

Table 1. Summary table of the top 10 gene families, in terms of total nodes. Of the 17761 total nodes over 53 families that were analyzed, only 25% were annotated. Abbreviations: CC (Cellular Component), MF (Molecular Function), BP (Biological Process).

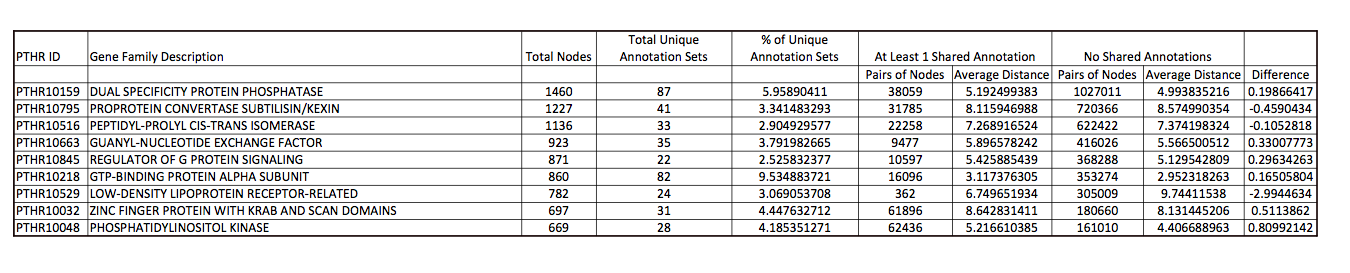


Table 2. Further analysis of the top 10 gene families, suggesting probable nonrandom sampling of nodes annotated. An average of 5% of the nodes have unique sets of annotations, and this can cause the phenomenon seen in the right hand columns – nodes with at least one shared annotation are further apart than those with none.

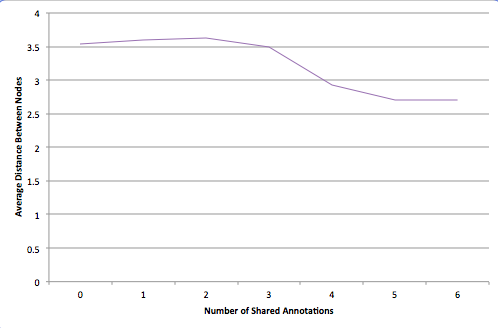


Figure 1. There is no negative correlation between average distance of nodes and the number of annotations they share, as would be expected from the current understanding of evolution.