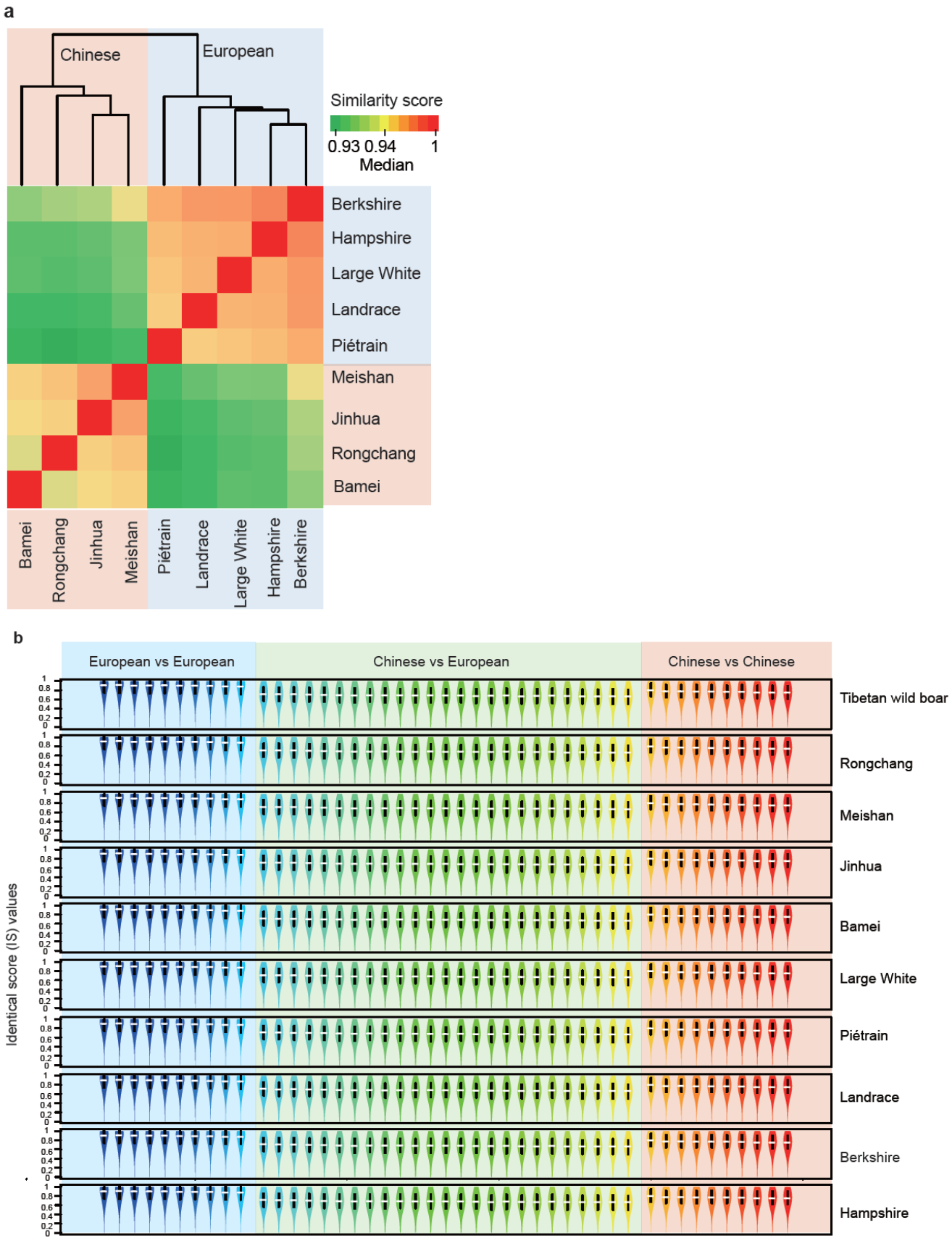

Supplemental Fig. S30. Pairwise similarity of the orphan reads (a) and the missing sequences (b) across ten assemblies. (a) Hierarchical classification of orphan reads based on pairwise similarity scores computed from the read set intersections as implemented in the Compareads tool (v.1.3)[7]. The orphan reads derived from Tibetan wild boar assembly were excluded, due to the idiosyncrasies of the different sequencing platforms (i.e. the Illumina HiSeq 2000 used for Tibetan wild boar and HiSeq 2500 used for other nine assemblies). **(b)** Violin plot of identity score (IS) for missing sequences across ten assemblies.