

Blood prognostic biomarker signatures for hemorrhagic cerebral cavernous malformations
(CCMs)

Jacob Croft^{1*}, Brian Grajeda^{2*}, Luis A Aguirre², Johnathan S Abou-Fadel¹,
Cameron C Ellis², Estevao Igor², Igor C. Almeida², Jun Zhang^{1,++}

Department of Molecular & Translational Medicine (MTM)¹
Texas Tech University Health Science Center El Paso (TTUHSCEP), El Paso, TX 79905 USA
Department of Biological Sciences², University of Texas at El Paso, El Paso, TX 79902 USA

Supplemental Materials

***Shared co-first authorship:** Authors who have worked together and contributed equally to this project.

⁺⁺All correspondence:

Jun Zhang, Sc.D., Ph.D.

Department of Molecular and Translational Medicine (MTM)

Texas Tech University Health Science Center El Paso

5001 El Paso Drive, El Paso, El Paso, TX 79905

Tel: (915) 215-4197, Email: jun.zhang2000@gmail.com

Supplemental Legends

Figure S1. A biological function analysis on the number of expressed proteins by CCM1 subjects compared to normal controls using group gene ontology (GO).

A-B. Group Gene Ontology (GO) analysis and pathway analysis were performed for biological profiling and systematic functional annotation of the identified group of DEPs in order to determine the statistical significance of a biological process. These key processes indicate the presence of proteins, but their effect on the study subject in both human (**A**) and mouse (**B**) cannot be determined solely based on this information. Supplemental figure 2 provides similar trends in molecular function (Suppl. Fig. 2A, 2C) and cellular processes (Suppl. Fig. 2B, 2D).

C-D. To confirm whether our identified DEPs are frequently associated with specific biological functions, GO enrichment analysis was performed on three distinct subtypes of GO: biological process, molecular function, and cellular component. The molecular function and cellular component subtypes can be observed in supplemental figure 2. It's important to note that these key nodes or processes indicate the presence of proteins, but their positive or negative effect on both human (**C**) and mouse (**D**) study subjects cannot be determined solely based on this information. **E-F.** Volcano plots exhibit differentially expressed proteins (DEPs) between CCM1-deficient subjects and healthy controls. This visualization illustrates proteins present in both control and confirmed CCM-deficient subjects, showcasing differentiations in protein expression in subject with deficient CCM1 compared to normal controls. A p-value, computed from a Student's t-test, is displayed using a $-\log_{10}$ transformation relative to the \log_2 fold change. This method helps to regulate infinite fold values for unique observations, thereby preventing distortion in the graph. The labeled data points within the plot represent fold changes in protein expressions for CCM1-deficient subjects compared to the control group in both humans (**E**) and mice (**F**). The significance of up- and down-regulated proteins from figures E and F was refined using a $-\log_{10}$ p-value threshold to

identify more significant findings.

Figure S2. Group Gene Ontology (GO) analysis on the molecular function and cellular component of the expressed proteins in both human and mouse models with CCM1 deficiency, compared to their respective healthy controls.

A. The group analysis of molecular function in the human CCM1 subject indicated that the majority of the expressed proteins were involved in binding processes. However, it's important to note that the number of proteins measured does not reflect whether these processes are positively or negatively regulated. Nonetheless, the higher number of proteins may suggest an increased demand due to decreased function observed in CCM1 diagnosis. **B.** The cellular component analysis of the human CCM1 subject revealed that the majority of the expressed proteins are involved in cellular anatomical entity. It's important to note that the number of proteins produced does not indicate whether their effects are positive or negative. However, we can speculate that the increased demand for these proteins may be a result of the observed changes in CCM1 diagnosis. **C.** Similar to the human CCM1 subject (Supp. Fig. 2A), the molecular function group analysis of the mouse CCM1 subject showed that the largest number of expressed proteins were involved in binding actions. However, it's important to note that these numbers do not reflect whether the effects of these proteins are positive or negative. Nonetheless, the increased demand for these proteins could be due to the negative effects observed in CCM1 diagnosis. **D.** Similarly, the cellular component group analysis of the mouse CCM1 subject showed that, like the affected human (Supp. Fig. 2B), most of the expressed proteins were involved in the cellular anatomical entity. **E.** This panel depicts the pathway of protein activity within the molecular function pathways of a human CCM1 patient, with highlighted nodes indicating where the measured proteins from the

collected samples interact with the pathway. It provides an expanded view of how the binding processes from Supp. Fig. 2A are affected. **F.** This panel illustrates the pathway of protein activity within the cellular components of cells in a human CCM1 subject, highlighting the affected areas that support the CCM1 diagnosis, as well as where the abundant proteins from Supp. Fig. 2B are affecting the cellular components. **G.** This molecular function pathway for the CCM1 mouse illustrates the areas where the harvested proteins affect the pathway, similar to what was observed in the human subject. The nodes represent the different stages of protein activity. As seen in both the mouse and human subjects, the binding activity is where these proteins are predominantly represented. Impaired binding processes are indicative of the CCM1 diagnosis, thus supporting the diagnosis in both laboratory and human subjects through replication. **H.** This cellular component pathway for the CCM1 mouse subject illustrates the areas where the harvested proteins affect the cellular components. These areas, if negatively affected, would support the diagnosis of CCM1, as they are the impaired areas anticipated by the cellular malformations observed in CCM1 diagnosis. The nodes represent different stages of protein activity within the pathway.

Figure S3. GSEA analysis demonstrates the pathways affected in both human and mouse CCM1 subjects.

A. The gene set enrichment analysis upset plots illustrate the pathways that are affected in the CCM1 human subject. The affected pathways support the diagnosis of CCM1 subjects by showing negative regulation of organismal processes, negative responses to binding of cells as observed in wound healing, and negative effects on the organization of cellular processes. This is

consistent with the biological process group GO as shown in supplemental figures 1A. **B.** This gene set enrichment analysis upset plot displays the pathways that are affected in the mouse CCM1 subject. Similar to the human subject, there is a decreased function of the membrane, which supports the diagnosis of CCM1 as the malformation of the plasma membrane results in oblong formations observed in this diagnosis.

Figure S4. GSEA analysis identified enriched pathways for the DEPs in the CCM1-deficiencies compared to their normal conditions, for both human and mouse subjects.

A. The pathways identified in the human CCM1 subjects were compared to age and gender-matched healthy controls, revealing negative regulation of all enriched pathways that are known to be affected by CCM1 deficiency. This supports the previous upset plots that displayed the negatively affected pathways, as well as providing additional insight into the results from the group GO molecular function and cellular components observed in Supp. Fig. S1A-B. **B.** The pathways identified through GSEA analysis of CCM1 mouse compared to a healthy control showed negative effects that signify the affected processes and further support the CCM1 diagnosis.

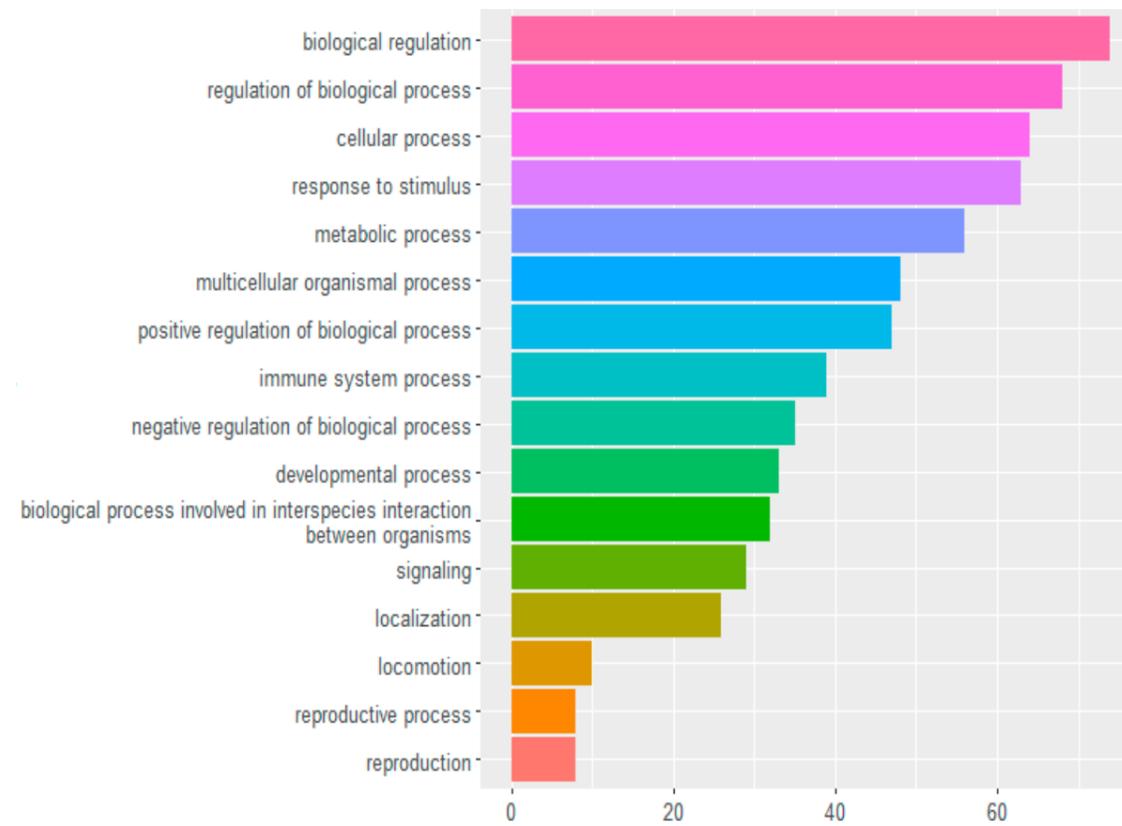
Figure S5. Enriched KEGG pathways reaffirm the existence of complement and coagulation cascades caused by CCM1 deficiency in both Human and Mouse subjects.

A. The KEGG enrichment pathways for the Human enriched genome shows the complement and coagulation, including ECM components. While this was identified as a KEGG pathway, it

in the human CCM1 patient it supports the degraded binding suggested by the GO enrichment where binding proteins were found (Fig. S2A). **B.** Similarly, the KEGG enrichment pathway analysis for the mouse Ccm1 subject highlighted the affected complement and coagulation cascades. This further supports our previous findings and strengthens the diagnosis of CCM1 deficiency.

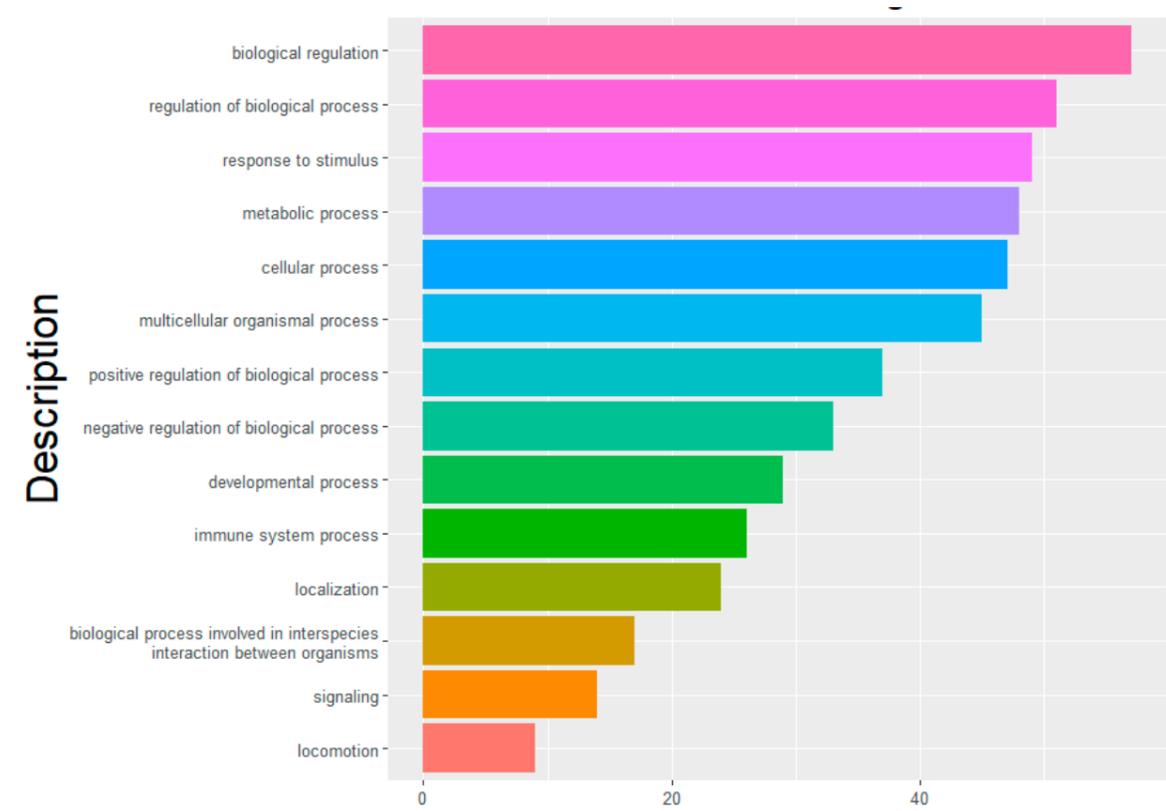
S1A

CCM1 vs Ctrl in Human



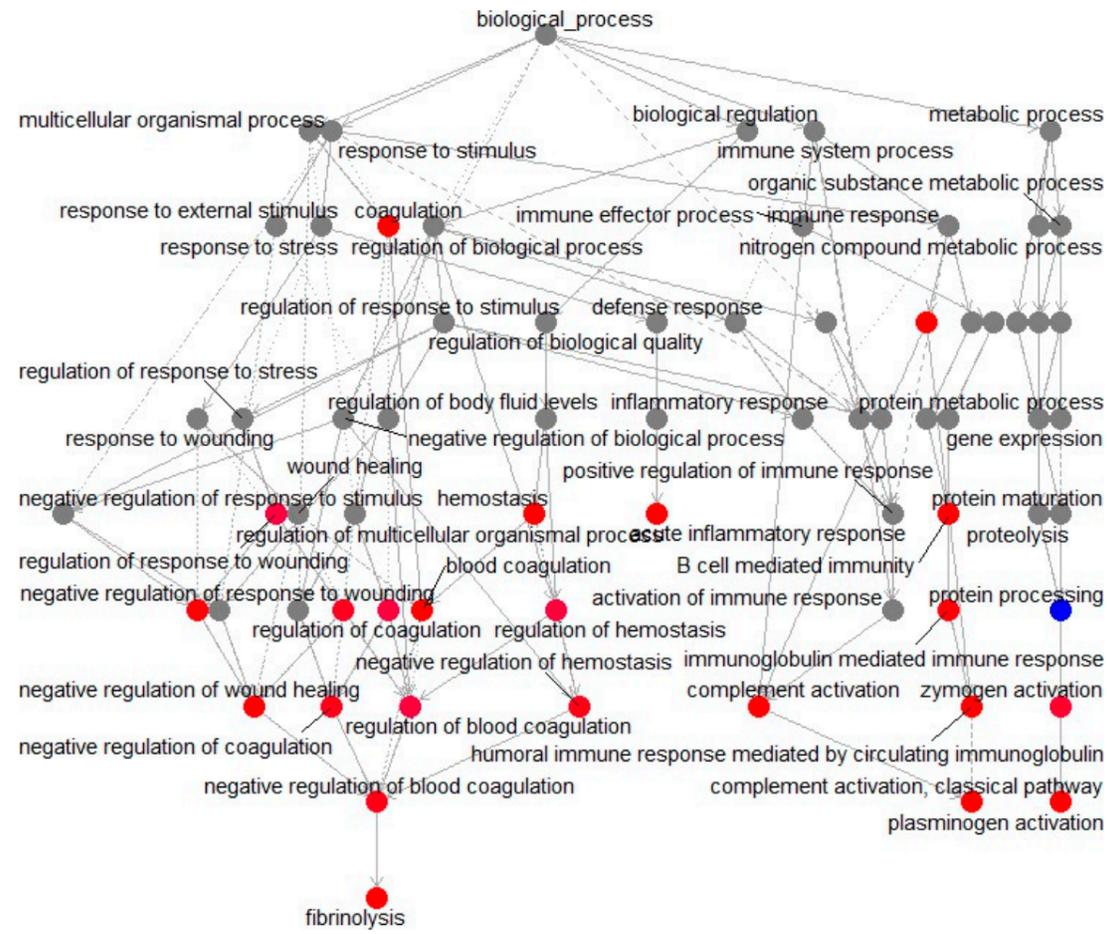
S1B

Ccm1 vs Ctrl in mouse

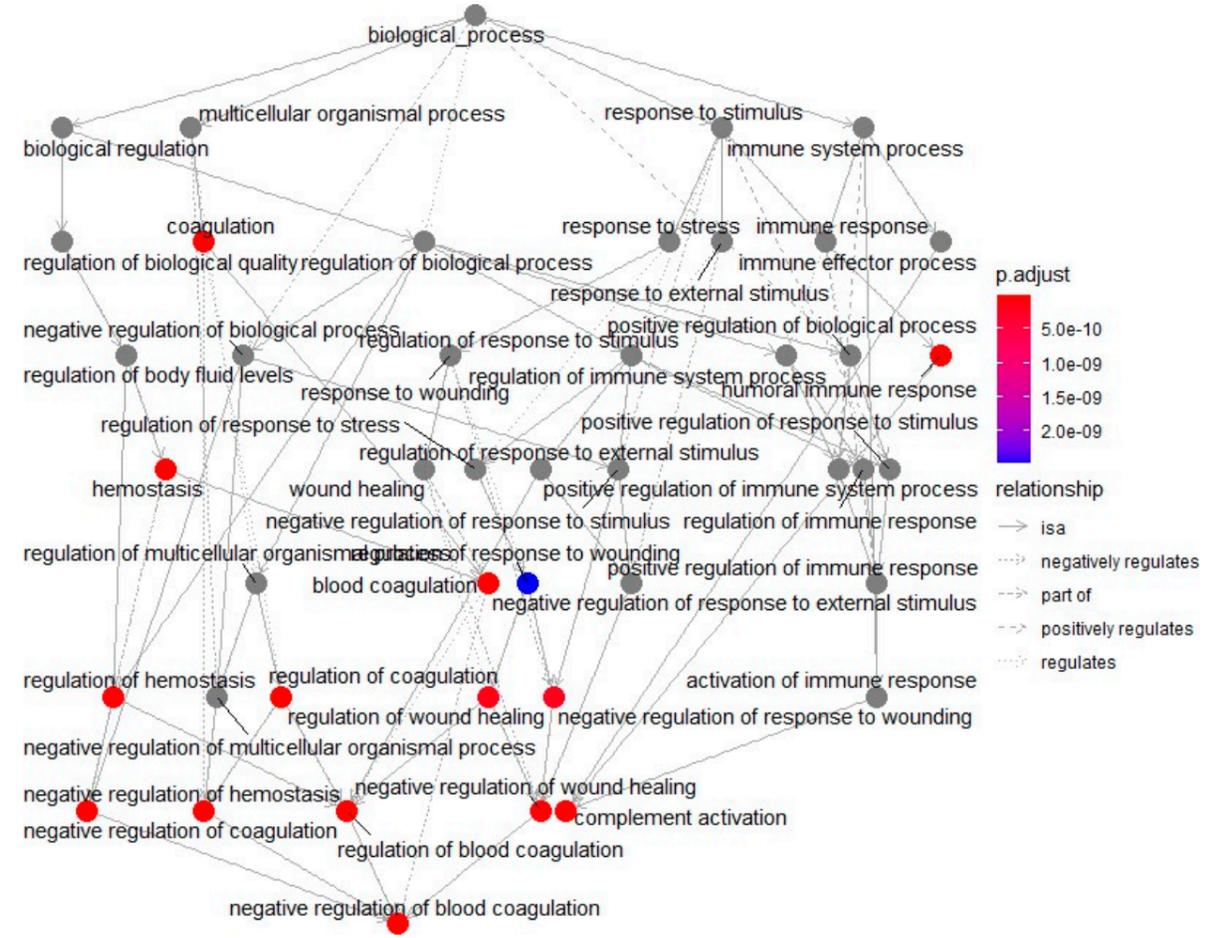


1. Group-GO Biological Process

S1C Human

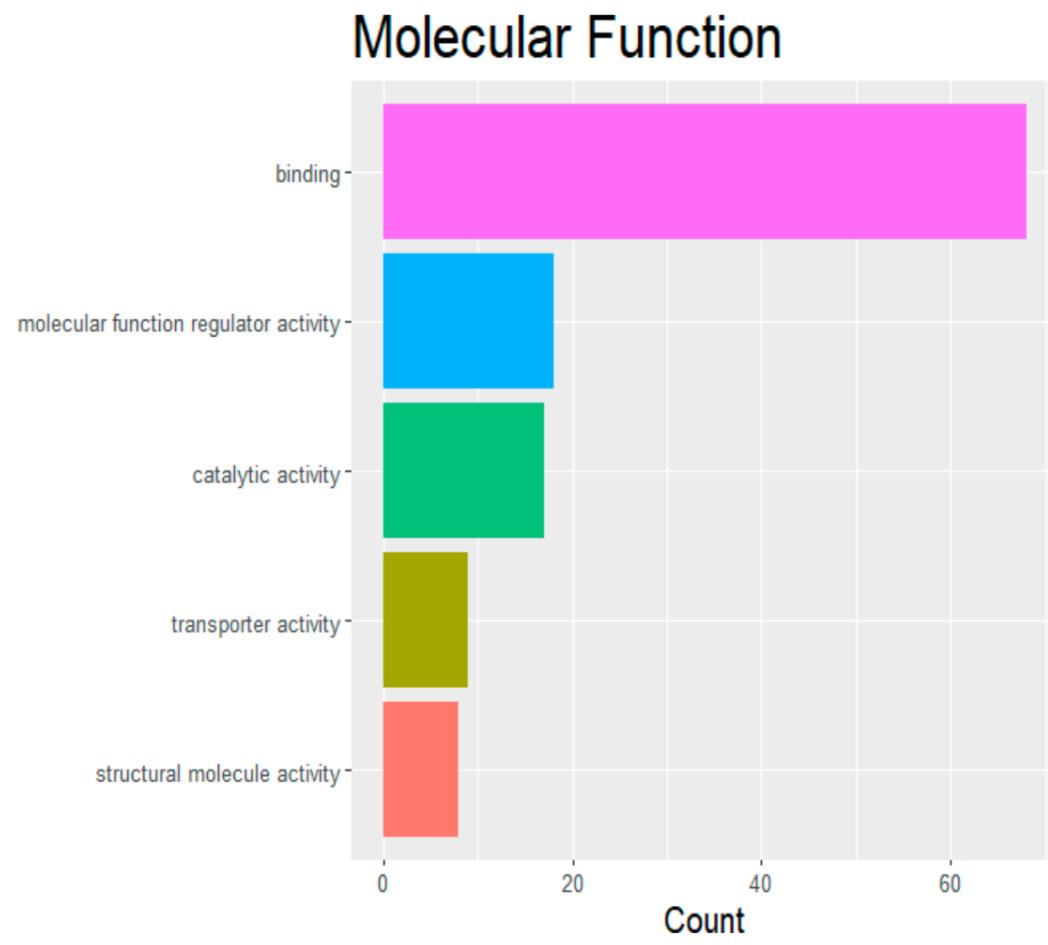


S1D Mouse

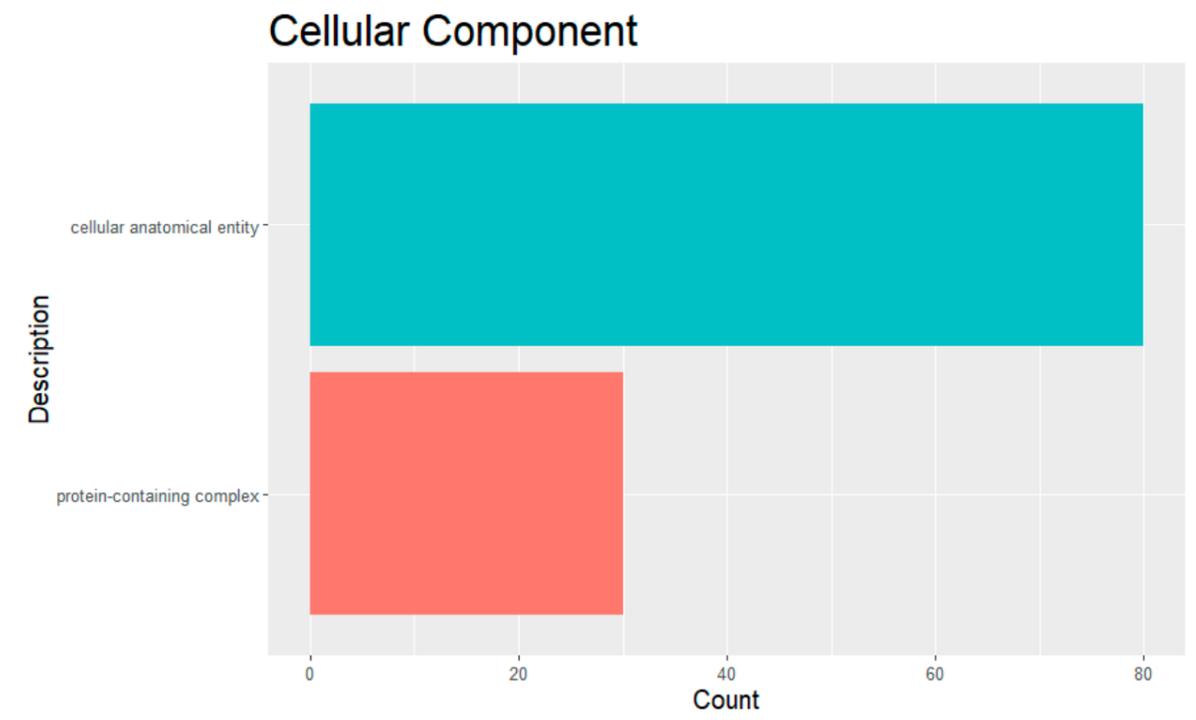


Enriched GO-Biological Process

S2A

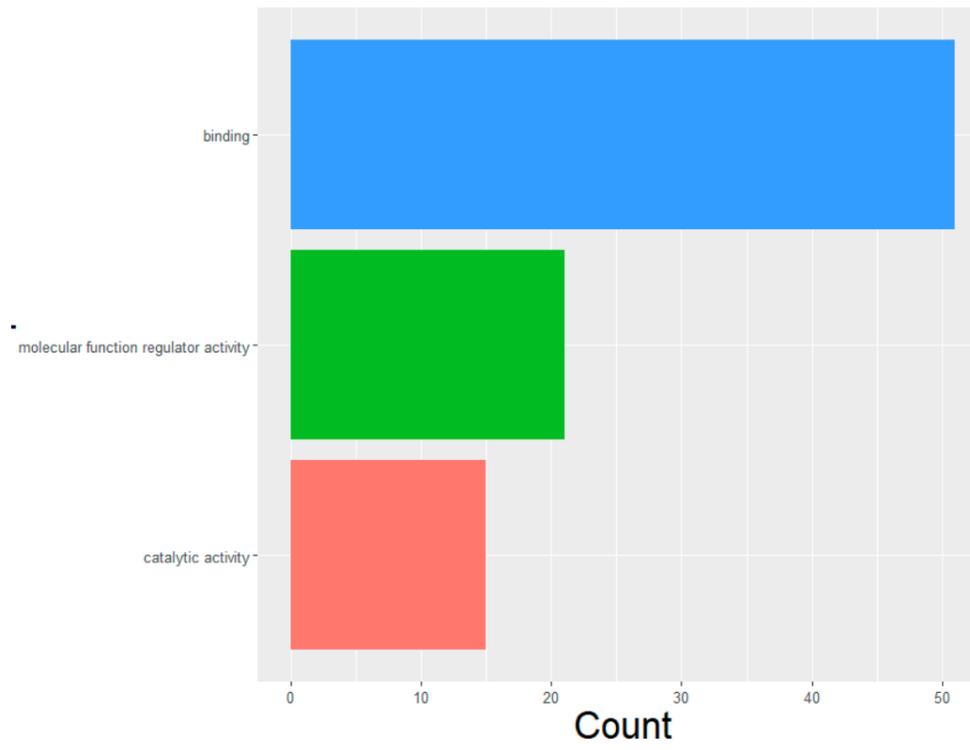


S2B.



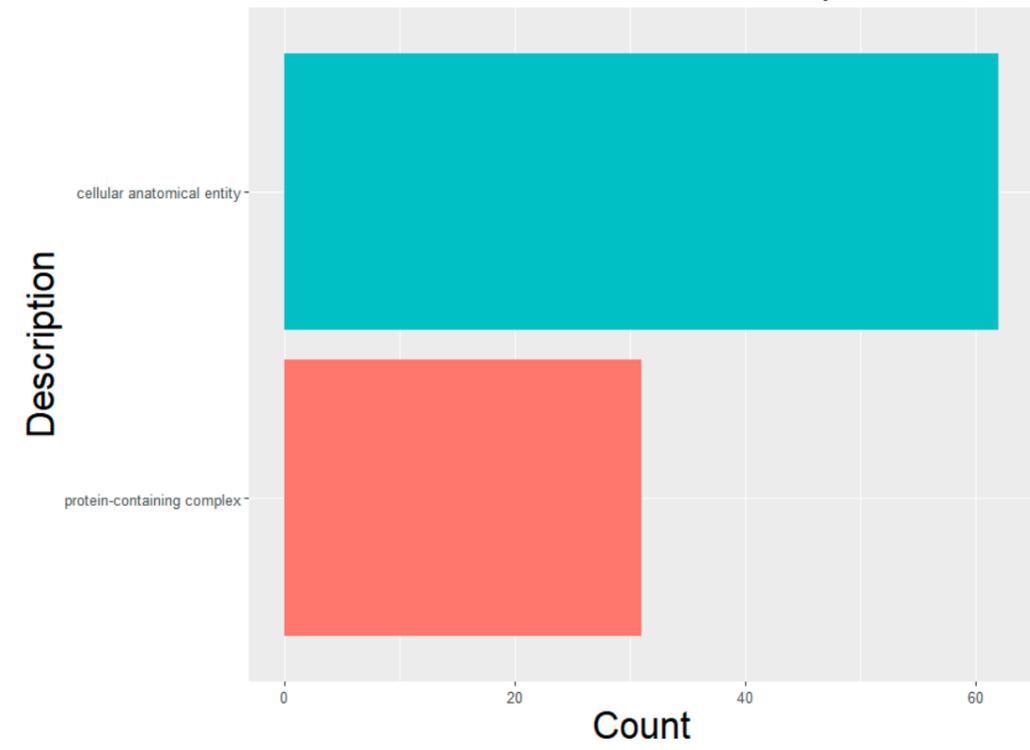
S2C.

CCM1 Vs. Ctrl Mouse Molecular Function

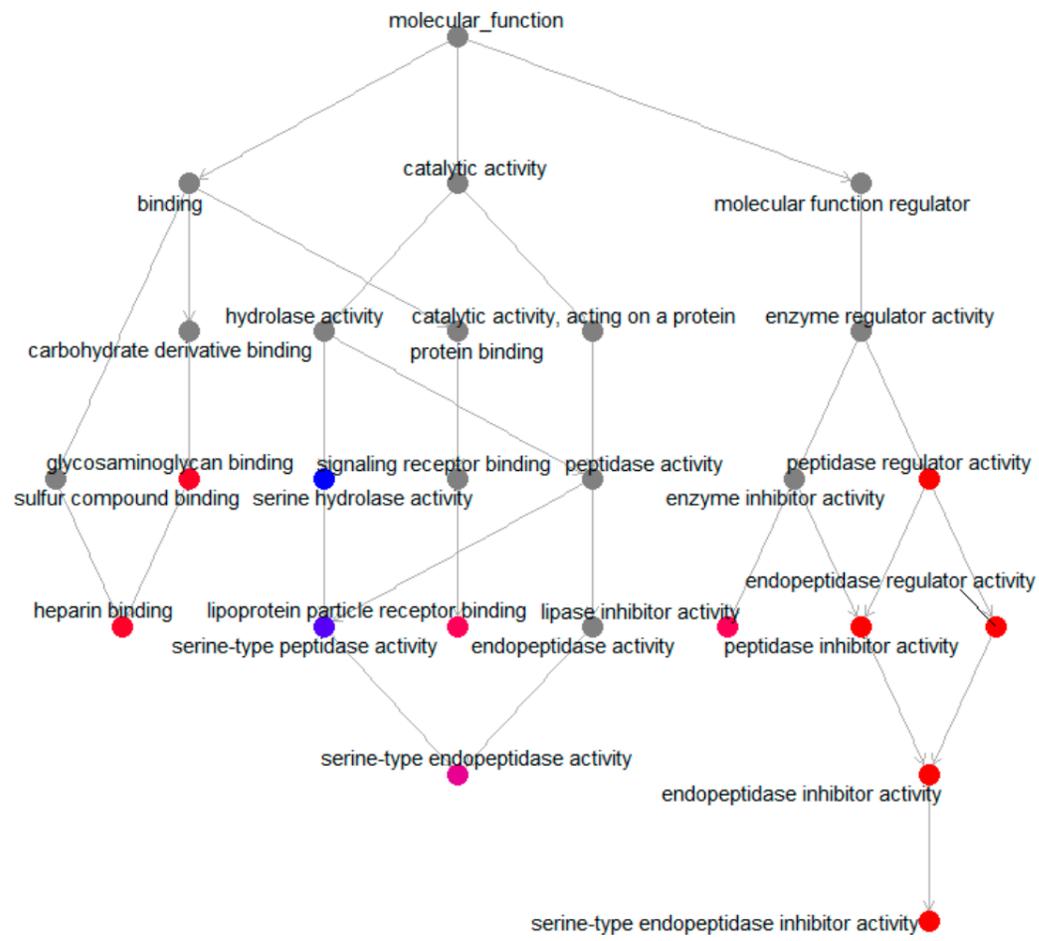


S2D.

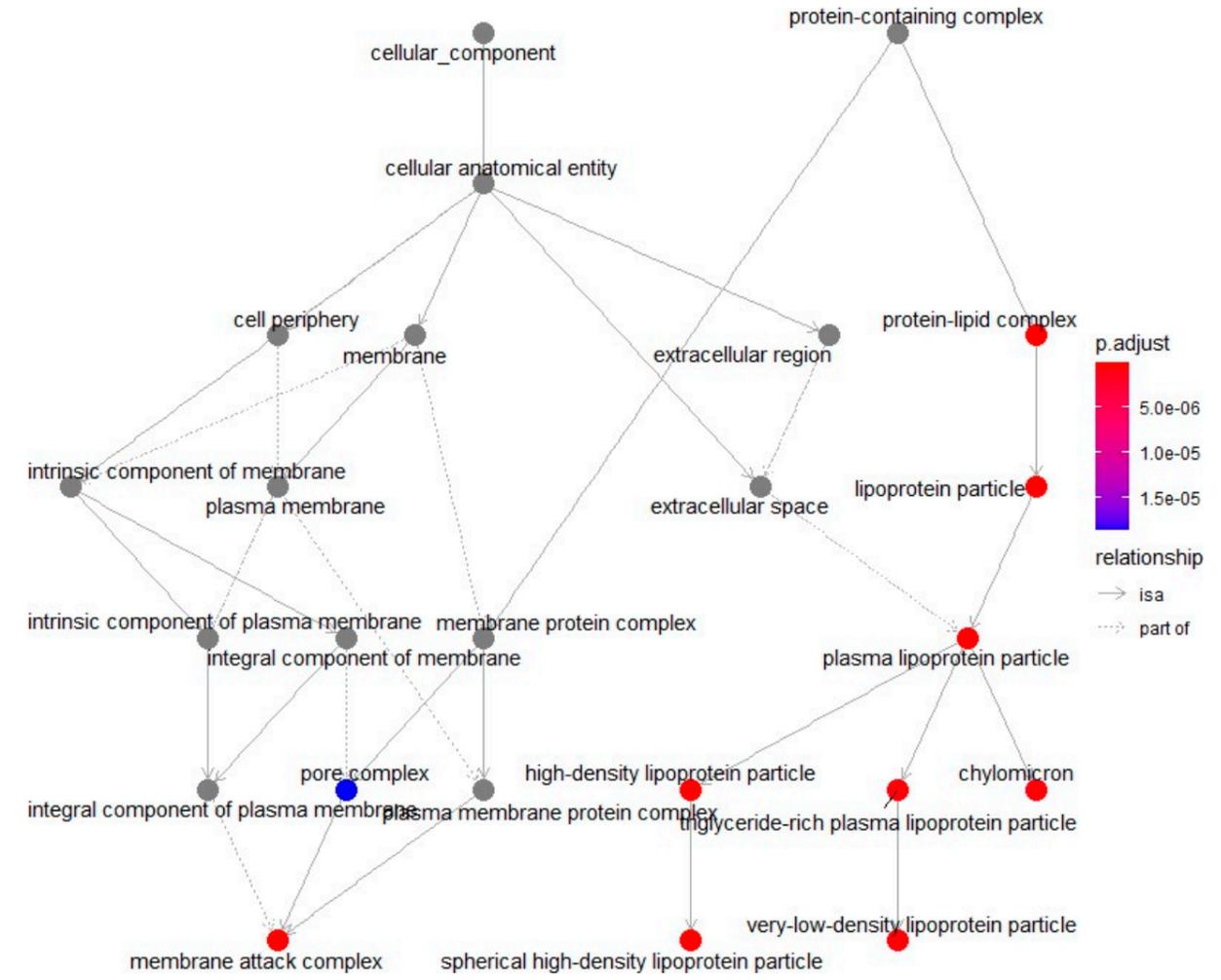
CCM1 vs. Ctrl Mouse Cellular Component



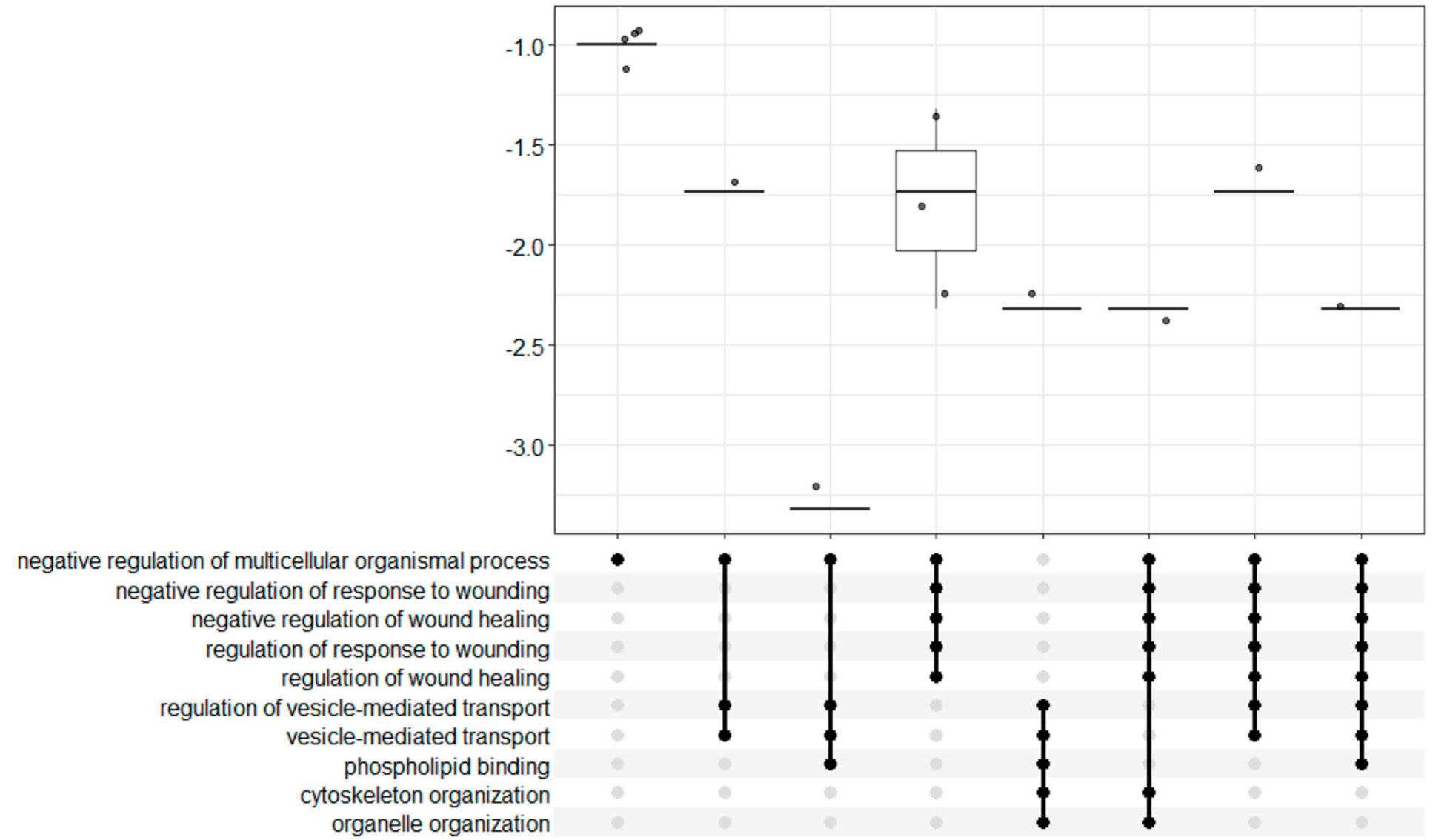
S2G. Mouse



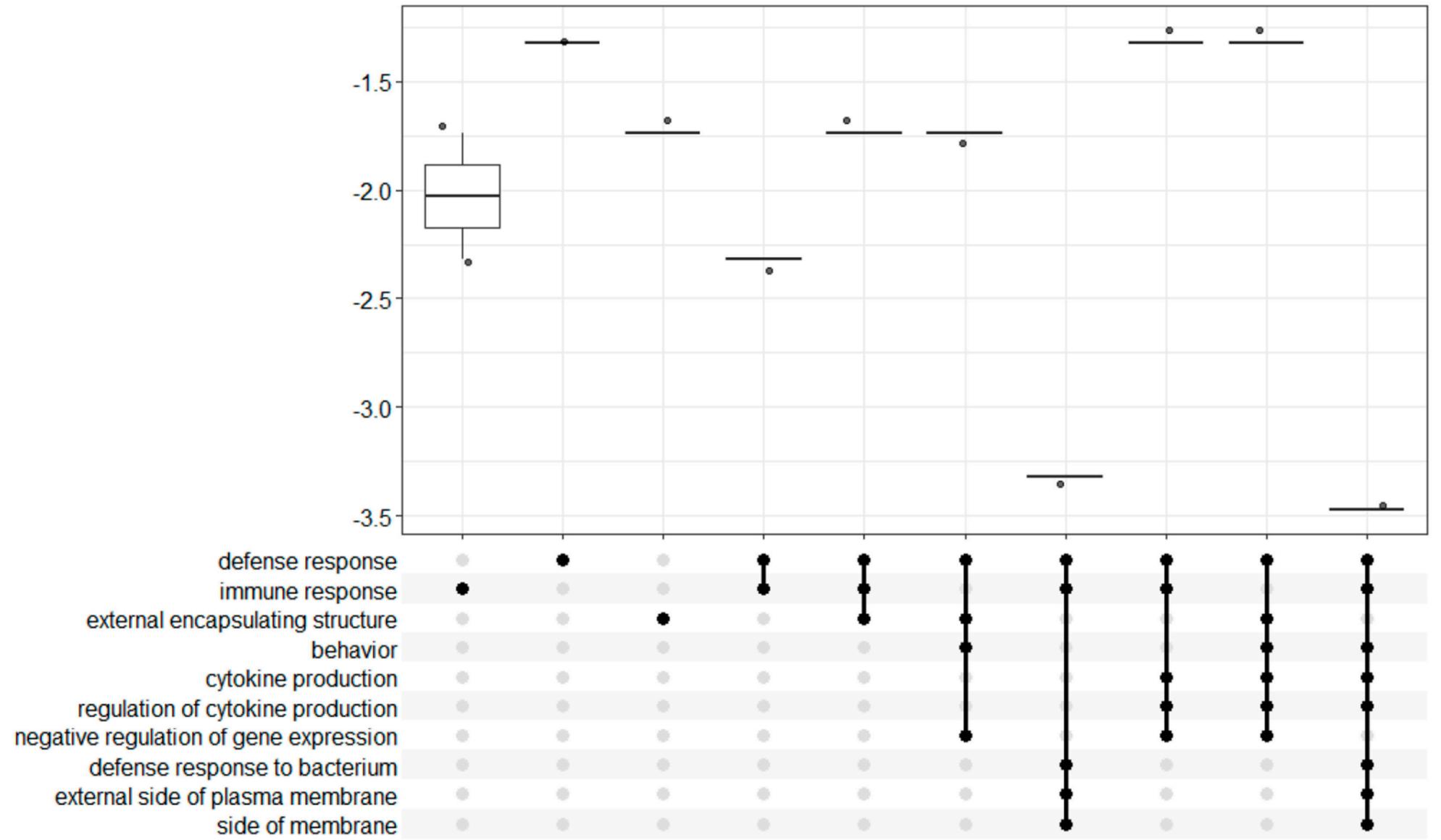
S2H.



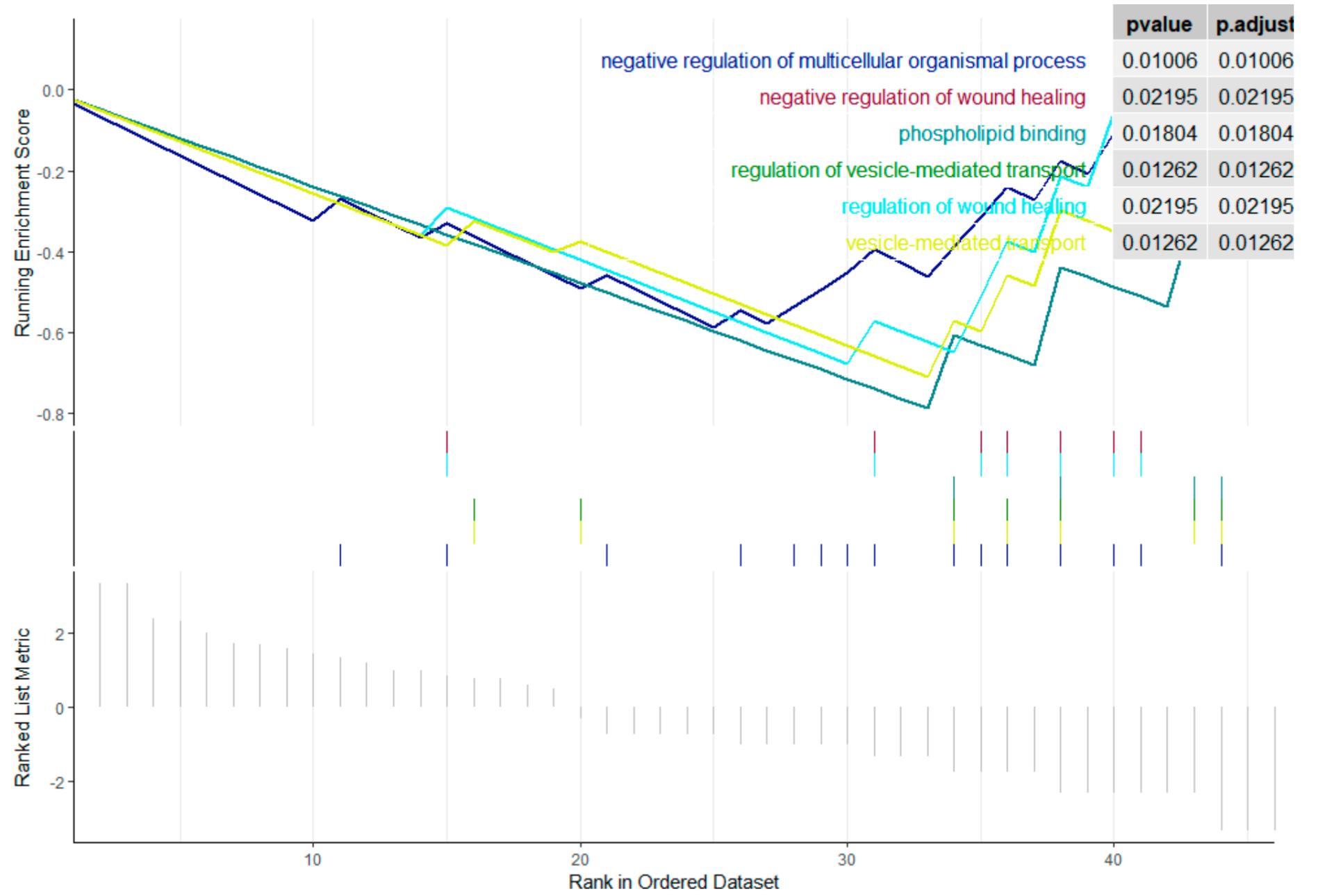
S3A.



S3B.



S4A.



S4B.

