Supplemental information for:

Population level analyses reveal contrasting landscapes of gene expression divergence and evolution between male and female transcriptome lineages in the house mouse

Kaizong Wei, Chen Xie, Aftab Ahmad, Lei Duan, Mingji Chu, Yuanxiao Gao, Diethard Tautz*, and Wenyu Zhang*

*Corresponding to: wyzhang@nwpu.edu.cn and tautz@evolbio.mpg.de

Supplemental Figures and Tables

Other supplemental data files stored separately for this manuscript

Data S1 Summary information of RNA-Seq data used in this study.

Data S2 Summary information of the number of genes with expression.

Data S3 Summary information of transcriptomic distances based on the (A) first and (B) second RNA-Seq datasets.

Data S4 Percentage of sex-biased genes in the four gene categories with distinct gene expression patterns.

Data S5 Statistics on the chromosomal distributions of gene categories with different expression patterns.

Data S6 Full list of significantly (A) enriched KEGG pathways and (B) GO BO terms of sex-specific DE genes.

Data S7 Full list of DE (A) protein-coding genes and (B) lncRNA genes for the first RNA-Seq dataset.

Data S8 Full list of DE (A) protein-coding genes and (B) lncRNA genes for the second RNA-Seq dataset.

Data S9 Full list of sex-biased (A) protein-coding genes and (B) lncRNA genes for the first RNA-Seq dataset.

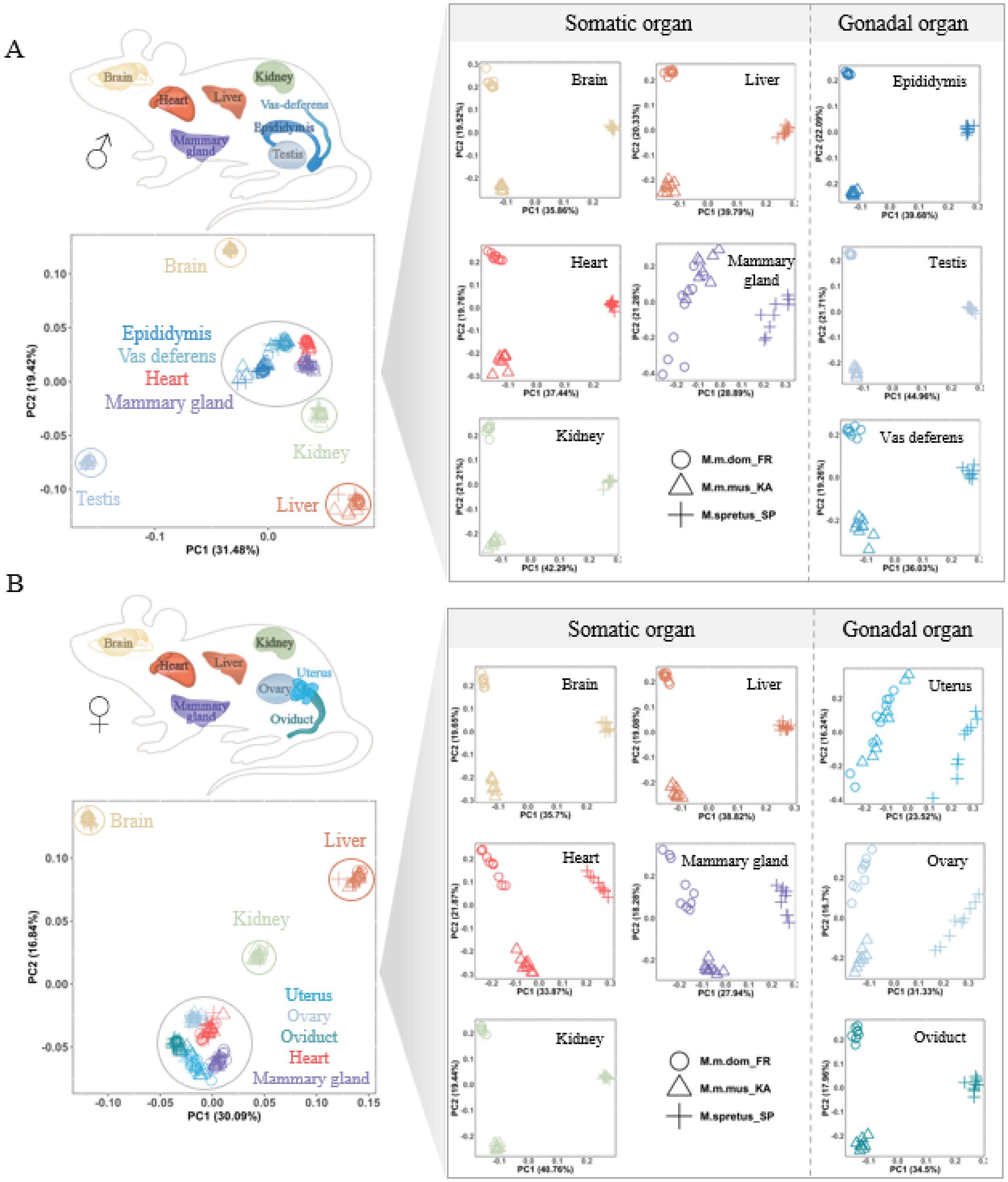

Figure S1 Sex- and organ-specific patterns of protein-coding gene expression divergence. Principal component analysis (PCA) of protein-coding gene expression profiles in (A) male and (B) female mice. Main panels display global clustering patterns across all organs, with insets showing organ-specific analyses. Principal components 1 and 2 (PC1/PC2) are shown, with variance explained indicated in axis labels.

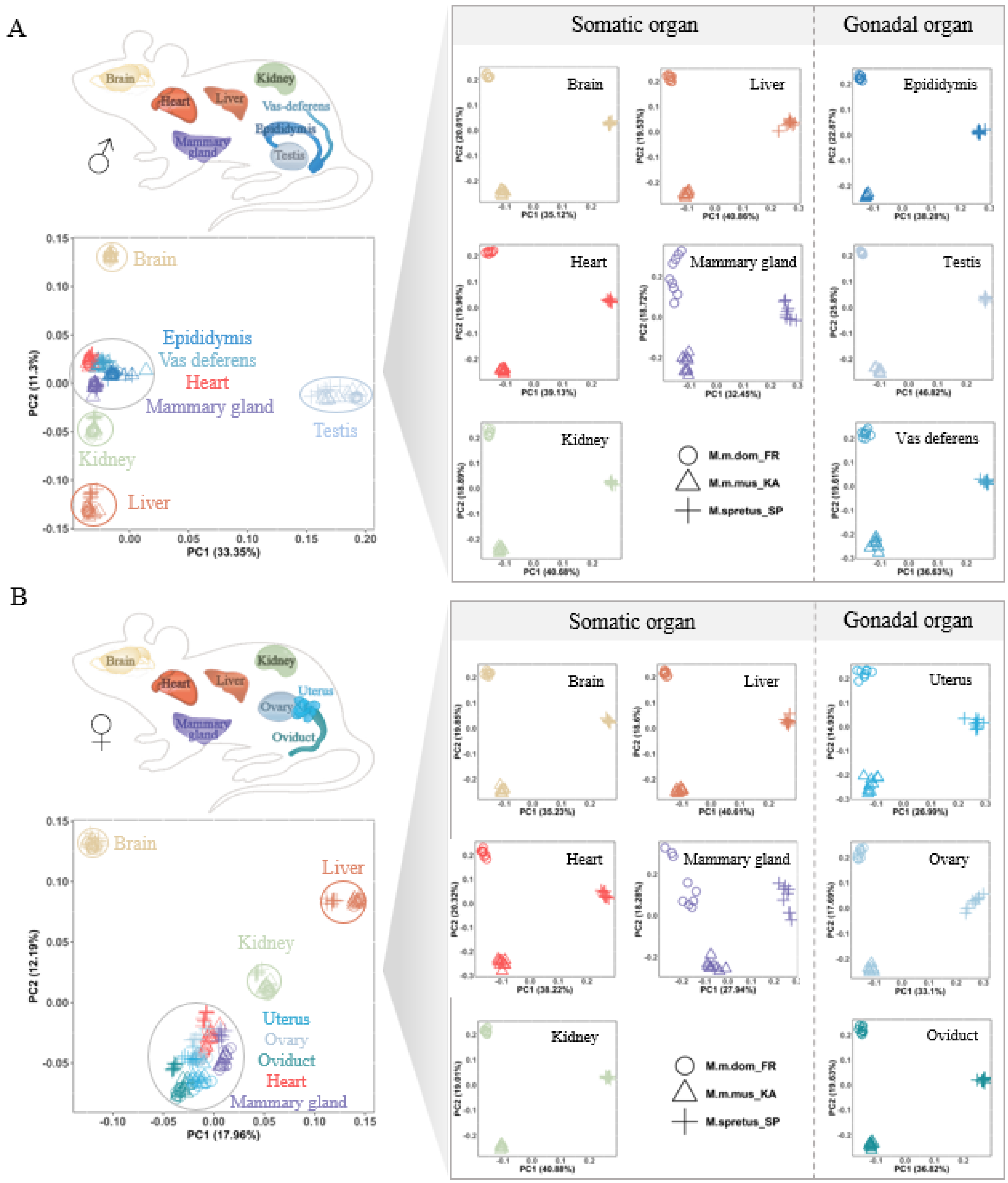

Figure S2 Sex- and organ-specific patterns of lncRNA gene expression divergence. Principal component analysis (PCA) of lncRNA gene expression profiles in (A) male and (B) female mice. Main panels display global clustering patterns across all organs, with insets showing organ-specific analyses. Principal components 1 and 2 (PC1/PC2) are shown, with variance explained indicated in axis labels.

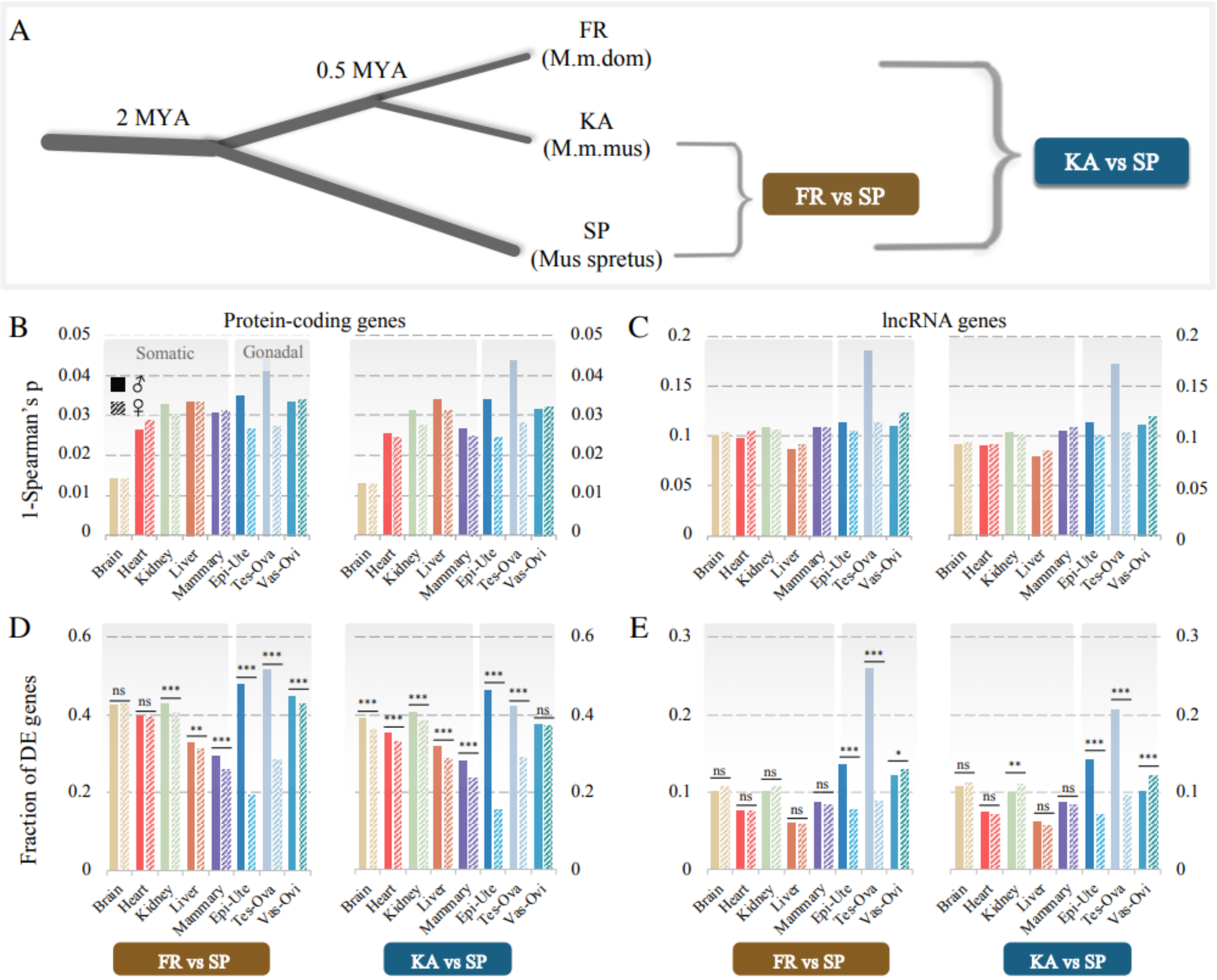

Figure S3 Comparative analysis of sex-specific gene expression evolution at inter-species comparison level. (A) Phylogenetic relationships and divergence times among three *Mus* taxa. Comparative analyses include FR vs SP and KA vs SP comparisons. Distributions of transcriptomic distance (1 - Spearman's rank correlation coefficient, ρ) for (B) protein-coding genes and (C) long non-coding RNA (lncRNA) genes. Male and female organ data were visually differentiated using solid-filled and hatched bars, respectively. Fraction of differentially expressed (DE) genes (FDR ≤ 0.05) for (D) protein-coding genes and (E) lncRNA genes. Significance of sex differences in DE gene counts per organ was assessed by Fisher's exact test (***: p-value ≤ 0.001; **: p-value ≤ 0.01; *: p-value ≤ 0.05; ns: p-value > 0.05). DE gene analysis was performed using edgeR. Organ abbreviations follow Figure 2.

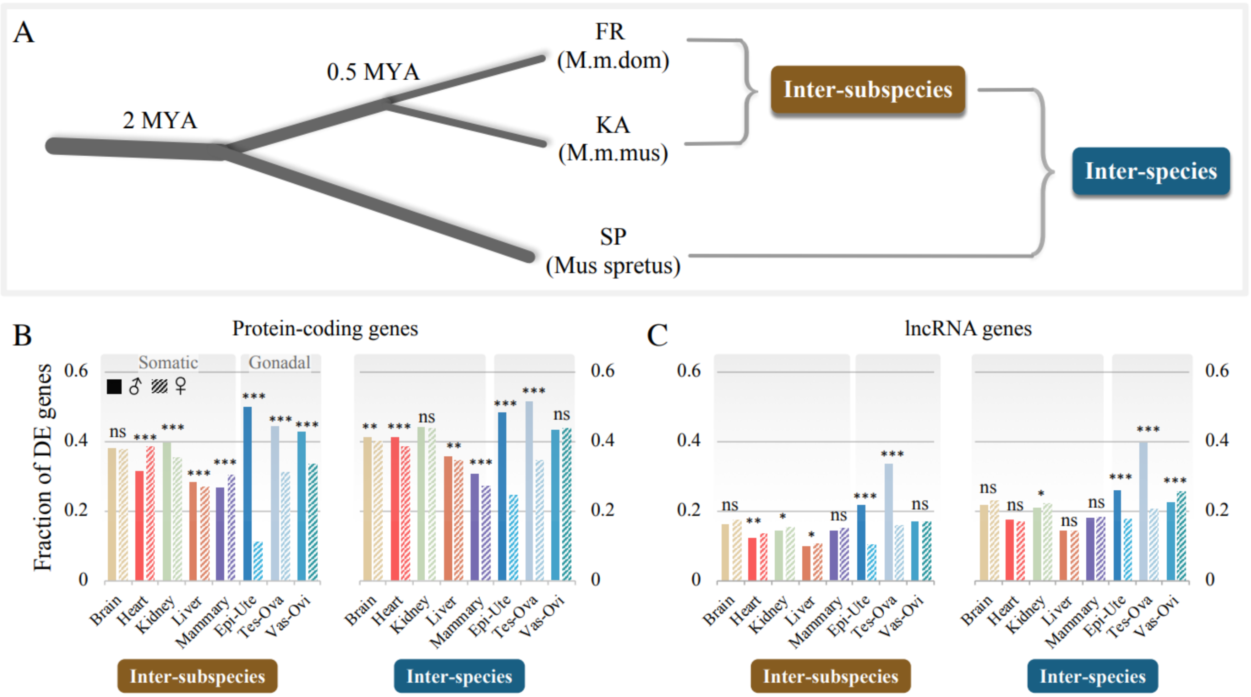

Figure S4 Comparative analysis of sex-specific DE gene expression with a different DE gene analysis tool. (A) Phylogenetic relationships and divergence times among three *Mus* taxa. Comparative analyses include inter-subspecies (*M. m. dom* versus *M. m. mus*) and inter-species (*Mus musculus* versus *Mus spretus*) comparisons. Differentially expressed (DE) gene counts (FDR ≤ 0.05) for (B) protein-coding genes and (C) lncRNA genes. DE analysis was performed using DESeq2. Male and female organ data were visually differentiated using solid-filled and hatched bars, respectively. Significance of sex differences in DE counts per organ was assessed by Fisher's exact test (***: p-value ≤ 0.001; **: p-value ≤ 0.01; *: p-value ≤ 0.05; ns: p-value > 0.05). Organ abbreviations follow Figure 2.

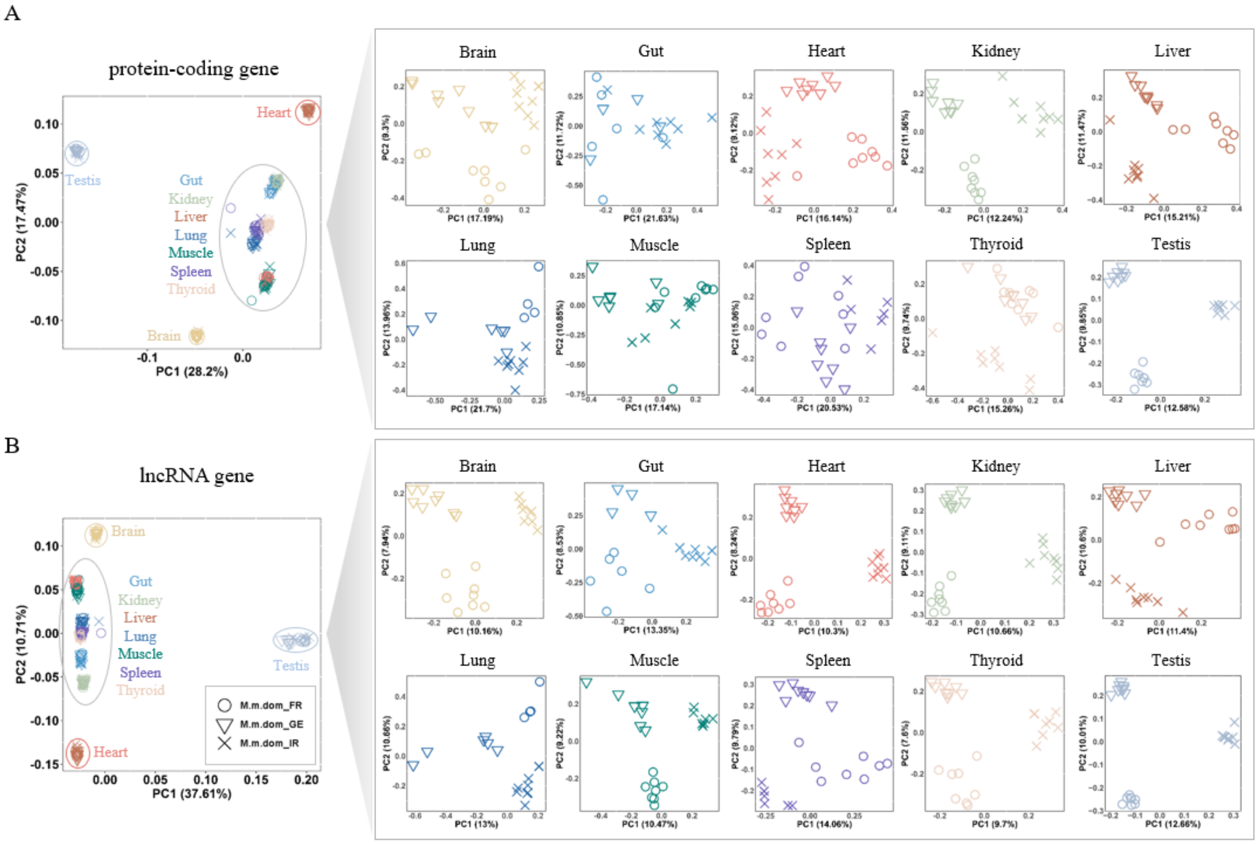

Figure S5 Organ-specific patterns of gene expression divergence in male transcriptome. Principal component analysis (PCA) of (A) protein-coding gene and (B) lncRNA gene expression profiles. Main panels display global clustering patterns across all organs, with insets showing organ-specific analyses. Principal components 1 and 2 (PC1/PC2) are shown, with variance explained indicated in axis labels.

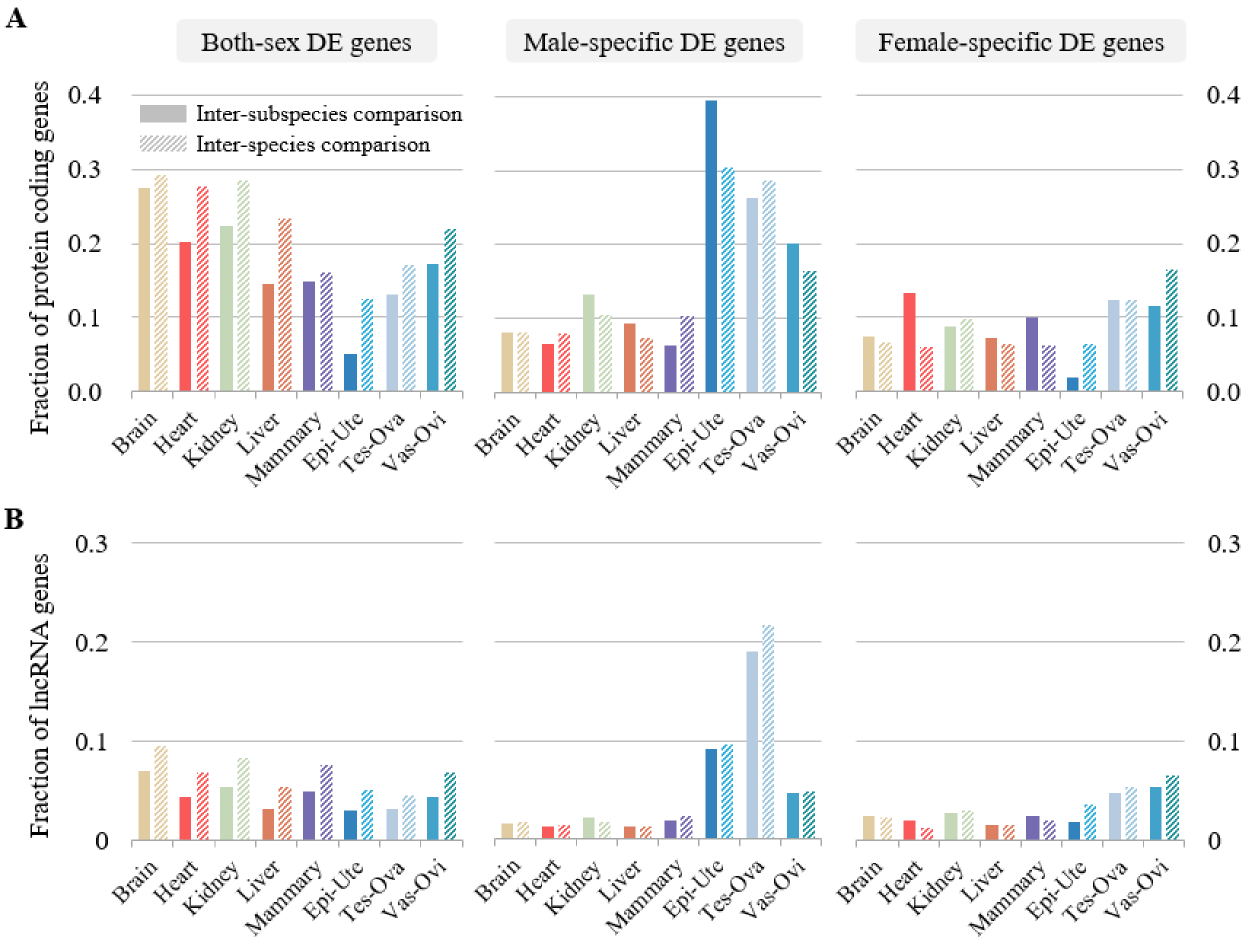

Figure S6 Distribution on the fraction of genes for three types of expression patterns. The fraction distribution patterns were visualized for (A) protein-coding and (B) lncRNA genes. Inter-subspecies and inter-species comparison data were visually differentiated using solid-filled and hatched bars, respectively. Definition of gene categories: i) both-sex DE genes, with significant gene expression level changes between taxa in both sexes; ii) male-specific DE genes, with significant gene expression level changes between taxa in male only; iii) female-specific DE genes, with significant gene expression level changes between taxa in female only. Organ abbreviations follow Figure 2.

Figure S7 Chromosomal distribution of genes for four types of expression patterns in inter-species comparisons. Percentages of genes located on autosomes, the X chromosome, and the Y chromosome are indicated adjacent to each bar. Definition of four types of gene expression patterns: i) non-DE genes, with significant gene expression level changes between taxa in neither sex; ii) both-sex DE genes, with significant gene expression level changes between taxa in both sexes; iii) male-specific DE genes, with significant gene expression level changes between taxa in male only; iv) female-specific DE genes, with significant gene expression level changes between taxa in female only. Organ abbreviations follow Figure 2.

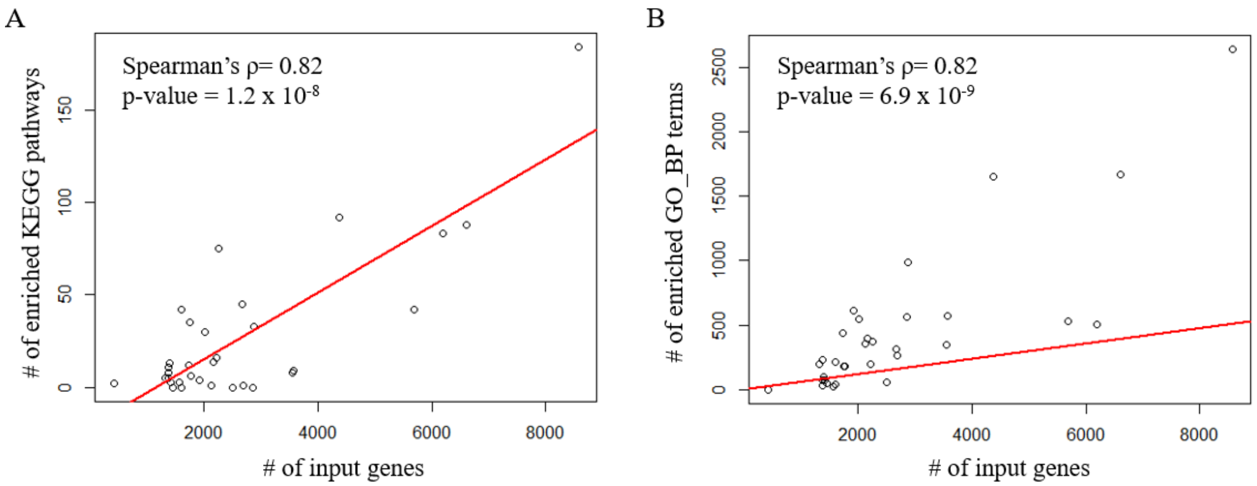

Figure S8 Correlation between the numbers of input genes and (A) KEGG pathways and (B) GO BP terms. These analyses were performed on the basis of male- and female-specific DE protein-coding genes across organs from both inter-subspecies and inter-species comparisons. The significant functional enrichment terms were defined as those passing statistical significance after multiple correction testing (FDR ≤ 0.05). The correlation significance tests were performed using Spearman's correlation tests.

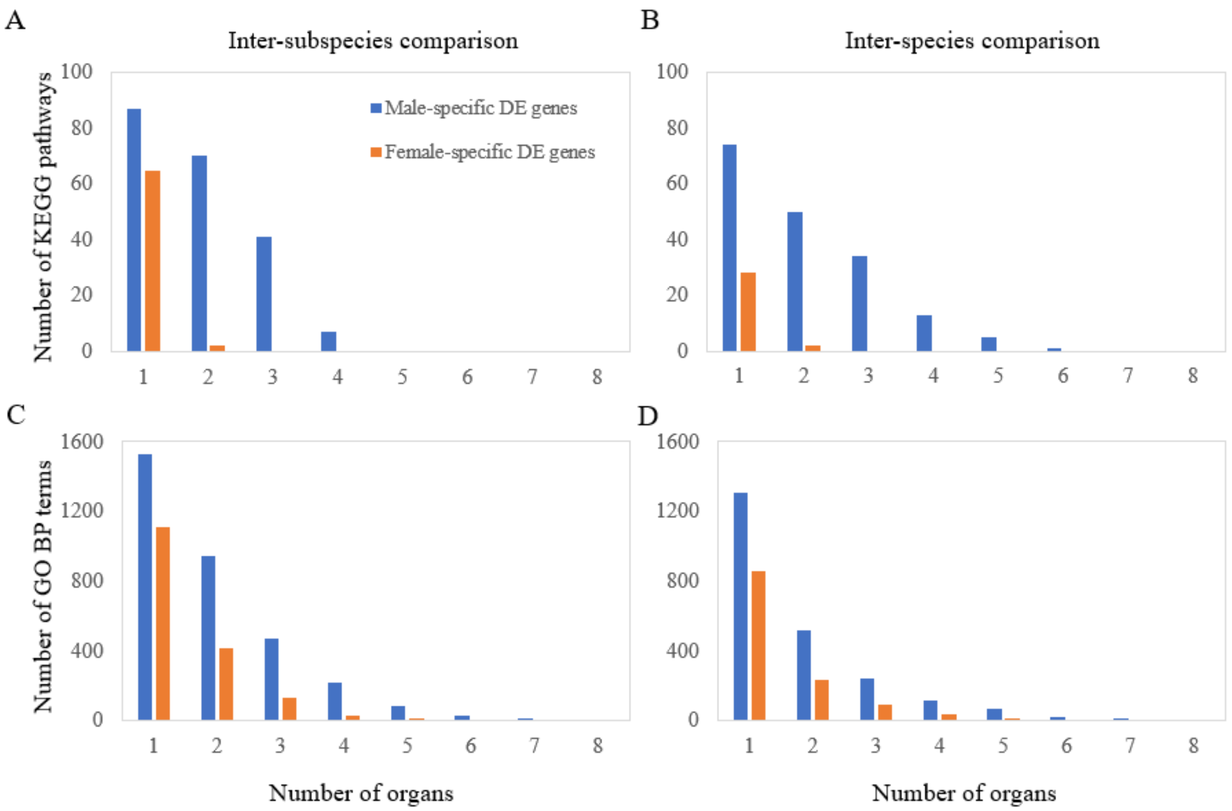

Figure S9 Distribution on the number of functional enrichment terms across organs for sex-specific DE genes. These analyses were performed on the basis of male- and female-specific DE gene across organs from both inter-subspecies and inter-species comparisons. Only functional enrichment terms (KEGG pathways and GO BP terms) passing statistical significance after multiple correction testing (FDR ≤ 0.05) were retained for analysis.

Supplementary Table S1 Statistics on the functional enrichment terms of sex-specific DE genes. Only the KEGG pathways and GO BP terms passing statistical significance after multiple correlation tests (FDR ≤ 0.05) were counted.

		# of sex-specific DE genes		**# of enriched KEGG pathways (FDR ≤ 0.05)**			**# of enriched GO BP terms (FDR ≤ 0.05)**		
		Male-specific DE genes	Female-specific DE genes	Male-specific DE genes	Female-specific DE genes	Overlap	Male-specific DE genes	Female-specific DE genes	Overlap
Brain	Inter-subspecies	1763	1596	6	0	0 (0%)	185	37	7 (6.3%)
	Inter-species	1750	1459	35	0	0 (0%)	178	46	10 (8.9%)
Heart	Inter-subspecies	1405	2881	13	33	1 (4.3%)	99	990	39 (7.2%)
	Inter-species	1734	1315	12	5	0 (0%)	442	199	40 (12.5%)
Kidney	Inter-subspecies	2857	1913	0	4	0 (0%)	566	616	90 (15.2%)
	Inter-species	2262	2132	75	1	0 (0%)	369	353	70 (19.4%)
Liver	Inter-subspecies	2024	1570	30	3	1 (6.1%)	547	21	7 (2.5%)
	Inter-species	1599	1381	42	8	1 (4%)	211	231	38 (17.2%)
Mammary	Inter-subspecies	1372	2173	11	14	1 (8%)	74	400	22 (9.3%)
	Inter-species	2221	1374	16	5	0 (0%)	195	35	1 (0.9%)
Epi-Ute	Inter-subspecies	8570	415	184	2	0 (0%)	2641	0	0 (0%)
	Inter-species	6602	1411	88	3	0 (0%)	1671	73	18 (2.1%)
Tes-Ova	Inter-subspecies	5680	2662	42	45	6 (13.8%)	532	314	46 (10.9%)
	Inter-species	6203	2683	83	1	0 (0%)	503	264	33 (8.6%)
Vas-Ovi	Inter-subspecies	4373	2503	92	0	0 (0%)	1655	53	22 (2.6%)
	Inter-species	3549	3565	8	9	0 (0%)	347	575	99 (21.5%)