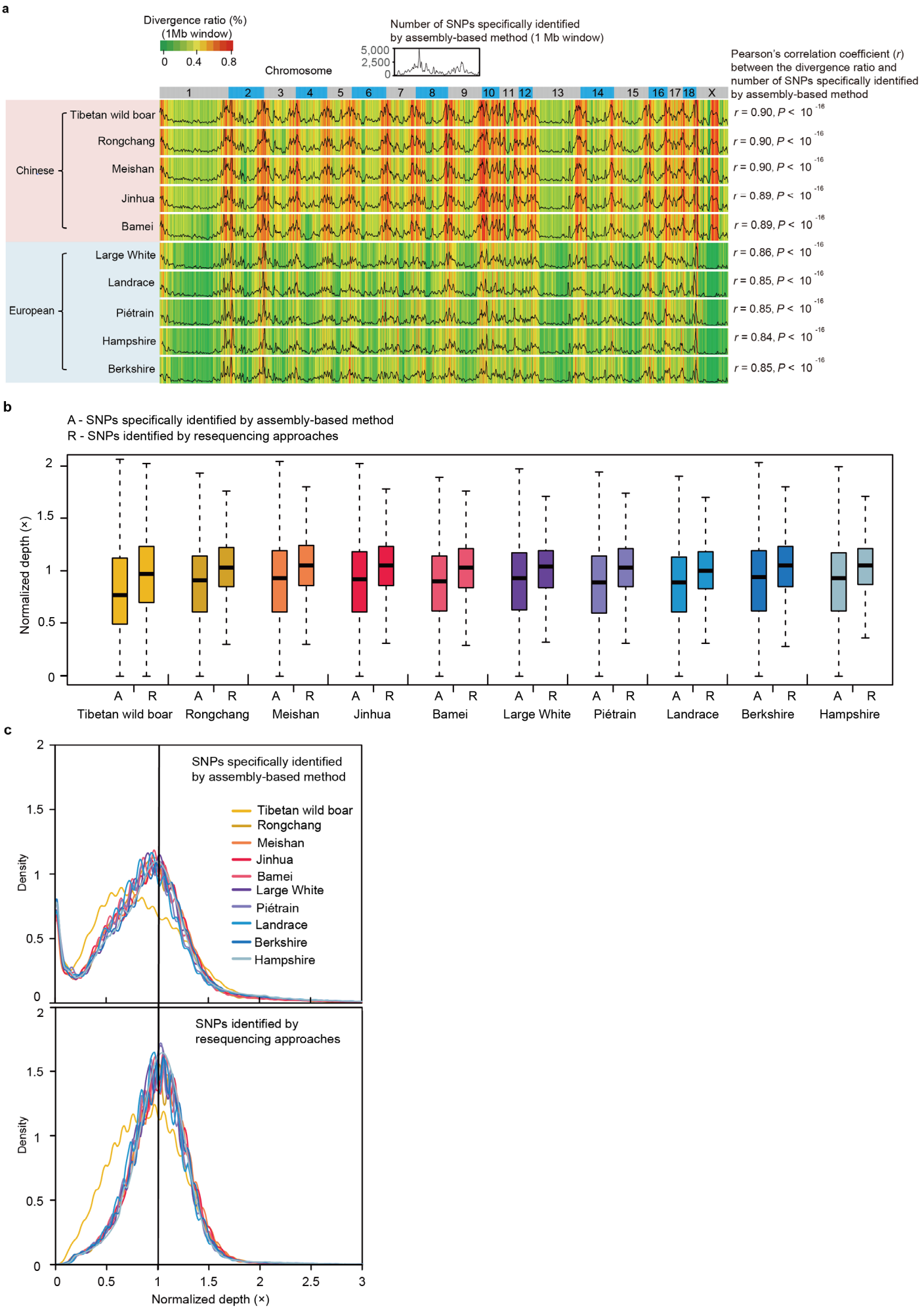

Supplemental Fig. S9. Comparison of SNPs identified by resequencing approaches and only by the

assembly-based method. (**a**) Distribution of single nucleotide divergence ratio (%) between the sequenced genomes to the reference pig genome in 2,603 non-overlapping windows in 1-Mb length of the whole genome. Divergence ratio (%) = (Number of homozygous SNPs + 0.5 × Number of heterozygous SNPs) / windows length × 100. We observed that SNPs identified only by the assembly-based method are mainly located in highly diverged regions between the sequenced genomes and the reference pig genome, which positively correlates with the divergence ratio (average Pearson's $r = 0.87$, $P < 10^{-16}$). (**b**) Box plot and (**c**) density distribution of SNPs depth by mapping short reads to the reference genome; the distribution was normalized by the average depth of 26,755 non-overlapping windows in 100 kb length of the whole genome. Compared to the SNPs identified by resequencing approaches, the assembly-based method specifically identified SNPs that exhibited relatively lower read depth by mapping short reads to the reference genome. Of the loci (~2,714) both covered by the Illumina's porcine 60K Genotyping BeadChip (v.2) and the assembly-based method specifically identified, more than 97% (~2,635) were consistent with the SNPs identified by the assembly-based method (**Supplemental Table S7**), demonstrating the high quality and reliability of SNP calls. These results suggest the improved power of variant detection of the assembly-based method, as opposed to the currently dominant resequencing approach, especially in divergent regions where unassembled short sequencing reads are difficult to be mapped.