Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A sub-organ-resolved pan-tissue single-cell atlas reveals cell-type-specific molecular regulation and xenotransplantation potential in pigs

Hongbin Pan

Pynnydx@163.com

yunnan agricultural university

Hong Hu

Yunnan Agricultural university

Xi Bai

Yunnan Agricultural university

Cao Deng

DNA Stories Bioinformatics Center https://orcid.org/0000-0001-9823-9121

Anjian Li

yunnan agricultural university

Yichao Mei

DNA Stories Bioinformatics Center

Taozong Guo

yunnan agricultural university

yuting yang

China Agricultural university

Shaokan Tan

DNA Stories Bioinformatics Center

Lei Sun

DNA Stories Bioinformatics Center

Fei Zhao

DNA Stories Bioinformatics Center

Xiangfang Zeng

China agricultural university

Shiyan Qiao

China agricultural university

Article

Keywords:

Posted Date: November 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11755645/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

A sub-organ-resolved pan-tissue single-cell atlas reveals cell-type-specific molecular regulation and xenotransplantation potential in pigs

Hong Hu[1,2,6], Xi Bai[1,3,6], Cao Deng[4,5,6], Anjian Li[1,6], Yichao Mei[5], Taozong Guo[1], Yuting Yang[1,2], Shaokan Tang[5], Lei Sun[5], Fei Zhao[5], Xiangfang Zeng[2,*], Shiyan Qiao[2,*], Hongbin Pan[1,*],

[1]Yunnan Provincial Key Laboratory of Animal Nutrition and Feed Science, Faculty of Animal Science and Technology, Yunnan Agricultural University, Kunming, China

[2]State Key Laboratory of Animal Nutrition and Feeding, Ministry of Agriculture and Rural Affairs Feed Industry Center, China Agricultural University, Beijing, China

[3] College of Animal Science, Anhui Science and Technology University, Chuzhou, China

[4]Key Laboratory of Animal Disease and Human Health of Sichuan Province, College of Veterinary Medicine, Sichuan Agricultural University, Chengdu, China

[5]Departments of Bioinformatics, DNA Stories Bioinformatics Center, Chengdu, China

[6]These authors contributed equally: Hong Hu, Xi Bai, Cao Deng, Anjian Li

[*]Correspondence: ziyangzxf@163.com, qiaoshiyan@cau.edu.cn, Pynnydx@163.com

Abstract

Pigs are core livestock species and critical large animal models for biomedicine, but existing porcine single-cell atlases are limited to single breeds and whole-organ resolution. Here, we built a multi-breed pan-porcine single-cell atlas with 9,484,620 cells from 1,001 samples, and performed the first single-nucleus transcriptomic profiling of 160 sub-organ sites across 31 tissues in Tibetan pigs (TBP), annotating 228 cell subtypes within 9 major lineages. We demonstrated that sub-organ heterogeneity is a core source of organ functional diversity across multiple physiological systems. Integrative analyses revealed that high-altitude adaptation in TBP arises from multi-scale transcriptional remodeling across organs, sub-organs and cell types; transcriptional coupling between muscle progenitors and myofibers underpins its superior meat quality; and gastrointestinal roughage tolerance depends on synergistic host-microbe interactions rather than microbiota alone. Cross-species comparison showed that TBP kidney tissues exhibit higher transcriptional conservation with humans than other pig breeds, supporting its potential for renal xenotransplantation, and we identified key immune genes as candidate editing targets. This atlas provides a valuable resource for livestock breeding, evolutionary research and translational medicine.

Introduction

Pigs (*Sus scrofa*) are a key species in the history of human livestock breeding and are the dominant source of meat in numerous countries and regions worldwide[1-4]. Pork contains key nutritional components such as high-quality proteins, unsaturated fatty acids, essential minerals, and various vitamins[5]. It is an indispensable provider of nutrition in the human diet and effectively meets nutritional needs for human growth, development, and metabolism[5]. Aside from their role in human nutrition, pigs have also been increasingly used in biomedical applications[6,7]. Owing to their numerous similarities with humans in terms of anatomical size and structure, physiological characteristics, immune responses, and genomic sequences and their relatively low breeding costs, pigs have become prominent models in translational medicine research[7]. Pigs are also crucial donor sources for human organ transplantation[8]. Recently, with continuous progress in gene-editing technologies, researchers have successfully bred various genetically modified pigs for use in organ transplantation[9,10]. As a high-quality livestock species for food production, model organism for translational medical research, and potential donor for human organ transplantation, pigs have irreplaceable academic value and application prospects in the agricultural economy and biomedical development.

Tibetan pigs (TBP) are a valuable local minipig breed on the Qinghai-Tibetan plateau; they are renowned as “jewels of the plateau” because of their superior meat quality and exceptional flavour, making them important contributors to the pig industry[11]. Through long-term natural selection and adaptive evolution, TBPs have developed unique biological characteristics supporting their survival in extremely harsh ecological environments, among which their adaptability to hypoxic and low-temperature plateau climates, tolerance of roughage, and resistance to disease are particularly

prominent[12, 13]. Considering the high biological similarity of pigs to humans, TBPs may therefore be used as valuable medical models for investigating plateau-related diseases such as high-altitude sickness and pulmonary arterial hypertension, as well as for screening drugs or exploring gene therapy applications targeting plateau-related diseases[14]. Furthermore, the organ size and physiological functions of miniature pigs are much more similar to those of humans than those of large pigs are, making the application of gene-editing technology for organ transplantation purposes more convenient and efficient[15-18].

In fields such as the optimization of livestock genetic breeding and the elucidation of regulatory mechanisms, single-cell sequencing technology, which has the ability to perform high-resolution analyses of information collected at the single-cell level, has become an essential technical support and research tool [19,20]. However, the single-cell transcriptomic studies conducted in livestock (pig and cattle) to date have focused primarily on macro-level analyses of entire organs, whereas the systematic exploration of suborgans as refined functional partitions, which will be needed to improve compatibility with humans for transplantation, remains relatively limited[21-23]. Especially, current single-cell atlas resources for pigs are largely restricted to individual breeds. Comprehensive cross-breed panoramic single-cell atlases are still lacking, calling for in-depth research. The construction of a pig single-cell atlas with enhanced anatomical resolution to delineate cellular diversity and transcriptional regulatory programs at the sub-organ level is therefore of great scientific and practical importance.

Here, this study integrates single-cell transcriptomic datasets of TBP, Large White pigs (LWP) and other local pig breeds. The scRNA-seq was performed on a total of 160 anatomical sub-sites from 31 tissues and organs of TBPs, and a pan-single-cell transcriptomic atlas of domestic pigs

across multiple pig breeds with sub-organ resolution was constructed. This study systematically characterizes the cellular composition and transcriptional heterogeneity of each sub-organ; combined with gene co-expression network, single-cell metabolic activity analysis and gene regulatory network, it dissects the molecular regulatory differences among different pig breeds and explores the cellular and molecular basis underlying the high-altitude adaptation and superior meat quality traits of TBPs; gastrointestinal cell-microbe association analysis was conducted to elucidate the potential mechanism by which TBP intestines adapt to high-fiber diets; further cross-species single-cell transcriptomic comparison between humans and pigs was carried out to evaluate the application potential of different pig breeds as human disease models and donors for xenotransplantation. The high-resolution multi-breed single-cell reference atlas constructed in this study can provide important data support for research on domestic pig genetic resource mining, molecular breeding, evolutionary medicine of plateau animals and comparative medicine.

Results

Construction of a pig pan-single-cell atlas across multiple organs and breeds

This study integrated various single-cell transcriptome data from pigs. The dataset consists of 9484620 cells and 1001 samples, with TBP contributing 2338851 cells and 296 samples, LWP contributing 4509033 cells and 434 samples, and other breeds contributing 2636736 cells and 271 samples **(Fig. 1)**. Among which, to construct a comprehensive single-cell transcriptomic atlas of TBPs at the suborgan level, snRNA-seq was performed on 31 tissues and organs of these pigs (comprising 160 suborgans) **(Fig. 2)**. After low-quality cells and doublets were filtered out, 2,338,851 cells remained for subsequent analyses. The average number of identified cells per suborgan among the 160 suborgans was 14,617. All high-quality cells from the tissues were

integrated using Scanpy (v1.10.3), and 228 subtle cell types were annotated via CellTypist (1.7.1)and CellID (1.10.1) for dataset; these cell types represented nine distinct cell lineages: Endothelial (959,229 cells), Epithelial (1,741,262 cells), Germline (140,656 cells), Lymphoid (1,610,995 cells), Mesenchymal (2,268,967), Myeloid (544,618 cells), Neural (2,077,314 cells), Stem (52,316 cells), and Unknown (89,263 cells) **(Extended Fig. 1-3, Supplementary Fig. S1, Supplementary Table S1-6)**. A Circos plot was generated to systematically visualize the global cell-type composition, transcriptional heterogeneity, and marker-gene expression signatures of the integrated pig pan-single-cell atlas **(Fig. 3)**. The central panel of the Circos plot displayed t-SNE embedding of nine major cell superclusters, comprising Epithelial, Mesenchymal, Lymphoid, Myeloid, Neural, Endothelial, Germline, Stem, and Unknown cell populations. Rings (f)–(n) visualized expression levels of the top 3 marker genes derived from each supercluster across all 228 cell types. As expected, these representative marker genes were highly enriched in their corresponding supercluster while exhibiting low expression in other superclusters, which substantiated the reliability of our cell-type annotation. This finding indicates the accuracy of the marker identification results and verifies the reliability of the cell type annotation **(Extended Fig. 4-5, Supplementary Table S7-8)**. Furthermore, Suborgans in systems such as the skin and nervous systems exhibited a clustered distribution, indicating strong intrinsic relevance among different suborgans within the same system. In contrast, the suborgans in the digestive and circulatory systems showed a dispersed distribution, suggesting more pronounced heterogeneity among the suborgans in these systems **(Extended Fig.6)**.

High regional heterogeneity of organs and tissues in TBPs

The suborgans within the same organ exhibited significant differences in transcriptional

heterogeneity **(Fig. 4a)**. In the heart, sub-organs of the heart muscle exhibited obvious differences in transcriptional heterogeneity. The wall of left atrium showed a relatively high overall level, while parts of the ventricular wall, cardiac septum and apex had lower values. For cardiac valves, the mitral valve displayed higher transcriptional heterogeneity than the tricuspid valve and aortic valve. Lymph node sub-regions had similar overall transcriptional heterogeneity, with significant differences detected in a few sub-region comparisons, suggesting a certain degree of divergence in transcriptional states across different lymph nodes. In the intestinal tract, the three sub-organs of the duodenum had similar levels of transcriptional heterogeneity. The jejunum similarly showed minor overall differences among its sub-regions. In contrast, sub-organs of the centrifugal colon showed the most marked transcriptional heterogeneity differences. The middle region displayed the highest and the posterior region the lowest level, and significant differences existed in all pairwise comparisons. The overall transcriptional heterogeneity was similar across the three sub-organs of the centripetal colon, with significant differences detected only in a few pairwise comparisons. In addition, the transcriptional heterogeneity of the anterior region was markedly higher than that of the posterior region in the rectum.

Brain suborgans were composed of heterogeneous cell-type clusters with region-specific differences in cell-type composition profiles across different suborgans **(Fig. 4b, Extended Fig.7)**. Cell annotation showed that all cells were clearly partitioned into 15 major cell subclusters with evident clustering patterns in two-dimensional space. Large clusters were formed by Cerebral Cortex Pyramidal Neuron, Interneuron and Glial cells, confirming high data quality and cellular type diversity. Furthermore, the cellular composition of the 28 sub organs in the TBP brain exhibits high heterogeneity, which can be classified into highly homogeneous, multi type balanced, and

intermediate types based on cell richness. TBP brain sub-organs were classified into three groups by cellular composition. Five highly homogeneous suborgans had low diversity: right cerebellar hemisphere cortex (interneurons, 0.91), cerebellar vermis (interneurons, 0.89; astrocytes, 0.07), olfactory bulb (interneurons, 0.82; astrocytes, 0.07; Neuroblast_Sensu_Vertebrata, 0.06), left cerebellar hemisphere cortex (interneurons, 0.74; neurons, 0.16), and thalamus (Neuroblast_Sensu_Vertebrata, 0.67; cerebral cortical pyramidal neurons, 0.18). The balanced group included cerebellar white matter (six cell types, no type >20%, highest diversity), mammillary body (dominant cell fraction, 0.31), and medulla oblongata (0.23). Other sub-organs showed intermediate features. Dominant cell types displayed marked regional specificity. The cerebellar system was enriched in interneurons, distinct from Others and ciliated cells in cerebellar white matter. Fifteen cortical suborgans were dominated by pyramidal neurons (0.31–0.53) with variable secondary populations. “Neuroblast_Sensu_Vertebrata” was the second major cell type in the frontal lobe (0.39) and olfactory tract (0.41), while astrocytes and radial glia were abundant in the hippocampal formation, temporal lobe and piriform lobe. The corpus callosum and oculomotor nerve were dominated by neural crest cells (0.30, 0.29); the oculomotor nerve contained uniquely high mesenchymal cells (0.14). The cerebral peduncle was enriched in radial glia (0.33), astrocytes (0.25) and ciliated cells (0.19). The medulla oblongata and its pyramid differed substantially: medulla oblongata had six evenly distributed cell types (radial glia and astrocytes, each 0.23), whereas the pyramid was dominated by Others (0.27). In summary, TBP brain suborgans exhibited regional cellular specialization: pyramidal neurons in cortical areas, interneurons in the cerebellum, and neural crest cells, radial glia and astrocytes in brainstem/nerve tracts, reflecting functional partitioning and microenvironmental heterogeneity.

The spatial heterogeneity and functional regionalization characteristics of cell types in TBP liver and its accessory sub organs were observed **(Supplementary Fig.S2)**. (1) Most hepatic regions showed high cellular homogeneity: the right triangular ligament was almost exclusively cholangiocytes (1.0), the proper hepatic artery was dominated by stellate cells (0.8), and bile ducts contained cholangiocytes (0.69) plus minor central venous LSECs. Several lobes and ligaments exhibited higher diversity; the left triangular ligament consisted of immune_mix (0.44), periportal hepatocytes (0.29) and others (0.17). (2) Cell types displayed distinct spatial localization: cholangiocytes enriched in the biliary tract and right triangular ligament, and stellate cells occupied vascular structures. Periportal hepatocytes and immune_mix distributed broadly across hepatic lobes and formed lobule units with central venous LSECs. Tem_Trm_cytotoxic_T cells were sparse in the proper hepatic artery and coronary ligament, while macrophages specifically accumulated in the quadrate lobe. The analysis of the cell-type composition characteristics of cardiac suborgans revealed significant heterogeneity between cardiovascular suborgans and myocardial tissue **(Supplementary Fig. S3)**. Additionally, the ventricular region is characterized by high-purity myocardial cells, while the atrial and valve regions are dominated by complex endothelial cell populations (exhibiting atrial or ventricular enrichment characteristics) and fibroblasts/stromal cells. In contrast, the observed differences in the cell-type composition among lymph node suborgans were relatively conserved, and it presents a highly conservative cell composition with B cells as the core **(Fig. 4c)**. All tested lymph nodes (mesenteric, parasternal, popliteal, prefemoral, submandibular) showed strong functional specialization as immune hubs and were predominantly occupied by BTAF1_High_B cells (63%–93%). Only minor cell populations were detected, including lymphatic endothelial cells (0.12–0.28), endothelial cells (0.15–0.31), and rare

BTAF1_High_T cells.

The intestinal sub-organs of TBPs exhibited prominent regional heterogeneity and functional gradients in cellular composition **(Extended Fig. 8)**, mirroring structural and microenvironmental transitions along the digestive tract from proximal nutrient absorption to distal excretion. Along the duodenal anterior–posterior axis, the proportion of Stromal cells decreased from 0.57 to 0.40, accompanied by elevated fractions of other cell populations, indicating cellular remodeling. The cecum showed the highest cellular diversity, with Stromal (0.27) and GdT cells (0.17) as major components. The middle centrifugal colon was dominated by myofibroblasts (0.69). In contrast, the middle jejunum exhibited extreme cellular homogeneity, where Stromal cells accounted for 0.96. The ileum was mainly composed of myofibroblasts (0.68) and Stromal cells (0.28). Distinct cellular variation was observed across colonic suborgans: the anterior centrifugal colon was dominated by Stromal cells (0.58), while the anterior centripetal colon had a high Stromal proportion (0.69) and its distal segments displayed greater cellular complexity. For the rectum, the anterior segment contained diverse cell populations (Stromal 0.41, Others 0.19, Myofibroblast 0.18), whereas the posterior rectum showed a more concentrated composition dominated by Stromal (0.33), Myofibroblast (0.33) and Others (0.20). Collectively, marked cellular shifts along the anterior–middle–posterior axis within individual intestinal organs demonstrate striking spatial heterogeneity of the intestinal cellular ecosystem in TBPs.

The muscular suborgans of TBPs exhibit considerable differences in cell-type richness and spatial heterogeneity. Specifically, certain skeletal muscles exhibited rich and balanced cellular compositions. For instance, Brachiocephalic Muscle contained Slow_Skeletal_Muscle (0.44), Fast_Skeletal_Muscle (0.17), and Mesenchymal_Stem (0.13); Deltoid Muscle (Acromial Part) was

dominated by PAX7_Low_MuSCs (0.43), Slow_Skeletal_Muscle (0.26), Others (0.20) and Lymphatic_Endothelial (0.11). In contrast, certain muscles exhibit a clear predominance of a single cell type: Fast_Skeletal_Muscle occupied 0.78 in the superficial pectoral, while Slow_Skeletal_Muscle accounted for 0.68 in Latissimus Dorsi Muscle. The relative abundance of Fast- and Slow_Skeletal_Muscle varied greatly across muscle suborgans, indicating distinct regional differences in myofiber-associated cell populations. Non-myofiber cell components also differed remarkably between muscles. Semimembranosus Muscle was mainly composed of Capillary_Endothelial (0.52) and Fibroblast (0.31). Rectus Femoris Muscle showed a balanced composition of multiple cell types. Masseter Muscle and Muscles of Tongue also harbored diverse cell populations with Capillary_Endothelial at 0.34 and 0.32, respectively, without a highly dominant cell population as observed in the superficial pectoral muscle **(Supplementary Fig. S4)**.

Supplementary Fig. S5 shows the similarities of transcriptional signatures between the suborgans within different organs, indicating the homogeneity and heterogeneity of the functional states of each suborgan. The different suborgans could be divided into several distinct similarity modules according to their anatomical sequence. In the gut, the transcriptional patterns within the Centripetal Colon (posterior) and duodenum (anterior, middle), as well as Rectum (Anterior) were highly similar. In the heart, Global transcriptional similarity was more scattered across cardiac suborgans relative to several other organs. While the mitral valve and aortic valve showed prominent valve-related transcriptional correlation, highly consistent transcriptional profiles were absent among atrial walls, atrioventricular walls, cardiac apex and valves. In the liver, the following suborgans exhibited the greatest similarities: Coronary Ligament of Liver, Hepatic Duct and Hepatic Artery Proper; Left Triangular Ligament of Liver and Right Triangular Ligament of Liver; Caudate

Lobe of Liver and Hepatoduodenal Ligament; Right Lateral and Left Lobe of Liver. In the lymphatic system, the parasternal and prefemoral lymph nodes showed the greatest similarity, whereas the transcriptional signatures among the other three suborgans differed significantly, suggesting that these suborgans may be involved in distinct immune functions **(Fig. 4d)**. In the muscle system, distinct similarity modules were again observed among the suborgans: (1) Subscapularis muscle、Latissimus dorsi muscle and rectus femoris muscle; (2) Pectoralis Major Muscle (Superficial Pectoral)、External Abdominal Oblique Muscle 和 and Semimembranosus Muscle; (3) Muscles of Tongue、Brachiocephalic Muscle、Longissimus Dorsi Muscle、Psoas Muscle and Masseter Muscle; (4) Muscles of Tongue and Brachiocephalic Muscle, Brachiocephalic Muscle and Longissimus Dorsi Muscle, Longissimus Dorsi Muscle and Psoas Muscle; and (5) Deltoid Muscle (Acromial Part)、Cervical Part of Trapezius Muscle and Deltoid Muscle (Scapular Part).

To clarify whether differences existed in different parts of the same tissue type, we performed differential expression analysis among shared cell types for the suborgans of the lymph, intestines, and muscle. The results revealed that suborgan-to-suborgan variation was more pronounced for Endothelial and BTAF1_High_B cell populations, whereas Lymphatic_Endothelial cells exhibited relatively stable profiles across suborgans **(Fig. 4e)**. Furthermore, the transcriptional states of shared cell types across the intestine and muscle showed marked suborgan specificity: (1) Myofibroblasts exhibited the greatest inter-suborgan variation; Centripetal_Colon (Posterior) contained abundant differentially expressed genes compared with multiple suborgans. By contrast, Stromal cells had modest overall variation, though Centrifugal_Colon (Anterior) and Cecum_Band showed relatively more differentially expressed genes versus other suborgans; (2) Capillary_Endothelial cells displayed the most prominent inter-suborgan differences, while the transcriptional profiles of

Fast_Skeletal_Muscle and Slow_Skeletal_Muscle were relatively conserved **(Supplementary Fig. S6-7)**.

Key organs and regulatory networks underlying the high-altitude adaptation of TBPs

To systematically elucidate the molecular regulatory mechanisms underlying the adaptability of TBPs to high-altitude environments, we integrated and analysed multidimensional data, including results from gene coexpression networks, single-cell metabolic activity analyses, and gene regulatory networks **(Fig. 5a-5d)**. Through gene coexpression network analysis, TBP and LWP may exhibit distinct gene co-expression regulatory patterns in the same organs. Specifically, In TBP, Muscle exhibited the strongest positive correlation with M3 and a moderate positive correlation with M2; Brain was positively associated with M4/M5, while organs including Lymph_nodes, Lung, Liver, Kidney, Intestine and Heart were mainly linked to M1/M2 with several negative correlations. The organs showing prominent interbreed differences included Muscle, Lymph_nodes, Lung, Liver, Kidney, Intestine, Heart and Brain: these organs generally displayed positive M2 and negative M1 correlations in TBP, whereas LWP showed the opposite trend. TBP Brain had strong positive correlations with M4/M5/M7, while LWP Brain was mainly associated with M6/M8. Although organs such as Uterus, Thyroid, Testis, Stomach, Spleen, Pituitary, Pancreas and Ovary showed relatively weak overall correlations, the magnitude and direction of module associations still varied between breeds. The breed differences were not limited to changes in correlation strength but also involved module switching and reversal of positive/negative correlation directions, indicating divergent regulatory patterns **(Fig. 5a, Extended Fig. 9-11, Supplementary Table S9-10)**.

Further analyses revealed that specific modules were associated with core physiological system functions，including cold resistance (5 modules) and low oxygen adaptation (8 modules), but no

modules related to DNA damage repair **(Fig. 5b, Extended Fig. 12, Supplementary Table S11)** . For cold- and hypoxia-adaptation-related modules: (1) No significant differential modules were detected in the brain; (2) In the heart, intestine, kidney, lymph nodes and muscle, TBP exhibited significant upregulation of M3 and M5 and downregulation of M1 and M8 in cold-adaptation modules, while M2, M3, M4, M5 were significantly upregulated and M1, M7, M8 significantly downregulated in hypoxia-adaptation modules, indicating the diversity and functional specialization of biological functions in different organs **(Supplementary Fig. S8, S9)**. The evaluation of single-cell metabolic activity further confirmed the functional specialization of the aforementioned organs. Distinct cell types showed marked differences in gene expression, while their metabolic networks exhibited both broadly shared properties and cell-specific signatures **(Extended Fig. 13)**. In contrast, Epithelial cells displayed enriched glucose metabolism, amino acid metabolism and glycan biosynthesis linked to absorption, secretion and barrier functions. Endothelial cells showed prominent energy and lipid metabolism, reflecting high requirements for energy supply and membrane maintenance, whereas mesenchymal cells were dominated by lipid and cellular structure-associated metabolism **(Fig. 5c)**.

The results of the gene regulatory network inference experiments provide key evidence regarding the molecular mechanisms underlying the high-altitude adaptability of TBPs an pigs at the transcriptional regulatory level. Shared organs of TBP and LWP displayed marked organ-specific regulon patterns, and these organ differences were clearly driven by key regulons. Regulon modules linked to neural development, neuronal differentiation and nervous system homeostasis (BHLHE41, DBX2, SOX10, RFX4, DLX2, TCF4) were specifically enriched in neural tissues, whereas myogenic regulons such as MYF6 were observed in muscle tissues. In addition,

regulon RSSZ profiles varied substantially across organs and breeds. Regulons including MAF, IRF8, BHLHE41, SOX9, NEUROD4, POU3F4, DLX2, SOX17, MYCN, MYF5, MYBL1, SPIB, EOMES, SOX10, SREBF2, NR5A2, PAX5, IKZF3 and RFX3 showed differential RSSZ enrichment across organ-by-breed groups, with each regulon exhibiting distinct organ and breed preferences **(Supplementary Fig. S10-11)**. In gene regulatory networks built for shared organs between TBP and LWP, regulon nodes had varied connectivity and importance measures; nodes CEBPB, NFIL3, TP63, SOX6, RFX4, BCLAF1, E2F8, ARNT2, ETS1, IRF7, IRF8 and PRRX1 exhibited prominent connectivity or high importance values **(Fig. 5d).** Furthermore, regulon specificity showed clear regional divergence across TBP intestinal suborgans. The duodenum, colonic regions, cecum band, and jejunum/ileum had distinct RSSZ enrichment patterns. Centrifugal colon and centripetal colon, as well as jejunum and ileum, shared partial similarity in regulon profiles, in agreement with the grouping from hierarchical clustering **(Extended Fig. 14)**.

Integrational analyses across multiple muscle sites elucidate the metabolic and transcriptional mechanisms governing excellent meat quality of TBPs

Different anatomical muscle sites exhibit specialized metabolic properties and synergistically form efficient and coordinated functional units for energy metabolism and nutrient regulation **(Fig. 6a,6b, Supplementary Table S12-13)**. Visualization analyses revealed that SM_TRC, SM_SUB, SM_LD and SM_PM had higher amino acid and fatty acid abundances than other samples, which are hallmarks of high-quality muscle. SM_GM exhibited enriched essential amino acids. According to heatmap-based sample clustering, muscle samples were categorized into high-, medium-, and low-quality groups: high quality (SM_TRC, SM_SUB, SM_LD, SM_PM), medium quality (SM_GM), and low quality (SM_DSP, SM_PS, SM_LDM, SM_RF, SM_SM).Variations in meat

quality among different muscles may arise from divergent metabolic pathway preferences **(Fig. 6c, Extended Fig. 15)**. Within amino acid, fatty acid, and energy metabolism pathways, PAX7_Low_MuSCs were uniquely present in high-quality muscle with consistently high metabolic activity. Ribosome_High_Skeletal_Muscle exhibited divergent metabolic profiles between quality groups. In high-quality muscle, its enhanced amino acid and fatty acid metabolism promotes intramuscular fat deposition and amino acid storage to sustain good meat quality. In low-quality muscle, increased Oxidative Phosphorylation activity drives oxidative catabolism of amino acids and fatty acids and diminishes net storage of flavor-related lipids.

DLX3, GATA5 and SOX8 showed high regulator activity (AUCell scores) and broad expression in Ribosome_High_Skeletal_Muscle and PAX7_Low_MuSCs, identifying them as core myogenic regulators broadly and highly active in muscle fibers **(Fig. 6d)**. Numerial checks verified their RSSZ peaks resided within their respective cell-type/quality groups. (1) Ribosome_High_Skeletal_Muscle from low-quality muscle was enriched in specific transcriptional regulators: PITX1 and PITX3 are validated fast-twitch myofiber identity determinants, and HES6 and CREB1 control myogenesis and muscle metabolism. (2) PAX7_Low_MuSCs and Ribosome_High_Skeletal_Muscle of high-quality muscle shared regulators with concordant regulatory features (EN2, MSC, ELF5, SOX7, TLX2, TBX5, LHX5, DBP). High-quality-specific regulatory modules showed high similarity between progenitors (PAX7_Low_MuSCs) and differentiated myofibers (Ribosome_High_Skeletal_Muscle). Such progenitor–mature myofiber coupling agrees with the finding that high-quality muscle has concurrent high amino acid/fatty acid metabolism and abundant intramuscular lipid and amino acid deposition **(Fig. 6e, Supplementary Table S14)**.

Integrational analyses of the gut microbiota elucidate the mechanisms underlying roughage tolerance in TBPs

We systematically integrated multidimensional technical strategies, including association analyses of gastrointestinal regions, cell-microbiota interactions, differential gene expression and pathway enrichment, to uncover the unique gastrointestinal adaptive signatures of TBPs under long-term environmental adaptation at a detailed level **(Fig. 7a–7d, Supplementary Fig. S12-18, Extended Fig. 16, Supplementary table S15-19)**. After filtering contaminating taxa, eleven resident gut genera (*Bacteroides*, *Blautia*, *Coprococcus*, etc.) were retained. Hierarchical clustering classified them into three functional clusters: (1) dominant colonic fermentative genera (*Bacteroides*, *Blautia*, *Coprococcus*, *Campylobacter*, *Bacillus*, *Bifidobacterium*), enriched in cecum and colon with low abundance in stomach and small intestine; (2) transitional genera (*Clostridium*, *Escherichia*, *Clostridioides*), abundant in jejunum and ileum and decreasing toward large intestine; (3) upper GI-dominant genera (*Enterococcus*, *Acinetobacter*), high in stomach and duodenum and declining distally **(Fig. 7a, Supplementary Fig. S12)**. Furthermore, we performed correlation analysis between the abundance of microbiota and the proportion of 24 gastrointestinal cell types in each sample. After filtering, 21 cell types retained significant associations with microbial profiles. The results showed that five cell types, including Pericyte, Lymphatic_Endothelial, Esophagus_Glandular, Enteroendocrine and Naive_B, exhibited significant correlations with microbial abundance **(Fig. 7b, Supplementary Fig. S13)**.

The `FindMarkers` function in Seurat was used to identify DEGs between TBP and LWP across Pericyte, Lymphatic_Endothelial, Esophagus_Glandular, Enteroendocrine, and Naive_B **(Supplementary Fig. S14–S18, Supplementary Table S15–S19)**. (1) A total of 2,140 DEGs were

identified in pericytes: Up-regulated genes were markedly enriched in axon guidance and synaptic signaling (GO:0007411), with representative genes including *ERBB4*, *DLG2*, *RBFOX1*, *CACHD1* and *TENM2*. These genes were further enriched in glutamatergic synapse pathway (*GRIA1*, *GRIK3*, *GRIK4*, *CACNA1A*) and Rho GTPase regulation (*TRIO*, *DOCK4*, *DOCK5*), suggesting a neuro-like phenotypic shift of pericytes and enhanced neuro-vascular coupling in TBP pigs. By contrast, down-regulated genes were predominantly related to protein-synthesis machinery (66 ribosomal proteins such as *RPS27*, *RPL4*, *RPL5*, and translation elongation factors *EEF1A1*, *EEF2*) and mitochondrial oxidative phosphorylation, alongside reduced iron-storage (*FTH1*, *FTL*) and inflammatory (*PTGS2*/*COX-2*) genes. (2) Lymphatic_Endothelial contained 4,953 DEGs (second most). Upregulated genes were enriched in protein phosphorylation, ErbB/EGF signaling, Rho GTPase cycling and adhesion molecule pathways with increased complement genes; downregulated genes were ribosomal, translation and oxidative phosphorylation genes, supporting active signaling coupled with restrained translation and energy metabolism. (3) Esophagus_Glandular yielded 4,153 DEGs. Upregulated genes were enriched in extracellular matrix remodeling and smooth muscle contraction, while mucus barrier genes (MUC5AC, TFF2, GKN1, GKN2) were downregulated, pointing to an adaptive strategy relying on mechanical strength rather than mucus lubrication. (4) Enteroendocrine possessed the largest DEG set (5,649). Downregulated hormone-related genes (*PYY*, *GIP*, *CCK*, *GAST*, *GCG*, *GHRL*, *SST*) plus ribosomal and oxidative phosphorylation genes; upregulated genes fell into PI3K-Akt survival, focal adhesion and cytoskeleton pathways. (5) Naive_B had 2,303 DEGs. Upregulated genes included MHC II genes (*HLA-DRA*, *HLA-DOB*), *FCRL3*, *FCRL1*, enriched in BCR signaling, FcεRI-dependent Ca^{2+} mobilization and phosphatidylinositol pathways. Downregulated genes included *CD79A*, *CD79B*, *B2M*, *PTGS2*,

ribosomal and translation genes.

DEGs across diverse intestinal cell types between the two pig breeds were mainly associated with substance transport, protein metabolism and homeostasis, cytoskeleton organization, signal transduction, cell adhesion and migration **(Extended Fig. 16)**. KEGG enrichment network revealed that AMPK, Insulin, FoxO, longevity regulating and ErbB signaling pathways were interconnected to form a regulatory module governing metabolism and energy status. Pathways including Integrin, Focal adhesion, Regulation of actin cytoskeleton and Tight junction constituted a network for cytoskeleton and cell junction. Immune pathways such as TNF, T cell receptor and C-type lectin receptor exhibited cell-type-specific differences. In addition, Protein processing in endoplasmic reticulum, Ubiquitin mediated proteolysis, Mitophagy and Proteasome pathways were involved in protein processing and organelle quality control. Collectively, intestinal differences between TBPs and LWPs need multiple interconnected functional layers **(Fig. 7c)**. Notably, nutrient- and energy-sensing pathways including AMPK, Insulin and FoxO were particularly prominent.

RSS values were used to quantify the enrichment and specificity of regulators across groups to further compare their specificity **(Fig. 7d)**. The TBP regulatory network comprised multiple core transcription factor modules with large target gene pools and moderate inter-module crosstalk. In LWP, the network was built around *ELK3*, *ELF3*, *BHLHE40* and *ATF3* with partial inter-module connections; ATF4 formed a separate, smaller module. The top five specific transcription factors of TBP were *MITF*, *ARID5B*, *ERG*, *ETV6* and *GATA4*. These five core regulators each formed independent target gene modules. *ARID5B*, *ERG* and *ETV6* displayed radiating network structures with abundant target genes, and *GATA4* and *MITF* also had large target gene networks. The top five specific transcription factors of LWP were *ATF3*, *ATF4*, *BHLHE40*, *ELF3* and *ELK3*. *ELK3*, *ELF3*,

BHLHE40 and *ATF3* established distinct radiating target gene networks, while *ATF4* had fewer target genes and a smaller network size.

Cross-species comparisons provide insights into the application of TBPs in translational medicine

To systematically present the characteristics of gene expression differences between humans and TBPs at the multiorgan level, a heatmap showing the significance of expression differences in Mendelian genetic disease-related genes across different organs was generated **(Fig. 8a, Supplementary Fig. S19).** No significant differences in the expression of 11 Mendelian genetic disease-related genes (*MSTN*, *OBSL*, *MC4R*, *SOX10*, *DCLRE1C*, *HOXA1*, *TYRP1*, *RHO*, *PNKP*, *BMP15*, and *CYP27B1*) were observed between human and TBP (small differences in expression levels). Compared with other transplanted organs, kidneys typically have a longer survival time in humans, of more than 270 days **(Fig. 8b)**. As shown in Fig. 8c, substantial transcriptomic divergence was observed in the heart, kidney and liver across humans, TBP and LWP ($P < 0.01$). Specifically, the kidney exhibited relatively low transcriptomic heterogeneity between TBP and humans, while the heart and liver showed lower heterogeneity between LWP and humans. Boxplots showing cell-type transcriptional heterogeneity depict transcriptomic differences of matched cell types among TBP, LWP and humans in the heart, liver and kidney **(Fig. 8d)**. In the kidney, epithelial cell types exhibited low transcriptomic heterogeneity between TBP and human ($P > 0.05$), whereas significant heterogeneity was observed between TBP and LWP ($P < 0.001$). In contrast, no significant difference was detected in hepatic Central_Venous_LSECs and Cholangiocytes between human and LWP. As shown in Fig. 8e, kidney tissues display considerable transcriptomic divergence across humans and multiple pig breeds ($P < 0.01$). Compared with LWP, Bama, DLY and Meishan pigs, TBP showed

no significant transcriptomic heterogeneity in kidney epithelial cell types relative to humans **(Fig. 8f)**. Key immune-associated genes that were differentially expressed between humans and TBPs were further identified and included *IL15*, *CXCL14*, *TNFSF13B*, *CCL14*, *TNFSF11*, and *MMP9* in the liver and *B2M*, *CORO1A*, and *CCL4* in the kidney **(Supplementary Table S20-21).**

Discussion

Pigs are a key livestock species in the animal husbandry industry and are a major source of animal protein[24]. Additionally, they serve as crucial large animal models for investigating complex human diseases and are ideal alternative donors for organ transplantation[7,24]. Their high homology with humans in terms of physiological characteristics, anatomical structure, genetic background, and organ size has established their core value in improving livestock production quality and advancing biomedical innovation[24]. Here, this study constructed a pan-porcine single-cell reference atlas comprising 9 484 620 cells and 1 001 samples. It completed the single-nucleus transcriptomic profiling of 31 tissues and organs from TBPs, covering a total of 160 sub-organ sites for the first time. After meticulous annotation, 228 cell subtypes were identified, which were grouped into nine major cell lineages. Cai et al. (2026) generated the all-from-one spatiotemporal single-cell atlas for pigs, profiling 2.56 million cells across 234 tissue sites from a pregnant sow and its male fetus, and annotated 890 cell types[25]. Wang et al. performed single-cell transcriptome sequencing of 20 pig tissues and successfully obtained 200,000 cells[21]. Yu et al. acquired 1.07 million single cells in a study focusing on single cells in the pig intestine[22]. Han et al. (2025) constructed the world's first multi-tissue single-cell expression atlas of dairy cows, which systematically covered 59 tissues, contained 1.79 million cells, and included 131 annotated cell types[23]. Our results substantially exceeded the number of single cells reported in previous studies and addressed the limitations of

traditional organ-level research, achieving a single-cell transcriptomic resolution for fine suborgan partitions. These findings lay an important foundation for the use of single-cell data to elucidate the cellular mechanisms underlying pig growth and development, disease occurrence, and regenerative repair.

At the suborgan level, the cell-type composition and transcription profiles exhibited significant system specificity and functional heterogeneity. The heterogeneous expression of DEGs within the same cell type across different suborgans indicates the precise regulation of cell function by the suborgan microenvironment[26]. Within a given cell type, the physiological demands of different suborgans (e.g., nutrient absorption pressure in different intestinal segments or metabolic load in different hepatic lobes) can drive the specific remodelling of gene expression, thereby realizing suborgan functional specialization[27]. On the basis of our systematic analyses at the suborgan level, we constructed a more comprehensive single-cell transcriptomic atlas of pigs than was previously available and obtained the key insight that suborgan heterogeneity is the core source of organ functional diversity. This provides precise targets for subsequent comprehensive research on the regulation of livestock growth and development as well as region-specific mechanisms of disease occurrence, such as intestinal segment-specific diseases and lobe-selective liver injury.

High-altitude hypoxia and cold stress endow TBPs with unique physiological and molecular adaptive phenotypes[12]. Previous studies mostly screened adaptive candidate loci based on bulk transcriptomics and population genomics, lacking mechanistic dissection at single-cell resolution[12,13,28,]. We observed module switching and reversal of correlation coefficients in co-expression patterns between TBPs and LWP in the same organs, indicating divergent transcriptional programs between breeds. Gene modules associated with cold tolerance and hypoxia adaptation

exhibited strong organ specificity: significant differential regulation was detected in the heart, intestine, kidney, lymph node and muscle, whereas no divergent modules were identified in the brain, suggesting that highland adaptation is not a systemic uniform response but involves organ-specific labor division under cold and hypoxic stress[13,29,30]. Single-cell metabolic analysis revealed specialized metabolic signatures across cell types: epithelial cells were enriched in glucose and amino acid metabolism, endothelial cells prioritized energy and lipid metabolism, and mesenchymal cells were dominated by lipid metabolism and pathways for cellular structural remodeling, supporting that cell-type division of labor orchestrates systemic metabolic reprogramming. The pySCENIC-based gene regulatory network analysis identified a panel of core regulators with breed- and organ-preferred activity, including *MAF*, *IRF8*, *SOX10* and *TCF4*, which play critical roles in cold tolerance and hypoxic adaptation[31-34]. Collectively, highland adaptation in TBPs arises from coordinated transcriptional remodeling at organ, sub-organ and cell-type levels rather than global alteration of a single pathway. This finding complements previous single-cell studies on the pituitary, lung and brain of TBPs, demonstrating that highland adaptation represents a complex regulatory process involving multiple cell populations [35,36].

Meat quality is an important economic trait of local pig breeds [37]. The muscles from different anatomical sites show substantial differences in myofiber subtypes, stroma and stem cell components [38]. In this study, skeletal muscles of TBPs were classified into high-quality, medium-quality and low-quality groups based on metabolic characteristics. PAX7_Low_MuSCs in high-quality muscles maintain sustained high metabolic activity. Ribosome_High_Skeletal_Muscle cells undergo metabolic reprogramming across meat quality groups: the metabolic activities of amino acids and fatty acids are elevated in high-quality muscles, which is conducive to intramuscular fat

deposition; oxidative phosphorylation is enhanced in low-quality muscles, accelerating lipid breakdown and consumption. Regulatory network analysis reveals that DLX3, GATA5 and SOX8 are core myogenic regulatory factors with broadly high activity in the myofiber lineage [39-41]. The progenitor cells and mature myofibers of high-quality muscles share a set of transcriptional regulatory modules, indicating that a synergistic transcriptional program exists during the differentiation of muscle progenitor cells into myofibers to maintain muscle development and meat quality homeostasis. Single-cell studies on livestock skeletal muscle have indicated that muscle stem cells, myofiber subtypes and stromal adipocyte progenitors jointly determine the meat quality phenotype [42]. This study further reveals that the regulatory coupling relationship between muscle progenitor cells and mature myofibers constitutes the molecular basis for the superior meat quality of TBPs, providing cellular-level evidence for elucidating meat quality formation in local pig breeds.

The interaction between the host and gut microbes is a key link for the gastrointestinal adaptation of plateau animals to plateau feeding conditions [43,44]. The colonized microbiota in the gastrointestinal tract of TBP can be classified into colon fermentation-dominant microbiota, transitional microbiota, and upper gastrointestinal tract-dominant microbiota, and the abundance of microbiota shows a gradient distribution along the anatomical positions of the gastrointestinal tract. Comparative analysis between TBP and LWP revealed that axon guidance and glutamatergic synaptic pathways were significantly upregulated in pericytes, suggesting that these cells underwent a neuron-like phenotypic transformation with enhanced neurovascular coupling [45]. The expression of satiety hormone genes such as *PYY, GIP* and *CCK* was downregulated in enteroendocrine cells, while the PI3K-Akt cell survival pathway was activated. This molecular signature may underpin the physiological phenotype of continuous foraging in TBP under low energy supply at high altitudes

and cold environment [46]. KEGG enrichment analysis indicated that nutrient-sensing pathways including AMPK, FoxO and insulin, as well as cytoskeleton, adherens junction and immune signaling networks, were markedly differentiated between the two pig breeds. Consistent with published findings on host-microbe interactions in the porcine gut, multiple non-epithelial cells in the gut can respond to microbial signals, rather than being restricted to epithelial cells and immune cells only [47]. This study demonstrates that gastrointestinal adaptation in TBPs is driven synergistically by host cell transcriptional reprogramming, spatially partitioned colonization of gut microbiota, and cell-microbe interactions. These results further confirm that the roughage tolerance trait is not determined solely by gut microbes [48]. Multiple cell types in the host gut adapt to high-fiber diets through transcriptional reprogramming in coordination with microbiota, providing theoretical support for roughage-tolerant livestock breeding from the perspective of host-microbe interactions.

Pig is an important model animal for translational medicine research [7]. Nevertheless, studies on disease models using TBPs remain relatively limited, and research dissecting their cellular composition and transcriptional features at single-cell resolution is particularly scarce. Our multi-organ comparison revealed striking conservation in the expression profiles of Mendelian disease-associated genes between human and TWP. Specifically, 11 well-documented disease genes [49, 50]—including *MSTN* (leg weakness and muscular hypertrophy), *OBSL1* (growth disorders, syndromic), *MC4R* (hyperphagia leading to hepatic steatosis), *SOX10* (Waardenburg syndrome), *DCLRE1C* (severe combined immunodeficiency disease), *HOXA1* (microtia), *TYRP1* (coat colour, brown), *RHO* (retinal atrophy, progressive), *PNKP* (embryonic lethality), *BMP15* (fecundity), and *CYP27B1* (vitamin D deficiency rickets, type IA)—showed no significant expression differences across major

organs between the two species. This transcriptional conservation implies that the core functions and regulatory logic of these genes are preserved across human and TWP, minimizing the risk of interspecies phenotypic mismatch when constructing corresponding genetic disease models.

Pigs are widely recognized as promising organ donors for xenotransplantation owing to their anatomical, physiological and genomic similarities to humans, yet the suitability of different pig breeds for transplantation varies substantially and remains incompletely characterized at single-cell resolution [8, 51]. Our cross-species transcriptomic analysis demonstrated substantial divergence in the heart, kidney and liver across humans, TBP and LWP, with organ-specific conservation patterns that have direct implications for xenotransplantation. Notably, the kidney exhibited the highest degree of transcriptomic conservation between humans and TBP among the three organs examined. At the cell-type level, kidney epithelial cells showed no significant heterogeneity between TBP and humans, whereas pronounced divergence was observed between TBP and LWP. When compared with multiple commercial and miniature pig breeds including LWP, Bama pigs, DLY and Meishan pigs, TBP consistently displayed the lowest transcriptomic heterogeneity in matched kidney epithelial cell types relative to humans. The high transcriptional similarity of kidney tissue likely contributes to better functional compatibility and reduced immune recognition, which may partially explain the superior clinical outcomes of renal xenotransplantation [52,53]. In contrast, the heart and liver exhibited greater transcriptomic divergence between TBP and humans, while LWP showed relatively lower heterogeneity in these organs. Specifically, hepatic central venous liver sinusoidal endothelial cells and cholangiocytes showed no significant difference between LWP and humans, indicating breed-specific patterns of molecular compatibility across organs. These findings suggest that no single pig breed is universally optimal for all organ xenotransplantation, and breed selection

should be tailored to the specific organ of interest [54]. We further identified key immune-related differentially expressed genes between humans and TBP in the liver and kidney. Among these, *B2M* encodes β2-microglobulin, an essential component of MHC class I molecules that plays a central role in xenogeneic immune recognition. Genetic ablation of *B2M* in donor pigs has been shown to reduce SLA class I expression and attenuate human T-cell-mediated rejection, representing a standard strategy for generating hypoimmunogenic xenografts [55]. The differential expression of chemokines such as CCL4 and CXCL14 further highlights the involvement of innate and adaptive immune pathways in cross-species incompatibility [10,56]. These candidate genes provide prioritized targets for subsequent gene editing efforts to improve immunological compatibility of Tibetan pig organs.

Conclusion

The present study constructed a suborgan-resolution pan-pig single-cell reference atlas comprising 9,484,620 cells and 1,001 samples. For the first time, single-nucleus transcriptomic profiling was performed on 31 tissues and 160 suborgan sites of TBPs. A total of 228 cell subtypes were annotated and grouped into 9 major cell lineages, verifying that suborgan heterogeneity serves as the core source of organ functional diversity. Breed comparison revealed that high-altitude adaptation in TBPs represents a transcription remodeling process coordinated across the organ–suborgan–cell multi-scale hierarchy. Transcriptional coupling between muscle progenitor cells and mature muscle fibers underpins their superior meat quality. The high-fiber adaptation of the gastrointestinal tract relies on synergistic interactions between host transcriptional reprogramming and gut microbiota, rather than the independent effects of microbes. Cross-species analysis indicated that the renal transcriptomic signatures of TBPs are more conserved with humans than those of other

pig breeds, conferring potential for use as donors in xenogeneic kidney transplantation, whereas the heart and liver carry substantial risks of molecular incompatibility. The immune differentially expressed genes identified in this study can provide targets for gene editing. In summary, this multi-breed pig single-cell atlas constitutes a valuable data resource for dissecting the mechanisms underlying high-altitude adaptation and meat quality formation in TBPs, and for advancing livestock molecular breeding, disease model establishment, and xenotransplantation research.

Methods

Ethical statement

The study was approved by the by the Animal Care and Use Committee of Yunan Agricultural University (No. 202503001)

Experiment design and single cell sequencing

One male and one female 600-day-old TBP were obtained from a TBP farm. In total, 325 samples were collected from 31 organs and 160 suborgans (**Supplementary Table S1-S2**). Fresh or cryopreserved tissue samples were dissociated using physical shearing, chemical lysis, and other methods. The resulting mixture was subsequently washed and purified to obtain nucleus suspensions. The suspensions were stained with AO/PI reagent and analysed using a Countstar Mira FL-S automatic cell fluorescence analyser to evaluate quality control indicators, including cell viability, aggregation rate, impurity content, and nucleus integrity. Qualified samples were used for single-cell library construction. Library construction was performed using the DNBelab C Series High-Throughput Single-Cell RNA Library Preparation Kit V3.0 according to the manufacturer's instructions. The main steps were as follows: (1) a water-in-oil droplet microreaction system was set up using a DNBelab C-TaiM 4 Single-Cell Droplet Generator, and cell lysis, mRNA capture,

and reverse transcription for cDNA synthesis were completed within the droplets; (2) emulsification was performed to recover the droplet contents, and oligo and cDNA products were obtained via magnetic bead separation; (3) the cDNA products were amplified and purified; (4) the oligo products were subjected to barcode labelling to construct an oligo library; (5) the cDNA products were subjected to fragmentation, end repair, and adapter ligation to construct a cDNA library; (6) an Invitrogen Qubit 4 fluorometer was used to determine the concentrations of the cDNA and oligo libraries, and a fragment analyser was used to determine library fragment sizes. Libraries meeting the following quality standards were used for subsequent steps: oligo library concentration >10 ng/μL with fragment lengths concentrated at 180 ± 10 bp; cDNA library concentration >10 ng/μL with fragment lengths distributed between 350 and 550 bp; (7) The cDNA and oligo libraries were denatured and cyclized separately to form single-stranded circular DNA; (8) rolling circle amplification was performed to generate DNA nanoballs (DNBs) containing multiple copies; and (9) finally, the DNBs were sequenced using a DNBSEQ-T7RS high-throughput sequencing platform to obtain high-quality single-cell transcriptomic sequencing data.

To construct a pan-porcine single-cell atlas and facilitate subsequent analyses, we integrated publicly available porcine single-cell transcriptomic datasets from published studies for joint analysis. These datasets cover embryonic tissues [57,58], liver [19], skin [59,60], skeletal muscle [61,62], gastrointestinal tract [22], testis [63], kidney [64,65], as well as previously reported pan-tissue single-cell atlases [21,25,66,67]. Detailed information of these datasets is provided in Supplementary Tables S5 and S6.

Cell type annotation and clustering

Expression quantification was performed on all single-cell samples using dnbc4tools_v2.1.3,

CellRanger_v9.0 and celescope_2.11.1. The reference genome was obtained from the Ensembl database, with reference genome version Sscrofa.11.1 and gene annotation file version Sus_scrofa.Sscrofa11.1.113.gtf.

Gene filtering: For gene filtering of each sample, genes expressed in at least 3 cells were retained.

Cell filtering: 1. The number of detected genes < 200. 2. Based on the total UMI per cell, cells were filtered using the 99th quantile. 3. Cells were retained only when the percentage of mitochondrial transcripts from mitochondrial genes was less than 5%. 4. Doublet filtering: The sc.external.pp.scrublet function in Scrublet_v0.2.3 (based on an expected doublet rate of 5%) was used to infer doublets, and cells with a doublet score above 0.25 were excluded from subsequent analyses.

Cell annotation: Annotation tools CellTypist and CellID were used. For cell type annotation, we collected published single-cell transcriptomic data from 34 human tissues including adipose tissue, adrenal gland, blood vessel, bone marrow, brain, bronchus, embryo, epididymis and esophagus to construct a custom cell type annotation reference dataset, and collected 9 official CellTypist reference datasets covering blood, heart, intestine, liver, lung and other tissues. We additionally collected cell type marker genes covering the remaining 9 tissues. CellTypist (1.7.1) and CellID (1.10.1) were then applied for cell type annotation. For correction of cell type annotation results, clustering analysis was conducted for each sample, and the annotation of each cell was corrected to the most abundant cell type within its corresponding cluster.

For cell clustering, we first selected the top 2000 highly variable genes using the sc.pp.highly_variable_genes function for principal component analysis, and the Harmony algorithm

was subsequently used to correct the sample batch effect. We first performed clustering via the Leiden method in scanpy (v1.10.3) (resolution=0.8), and the t-SNE method was adopted to visualize cells after dimensionality reduction for clustering.

Identification of marker genes

To identify marker genes for each cell cluster, differential expression analysis was performed using the scanpy.tl.rank_genes_groups function in Scanpy (version 1.10.3). We employed a one-vs-rest strategy using the non-parametric Wilcoxon rank-sum test to compare the gene expression profile of each individual cluster against all remaining cells. P-values were adjusted for multiple testing using the Benjamini-Hochberg (BH) correction to control the false discovery rate (FDR). Genes were considered as statistically significant marker genes if they met the following criteria: an adjusted p-value < 0.01, a minimum $\log_2$(fold-change) > 2.

Cell cycle analysis

The *score_genes* function in Scanpy was used for the binary classification of cell cycle states based on two gene sets: cell cycle marker genes for G1, S, G2, and M phases (to determine cycling status) and G0 phase marker genes (to determine quiescent status). The cell cycle index was determined by calculating the logarithmic ratio (log10) of the number of cycling cells to the number of quiescent cells for each cell type. The calculation of this index relied on high-confidence cell cycle marker genes from G1-M phases (G1/S phase: *CDCA7L* [*CEP57*]; S phase: *ABHD10*, *CCDC14*, *CDKN2AIP*, *NT5DC1*, *SVIP*, and *PTAR1*; G2 phase: *YEATS4* and *DCTPP1* [*ANKRD36C*]; G2/M phase: *SMC4*, *TMPO*, *LMNB1*, and *HINT3*; M phase: *HMG20B* and *HPS4* [*HMGB3*]) to identify cycling cells and G0 phase marker genes (*CDKN1A*, *CDKN1B*, and *CDKN1C*) to identify quiescent cells[68].

Transcriptional similarity among organs and tissues

MetaNeighbor (v1.22.0) was used to calculate the area under the receiver operating characteristic curve (AUROC) between suborgans[69]. The gene expression matrix of the organs or tissues was first constructed with a SingleCellExperiment object as the input. The variable gene function was subsequently used to screen for highly variable genes, which were subsequently input into the MetaNeighborUS function (with fast_version = T). On the basis of the obtained AUROC similarity matrix, a visualization heatmap was generated using Complex Heatmap[70]. To unify the colour scale range of the heatmap, the AUROC values were subjected to 0–1 max–min normalization.

Transcriptional heterogeneity calculation

In this study, a standardized method was used to calculate single-cell transcriptional heterogeneity. First, a rigorous quality control scheme was applied to the raw single-cell RNA sequencing data, and a two-step downsampling strategy was used to eliminate technical biases caused by differences in cell number and sequencing depth. (1) The downsampling cell thresholds used for different treatments are as follows: a threshold of 300 was used for cell-type transcriptional heterogeneity (Fig.3 c); a threshold of 10000 was used for organ transcriptional heterogeneity (Fig.8 b-f); a threshold of 5000 was used for suborgan transcriptional heterogeneity (Fig.4 a). (2) Based on the gene × cell matrix generated in the first step, further downsampling was conducted to ensure identical cell numbers and total UMI counts across the groups. To reduce technical noise and highlight biological variations, all the genes were subsequently divided into 10 bins according to their average expression levels. In each bin, 10% of the genes with the lowest coefficient of variation (standard deviation/mean) were selected to form a stably expressed signature gene set. This

signature gene set was employed to calculate the Euclidean distance between each pair of cells, and the results were used to measure transcriptional heterogeneity in each cell[71].

Metabolic activity analysis

To compare metabolic activity among the different cell types, we referenced the published Single-Cell-Metabolic-Landscape pipeline (https://github.com/LocasaleLab/Single-Cell-Metabolic-Landscape)[72]. Because the TBP atlas contains 2,338,851 cells, we initially used the aggregateReference function in the SingleR package (power = 0.7) to merge gene expression data by cell type and generate a meta-cell gene expression matrix, improving computational efficiency. After excluding cell types corresponding to fewer than four cells, the activity score of each metabolic pathway for each cell type was calculated using our original pipeline. Additionally, we used ALRA (a faster alternative) instead of scImpute in the original pipeline for missing-value imputation[73,74]. The screening criteria for low-dropout genes were as follows: a gene was expressed in at least 75% of the cell types, and in each cell type, it was expressed in at least 75% of the cells. Finally, permutation tests were used to evaluate the statistical significance of pathway activity in each cell type.

The analyses of differentially expressed genes (DEGs)

DEGs separately using the rank_genes_groups function in Scanpy (v1.10.3) with the Wilcoxon parameter method. Genes were considered as statistically significant marker genes if they met the following criteria: an adjusted p-value < 0.001, a minimum $\log_2$(fold-change) > 2. DEGs separately using the rank_genes_groups function in Scanpy (v1.10.3) with the Wilcoxon parameter method. Genes were considered as statistically significant marker genes if they met the following criteria: an adjusted p-value < 0.001, a minimum $\log_2$(fold-change) > 2.

Construction of coexpression networks

hdWGCNA was used to identify characteristic gene modules in the studied organ systems. The specific workflow was as follows. First, genes expressed in at least 5% of the cells were filtered to construct a WGCNA object. Metacells were constructed using the "MetacellsByGroups" function to reduce data sparsity while retaining cellular heterogeneity. The "TestSoftPowers" function was subsequently used to calculate the optimal soft power threshold for constructing the coexpression network, and a soft threshold of four was determined to best fit the network structure in this study. The "ModuleEigengenes" function was used to calculate the module eigengenes. Finally, UMAP was used to embed the topological overlap matrix of the coexpression network into a two-dimensional manifold for its visualization[75].

Inference of gene regulatory networks

PySCENIC was used to construct gene regulatory networks for specific organs and organ systems. The GRBboost2 algorithm was initially used to infer gene regulatory networks, followed by a cis-regulatory motif analysis for each coexpression module using the "pyscenic ctx" function. Subsequently, AUCell was used to score the target genes of transcription factors. Finally, the "calcRSS" function identified transcription factors that were specific to each organ and suborgan system. A transcription factor was considered significant when its AUCell score was > 0.1 and its Z-score-normalized regulon specificity score (RSSZ) was > 1.0. A higher RSSZ value indicates that the regulator is more specifically expressed in the corresponding organ, where positive values indicate high expression, and negative values indicate low expression[76].

Analysis of nutritional components in TBP muscle

The T700 metabolomics platform was used to detect nutritional components in samples from

10 muscle sites (n=40). The detection indicators included total amino acids, essential amino acids, umami amino acids, sweet amino acids, total fatty acids, polyunsaturated fatty acids and monounsaturated fatty acids. The T700 metabolomics detection of muscle samples was completed by Wuhan Metware Biotechnology Co., Ltd (Wuhan, China).

Metagenomic sequencing of gut microbiota in TWP pigs

Microbial genomic DNA was extracted from stool samples using the E.Z.N.A.® Stool DNA Kit (Omega Bio-tek, Norcross, Georgia, USA). After metagenomic shotgun sequencing libraries were constructed, next-generation high-throughput sequencing was performed on the DNBSEQ-T7 platform (BGI, Shenzhen, China) with a 150 bp paired-end read mode. Raw data processing, sequence alignment, host contamination removal, and taxonomic annotation were all carried out following the same protocol [77].

MaAsLin 3 (Microbiome Multivariable Associations with Linear Models 3) is an R package for association analysis of microbiome, meta-omic features (taxa, genes, pathways), and clinical/environmental metadata, built on hurdle two-part generalized linear/mixed models. This method first performs TSS normalization and log transformation on microbial features. It uses logistic regression to model feature prevalence, and linear/log-linear models to fit non-zero abundance. It supports the configuration of fixed effects, random effects, interaction terms, and is suitable for cross-sectional and longitudinal cohorts. This framework can distinguish true zeros from technical zeros, and separate two signals: microbial detection absence and abundance alteration. The software corrects compositional bias in relative abundance using the median coefficient comparison method, and can also model with spike-in and qPCR absolute quantification data. It outputs effect sizes and P-values for each association pair at both prevalence and abundance levels,

followed by FDR correction to reduce false positives caused by confounding factors and compositional effects, so as to identify microbial features independently associated with phenotypes.

Data availability

The single cell RNA-sequencing data generated in this study have been deposited in the National Genomics Data Center (NGDC) database, with accession number CRA053414 (https://ngdc.cncb.ac.cn/gsa/s/1wqziUBI) and CRA031730 (https://ngdc.cncb.ac.cn/gsa/s/PYy3TzGe).

Acknowledgements

This work was supported by the Nature Science Foundation of China (No. U1802234, No. 32160795), the "Xing Dian Talent Support Program" of Yunnan Province (XDYC-CYCX-2022-0029 and XDYC-QNRC-2023-0403), and the special Fund for Yunnan Agriculture Research System.

Author contributions

Conceptualization, H.H., X.B., C.D., X.F.Z., S.Y.Q., and H.B.P; Methodology, H.H., X.B., C.D., Y.C.M., A.J.L., T.Z.G., Y.T.Y, S.K.T., L.S., and F.Z.; Sequencing, C.D., Y.C.M., S.K.T., L.S., and F.Z; Writing-Original Draft, H.H., X.B., C.D., Y.C.M., and H.B.P.; Writing-review & editing, all authors; Funding acquisition, H.H., X.B., X.F.Z., S.Y.Q., and H.B.P; Resources, H.H., X.B., X.F.Z., S.Y.Q., and H.B.P; Supervision, H.H., X.B., X.F.Z., S.Y.Q., and H.B.P.

Competing interests

The authors declare no competing interests.

References

1. Chen, C. et al. Expanded catalog of microbial genes and metagenome-assembled genomes from

the pig gut microbiome. *Nat. Commun.* **12**, 1106 (2021).

2. Zhou, Y. et al. Characterization of the pig lower respiratory tract antibiotic resistome. *Nat. Commun.* **14**, 4868 (2023).

3. Jiang, G., Hu, H. & Wang, Y. How do livestock environmental regulations promote the vertical integration of the livestock industry chain? Evidence from Chinese-listed livestock enterprises. *J. Clean. Prod.* **413**, 137508 (2023).

4. Sievert, K., Lawrence, M., Parker, C. & Baker, P. How power in corporate-industrial meat supply chains enables negative externalities: Three case studies from Brazil, the US, and Australia. *One Earth* **7**, 1424-1441 (2024).

5. Wang, L. et al. Single-nucleus and bulk RNA sequencing reveal cellular and transcriptional mechanisms underlying lipid dynamics in high marbled pork. *Npj Sci. Food* **7**, 23 (2023).

6. Yang, H. et al. ABO genotype alters the gut microbiota by regulating GalNAc levels in pigs. *Nature* **606**, 358-367 (2022).

7. Lunney, J. K. et al. Importance of the pig as a human biomedical model. *Sci. Transl. Med.* **13**, eabd5758 (2021).

8. Tao, K. S. et al. Gene-modified pig-to-human liver xenotransplantation. *Nature* **641**, 1029-1036 (2025).

9. He, J. et al. Pig-to-human lung xenotransplantation into a brain-dead recipient. *Nat. Med.* **31**,3388-3393 (2025).

10. Schmauch, E. et al. Integrative multi-omics profiling in human decedents receiving pig heart xenografts. *Nat. Med.* **30**, 1448-1460 (2024).

11. Wang, J. et al. Tandem mass tag-labeled quantitative proteomic analysis of tenderloins between

Tibetan and Yorkshire pigs. *Meat Sci.* **172**, 108343 (2021).

12. Li, X. et al. Adaptation of Diqing Tibetan pigs to hypoxic and cold environments through extramedullary hematopoiesis and uncoupled thermogenesis in the liver. *BMC Biol.* **23**, 181 (2025).

13. Lai, D. et al. Integrative genomic, transcriptomic and epigenomic analysis reveals cis-regulatory contributions to high-altitude adaptation in Tibetan pigs. *Mol. Biol. Evol.* **42**, msaf169 (2025).

14. Wang, Z. et al. Whole-genome analysis reveals distinct adaptation signatures to diverse environments in Chinese domestic pigs. *J. Anim. Sci. Biotechno.* **15**, 97 (2024a).

15. Wang, Y. et al. Pig-to-human kidney xenotransplants using genetically modified minipigs. *Cell Rep. Med.* **5**, 101744 (2024b).

16. Matar, A. J., Sachs, D. H. & Duran-Struuck, R. The MHC-characterized miniature swine: lessons learned from a 40-year experience in transplantation. *Transplantation* **106**, 928-937 (2022).

17. Zhang, M. et al. Reproducible porcine model for kidney allotransplantation of low weight miniature pig. *Transl. Androl. Urol.* **11**, 439-450 (2022).

18. Lu, T., Yang, B., Wang, R. & Qin, C. Xenotransplantation: current status in preclinical research. *Front. Immunol.* **10**, 3060 (2020).

19. Rao, L., Cai, L. & Huang, L. Single-cell dynamics of liver development in postnatal pigs. *Sci. Bull.* **68**, 2583-2597 (2023).

20. Xiao, Y. Y. et al. Single-cell profiling of the pig cecum at various developmental stages. *Zool. Res*. **45**, 55-68 (2024).

21. Wang, F. et al. Endothelial cell heterogeneity and microglia regulons revealed by a pig cell landscape at single-cell level. *Nat. Commun.* **13**, 3620 (2022).

22. Yu, P. et al. IPGCA: A Comprehensive Single Cell Atlas of 1 074 127 Porcine intestinal cells revealing cellular dynamics, genetic regulation, and cross‐species conservation. *Adv. Sci.,* e07882 (2025).

23. Han, B. et al. A multi-tissue single-cell expression atlas in cattle. *Nat. Genet.* **57**,2546-2561 (2025).

24. Teng, J. et al. A compendium of genetic regulatory effects across pig tissues. *Nat. Genet.* **56**, 112-123 (2024).

25 Cai, L. et al. Family single-cell atlases reveal pig pregnancy and fetal growth restriction critical cell types. *Science*. **26**, eadu7265 (2026).

26. Qu, F. et al. Three-dimensional molecular architecture of mouse organogenesis. *Nat. Commun*. **14**, 4599 (2023).

27. Dworak, H., Rozmaric, T., Grillari, J. & Ogrodnik, M. Cells of all trades–on the importance of spatial positioning of senescent cells in development, healing and aging. *Febs Lett.* **599**, 2087-2106 (2025).

28. Li, P. et al. Comprehensive transcriptome analysis reveals MSTRG.19853.1/ssc-miR-361-3p/NPPA axis is related to hypoxic adaptation in Tibetan pigs. *BMC Genomics*. **26**, 5 (2025).

29. Ilardo, M. A. et al. Physiological and genetic adaptations to diving in sea nomads. *Cell*. **173**, 9-580 (2018).

30. Schipke, J. D., Limper, U., & Tetzlaff, K. Adaptations to breath-hold diving: from traditional divers to elite athletes. *Diving Hyperb Med.* **56**, 365-370 (2026).

31. Liu, K., Wang, F., Wang, S., Li, W., & Ye Q. Mangiferin attenuates myocardial ischemia-reperfusion injury via MAPK/Nrf-2/HO-1/NF-κB in vitro and in vivo. *Oxid Med Cell Longev.*

13, 7285434 (2019).

32. Dzwigonska, M., et al. Hypoxic stress dysregulates functions of glioma-associated myeloid cells through epigenomic and transcriptional programs. *Cell Rep.* **44**,116222 (2025).

33. Li, Q., et al. Modulation of Sox10, HIF-1α, Survivin, and YAP by Minocycline in the Treatment of Neurodevelopmental Handicaps following Hypoxic Insult. *Am J Pathol.* **185**, 2364-78 (2015).

34. Jiao H, et al. Uncovering the chromatin-mediated transcriptional regulatory network governing cold stress responses in fish immune cells. *J Genet Genomics*. **52**, 1046-1057 (2025).

35. Li, H. et al. Integrating short-read and long-read single-cell transcriptomics of pig pituitary reveals mechanisms of high-altitude hypoxia adaptation. *PLoS Genet*. **22**, e1012267 (2026).

36. Chang, Y. et al. Single-cell long- and short-read transcriptomics sheds light on high-altitude adaptation in the porcine cerebral cortex. *Transl Psychiatry*. **16**, 565 (2026).

37. Huang, Y. et al. A large-scale comparison of meat quality and intramuscular fatty acid composition among three Chinese indigenous pig breeds. *Meat Sci*. **168**, 108182 (2020).

38. Micheli, A. J., Spector, J. A., Elemento, O., & Cosgrove, B. D. A reference single-cell transcriptomic atlas of human skeletal muscle tissue reveals bifurcated muscle stem cell populations. *Skelet Muscle*. **10,** 19 (2020).

39. Lee, K. et al. A Whole genome association study on meat quality traits using high density snp chips in a cross between Korean native pig and Landrace. *Asian-Australas J Anim Sci.* **25**, 1529-1539 (2012).

40. Guimarães D. et al. Concerted regulation of skeletal muscle metabolism and contractile properties by the orphan nuclear receptor Nr2f6. *J Cachexia Sarcopenia Muscle*. **15,** 1335-

1347 (2024).

41. Jin, S. et al. Genetic insights into large and small body size variation in Diqing Tibetan pigs: A comparative genomic analysis. *Genomics* **118,** 111200 (2026).

42. Xu, D. et al. Single-cell RNA-sequencing provides insight into skeletal muscle evolution during the selection of muscle characteristics. *Adv Sci*. **10**, e2305080 (2023).

43. Liu, K., Yang, J. & Yuan, H. Recent progress in research on the gut microbiota and highland adaptation on the Qinghai-Tibet Plateau. *J Evol Biol.* **34**, 1514-1530 (2021).

44. Wang, X.et al. Seasonal variations in composition and function of gut microbiota in grazing yaks: Implications for adaptation to dietary shift on the Qinghai-Tibet plateau. *Ecol Evol.* **14**, e70337 (2024).

45. Niciu, M. J., Kelmendi B. & Sanacora G. Overview of glutamatergic neurotransmission in the nervous system. *Pharmacol Biochem Behav.* **100**, 656-664 (2012).

46. Zhang, Y. et al. Porcine gut microbiota in mediating host metabolic adaptation to cold stress. *NPJ Biofilms Microbiomes.* **8**, 18 (2022).

47. Shekhar, S., Šrůtková, D. & Schwarzer M. Microbial metabolites and MAMPs: orchestrators of intestinal stem cell function. *Gut Microbes.* **18,** 2755115 (2026).

48. Wang, Y. et al. From genomes to phenomes: Multi-omics dissection of positive and balancing selection in Jinhua pig. *Genomics* **118**, 111218 (2026).

49. Amberger, J. S., Bocchini, C. A., Schiettecatte, F., Scott, A. F., & Hamosh, A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* **243**, D789-798 (2015).

50. Hou, N., Du, X., & Wu, S. Advances in pig models of human diseases. *Animal Model Exp Med.*

5, 141-152 (2022).

51. Montgomery, R. A. et al. Physiology and immunology of pig-to-human decedent kidney xenotransplant. *Nature* **650**, 218–229 (2026).

52 Cheung, M. D. et al. Spatiotemporal immune atlas of a clinical-grade gene-edited pig-to-human kidney xenotransplant. *Nat. Commun.* **15**, 3140 (2024).

53.Schmauch, E. et al. Multi-omics analysis of a pig-to-human decedent kidney xenotransplant. *Nature* **650**, 205–217 (2026).

54. Ali, A., Kurome, M., Kessler, B., Kemter, E., & Wolf, E. What genetic modifications of source pigs are essential and sufficient for cell, tissue, and organ xenotransplantation? *Transpl Int*. **37**, 13681 (2024).

55. Xu, K. et al. Production of Triple-Gene (GGTA1, B2M and CIITA)-Modified Donor Pigs for Xenotransplantation. *Front. Vet. Sci.* **9**, 848833 (2022).

56. Friedrich, V. et al. Deep learning-aided inter-species-comparison reveals shared and distinct molecular patterns in cynomolgus monkey and humans following non-specific T cell activation. *Front Immunol.* **16,** 1603716 (2025).

57. Chen, M. et al. Integration of single-cell transcriptome and chromatin accessibility of early gonads development among goats, pigs, macaques, and humans. *Cell Rep.* **41,** 111587 (2022).

58 Simpson, L. A single-cell atlas of pig gastrulation as a resource for comparative embryology. *Nat Commun*. **15**, 5210 (2024).

59. Zou, Q. et al. A single-cell transcriptome atlas of pig skin characterizes anatomical positional heterogeneity. *Elife* **12**, e86504 (2023).

60. Zheng, T. et al. A single-cell transcriptome atlas of pig skin reveals cellular heterogeneity from

embryonic development to postnatal aging. *BMC Biol*. **23**, 275 (2025).

61. Cai S, et al. Integrative single-cell RNA-seq and ATAC-seq analysis of myogenic differentiation in pig. *BMC Biol.* **21,** 19 (2023).

62. Chen, Z. et al. Single-nucleus RNA sequencing reveals cellular transcriptome features at different growth stages in porcine skeletal muscle. *Cells* **14**, 37 (2025).

63. Wang, X. et al. Single-cell transcriptomic and cross-species comparison analyses reveal distinct molecular changes of porcine testes during puberty. *Commun Biol*. **7**, 1478 (2024).

64. Yao, T. et al. Dynamic single-nucleus transcriptomic profiling of kidney development in Bamaxiang pigs and wild boar. *Zool Res*. **47**, 58-71 (2026).

65. Yao. T., Li, N., & Huang, L. S. Integrated single-cell transcriptomic map of pig kidney cells across various periods and anatomical sites. *Zool Res*. **46**, 469-482 (2025).

66. Chen, L. et al. A dataset of single-cell transcriptomic atlas of Bama pig and potential marker genes across seven tissues. *BMC Genom Data*. **26,** 16 (2025).

67. Chen, L. et al. Construction of a multitissue cell atlas reveals cell-type-specific regulation of molecular and complex phenotypes in pigs. *Adv Sci*. **13**, e04961 (2026).

68. Consortium, T. S. et al. The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science* **376**, eabl4896 (2022).

69. Crow, M., Paul, A., Ballouz, S., Huang, Z. J. & Gillis, J. Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. N*at. Commun.* **9**, 884 (2018).

70. Gu, Z. Complex heatmap visualization. *iMeta* **1**, e43 (2022).

71. Lai, Y. et al. Multimodal cell atlas of the ageing human skeletal muscle. *Nature* **629**, 154–164

(2024).

72. Xiao, Z., Dai, Z. & Locasale, J. W. Metabolic landscape of the tumor microenvironment at single cell resolution. *Nat. Commun.* **10**, 3763 (2019).

73. Linderman, G. C. et al. Zero-preserving imputation of single-cell RNA-seq data. *Nat. Commun.* **13**, 192 (2022).

74. Li, W. V. & Li, J. J. An accurate and robust imputation method scImpute for single-cell RNA-seq data. *Nat. Commun.* **9**, 997 (2018).

75. Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E. & Swarup, V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep. Methods* **3**, 100498 (2023).

76. Van De Sande, B. et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* **15**, 2247–2276 (2020).

77. Hu, H.et al. Mulberry branch fiber improved lipid metabolism and egg yolk fatty acid composition of laying hens via the enterohepatic axis. *Microbiome* **12**, 73 (2024).

Figures

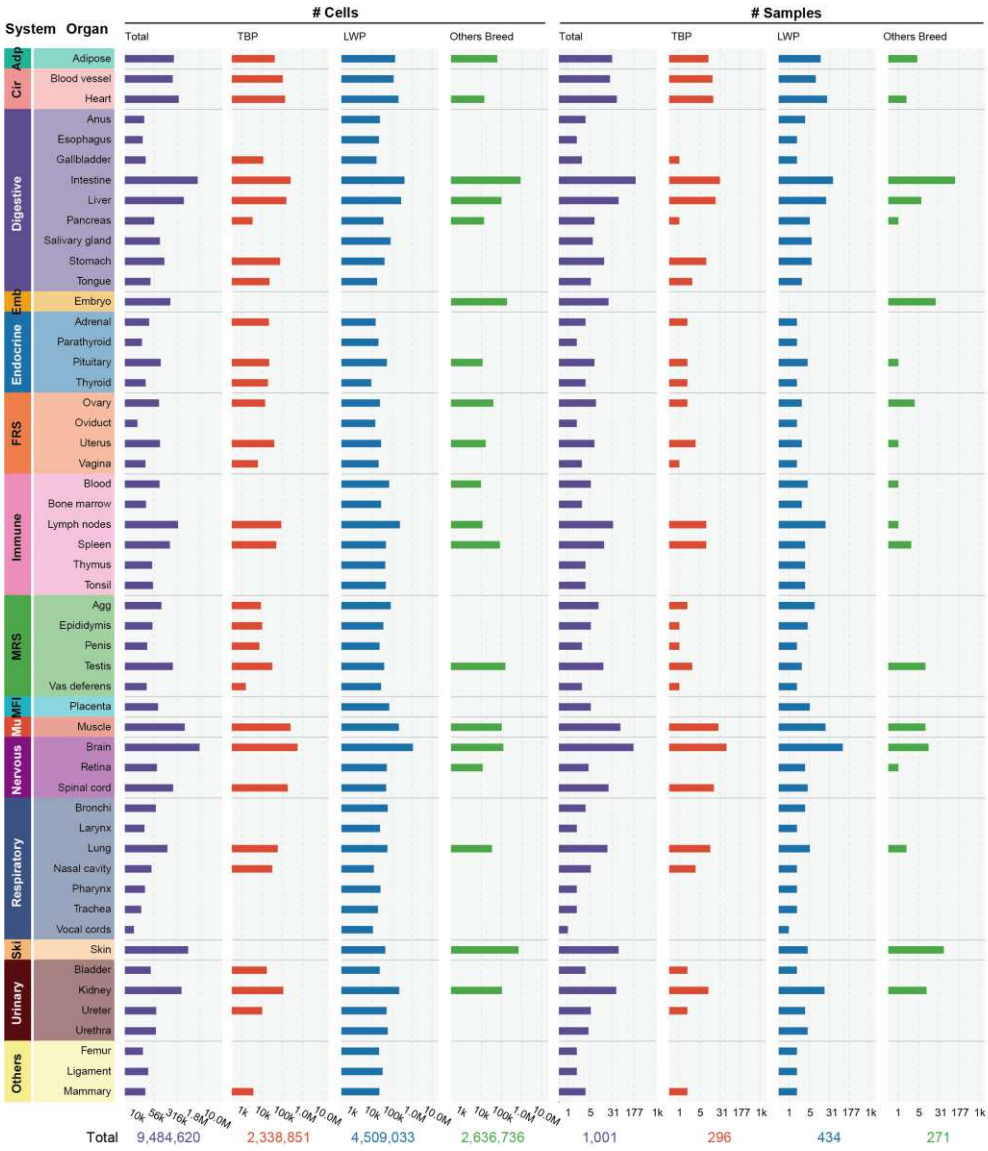


Figure 1 Overview of cell and sample numbers across organs/tissues in the integrated dataset.

This figure summarizes the cell count and sample size for each organ/tissue in the integrated single-cell transcriptomic dataset. Organs on the left are grouped by tissue systems, including Adipose (Adp), Circulatory (Cir), Digestive, Embryo (Emb), Endocrine, Female Reproductive (FRS), Immune, Male Reproductive (MRS), Maternal fetal interface (MFI), Muscle (Mu), Nervous, Respiratory, Skin (Ski), Urinary, and Others, with distinct colors assigned to each system. The two central panels display cell numbers (Cells, left) and sample numbers (Samples, right), respectively.

Within each panel, statistics are grouped by three pig breeds plus the total: Total (purple), TBP (red), LWP (blue), and Other Breeds (green). The x-axis is on a logarithmic scale (Cells: 1k, 10k, 100k, 1M, 10M; Samples: 1, 5, 31, 177, 1k). This dataset contains a total of 9,484,620 cells and 1,001 samples, among which TBP contributes 2,338,851 cells and 296 samples, LWP contributes 4,509,033 cells and 434 samples, and other breeds contribute 2,636,736 cells and 271 samples. The organ abbreviation Agg in the figure denotes the accessory genital gland.

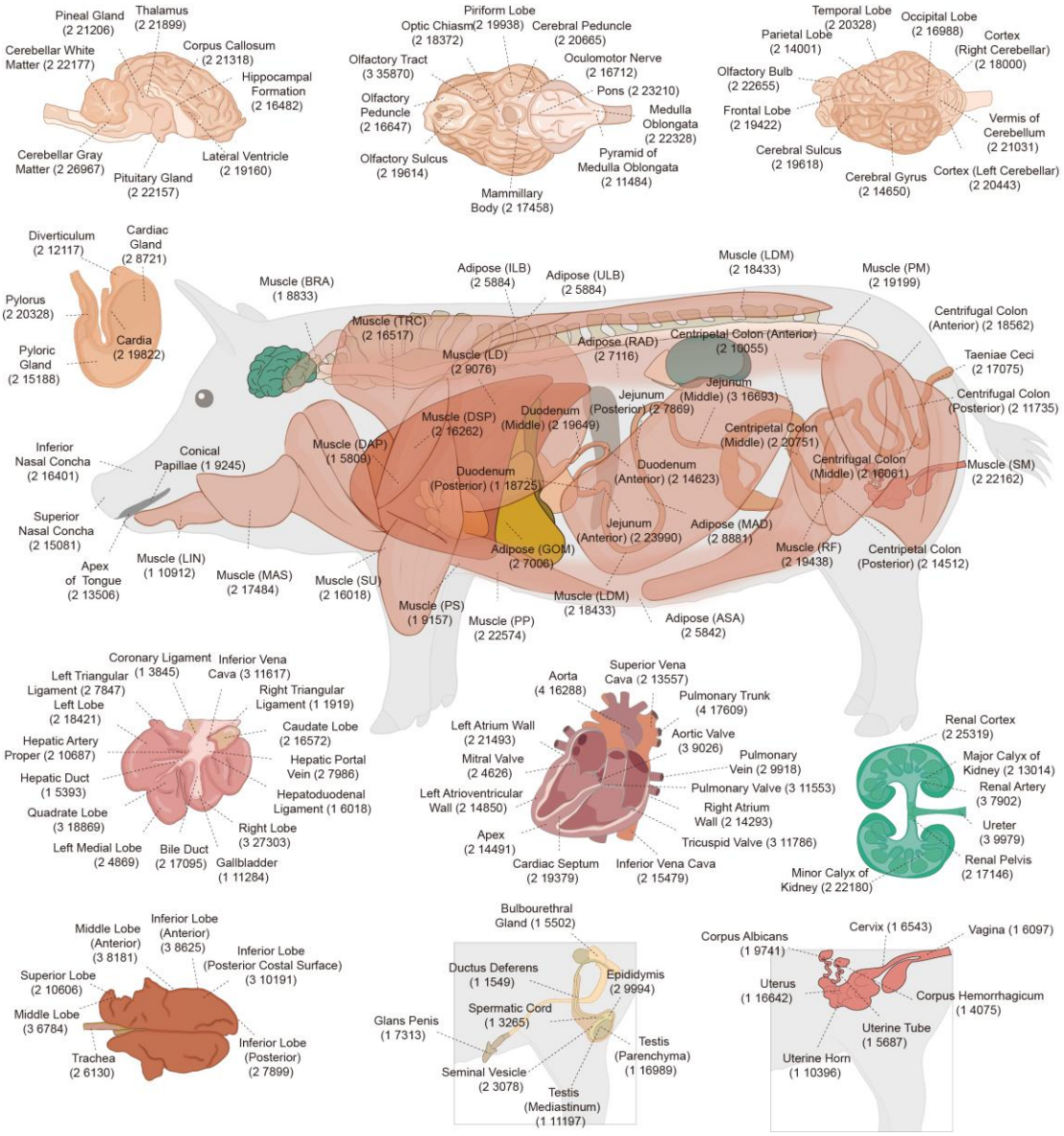


Figure 2 A Single-cell atlas across multiple tissues in Tibetan pigs.

This figure illustrates the major organs; tissues and their fine anatomical suborgans covered during the construction of the single-cell transcriptomic atlas of Tibetan pigs (TBP). A whole-body anatomical schematic of the pig is placed at the center to mark the spatial locations of various organ systems in vivo. Magnified substructures of key organs, including the brain (three views: sagittal, ventral and dorsal), stomach, liver, heart, kidney, lung, male reproductive system and female reproductive system, are displayed around the central schematic. Numbers in parentheses following each anatomical structure indicate the number of collected samples and recovered cells for the

corresponding sub-organ. This dataset covers multiple systems including digestive, nervous, circulatory, respiratory, urinary, immune, endocrine, muscle, adipose and reproductive systems, comprising 31 organ/tissue categories and 160 refined sub-organs/suborgans, with a total of 296 samples and 2,338,851 cells.



Figure 3 Circos Plot Depicting Multi-layered Genomic and Cellular Features and Uniform manifold approximation and projection (UMAP) visualization of Tibetan Pigs.

This Circos plot illustrates the overall cell type composition, transcriptional heterogeneity, and marker gene expression patterns of the porcine single-cell atlas. At the center is a t-SNE visualization of nine major cell superclusters, including Epithelial, Mesenchymal, Lymphoid, Myeloid, Neural, Endothelial, Germline, Stem, and Unknown cells. From the inner to the outer rings (a–n): a. Cell type annotations for 228 fine-grained cell types; b. Supercluster assignment for the nine major superclusters; c. Transcriptional heterogeneity: the degree of transcriptional

heterogeneity of each cell type; d. Cell cycle: cell cycle state distribution across cell types; e. Organ system proportion: the cellular proportion of each organ system within every supercluster; f–n. Marker gene expression: expression levels of signature genes for the nine superclusters (Epithelial (f), Mesenchymal (g), Lymphoid (h), Myeloid (i), Neural (j), Endothelial (k), Germline (l), Stem (m), and Unknown (n)) across the 228 cell types. The top three marker genes within each supercluster are displayed. These marker genes are highly expressed in their corresponding supercluster while showing low expression in others, validating the reliability of cell type annotation. Abbreviations in ring a: EGCN, Extratelencephalic-Glutamatergic-Cortical-Neuron; TTMC, Tracheobronchial-Tree-Multiciliated-Columnar; CMCABT, Central-Memory-CD4^{+}-Alpha-Beta-T.

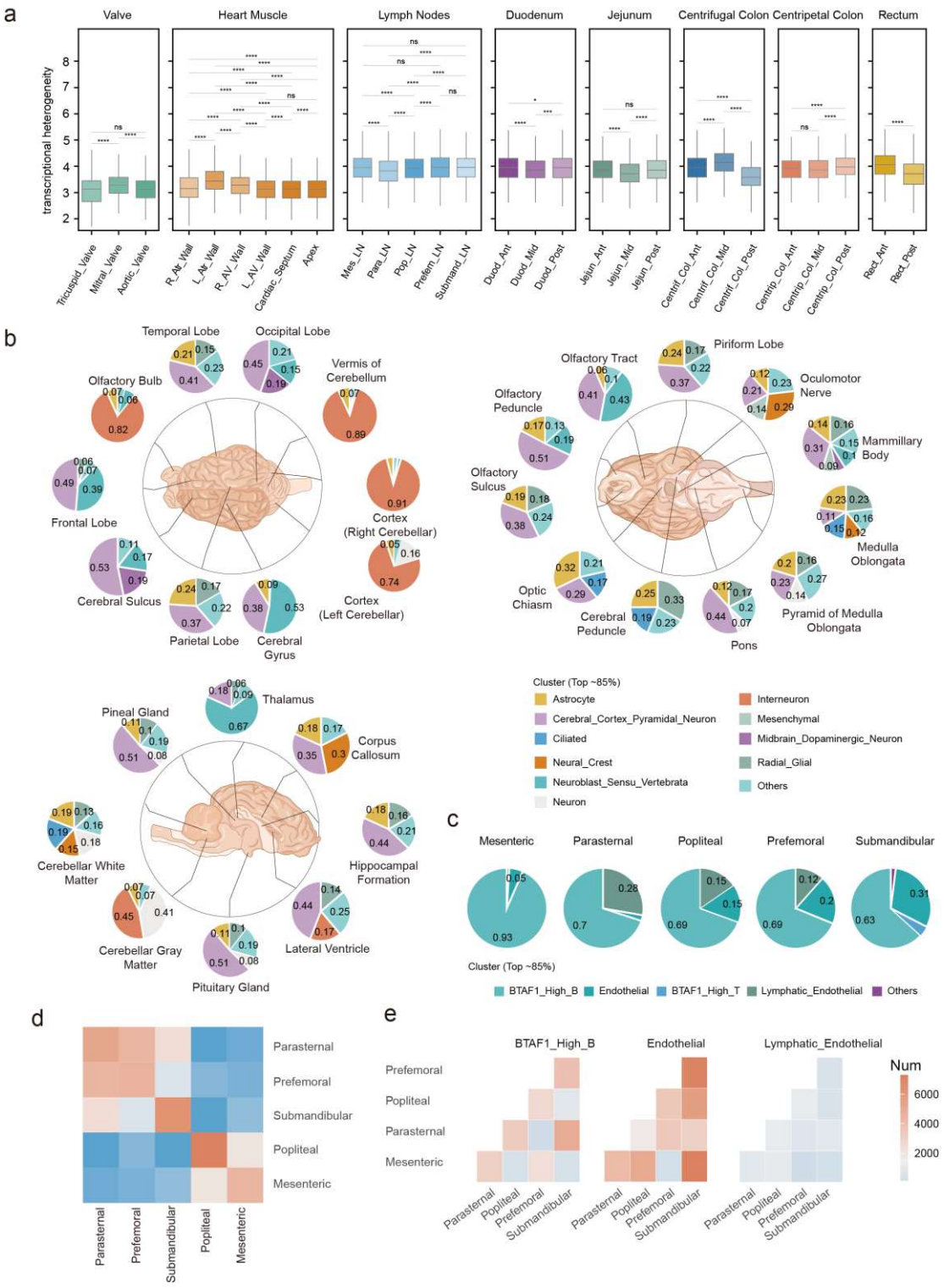


Figure 4 High regional heterogeneity of organs and tissues in Tibetan pigs.

(a) Suborgan-level transcriptional heterogeneity in TBP. This panel compares transcriptional heterogeneity of representative suborgans across multiple organ systems including cardiovascular, immune and digestive systems, such as heart valve, heart muscle, lymph nodes, duodenum, jejunum,

centrifugal colon, centripetal colon and rectum. Transcriptional heterogeneity is visualized by boxplots, where the X-axis represents distinct sub-organs and the Y-axis denotes pairwise Euclidean distance between cells, used as a metric for single-cell transcriptional heterogeneity. **(b)** Cell type composition of brain sub-organs. Cell type analysis was performed for cortical lobes (frontal, parietal, temporal, occipital, cingulate, piriform lobes, etc.), deep brain structures (hippocampus, thalamus, fornix, corpus callosum, etc.) and cerebellar compartments (vermis, cerebellar gray matter, cerebellar white matter, cerebellar hemisphere cortex, etc.). Pie charts show the relative proportions of major cell types (Astrocyte, pyramidal neuron, interneuron, midbrain dopaminergic neuron, radial glia, etc.) within each suborgan. Only cell types accounting for the top ~85% of cells are displayed, with the remaining pooled into "Others". **(c)** Cell type composition of lymph node sub-organs. Pie charts illustrate the relative proportions of major cell types (BTAF1-High_B, endothelial cell, BTAF1-High_T, lymphatic endothelial cell, etc.) in mesenteric, parasternal, popliteal, prefemoral and mandibular lymph nodes (top ~85% cell types shown). **(d)** Transcriptional similarity among lymph node sub-organs. The heatmap displays transcriptional similarity across five lymph nodes (mesenteric, parasternal, popliteal, prefemoral and mandibular) based on cell type composition and gene expression profiles. Color gradient from blue to red indicates increasing similarity, to evaluate global transcriptional resemblance between different lymph nodes. **(e)** Identification of differentially expressed genes among lymph node sub-organs. This heatmap shows the number of DEGs across three shared cell types (BTAF1-High_B, endothelial cells and lymphatic endothelial cells) in four lymphatic sub-organs (mandibular lymph node excluded). Color intensity scales with the number of DEGs; darker colors represent more DEGs (adjusted *p* < 0.01 and |logFC| > 1).

Abbreviations in panel (a):R_Atr_Wall（Wall_of_Right_Atrium）; L_Atr_Wall（Wall_of_Left_Atrium）; R_AV_Wall（Right_Atrioventricular_Wall）; L_AV_Wall（Left_Atrioventricular_Wall）; Mes_LN（Mesenteric_Lymph_Nodes）; Para_LN（Parasternal_Lymph_Nodes）; Pop_LN（Popliteal_Lymph_Nodes）; Prefem_LN（Prefemoral_Lymph_Nodes）; Submand_LN（Submandibular_Lymph_Nodes）; Duod_Ant（Duodenum(Anterior)）; Duod_Mid（Duodenum(Middle)）; Duod_Post（Duodenum(Posterior)）; Jejun_Ant（Jejunum(Anterior)）; Jejun_Mid（Jejunum(Middle)）; Jejun_Post（Jejunum(Posterior)）; Centrif_Col_Ant（Centrifugal_Colon(Anterior)）; Centrif_Col_Mid（Centrifugal_Colon(Middle)）; Centrif_Col_Post（Centrifugal_Colon(Posterior)）; Centrip_Col_Ant（Centripetal_Colon(Anterior)）; Centrip_Col_Mid（Centripetal_Colon(Middle)）; Centrip_Col_Post（Centripetal_Colon(Posterior)）; Rect_Ant（Rectum(Anterior)）; Rect_Post（Rectum(Posterior)）

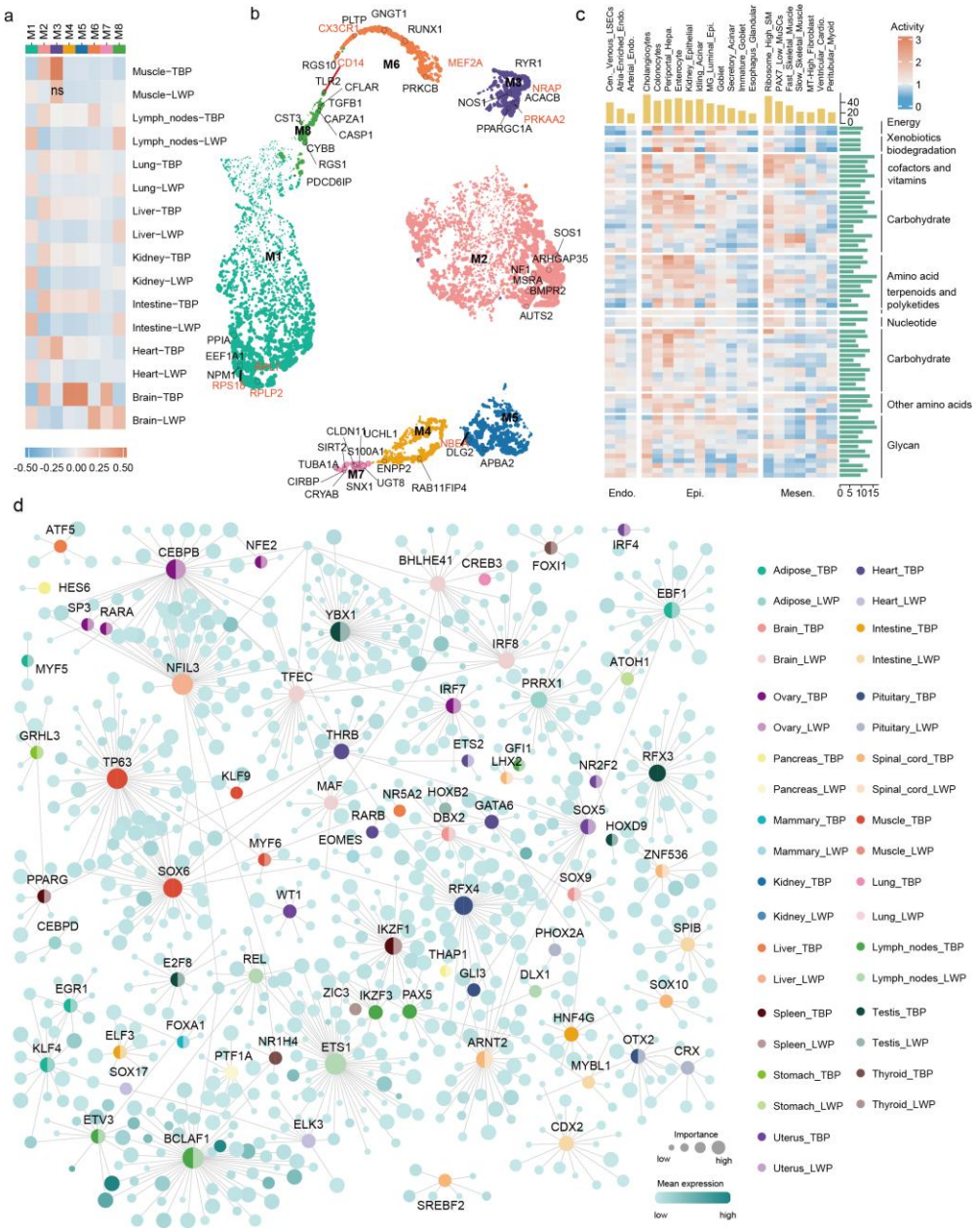


Figure 5 Comparative analysis of organ differences and regulatory network construction between TBP and LWP.

(a) Correlation scores and significance levels of each module across organs in two breeds (TBP and LWP). “ns” denotes non-significant correlation; blocks without “ns” are significant. Blue indicates negative correlation, red indicates positive correlation, and color intensity represents the magnitude of correlation. (b) hdWGCNA network. This plot shows associations between hub genes related to cold tolerance and hypoxia adaptation across modules. Black represents hypoxia adaptation, and

red represents cold tolerance. (c) Heatmap of metabolic pathway activity scores for Endothelial, Epithelial and Mesenchymal cell types. (d) Major organ regulatory network of TBP and LWP. Abbreviations in panel (c): Endo., Endothelial; Hepa., Hepatocytes; Epi., Epithelial; SM, Skeletal Muscle; Cardio., Cardiomyocytes.

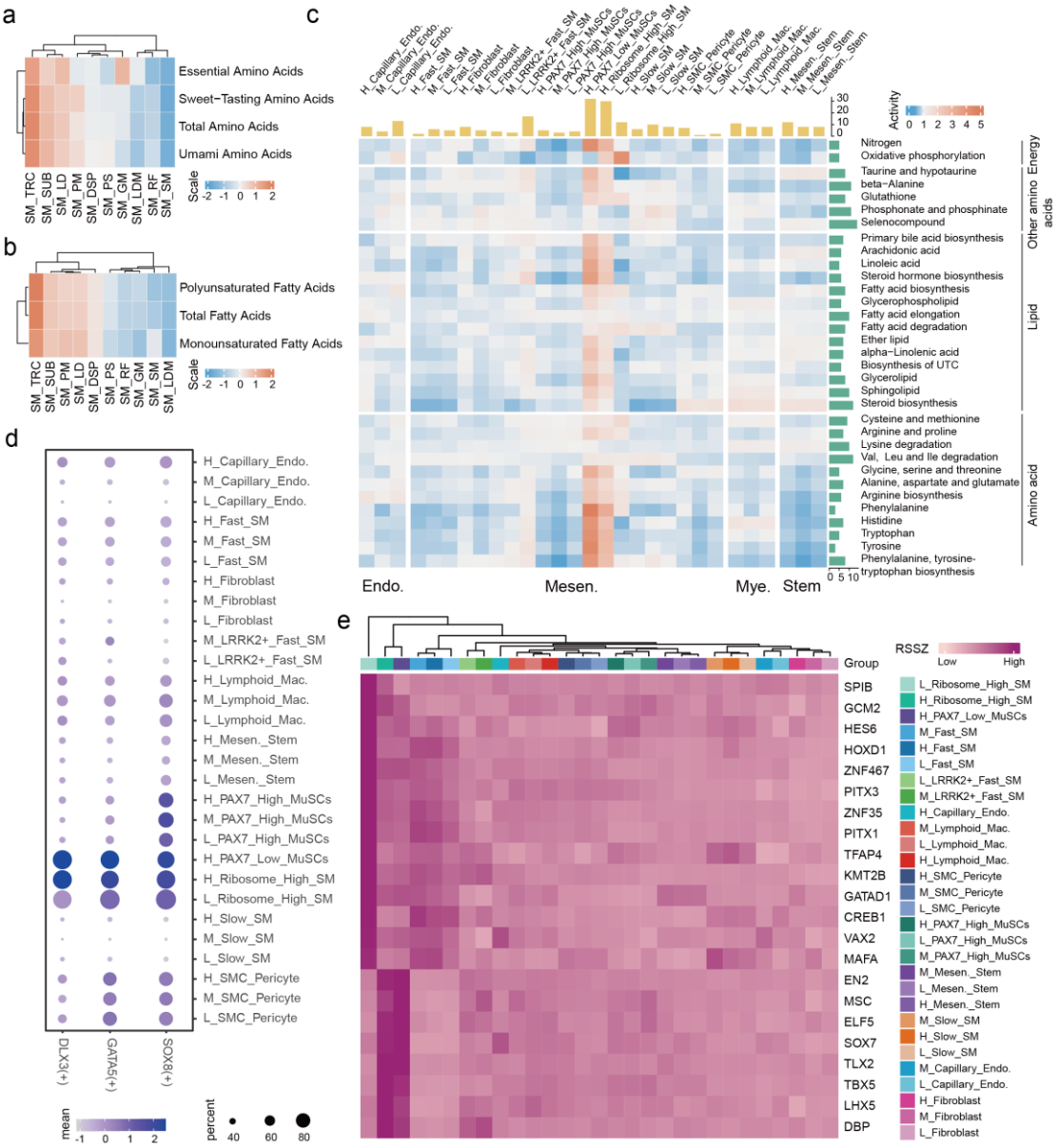


Figure 6 Analysis of muscle metabolic activity and regulatory network inference in Tibetan pigs (TBP).

(a) Heatmap showing the contents of amino acids across different samples. (b) Heatmap displaying the contents of fatty acids across different samples. (c) Heatmap of activity scores for amino acid, fatty acid, and energy metabolic pathways. (d) AUCell activity scores of metabolic pathways in different cell types. (e) Heatmap of cell-type-specific regulatory factors across distinct cell types. Abbreviations for panels (a–e): SM_TRC, Cervical_Part_of_Trapezius_Muscle; SM_SUB,

Subscapularis_Muscle; SM_LD, Latissimus Dorsi Muscle; SM_PM, Psoas Major Muscle; SM_DSP, Deltoid Muscle; SM_PS, Pectoralis Major Muscle; SM_GM, Gluteus Medius Muscle; SM_LDM, Longissimus Muscle; SM_RF, Rectus Femoris Muscle; SM_SM, Semimembranosus Muscle; Endo., Endothelial; Hepa., Hepatocytes; Epi., Epithelial; SM, Skeletal Muscle; Cardio., Cardiomyocytes; Mesen., Mesenchymal; Mye., Myeloid; Mac., Macrophage; UTC, unsaturated fatty acids.

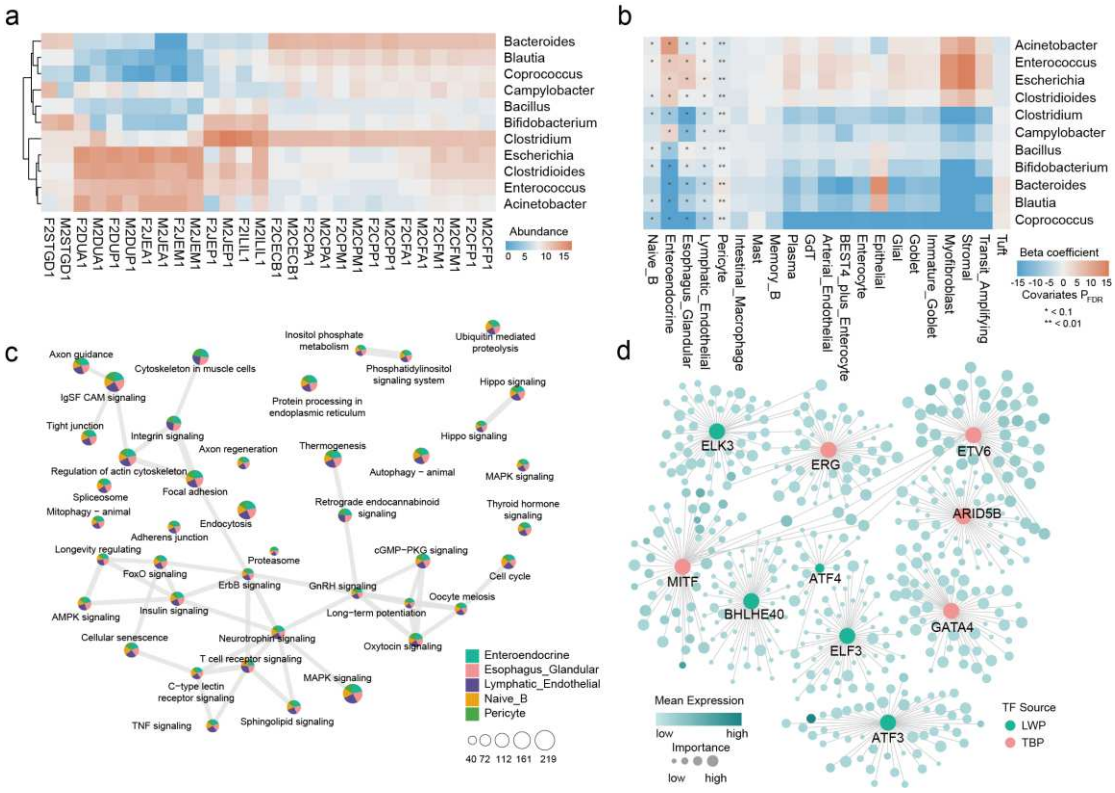


Figure 7 Gut microbiota enrichment analysis and regulatory network inference in Tibetan pigs (TBP).

(a) Heatmap showing the abundance of microbial taxa across digestive tract samples. (b) Heatmap illustrating associations between digestive tract cell types and filtered microbiota. (c) KEGG pathway enrichment network of differentially expressed genes between TBP and LWP in five cell types, including Pericyte, Lymphatic_Endothelial, Esophagus_Glandular, Enteroendocrine, and Naive_B. (d) Regulatory network of the top five specific transcription factors ranked by RSSZ values in TBP and LWP.

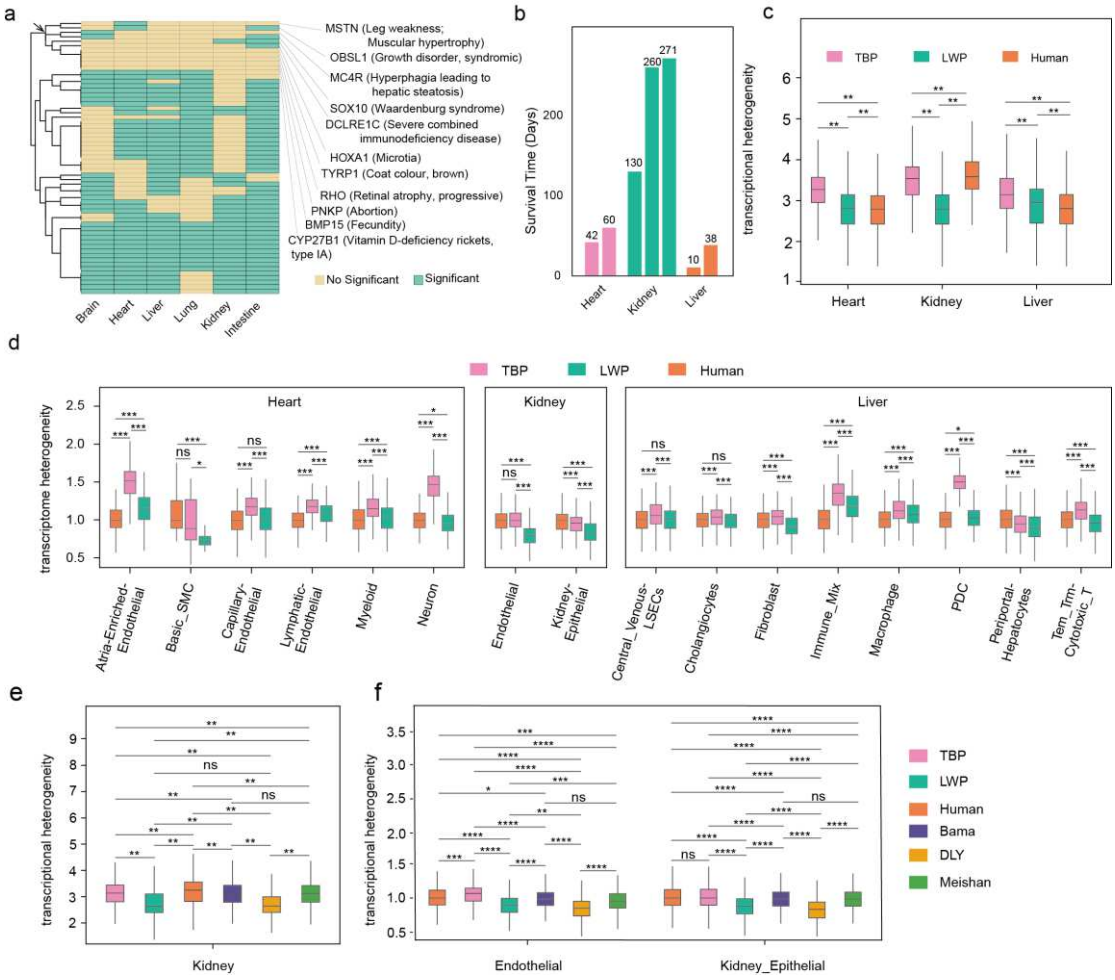
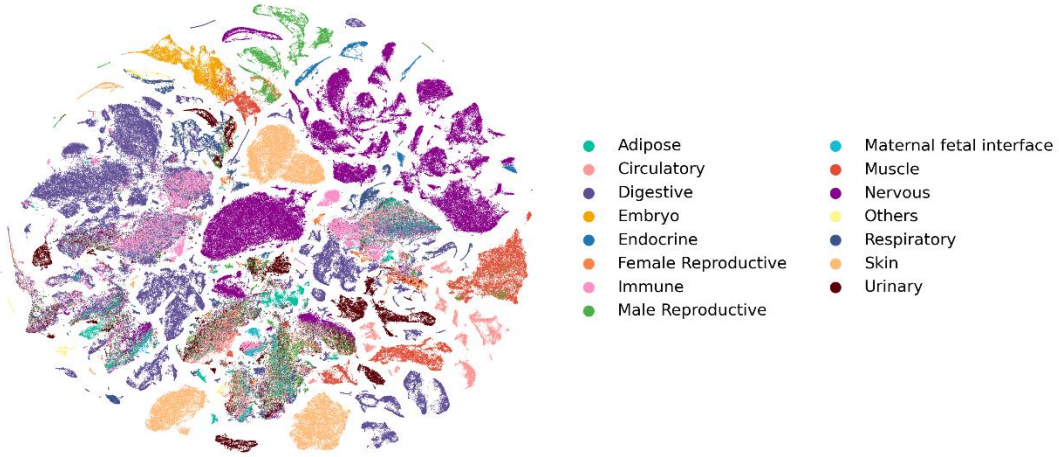


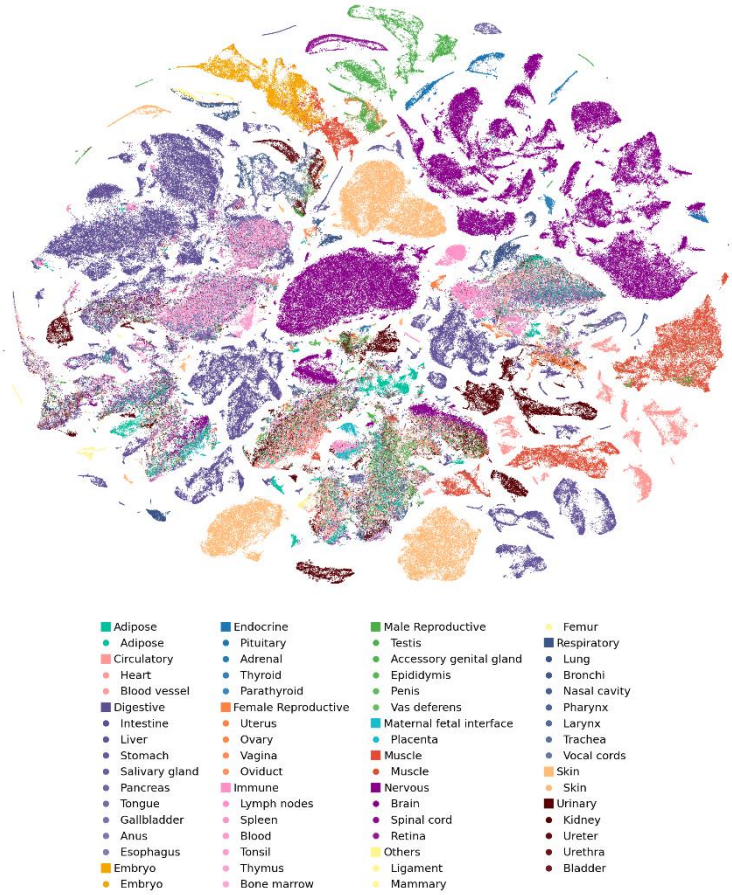
Figure 8 Cross-species comparisons provide insights into the application of Tibetan pigs in translational medicine.

(a) Heatmap showing the significance of differential expression of Mendelian disease genes across different organs between human and Tibetan pig (TBP). (b) Bar plot of pig-human xenotransplant survival days. (c) Boxplots of transcriptional heterogeneity in the heart, kidney and liver among TBP, LWP and human. (d) Boxplots of transcriptional heterogeneity across distinct cell types in TBP, LWP and human. (e) Boxplots of kidney transcriptional heterogeneity among TBP, LWP, human, Bama Xiang pig, DLY pig and Meishan pig. (f) Boxplots of transcriptional heterogeneity of shared cell types (endothelial and epithelial cells) in the kidney among TBP, human, Bama Xiang pig, DLY pig and Meishan pig. **: $P < 0.01$; ***: $P < 0.0001$

Extended Data Figures

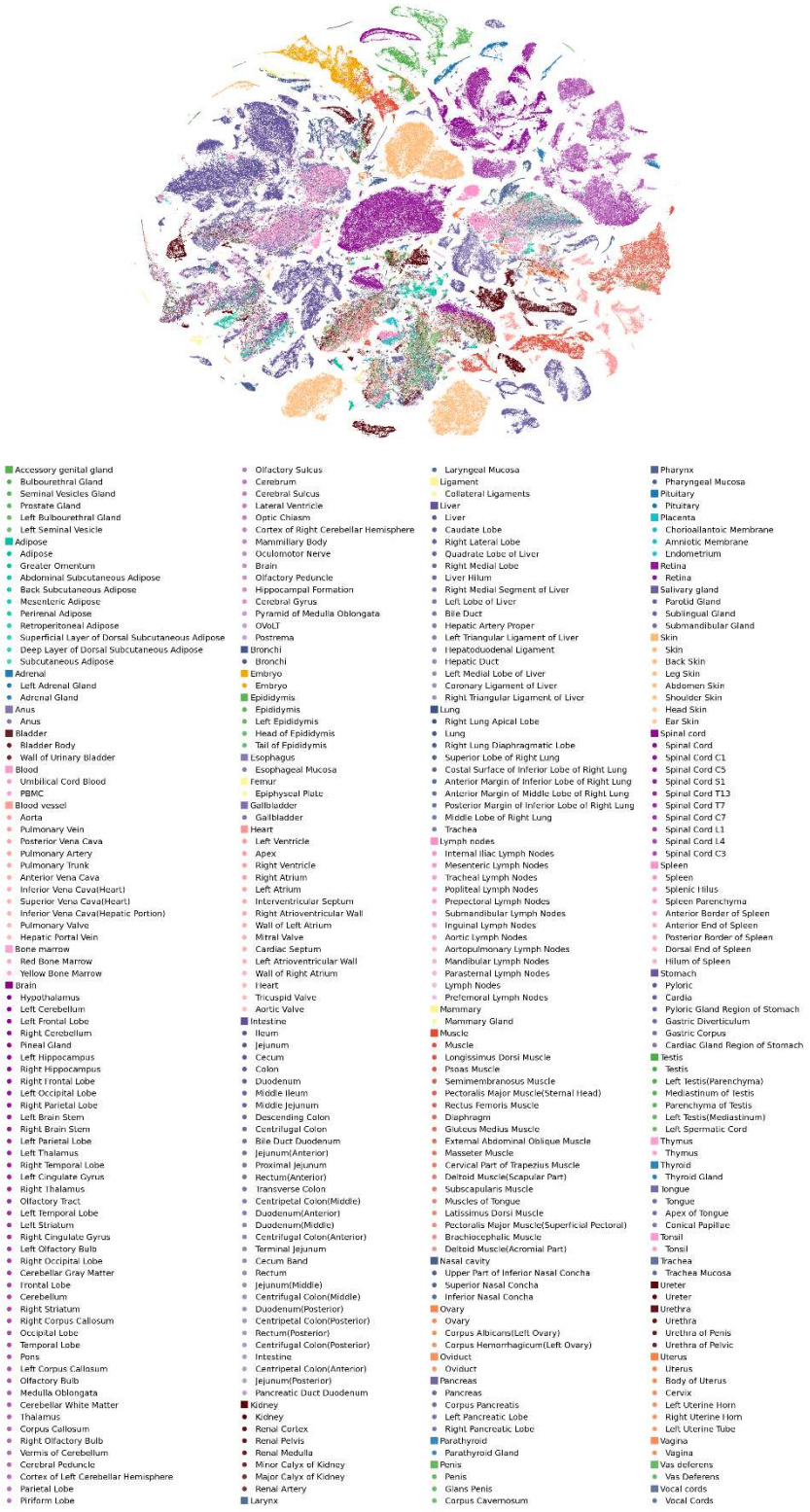


Extended Figure 1 *t-SNE* plot of single cells from porcine tissue.
Cells are colored and classified according to 15 tissue systems; "Others" denotes cells originating from systems not listed.



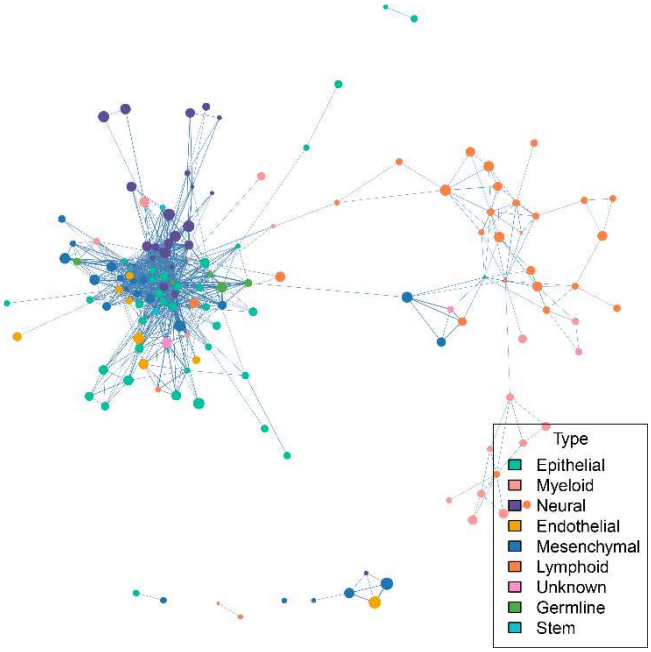
Extended Figure 2 *t-SNE* plot of single cells from porcine organs.

Cells are colored and classified according to organ origin, with a total of 52 organs included.



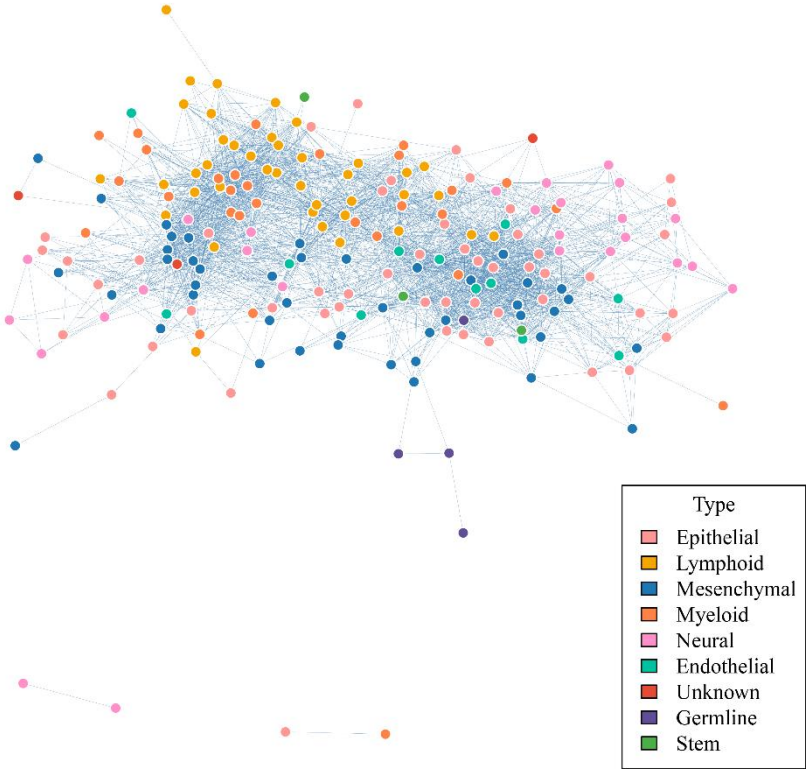
Extended Figure 3 *t-SNE* plot of single cells from porcine sub-organs.

Cells are colored and classified according to sub-organ origin, with a total of 296 sub-organs included.



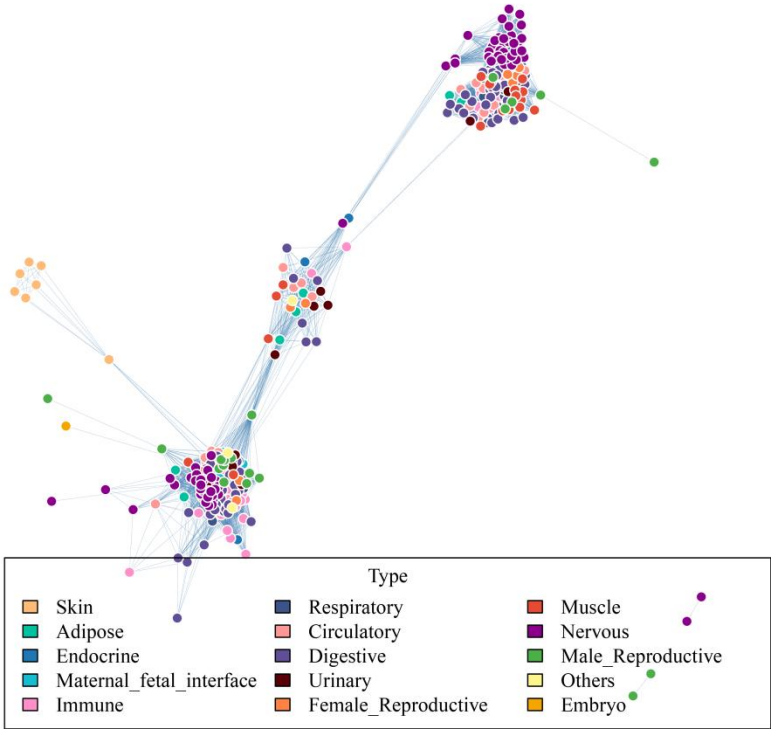
Extended Figure 4 Marker gene network diagram of cell types.

Each node represents a cell type; node size corresponds to the cell number of each cell type. Edges indicate shared marker genes between two cell types, and edge thickness reflects the number of overlapping marker genes. Node color denotes the assigned Supercluster.



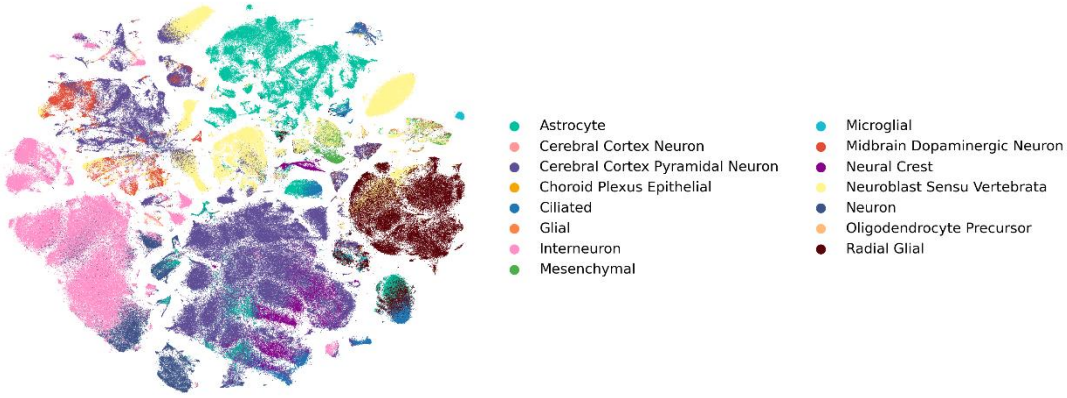
Extended Figure 5 Cell-type correlation network.

The average expression of each gene was calculated for all cells within each cell type/sub-organ to construct a gene × cell type/sub-organ average expression matrix. Spearman's rank correlation coefficients were computed to evaluate expression correlations across cell types/sub-organs. Each node represents a distinct cell type. Only correlations with $p \leqslant 0.01$ and $|r| \geqslant 0.9$ were retained.

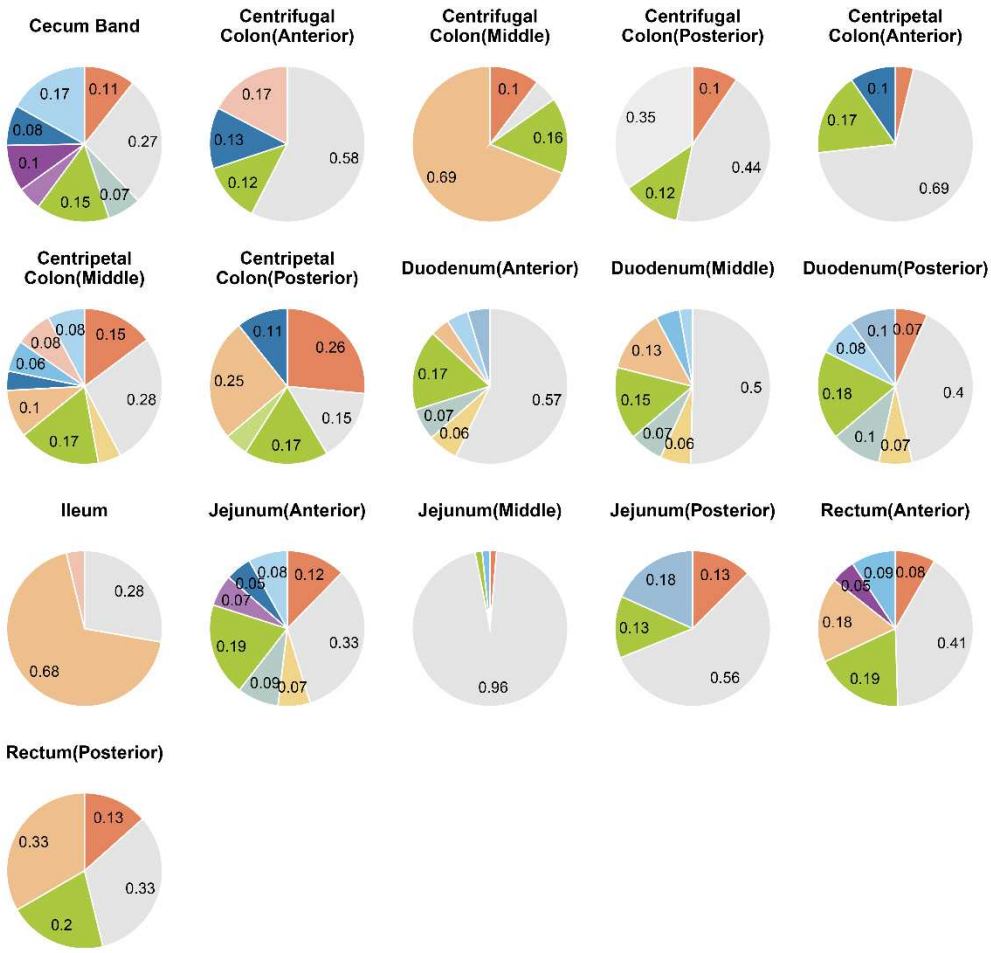


Extended Figure 6 Sub-organ correlation network.

Each node represents a distinct sub-organ. Correlations were calculated based on gene expression profiles, and only correlations with $p \leqslant 0.01$ and $|r| \geqslant 0.9$ were retained. The network covers 15 tissue systems.

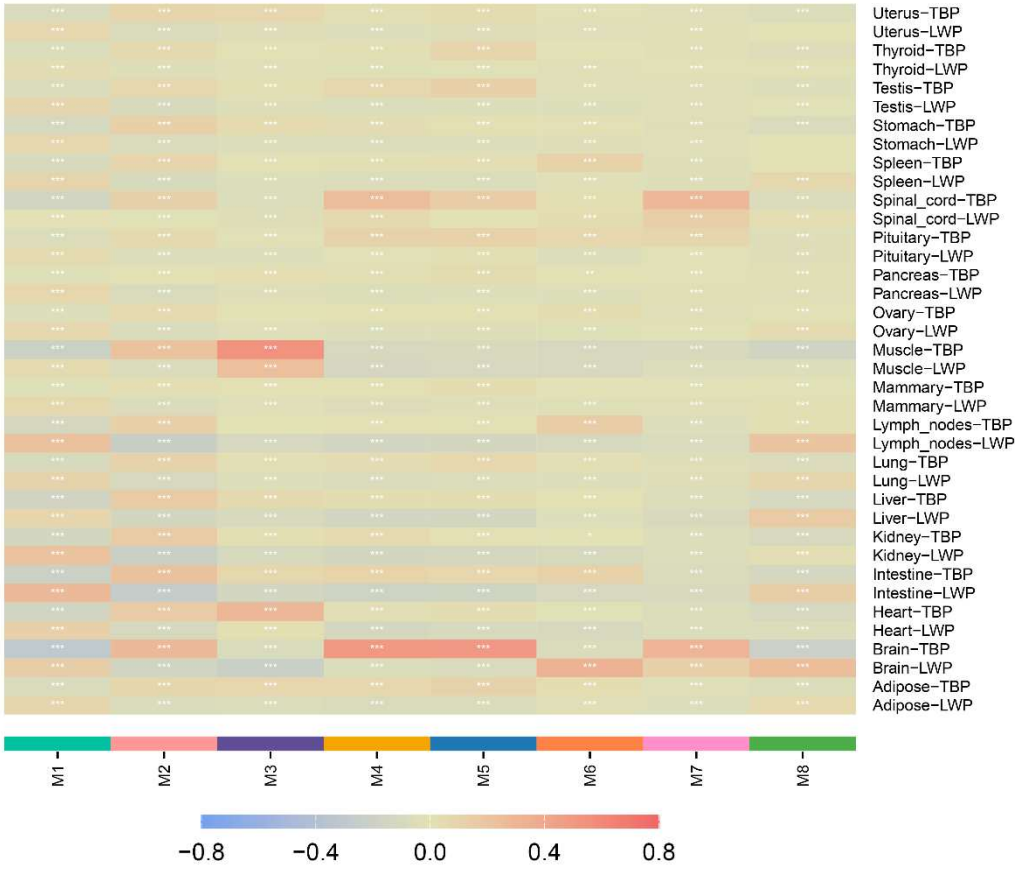


Extended Figure 7 *t-SNE* plot of annotated brain cells from Tibetan pigs.
Different colors denote distinct cell types, corresponding to 15 cell types in total.



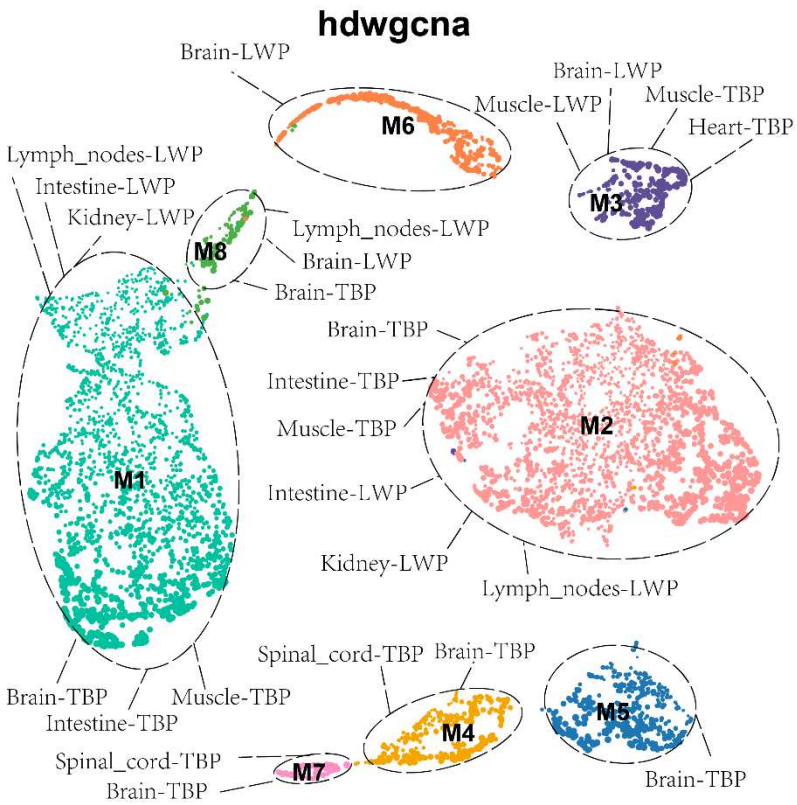
Extended Figure 8 Cell type composition of sub-organs in the intestine.

"Others" represents a pooled group of rare cell types. Briefly, cell types were sorted in descending order by cell count; only cell types accounting for the top 85% of total cells are displayed, and the remaining 15% are grouped into "Others".

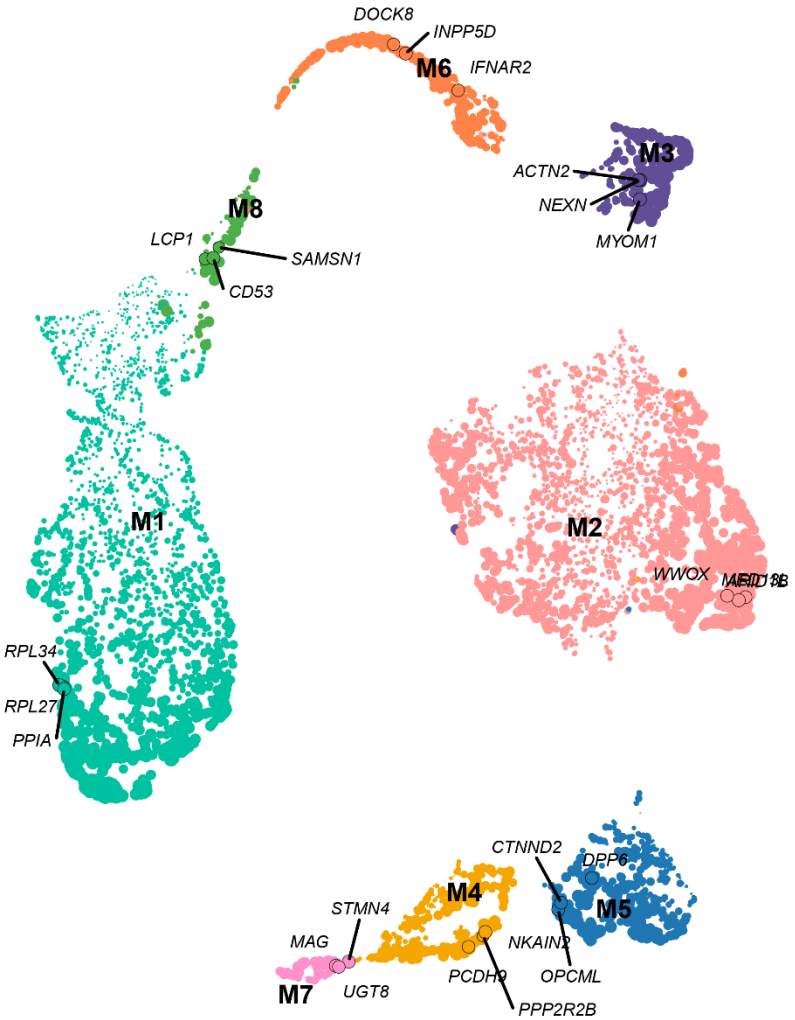


Extended Figure 9 Correlation scores and significance of each module in various organs between TBP and LWP.

Asterisks (*) denote significance, and color intensity indicates correlation strength. Blue indicates a negative correlation, and red indicates a positive correlation.

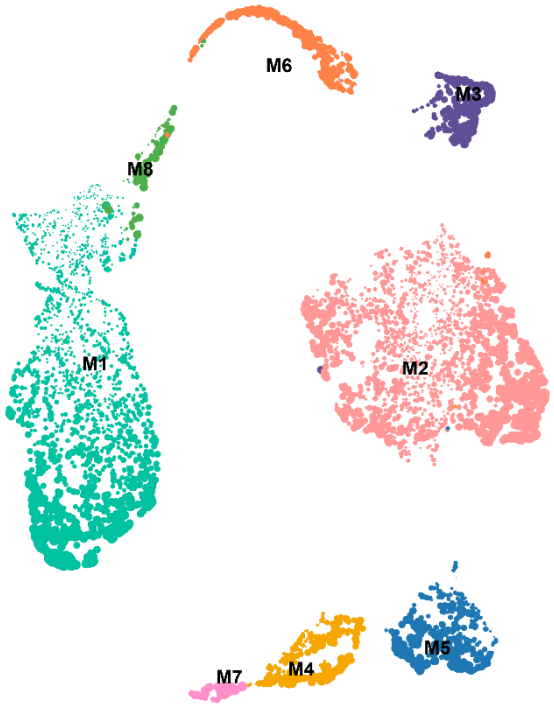


Extended Figure 10 Association plot between each module and organs.
Only organ-module pairs with absolute correlation values > 0.2 were retained.



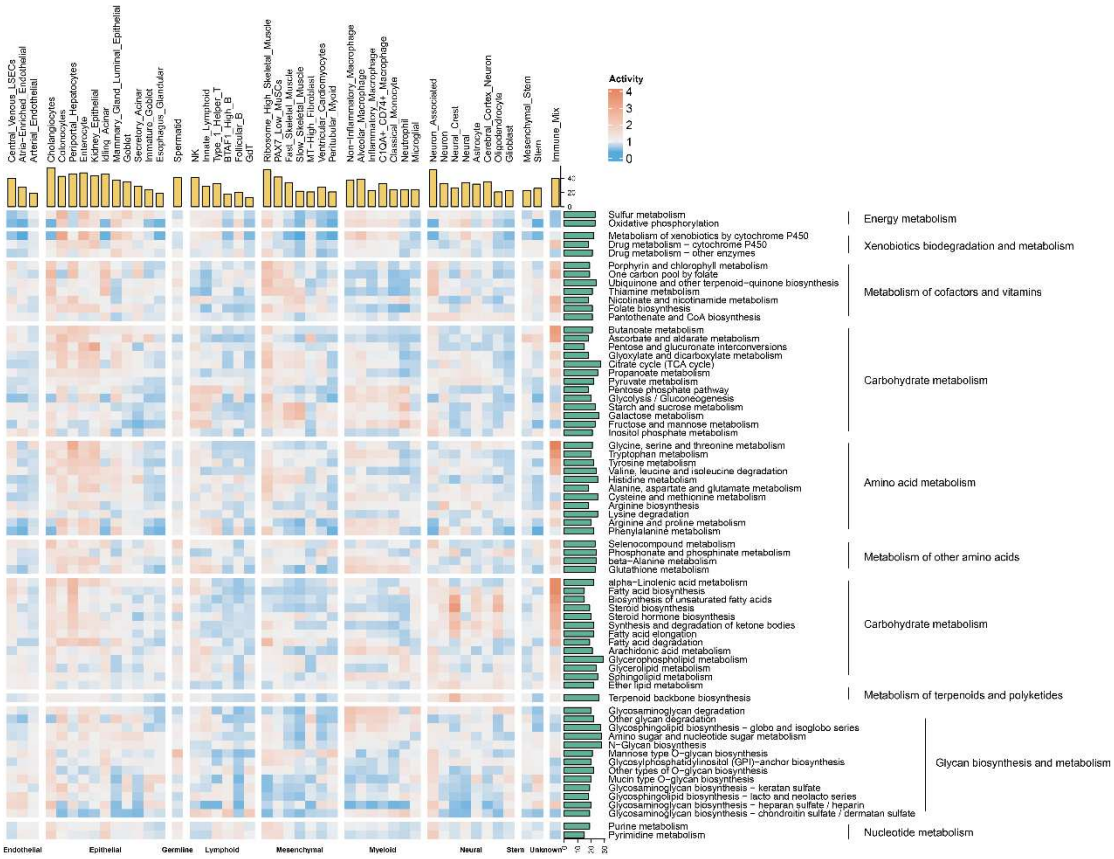
Extended Figure 11 The top three hub genes in each module (genes occupying core positions within the same gene co-expression module)

Different colors represent different gene modules. Within each module, circles indicate the top three hub genes.

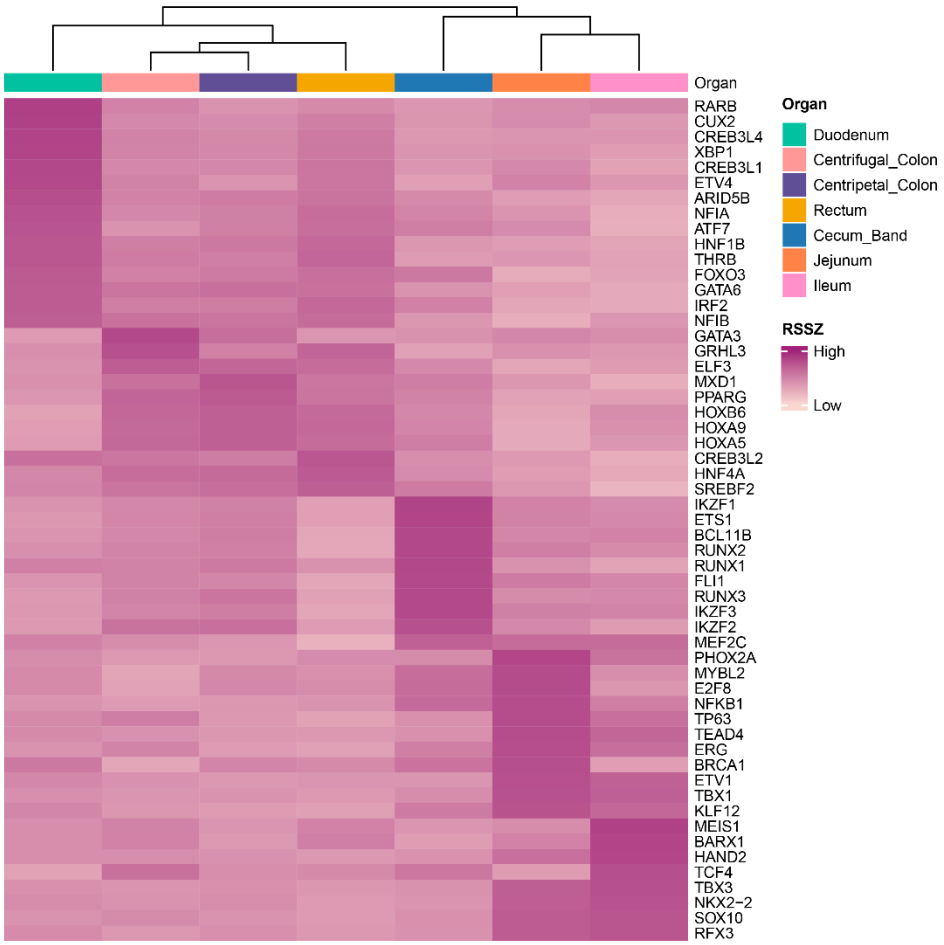


Extended Figure 12 Visualization of hub genes related to DNA damage repair within the module

Different colors represent different gene modules. Within each module, circles indicate the top three hub genes.

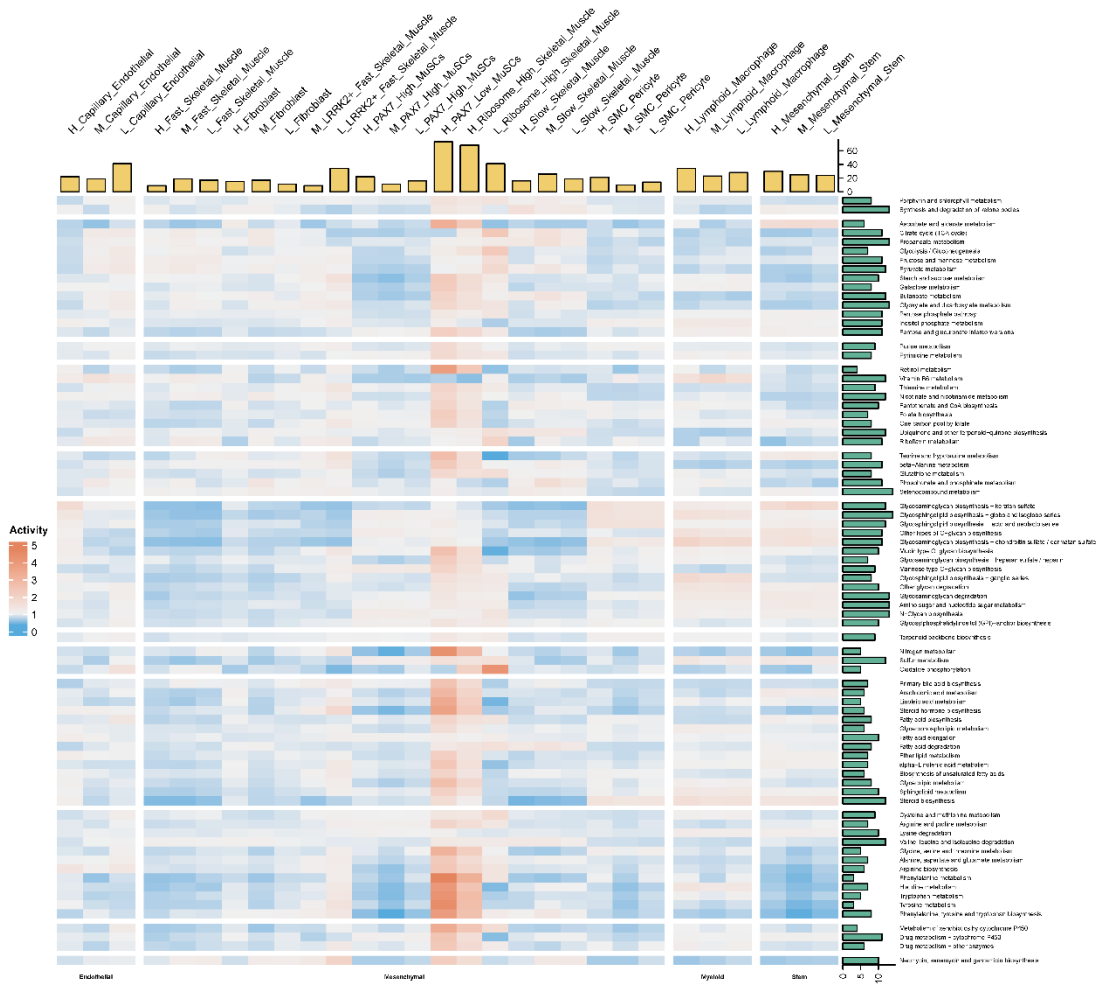


Extended Figure 13 Heatmap of metabolic pathway activity scores across cell types.
Expression levels of metabolism-related genes were calculated to evaluate differences in metabolic activity among cell types. The heatmap displays the top three cell types with the highest activity for each pathway. Columns are grouped by major cell classes. The top bar plot shows the number of significantly upregulated pathways in each cell type ($p < 0.01$, activity score > 1), and the right bar plot indicates the number of cell types with significantly high expression for each pathway.

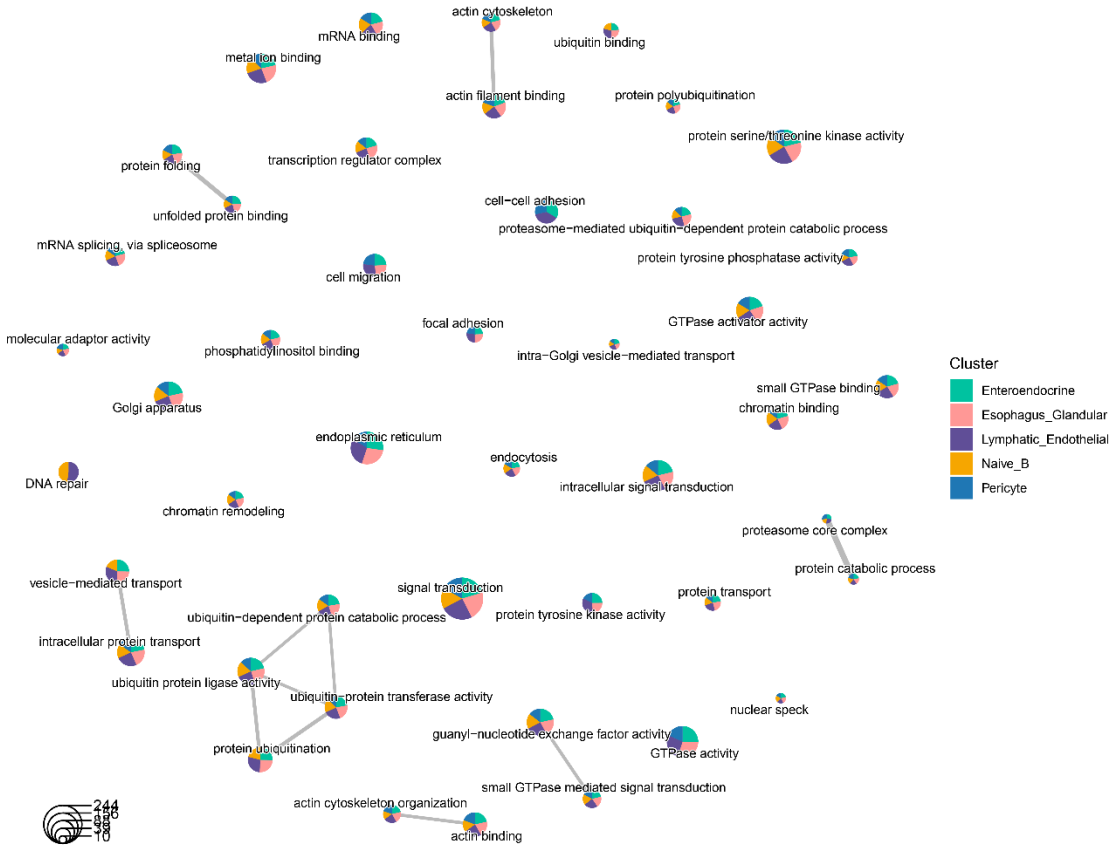


Extended Figure 14 Heatmap of RSSZ values for regulators across digestive tract organs in TBP.

A transcription factor was considered significant when its AUCell score > 0.1 and Z-normalized RSSZ > 1.0. Higher RSSZ values indicate stronger expression specificity of the corresponding regulator in the given organ; positive values represent high expression, while negative values represent low expression.



Extended Figure 15 Heatmap of activity scores for all metabolic pathways. Cell type prefixes H, L and M denote high, low and medium pork quality. Increased red intensity in the heatmap represents higher activity.



Extended Figure 16 GO feature enriches network diagrams. Based on intestinal single-cell transcriptomic data, DEGs were characterized within identical cell types between highland Tibetan pigs and lowland Landrace pigs, with subsequent GO enrichment and network analyses. Five cell types are distinguished by different colors.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables.xlsx
- SupplementaryFigures.pdf