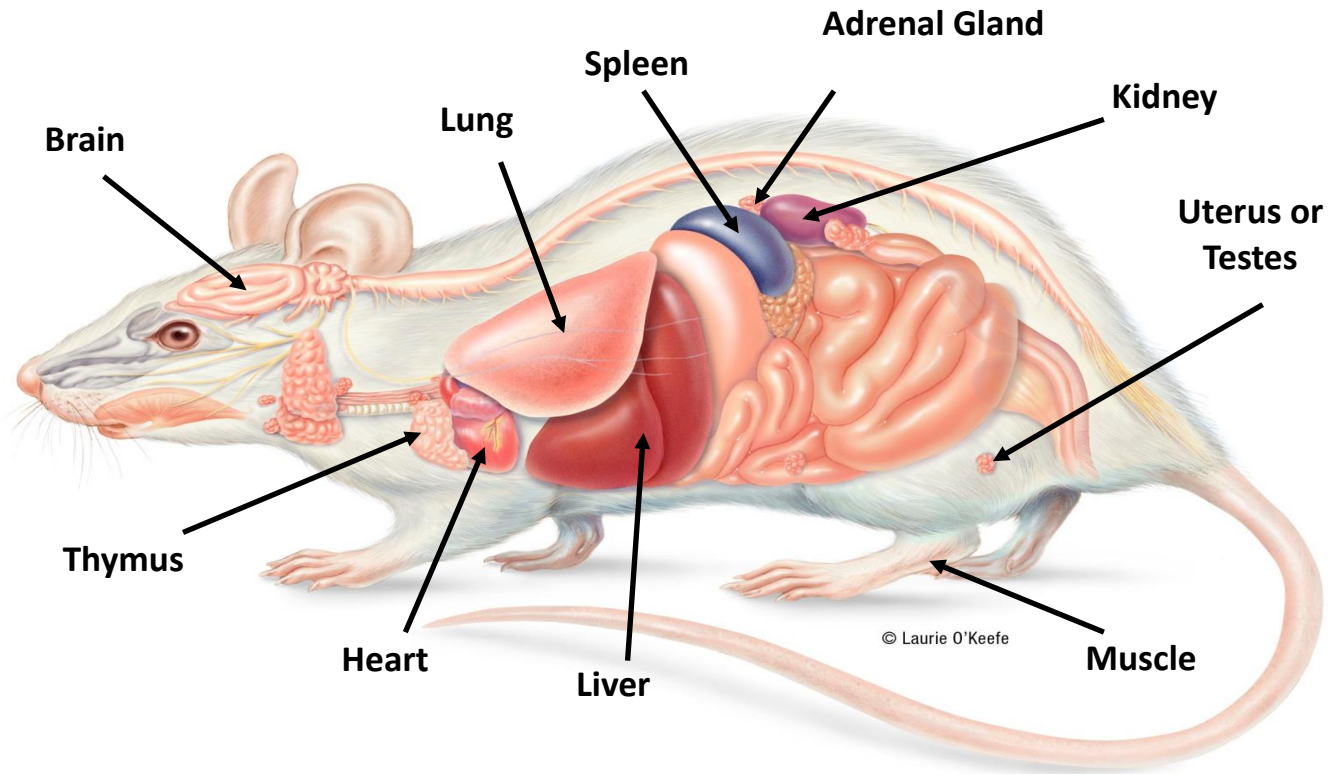
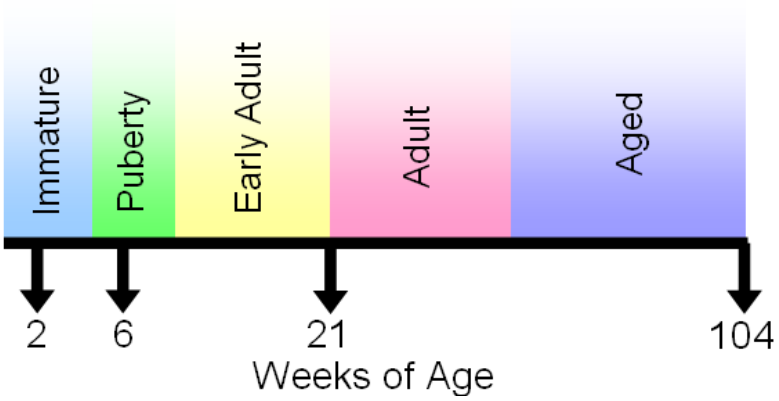
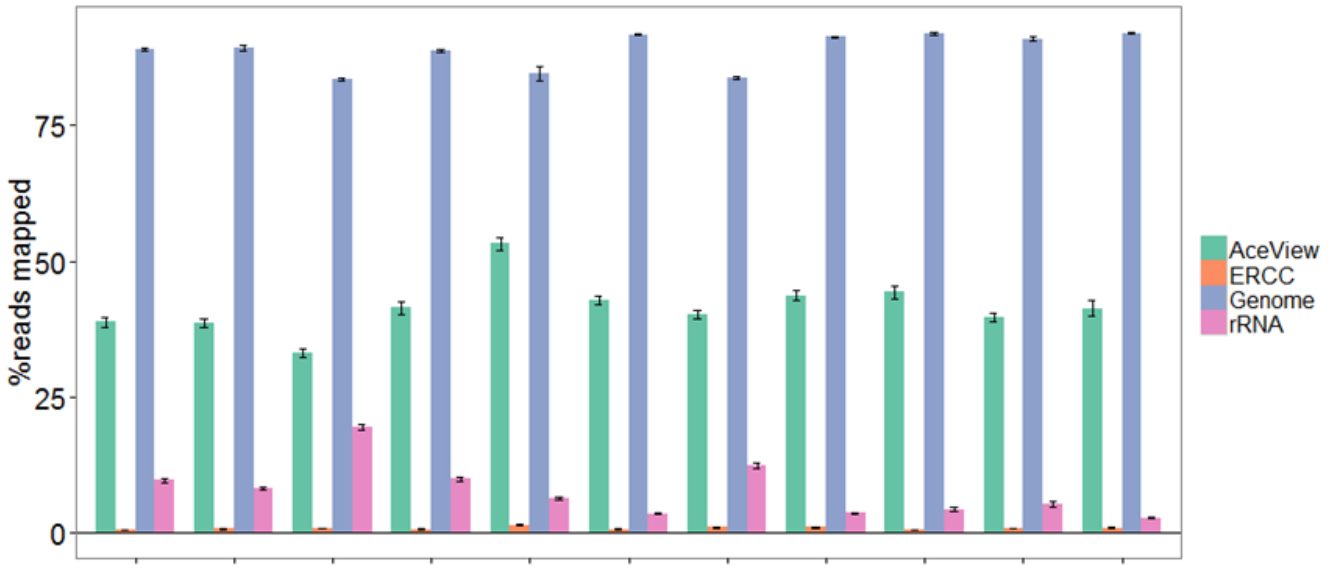
A rat RNA-Seq transcriptomic Bodymap across 11 organs and 4 developmental stages. Yu Y *et al.*

Supplementary Figures and Tables



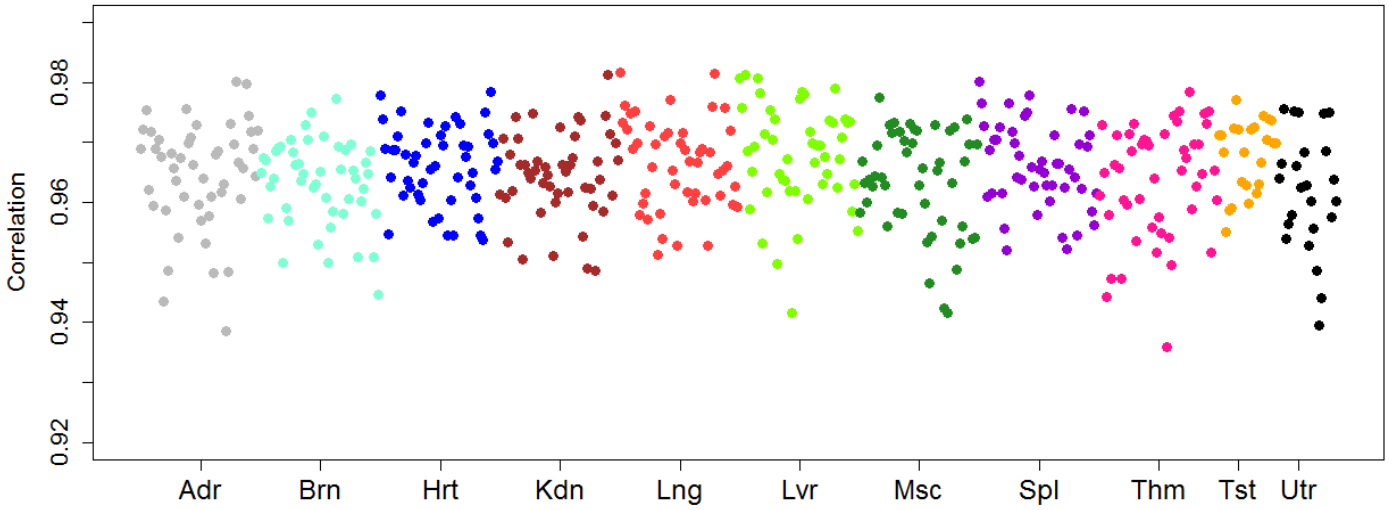


Supplementary Figure S1. Rat transcriptomic BodyMap study design. Diagram of organs harvested from Fischer 344 rats and the selected four developmental stages at which organs were harvested. Total number of RNA samples: 10 organs per rat x 32 rats (2 sexes x 4 developmental stages x 4 replicates) = 320 for RNA-Seq.

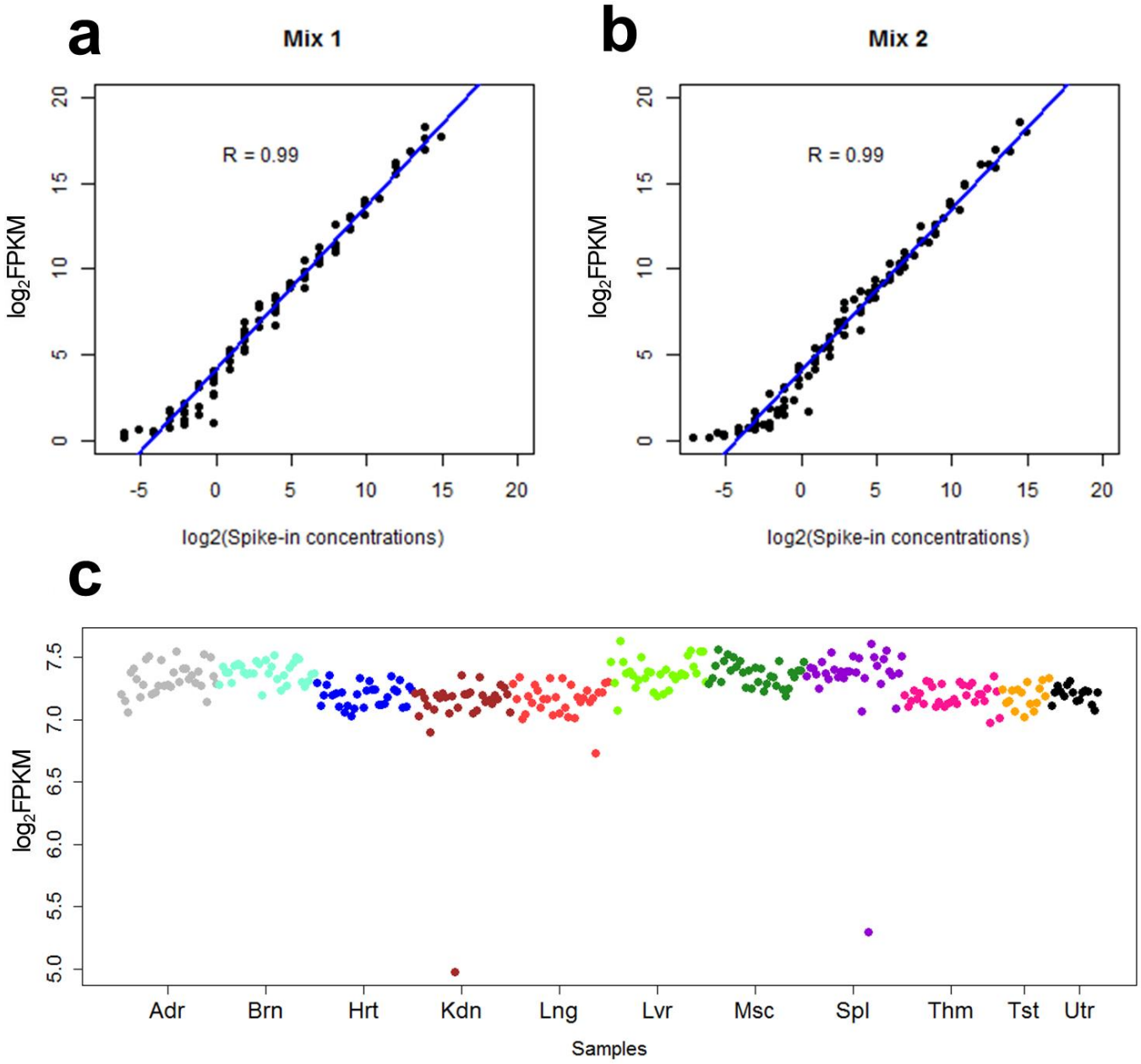


	Adr	Brn	Hrt	Kdn	Lvr	Lng	Msc	Spl	Tst	Thm	Utr
AceView	38.81	38.63	33.14	41.46	53.25	42.88	40.13	43.75	44.32	39.6	41.34
ERCC	0.66	0.78	0.87	0.74	1.51	0.76	1.06	1.05	0.69	0.91	0.95
Genome	89.06	89.27	83.57	88.74	84.66	91.72	83.77	91.35	91.97	90.89	91.93
rRNA	9.75	8.21	19.57	9.97	6.43	3.6	12.4	3.67	4.4	5.33	2.86

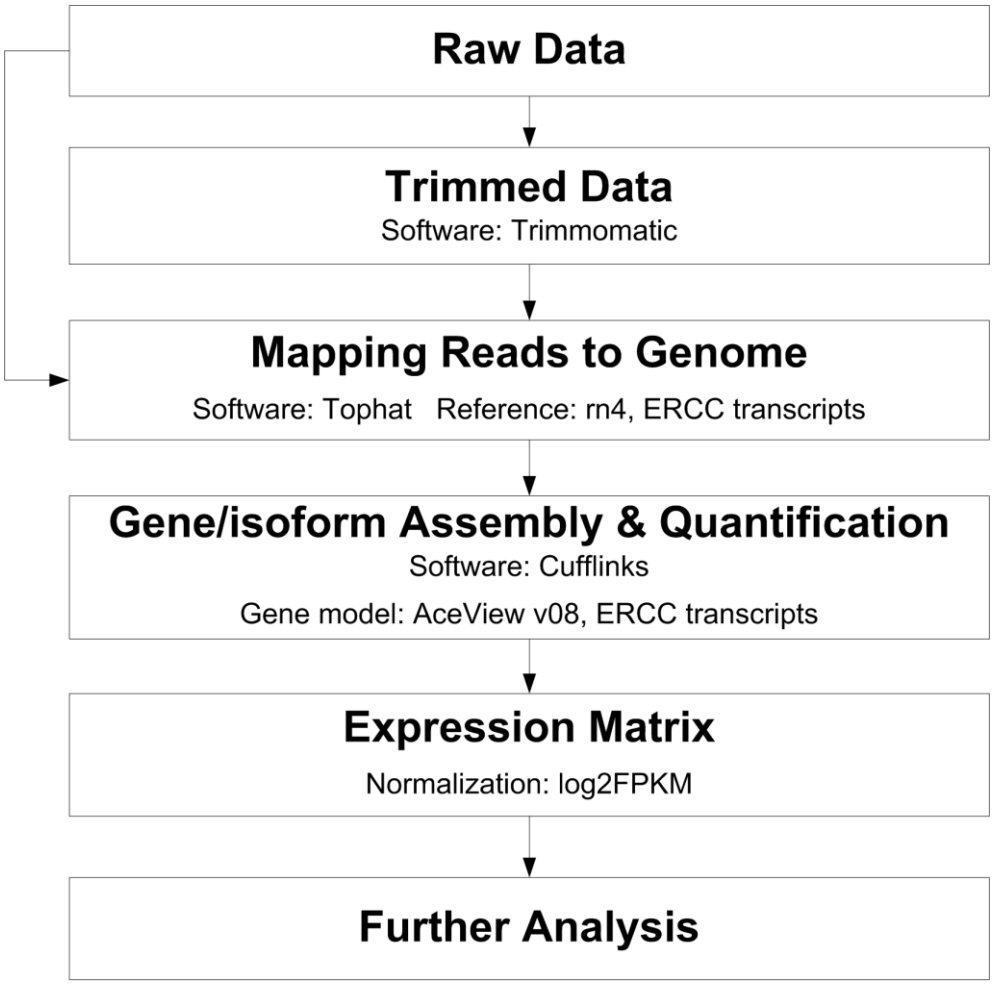
Supplementary Figure S2. The percentage of reads mapped to ERCC transcripts, rat rRNA, AceView genes, and genomic regions. The proportions of reads mapped (ratio, Mean ± SE; N=16 in uterus and testes, N=32 in the other 9 organs) to AceView exons (green), ERCC transcripts (orange), genomic regions (blue), and rRNA (pink) in each organ. Adr, adrenal; Brn, brain; Hrt, heart; Kdn, kidney; Lng, lung; Lvr, liver; Msc, skeletal muscle; Spl, spleen; Thm, thymus; Tst, testis; and Utr, uterus.



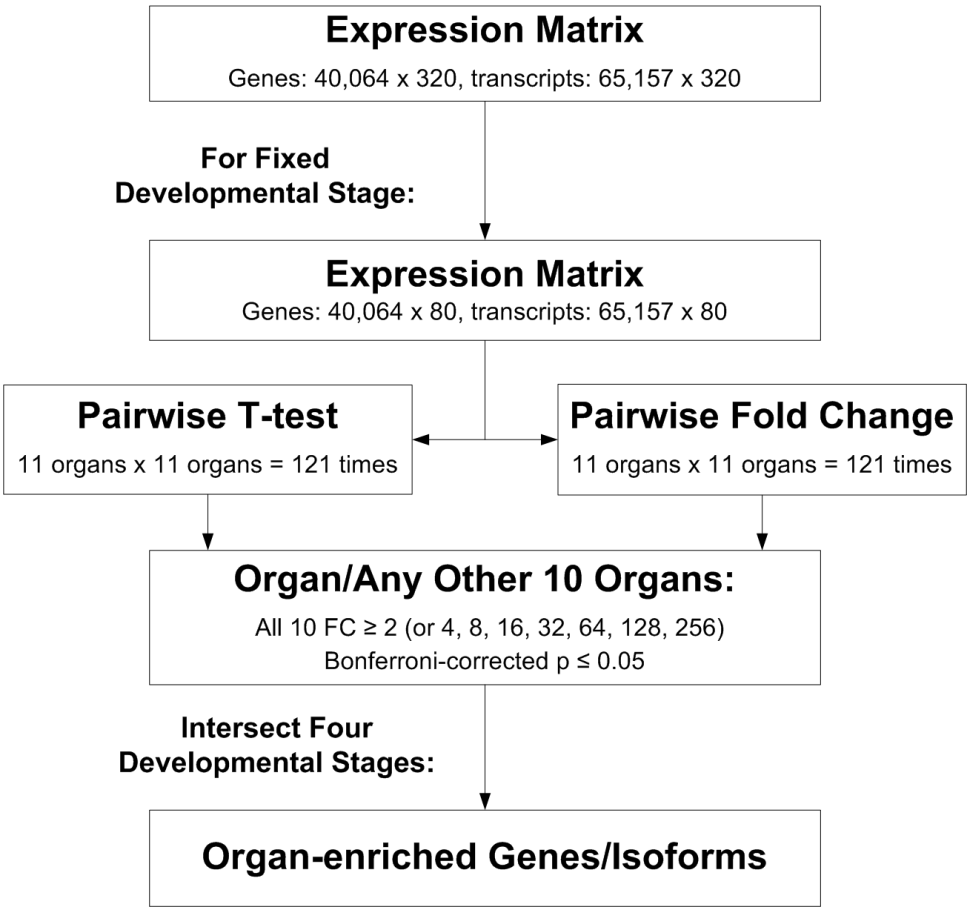
Supplementary Figure S3. The distribution of the pair-wise correlation (R) between any two biological replicates of the same condition (same organ x age x sex). Y-axis represents squared Pearson correlation between any two biological replicates (log_2FPKM). X-axis represents a total of 6*80=480 values in 11 organ types. There are four biological replicates for each of the 80 sample groups in this study, resulting in six pair-wise comparisons of biological replicates per sample group.



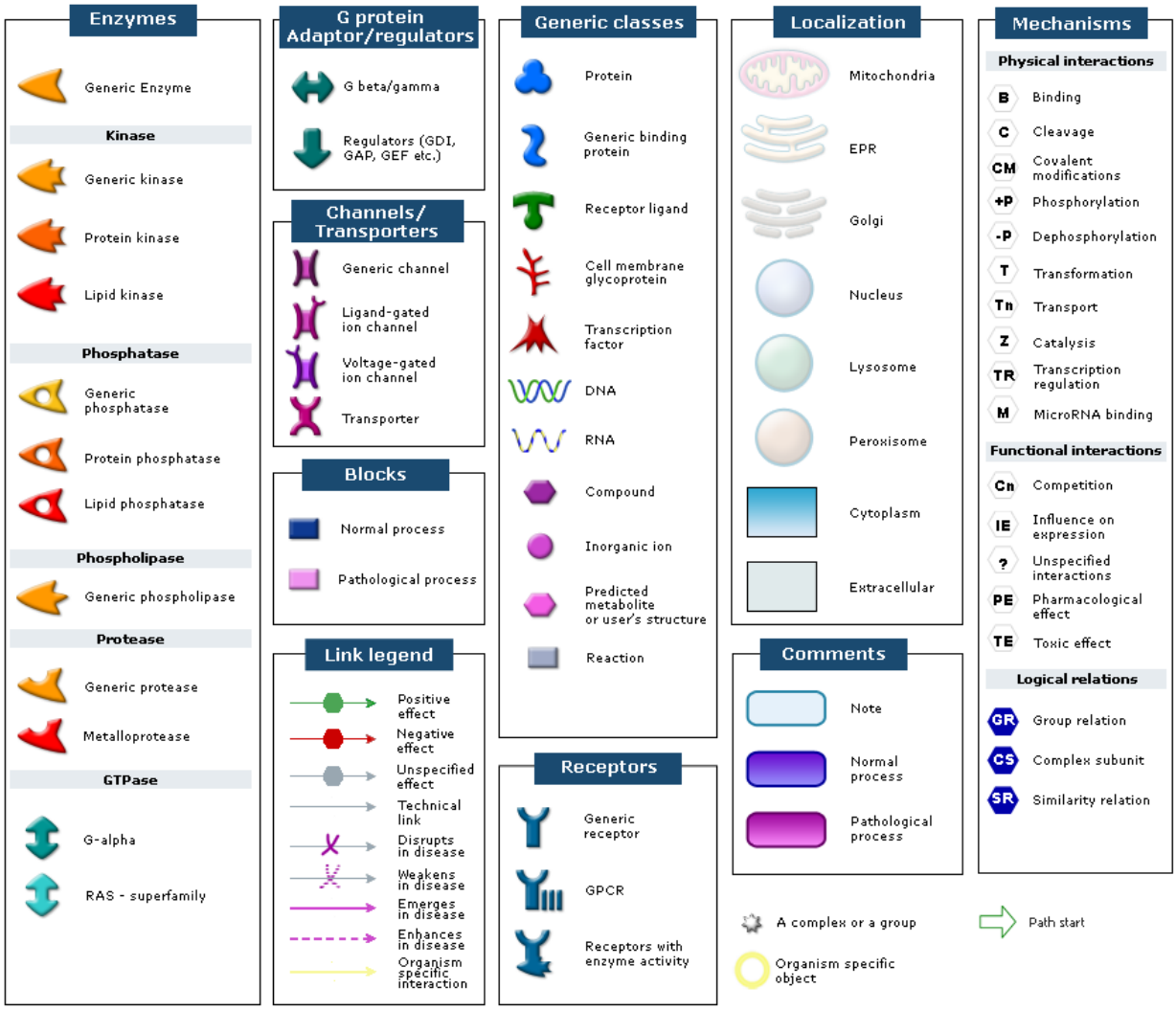
Supplementary Figure S4. ERCC performance in this study. (**a**) Scatterplot of averaged $\log_2$(FPKM) across 160 samples spiked with ERCC mixture 1 vs. $\log_2$(spike-in concentrations); (**b**) Scatterplot of averaged $\log_2$(FPKM) across 160 samples spiked with ERCC mixture 2 vs. $\log_2$(spike-in concentrations); (**c**) Averaged expression value of 92 ERCC transcripts for each of the 320 samples. Y-axis represents the averaged $\log_2$(FPKM) and X-axis represents 320 samples in this study.



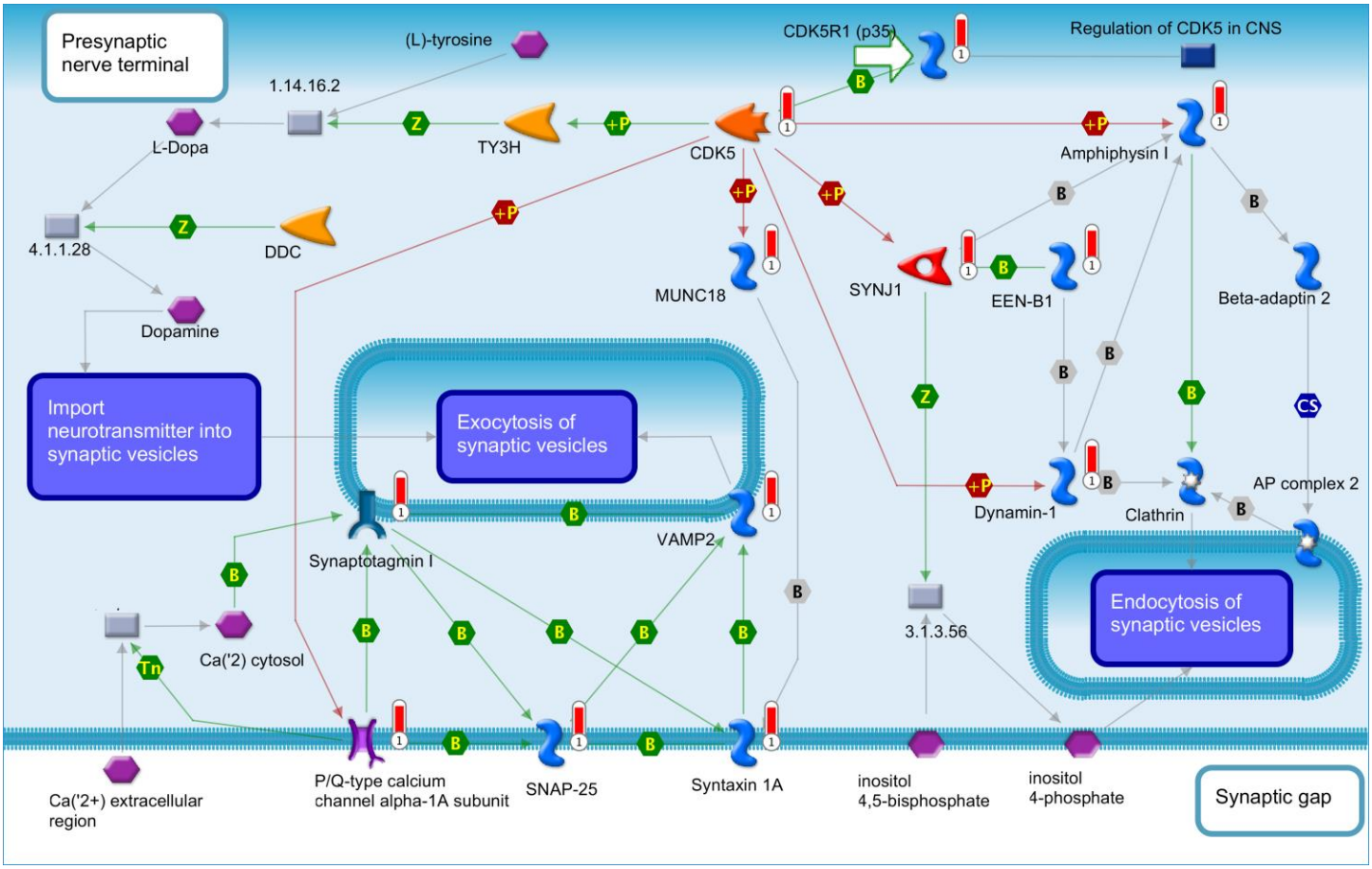
Supplementary Figure S5. Mapping pipeline. In total, 320 samples together with technical replicates were used in this study. Data were first trimmed using Trimmomatic. We used the rat transcriptome from AceView v08, which includes 40,064 genes, as a reference. In addition, the UCSC rn4 rat genome was used as a reference genome. Reads were aligned to the rat genome and transcriptome using TopHat v2.0.4 allowing a maximum of 2 mismatches in the alignment. The default parameter settings were used. Alignment results were then processed using Cufflinks for gene/transcript assembly and quantification. For samples with 2–3 technical replicates, averaged FPKM (Fragment Per Kilobase per Million mapped reads) values were used. To avoid infinite values, a pseudo-count of one was added to the FPKM value of each gene before $\log_2$ transformation.



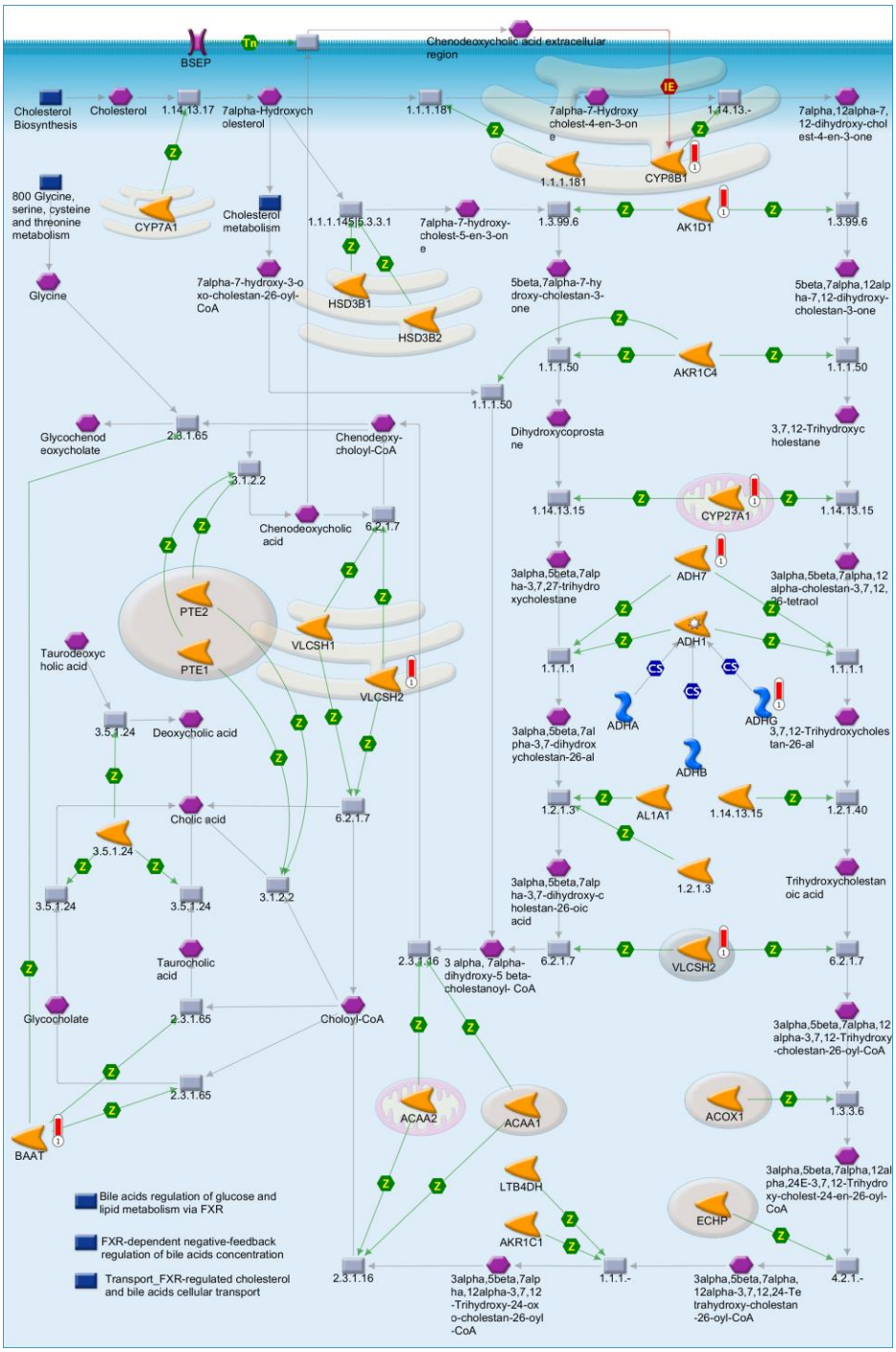
Supplementary Figure S6. Workflow to identify organ-enriched genes. A total of 320 samples were separated into 11 groups based on organ. Differentially expressed gene analysis was performed by pair-wise comparison of any 2 of the 11 organs. A gene was defined as organ-enriched if its expression level was at least 2-fold higher than that in any other organ, using a Bonferroni-corrected $p \leq 0.05$ across the four developmental stages.



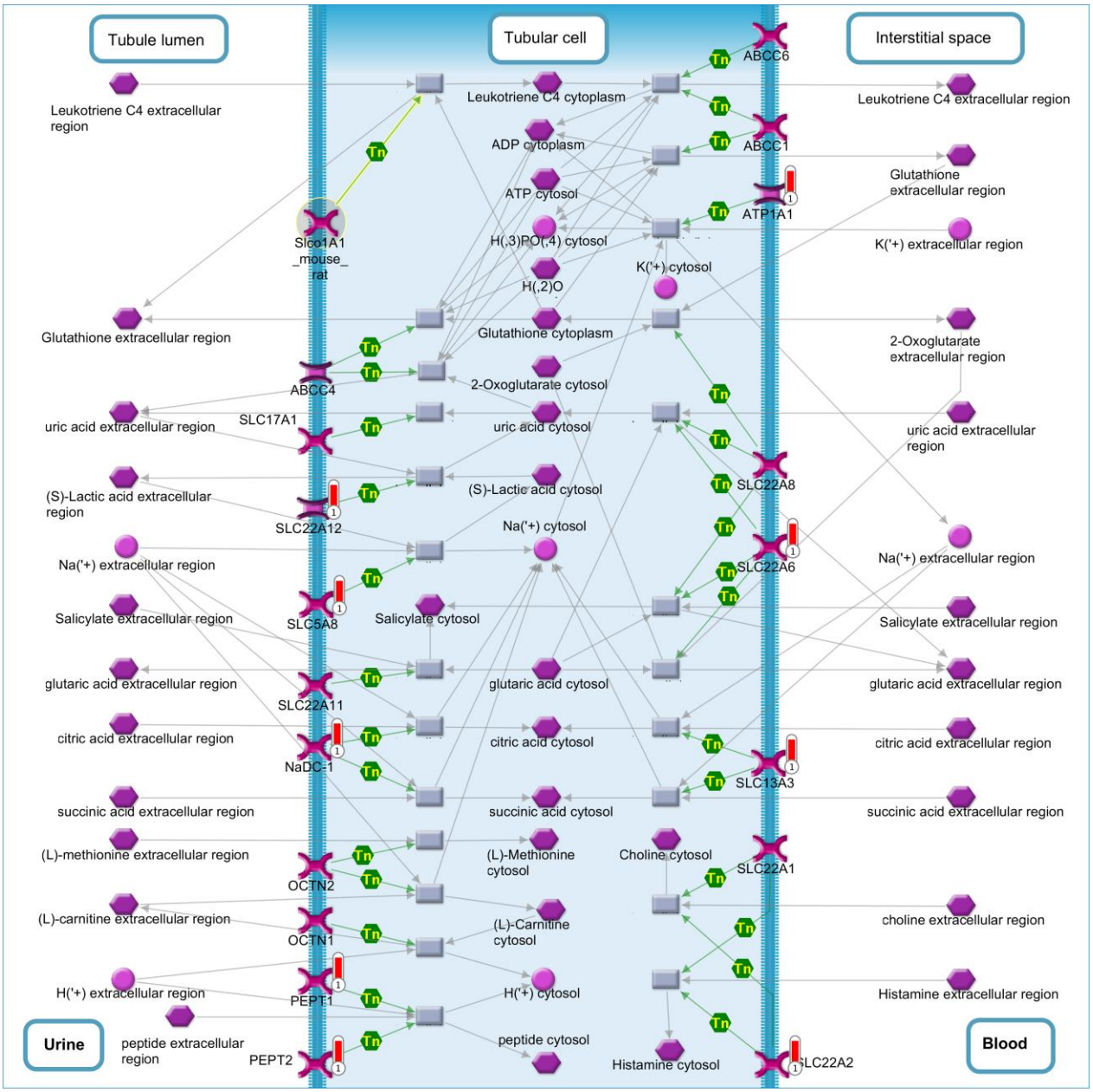
Supplementary Figure. Legends to Thompson Reuters Pathway Maps (for Suppl. Figs. S7, S8, and S9)



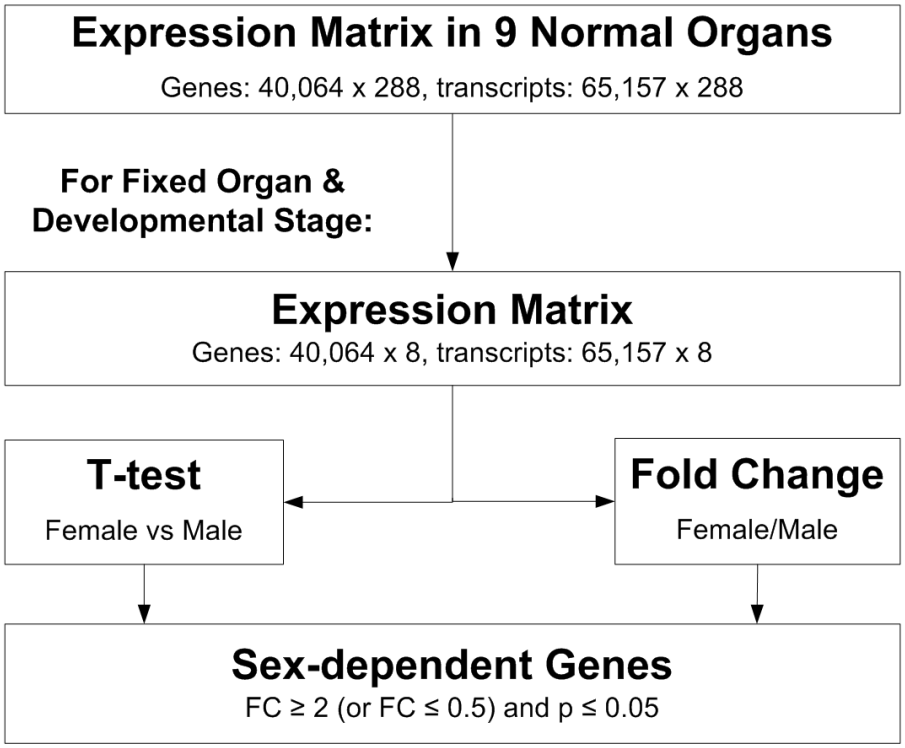
Supplementary Figure S7. Brain-enriched genes were involved in the role of CDK5 in presynaptic signaling. Genes identified as specifically enriched in the brain were uniquely overrepresented in the cyclin-dependent kinase 5 (CDK5) presynaptic signaling pathway map from Thompson Reuters Canonical Pathway Maps ontology. Enriched genes were highlighted with red thermometers on the map. CDK5 is a member of the small serine/threonine cyclin-dependent kinase family. CDK5 and its neuronspecific activator, cyclin-dependent kinase 5 regulatory subunit 1 (CDK5R1 (p35)), are all abundantly expressed in the adult brain. Moreover, they are present in subcellular fractions enriched for synaptic membranes, and localize to pre- and post-synaptic compartments, thus they can regulate both pre- and post-synaptic functions.



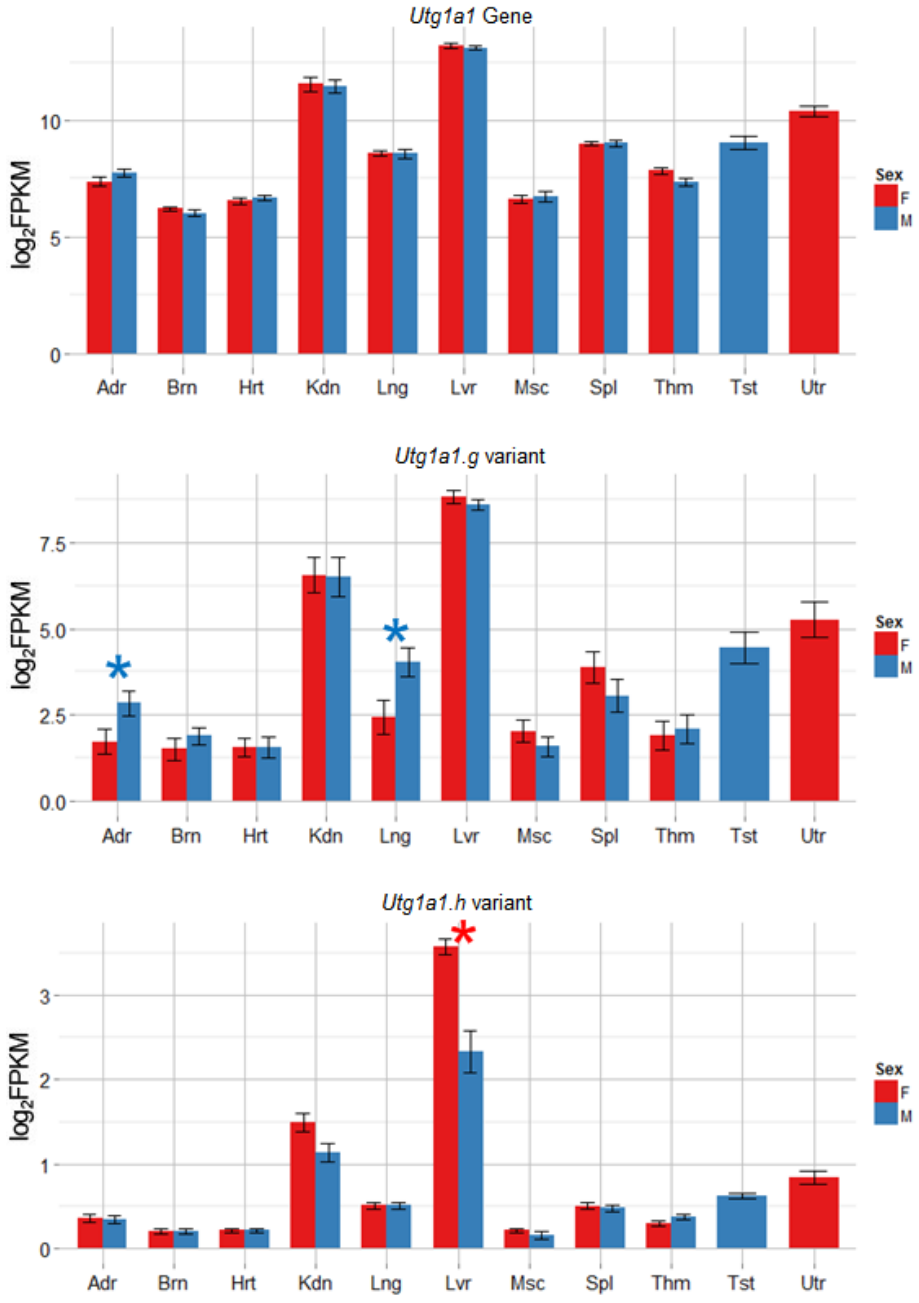
Supplementary Figure S8. Liver-enriched genes were involved in bile acid biosynthesis pathway. Genes identified as specifically enriched in the liver were uniquely overrepresented on the bile acid biosynthesis pathway map from Thompson Reuters Canonical Pathway Maps ontology. Enriched genes were highlighted with red thermometers on the map. Bile acids (BAs: cholic acid, deoxycholic acid, chenodeoxycholic acid, etc.) are amphipathic steroidal compounds derived from the enzymatic catabolism of cholesterol in hepatocytes.



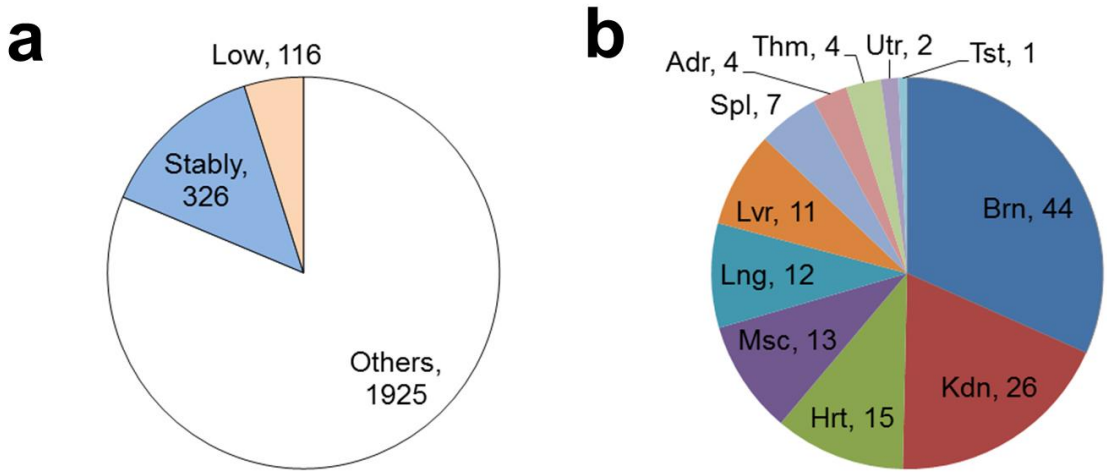
Supplementary Figure S9. Kidney-enriched genes were involved in renal secretion of organic electrolytes. Genes identified as specifically and uniquely enriched in the kidney were uniquely overrepresented in the renal secretion of organic electrolytes pathway map from Thompson Reuters Canonical Pathway Maps ontology. Enriched genes were highlighted with red thermometers on the map. Renal secretion of organic cations (including bases) and organic anions (including acids) has physiological, pharmacological, and toxicological importance. The proximal tubule is the primary site of renal electrolyte secretion.



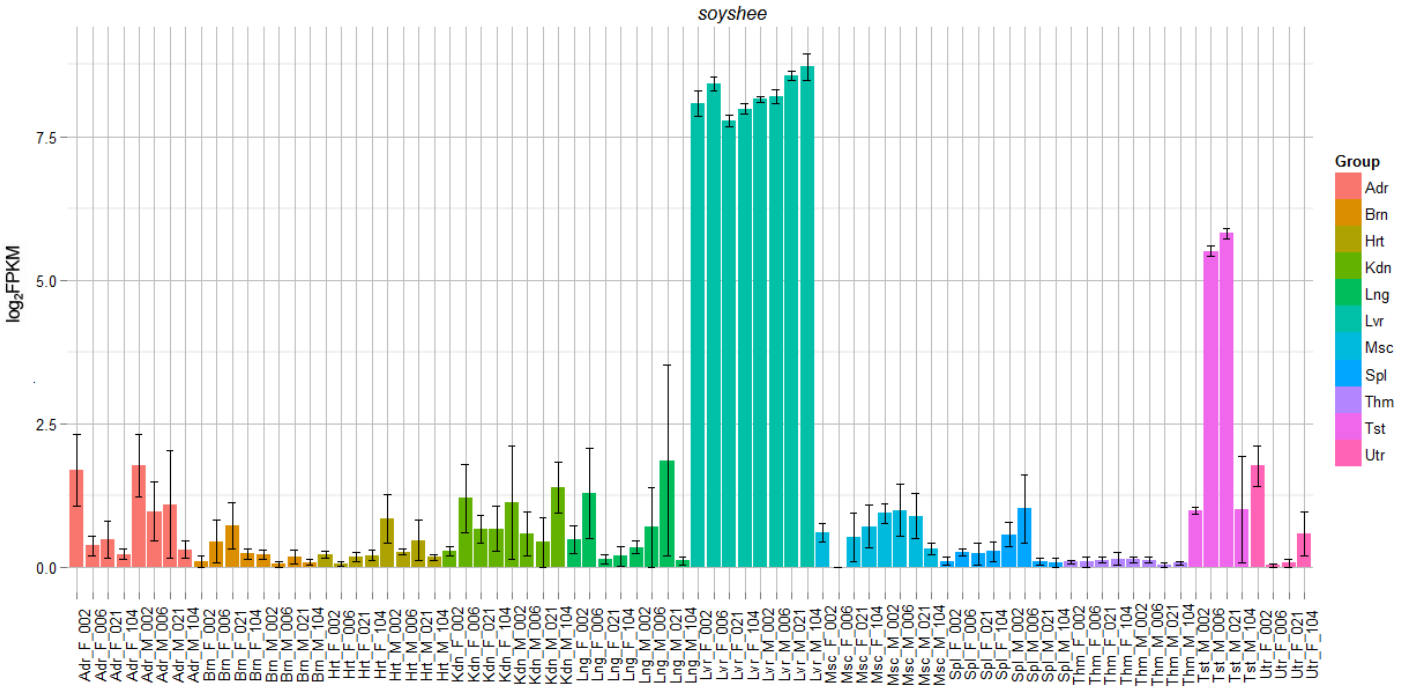
Supplementary Figure S10. Scheme for identifying sex-specific genes in each organ at each developmental stage. All 288 samples from 9 organs (except uterus and testis) were used to identify sex-specific genes. For each organ, samples were separated into four groups on the basis of developmental stage. Fold change and t-test *p*-value were calculated between males and females in each organ across the four stages. Sex-specific genes were those that had a fold change ≥ 2 (or ≤ 0.5) and *p*-value ≤ 0.05 for any organ at any development stage.



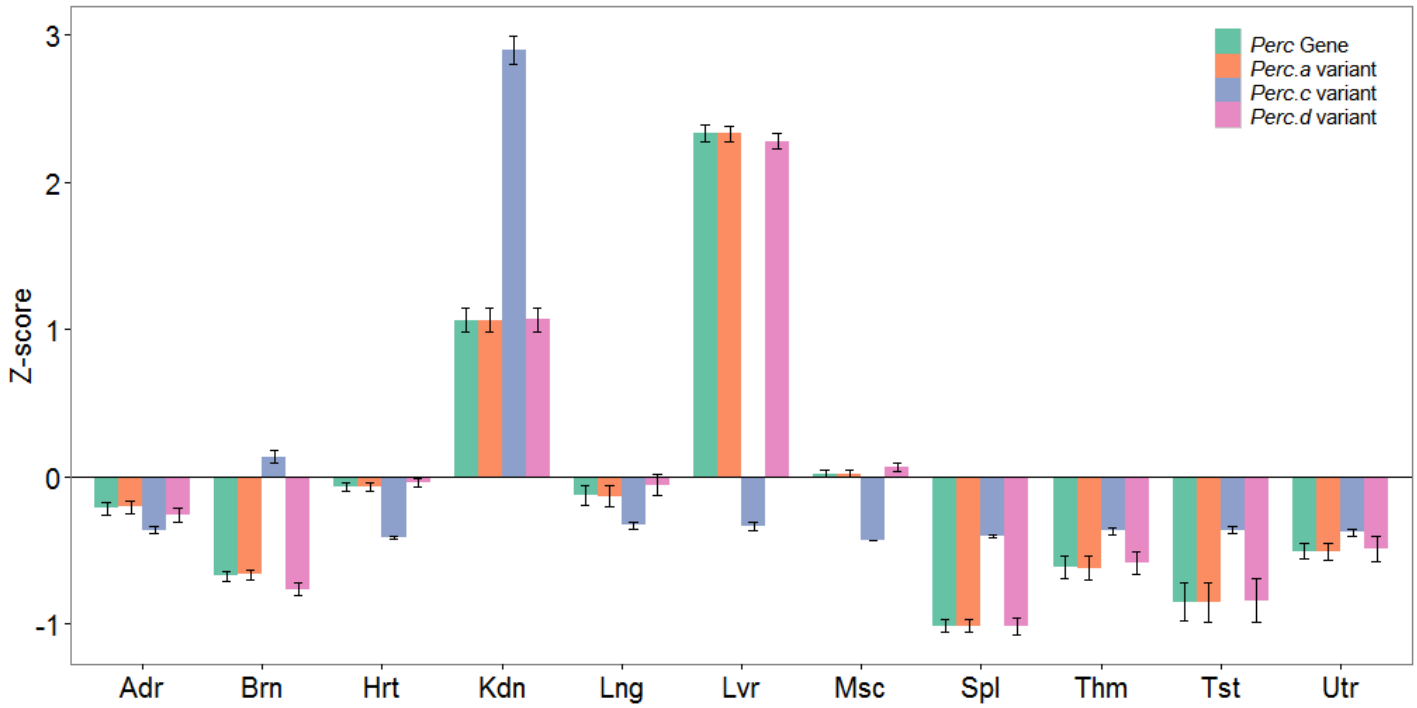
Supplementary Figure S11. Sex-specific alternatively spliced variant expression of *Ugt1a* in rats. Bars: expression (log_2FPKM, Mean ± SE; N=16 in uterus and testes, N=32 in the other 9 organs) of gene *Ugt1a* (upper), its isoform *g* (middle) and isoform *h* (bottom) in each organ in females (red) and males (blue). Stars: sex-specific (female > male, red; female < male, blue) and differently expressed in the organ. For any given organ and transcript variant, a gene/transcript variant with an FC between males and females ≥ 2 (or ≤ 0.5) with a *p*-value ≤ 0.05 (Student's t-test) was considered as a sex-specific gene/transcript variant. Adr, adrenal; Brn, brain; Hrt, heart; Kdn, kidney; Lvr, liver; Lng, lung; Msc, skeletal muscle; Spl, spleen; Tst, testis; Thm, thymus; and Utr, uterus.



Supplementary Figure S12. Expression of spliced noncoding genes in AceView. (**a**) Distribution of 2,367 spliced noncoding genes based on gene expression level. (**b**) Cross-organ distribution of 139 organ-enriched noncoding genes. A gene that was expressed at FPKM $\geqslant$ 1 in all eleven organs, across all four developmental stages, in both males and females was defined as a commonly-expressed gene. A gene that was FPKM<1in all eleven organs, across all four developmental stages in both males and females, was defined as a "low" expressed gene. Adr, adrenal; Brn, brain; Hrt, heart; Kdn, kidney; Lvr, liver; Lng, lung; Msc, skeletal muscle; Spl, spleen; Tst, testis; Thm, thymus; and Utr, uterus.



Supplementary Figure S13. Expression profile of the spliced novel non-coding gene *soyshee*. Y-axis shows the log_2FPKM expression level (Mean ± SE; N=4). Adr, adrenal; Brn, brain; Hrt, heart; Kdn, kidney; Lvr, liver; Lng, lung; Msc, skeletal muscle; Spl, spleen; Tst, testis; Thm, thymus; and Utr, uterus.



Supplementary Figure S14. Expression of the *Pecr* gene and its alternatively spliced variants in rats. The Y-axis represents the Z-score (Mean 0, SD 1) of $\log_2$FPKM (Mean ± SE; N=16 in uterus and testes, N=32 in the other 9 organs) and the X-axis shows the *Pecr* gene and its transcript variant *Pecr.a*, *Pecr.c*, and *Pecr.d* (*Pecr* gene in green; *Pecr.a* in orange; *Pecr.c* in blue; and *Perc.d* in pink) in each organ. Adr, adrenal; Brn, brain; Hrt, heart; Kdn, kidney; Lvr, liver; Lng, lung; Msc, skeletal muscle; Spl, spleen; Tst, testis; Thm, thymus; and Utr, uterus.

Supplementary Tables

Supplementary Table S1. Significantly enriched GeneGo canonical pathway maps based on the commonly-expressed genes in all organs across the developmental stages.

Maps	# Total Genes	# On Map	p Value*	Min FDR
Oxidative phosphorylation	59	58	1.39E-07	1.06E-04
GTP-XTP metabolism	44	43	1.34E-05	3.69E-03
Cytoskeleton remodeling	65	61	1.46E-05	3.69E-03
Cell cycle_Influence of Ras and Rho proteins on G1/S Transition	40	39	4.45E-05	7.94E-03
Transport_Intracellular cholesterol transport in norm	38	37	8.07E-05	7.94E-03
Ubiquinone metabolism	38	37	8.07E-05	7.94E-03
Transport_Clathrin-coated vesicle cycle	52	49	8.21E-05	7.94E-03
Cytoskeleton remodeling_TGF, WNT and cytoskeletal remodeling	76	69	8.39E-05	7.94E-03
Transcription_Sin3 and NuRD in transcription regulation	28	28	1.30E-04	1.09E-02
NRF2 regulation of oxidative stress response	43	41	1.54E-04	1.17E-02

*The enrichment analysis was performed using a hypergeometric test to find over-representation of genes. Only statistically significant pathways with p-values < 0.05 were considered. The Benjamini–Hochberg FDR method was used to adjust the p-values for multiple testing.

Supplementary Table S2. The number of organ-enriched genes* based on different fold changes.

FC	Adr	Brn	Hrt	Kdn	Lvr	Lng	Msc	Spl	Tst	Thm	Utr	Total
2	107	1,401	144	386	454	232	306	183	25	62	113	3,413
4	34	778	67	204	319	87	153	59	18	38	23	1,780
8	20	392	39	132	235	36	93	23	13	14	3	1,000
16	12	186	24	82	190	23	72	15	10	7	1	622
32	7	81	11	46	144	18	46	9	7	2	0	371
64	6	32	4	32	65	13	21	4	6	1	0	184
128	3	11	0	21	6	11	2	0	2	1	0	57
256	1	3	0	10	1	7	2	0	0	0	0	24

*****A gene was considered organ-enriched if its expression level (FPKM) was at least 2-fold higher than that in any other organs with a Bonferroni-corrected *p*-value ≤ 0.05 across the four developmental stages. FC: fold change.

Supplementary Table S3. The most characteristic pathways and biological processes identified in organ-enriched genes.

Tissue	Characteristic pathways and processes
Brain	Neurophysiological processes: dopamine signaling, CDK5 signaling, GABA signaling.
Adrenal gland	Pathways dominated by those involving biosynthesis and metabolism, e.g. metabolism of catecholamines.
Heart	Significant pathways were associated with muscle contraction, signal transduction and regulation of cardiac hypertrophy.
Kidney	Top pathways were associated with regulation of renal functions, e.g. secretion of drugs. Other pathways included those associated with metabolism.
Liver	Significant pathways were associated with various metabolic processes e.g. fatty acid oxidation, bile acid biosynthesis.
Lung	Top pathways were directly associated with cell adhesion e.g. endothelial cell contacts. Other developmental pathways were also present, e.g. BMP signaling and NOTCH induced EMT.
Muscle	Significant pathways were mostly enriched for processes involving glycolysis and gluconeogenesis, and also regulation of acetly CoA carboxylase 2 activity in muscle.
Spleen	Significant pathways were mostly enriched for those associated with immune response.
Testes	Significant pathways were associated with glutathione metabolism.
Thymus	Significant pathways were associated with various immune related processes and signaling pathways.
Uterus	Significant pathways were associated cytoskeletal removeling and developmental processes.

In general, selection and ranking of the pathways was highly consistent with the organ function. Conversely, genes defined as commonly-expressed tended to be enriched in non-organ specific pathways.

Supplementary Table S4. Numbers of development-dependent differentially expressed genes in the various organs.

Organ	002_006	002_021	002_104	006_021	006_104	021_104	Unique
Adr	1,689	2,214	2,655	630	1,399	862	3,679
Brn	749	1,189	1,252	649	540	728	2,211
Hrt	1,073	1,593	1,847	652	844	491	2,633
Kdn	1,984	2,521	2,408	650	959	739	3,561
Lvr	1,980	2,669	2,601	1,169	1,502	1,283	4,220
Lng	850	1,594	2,340	594	1,253	965	3,185
Msc	1,720	3,035	3,906	1,267	2,430	1,191	5,064
Spl	1,623	2,245	1,900	764	1,079	1,060	3,554
Tst	11,989	12,587	6,591	1,133	12,348	12,708	16,186
Thm	695	1,232	5,554	1,059	5,390	4,302	6,499
Utr	2,662	3,075	3,033	1,306	1,696	1,271	4,847

To identify development-dependent genes in each organ, we used a fold-change ≥ 2 (or ≤ 0.5) plus Bonferroni-adjusted *p*-value ≤ 0.05 based on an ANOVA model to select differentially expressed genes among the developmental stages. The first 6 columns show the number of DEGs in comparison between two developmental stages for the organ. The last column shows the total number of DEGs for the organ. Adr, adrenal; Brn, brain; Hrt, heart; Kdn, kidney; Lvr, liver; Lng, lung; Msc, skeletal muscle; Spl, spleen; Tst, testis; Thm, thymus; and Utr, uterus.

Supplementary Table S5. Number of genes that were female- or male-dominant.

Organ	Week 2		Week 6		Week 21		Week 104	
	Female*	Male	Female	Male	Female	Male	Female	Male
Adr	166	350	317	236	214	469	221	397
Brn	265	191	189	157	206	448	200	127
Hrt	176	462	224	122	199	153	97	505
Kdn	194	291	446	219	528	397	199	265
Lvr	322	435	328	267	715	495	295	377
Lng	312	321	182	90	253	131	156	359
Msc	143	763	233	231	203	412	427	127
Spl	186	308	161	139	249	200	195	174
Thm	157	288	150	207	202	240	326	240

*Female: number of genes that were female-dominant, DEG between female and male: FC $\geq$ 2 (Female/Male) with T-test *p*-value $\leq$ 0.05. Male-dominant genes were calculated in a similar manner; FC $\leq$ 0.5 (Female/Male) with T-test *p*-value $\leq$ 0.05.

All 288 samples (except uterus and testis) were separated on the basis of developmental stages and organ. Fold change and t-test *p*-value were calculated between males and females for each organ and developmental stage. For any organ at any development stage, genes with a fold change $\geq$ 2 and *p*-value $\leq$ 0.05 were considered sex-specific.

Supplementary Table S6. The 92 ERCC transcripts and their NCBI accession numbers.

ERCC_ID	GenBank	ERCC_ID	GenBank	ERCC_ID	GenBank	ERCC_ID	GenBank
ERCC-00002	DQ459430	ERCC-00046	DQ516748	ERCC-00085	DQ883669	ERCC-00138	DQ516777
ERCC-00003	DQ516784	ERCC-00048	DQ883671	ERCC-00086	DQ516791	ERCC-00142	DQ883646
ERCC-00004	DQ516752	ERCC-00051	DQ516740	ERCC-00092	DQ459425	ERCC-00143	DQ668362
ERCC-00009	DQ668364	ERCC-00053	DQ516785	ERCC-00095	DQ516759	ERCC-00144	DQ854995
ERCC-00012	DQ883670	ERCC-00054	DQ516731	ERCC-00096	DQ459429	ERCC-00145	DQ875386
ERCC-00013	EF011062	ERCC-00057	DQ668366	ERCC-00097	DQ516758	ERCC-00147	DQ516790
ERCC-00014	DQ875385	ERCC-00058	DQ459418	ERCC-00098	DQ459415	ERCC-00148	DQ883642
ERCC-00016	DQ883664	ERCC-00059	DQ668356	ERCC-00099	DQ875387	ERCC-00150	DQ883659
ERCC-00017	DQ459420	ERCC-00060	DQ516763	ERCC-00104	DQ516815	ERCC-00154	DQ854997
ERCC-00019	DQ883651	ERCC-00061	DQ459426	ERCC-00108	DQ668365	ERCC-00156	DQ883643
ERCC-00022	DQ855004	ERCC-00062	DQ516786	ERCC-00109	DQ854998	ERCC-00157	DQ839618
ERCC-00024	DQ854993	ERCC-00067	DQ883653	ERCC-00111	DQ883685	ERCC-00158	DQ516795
ERCC-00025	DQ883689	ERCC-00069	DQ459421	ERCC-00112	DQ459422	ERCC-00160	DQ883658
ERCC-00028	DQ459419	ERCC-00071	DQ883654	ERCC-00113	DQ883663	ERCC-00162	DQ516750
ERCC-00031	DQ459431	ERCC-00073	DQ668358	ERCC-00116	DQ668367	ERCC-00163	DQ668359
ERCC-00033	DQ516796	ERCC-00074	DQ516754	ERCC-00117	DQ459412	ERCC-00164	DQ516779
ERCC-00034	DQ855001	ERCC-00075	DQ516778	ERCC-00120	DQ854992	ERCC-00165	DQ668363
ERCC-00035	DQ459413	ERCC-00076	DQ883650	ERCC-00123	DQ516782	ERCC-00168	DQ516776
ERCC-00039	DQ883656	ERCC-00077	DQ516742	ERCC-00126	DQ459427	ERCC-00170	DQ516773
ERCC-00040	DQ883661	ERCC-00078	DQ883673	ERCC-00130	EF011072	ERCC-00171	DQ854994
ERCC-00041	EF011069	ERCC-00079	DQ883652	ERCC-00131	DQ855003		
ERCC-00042	DQ516783	ERCC-00081	DQ854991	ERCC-00134	DQ516739		
ERCC-00043	DQ516787	ERCC-00083	DQ516780	ERCC-00136	EF011063		
ERCC-00044	DQ459424	ERCC-00084	DQ883682	ERCC-00137	DQ855000		