

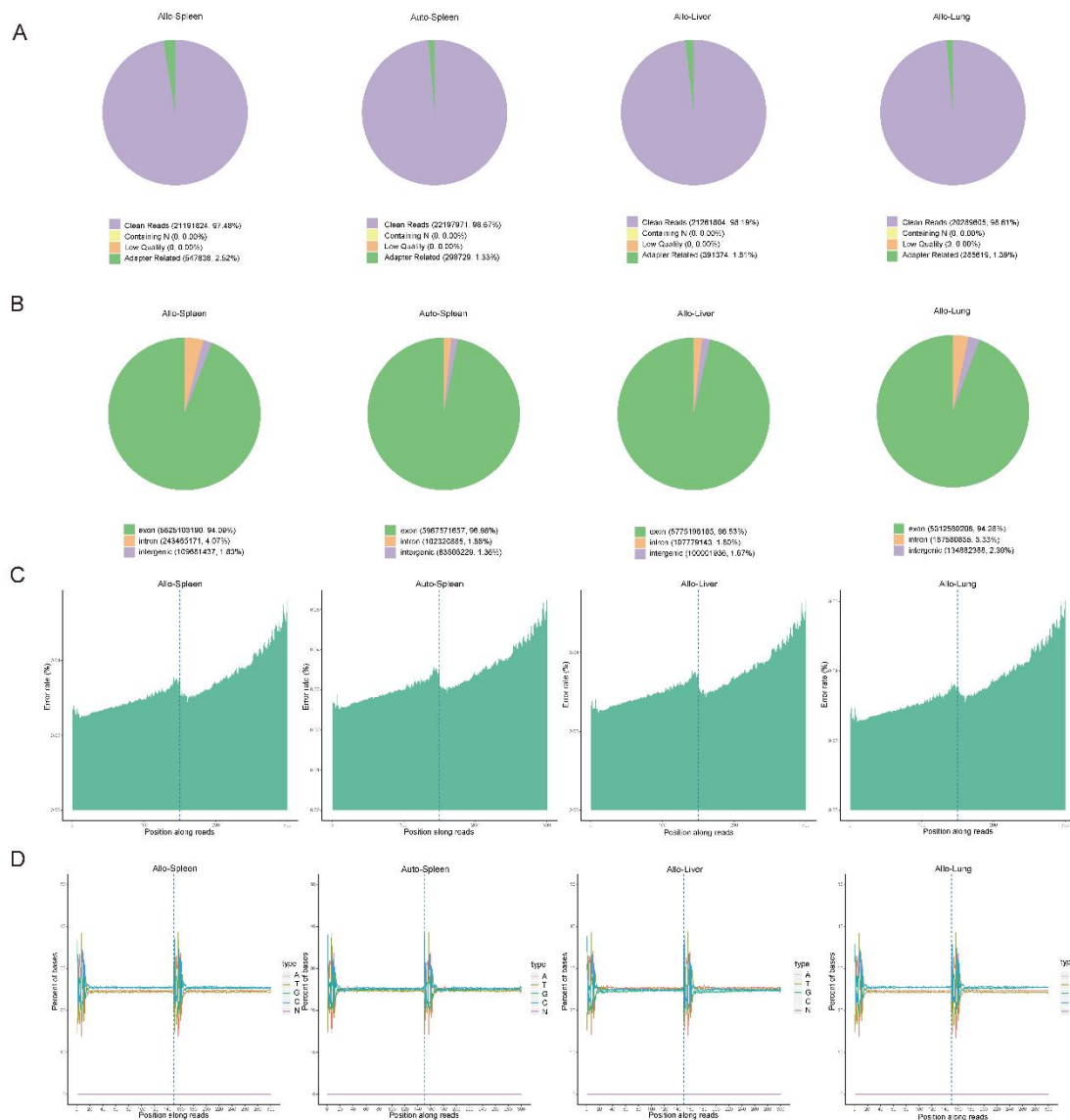
# Supplementary Figures

## **The potential genes mediate the pathogenicity of allogeneic CD4<sup>+</sup>T cell in aGVHD mouse model**

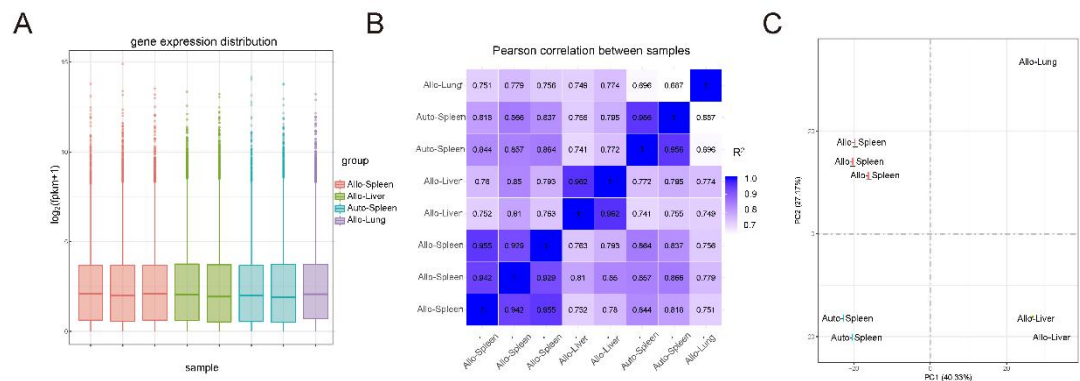
Zheng-yu Yu<sup>1</sup>, Chen-chen Qin<sup>1</sup>, Min Cao<sup>1</sup>, Xiao-ya He<sup>1</sup>,  
Han-yun Ren\* & Hui-hui Liu\*

1. Department of Hematology, Peking University First Hospital, Beijing, China

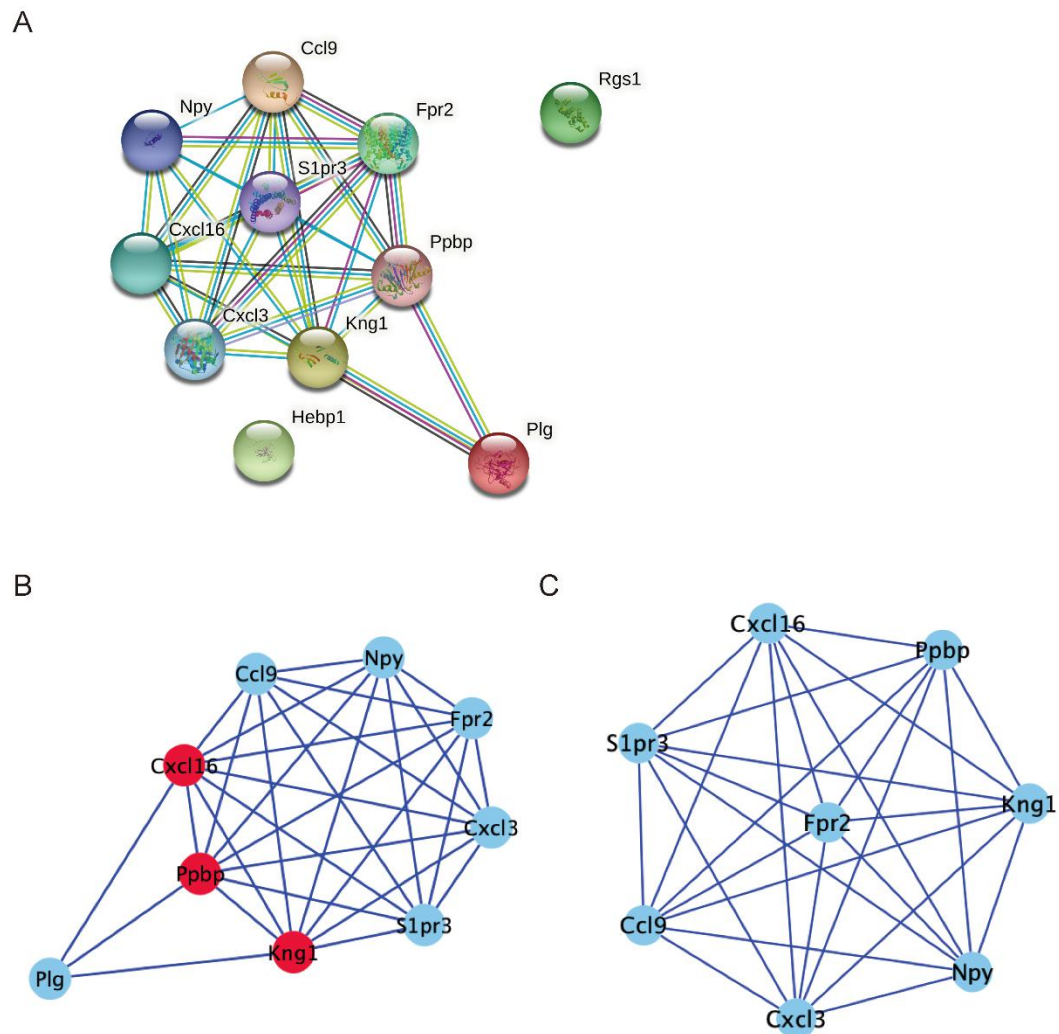
\*Corresponding author. Address: Department of Hematology, Peking University First Hospital, No. 8 Xishiku Street, Beijing 100034, China. Tel: 86-10-83572526.



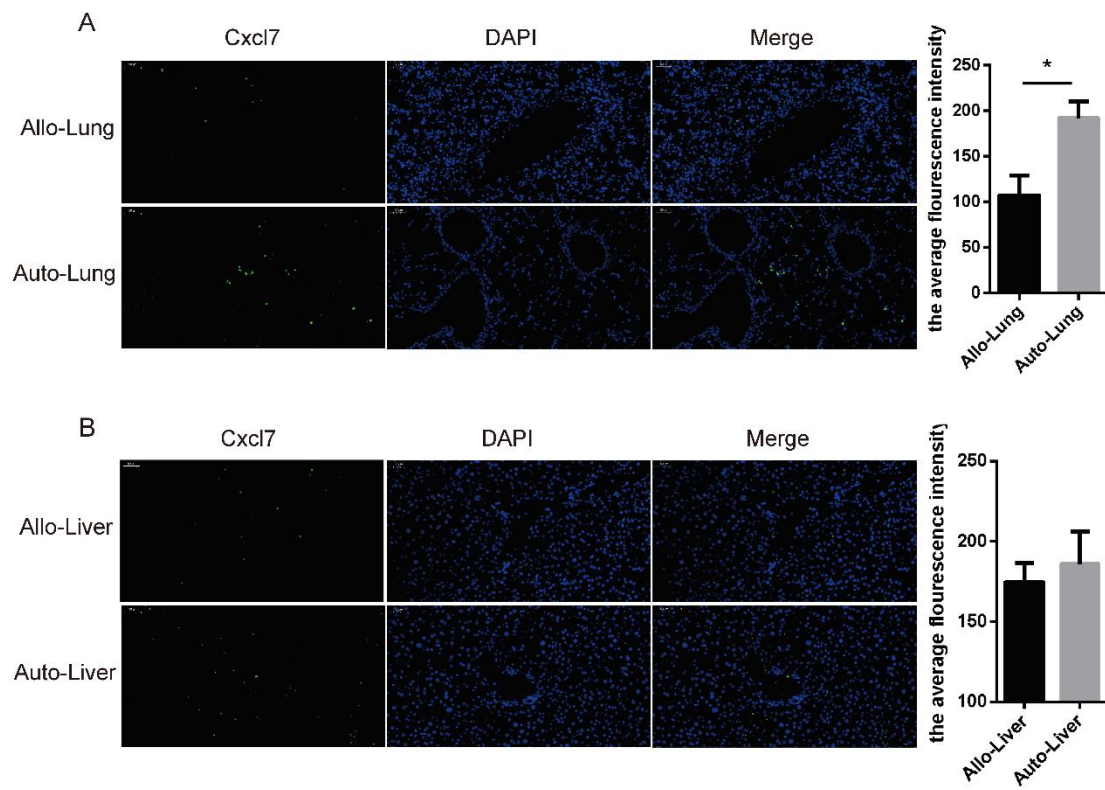
**Supplementary Figure S1.** The quality of sequencing data. **(A)** Classification of raw reads. Clean reads for the percent of clean reads; Adapter related for the percent of reads with adapter; Containing N for the percent of reads with N bases; Low quality for percent of reads with low sequencing quality. **(B)** Percent of genome regions. Exon for the number of exons reads and the percent to the clean reads; intron for the number of intron reads and the percent to the clean reads; intergenic for the number of intergenic regions and the percent to the clean reads. **(C)** Error rate distribution along reads. **(D)** Bases content along reads. The base distribution of two reads of double-terminal sequencing sequence was presented on the left and right sides.



**Supplementary Figure S2.** The quantitative analysis of sequencing data. **(A)** The gene expression distribution. FPKM corrected for sequencing depth and gene length. The maximum, upper quartile median, lower quartile, and minimum (top to down) values of gene expression in each sample were presented. **(B)** The Pearson correlation between samples. The  $R^2$  value and their expression pattern were positive correlation. **(C)** Principal component analysis. PCA used the calculation method of linear algebra to reduce the dimension and extract the principal components of tens of thousands of gene variables.



**Supplementary Figure S3.** Bioinformatics analysis of protein quantitative identification in Allo-Spleen, Auto-Spleen, Allo-Lung and Allo-Liver. (A) PPI network of genes related to immunoregulation. *Pbp* was the same as *Cxcl7*. (B) The hub genes were highlighted in red. (C) The module comprises 8 nodes and 28 edges.



**Supplementary Figure S4.** Quantification of the histology staining shown. (A) The left is one representative section per group. The right is the number of Cxcl7 per scale in the liver and lung respectively. Cxcl7 were detected by colocalization of Cxcl7 (green) and dapi(blue). Cxcl7 deposition was quantified on a 0-3 scale to determine the amount of antibody in the tissues.