**Table S1. Gene markers for cluster annotation.** This list defines each subpopulation of thyroid organoids using scTyper.

| Cell type | Genes |
| --- | --- |
| Immature Thyrocytes | PAX8, FOXE1, NKX2-1, HHEX, TSHR, TG |
| Mature Thyrocytes | PAX8, FOXE1, NKX2-1, HHEX, TSHR, TG, TPO |
| Thyroid progenitors | PAX8, FOXE1, NKX2-1, HHEX |
| Goblet cells | AGR2, AQP3, ATOH1, BACE2, CDON, CDX2, CLCA1, CREB3L1, FCGBP, GALNT12, GUCA2A, ITLN1, KLK1, KRT20, KRT7, LRRC26, MANF, MUC13, MUC2, MUC4, MUC5B, NLRP6, PDIA5, PHGR1, PLA2G10, REP15, SLC9A8, SPDEF, SPINK4, TFF3, TPSG1, ZG16 |
| Embryonic stem cells | ABCG2, ANO6, BCL3, BNIP3, CD24, CD59, CD9, CDH1, CDK8, CTNNB1, DNMT3B, DPPA2, DPPA3, DPPA4, EPCAM, ESRRB, FBXO15, FUT4, FZD1, GAL, GDF3, GJB1, GJB4, GJC1, HES1, HHEX, HMGA2, HOXB5, IL6ST, ITGA4, ITGA6, ITGB1, KCNIP3, KIT, KITLG, KLF4, L1TD1, LEF1, LIFR, LMNA, MYC, NACC1, NANOG, NR6A1, PCGF2, PECAM1, PITX2, PIWIL1, PIWIL2, PIWIL4, PML, PODXL, POU5F1, PRDM5, PROM1, PUM2, SALL4, SLC46A2, SMAD1, SMAD2, SMAD3, SMAD4, SMAD5, SMAD9, SOX15, SOX2, STAT3, SUMO2, TAF8, TDGF1, TEX19, THY1, TRIM28, TRIM6, ZFP42, ZFX, ZIC1 |
| Airway progenitor cells | ABI3BP, AQP3, DAPL1, GSTM2, HPGD, ICAM1, KRT14, KRT15, KRT5, PHLDA3, RPS18, SDC1 |
| Epithelial cells | CD24, CEACAM1, ST6GAL1, ITGB4, IL1R1, PROM1, CDH1, KRT1, KRT7, MUC1, ICAM1, KRT14, KRT5, ITGAL, CD2, KLK3, ITGA5, ITGA4, ITGA2, KRT3, KRT16, SCNN1A, KRT15, ITGA1, KRT2, SCNN1B, SCNN1D, SCNN1G, IFI16, BOK, NKD1, FZD6, DKK3, NRP2, SFRP5, RAI14, DEFB1, KLK1, AGR2, APOA1, GPA33, ANPEP, CRYBA1, BMI1, BRCA1, MUC16, CEACAM5, CTSE, SCGB1A1, EXO1, FOXA1, GABRP, GGT1, SFN, KRT13, LTF, SLC46A2, KLK10, P2RX7, CDKN2A, TP63, CDH3, PSCA, AGER, ZFP42, SPRR1B, SI, TTR, TM4SF20, TSTD1, SYCN, HBEGF, PIGR, MUC13, SELENBP1, ELF3, TSPAN1, GUCA2A, PHGR1, LYPD8, LGALS4, GATA2, SEC23B, TSPAN8, DLX5, DGAT2, ITPR2, THRSP, PLA2G4A, SLC25A48, PGR, FERMT1, EHF, PLEKHS1, CDKL1, MECOM, MSX1, RNF128, ANLN, CKAP2, HMMR, KIF15, CKAP2L, KIF20B, HIRIP3, INCENP, KIF23, PRC1, ECT2, CXCL10, CXCL8, CCL20, CXCL17, PRG4, ALOX15, F5, EMILIN2, SPTSSB, FMO5, IVL, VSIG2, AQP3, PAQR5, EPCAM, CLDN1, OCLN, MUC5AC |
| Cardiovas-cular cells | ACTA2, CNN1, CALD1, DES |
| Fibroblasts | IL1R1, FAP, FLI1, CELA1, LOX, PDGFRB, P4HA1, UCP2, CCR2, ITGAL, FGR, HCK, TNFRSF1B, PRKCD, ENO3, ABI3, TREML4, PIP4K2A, CD300E, SERPINB10, CTHRC1, TBX18, COL15A1, GJB2, IL34, EDN3, SLC6A13, VTN, ITIH5, LUM, DPT, POSTN, PENK, MMP14, COL6A2, FABP4, ASPN, ANGPTL2, EFEMP1, SCARA5, IGFBP3, COPZ2, DPEP1, ADAMTS5, COL5A1, CD248, PI16, PAMR1, TNXB, MMP2, COL14A1, CLEC3B, IGFBP6, COL5A2, FBN1, MFAP5, FKBP10, PALLD, WIF1, SNHG18, CDH11, PTCH1, ARAP1, FBLN2, IGF1, PRRX1, FKBP7, OAF, COL6A3, CTSK, DKK1, C1S, RARRES2, GREM1, SPON2, TCF21, PCSK6, COL8A1, ENTPD2, CXCL8, CXCL3, IL6, CYP1B1, COL13A1, ADAMTS10, CCL11, ADAM33, COL4A3, COL4A4, LAMA2, ACKR3, CD55, FBLN7, FIBIN, THBS2, NOV, PTX3, MMP3, LRRK1, HGF, FRZB, COL12A1, COL7A1, MEOX1, PRG4, PKD2, CCL19, NNMT, FOXF1, HAS1, CTGF, ERCC1, WISP1, TWIST2, RIPK3, DDR2, ELN, FN1, HHIP, FMO2, COL1A2, COL3A1, VIM, FSTL1, GSN, SPARC, S100A4, NT5E, COL1A1, MGP, NOX4, THY1, CD40, SERPINH1, CD44, PDGFRA, EN1, DCN, CEBPB, EGR1, FOSB, FOSL2, HIF1A, KLF2, KLF4, KLF6, KLF9, NFAT5, NFATC1, NFKB1, NR4A1, NR4A2, PBX1, RUNX1, STAT3, TCF4, ZEB2, LAMC1, MEDAG, LAMB1, DKK3, TBX20, MDK, GSTM5, NGF, VEGFA, FGF2, P4HTM, CKAP4, INMT, CXCL14 |

图示, 示意图

描述已自动生成

**Figure S1. Dot plots for fold enrichment in cluster “airway progenitor cells”.** The GO enrichment visualisation was based on the fold enrichment, calculated as the gene ratio (the fraction of genes within a specific gene set that are differentially expressed) divided by the background ratio (the fraction of differentially expressed genes among all genes examined).

图示

描述已自动生成

**Figure S2. Dot plots for gene ratio in cluster “airway progenitor cells”.** The GO enrichment visualisation was based on the gene ratio, which the fraction of genes within a specific gene set that are differentially expressed.

图表

描述已自动生成

**Figure S3. Dot plots for fold enrichment in cluster “epithelial cells”.** The GO enrichment visualisation was based on the fold enrichment, calculated as the gene ratio (the fraction of genes within a specific gene set that are differentially expressed) divided by the background ratio (the fraction of differentially expressed genes among all genes examined).

图表, 散点图

描述已自动生成

**Figure S4. Dot plots for gene ratio in cluster “epithelial cells”.** The GO enrichment visualisation was based on the gene ratio, which the fraction of genes within a specific gene set that are differentially expressed.

图表

低可信度描述已自动生成

**Figure S5. Dot plots for fold enrichment in cluster “fibroblasts”.** The GO enrichment visualisation was based on the fold enrichment, calculated as the gene ratio (the fraction of genes within a specific gene set that are differentially expressed) divided by the background ratio (the fraction of differentially expressed genes among all genes examined).

图表

中度可信度描述已自动生成

**Figure S6. Dot plots for gene ratio in cluster “fibroblasts”.** The GO enrichment visualisation was based on the gene ratio, which the fraction of genes within a specific gene set that are differentially expressed.