



**Supplemental Fig. S38. Expression of missing genes quantified by orphan reads after initial mapping of RNA-seq onto the reference genomes.** Summary of mapping ratio **(a)** and unmapped RNA-seq reads **(b)** of 96 mRNA-seq libraries (7 to 10 libraries for each of 10 pig breeds). For OEA reads, only one end was counted. **c**, Evidence of transcription of missing genes. About 0.59 G 'orphan reads' against the reference genome from each of 96 pair-end RNA-seq libraries (7 to 10 libraries for each of ten breeds) were retrieved, and mapped onto the missing genes in their respective assemblies and 1,000 randomly selected annotated genes in reference genome (i.e., negative controls). Consequently, an average of 91.51% (845 of 923) missing genes in each of ten assemblies with  $\log_2$ -transformed FPKM expression values greater than 0.3 in at least one library were identified. FPKM, expression values denoted as fragments per kb of transcript per Mb orphan reads. RNA-seq data for four tissues of Tibetan wild boar (heart, kidney, liver and lung) were downloaded from the NCBI-Gene Expression Omnibus (GEO) database under accession code GSE43892.