Cell Reports, Volume *40*

Supplemental information

A single-cell analysis of thymopoiesis and thymic iNKT cell development in pigs

Weihong Gu, Darling Melany C. Madrid, Sebastian Joyce, and John P. Driver

Supplemental materials

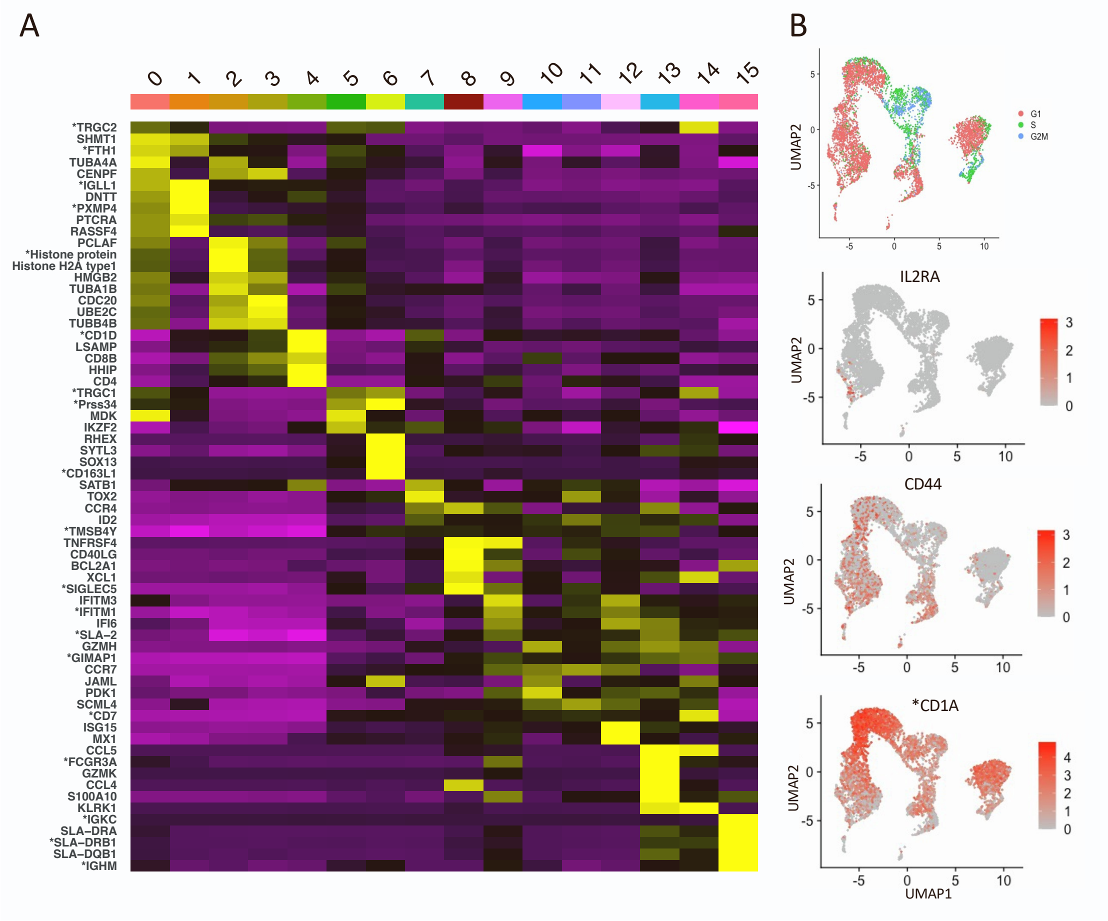

Figure S1. The transcriptional profile of pig thymocytes, related to Figure 1. (A) Heatmap showing row-scaled mean expression of the five most differentially expressed genes per cluster (clusters from Figure 1A). **(B)** UMAP plot showing cell cycling phases in pig thymocytes. The same UMAP plot is used to show cells with high expression of *IL2RA*, *CD44*, and **CD1A*. Asterisks indicate non-annotated genes (described in Table S13).

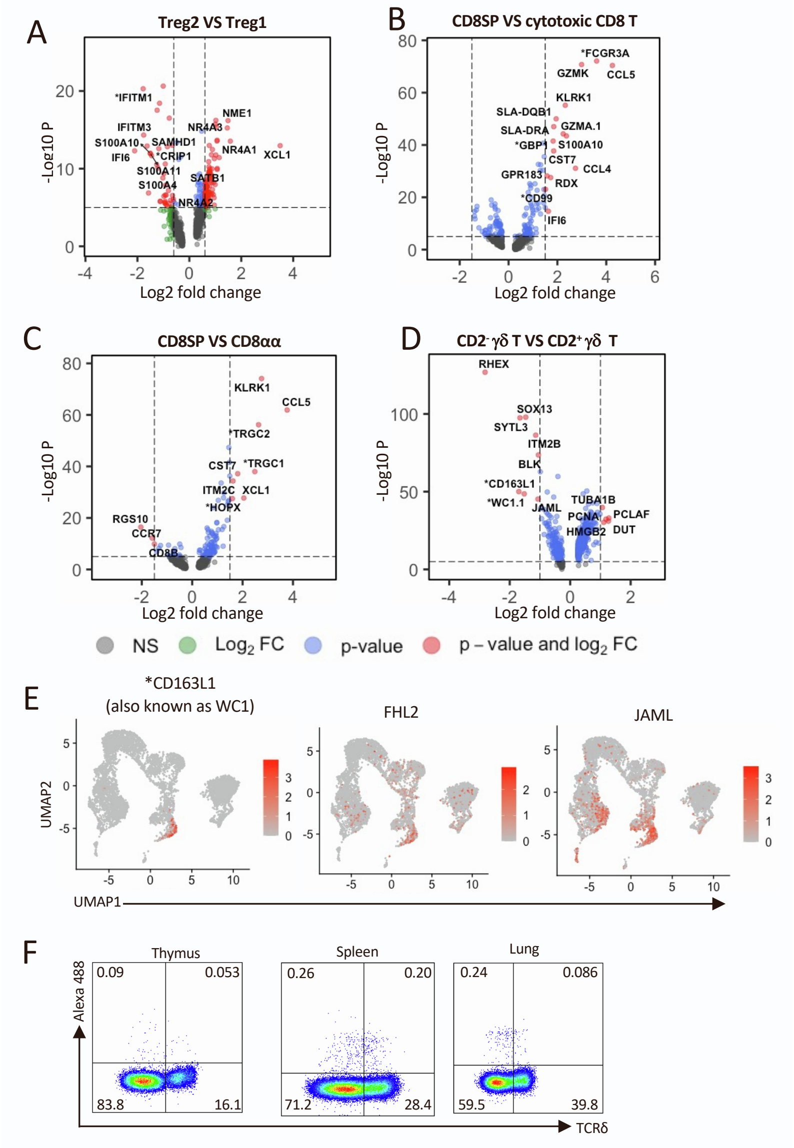

Figure S2. Analysis of differentially expressed genes in unconventional T cells, related to Figure 2. (A-D) Volcano plots showing differentially expressed genes between Treg2 cells and Treg1 cells **(A)**, CD8SP cells and cytotoxic CD8 T cells **(B)**, CD8SP cells and CD8αα T cells **(C)**, and CD2$^-$ γδ T cells and CD2$^+$ γδ T cells **(D).** The x-axis and y-axis respectively represent log2(fold change) and -log10(p-value). **(E)** UMAP plot showing the expression pattern of selected marker genes associated with CD2$^-$ γδ T cells. **(F)** No primary antibody controls for JAML staining. Thymocytes, splenocytes, and lung leukocytes were membrane labeled with antibodies against CD3, TCR1 δ chain, and rabbit IgG, and analyzed according to the Method Details. Asterisks indicate non-annotated genes (described in Table S13).

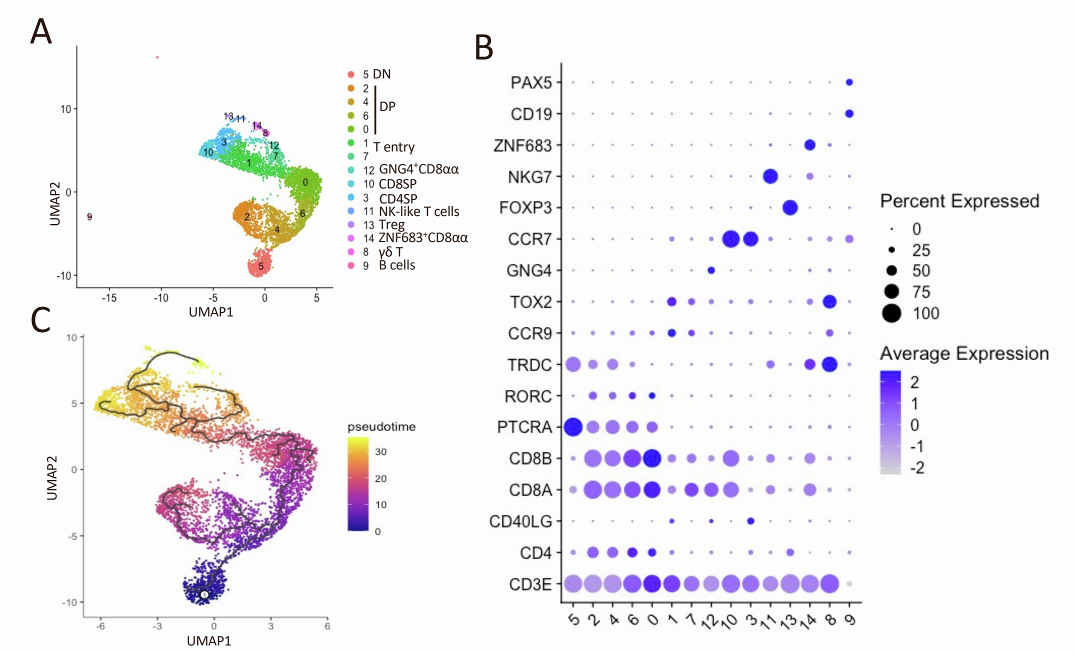

Figure S3. scRNA-seq analysis of human thymocytes, related to Figure 4. (A) UMAP plot of human CD34$^-$ thymocytes from the original public dataset (Le et al., 2020). **(B)** Dot plot showing the *Z*-scored mean expression of selected marker genes in clusters from (A). **(C)** Pseudotime trajectory of human thymocyte development created in Monocle 3 using clusters from (A), excluding γδ T cells and B cells.

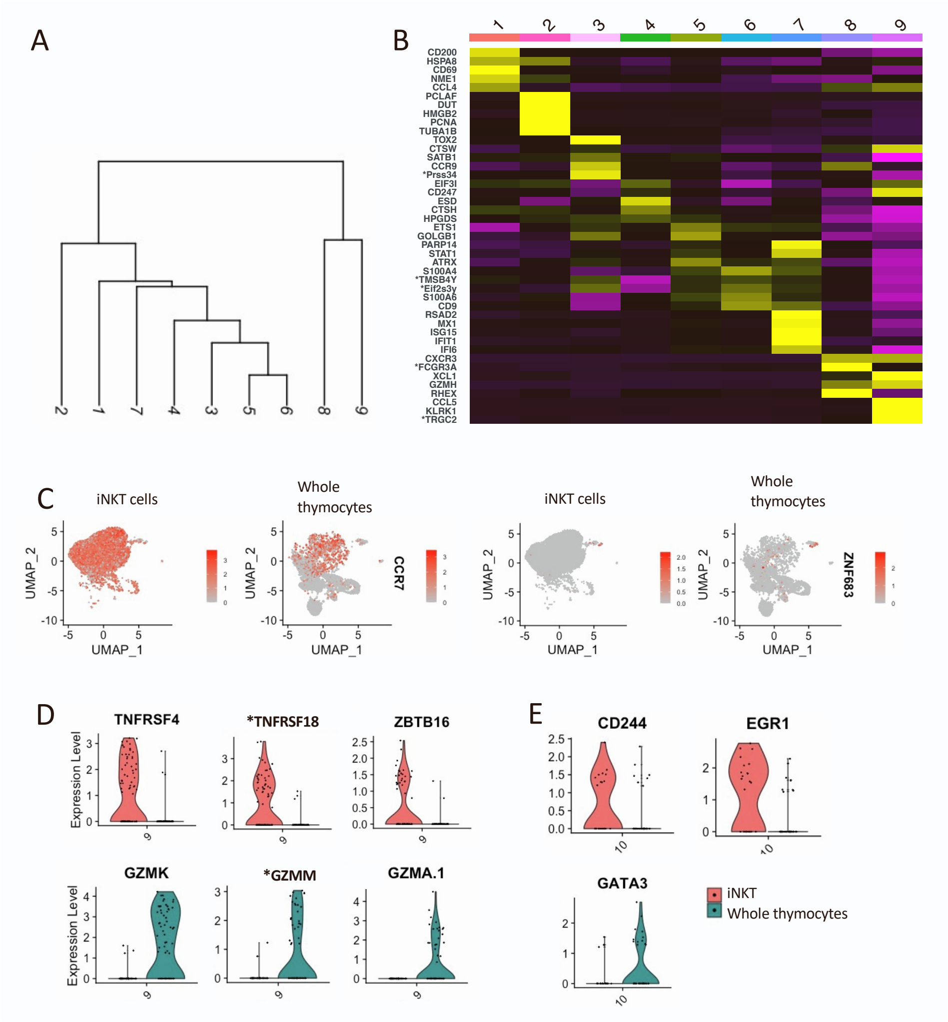

Figure S4. Transcriptional profile of pig iNKT cells and integrative analysis between iNKT cells and whole thymocytes, related to Figure 5. (A) Dendrogram showing similarities among iNKT cell clusters. **(B)** Heatmap showing row-scaled mean expression of the five most differentially expressed genes per cluster (clusters from Figure 5B). **(C)** UMAP plots showing pig iNKT cells and whole thymocytes (from integrative data in Figure 5E) with high expression of *CCR7* and *ZNF683*. **(D-E)** Violin plots showing expression of selected genes that are differentially expressed between iNKT-sw1 and cytotoxic CD8 T cells (cluster 9 in Figure 5E) **(D)**, and between iNKT-sw2 and CD8αα T cells (cluster 10 in Figure 5E) **(E)**. Asterisks indicate non-annotated genes (described in Table S13).

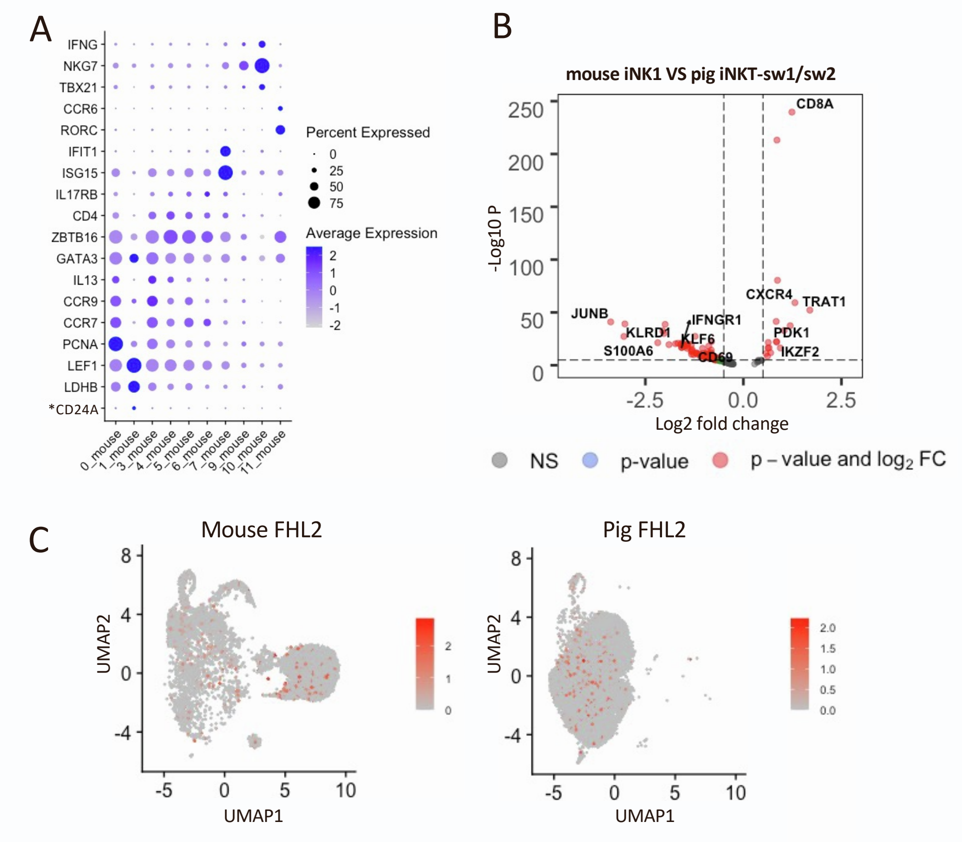

Figure S5. Comparisons of gene expression between mouse and pig iNKT cells, related to Figure 6. (A) Dot plot showing the *Z*-scored mean expression of selected marker genes in mouse iNKT cell clusters displayed in Figure 6A. **(B)** Volcano plot (from Figure 6A) showing differentially expressed genes between mouse iNKT1 cells (clusters 9 and 10 combined) and pig iNKT-sw1/sw2 cells (cluster 8) from Figure 6A. **(C)** UMAP showing the expression of *FHL2* in mouse and pig iNKT cell clusters displayed in Figure 6A. Asterisks indicate non-annotated genes (described in Table S13).

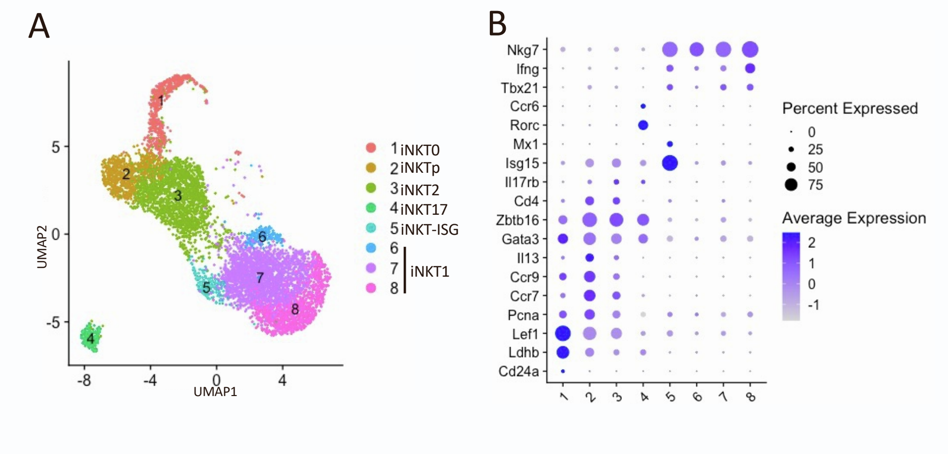

Figure S6. scRNA-seq analysis of mouse iNKT cells, related to Figure 7. (A) UMAP plot of murine iNKT cells from the original public dataset (Harsha Krovi et al., 2020). Corresponds to the mouse dataset used to generate Figure 7. **(B)** Dot plot showing the *Z*-scored mean expression of selected marker genes in mouse iNKT cell clusters displayed in A.

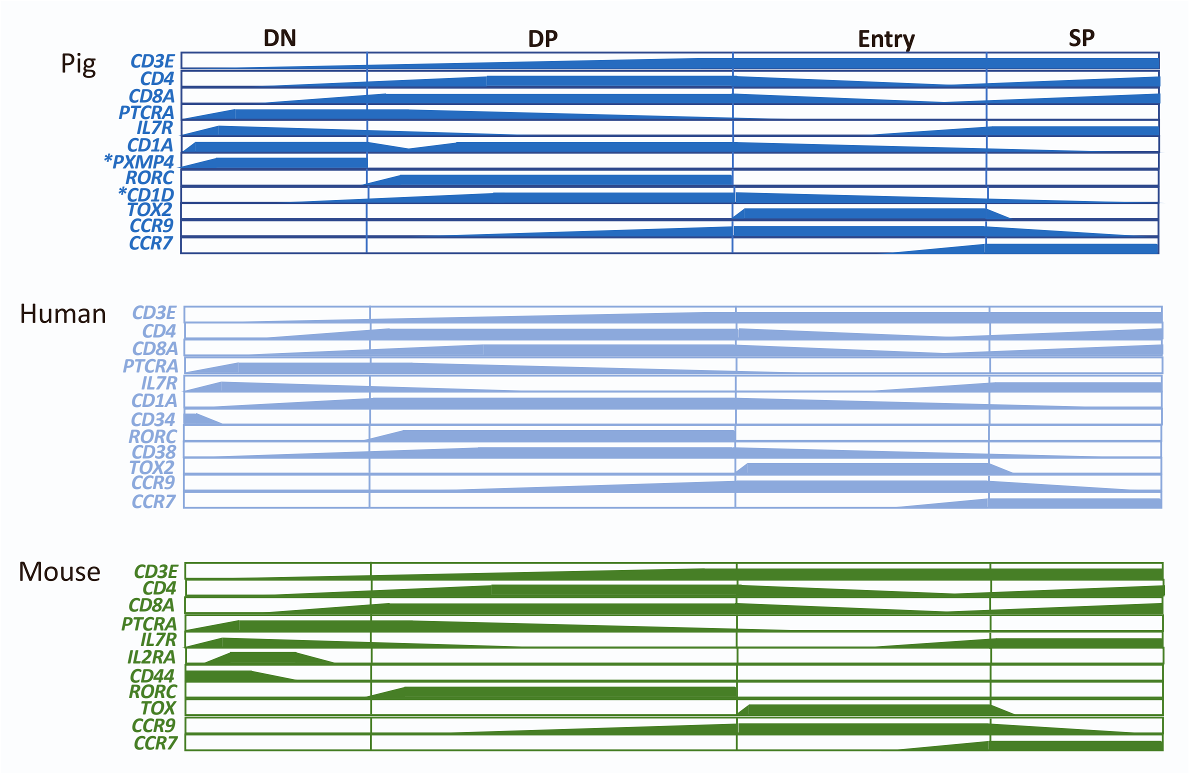

Figure S7. Stage-specific patterns of classical thymocyte development markers in pigs, humans, and mice, related to Figures 1 and 4.

Table S13. Non-annotated genes, related to Figures 1-7, S1, S2, S4, and S5. For genes displayed in the text/figures that were not annotated in the Sscrofa 11.1 reference genome, we replaced the original Ensembl IDs with the most recent available gene symbol or gene description in NCBI (https://www.ncbi.nlm.nih.gov/gene/) or the homologous human or mouse gene symbol from Ensemble (https://www.ensembl.org/biomart/martview/). The gene annotated as *HLA-DRA* was converted to the pig-specific gene symbol *SLA-DRA*.

Ensembl IDs	**Replaced gene symbol**
ENSSSCG00000026217	KLRD1
ENSSSCG00000038351	TNFRSF18
ENSSSCG00000036618	FCGR3A
ENSSSCG00000005095	PRKCH
ENSSSCG00000034625	PXMP4
ENSSSCG00000006452	CD1D
ENSSSCG00000025644	CD1A
ENSSSCG00000009240	PLAC8
ENSSSCG00000033465	HIVEP3
ENSSSCG00000032810	GZMM
ENSSSCG00000009746	RAN
ENSSSCG00000029668	IL2RB
ENSSSCG00000029596	T-cell receptor gamma chain C region 1-like (TRGC1) (homologous to human TRGV10 and mouse Trgc4)
ENSSSCG00000023584	T-cell receptor gamma-2 chain C region-like (TRGC2) (homologous to human TRGV9)
ENSSSCG00000001229	SLA-2
ENSSSCG00000033909	GIMAP1
ENSSSCG00000001396	SLA-8
ENSSSCG00000037364	CD7
ENSSSCG00000022322	BCL2L11
ENSSSCG00000016737	PPIA
ENSSSCG00000021591	RBM3

ENSSSCG00000020808	CCT6A
ENSSSCG00000011676	ATP1B3
ENSSSCG00000014031	HNRNPAB
ENSSSCG00000037142	CRIP1
ENSSSCG00000014565	IFITM1
ENSSSCG00000024973	GBP1
ENSSSCG00000048323	CD99
ENSSSCG00000008898	HOPX
ENSSSCG00000034914	CD163L1 (also known as WC1)
ENSSSCG00000031085	Antigen WC1.1
ENSSSCG00000008032	Prss34
ENSSSCG00000033734	TMSB4Y
ENSSSCG00000012178	Eif2s3y
ENSSSCG00000031023	Cd24a
ENSSSCG00000032591	IFITM1-like
ENSSSCG00000014540	FTH1
ENSSSCG00000010077	IGLL1
ENSSSCG00000033550	Histone domain-containing protein
ENSSSCG00000033451	Histone H2A type 1
ENSSSCG00000003236	SIGLEC5/ SIGLEC14
ENSSSCG00000036224	IGKC
ENSSSCG00000001455	SLA-DRB1
ENSSSCG00000037775	IGHM
ENSSSCG00000014997	CASP4
ENSSSCG00000027196	GIMAP6
HLA-DRA	SLA-DRA