

Supplemental Tables

Table S1 A robustness test for the TSG distribution in forest and prairie domains

	Liver	Cortex	Hippo	Lung	LV	Spleen	Ovary	Adrenal	Aorta	Panc
F_t	407	610	555	246	136	371	132	167	51	323
F_{nt}	14663	14460	14515	14824	14934	14699	14938	14903	15019	14747
P_t	151	284	261	144	43	160	35	62	18	118
P_{nt}	3192	3059	3082	3199	3300	3183	3308	3281	3325	3225
P -val	1.3e-7	8.3e-24	1.8e-12	8.1e-19	-	9.7e-12	-	7.4e-4	-	6.1e-6

Tissue-specific genes are identified based on the criterion $s_i^t > 1$. F_t , F_{nt} , P_t and P_{nt} represent the number of forest TSG, forest non-TSG, prairie TSG and prairie non-TSG in corresponding tissue, respectively. Fisher's exact test was performed and "-" indicates the corresponding p-value is larger than 0.05. Hippo=hippocampus, LV=left ventricle, Panc=pancreas.

Table S2 Functional module information

Tissue name	Chromatin	Gene number	Location/bp
Liver	chr4	11 ^a /37 ^b	69215613-72897763
Spleen	chr19	12/27	54711241-55450974
Adrenal	chr6	4/6	52535852-52859665
Pancreas	chr1	10/20	101702596-108507545

^aThe number of tissue-specific genes belonging to the corresponding tissue in this functional module (prairie domain). ^bThe number of all genes in this functional module. For instance, 11 out of 37 genes located in the liver functional module are liver-specific genes.

Table S2-1 Representative liver-specific genes and related functions in liver functional module

<i>UGT2B10</i> , <i>UGT2B17</i> , <i>UGT2B15</i> , <i>UGT2A3</i> , <i>UGT2B7</i>	Drug metabolism; porphyrin and chlorophyll metabolism
GC	Vitamin D transport, storage and metabolism

Table S2-2 Representative spleen-specific genes and related functions in spleen functional module

<i>LILRA4</i>	Encoding the immunoglobulin-like cell surface protein and Modulating the immune response
<i>KIR2DL3, KIR3DX1, KIR3DL1</i>	Killer cell immunoglobulin-like receptors
<i>NCR1</i>	Contributing to the increased efficiency of activated natural killer cells
<i>NLRP7</i>	Feedback regulator of caspase-1-dependent interleukin 1 beta secretion

Table S2-3 Representative adrenal-specific genes and related functions in adrenal functional module

<i>GSTA7P, GSTA1, GSTA2, GSTA3</i>	Adding glutathione to electrophilic compounds, an important step in detoxification of these compounds
------------------------------------	--

Table S2-4 Representative pancreas-specific genes and related functions in pancreas functional module

<i>AMY1A, AMY1B, AMY1C, AMY2A</i>	Catalyzing the first step in digestion of dietary starch and glycogen
-----------------------------------	--

Table S3. Examples of high order correlation between F-P gene pairs and TFs in liver

Forest gene	Prairie gene	TF
<i>CYP4A11</i>	<i>CYP2J2</i>	NR1I3
<i>ACADM</i>	<i>GBP7</i>	NR1I3
<i>AK4</i>	<i>GBP7</i>	NR1I3
<i>TMEM56</i>	<i>DPYD</i>	ESR1
<i>CA14</i>	<i>FMO3</i>	ARID3C

In each example, the forest and prairie genes are not only highly correlated but forming strong spatial contact in liver. Furthermore, the expression level of corresponding TF correlates with not only the individual forest and prairie genes in liver, but the correlation level (in different tissues) between them.

Table S4 Threshold and size of Hi-C data (chr1) in ten tissues

	Liver	Cortex	Hippocampus	Lung	LV ^a	Spleen	Ovary	Adrenal	Aorta	Pancreas
α^t	1.8960	0.4272	0.4790	0.2249	2.0133	0.4049	0.3242	0.3620	1.3111	0.3754
Before ^b	4371504	1588909	1780322	1040669	2626072	1533076	1406427	1385049	3549393	1476175
After ^c	1187231	1563662	1726100	1034985	1488972	1507728	1393523	1367474	1569774	1432612

^aLeft ventricle. ^bThe number of non-zero elements in the Hi-C data of each tissue. ^cThe number of elements, the value of which is larger than α^t .

Supplemental Figures

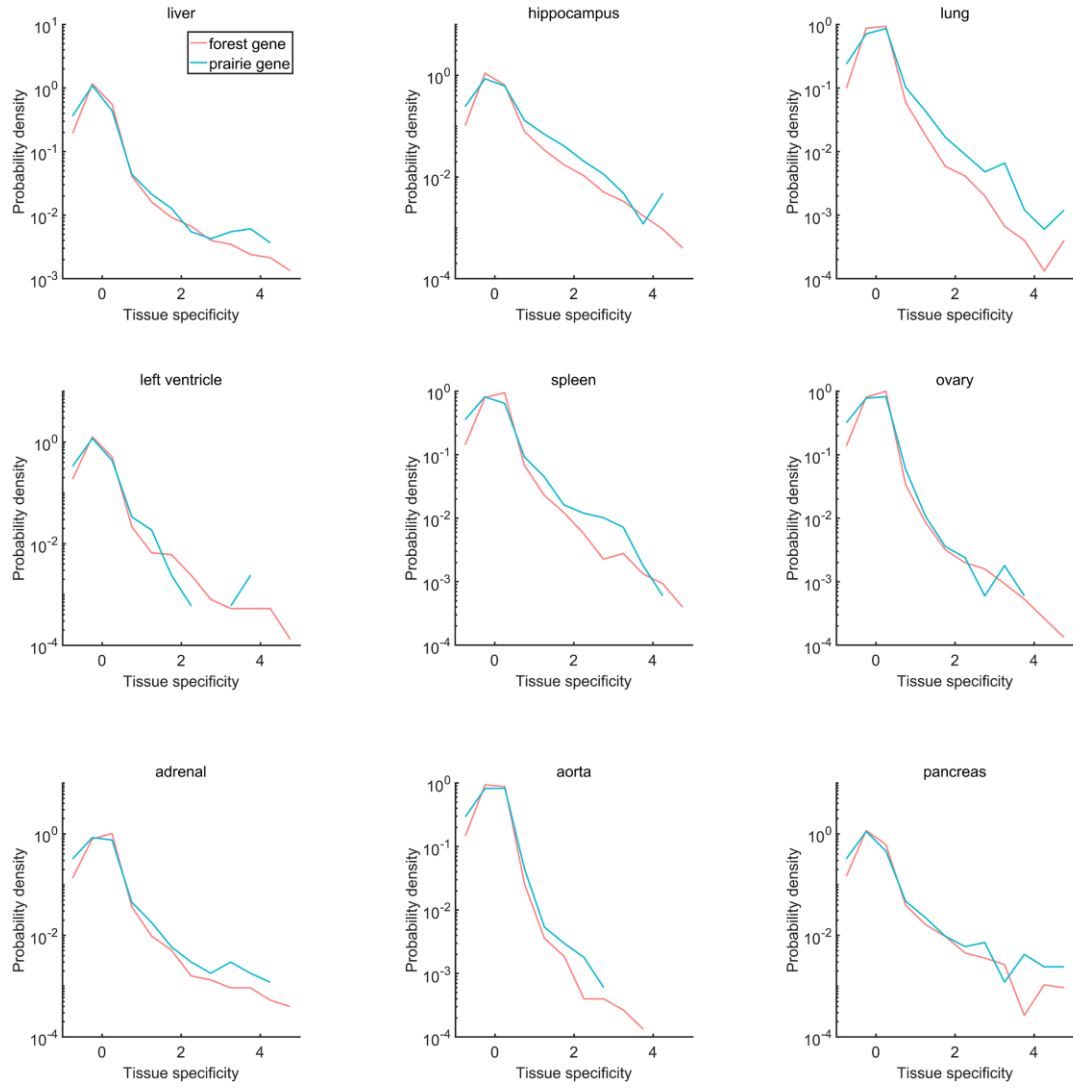


Fig. S1. The probability density distribution of tissue specificities of forest and prairie genes in nine tissues.

The range of gene tissue specificity was chosen as $[-1,5]$, resulting in the missing of several genes in some tissues (the total number of forest and prairie genes is 18413). Related to **Fig. 1A**.

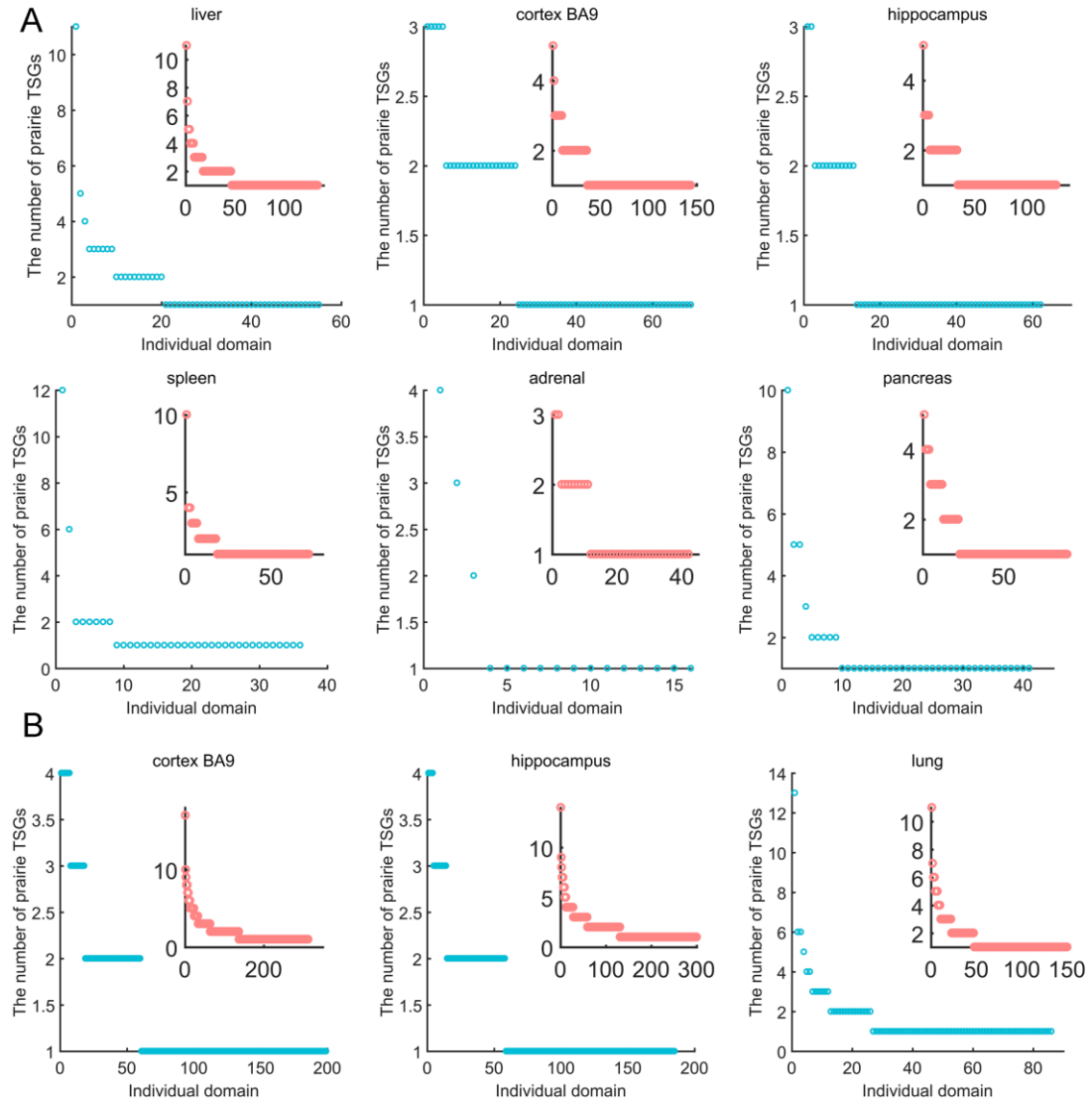


Fig. S2. The number of TSGs relevant to one tissue in individual prairie domains (inner, forest). Therefore, in each figure, the data point represents the number of relevant TSGs in one certain prairie/forest domain. We identified several prairie domains, named functional modules that significantly harbor TSGs belonging to one tissue, in liver, spleen, adrenal and pancreas. In contrast, two brain tissues, cortex BA9 and hippocampus, exhibit the broader and more uniform distribution of relevant prairie TSGs and thereby do not possess functional modules under the two criteria in TSG identification: $s_i^t > 2$ (A) and $s_i^t > 1$ (B). Related to **Fig. 1C**.

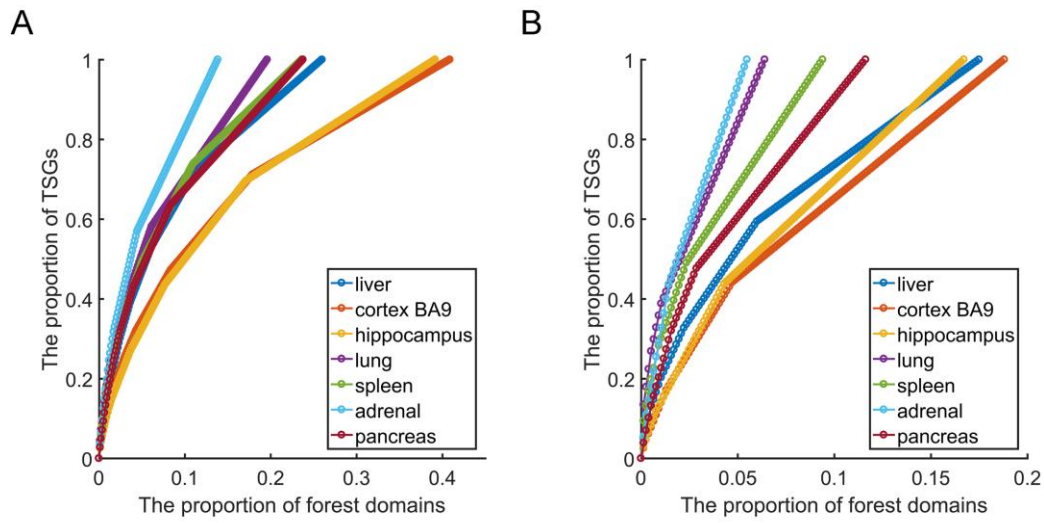


Fig. S3. “Saturation curve” for forest TSGs under the criteria $s_i^t > 1$ (A) and $s_i^t > 2$ (B). Related to Fig.

1C.

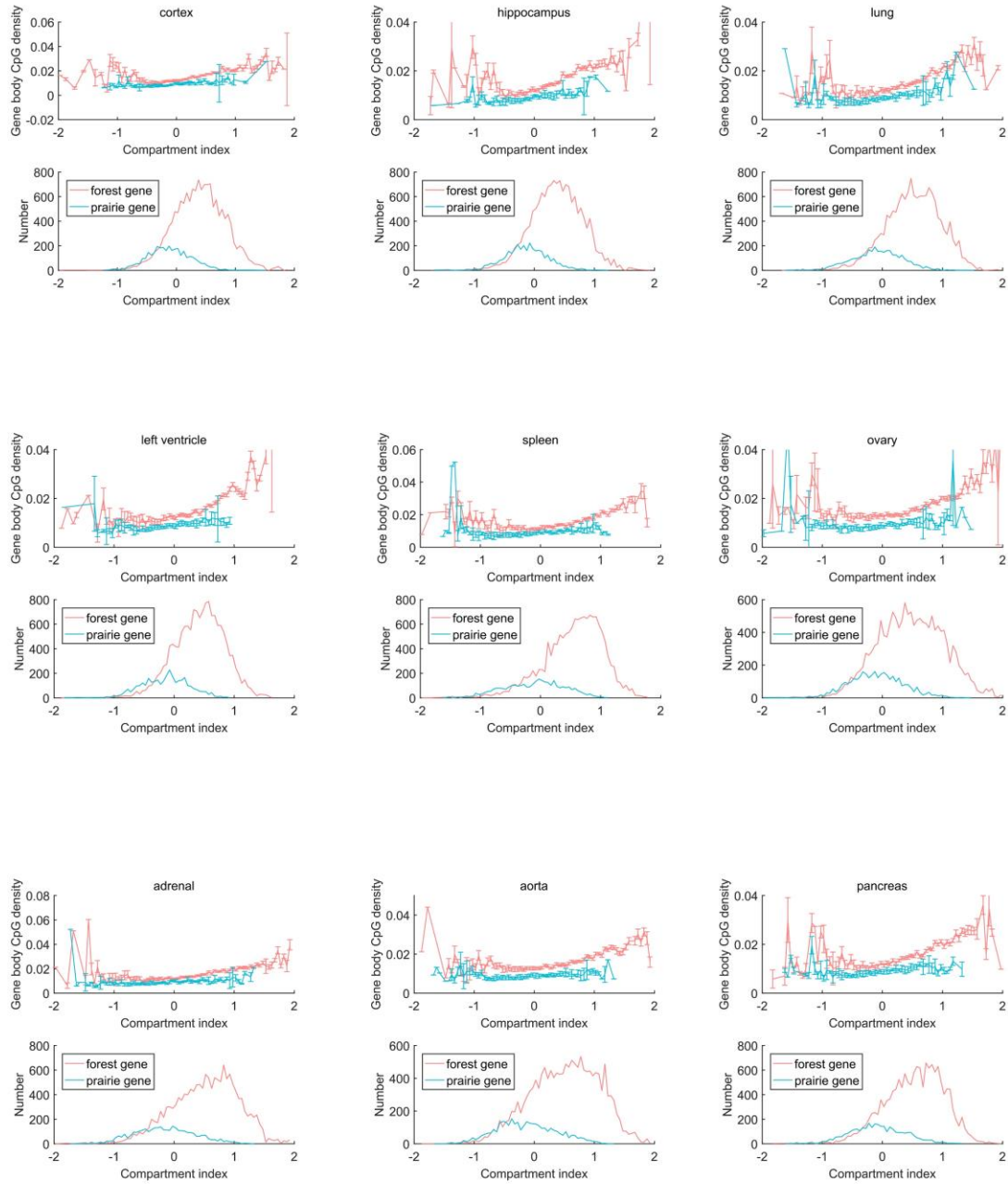


Fig. S4. The relation between compartment index and gene body CpG density in nine tissues. The range of compartment index was chosen as $[-2, 2]$, resulting in the missing of several data points in some tissues. For forest genes, the Pearson correlation coefficient was larger than 0.2 and smaller than 0.3 (the corresponding p -value $< 1e-100$) in these nine tissues. For prairie genes, the Pearson correlation coefficient was larger than 0.05 and smaller than 0.11 (the corresponding p -value < 0.005) in eight tissues except for aorta. Related to **Fig. 2B**.

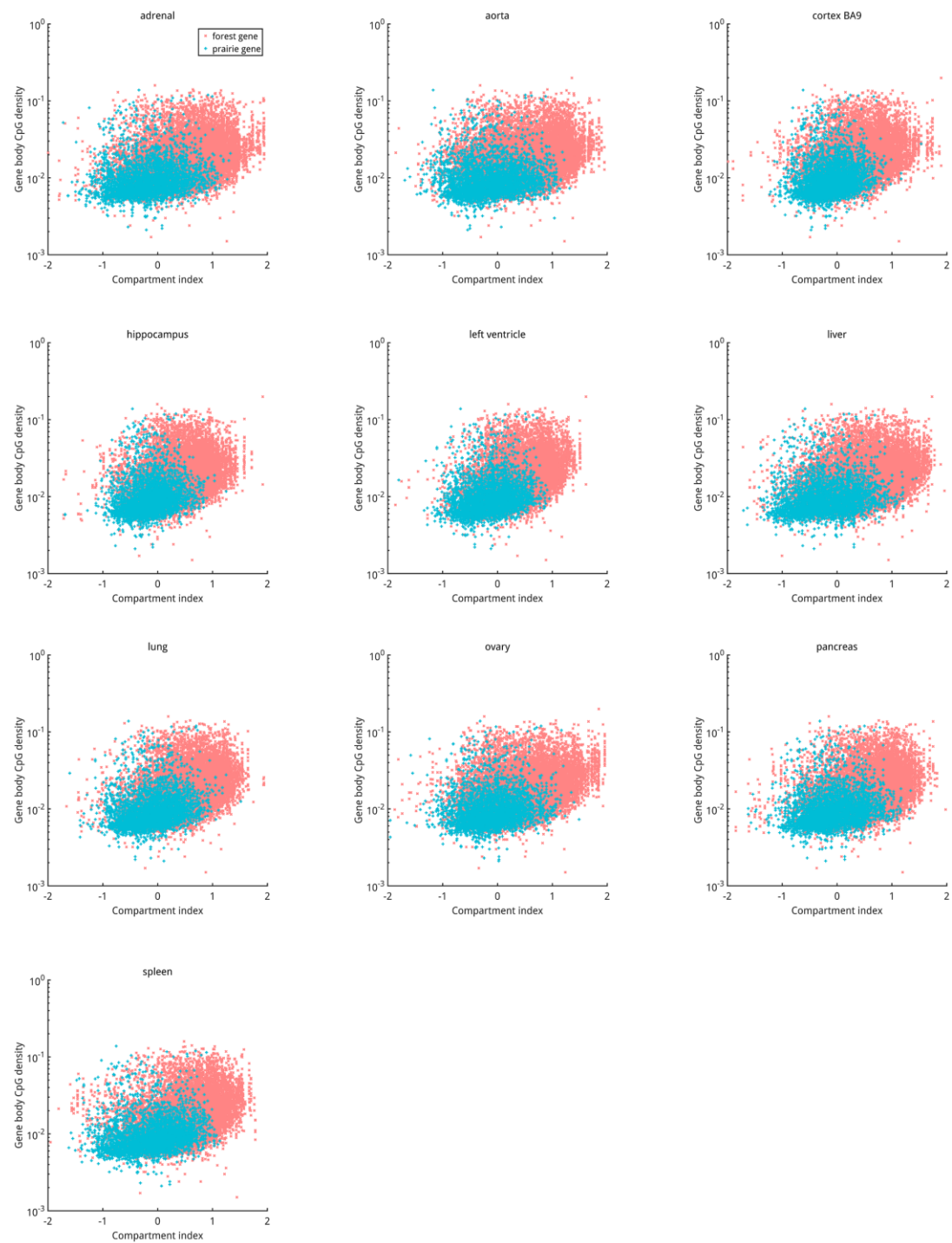


Fig. S5. Scatter plot for the relation between compartment index and gene body CpG density. Related to **Fig. 2B** and **S4**.

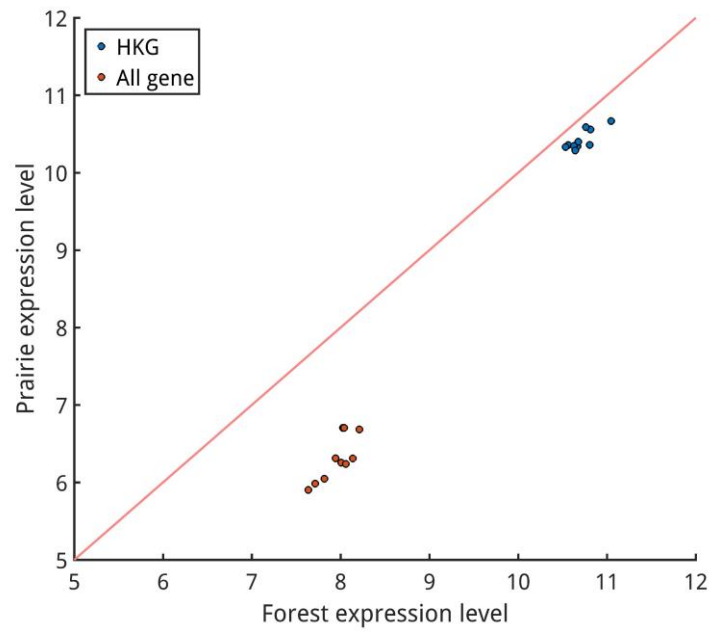


Fig. S6. The difference of gene/HKG expression levels between forest and prairie. Ten tissues, liver, cortex BA9, hippocampus, lung, left ventricle, spleen, ovary, adrenal, aorta and pancreas, were included and each data point represents the average expression level in one of the ten tissues.

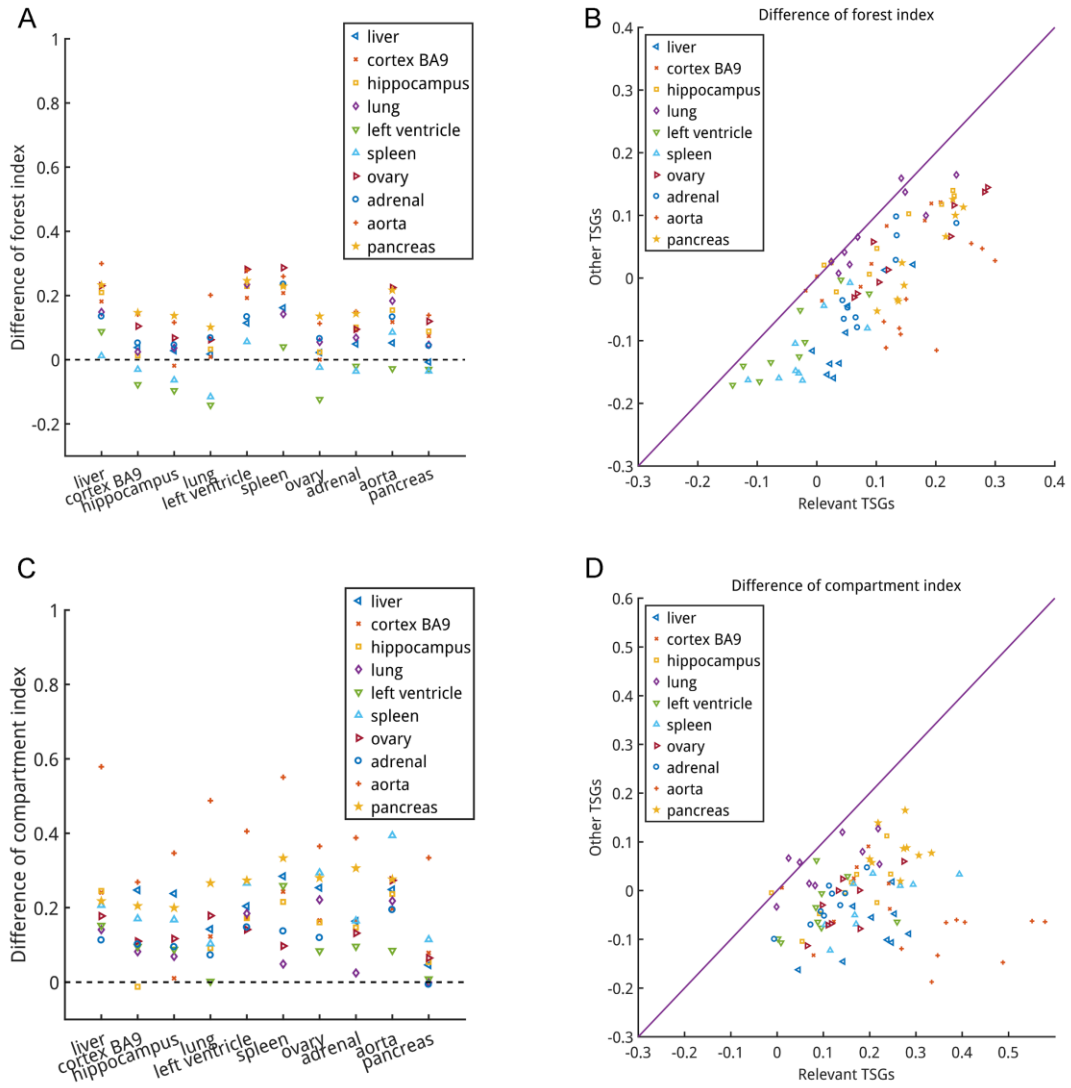


Fig. S7. A robustness test for the relation between prairie gene regulation and genome organization. The criterion we used here in TSG identification is $s_i^t > 1$. Related to **Fig. 3B, 3C**.

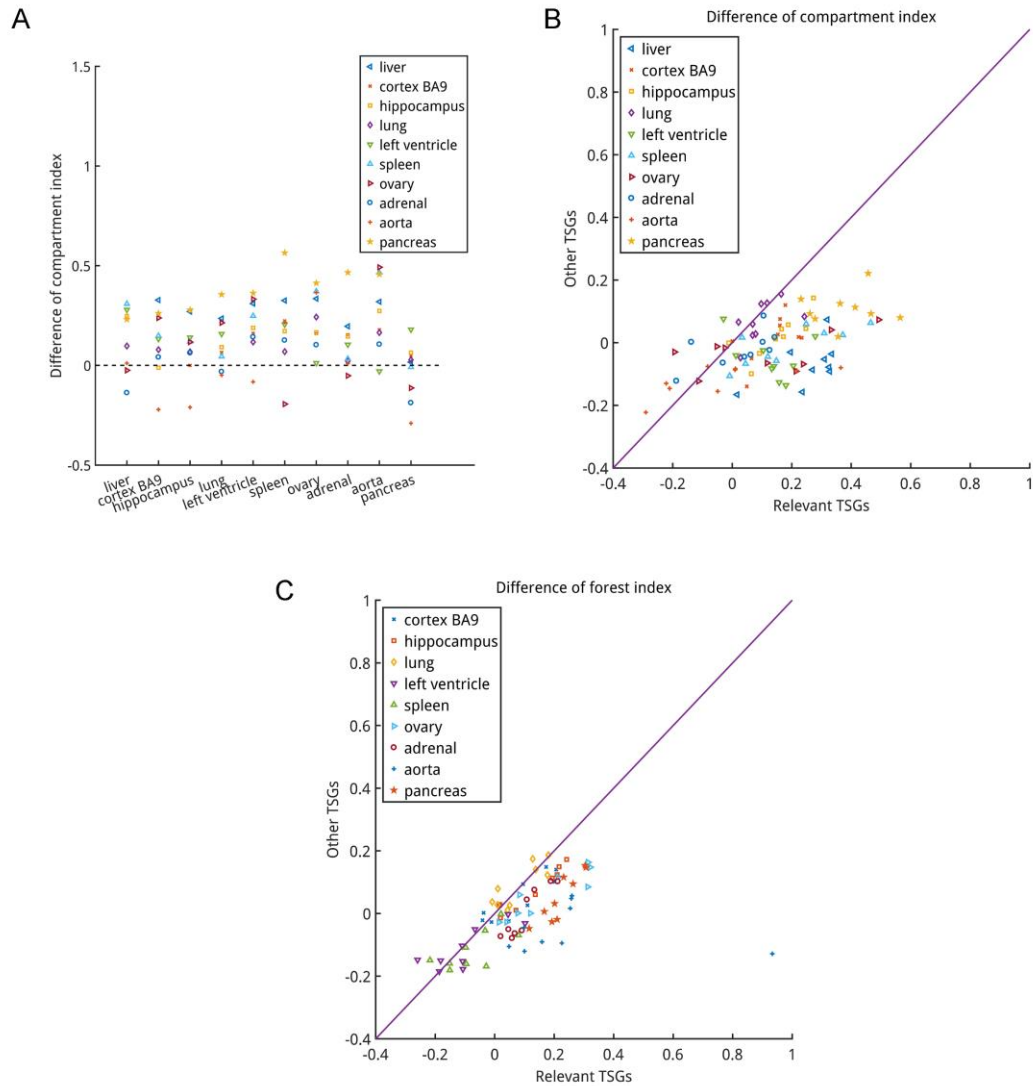


Fig. S8. The regulation of prairie TSGs is strongly associated with 3D genome organization. **A:** Compartment index change of relevant prairie TSGs (TSGs belonging to one certain tissue, e.g., liver prairie TSGs) from the nine control tissues (e.g. spleen) to the relevant tissue (e.g. liver). The value behind each symbol represents the median compartment index difference calculated between relevant tissue (illustrated in legend) and control tissue (labeled in x-axis) for relevant tissue's prairie TSGs. For instance, pancreas (five-pointed star) has nine values (median compartment index differences), calculated between pancreas and nine control tissues for the prairie TSGs of pancreas, respectively. Similar to forest index, almost all values are positive, indicating that when tissue type changes from T1 to T2, at least half of prairie TSGs belonging to T2 enter a more compartment A environment for activation. **B-C:** Comparison of **(B)** compartment and **(C)** forest index changes between relevant prairie TSGs (prairie TSGs of one certain tissue, e.g. liver prairie TSGs) and prairie TSGs specific to other tissues (complementary prairie TSGs, e.g., non-liver prairie TSGs) when tissue type changes from one other tissue (e.g. spleen) to

relevant tissue (e.g. liver). For instance, in **Fig. S8B**, pancreas (five-pointed star) has nine data points, in each data point, the corresponding x ('Relevant TSGs') and y ('Other TSGs') values represent the compartment index changes of pancreas prairie TSGs and non-pancreas prairie TSGs from one of the nine control tissue to pancreas, respectively. In each figure the median values were shown. In such a process, we observed a higher and more positive value of compartment/forest index change for relevant TSGs, indicating a more favorable transcription environment they entered, compared to TSGs of the other tissues. Together, these results emphasize the tight link between high order chromatin structure and prairie gene expression. Related to **Fig. 3B and 3C**.

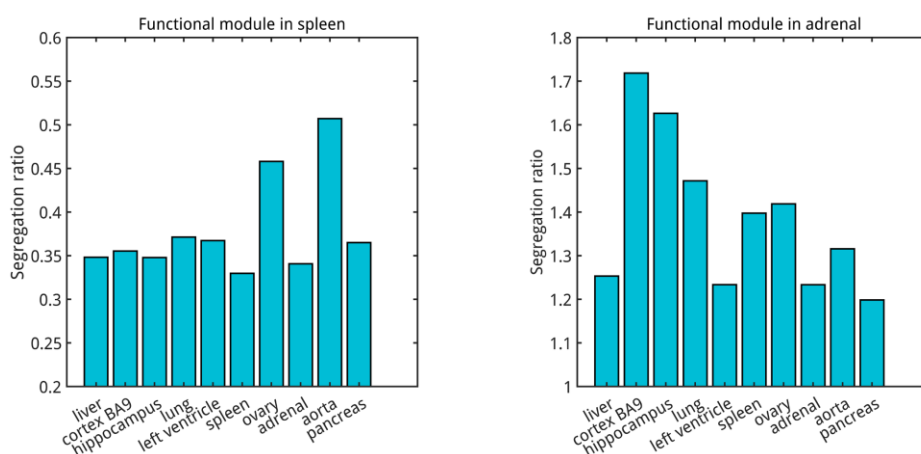


Fig. S9. The segregation ratio of spleen and adrenal functional modules in ten tissues. We noticed that the segregation ratio of adrenal functional module in adrenal is a litter bigger than that in pancreas and then check the tissue specificity of genes residing in this domain in pancreas. The results revealed that three out of six genes also possess high tissue specificity in pancreas: *GSTA7P*, 8.73; *GSTA2*, 4.69; *GSTA1*, 1.80. Therefore, similar to adrenal, these genes also need to intermingle with forest regions for activation in pancreas. Related to **Fig. 3E**.

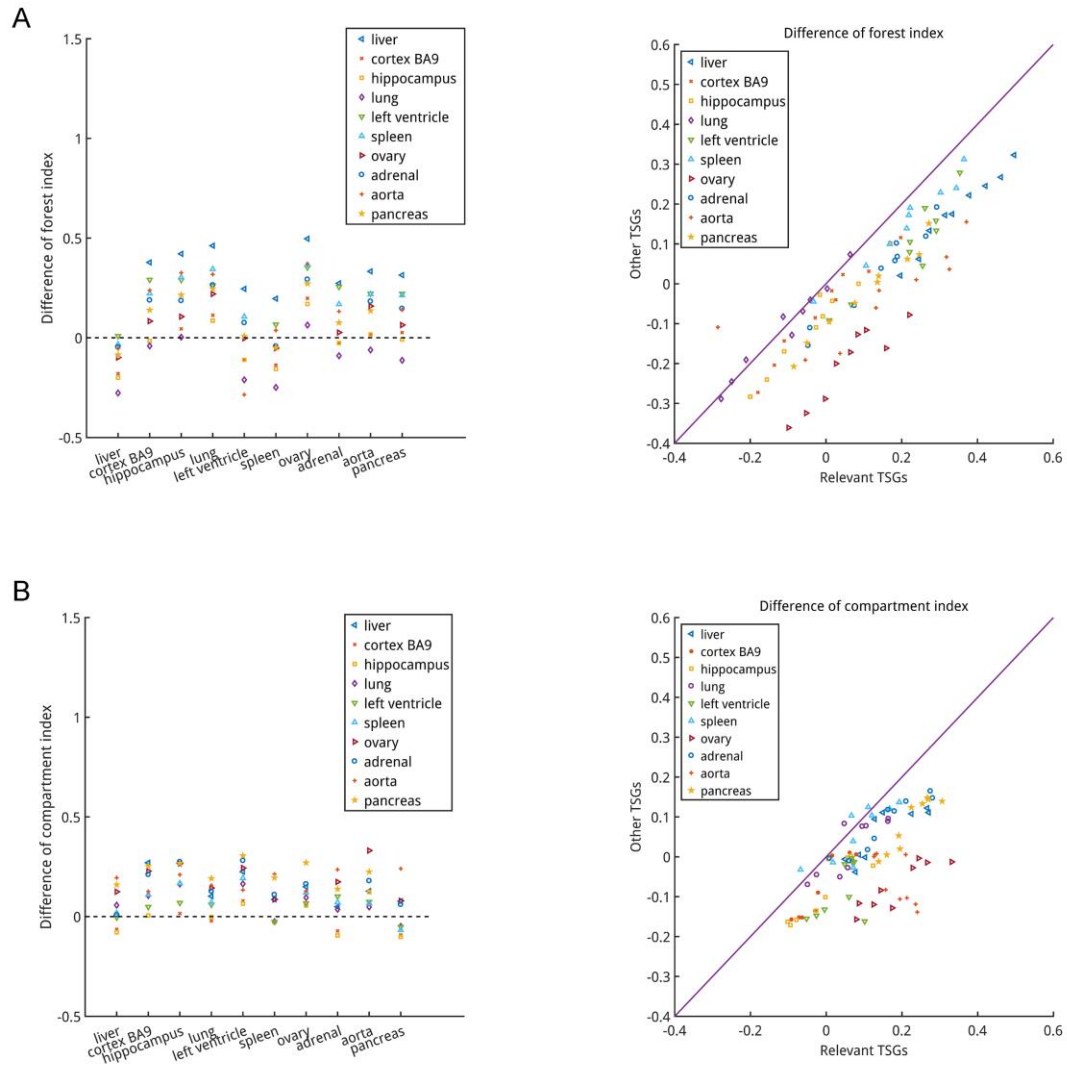


Fig. S10. The regulation of forest TSGs is also associated with genome organization. **(A-B)** Akin to prairie TSGs, the regulation of forest TSGs are also related to **(A)** forest and **(B)** compartment index changes, i.e., the chromatin structure reorganization (median values were shown).

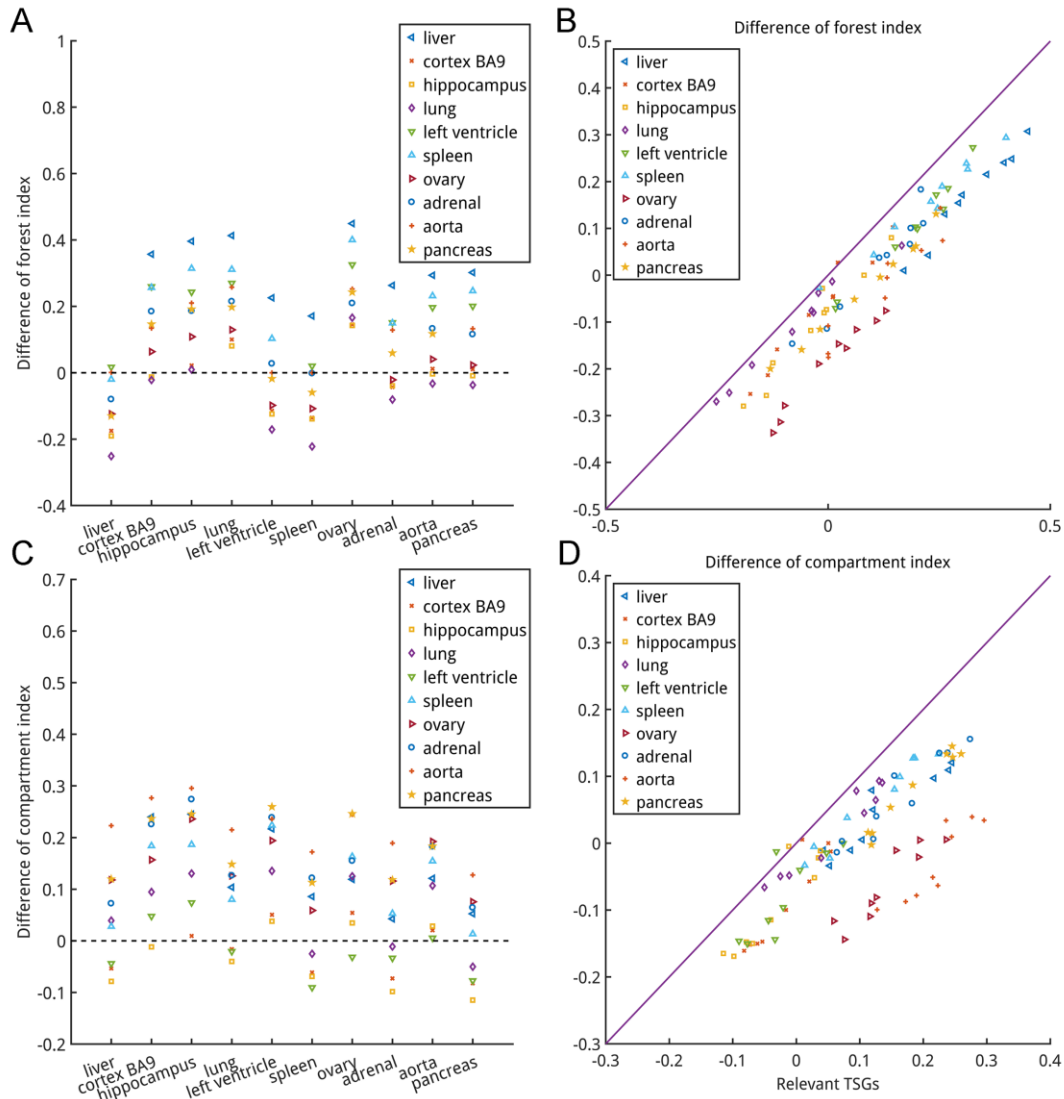


Fig. S11. A robustness test for the relation between forest gene regulation and genome organization. The criterion we used here in TSG identification is $s_i^t > 1$. Related to **Fig. S10**.

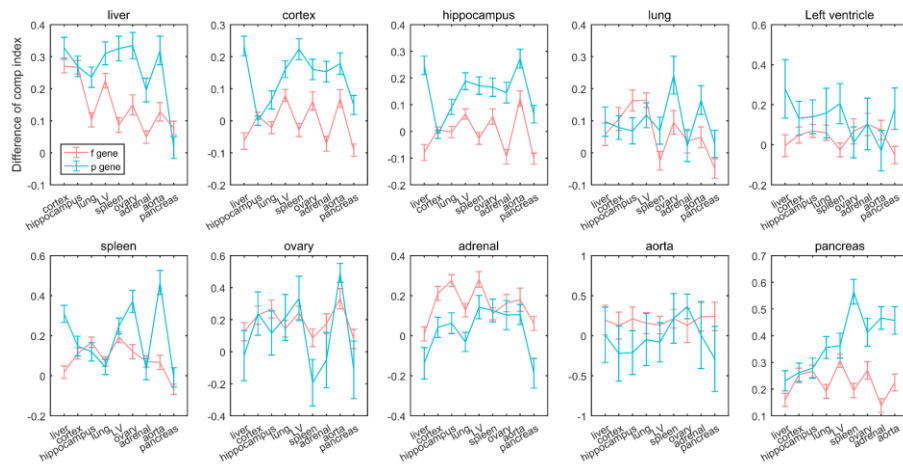


Fig. S12. The activation of prairie TSGs depends more on specific chromatin structure reorganization

(movement toward compartment A). We further compared the compartment index changes of forest and prairie TSGs relevant to one tissue (e.g., liver) when tissue type changes from control tissue (e.g., spleen) to relevant tissue (e.g., liver) and found such extent of prairie TSGs is larger than forests in most cases. The title of each boxplot represents the relevant tissue and the nine control tissues are labeled in x-axis. Each data point represents the median value of compartment index changes. Related to **Fig. 3D**.

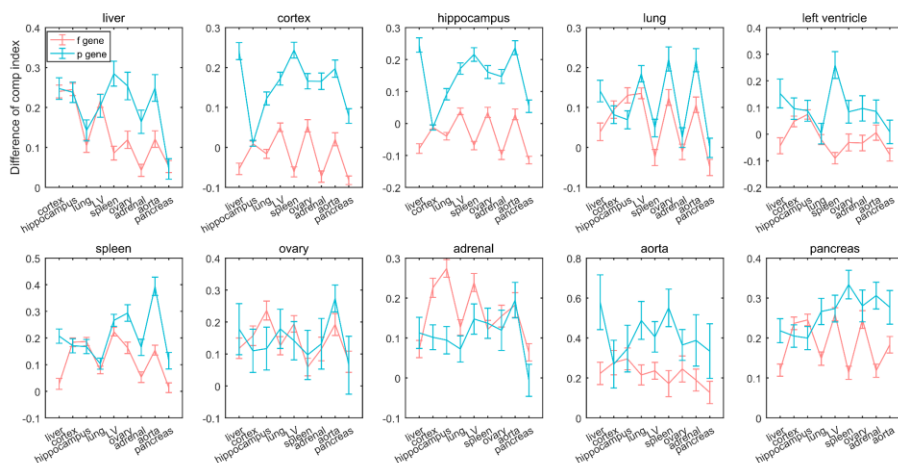


Fig. S13. A robustness test, related to **Fig. S12**. The criterion we used here in TSG identification is $s_t^f > 1$.

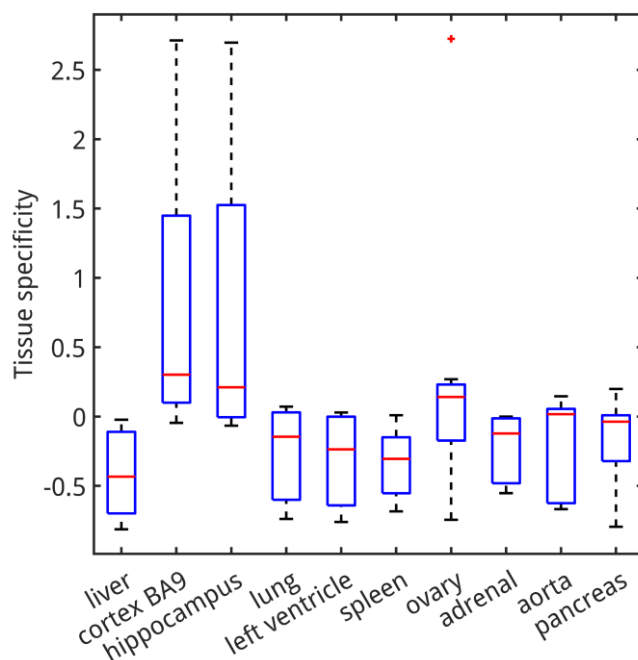


Fig. S14. The tissue specificities of eight representative forest genes showing the most negative correlation between gene expression level and forest index in ten tissues.

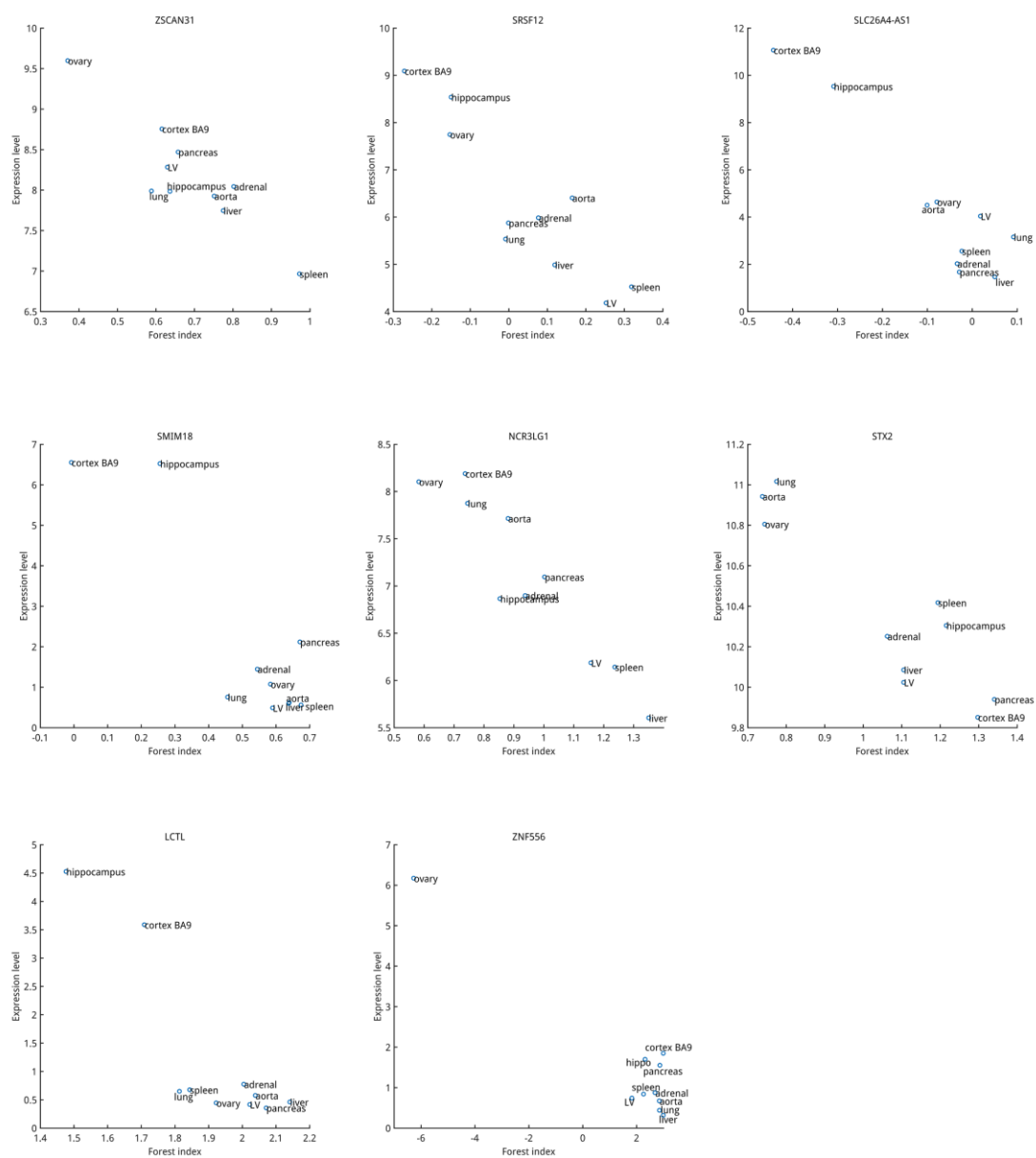


Fig. S15. Detailed information about the expression level and forest index of eight forest genes in ten tissues. Related to **Fig. S14**. LV=left ventricle.

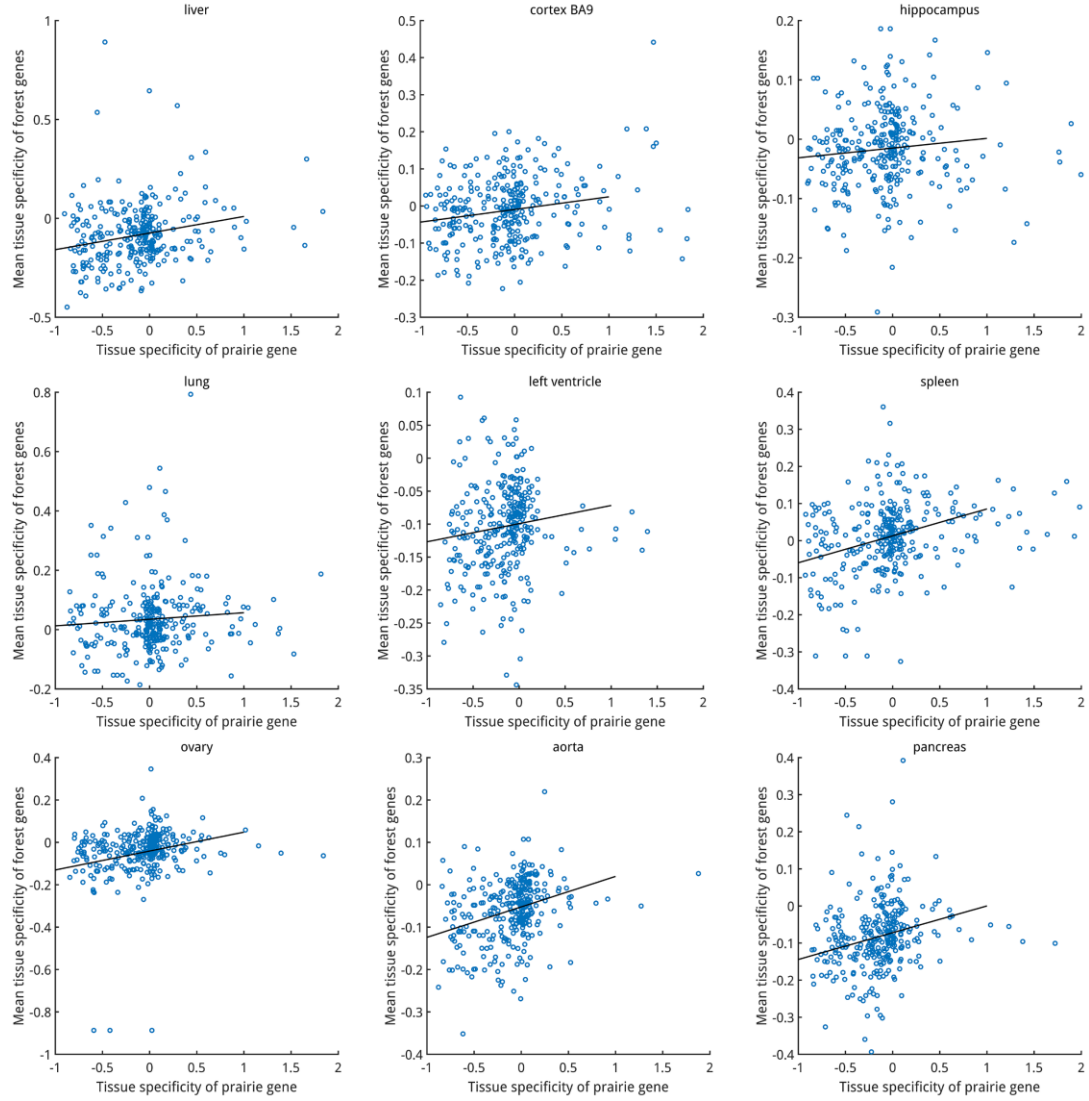


Fig. S16. The relation between tissue specificity of prairie gene and the mean tissue specificity of its highly contacted forest genes in nine tissues. All prairie genes of chr1 (the number is 317) were selected for calculation. When drawing these figures, the range of prairie gene tissue specificity (x-axis) was selected within $[-1, 2]$ and 95.9%~100% prairie genes were retained in these nine tissues. Related to **Fig. 3F**.

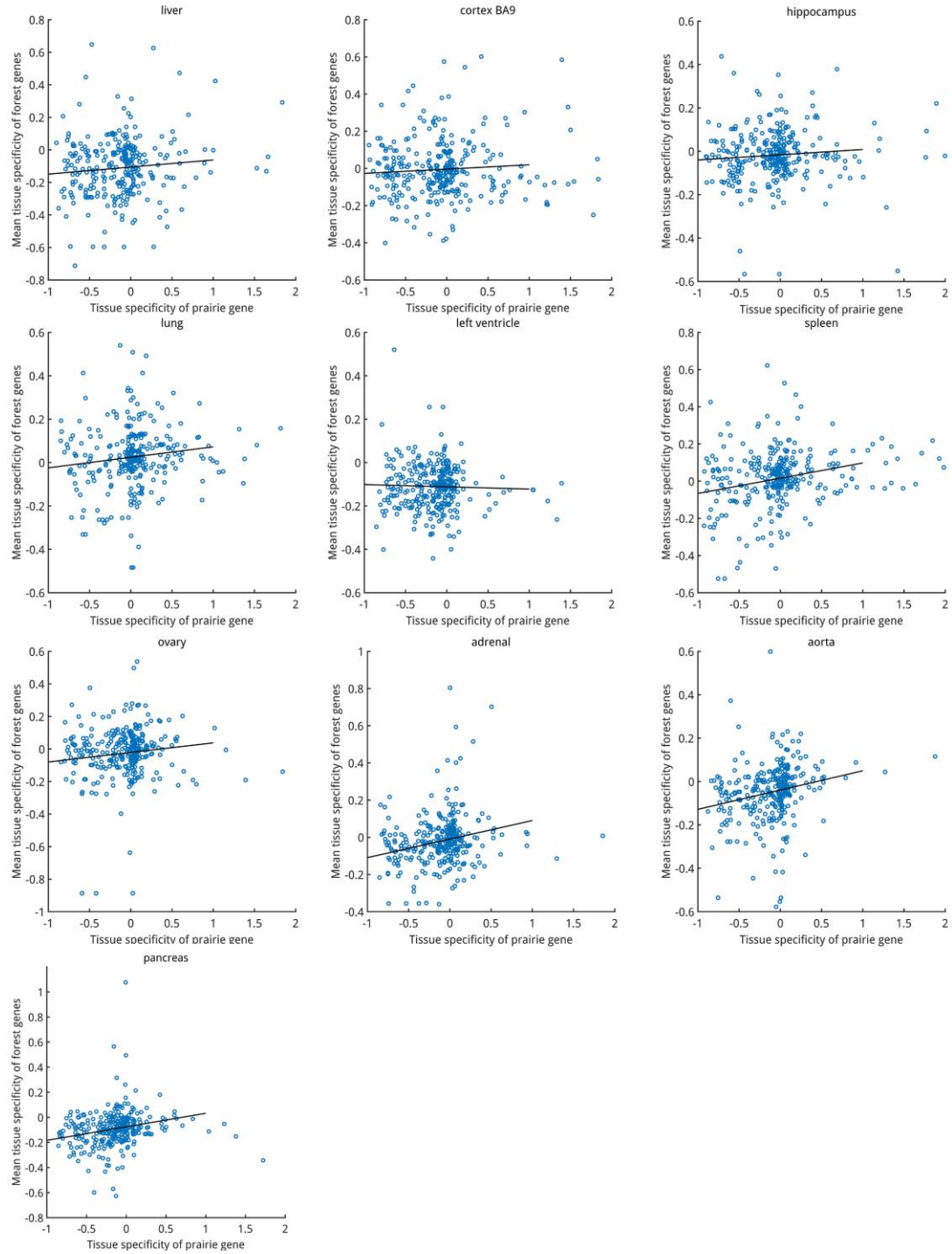


Fig. S17. A robustness test for the relation between tissue specificity of prairie gene and the mean tissue specificity of its highly contacted forest genes. Here the criterion we used in the identification of highly contacted gene pairs is 90th percentile (top10%, see methods). Related to **Fig. 3F** and **S16**.

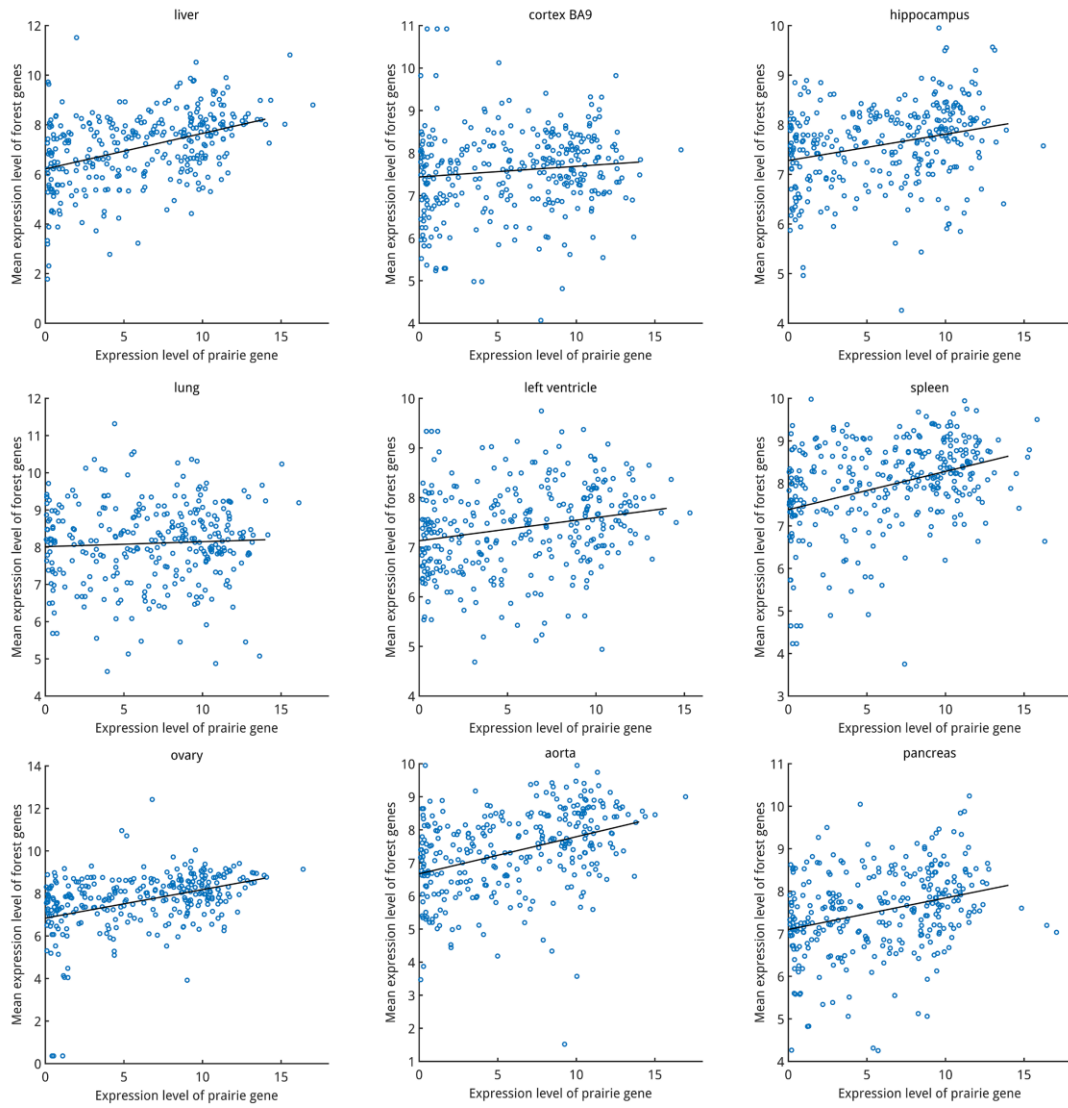


Fig. S18. The relation between expression level of prairie gene and the mean expression level of its highly contacted forest genes in nine tissues. Related to **Fig. 3F**.

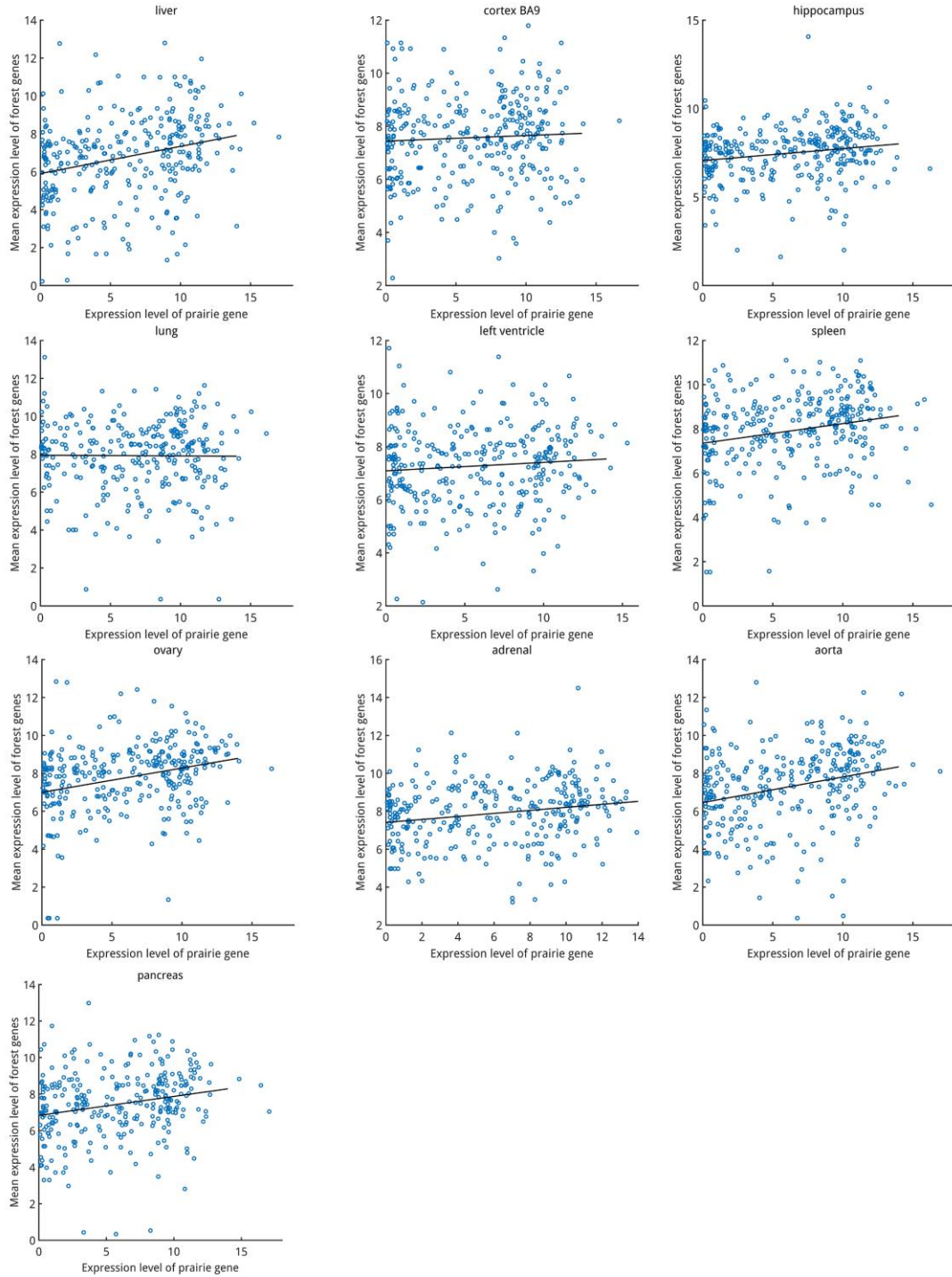


Fig. S19. A robustness test for the relation between the expression level of prairie gene and the mean expression level of its highly contacted forest genes. The criterion we used here in identifying highly contacted gene pairs is 90th percentile (top10%, see methods). Related to **Fig. 3F** and **S18**.

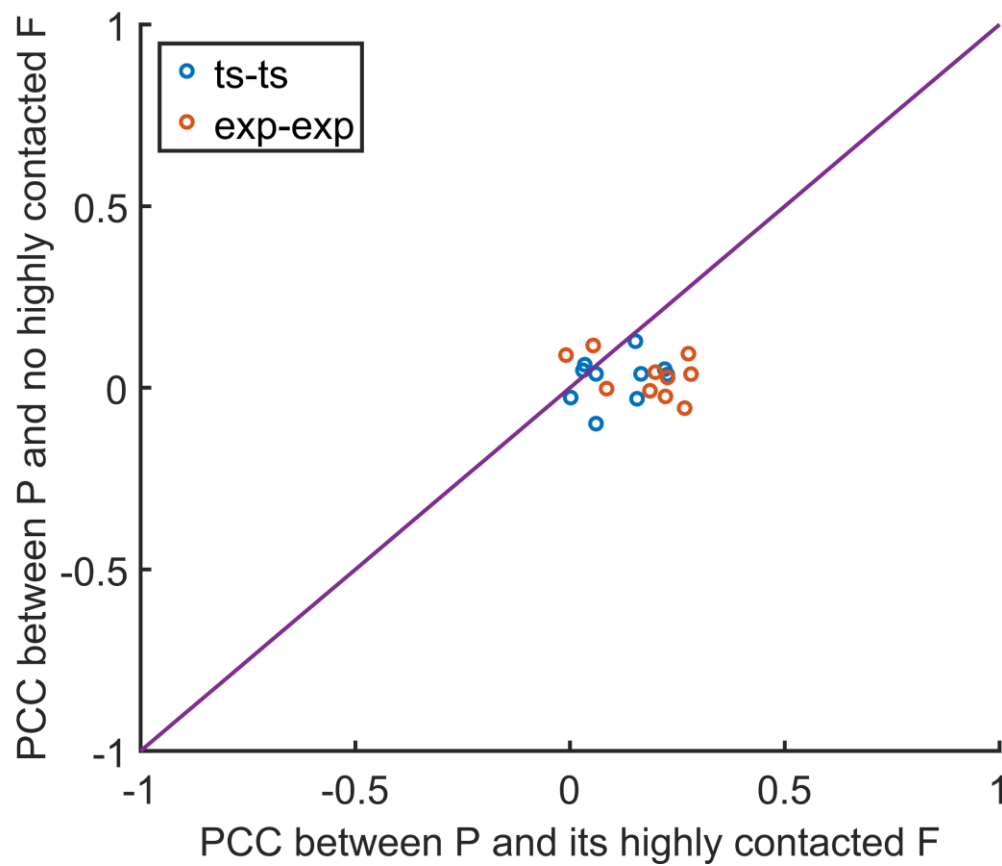


Fig. S20. A robustness test for the comparison between PCCs (Pearson correlation coefficients) calculated from Fig. S17 and S19 (x-axis) and the control analysis (correlation between prairie genes and randomly chosen forest genes that are not in strong contact with the former, y-axis). As mentioned in Fig. S17 and S19, the criterion in identifying gene pairs in strong contact is 90th percentile (top10%, see methods). For Fig. S19 (exp-exp), P -val < 0.001 in seven tissues except for cortex, lung and left ventricle, while P -val > 0.1 in nine tissues except for cortex (0.042) in the corresponding control analysis. For Fig. S17 (ts-ts), P -val < 0.01 in five tissues except for cortex, hippocampus, lung, left ventricle and ovary, while P -val > 0.25 in eight tissues except for liver (0.029) and lung (0.084) in the corresponding control analysis.

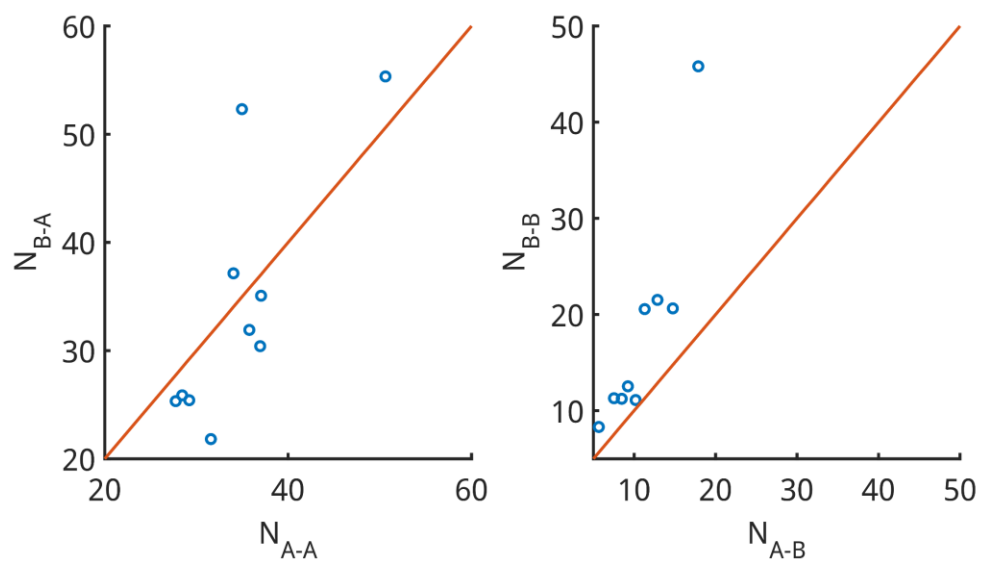


Fig. S21. A robustness test for the association between gene co-regulation and compartmentalization.

The criterion we used here in identifying highly correlated gene pairs is 99.9th percentile (see methods).

Related to **Fig. 4A**.

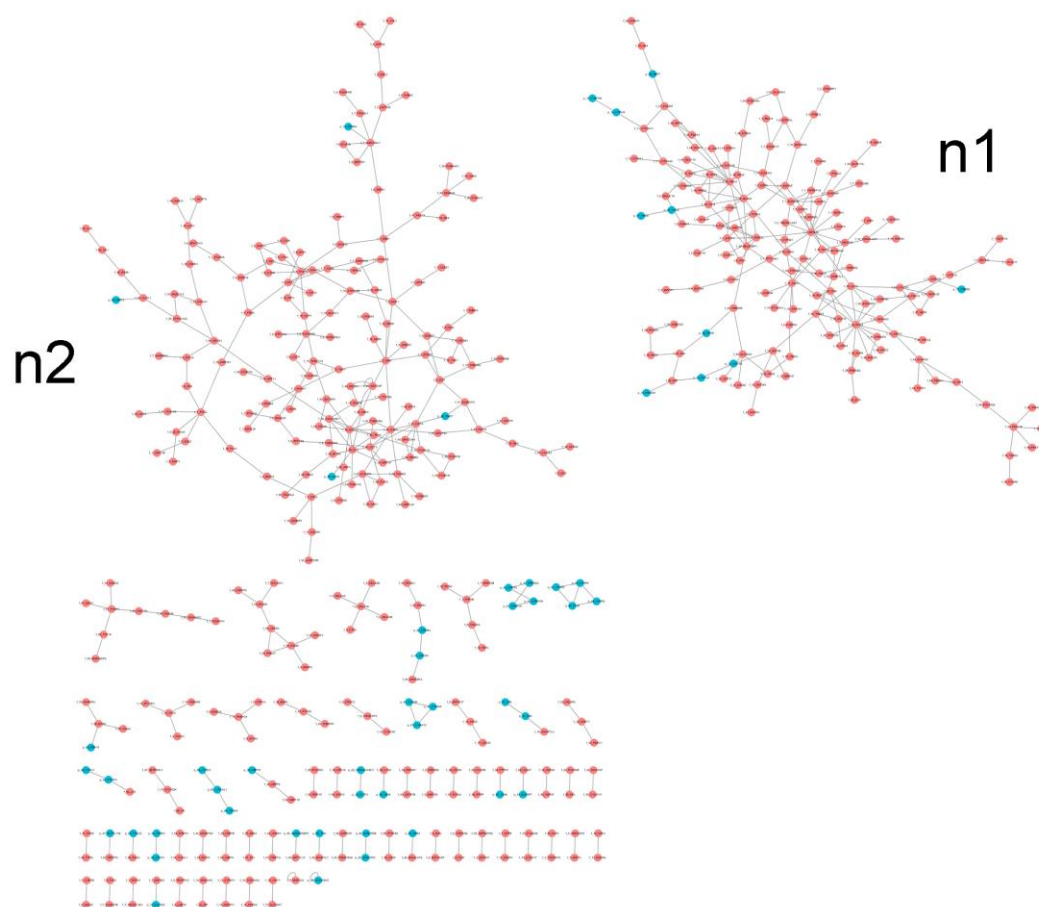


Fig. S22. Gene network in liver. Red circle, forest gene; blue circle, prairie gene. n1, liver-specific sub-network; n2, the more generic sub-network. Related to **Fig. 4B**.

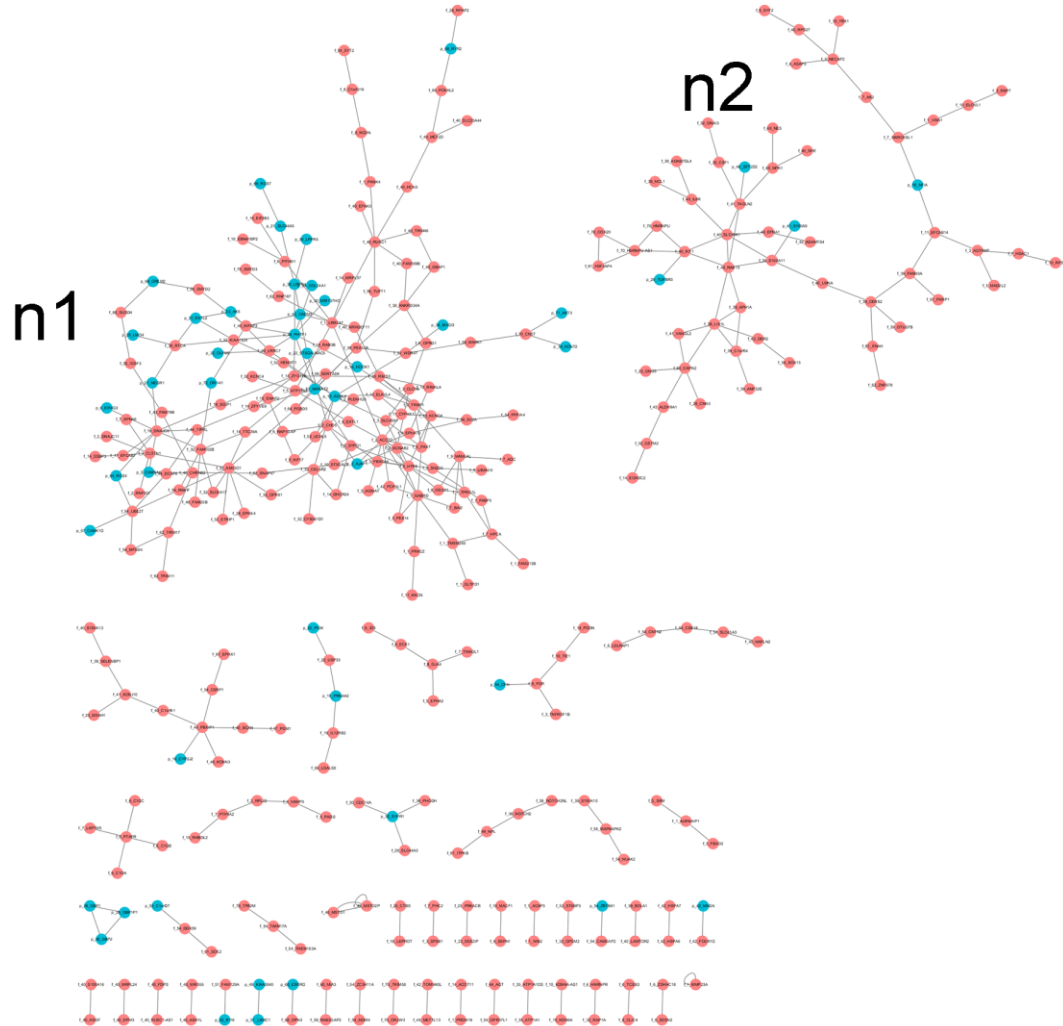


Fig.S23. Gene network in cortex BA9. Red circle, forest gene; blue circle, prairie gene. n1, cortex-specific sub-network; n2, more generic sub-network. Related to **Fig. 4B**.

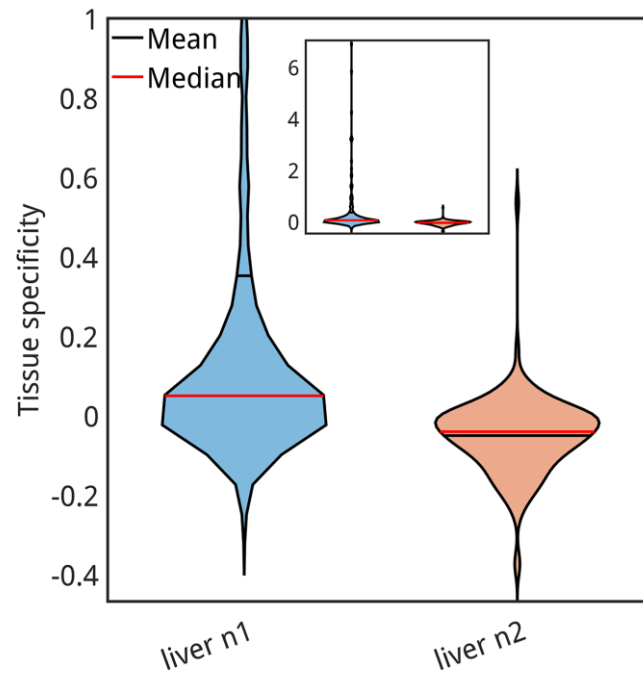


Fig. S24. The tissue specificity distribution of two liver sub-networks. Inner, the intact and unamplified figure. Related to **Fig. 4C**.

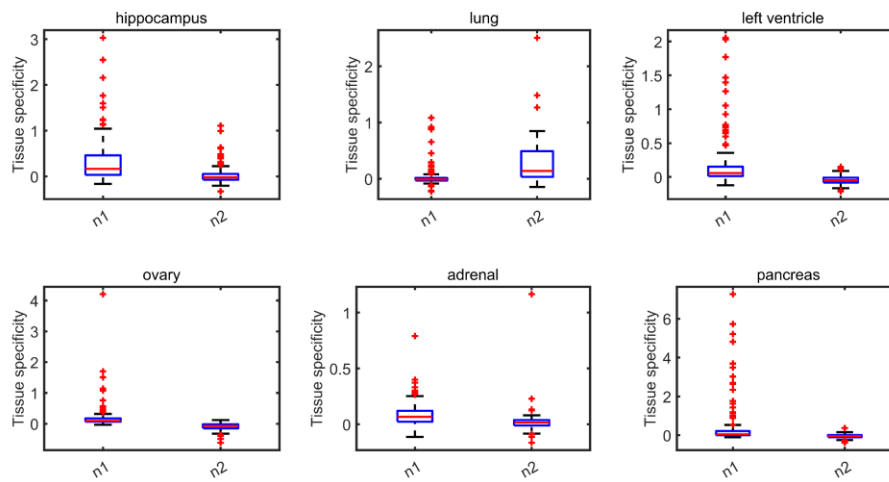


Fig. S25. Akin to liver and cortex, the expression+Hi-C networks of hippocampus, lung, left ventricle, ovary, adrenal and pancreas all showed a tendency of roughly containing two independent sub-networks, the tissue specificity distribution of which is distinctly different. $P\text{-val} < 1e-6$ in hippocampus, left ventricle, ovary and pancreas. $P\text{-val} = 0.017$ and 0.068 in lung and adrenal, respectively.

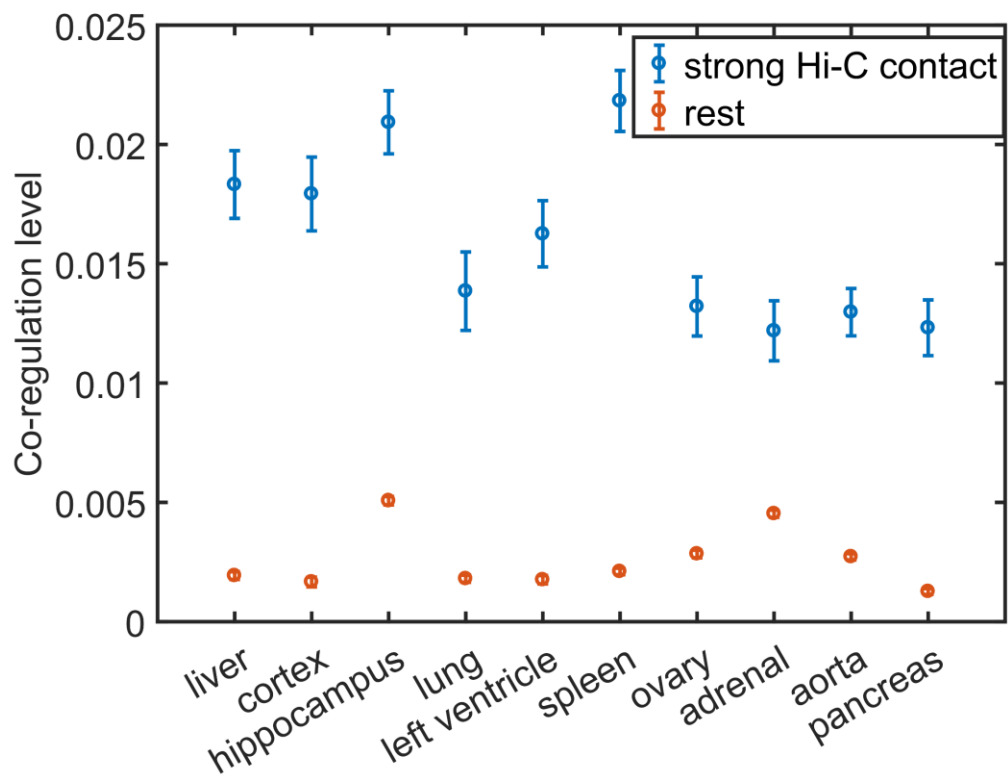


Fig. S26. A robustness test for how spatial contact affects gene co-regulation level. The criterion we used here in identifying highly contacted gene pairs is 90th percentile (top 10%, see methods) and average values were shown. Related to **Fig. 4E**.

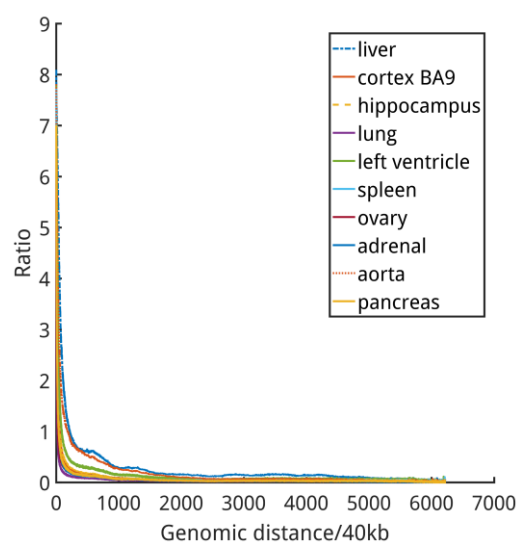


Fig. S27. The ratio calculated between the number of non-zero and zero Hi-C elements given one certain genomic distance. Hi-C data of ten tissues of chr1 was used.