Article

Sex differences orchestrated by androgens at single-cell resolution

Fei Li[1,10], Xudong Xing[2,10], Qiqi Jin[1,3,4,10], Xiang-Ming Wang[2,10], Pengfei Dai[1,3,10], Ming Han[1,3,10], Huili Shi[1,3], Ze Zhang[5], Xianlong Shao[1,3], Yunyi Peng[1,3], Yiqin Zhu[1], Jiayi Xu[6], Dan Li[7], Yu Chen[7], Wei Wu[8], Qiao Wang[8], Chen Yu[9 ✉], Luonan Chen[1,4,5 ✉], Fan Bai[2 ✉] & Dong Gao[1,9 ✉]

Sex differences in mammalian complex traits are prevalent and are intimately associated with androgens[1–7]. However, a molecular and cellular profile of sex differences and their modulation by androgens is still lacking. Here we constructed a high-dimensional single-cell transcriptomic atlas comprising over 2.3 million cells from 17 tissues in *Mus musculus* and explored the effects of sex and androgens on the molecular programs and cellular populations. In particular, we found that sex-biased immune gene expression and immune cell populations, such as group 2 innate lymphoid cells, were modulated by androgens. Integration with the UK Biobank dataset revealed potential cellular targets and risk gene enrichment in antigen presentation for sex-biased diseases. This study lays the groundwork for understanding the sex differences orchestrated by androgens and provides important evidence for targeting the androgen pathway as a broad therapeutic strategy for sex-biased diseases.

Sex differences are prevalent in human development, physiology and disease. Several important attempts have been made to explore transcriptomic sex differences using bulk RNA sequencing (RNA-seq)[2–6]. The investigation of physiological sex differences at single-cell resolution has the potential to elucidate the biological foundations of sex-biased diseases and to support the development of sex-informed diagnostic and therapeutic strategies.

Phenotypic sex differences are shaped by the combined effects of endogenous factors and exogenous environmental exposures[1]. Hormones, especially androgens and oestrogens, are thought to contribute substantially to sex-biased phenotypes. The actions of androgens are mediated by binding to androgen receptor (AR), a ligand-dependent nuclear transcription factor[8]. Drugs targeting AR as well as androgen synthesis to inhibit the androgen pathway, including enzalutamide and abiraterone, have brought great benefits to patients with prostate cancer[9,10]. More importantly, recent studies have demonstrated that female-biased responses to anti-PDL1 treatment in colon cancer and BRAF/MEK-targeted therapy in melanoma are associated with AR activity[11,12]. These studies highlight the need to systematically explore how androgens modulate sex differences and to investigate the application potential of targeting the androgen pathway for sex-biased diseases. Here we delineate the molecular and cellular programs that are modulated in response to androgens in the context of sex differences at single-cell resolution.

Cell composition responses to androgens

To systematically characterize sex differences and their modulation by androgens, we performed androgen deprivation via castration surgery (male castration (MC)) or sham operation (male sham (MS)) in adult male C57BL/6 mice and androgen supplementation via sustained release of dihydrotestosterone (DHT) (female DHT (FD)) or sham operation (female sham (FS)) in adult female C57BL/6 mice (Fig. 1a).

We observed that AR-positive (AR⁺) cells were prevalent across tissues in both male and female mice, and also validated this result by using a systemic AR-knockout mouse model[13] (Extended Data Fig. 1), forming a cellular basis for the function of androgens throughout the body. Following 4 weeks of treatment, androgen supplementation promoted body weight gain, whereas androgen deprivation had the reverse effect (Extended Data Fig. 2a). In addition to the prostate, the lacrimal gland and salivary gland exhibited increased weight after androgen supplementation and decreased weight after androgen deprivation (Extended Data Fig. 2b–g). By contrast, the weight and size of the thymus, spleen and stomach exhibited a negative association with androgens (Extended Data Fig. 2b–g). These results revealed the tissue-specific responses to androgen supplementation or deprivation.

To explore sex differences and their modulation by androgens at the cellular and molecular levels, we performed single-cell transcriptome profiling of tissues from MS, MC, FS and FD mice (Fig. 1a–c). In each

[1]Key Laboratory of Multi-Cell Systems, Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai, China. [2]Biomedical Pioneering Innovation Center (BIOPIC), Peking-Tsinghua Center for Life Sciences, Beijing Advanced Innovation Center for Genomics (ICG), School of Life Sciences, Peking University, Beijing, China. [3]University of Chinese Academy of Sciences, Beijing, China. [4]School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [5]Key Laboratory of Systems Health Science of Zhejiang Province, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China. [6]Shanghai Normal University, Shanghai, China. [7]Human Oncology and Pathogenesis Program, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [8]Key Laboratory of Medical Molecular Virology (MOE/NHC/CAMS), Shanghai Institute of Infectious Disease and Biosecurity, Shanghai Frontiers Science Center of Pathogenic Microbes and Infection, School of Basic Medical Sciences, Shanghai Medical College, Fudan University, Shanghai, China. [9]Institute of Cancer Research, Shenzhen Bay Laboratory, Shenzhen, China. [10]These authors contributed equally: Fei Li, Xudong Xing, Qiqi Jin, Xiang-Ming Wang, Pengfei Dai, Ming Han. ✉e-mail: yu@szbl.ac.cn; lnchen@sibcb.ac.cn; fbai@pku.edu.cn; dong.gao@sibcb.ac.cn
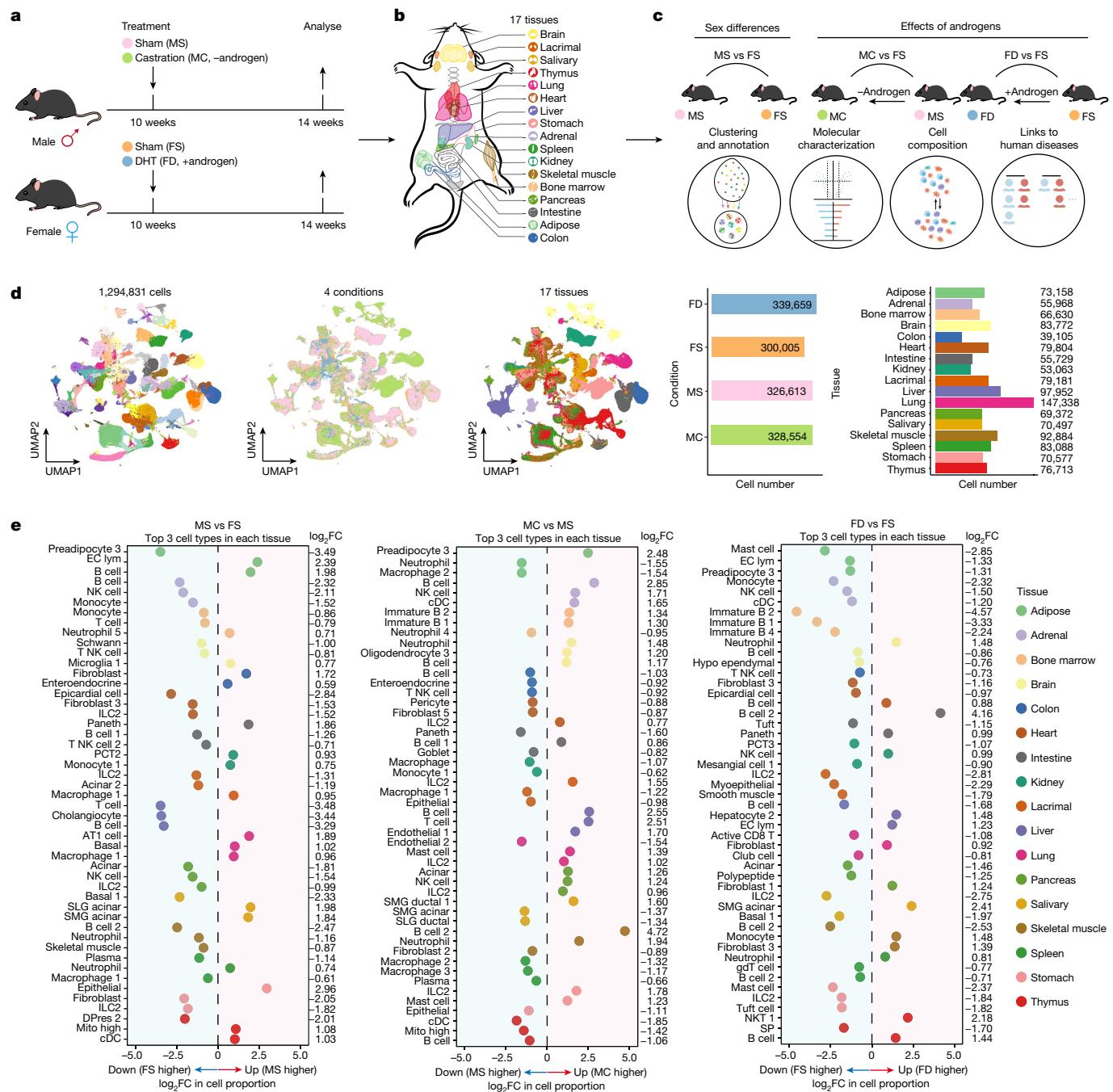
Fig. 1 | A single-cell transcriptomic atlas of sex differences and androgen effects. a, Schematic of the experimental design to explore sex differences and their modulation by androgens. **b**, Construction of the single-cell transcriptomic landscape based on 17 tissues. **c**, Schematic of three comparisons to explore sex differences (MS versus FS) and the effects of androgen deprivation (MC versus MS) or supplementation (FD versus FS) (top). **d**, UMAP plot showing cell clusters (left), condition information (middle) and tissue information (right). The bar plots illustrate the distribution of cell numbers for each condition and tissue. In each condition, each tissue contained at least three biological replicates (three mice). **e**, Lollipop plot showing the top three cell types (cell proportion of more than 1%) from each tissue with the largest proportional differences on three comparisons. cDC, conventional dendritic cell; EC lym, lymphatic endothelial cell; FC, fold change; Hypo ependymal, hypothalamic ependymal cell; DPres, double-positive thymocyte undergoing rearrangement; DPblast, blast double-positive thymocyte; gdT, γδ T cell; Mito, mitochondria signature-high; NK, natural killer; PCT, proximal convoluted tubule; plasma, terminally differentiated plasma B cell; SLG, sublingual gland; SMG, submandibular gland; SP, single-positive thymocyte.

condition, each tissue contained at least three biological replicates (three mice) (Supplementary Table 1). A total of 1,294,831 qualified cells with a mean of 2,070 detected genes across 17 tissues were obtained and subsequently annotated (Fig. 1d, Supplementary Table 1 and Supplementary Fig. 1). Both expected tissue-common and tissue-specific cell types were identified (Extended Data Fig. 3). We also developed a

web tool (https://casadbtools.com/) to provide customized visualization and analyses of our single-cell RNA-seq (scRNA-seq) dataset.

We compared the cellular proportions between MS and FS mice (MS versus FS, sex differences), MC and MS mice (MC versus MS, the effects of androgen deprivation) and FD and FS mice (FD versus FS, the effects of androgen supplementation) (Fig. 1c and Supplementary
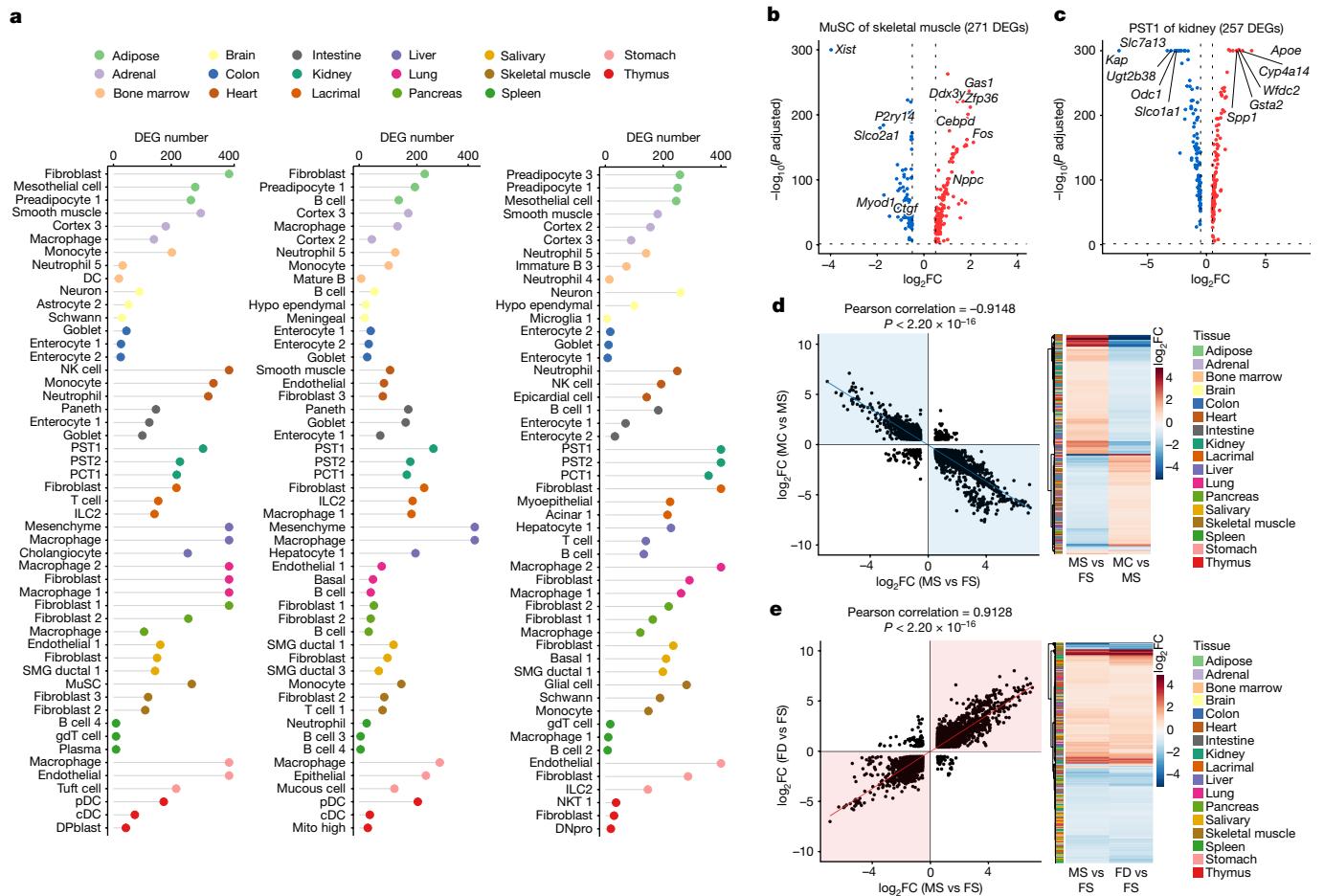
Fig. 2 | Molecular characterization of sex differences and androgen effects. a, Lollipop plot showing the top three cell types from each tissue with the largest numbers of DEGs based on the comparisons of MS versus FS (left), MC versus MS (middle) and FD versus FS (right). The upper *x* axis limit was set to 400. DNpro, double-negative progenitor thymocyte; MuSC, muscle stem cell; pDC, plasmacytoid dendritic cell; PST, proximal straight tubule. **b,c**, Volcano plot showing the DEGs in MuSCs of skeletal muscle between MS and FS (**b**) and the DEGs in PST1 cells of the kidney between MC and MS (**c**). The top five upregulated and downregulated DEGs, and the sex chromosome genes *Xist* and *Ddx3y* are labelled. DEGs were defined by Wilcoxon rank sum test, and *P* values

were adjusted by the Bonferroni correction using the FindMarkers function implanted in Seurat. The horizontal and vertical dashed lines represent *P* adjusted = 0.05 and |log₂FC = 0.5|, respectively. **d,e**, Scatter plot showing the correlation of log₂FC in DEGs between sex differences and MC versus MS (**d**) or between sex differences and FD versus FS (**e**) based on all the cell types from 17 tissues (left). Heatmap denoting the log₂FC in the overlapping DEGs from a single cell type in the comparisons of MS versus FS and MC versus MS (**d**) or MS versus FS and FD versus FS (**e**); DEGs are coloured according to their tissue of origin (right). In **d,e**, *P* values are shown; the test statistic is based on Pearson's product moment correlation coefficient and follows a *t* distribution.

Table 2). We selected the top three cell types with the largest proportional changes (absolute value, increase or decrease, false discovery rate (FDR) < 0.05) in each tissue in the MS versus FS, MC versus MS and FD versus FS comparisons (Fig. 1e). The top three cell types with the largest proportional differences between sexes tended to vary across tissues (Fig. 1e). For example, in the liver, the proportions of T cells and B cells were markedly lower in male mice than in female mice; in the adipose, lymphatic endothelial cells exhibited a higher proportion in male mice (Fig. 1e).

In general, sex differences (MS versus FS) in cell proportions were significantly positively correlated with the effects of androgen supplementation (FD versus FS) (Pearson correlation = 0.4625, *P* = 7.14 × 10⁻¹⁴) but significantly negatively correlated with the effects of androgen deprivation (MC versus MS) (Pearson correlation = −0.6925, *P* < 2.20 × 10⁻¹⁶) (Extended Data Fig. 4a–c), as exemplified by the salivary gland (Extended Data Fig. 4d,e). At the cell-type level, in the salivary gland, the proportion of submandibular gland acinar cells, which was male biased, was decreased by androgen deprivation in male mice and effectively increased by androgen supplementation in female mice (Extended Data Fig. 4f–h). These findings demonstrated

that androgens could serve as one of the major contributors to the sex differences in tissue cell composition.

Molecular responses to androgens

We identified differentially expressed genes (DEGs) (*P* adjusted < 0.05 and |log₂ fold change (FC)| > 0.5) in each cell type across 17 tissues between MS and FS mice (hereafter referred to as sex-biased DEGs) (Supplementary Table 3). The DEGs primarily reflected biological differences rather than the biases introduced by cell numbers or cell proportions (Supplementary Fig. 2). We determined the top three sex-biased cell types for each tissue by ranking the cell types according to the number of DEGs (Fig. 2a and Supplementary Fig. 3a). For example, in muscle stem cells of the skeletal muscle, *Xist* and *Ddx3y*, two sex chromosome genes, were sex-biased DEGs as expected (Fig. 2b). We next performed biological pathway enrichment analyses of sex-biased DEGs in each cell type, considering the direction of the sex effect (Gene Ontology: biological process, *P* < 0.01 and *q* < 0.01; Supplementary Table 4). These biological pathways were further ranked according to the number of shared cell types. Some expected pathways shared by multiple cell

Article

types or tissues were observed: steroid metabolic processes driven by male-biased DEGs (Extended Data Fig. 5a) and inactivation of the X chromosome in which female-biased DEGs were functionally enriched (Extended Data Fig. 5b).

We next identified the DEGs induced by androgen deprivation between MC and MS mice or by androgen supplementation between FD and FS mice (Fig. 2a, Supplementary Table 3 and Supplementary Fig. 3b,c). Androgen deprivation significantly downregulated *Kap* expression in the proximal straight tubule (PST1) cells of the kidney (Fig. 2c), consistent with a previous study[14]. By determining which cell types are susceptible to androgens in each tissue, we observed that in the kidney, PST1 cells not only displayed the largest number of sex-biased DEGs but also exhibited the strongest response to both androgen deprivation and supplementation (Fig. 2a). Through pathway enrichment analyses, we evaluated the cross-cell-type effects of androgens on biological functions and found that the upregulated DEGs elicited by androgen deprivation were functionally associated with T cell activation, whereas the downregulated DEGs were involved in neutrophil chemotaxis and leukocyte migration (Extended Data Fig. 5c,d). Consistently, androgen supplementation exerted the opposite effects on these biological pathways (Extended Data Fig. 5e,f). Further examination at the gene level revealed *S100a9* and *S100a8* in the 'neutrophil chemotaxis' pathway (Extended Data Fig. 5g) and *Egr1* in the 'T cell activation' pathway (Extended Data Fig. 5h), which were regulated by androgens in multiple cell types.

The magnitude of the expression differences among androgen deprivation-induced DEGs was significantly negatively correlated with the magnitude of the sex differences in the expression of these genes (Pearson correlation = −0.9148, $P < 2.20 × 10^{-16}$) (Fig. 2d). Consistently, the magnitude of the expression differences among androgen supplementation-induced DEGs was significantly positively correlated with the magnitude of the sex differences in the expression of these genes (Pearson correlation = 0.9128, $P < 2.20 × 10^{-16}$) (Fig. 2e). Performing similar analyses exclusively on the DEGs located on autosomes yielded consistent results (Extended Data Fig. 5i,j).

Androgens potentially function in biological processes by influencing other hormones. To evaluate this, we performed DHT supplementation (female ovariectomy DHT (FOD)) or vehicle supplementation (female ovariectomy vehicle (FO)) in ovariectomized female mice (Extended Data Fig. 6a). We then performed single-cell transcriptome profiling of the 17 tissues of FO and FOD mice, and finally obtained a total of 701,925 qualified cells (346,546 cells for FO and 355,379 cells for FOD; Supplementary Table 5). Cell-type annotation for each tissue was performed by leveraging the high-quality reference data built from the four conditions of FD, FS, MS and MC. We demonstrated that androgens exhibited consistent effects on the tissue weight (Extended Data Fig. 6b,c), cell composition (Extended Data Fig. 6d) and gene expression (Extended Data Fig. 6e) between intact and ovariectomized female mice. Castration may also have other effects, although it will indeed remove androgens. We added DHT back to castrated male mice (male castration DHT supplementation (MCD)) (Extended Data Fig. 6f) and performed scRNA-seq on the 17 tissues of MCD mice, resulting in a total of 340,675 qualified cells (Supplementary Table 5). Despite some heterogeneity among tissues, from the perspectives of tissue weight (Extended Data Fig. 6g,h), cell composition (Extended Data Fig. 6i) and gene expression (Extended Data Fig. 6j), the addition of androgen back to castrated male mice generally reversed the effects of castration.

Androgen-associated sex-biased molecules

Enhanced therapeutic effects of immune therapy and targeted therapy induced by androgen pathway inhibition have been identified[11,12]. Thus, the sex-biased DEGs regulated by androgens may represent potential biomarkers and therapeutic targets for androgen pathway inhibition to treat sex-biased diseases. We defined androgen-associated sex-biased

DEGs (AASB-DEGs) for each cell type across the 17 tissues, comprising positive AASB-DEGs (the expression levels of which were male biased and positively regulated by androgens) and negative AASB-DEGs (the expression levels of which were female biased and negatively regulated by androgens) (Fig. 3a, Extended Data Fig. 7a and Supplementary Table 6). For example, in adipose adipocytes, there were 15 negative AASB-DEGs and 6 positive AASB-DEGs (Fig. 3b). Of note, *Fasn* is a negative AASB-DEG, and its expression was female biased and upregulated by androgen deprivation in male mice and downregulated by androgen supplementation in female mice (Fig. 3b,c), consistent with previous studies[15].

The number of AASB-DEGs varied across cell types. Among these cell types, PST1 cells in the kidney not only showed the largest number of AASB-DEGs (182) but also exhibited a relatively higher percentage of AASB-DEGs among the sex-biased DEGs (58.7%) (Fig. 3d). Moreover, *Ar* expression significantly positively correlated with the numbers of AASB-DEGs (Pearson correlation = 0.3096, $P = 7.55 × 10^{-5}$) (Extended Data Fig. 7b,c). These findings demonstrated that androgens sensitively modulated sex differences in these cell types.

Although AASB-DEGs exhibited a skewed distribution across tissues, we found 93 positive and 81 negative AASB-DEGs that were shared by at least two tissues (Fig. 3e,f). These AASB-DEGs included cell growth and differentiation factors, such as *Egfr*, *Fos* and *Fosb*, and immune regulators, such as *Ccl3*, *Il33* and *Ccl4* (Fig. 3e,f). Some expected AASB-DEGs were also observed, such as *Ace* and *Hsd11b1* (Fig. 3e and Extended Data Fig. 7d,e), consistent with previous studies[2,16]. We then performed pathway enrichment analyses of AASB-DEGs in each cell type (Supplementary Table 6). We found 50 and 145 shared biological pathways in which positive or negative AASB-DEGs, respectively, were functionally involved. The 50 positive biological pathways included 'drug catabolic process', 'hormone catabolic process' and 'negative regulation of peptidase activity' pathways, and the 145 negative biological pathways included 'fatty acid metabolic process', 'purine nucleotide metabolic process' and 'acyl-CoA metabolic process' pathways (Extended Data Fig. 7f). These AASB-DEGs might represent pivotal biomarkers and attractive therapeutic targets for precision medicine targeting the androgen pathway.

Androgen-associated immune responses

To examine the effects of sex and androgens on immune cell composition, we extracted all the immune cells in our atlas and further reconstructed the immune compartment, ultimately resulting in nine major cell types (Fig. 4a,b and Extended Data Fig. 8a). We observed that sex differences in immune cell populations were significantly positively correlated with androgen supplementation-induced changes (Pearson correlation = 0.6399, $P = 4.22 × 10^{-16}$) but significantly negatively correlated with androgen deprivation-induced changes (Pearson correlation = −0.7372, $P < 2.20 × 10^{-16}$) (Fig. 4c and Extended Data Fig. 8b–d).

We next identified the sex-biased immune cell types that exhibited coordinated responses to androgen deprivation and supplementation (AASB-immune cell types) and found that the proportions of 9 cell types were positively associated with androgens (positive AASB-immune cell types), and that those of 13 cell types were negatively associated with androgens (negative AASB-immune cell types) (Fig. 4d). Seven of nine positive AASB-immune cell types (7 of 9) were myeloid cells, whereas all the 13 negative AASB-immune cell types (13 of 13) were lymphoid cells (Fig. 4d). Neutrophils were a positive AASB-immune cell type in the kidney, lacrimal gland and liver (Fig. 4d), consistent with the neutropenia observed in AR-knockout mice and the castration-induced restoration of neutrophil counts[17].

Among these AASB-immune cell types, we noticed that group 2 innate lymphoid cells (ILC2s) highly expressing *Gata3*, *Areg* and *Rora* (Fig. 4e), which were identified in the 2010s[18–23], were a negative AASB-immune cell type shared by seven tissues: the heart, lacrimal
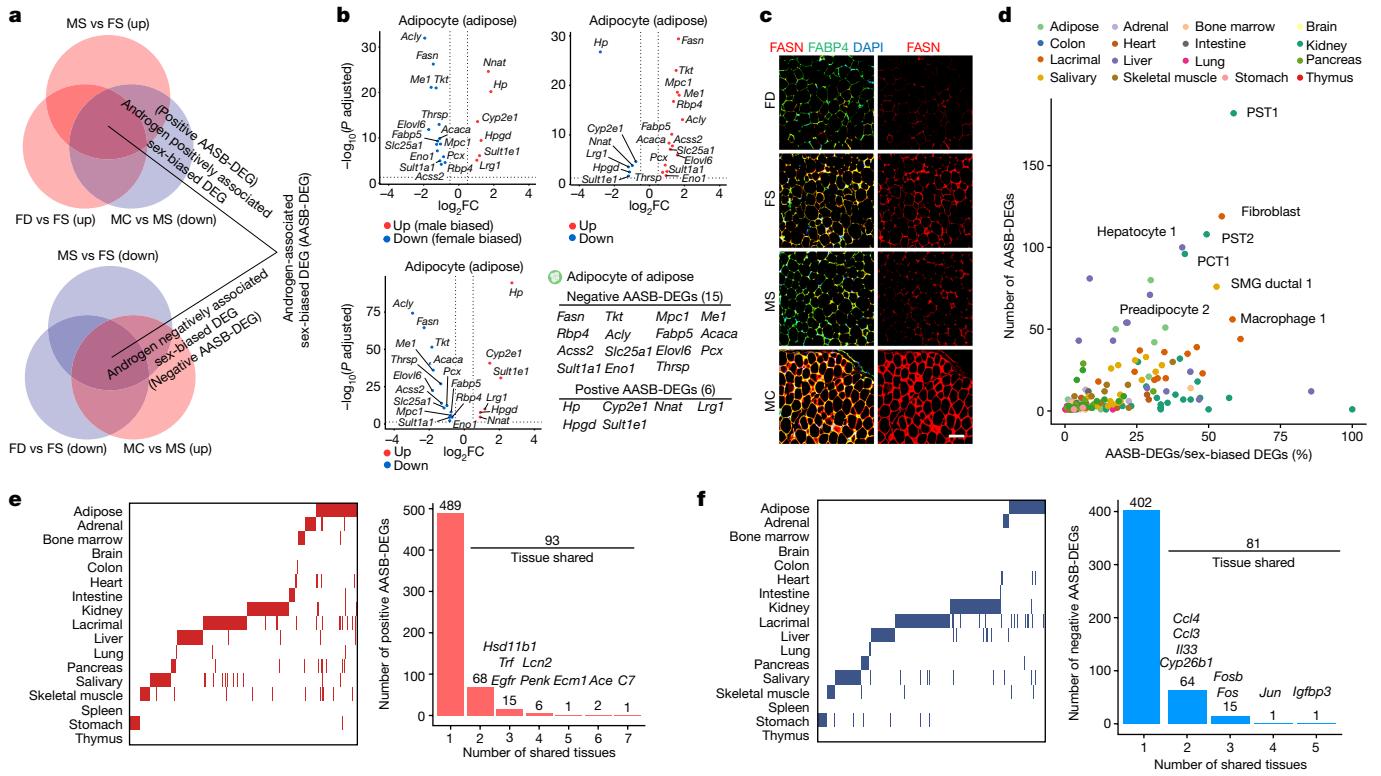
Fig. 3 | Identification of AASB gene expression. a, Venn diagram showing the identification of AASB-DEGs, comprising positive AASB-DEGs and negative AASB-DEGs. **b**, Volcano plot showing the 21 AASB-DEGs in adipocytes from adipose tissue based on the comparisons of MS versus FS (top left), MC versus MS (top right) and FD versus FS (bottom left). DEGs were defined by Wilcoxon rank sum test and P values were adjusted by the Bonferroni correction using the FindMarkers function implanted in Seurat. The horizontal and vertical dashed lines represent P adjusted = 0.05 and $|\log_2 FC| = 0.5$, respectively. **c**, Immunofluorescence staining for FASN, FABP4 (an adipocyte marker) and

DAPI in adipose adipocytes. The scale bar represents 50 μm. Each experiment was repeated independently with similar results for three times. **d**, Dot plot showing the number of AASB-DEGs (y axis) and the percentage of AASB-DEGs among the sex-biased DEGs (x axis) in each cell type across tissues. The top cell types with the largest numbers of AASB-DEGs are labelled. **e,f**, Heatmap showing the positive (**e**) and negative (**f**) AASB-DEGs across tissues (left). The bar plot denoting the categorization of the positive (**e**) and negative (**f**) AASB-DEGs according to the number of shared tissues (right).

gland, liver, pancreas, salivary gland, spleen and stomach (Fig. 4f,g). The proportions of ILC2s were female biased in multiple tissues, which increased after androgen deprivation and decreased after androgen supplementation (Fig. 4g). The female-biased and androgen-regulated ILC2 proportion was consistent with previous studies[24–26]. ILC2s possessed the highest expression levels of Ar among the nine major immune cell types (Fig. 4e), which might partially account for the exceptional responses of ILC2s to androgen perturbations.

Linkages to sex-biased diseases

The identification of putative causal genes for sex-biased diseases will improve the understanding of their pathological underpinnings. We extracted the 214 diagnoses-main ICD10 disease phenotypes from the UK Biobank catalogue[27] that were relevant for the 17 tissues in our atlas for downstream analyses (Fig. 5a and Extended Data Fig. 9a). On the basis of the information in the UK Biobank, we compared the incidence rates of these 214 diseases between male and female individuals. Consequently, 119 diseases with a sex-biased prevalence (sex-biased diseases) (P adjusted < 0.05) and 95 diseases with a comparable prevalence between the sexes (sex-comparable diseases) were defined (Fig. 5b, Extended Data Fig. 9b and Supplementary Table 7). Among these 119 sex-biased diseases, malignant neoplasms of the stomach (C16), kidney (C64) and lung (C34) were male biased, whereas asthma (J45) was female biased (Fig. 5b), consistent with their clinical prevalence[7,28].

We next identified the risk single-nucleotide polymorphisms (SNPs) that were significantly different (FDR < 0.05) between cases

and controls and then linked these risk SNPs to the associated risk genes (Supplementary Table 7). The risk genes shared by multiple sex-biased diseases may represent the essential components that drive the pathogenesis of sex-biased diseases. For the 119 sex-biased diseases, even though the majority of risk genes were disease specific (4,737 of 7,426, 63.79%), 8.91% of risk genes (662 of 7,426) were shared by more than three diseases (disease-shared risk genes) (Fig. 5c). However, for the 95 sex-comparable diseases, the percentage of disease-shared risk genes was only 2.34% (97 of 4,154), significantly lower than that of the sex-biased diseases ($P < 2.20 \times 10^{-16}$, proportion test) (Fig. 5c). The 662 disease-shared risk genes for sex-biased diseases were functionally associated with antigen processing, regulation of interferon, cell–cell adhesion and synapse organization (Extended Data Fig. 9c,d). The top disease-shared risk genes of sex-biased diseases exhibited an enrichment of major histocompatibility complex (MHC) genes, as exemplified by HLA-$DQA1$, an MHC class II α-chain paralogue, were shared by 17 sex-biased diseases (Fig. 5d).

By exploring our scRNA-seq data, we demonstrated that some MHC genes were sex-biased DEGs (Extended Data Fig. 10a,b); for example, $H2$-Aa (known as HLA-$DQA1$ in humans) was a female-biased DEG in liver macrophages (Extended Data Fig. 10a). In addition, we also found that some MHC genes could be regulated by androgens (Extended Data Fig. 10c–f), as exemplified by the androgen supplementation-induced upregulation of $H2$-$Eb1$ (known as HLA-$DRB1$ in humans) in bone marrow immature B cells and $H2$-$Ab1$ (known as $H2AB1$ in humans) in spleen macrophages (Extended Data Fig. 10e), further experimentally validated by quantitative PCR with reverse transcription (qRT–PCR) (Extended
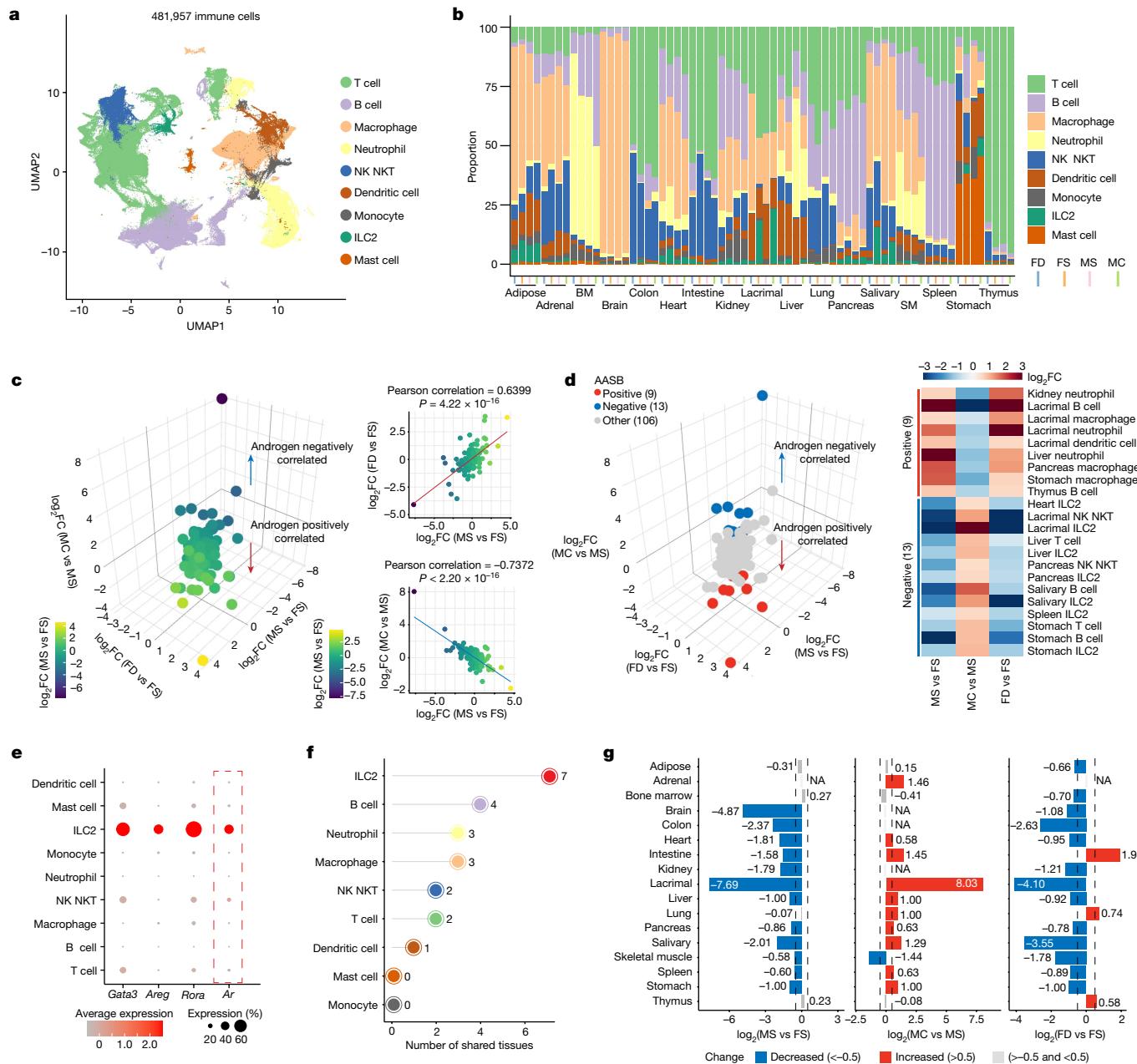
Fig. 4 | Androgens modulate sex-biased immune compartments. a, UMAP plot showing the nine major cell types of the immune system across 17 tissues from MS, MC, FS and FD mice. **b**, Cell proportions of the nine major immune cell types. BM, bone marrow; SM, skeletal muscle. **c**, Scatter plot showing the $\log_2$-transformed fold change in cell proportions based on three comparisons (left). The Pearson correlation analyses to evaluate the relationship between sex differences and the effects of androgen supplementation (top right) or between sex differences and androgen deprivation (bottom right) on immune compartments. P values are shown; the test statistic is based on Pearson's product moment correlation coefficient and follows a t distribution. **d**, Scatter plot showing the rules to define AASB-immune cell types (left). A heatmap denoting the changes in the proportions of AASB-immune cell types (right) is also shown. **e**, Dot plot showing the expression levels of *Ar* (red dashed rectangle) and ILC2 signatures, including *Gata3*, *Areg* and *Rora*. **f**, Lollipop plot showing the number of tissues sharing the target AASB-immune cell types. **g**, Bar plots showing the differences in ILC2 proportions based on three comparisons. The light grey and NA indicates the comparison was performed between two cell types with cell numbers less than 10.

Data Fig. 10g,h). These results revealed an essential role of MHC genes in sex differences in human diseases, and some of these effects could be modulated by targeting the androgen pathway.

Our single-cell atlas can help to identify cellular targets for each sex-biased disease from the expression patterns of risk genes. On the basis of our scRNA-seq data of the 17 tissues, we performed cross-species analyses to establish the associations between cell types and sex-biased diseases by determining disease enrichment scores (Extended Data Fig. 10i and Supplementary Table 8). On the basis of these data, we categorized these sex-biased diseases into five major groups, which corresponded to the five major cell lineages across tissues: 'Immune 1' (macrophages, dendritic cells and B cells), 'Immune 2' (T cells and NK cells), 'Immune 3' (neutrophils and monocytes), 'Strom-adipo' (mesenchymal cells, endothelial cells and adipocytes) and 'neural' (neural cells) (Fig. 5e).

The expected associations were observed; for example, heart disorders (I44 and I50) were classified as the Strom-adipo group in fibroblasts and endothelial cells, and gastritis and duodenitis (K29) belonged to the Immune 3 group in neutrophils and monocytes (Supplementary Table 8), mirroring their relevant clinical features. Our results also
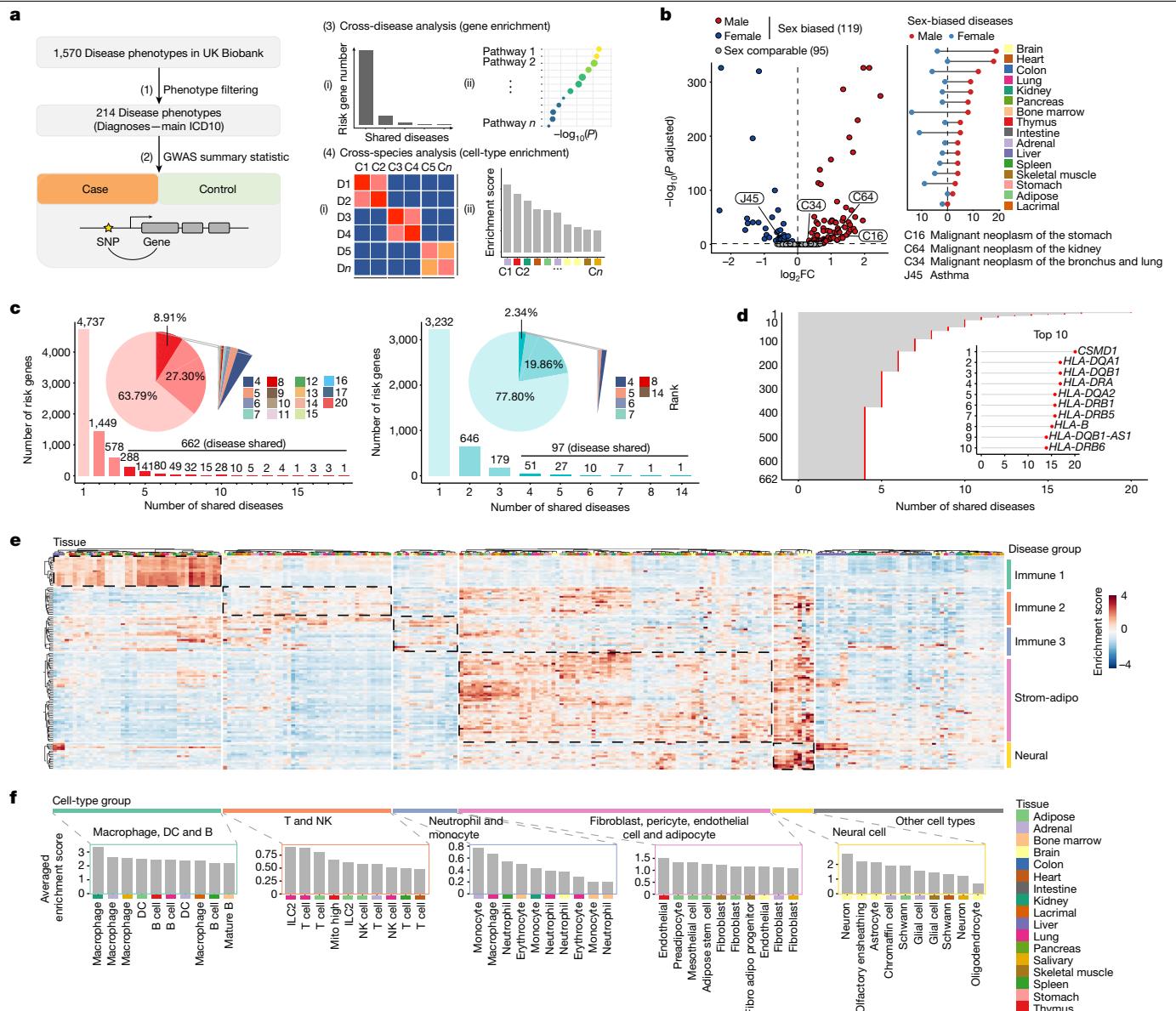
Fig. 5 | Cell-type and risk gene enrichment for sex-biased diseases.
a, Schematic of the analysis pipeline for cell-type enrichment based on cross-species integration and gene enrichment based on cross-disease analyses. GWAS, genome-wide association study. **b**, Volcano plot showing the definition of sex-biased diseases (left) and a lollipop plot showing the number of sex-biased disease phenotypes for each of the 17 tissues (right). The horizontal and vertical dashed lines represent P adjusted = 0.05 and $\log_2$FC = 0, respectively. **c**, Bar plot and pie plot denoting the categorization of risk genes according to the number of sex-biased diseases (left) or sex-comparable diseases (right). **d**, Top ten disease-shared risk genes of sex-biased diseases. **e**, Heatmap denoting the enrichment scores of sex-biased diseases across all the cell types from 17 tissues in our single-cell atlas. Sex-biased diseases were partitioned into five major groups according to their cell-type enrichment. The dashed boxes indicate the the association between disease group and cell-type group. **f**, The cell types with the top ten highest average enrichment scores for each disease group.

identified novel cell-type-disease associations; for instance, macrophages were associated with non-insulin-dependent diabetes mellitus (E11) (Supplementary Table 8). The Immune 1 group was strongly associated with macrophages and dendritic cells (average enrichment score larger than 2) across multiple tissues (Fig. 5f). As both macrophages and dendritic cells are known to highly express MHC genes, this finding was consistent with the functional enrichment of antigen presentation among disease-shared risk genes and suggested that these cell types may serve as potential cellular targets for sex-biased diseases.

Discussion

Several important studies have offered insights into sex differences[2–6]. In addition to characterizing sex differences, uncovering their underlying

mechanisms, such as modulation by androgens, will provide insights into clinical translation for sex-biased diseases. Our study systematically explored the modulation of sex differences by androgens from multiple perspectives and identified essential sex-biased molecules and cell types that are regulated by androgens and may drive sex-biased human phenotypes.

There is now increasing recognition that the sex-biased magnitude of clinical benefits from drugs can be driven by the androgen pathway in various human diseases[11,12]. Thus, identifying sex-biased genes that can be regulated by androgens through defining AASB-DEGs in each cell type will facilitate precision medicine through targeting the androgen pathway. Our study aimed to explore the androgen functions that are shared by both sexes, but sex-specific responses to androgens do exist, as exemplified by the clinical phenotype that women with androgen

Article

excess and men with androgen deficiency have some similar metabolic phenotypes, such as obesity. Thus, future insights are needed to elucidate the sex-dependent functions of androgens.

Sex-biased immune compartments contribute to the sex differences in the prevalence and severity of various diseases[1,29,30]. Our study highlights some immune cell types showing sex-different proportions and can be modulated by androgens. Of note, the proportion of ILC2s was female biased and was negatively regulated by androgens in multiple tissues. As ILC2s function in a wide variety of diseases[18–20,31], especially in cancers, targeting the androgen pathway to modulate ILC2s and facilitate immunotherapeutic strategies may be broadly applicable. Furthermore, this atlas could serve as a useful resource for the community to deeply dive into the functional association between androgens and immune cell types, including ILC2.

Androgen perturbation-induced dysregulation of other hormones, such as oestrogen and progesterone, which are mainly produced by the ovary. In addition, consistent with previous studies in humans[32,33] and in rats[34], corticosterone levels were also affected by androgen perturbations in mice (Extended Data Fig. 11). Although we have demonstrated that androgens similarly impacted the tissue weight, cell composition and gene expression between intact and ovariectomized mice, it is important to note that some effects of androgens reflected in our study may be functionally associated with androgen-induced dysregulation of other hormones to some extent. Sex hormones fluctuate during the oestrous cycle of female mice, and it would be interesting to perform single-cell profiling of female mice at different stages of the oestrous cycle in the future. Moreover, we have developed a computational pipeline to dissect the primary and secondary effects of the androgen–AR axis and have provided the related data in our web tool (https://casadbtools.com/andr_effect).

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-07291-6.

1. Khramtsova, E. A., Davis, L. K. & Stranger, B. E. The role of sex in the genomics of human complex traits. *Nat. Rev. Genet.* **20**, 173–190 (2019).
2. Oliva, M. et al. The impact of sex on gene expression across human tissues. *Science* **369**, eaba3066 (2020).
3. Lopes-Ramos, C. M. et al. Sex differences in gene expression and regulatory networks across 29 human tissues. *Cell Rep.* **31**, 107795 (2020).
4. Naqvi, S. et al. Conservation, acquisition, and functional impact of sex-biased gene expression in mammals. *Science* **365**, eaaw7317 (2019).
5. Gershoni, M. & Pietrokovski, S. The landscape of sex-differential transcriptome and its consequent selection in human adults. *BMC Biol.* **15**, 7 (2017).
6. Yang, X. et al. Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res.* **16**, 995–1004 (2006).
7. Mauvais-Jarvis, F. et al. Sex and gender: modifiers of health, disease, and medicine. *Lancet* **396**, 565–582 (2020).
8. Watson, P. A., Arora, V. K. & Sawyers, C. L. Emerging mechanisms of resistance to androgen receptor inhibitors in prostate cancer. *Nat. Rev. Cancer* **15**, 701–711 (2015).
9. Beer, T. M. et al. Enzalutamide in metastatic prostate cancer before chemotherapy. *N. Engl. J. Med.* **371**, 424–433 (2014).
10. Schmidt, K. T., Huitema, A. D. R., Chau, C. H. & Figg, W. D. Resistance to second-generation androgen receptor antagonists in prostate cancer. *Nat. Rev. Urol.* **18**, 209–226 (2021).
11. Vellano, C. P. et al. Androgen receptor blockade promotes response to BRAF/MEK-targeted therapy. *Nature* **606**, 797–803 (2022).
12. Yang, C. et al. Androgen receptor-mediated CD8+ T cell stemness programs drive sex differences in antitumor immunity. *Immunity* **55**, 1–16 (2022).
13. De Gendt, K. et al. A Sertoli cell-selective knockout of the androgen receptor causes spermatogenic arrest in meiosis. *Proc. Natl Acad. Sci. USA* **101**, 1327–1332 (2004).
14. Pihlajamaa, P. et al. Tissue-specific pioneer factors associate with androgen receptor cistromes and transcription programs. *EMBO J.* **33**, 312–326 (2014).
15. Hartig, S. M. et al. Feed-forward inhibition of androgen receptor activity by glucocorticoid action in human adipocytes. *Chem. Biol.* **19**, 1126–1141 (2012).
16. Freshour, J. R., Chase, S. E. & Vikstrom, K. L. Gender differences in cardiac ACE expression are normalized in androgen-deprived male mice. *Am. J. Physiol. Heart Circ. Physiol.* **283**, H1997–H2003 (2002).
17. Chuang, K. H. et al. Neutropenia with impaired host defense against microbial infection in mice lacking androgen receptor. *J. Exp. Med.* **206**, 1181–1199 (2009).
18. Spits, H. & Mjosberg, J. Heterogeneity of type 2 innate lymphoid cells. *Nat. Rev. Immunol.* **22**, 701–712 (2022).
19. Jacquelot, N., Seillet, C., Vivier, E. & Belz, G. T. Innate lymphoid cells and cancer. *Nat. Immunol.* **23**, 371–379 (2022).
20. Ebbo, M., Crinier, A., Vely, F. & Vivier, E. Innate lymphoid cells: major players in inflammatory diseases. *Nat. Rev. Immunol.* **17**, 665–678 (2017).
21. Buonocore, S. et al. Innate lymphoid cells drive interleukin-23-dependent innate intestinal pathology. *Nature* **464**, 1371–1375 (2010).
22. Neill, D. R. et al. Nuocytes represent a new innate effector leukocyte that mediates type-2 immunity. *Nature* **464**, 1367–1370 (2010).
23. Spits, H. et al. Innate lymphoid cells—a proposal for uniform nomenclature. *Nat. Rev. Immunol.* **13**, 145–149 (2013).
24. Cephus, J. Y. et al. Testosterone attenuates group 2 innate lymphoid cell-mediated airway inflammation. *Cell Rep.* **21**, 2487–2499 (2017).
25. Laffont, S. et al. Androgen signaling negatively controls group 2 innate lymphoid cells. *J. Exp. Med.* **214**, 1581–1592 (2017).
26. Blanquart, E. et al. Targeting androgen signaling in ILC2s protects from IL-33-driven lung inflammation, independently of KLRG1. *J. Allergy Clin. Immunol.* **149**, 237–251.e12 (2022).
27. Backman, J. D. et al. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature* **599**, 628–634 (2021).
28. Costa, A. R. et al. The sex bias of cancer. *Trends Endocrinol. Metab.* **31**, 785–799 (2020).
29. Klein, S. L. & Flanagan, K. L. Sex differences in immune responses. *Nat. Rev. Immunol.* **16**, 626–638 (2016).
30. Natri, H., Garcia, A. R., Buetow, K. H., Trumble, B. C. & Wilson, M. A. The pregnancy pickle: evolved immune compensation due to pregnancy underlies sex differences in human diseases. *Trends Genet.* **35**, 478–488 (2019).
31. Bal, S. M., Golebski, K. & Spits, H. Plasticity of innate lymphoid cell subsets. *Nat. Rev. Immunol.* **20**, 552–565 (2020).
32. Alyamani, M. et al. Deep androgen receptor suppression in prostate cancer exploits sexually dimorphic renal expression for systemic glucocorticoid exposure. *Ann. Oncol.* **31**, 369–376 (2020).
33. Li, J. et al. Aberrant corticosteroid metabolism in tumor cells enables GR takeover in enzalutamide resistant prostate cancer. *eLife* **6**, e20183 (2017).
34. Seale, J. V., Wood, S. A., Atkinson, H. C., Harbuz, M. S. & Lightman, S. L. Gonadal steroid replacement reverses gonadectomy-induced changes in the corticosterone pulse profile and stress-induced hypothalamic–pituitary–adrenal axis activity of male and female rats. *J. Neuroendocrinol.* **16**, 989–998 (2004).

Methods

Mice

Mice were kept in the specific pathogen-free animal facility at the Center for Excellence in Molecular Cell Science (CEMCS) with autoclaved food, bedding and water. Animals were housed at room temperature (23 ± 2 °C) at a humidity of 30–70% on a 12-h light/12-h dark cycle (6:00–18:00). Mice were generated through standard mouse breeding procedures within the CEMCS animal facility. All mouse studies were approved by the Institutional Animal Care and Use Committee (IACUC) of the CEMCS. Mice were bred and maintained according to Shanghai Laboratory Animal Center Institutional Animal Regulations. C57BL/6 mice were purchased from LINGCHANG Biotech. Using the systemic AR-knockout mouse model, *Rosa26*[CreERT2/+];*Ar*[flox]/Y, mice was generated by crossing *Ar*[flox]/Y mice with *Rosa26*[CreERT2/+] mice, followed by tamoxifen administration.

In vivo treatment

Male mice were castrated (MC) or sham-operated (MS) at the age of 10 weeks using standard procedures. DHT pellets (12.5 mg per pellet; NA-161, Innovative Research of America) (FD) or vehicle cholesterol pellets (sham operation (FS)) were placed subcutaneously in female mice at the age of 10 weeks. Female mice were ovariectomized followed by DHT pellets (FOD) or vehicle cholesterol pellets (FO) supplementation at the age of 10 weeks. Male mice were castrated followed by DHT supplementation (MCD). Following 4 weeks of treatment, tissues were dissociated into single-cell suspensions for scRNA-seq at the age of 14 weeks. Mice were allocated randomly to the experimental groups. The minimum sample size for each group was three, which proved sufficient to determine statistically meaningful and reproducible results.

Fluorescence-activated cell sorting

For the bone marrow immature B cells, bone marrow was pushed out into a 15-ml tube with a syringe using 5 ml of cold PBS and applied to a 70-µm cell strainer. After the removal of red blood cells, bone marrow cells were applied to a 70-µm cell strainer again to collect single cells. The bone marrow cells were stained with B220-PerCP-Cy5.5 (1:200; 552771, BD Pharmingen), CD19-APC/Cy7 (1:200; 557655, BD Pharmingen), CD93 (AA4.1)-PE (1:500; 558039, BD Pharmingen), IgM-APC (1:250; 17-5790-82, Thermo Fisher Scientific) and MHC class II I/A-FITC (1:500; 11-5321-82, Thermo Fisher Scientific). Bone marrow immature B cells (CD19+/B220+/AA4.1+/IgM+) were then sorted out. For spleen macrophages, the spleen was transferred into a 70-µm cell strainer placed in a 10-cm dish with cold PBS and then ground into single cells by the piston of a 20-ml syringe. After the removal of red blood cells, spleen cells were stained with CD45-APC (1:500; 17-0451-83, Thermo Fisher Scientific), F4/80-PE (1:500; 12-4801-82, Thermo Fisher Scientific) and Ly6G-PE/Cy7 (1:500; 25-9668-82, Thermo Fisher Scientific). Spleen macrophages (CD45+/Ly6G-/F4/80+) were then sorted out. Fluorescence-activated cell sorting (FACS) data were analysed by FlowJo (v10.0.7).

Histological and immunostaining assays

Mouse tissues were fixed with 4% paraformaldehyde overnight at 4 °C as previously reported[35]. Fixed tissues were washed with cold PBS three times and then dehydrated sequentially with 75%, 95% and 100% ethanol for 1.5 h each. Tissues were cleared in xylene for 22 min at 55 °C and then immersed in paraffin three times for 2 h each at 65 °C. For immunohistochemistry, freshly sliced 2-µm paraffin sections were dewaxed and subjected to antigen retrieval by boiling for 45 min in 0.01 M citrate buffer. Endogenous peroxidase activity was quenched by immersing the slides in 3% H_2O_2 for 20 min. The slides were then blocked in 2% BSA for 1 h and stained with primary antibodies in blocking buffer at 4 °C overnight or at room temperature for 2 h. The slides were then incubated with biotin-labelled goat anti-rabbit IgG (1:1; SP-9001 reagent 3, ZSGB-BIO) for 20 min at room temperature and horseradish peroxidase-labelled streptomyces (1:1; SP-9001 reagent 4, ZSGB-BIO) for 20 min at room temperature and stained with DAB (Vector Laboratories). The primary antibody anti-AR (1:500; ab108341, Abcam) was used. For immunofluorescence, sections were dewaxed and subjected to antigen retrieval by boiling for 45 min in 0.01 M citrate buffer. Then, sections were blocked in 5% goat serum for 1 h at room temperature, stained with primary antibodies at 4 °C overnight, washed with PBST three times and incubated with secondary antibodies for 1 h at room temperature. The sections were then washed with PBST three times and stained with DAPI (Thermo) for 7 min. The following primary antibodies were used: anti-FABP4 (1:500; ab13979, Abcam) and anti-FASN (1:500; T56597, Abmart). The following secondary antibodies were used: goat anti-rabbit IgG, Alexa Fluor 488 (1:500; A-11034, Thermo Fisher Scientific) and goat anti-rabbit IgG, Alexa Fluor 647 (1:500; A-21247, Thermo Fisher Scientific). The quantification analysis of immunofluorescence was conducted by at least two members of our laboratory who did not know the samples and antibodies used.

qRT–PCR (SYBR)

Total RNA was extracted with TRIzol reagent (15596018, Ambion). The solution was mixed well by pipetting several times and lysed at room temperature for 30 min. Then, a 1/5 volume of chloroform was added, and the mixture was vortexed for 15 s. The mixture was incubated for 2 min at room temperature and centrifuged at 13,000*g* for 15 min at 4 °C. The aqueous phase was transferred into a new tube, and an equal volume of isopropanol was added. The mixture was centrifuged at 13,000*g* for 10 min at 4 °C. The supernatant was discarded, and the pellet was resuspended in 75% ethanol and centrifuged at 13,000*g* for 7 min at 4 °C. The supernatant was then carefully removed and discarded. For RNA extraction from FACS-sorted cells, approximately 3,500 cells were sorted into 500 µl of TRIzol LS reagent (10296028, Ambion). Then, a 1/5 volume of chloroform was added, and the mixture was vortexed for 15 s. The mixture was incubated for 2 min at room temperature and centrifuged at 12,000*g* for 15 min at 4 °C. The aqueous phase was transferred into a new tube, and two volumes of 70% ethanol were added and mixed gently followed by a short centrifugation. A total of 700 µl of the mixture was transferred to a RNeasy Mini column placed in a 2-ml collection tube and centrifuged at 8,000*g* for 30 s. The flow-through was discarded and the above step was repeated to process all the remaining volume of the mixture. The following steps were performed according to the instructions of the RNeasy Mini Kit (74106, Qiagen). The pellet was resuspended in 40 µl of nuclease-free water. Reverse transcription was further performed with PrimeScript RT Master Mix (RR036A, Takara) with 400 ng (half of the RNA amount for FACS-sorted cells) of total RNA as input. qRT–PCR was conducted with SYBR qPCR Mix (1184ES08, Yeasen) using the manufacturer's protocol. The primer sequences are listed below: Mouse *Actb*-F: 5′-CATTGCTGACAGGATGCAGAAGG-3′; mouse *Actb*-R: 5′-TGCTGGAAGGTGGACAGTGAGG-3′; mouse *Il5*-F: 5′-GATGAGGCTTCCTGTCCCTACT-3′; mouse *Il5*-R: 5′-TGACAGGTTTTGGAATAGCATTTCC-3′; mouse *H2-Eb1*-F: 5′-TCCAGTGGCTTTGGTCAGAGAC-3′; mouse *H2-Eb1*-R: 5′-CAGGTTCTCCTCCAGGTTGTAG-3′; mouse *H2-Ab1*-F: 5′-GTGTGCAGACACAACTACGAGG-3′; and mouse *H2-Ab1*-R: 5′-CTGTCACTGAGCAGACCAGAGT-3′.

Measurement of serum corticosterone

Mouse blood was obtained and centrifuged at 4,000 rpm for 20 min at room temperature. The supernatant serum was saved for the measurement of serum corticosterone by the combination of liquid chromatography with mass spectrometry.

Library preparation for scRNA-seq data

scRNA-seq libraries were prepared using the Chromium Next GEM Single Cell 3′ Reagent Kit v3.1 (10X Genomics) according to the

Article

manufacturer's instructions. The target cell recovery for each library was 8,000. The constructed libraries were further sequenced on an Illumina NovaSeq 6000 (PE150) platform with up to 100 GB of data per library.

Data preprocessing and quality control for scRNA-seq data

Sequencing output from the NovaSeq system was demultiplexed using bcl2fastq (v2.20.0.422) to convert BCL files to FASTQ files. FASTQ data were mapped to the mouse genome, a precreated mRNA reference of mm10, and converted into a count matrix for further analyses by using the cellranger count function in the Cell Ranger software (v6.0.2). The resulting expression matrices were processed individually in R (v4.1.0) using the CreateSeuratObject function of Seurat (v4.0.3)[36]. All the objects from the four conditions (MS, MC, FS and FD) for each tissue were further merged together. We used the emptyDrops function of DropletUtils (v1.18.1)[37] to identify cells from empty droplets and remove barcode-swapped pseudo-cells from the gene expression matrices. In addition, DoubletFinder (v2.0.3) was applied to identify and remove doublets based on the expected doublet rate derived from the loading rate. Low-quality cells were further identified and excluded using tissue-specific thresholds according to the number of genes detected and the mitochondrial gene ratios.

Clustering and annotation of cell types

Data normalization, dimensionality reduction, clustering and identification of cell types were also conducted using Seurat (v4.0.3). (1) Filtered gene–barcode matrices were normalized with the NormalizeData function using the 'LogNormalize' method. (2) The FindVariableFeatures function was utilized to identify the top 2,000 variable genes using the default 'vst' method. (3) The FindIntegrationAnchors function was applied to identify anchors. Subsequently, the IntegrateData function was used to integrate multiple datasets. (4) The default assay for downstream analyses was set to 'integrated'. After data scaling with the ScaleData function, dimensional reduction analysis was performed using the RunPCA and RunUMAP functions. Subsequently, the FindNeighbors and FindClusters functions were used to cluster cells. We used clustertree (v0.5.0)[38] to determine the appropriate resolution. (5) Analyses of DEGs for each cluster were performed by using the Wilcoxon rank-sum test as implemented in the FindAllMarkers function, based on which cell clusters were annotated according to the representative cell markers.

Integration of all the datasets across tissues

Datasets generated from the 17 tissues were merged into one Seurat object with the merge function of Seurat. To further identify variable genes for downstream scaling, the union of anchor genes in each of the 17 tissues (17 anchor gene sets) was taken, resulting in 7,805 genes. Among these 7,805 genes, only the 3,482 genes that were expressed in all 17 tissues were retained for downstream analyses. After data scaling with the ScaleData function based on these 3,482 genes, dimensional reduction analysis was performed using the RunPCA and RunUMAP functions. The FindNeighbors and FindClusters functions were applied to cluster cells.

Cell proportion analysis

For each cell type in each sample, the number of cells was determined and divided by the total cell number of this sample to calculate the cell-type proportions. On the basis of these cell-type proportions of samples, the final cell-type proportion of each condition was defined as the average percentage of the cell-type proportions of all samples belonging to this condition. To evaluate the differences in cell proportion of one cell type between two conditions, the averaged cell-type proportion of one condition is divided by the averaged cell-type proportion of another condition, and this ratio was then transformed by $\log_2$ to generate the $\log_2$FC value. For each tissue, the top three cell types (cell proportion of more than 1%) showing the largest differences (the highest $|\log_2$FC$|$ values with FDR < 0.05 calculated by the scProportion test[39]) in proportions between the two different conditions are shown.

Definition of DEGs

The cell types that were missing or had fewer than three cells in the comparison groups (MS versus FS, MC versus MS and FD versus FS) were excluded from downstream analyses. In each of the same cell types among different conditions, DEGs were defined by the Wilcoxon rank-sum test using the FindMarkers function, the 'logfc. threshold' parameter of which was set to 0.5, resulting in DEGs with both $|\log_2$FC$| > 0.5$ and adjusted $P < 0.05$ ($|\log_2$FC$| > 0.5$ and $P < 0.05$). P values were adjusted by the Bonferroni correction implanted in the Seurat package. The DEGs defined by such criteria primarily reflected biological differences rather than the biases introduced by cell numbers or cell proportions. Sex-biased DEGs were composed of male-biased DEGs that showed significantly higher expression levels in MS mice and female-biased DEGs that showed significantly higher expression levels in FS mice.

Pathway enrichment analyses

Pathway enrichment was performed by using the enrichGO function in the clusterProfiler package (v3.12.0)[40,41] based on the DEGs for different comparisons in each cell type (P value cut-off = 0.01, q value cut-off = 0.01). P values were adjusted by the Benjamini–Hochberg method in the clusterProfiler package.

Definition of AASB-DEGs

The cell types present in all the MS, MC, FS and FD mice were retained for downstream analyses. AASB-DEGs in each cell type across the 17 tissues comprised positive AASB-DEGs, the expression levels of which were male biased and positively associated with androgens (the intersection of male-biased DEGs (MS versus FS), including upregulated DEGs in response to androgen supplementation (FD versus FS) and downregulated DEGs in response to androgen deprivation (MC versus MS)), and negative AASB-DEGs, the expression levels of which were female biased and negatively associated with androgens (the intersection of female-biased DEGs (between MS and FS), including downregulated DEGs in response to androgen supplementation (between FD and FS) and upregulated DEGs in response to androgen deprivation (between MC and MS)). The cell–AASB-DEG networks were constructed and visualized in Cytoscape (v3.9.1)[42]. Cell types and AASB-DEGs were selected as nodes, and the corresponding $\log_2$FC values were selected as edges. The nodes of cell type were coloured according to their tissue of origin, and edges were coloured according to whether their corresponding $\log_2$FC values were positive numbers (positive numbers in red; negative numbers in blue). Network styles were modified from the 'sample 1' style provided in Cytoscape. The cell-biological pathway networks were constructed and visualized in a similar manner.

Immune compartment analysis

Immune cell clusters from each of the 17 tissues were extracted and merged together to generate one Seurat object. The RunHarmony function in the Harmony package (v1.0) was used to remove batch effects from multiple datasets[43]. Dimension reduction and cell clustering were performed as described above. Non-immune cell clusters were identified and removed from downstream analyses. Subsequently, clusters were categorized into nine major cell types according to the representative marker genes. For each tissue, the number of cells of each cell type in the different conditions (MS, MC, FS and FD) was determined and divided by the total number of cells in the same group to calculate the cell-type proportion. The $\log_2$FC (MS versus FS) values were calculated to identify the sex differences

in the immune compartment between the MS and FS conditions, and the $\log_2$FC (MC versus MS) or $\log_2$FC (FD versus FS) values between the MC and MS conditions or the FD and FS conditions, respectively, were calculated to evaluate the effects of androgens on the immune system. AASB-immune cell types were composed of positive AASB-immune cell types, the proportions of which were male biased ($\log_2$FC (MS versus FS) > 0.5), increased by androgen supplementation ($\log_2$FC (FD versus FS) > 0.5) and decreased by androgen deprivation ($\log_2$FC (MC versus MS) < (−0.5)), and negative AASB-immune cell types, the proportions of which were female biased ($\log_2$FC (MS versus FS) < (−0.5)), decreased by androgen supplementation ($\log_2$FC (FD versus FS) < (−0.5)) and increased by androgen deprivation ($\log_2$FC (MC versus MS) > 0.5).

Integrative analyses of human sex-biased diseases

(1) For the identification of sex-biased diseases, the UK Biobank GWAS catalogue was downloaded, and a total of 376 disease phenotypes (labelled as Diagnoses-main ICD10) were specifically extracted. These traits were diagnosed according to the authoritative International Classification of Diseases, avoiding a scenario in which disease types are vague. Among the 376 disease phenotypes, 214 phenotypes referring to the 17 tissues in our atlas were retained for downstream analysis. Next, prop.test and p.adjust (the Bonferroni correction) in the stats R package (v4.1.0) were used to compare the incidence rates of each of the 214 diseases between males and females, generating 119 sex-biased disease phenotypes (P adjusted < 0.05) and 95 sex-comparable disease phenotypes. (2) For the definition of risk genes, the GWAS summary statistics of each phenotype were used as input to calculate the FDR (the Benjamini–Hochberg procedure) values for SNPs using the function p.adjust in the stats R package, and the SNPs with low-confidence variants (expected_case_minor_AC < 25 or minor_AF < 0.001) were filtered out for downstream analyses. Linkage of risk SNPs (FDR < 0.05) to risk genes was performed by using ANNOVAR[44]. Duplicated gene names were further removed. (3) For pathway enrichment analysis of the disease-shared risk genes, pathway enrichment analysis was performed by using the enrichGO function implemented in the clusterProfiler package (v3.12.0)[40,41] based on the disease-shared risk genes (P value cut-off = 0.01, q value cut-off = 0.01). (4) For cell-type enrichment, in each tissue, we combined the cell types belonging to the same lineage as one input cell type for downstream analyses. The AddModuleScore function in Seurat[36] was performed on the mouse homologues of risk genes of each sex-biased disease. The values calculated from the above step were further scaled to unit length by using the decostand function in the vegan package (v2.6.2), generating the disease enrichment scores. The clustering of rows (sex-biased diseases) and columns (cell types) based on the above constructed disease enrichment score matrix was performed by using the pheatmap function in the pheatmap package (v1.0.12) with the ward.D2 cluster method and Euclidean distance.

Mapping and annotating of three additional conditions (FO, FOD and MCD)

Integrated references of each tissue based on four conditions (MS, MC, FS and FD) were first built. To annotate the cell types in each tissue in three additional conditions (FO, FOD and MCD), the query dataset, which was merged from the expression matrix of the FO, FOD and MCD conditions, was constructed. Second, the FindTransferAnchors function from the Seurat package was applied to find a set of anchors between the built reference and the query dataset. These anchors were then passed to the TransferData function to transfer data from the reference to query object. Finally, the MapQuery function was used to enable the projection of a query onto the reference UMAP structure. All details regarding these analyses can be found in the website tutorial (https://satijalab.org/seurat/archive/v4.3/integration_rpca).

Statistical analyses

R (v4.1.0) and GraphPad Prism (v8.0.2) were used for statistical analysis. The difference among the experimental groups was assessed using two-tailed Student's t-test or Wilcoxon rank-sum test for comparing two groups as described in the figure legends or Methods. Correlations were estimated by Pearson correlation. Data with replicates are presented as the mean ± s.e.m.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Customizable visualizations and analyses can be executed in our web tool (https://casadbtools.com). The raw data for scRNA-seq have been deposited in the GSA (https://ngdc.cncb.ac.cn/gsa/) under CRA006610. The processed data for scRNA-seq have been deposited in the OMIX under OMIX001083. The mm10 genome reference was obtained from: https://ftp.ensembl.org/pub/release-93/fasta/mus_musculus/dna/Mus_musculus.GRCm38.dna.primary_assembly.fa.gz. Source data are provided with this paper.

Code availability

The codes used for analysing and visualizing the data in this study are available in the GitHub repository (https://github.com/lifei176/Single-cell-atlas-of-sex-differences-and-androgen-effects)[45] and in Zenodo (https://doi.org/10.5281/zenodo.10784644)[46].

35. Li, F. et al. Distinct mechanisms for TMPRSS2 expression explain organ-specific inhibition of SARS-CoV-2 infection by enzalutamide. *Nat. Commun.* **12**, 866 (2021).
36. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
37. Lun, A. T. L. et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* **20**, 63 (2019).
38. Zappia, L. & Oshlack, A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *Gigascience* **7**, giy083 (2018).
39. Miller, S. A. et al. LSD1 and aberrant DNA methylation mediate persistence of enteroendocrine progenitors that support BRAF-mutant colorectal cancer. *Cancer Res.* **81**, 3791–3805 (2021).
40. Wu, T. et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation* **2**, 100141 (2021).
41. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
42. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
43. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
44. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
45. Li, F. Single cell atlas of sex differences and androgen effects. *GitHub* https://github.com/lifei176/Single-cell-atlas-of-sex-differences-and-androgen-effects (2024).
46. Li, F. Single cell atlas of sex differences and androgen effects. *Zenodo* https://doi.org/10.5281/zenodo.10784644 (2024).

Extended Data Fig. 1 | AR-positive cells are prevalent across tissues.
a, **b**, Immunochemical staining for AR across tissues in male mice (**a**) and female mice (**b**). **c**, Schematic for the generation of the systemic AR knockout in male mice. **d**, Immunochemistry staining on AR across multiple tissues of *Rosa26*^CreERT2/+^; *Ar*^flox^/Y male mice 2 weeks post tamoxifen or corn oil treatment. Scale bar represents 50 μm. In **a**, **b** and **d**, experiments were repeated independently with similar results for three times.
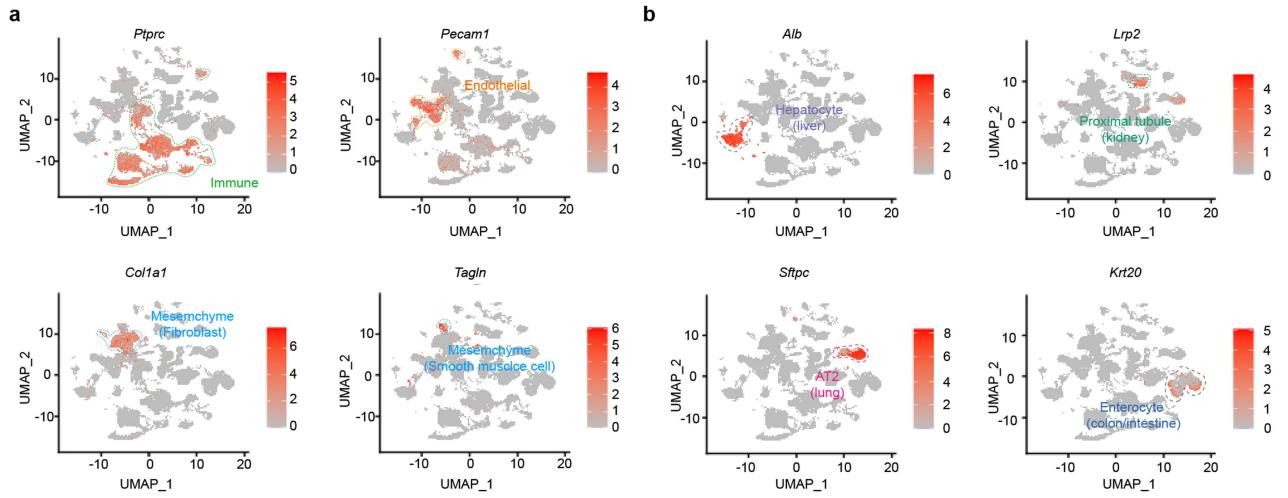
Article

Extended Data Fig. 2 | Effects of androgens on body weight and tissue weight. a, Quantification of the changes in body weight of MS (n = 15), MC (n = 16), FS (n = 27) and FD (n = 29) mice. n denotes the number of mice. Data are mean ± s.e.m. **b**, Volcano plot showing the weight changes in each tissue induced by androgen deprivation in male mice. The tissues with decreased weight in response to androgen deprivation are marked in blue, while the tissues with increased weight in response to androgen deprivation are marked in red. **c**, Heatmap showing the weight changes in each tissue induced by androgen deprivation in male mice. **d**, Volcano plot showing the changes in the weight of each tissue induced by androgen supplementation in female mice.

The tissues with decreased weight in response to androgen supplementation are marked in blue, while the tissues with increased weight in response to androgen supplementation are marked in red. **e**, Heatmap showing the weight changes in each tissue induced by androgen supplementation in female mice. **f**, Summary of the effects of androgens on the weight of each tissue. **g**, Representative tissues, including the salivary gland (androgen-promoting, left), thymus (androgen-suppressing, middle) and liver (nonsignificant changes, right) tissue of MS, MC, FS and FD mice. In **b**, **d**, statistical comparisons were performed using two-tailed Student's t test.

a

Ptprc

Endothelial

Immune

Pecam1

Col1a1

Mesemchyme
(Fibroblast)

Tagln

Mesemchyme
(Smooth musclce cell)

b

Alb

Hepatocyte
(liver)

Lrp2

Proximal tubule
(kidney)

Sftpc

AT2
(lung)

Krt20

Enterocyte
(colon/intestine)
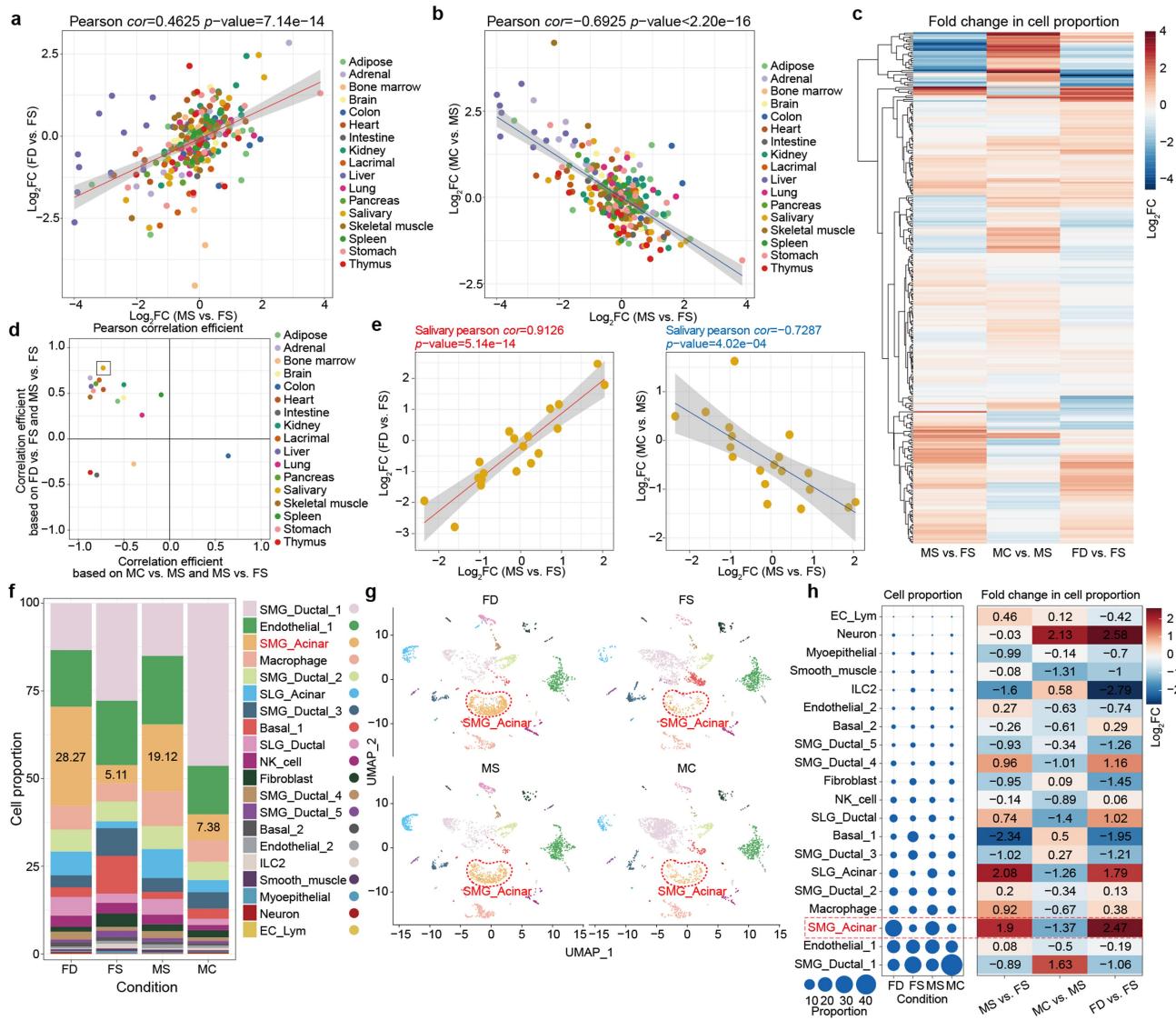
Extended Data Fig. 3 | Expression levels of representative markers.
a, Feature plot showing the RNA expression levels of representative markers
for tissue-common cell lineages, including *Ptprc* for immune cells, *Pecam1* for
endothelial cells, *Col1a1* for fibroblasts and *Tagln* for smooth muscle cells.

b, Feature plot showing the RNA expression levels of representative markers
for tissue-specific cell lineages, including *Alb* for hepatocytes in the liver, *Lrp2*
for proximal tubule cells in the kidney, *Sftpc* for AT2 cells in the lung and *Krt20*
for enterocytes in the colon and intestine.
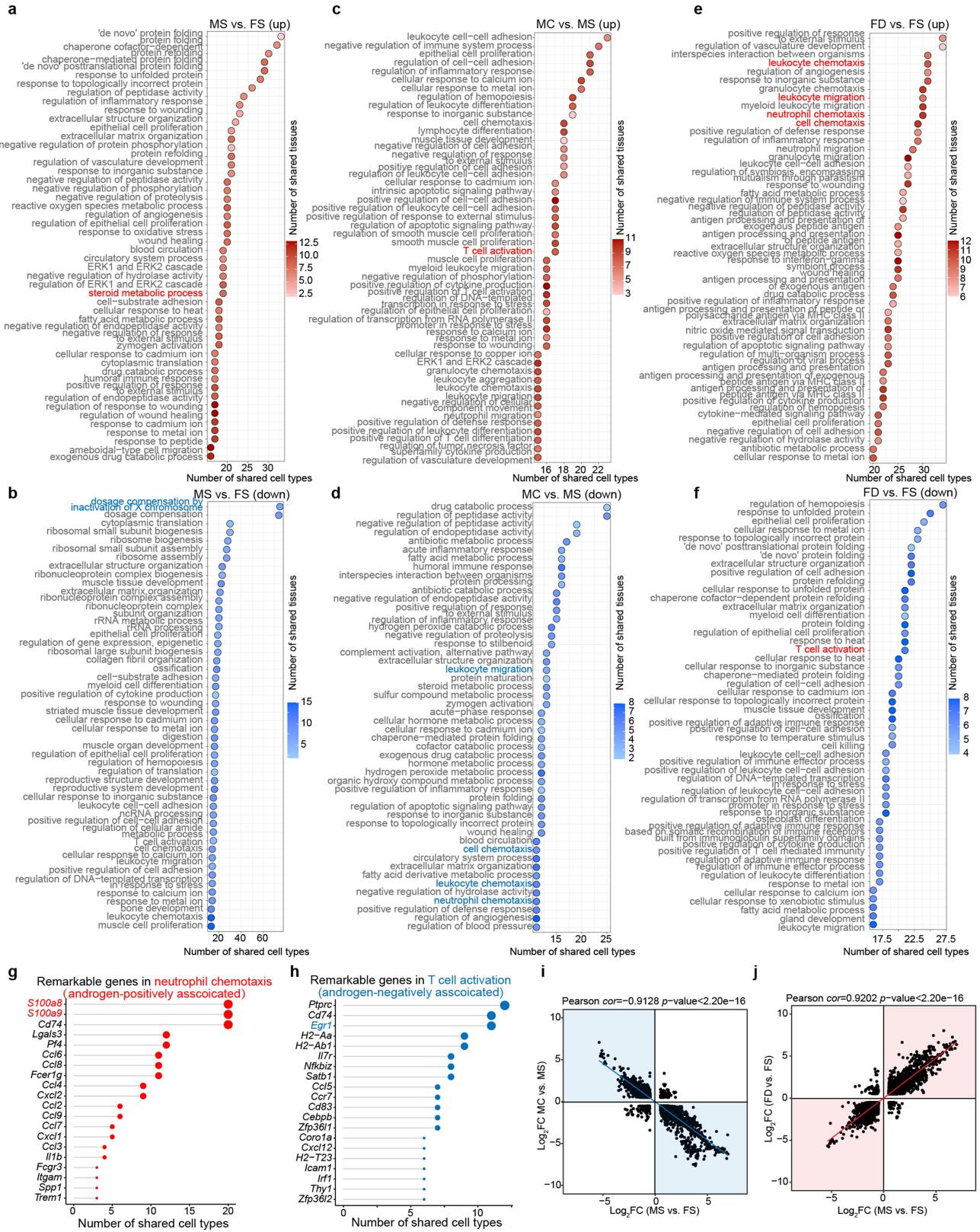
Extended Data Fig. 4 | Androgens modulate sex-biased cellular populations.
a, Scatter plot showing the correlation between sex differences (log$_2$(MS vs. FS)) and the differences in the proportions of all the cell types across tissues between FD and FS (log$_2$(FD vs. FS)). **b**, Similar to (**a**), but for the correlation between log$_2$(MS vs. FS) and log$_2$(MC vs. MS). **c**, Heatmap denoting the impacts of sex and androgens on the proportions of all cell types across tissues. The color key represents the log$_2$-transformed fold change in the cell proportions. **d**, Dot plot showing the Pearson correlation efficiency calculated based on sex differences (log$_2$(MS vs. FS)) and the differences in cell composition between MC and MS (log$_2$(MC vs. MS)) (x-axis), as well as sex differences (log$_2$(MS vs. FS)) and the differences in cell composition between FD and FS (log$_2$(FD vs. FS)) (y-axis). The dot color represents the tissue type. **e**, Scatter plot showing the correlation between sex differences (log$_2$(MS vs. FS)) and the differences in

salivary cell composition between FD and FS (log$_2$(FD vs. FS), left) or the differences in salivary cell composition between MC and MS (log$_2$(MC vs. MS), right). **f**, Bar plot denoting the cellular composition of salivary gland across four conditions. **g**, UMAP plot showing the cell composition of salivary gland based on four conditions involving sex and androgens. Two thousand cells for each condition were randomly extracted for visualization. **h**, Bubble plot (left) and heatmap (right) denoting the impacts of sex and androgens on the proportions of all salivary cell types. The size indicates the proportion. The color key represents the log$_2$-transformed fold change in cell proportions. In **a**, **b**, and **e**, *P* values are shown; the test statistic is based on Pearson's product moment correlation coefficient and follows a *t* distribution; level of confidence interval is 0.95.
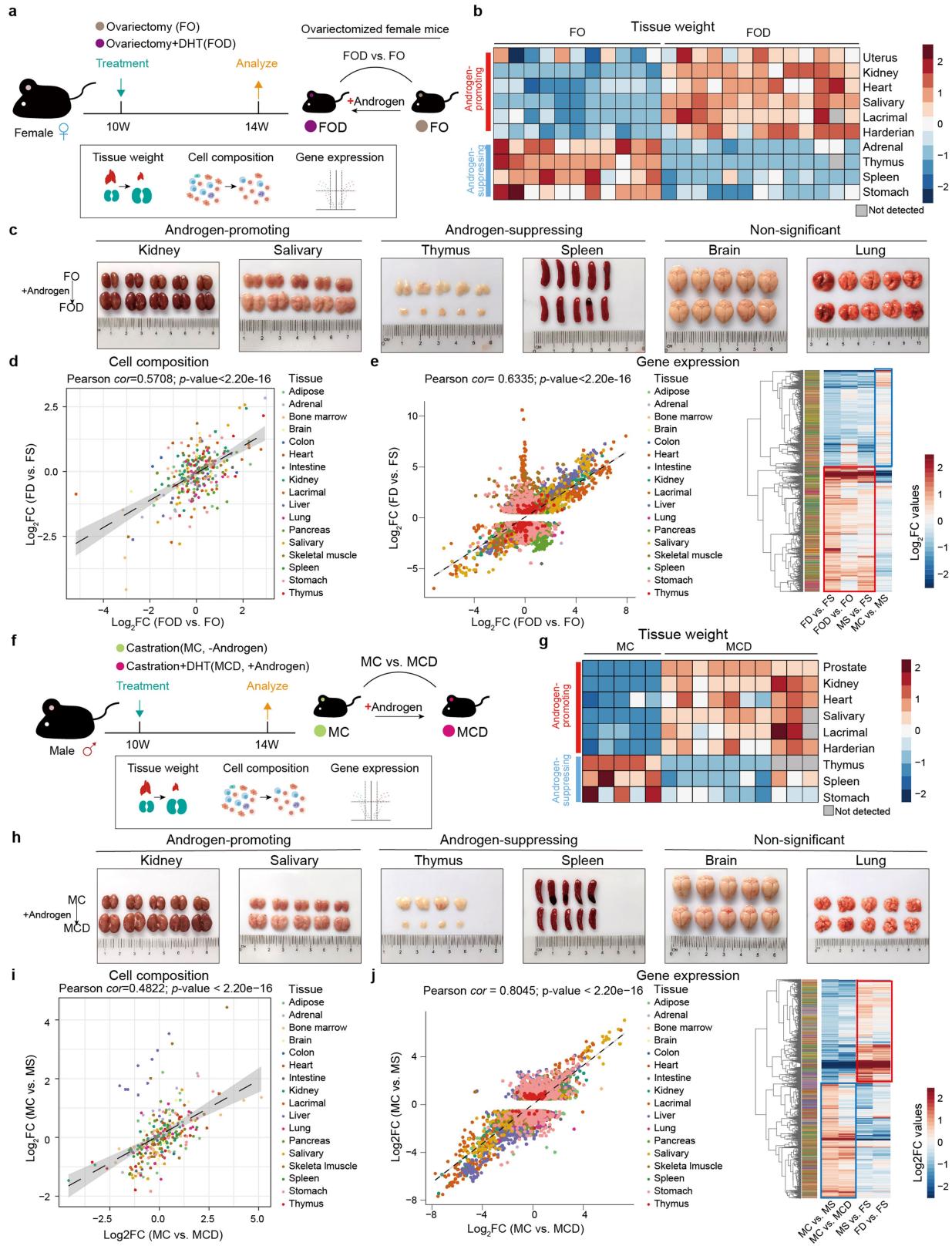
Extended Data Fig. 5 | See next page for caption.

Article

Extended Data Fig. 5 | Cross-cell type effects of sex and androgens on biological pathways. a-f, Dot plot showing the top 50 cross-cell type biological pathways based on the upregulated DEGs (**a**) and downregulated DEGs (**b**) between MS and FS, the upregulated DEGs (**c**) and downregulated DEGs (**d**) between MC and MS, and the upregulated DEGs (**e**) and downregulated DEGs (**f**) between FD and FS. The x-axis represents the number of cell types sharing the target biological pathway. Gradient color represents the number of tissues sharing the target biological pathway. **g**, **h**, Lollipop plot showing the remarkable androgen-associated genes in the "neutrophil chemotaxis" pathway (**g**) and the "T-cell activation" pathway (**h**). The x-axis represents the number of cell types sharing the target gene. **i**, Scatter plot showing the correlation of $\log_2 FC$ in the DEGs which locates on autosomes between sex differences and MC vs. MS based on all of the cell types from 17 tissues. **j**, Scatter plot showing the correlation of $\log_2 FC$ in the DEGs which locates on autosomes between sex differences and FD vs. FS based on all of the cell types from 17 tissues. In **i**, **j**, P values are shown; the test statistic is based on Pearson's product moment correlation coefficient and follows a t distribution.
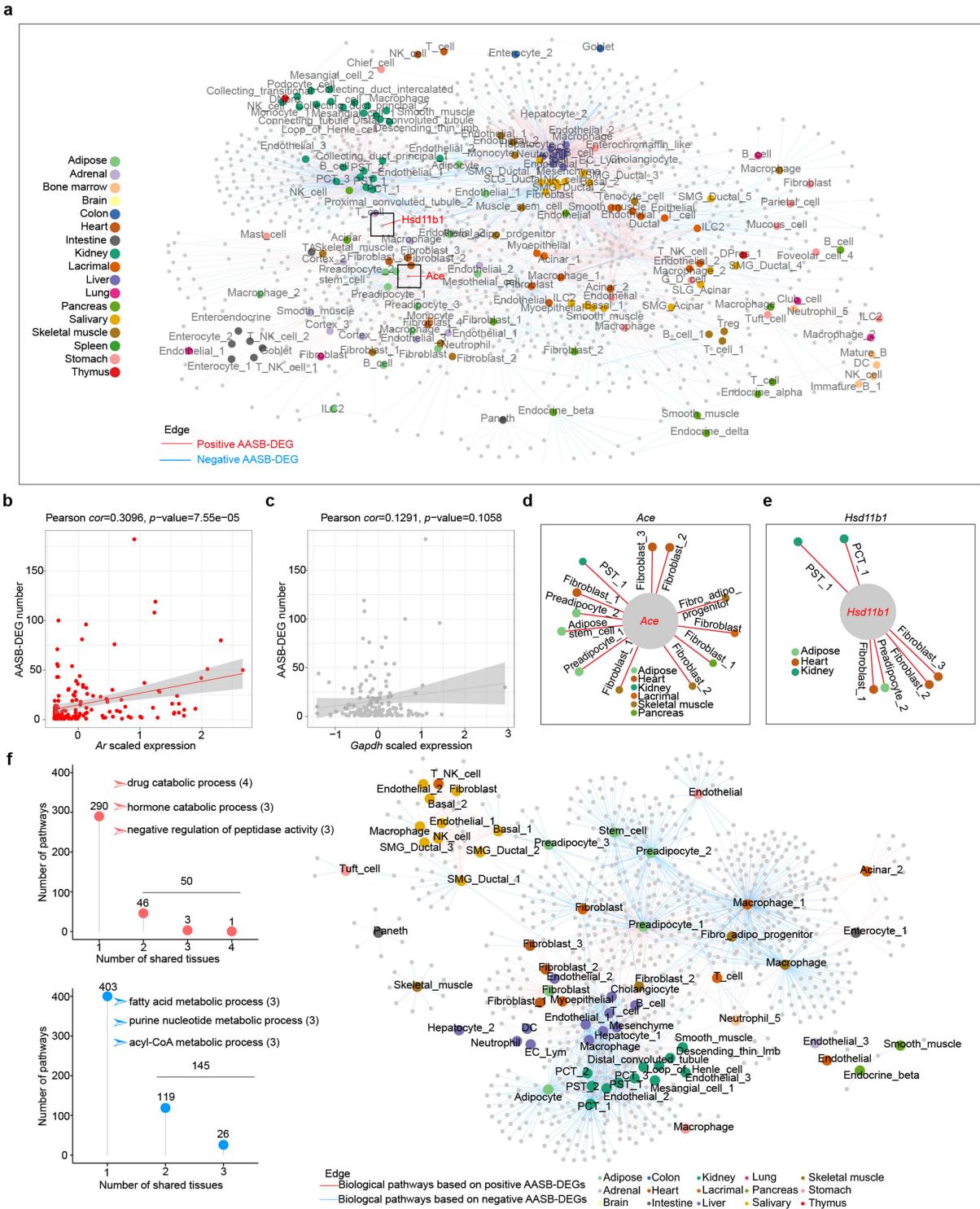
Extended Data Fig. 6 | See next page for caption.

Article

Extended Data Fig. 6 | The effects of androgen supplementation on ovariectomized female mice and castrated male mice. a, Schematic diagram of the experimental design to investigate the effects of androgens on the tissue weight, gene expression and cell composition in ovariectomized female mice (FOD vs. FO). **b**, Heatmap showing the tissue weight changes induced by androgen supplementation in ovariectomized female mice. **c**, Representative tissues, including the kidney and salivary gland (androgen-promoting), thymus and spleen (androgen-suppressing), and brain and lung (nonsignificant changes) of FO and FOD mice. **d**, Scatter plot showing the correlation of androgen supplementation-induced cell proportional differences between intact (FD vs. FS) and ovariectomized (FOD vs. FO) female mice across tissues. Level of confidence interval is 0.95. **e**, Scatter plot showing the correlation of DEG expression differences between FD vs. FS and FOD vs. FO mice (left). Heatmap denoting the log$_2$FC in the expression of the DEGs from a single cell type across the four comparisons, where DEGs are colored according to their tissue of origin (right). **f**, Similar to (**a**) but for castrated male mice (MCD vs. MC). **g**, Similar to (**b**) but for androgen supplementation in castrated male mice. **h**, Similar to (**c**) but for MC and MCD mice. **i**, Scatter plot showing the correlation of cell proportional differences between the comparisons of MC vs MCD and MC vs. MS across tissues. Level of confidence interval is 0.95. **j**, Similar to (**e**) but for comparison between MC vs. MCD and MC vs. MS mice. In **d**, **e**, **i**, and **j**, P values are shown; the test statistic is based on Pearson's product moment correlation coefficient and follows a t distribution. In each condition (FOD, FO and MCD), each tissue contained at least three biological replicates (three mice).
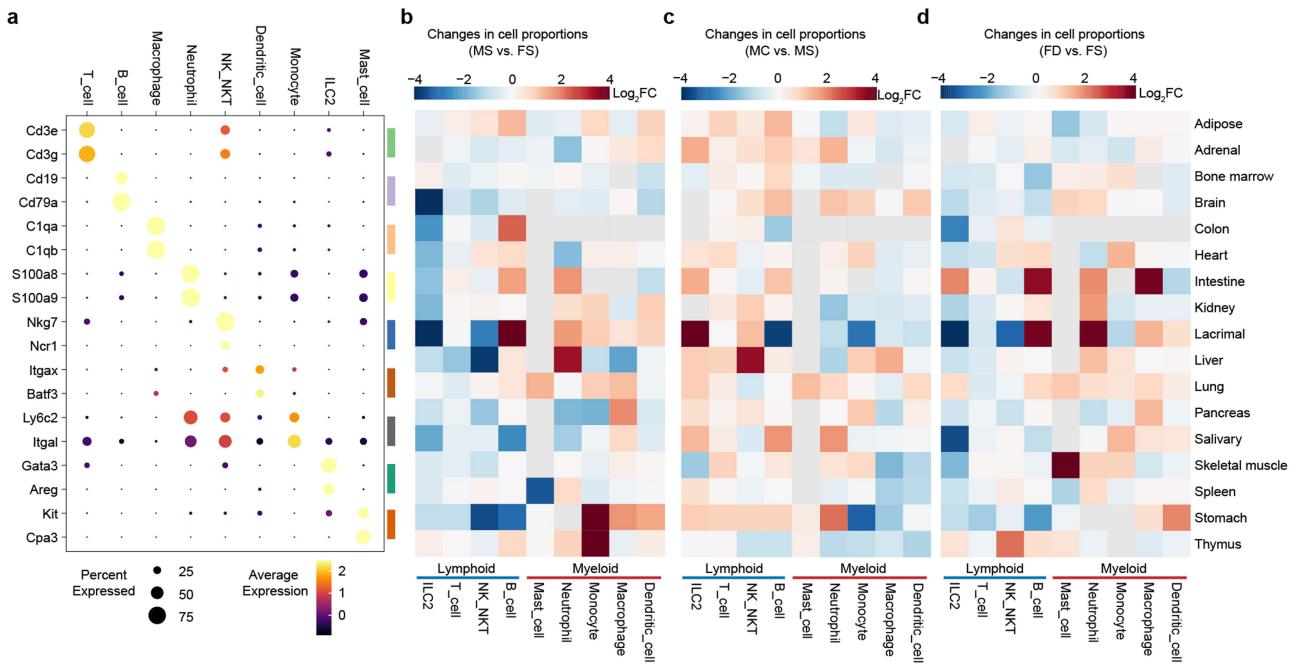
Extended Data Fig. 7 | AASB-DEGs and their functionally enriched pathways across tissues. a, Network of all AASB-DEGs in each cell type across tissues. Red edges and blue edges represent the positive AASB-DEGs and the negative AASB-DEGs, respectively. The gray node represents the AASB-DEG, and the colored node represents the cell type. **b**, Scatter plot showing the correlation between the numbers of AASB-DEGs and *Ar* RNA expression. **c**, Scatter plot showing the correlation between the numbers of AASB-DEGs and *Gapdh* RNA expression. **d**, **e**, Representative cross-tissue AASB-DEGs and their relevant cell types, including *Ace* (positive AASB-DEG, **d**) and *Hsd11b1* (positive AASB-DEG, **e**). **f**, Lollipop plot denoting the shared biological pathways across tissues based on positive (left top) or negative (left bottom) AASB-DEGs. Network of significantly enriched biological pathways based on AASB-DEGs in each cell type across tissues (right). In **b**, **c**, *P* values are shown; the test statistic is based on Pearson's product moment correlation coefficient and follows a t distribution; level of confidence interval is 0.95.
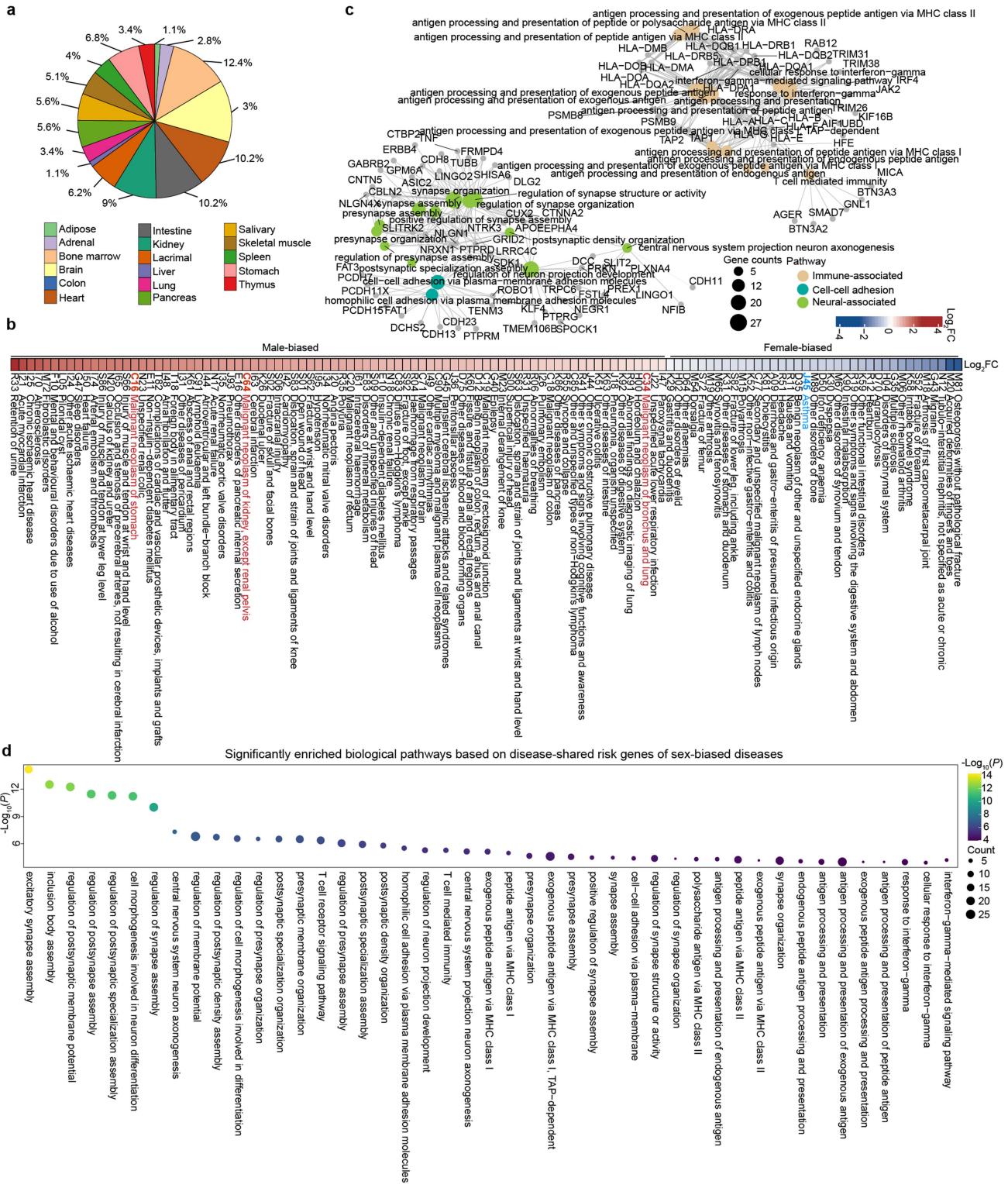
Extended Data Fig. 8 | Dissection of the effects of androgens on the immune compartment. a, Dot plot showing the expression patterns of representative markers for nine major immune cell types. The dot size represents the percentage of cells with detectable gene expression. The color key represents the expression levels. **b-d**, Heatmap denoting sex differences (**b**) and the effects of androgen deprivation (**c**) and androgen supplementation (**d**) on immune cell proportions across tissues. The color key represents the log₂-transformed fold change in the cell proportion. Light gray indicates that the comparison was performed on two cell types with cell numbers less than 10.
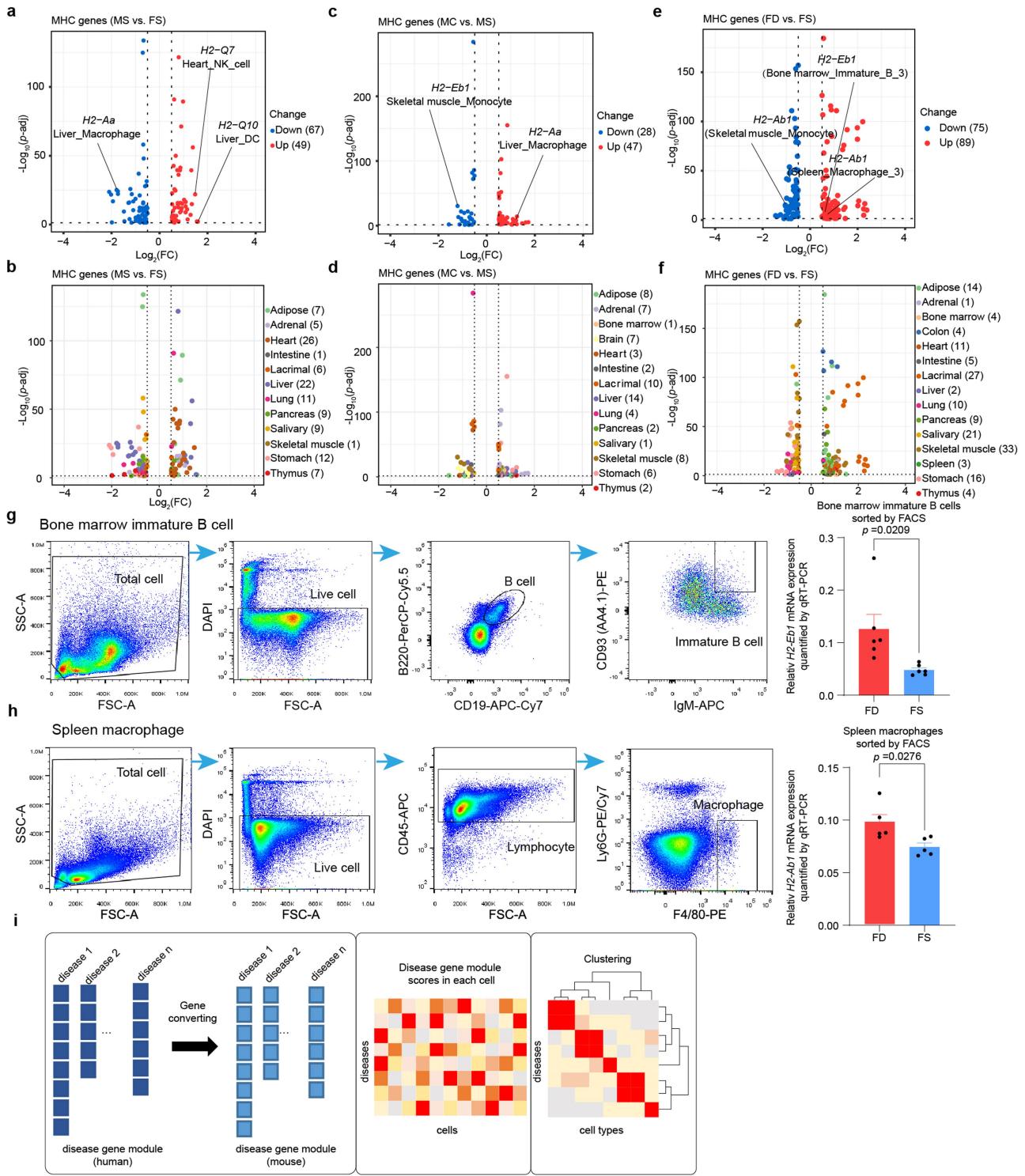
Extended Data Fig. 9 | Enriched biological pathways based on disease-shared risk genes. a, Pie chart showing the categorization of sex-biased diseases in the 17 tissues in our atlas. **b**, Heatmap denoting the differences in the incidence rates of sex-biased diseases between males and females. Representative male-biased and female-biased diseases are marked in red and blue, respectively. **c**, Network denoting the top 30 significantly enriched biological pathways based on the 662 risk genes that were shared by multiple sex-biased diseases. Node size represents gene number. Distinct biological pathway modules are marked in different colors. **d**, Dot plot denoting the significantly enriched biological pathways based on the 662 disease-shared risk genes of sex-biased diseases. Significantly enriched biological pathways were defined by the hypergeometric distribution and *P* values were adjusted by the Benjamini-Hochberg method (BH) by using *EnrichGO* function implanted in the clusterProfiler.
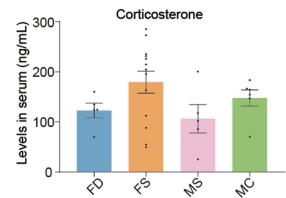
Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Sex and androgens influence MHC gene expressions. a, **b**, Volcano plot denoting the MHC genes that showed significant differential expression (**a**) and their tissue of origin (**b**) based on the comparison of MS vs. FS. DEGs were defined by Wilcoxon Rank Sum test and *P* values were adjusted by the Bonferroni correction using the *FindMarkers* function implanted in Seurat. **c**, **d**, Volcano plot denoting the MHC genes that showed significant differential expression (**c**) and their tissue of origin (**d**) based on the comparison of MC vs. MS. DEGs were defined by Wilcoxon Rank Sum test and *P* values were adjusted by the Bonferroni correction using the *FindMarkers* function implanted in Seurat. **e**, **f**, Volcano plot denoting the MHC genes that showed significant differential expression (**e**) and their tissue of origin (**f**) based on the comparison of FD vs. FS. DEGs were defined by Wilcoxon

Rank Sum test and *P* values were adjusted by the Bonferroni correction using the *FindMarkers* function implanted in Seurat. **g**, Gating strategy for bone marrow immature B cells (CD19$^+$/B220$^+$/AA4.1$^+$/IgM$^+$) and relative *H2-Eb1* mRNA expression level quantified by qRT–PCR in the FACS-sorted bone marrow immature B cells of FD and FS mice. n = 6 mice examined per group. **h**, Gating strategy for spleen macrophages (CD45$^+$/Ly6G$^+$/F4/80$^+$) and relative *H2-Ab1* mRNA expression level quantified by qRT–PCR in the FACS-sorted spleen macrophages of FD and FS mice. n = 5 mice examined per group. **i**, Schematic to explain how we categorized human sex-biased diseases into five major groups using our scRNA-seq data. In **g**, **h**, Data are mean ± s.e.m. *P* values are shown; statistical comparisons were performed using two-tailed Student's t test.

Extended Data Fig. 11 | Serum levels of corticosterone. Serum levels of corticosterone in FD, FS, MS and MC mice (FD: n = 5 mice; FS: n = 13 mice; MS: n = 5 mice; MC: n = 6 mice). n denotes the number of mice. Data are mean ± s.e.m.

nature portfolio

Corresponding author(s): Dong Gao

Last updated by author(s): Mar 6, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The tissue-specific dissociation protocols to generate single-cell suspension have been detailedly described in the Supplementary Methods of our manuscript. Single-cell RNA-seq libraries were prepared using the Chromium Next GEM Single Cell 3' Reagent Kit v3.1 (10x Genomics) according to the manufacturer's instructions. The target cell recovery for each library was 8,000 cells. The constructed libraries were further sequenced on an Illumina NovaSeq 6000 (PE150) platform.
Data analysis	The detailed methods for data pre-processing, analysis and the analysis softwares have been provided in the Methods of our manuscript. For single-cell data analysis, sequencing output from the NovaSeq system was demultiplexed using bcl2fastq (version 2.20.0.422). Single-cell expression matrix was generated by Cell Ranger software (version 6.0.2) and processed by Seurat (version 4.0.3) in R (version 4.1.0). Empty droplets and doublets were identified and removed by DropletUtils (version 1.18.1) and DoubletFinder (version 2.0.3) respectively. The resolution of clustering was determined by clustertree (version 0.5.0). Immune cell data was integrated by Harmony package (version 1.0). The disease enrichment scores were generated by scaling to unit length by using the decostand function in the vegan package (version 2.6.2). Pathway enrichment analyses were performed by clusterProfiler (version 3.12.0). Gene network was visualized by Cytoscape (version 3.9.1). Heatmap was generated by pheatmap (version 1.0.12). For fluorescence-activated cell sorting (FACS) analysis, cell identification and quantification were analyzed by FlowJo (version 10.0.7). R (version 4.1.0) and GraphPad Prism (version 8.0.2) were used for statistical analysis. In addition, we have also constructed a webtool application for customized data analyses and visualizations at: https://casadbtools.com. The codes used for analyzing and visualizing the data in this study are available in GitHub repository (https://github.com/lifei176/Single-cell-atlas-of-sex-differences-and-androgen-effects) and Zenodo (https://doi.org/10.5281/zenodo.10784644).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about <u>availability of data</u>

All manuscripts must include a <u>data availability statement</u>. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our <u>policy</u>

The raw data for single-cell RNA-seq has been deposited in GSA (https://ngdc.cncb.ac.cn/gsa/) under CRA006610. The processed data for single-cell RNA-seq has been deposited in OMIX (https://ngdc.cncb.ac.cn/omix/releaseList) under OMIX001083. The mm10 genome reference was obtained from: https://ftp.ensembl.org/pub/release-93/fasta/mus_musculus/dna/Mus_musculus.GRCm38.dna.primary_assembly.fa.gz.

Human research participants

Policy information about <u>studies involving human research participants and Sex and Gender in Research.</u>

Reporting on sex and gender	We have carefully used the terms of "sex" and "gender" in our manuscript.
Population characteristics	Our study did not involve human research participants.
Recruitment	Our study did not involve human participants recruitment.
Ethics oversight	Our study did not need ethics oversight in relation to human research.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see <u>nature.com/documents/nr-reporting-summary-flat.pdf</u>

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size for each data has been stated in their relevant figure legends. The minimum sample size for each group is three (n=3), which proved sufficient to determine statistically meaningful and reproducible results.
Data exclusions	Low-quality single-cell RNA-seq data was excluded due to a technical issue on tissue dissociation .
Replication	All experiments were repeated multiple times as stated in the individual figure legends. All the experiments at replication were successful. For scRNA-seq, in each condition, each tissue contained at least three biological replicates (three mice). Immunofluorescence assays, immunochemical staining assays, FACS assays and qRT-PCR assays contained at least three biological replicates.
Randomization	Mice were allocated randomly to the experimental groups. Fields with useful information for histology analysis by immunofluorescence and immunochemistry were selected randomly. Mice with different treatments for single-cell RNA-seq data were randomly selected.
Blinding	The quantification of immunofluorescence was conducted by at least two members of our lab who did not know the samples and antibodies we used. Other experiments does not require investigators to be blinded to group allocation during data collection and/or analysis as they were much more quantitative. Samples were analyzed with the same protocol by different investigators. The data collection was not subjective.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	The following antibodies for immunohistochemistry were used: anti-AR (Abcam, ab108341), biotin-labeled goat anti-rabbit IgG (ZSGB-BIO, SP-9001 reagent 3) and HRP-labeled streptomyces (ZSGB-BIO, SP-9001 reagent 4); The following antibodies for immunofluorescence were used: anti-FABP4 (Abcam, ab13979), anti-FASN (Abmart, T56597), goat anti-rabbit IgG, Alexa Fluor™ 488 (Thermo Fisher Scientific, A-11034) and goat anti-rabbit IgG, Alexa Fluor™ 647 (Thermo Fisher Scientific, A-21247); The following antibodies for FACS were used: B220-PerCP-Cy5.5 (BD Pharmingen, 552771), CD19-APC/Cy7 (BD Pharmingen, 557655), CD93 (AA4.1)-PE (BD Pharmingen, 558039), IgM-APC (Thermo Fisher Scientific, 17-5790-82), MHC Class II I/A-FITC (Thermo Fisher Scientific, 11-5321-82), CD45-APC (Thermo Fisher Scientific, 17-0451-83), F4/80-PE (Thermo Fisher Scientific, 12-4801-82), and Ly6G-PE/Cy7 (Thermo Fisher Scientific, 25-9668-82).
Validation	The validation of each primary antibody was conducted by manufactures. The validations of each antibody according to the manufactures website are listed below: The AR antibody (Abcam, ab108341) previously validated in mouse for IHC as stated by the manufactures on their website (https://www.abcam.cn/products/primary-antibodies/androgen-receptor-antibody-er1792-chip-grade-ab108341.html). The FABP4 antibody (Abcam, ab13979) previously validated in mouse for IF as stated by the manufactures on their website (https://www.abcam.cn/products/primary-antibodies/fabp4-antibody-ab13979.html). Also validated in Jiang Z et al., Cell Rep, 2021. The FASN antibody (Abmart, T56597) previously validated in mouse for IF as stated by the manufactures on their website (http://www.ab-mart.com.cn/page.aspx?node=%2077%20&id=%202265). The B220-PerCP-Cy5.5 antibody (BD Pharmingen, 552771) previously validated in mouse for FACS as stated by the manufactures on their website (https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/percp-cy-5-5-rat-anti-mouse-cd45r-b220.552771). The CD19-APC/Cy7 antibody (BD Pharmingen, 557655) previously validated in mouse for FACS as stated by the manufactures on their website (https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/apc-cy-7-rat-anti-mouse-cd19.557655). The CD93 (AA4.1)-PE antibody (BD Pharmingen, 558039) previously validated in mouse for FACS as stated by the manufactures on their website (https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/pe-rat-anti-mouse-cd93-early-b-lineage.558039). The IgM-APC antibody (Thermo Fisher Scientific, 17-5790-82) previously validated in mouse for FACS as stated by the manufactures on their website (https://www.thermofisher.cn/cn/zh/antibody/product/IgM-Antibody-clone-II-41-Monoclonal/17-5790-82). The MHC Class II I/A-FITC antibody (Thermo Fisher Scientific, 11-5321-82) previously validated in mouse for FACS as stated by the manufactures on their website (https://www.thermofisher.cn/cn/zh/antibody/product/MHC-Class-II-I-A-I-E-Antibody-clone-M5-114-15-2-Monoclonal/11-5321-82). The CD45-APC antibody (Thermo Fisher Scientific, 17-0451-83) previously validated in mouse for FACS as stated by the manufactures on their website (https://www.thermofisher.cn/cn/zh/antibody/product/CD45-Antibody-clone-30-F11-Monoclonal/17-0451-83). The F4/80-PE antibody (Thermo Fisher Scientific, 12-4801-82) previously validated in mouse for FACS as stated by the manufactures on their website (https://www.thermofisher.cn/cn/zh/antibody/product/F4-80-Antibody-clone-BM8-Monoclonal/12-4801-82). The Ly6G-PE/Cy7 antibody (Thermo Fisher Scientific, 25-9668-82) previously validated in mouse for FACS as stated by the manufactures on their website (https://www.thermofisher.cn/cn/zh/antibody/product/Ly-6G-Antibody-clone-1A8-Ly6g-Monoclonal/25-9668-82).

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	C57BL/6 mice were purchased from LINGCHANG Biotech. Male mice were castrated (MC) or sham-operated (MS) at the age of 10 weeks using standard procedures. Dihydrotestosterone (DHT) pellets (12.5 mg per pellet, Innovative Research of America, NA-161) (FD) or vehicle cholesterol pellets (sham operation, FS) were placed subcutaneously in female mice at the age of 10 weeks. Female mice were ovariectomized followed by DHT pellets (FOD) or vehicle cholesterol pellets (FO) supplementation at the age of 10 weeks. Male mice were castrated followed by DHT supplementation (MCD). Following 4 weeks of treatment, tissues were dissociated into single-cell suspensions for scRNA-seq at the age of 14 weeks. Rosa26CreERT2/+; Arflox/Y mice (C57BL/6 background) were treated with tamoxifen at the age of 8 weeks and collected at the age of 10 weeks.
Wild animals	Our study dose not include wild animals.
Reporting on sex	Our study aims to systematically explore the physiological sex differences and their regulations by androgens. We have clearly reported the sex of each mice used in our study.

Field-collected samples	This study did not include samples collected from the field.
Ethics oversight	All mouse studies were approved by the Center for Excellence in Molecular Cell Science (CEMCS) (Shanghai Institute of Biochemistry and Cell Biology, SIBCB) Animal Care and Use Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	For the bone marrow immature B cells, bone marrow was pushed out into a 15-mL tube with a syringe using 5 mL of cold PBS and applied to a 70-μm cell strainer. After the removal of red blood cells, bone marrow cells were applied to a 70-μm cell strainer again to collect single cells. For spleen macrophages, the spleen was transferred into a 70-μm cell strainer placed in a 10-cm dish with cold PBS and then ground into single cells by the piston of a 20-mL syringe. After the removal of red blood cells, spleen cells were stained.
Instrument	FACS assays were performed on a Sony MA900.
Software	Downstream quantification analyses were performed by using FlowJo (version 10.0.7).
Cell population abundance	Cell purity of target cells was over 90% as indicated by Sony MA900 when sorting these cells out.
Gating strategy	Gating strategy to sort out immature B cells from the total live cells in the bone marrow was based on CD19+/B220+/AA4.1+/IgM+. Gating strategy to sort out macrophages from the total live cells in the spleen was based on CD45+/Ly6G-/F4/80+.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.