Figure S1: Domain architectures of the RGS-family proteins. The domain architectures were identified using HMMER3 profile HMM search against the Pfam database (release 27). Proteins are listed for each of the 9 clusters identified by MOCASSIN-prot. The image for each protein was generated using the domain graphic generator on the Pfam website (http://pfam.xfam.org/generate graphic).
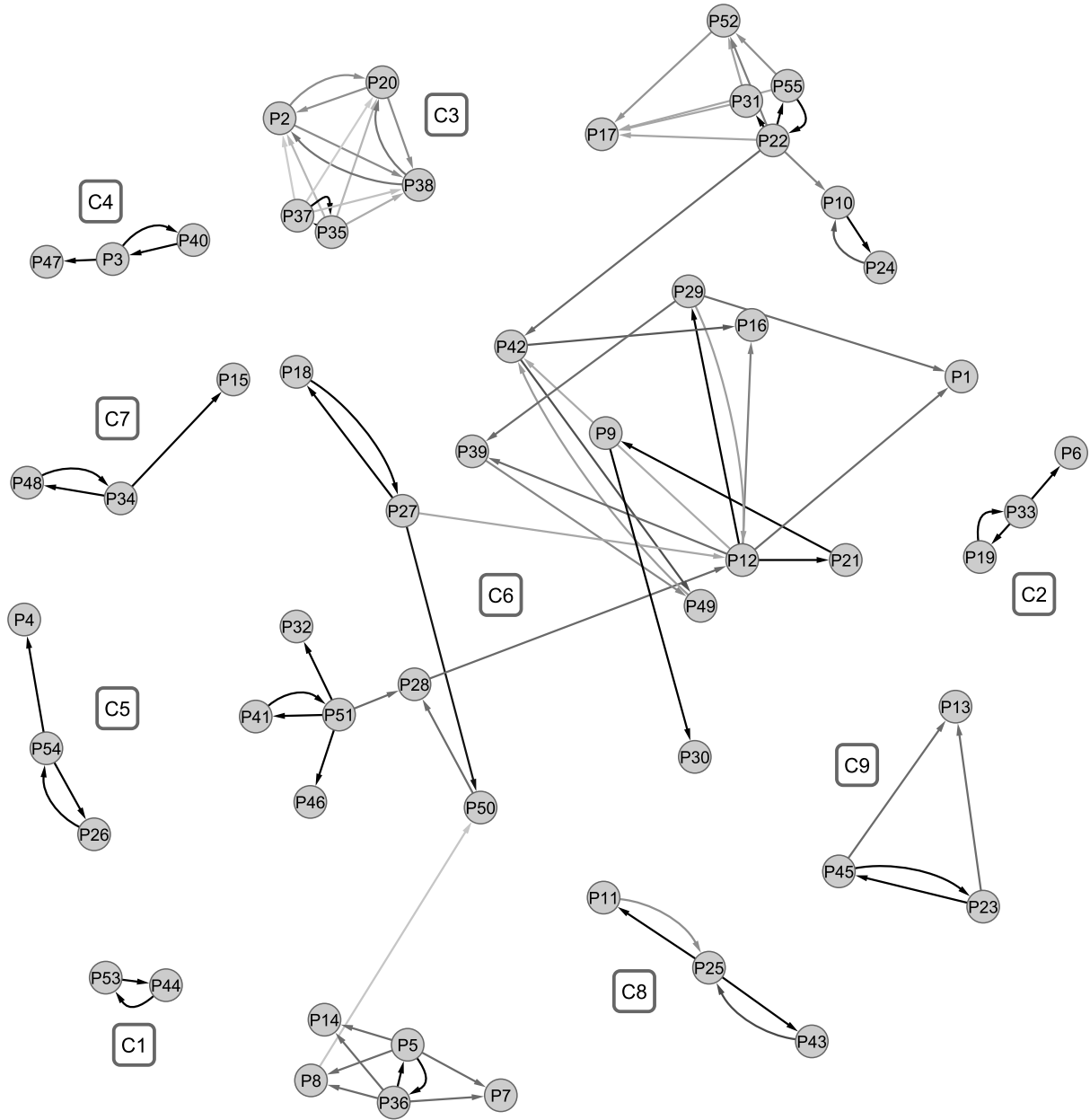
Figure S2. Primary MOCASSIN-prot network constructed from the 55 RGS-family proteins. Nine clusters were identified and labeled in descending order (from C1 to C9) according to their average optimal objective value. The nodes represent distinct proteins, and the edges are directed so that the incoming edge weights of each node sum to 1. The edge color indicates the edge weight, with darker (black) edges indicating high edge weights and lighter (gray) edges indicating low weights (note that edge weights are normalized to each reference protein). The optimal objective value for each protein is represented in the network by the length of its incoming edges, with longer edges corresponding to small objectives values. Since each protein has one optimal objective value, all incoming edges to a given node have the same edge length. The prefuse force-directed layout in Cytoscape was used to visualize the network using this edge length convention. Note that Cytoscape uses the Barnes-Hut approximation algorithm to produce the "best" node layout such that each edge has a specified length, and therefore in some cases the incoming edges are not exactly the same length.
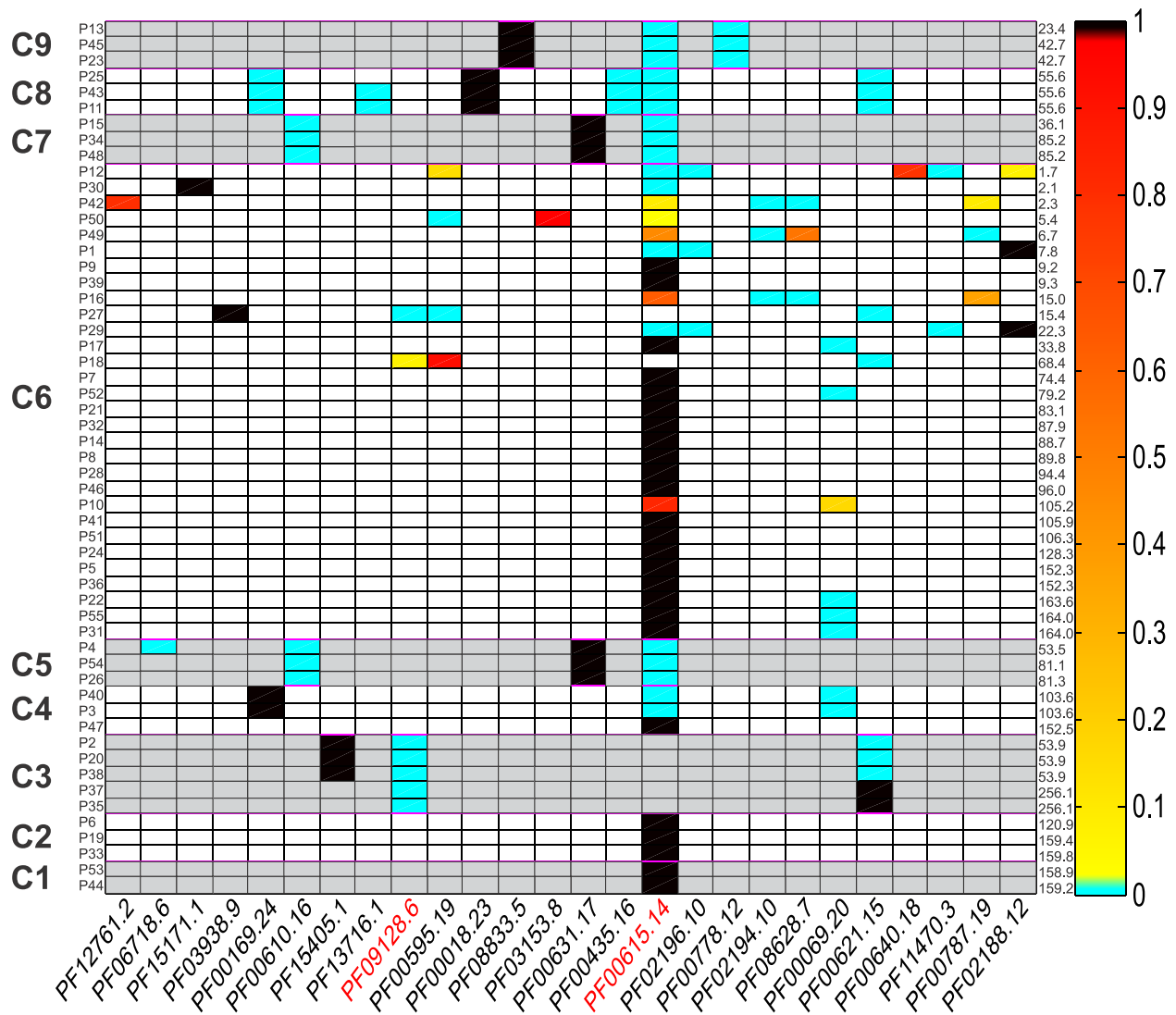
Figure S3. Diversity profile of the RGS-family proteins generated by MOCASSIN-prot. Each row represents one of the 55 RGS proteins, while each column represents one of the 26 Pfam domains (RGS and RGS-like domains are shown in red). Clusters (1-9) are separated by magenta lines and alternating white/gray background. Each cell is color-coded based on the diversity weight **x** from the LP problem (eq. 1): black for $x_i = 1$, cyan for $x_i = 0$ (but domain exists), and from light yellow to red for $0 < x_i < 1$. Blank cells indicate domain absence. Proteins in each cluster (P1-P55) are arranged according to their optimal objective value, with the largest appearing as the lowest row in the cluster. The objective value for each protein is shown on the right.
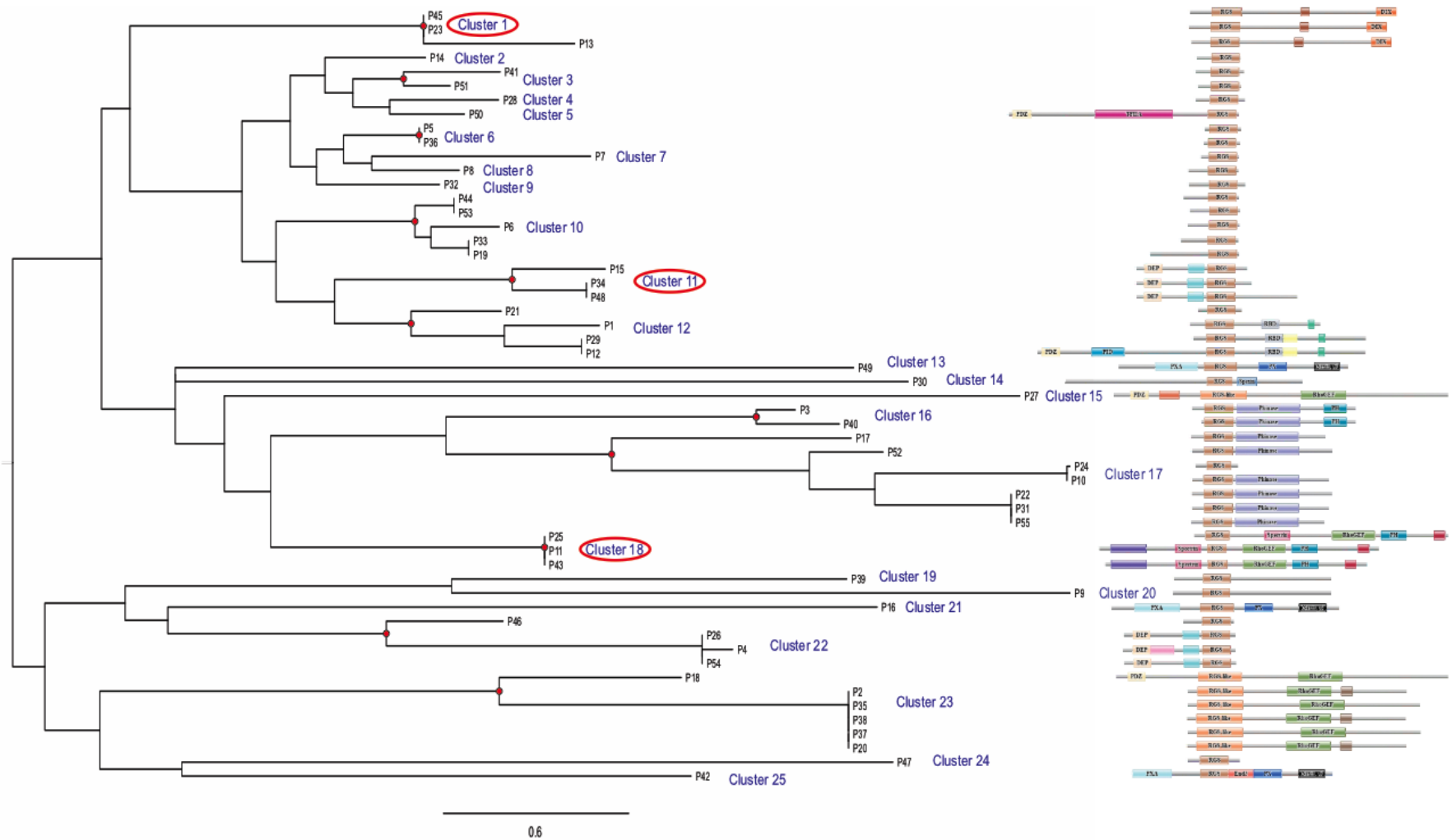
Figure S4: Maximum-likelihood phylogeny of RGS-family proteins. Nodes with at least 70% bootstrap support, shown by a red dot, define phylogenetic clusters. Clusters are labeled (Cluster 1-25). The domain architecture for each protein is shown on the right. Cluster labels with red circles indicate clusters that were identical to a cluster in the MOCASSIN-prot clustering in Figure S2.
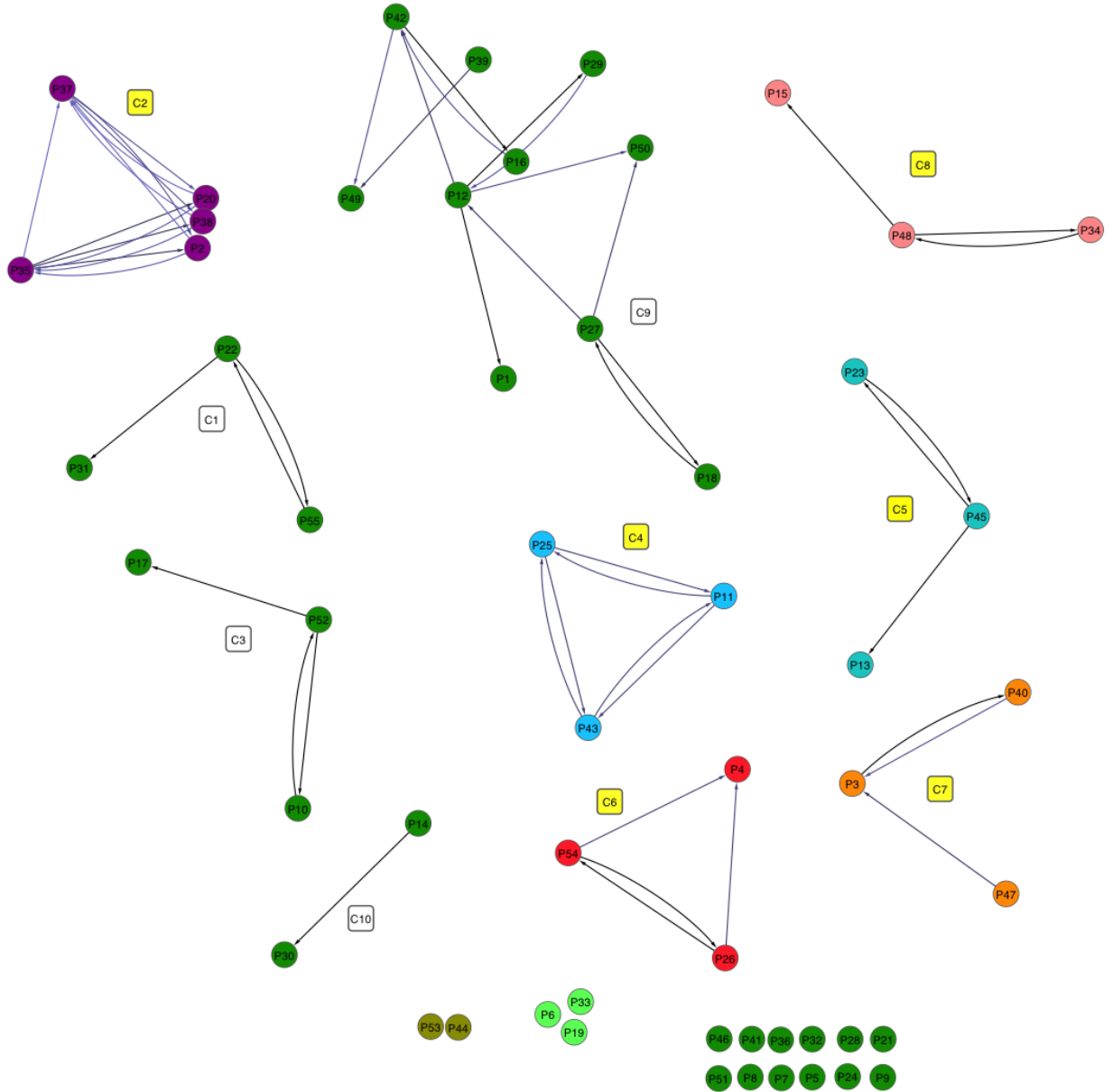
Figure S5: Secondary MOCASSIN-prot network constructed from the RGS-family proteins. From the 55 mouse RGS-family proteins, 10 clusters and 17 isolated proteins were identified using a second iteration of MOCASSIN-prot. Proteins are colored according to the cluster they belonged to in the primary network. Yellow cluster labels indicate clusters identical to those found in the primary network. Note that clusters in the network are labeled in descending order according to their average optimal objective value in this analysis (independent of the optimal objective values from the primary network).