

Additional STRING results for *BicBin*

Miranda van Uitert, Wouter Meuleman and Lodewyk Wessels*

We have validated most of the interesting biclusters that we found using the biclustering algorithm *BicBin*, with the STRING database. In Figures 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 we show the STRING output of the GO enriched and highly dissimilar biclusters among the top 20 biclusters obtained with $\alpha = 0.5$ and $\beta = 0.6$. A particular nice example is bicluster 20, see Figure 10, where all but two transcription factors are already connected. We hypothesize that these two transcription factors, Err1 and Stat6 could be involved in similar processes as the other transcription factors. This bicluster is enriched for several GO categories, such as "GO:0005509: calcium ion binding", "GO:0008270: zinc ion binding", and "GO:0030326: embryonic limb morphogenesis". We speculate that Err1 or Stat6 might be involved in regulating these processes. Finally, we show the first bicluster of the other 3 α - β combinations in Figures 11, 12 and 13.

*Email: l.wessels@nki.nl.

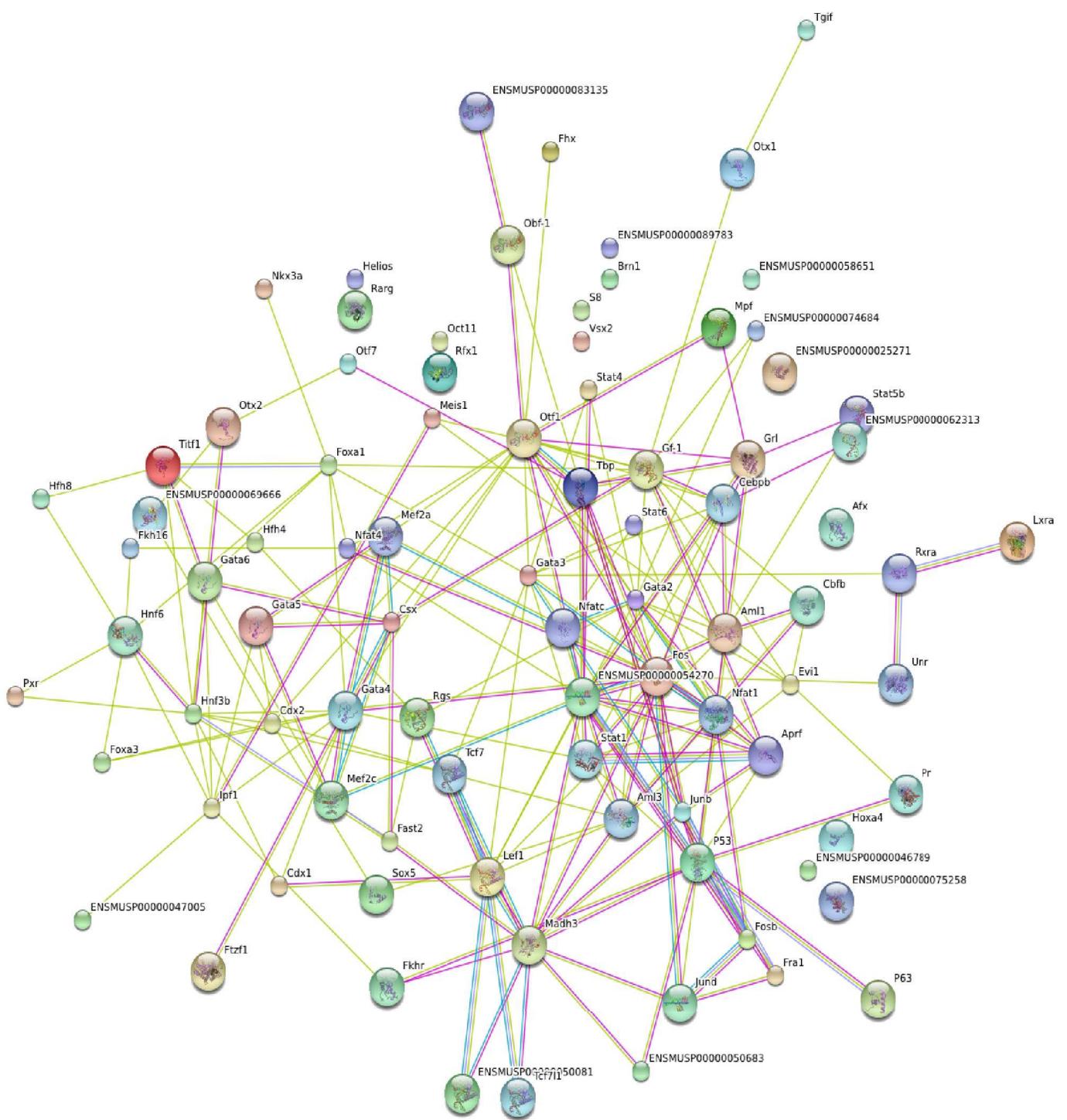


Figure 1: STRING output for the 2nd bicluster ($\alpha = 0.5$, $\beta = 0.6$). Transcription factors are represented by nodes, and a variety of associations between the nodes are represented by the colored edges. The following associations between nodes are shown in the graph: 1) 'neighborhood' indicating frequent occurrence of genes in the same genomic neighborhood (dark green); 2) 'co-occurrence' of linked orthologous groups across species (dark blue); 3) 'gene fusion' indicating whether fusion occurred (red); 4) 'co-expression' representing evidence of mRNA co-expression of the associated genes (black); 5) 'experiments' representing experimental evidence for an interaction (purple); 6) 'databases' representing evidence from pathway databases for association (turquoise) and 7) 'text mining' representing co-publication information (light green).

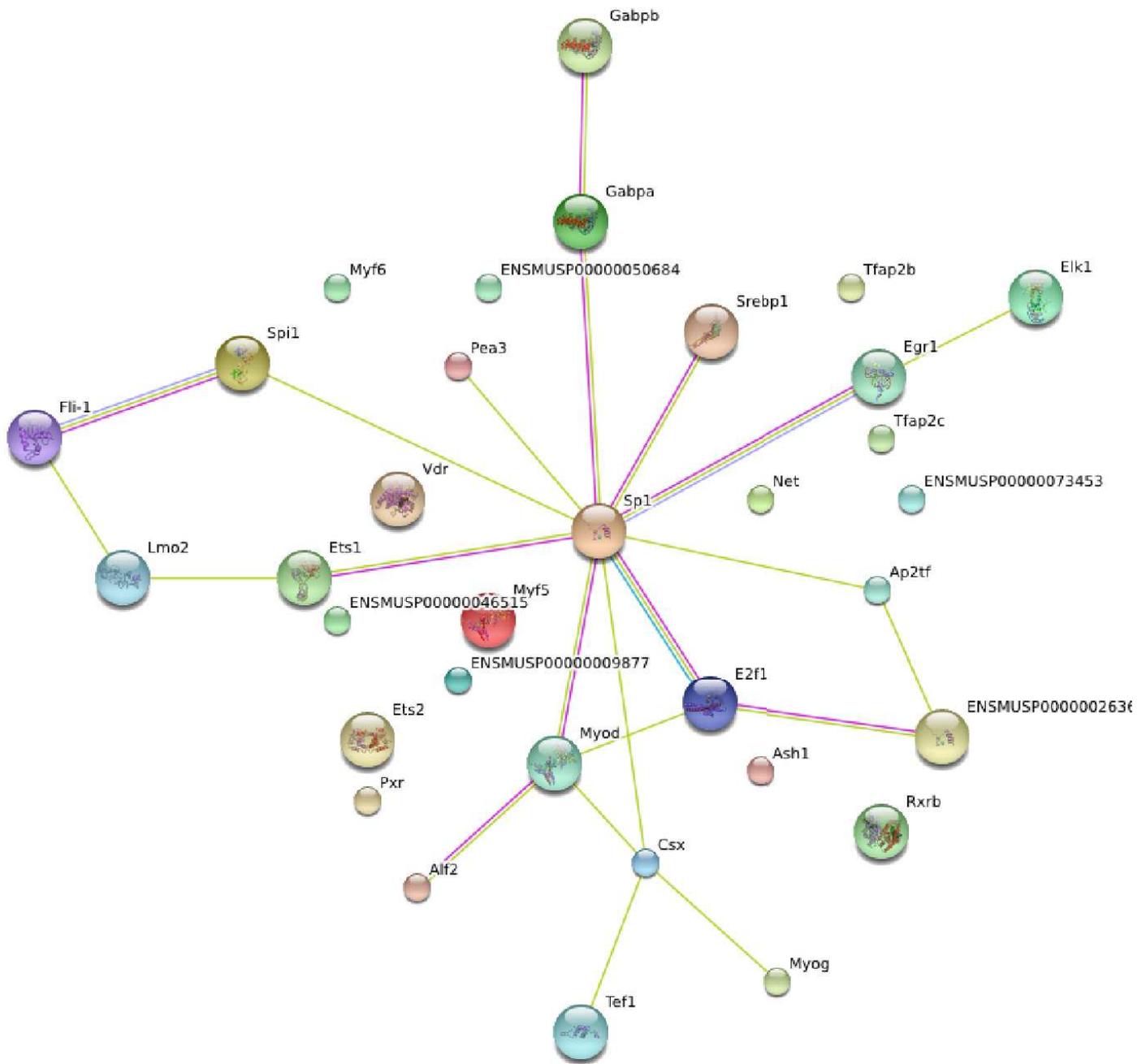


Figure 2: STRING output for the 3rd bicluster ($\alpha = 0.5, \beta = 0.6$). Transcription factors are represented by nodes, and a variety of associations between the nodes are represented by the colored edges. The following associations between nodes are shown in the graph: 1) 'neighborhood' indicating frequent occurrence of genes in the same genomic neighborhood (dark green); 2) 'co-occurrence' of linked orthologous groups across species (dark blue); 3) 'gene fusion' indicating whether fusion occurred (red); 4) 'co-expression' representing evidence of mRNA co-expression of the associated genes (black); 5) 'experiments' representing experimental evidence for an interaction (purple); 6) 'databases' representing evidence from pathway databases for association (turquoise) and 7) 'text mining' representing co-publication information (light green).

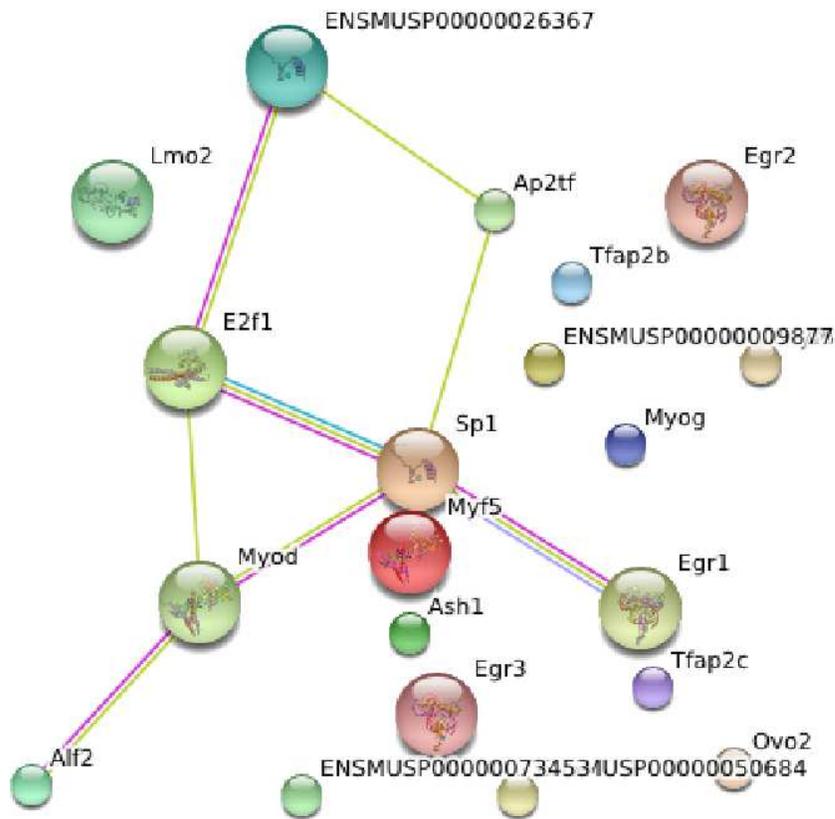


Figure 3: STRING output for the 5th bicluster ($\alpha = 0.5$, $\beta = 0.6$). Transcription factors are represented by nodes, and a variety of associations between the nodes are represented by the colored edges. The following associations between nodes are shown in the graph: 1) 'neighborhood' indicating frequent occurrence of genes in the same genomic neighborhood (dark green); 2) 'co-occurrence' of linked orthologous groups across species (dark blue); 3) 'gene fusion' indicating whether fusion occurred (red); 4) 'co-expression' representing evidence of mRNA co-expression of the associated genes (black); 5) 'experiments' representing experimental evidence for an interaction (purple); 6) 'databases' representing evidence from pathway databases for association (turquoise) and 7) 'text mining' representing co-publication information (light green).

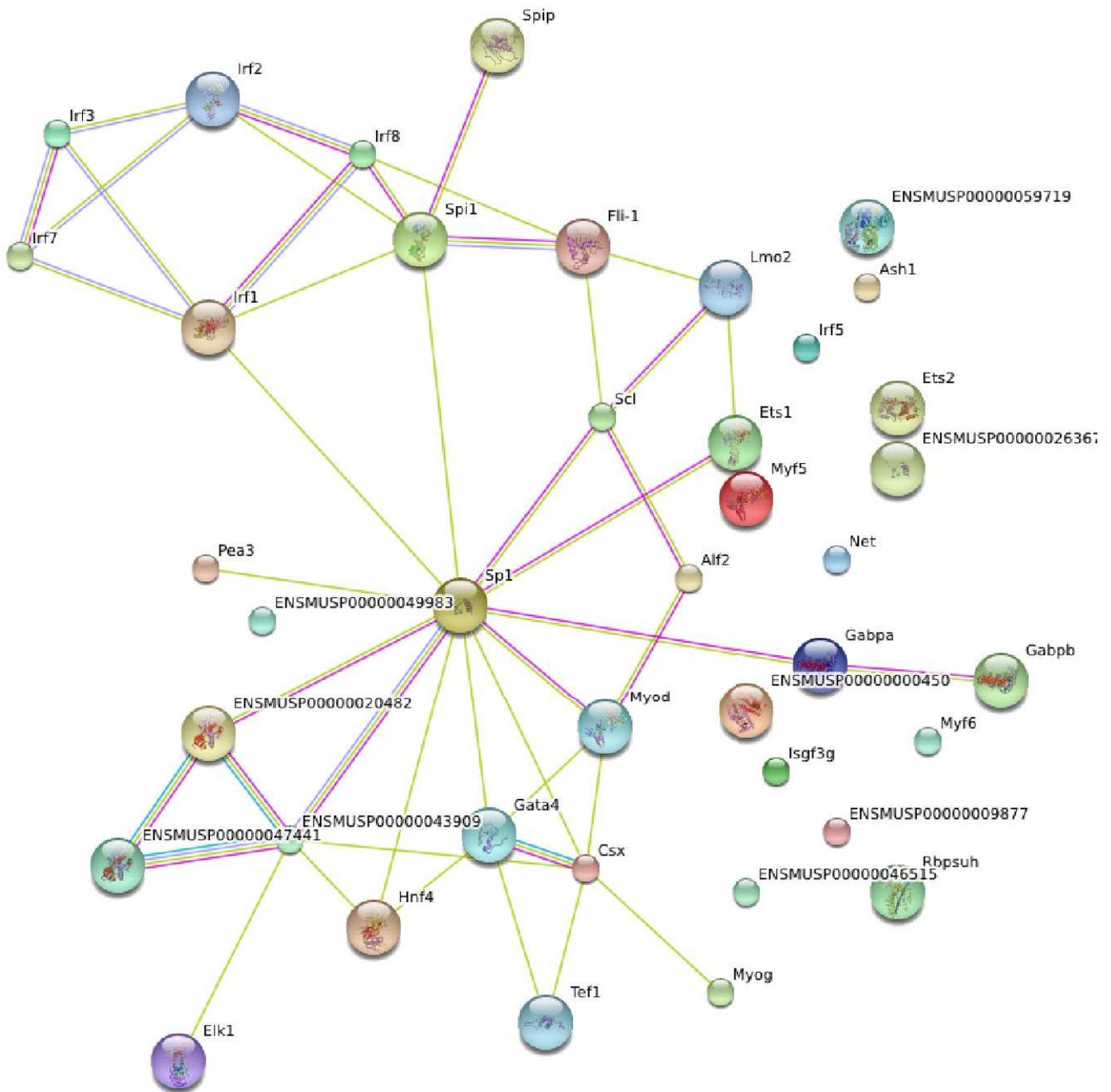


Figure 4: STRING output for the 6th bicluster ($\alpha = 0.5, \beta = 0.6$). Transcription factors are represented by nodes, and a variety of associations between the nodes are represented by the colored edges. The following associations between nodes are shown in the graph: 1) 'neighborhood' indicating frequent occurrence of genes in the same genomic neighborhood (dark green); 2) 'co-occurrence' of linked orthologous groups across species (dark blue); 3) 'gene fusion' indicating whether fusion occurred (red); 4) 'co-expression' representing evidence of mRNA co-expression of the associated genes (black); 5) 'experiments' representing experimental evidence for an interaction (purple); 6) 'databases' representing evidence from pathway databases for association (turquoise) and 7) 'text mining' representing co-publication information (light green).



Figure 5: STRING output for the 7th bicluster ($\alpha = 0.5$, $\beta = 0.6$). Transcription factors are represented by nodes, and a variety of associations between the nodes are represented by the colored edges. The following associations between nodes are shown in the graph: 1) 'neighborhood' indicating frequent occurrence of genes in the same genomic neighborhood (dark green); 2) 'co-occurrence' of linked orthologous groups across species (dark blue); 3) 'gene fusion' indicating whether fusion occurred (red); 4) 'co-expression' representing evidence of mRNA co-expression of the associated genes (black); 5) 'experiments' representing experimental evidence for an interaction (purple); 6) 'databases' representing evidence from pathway databases for association (turquoise) and 7) 'text mining' representing co-publication information (light green).

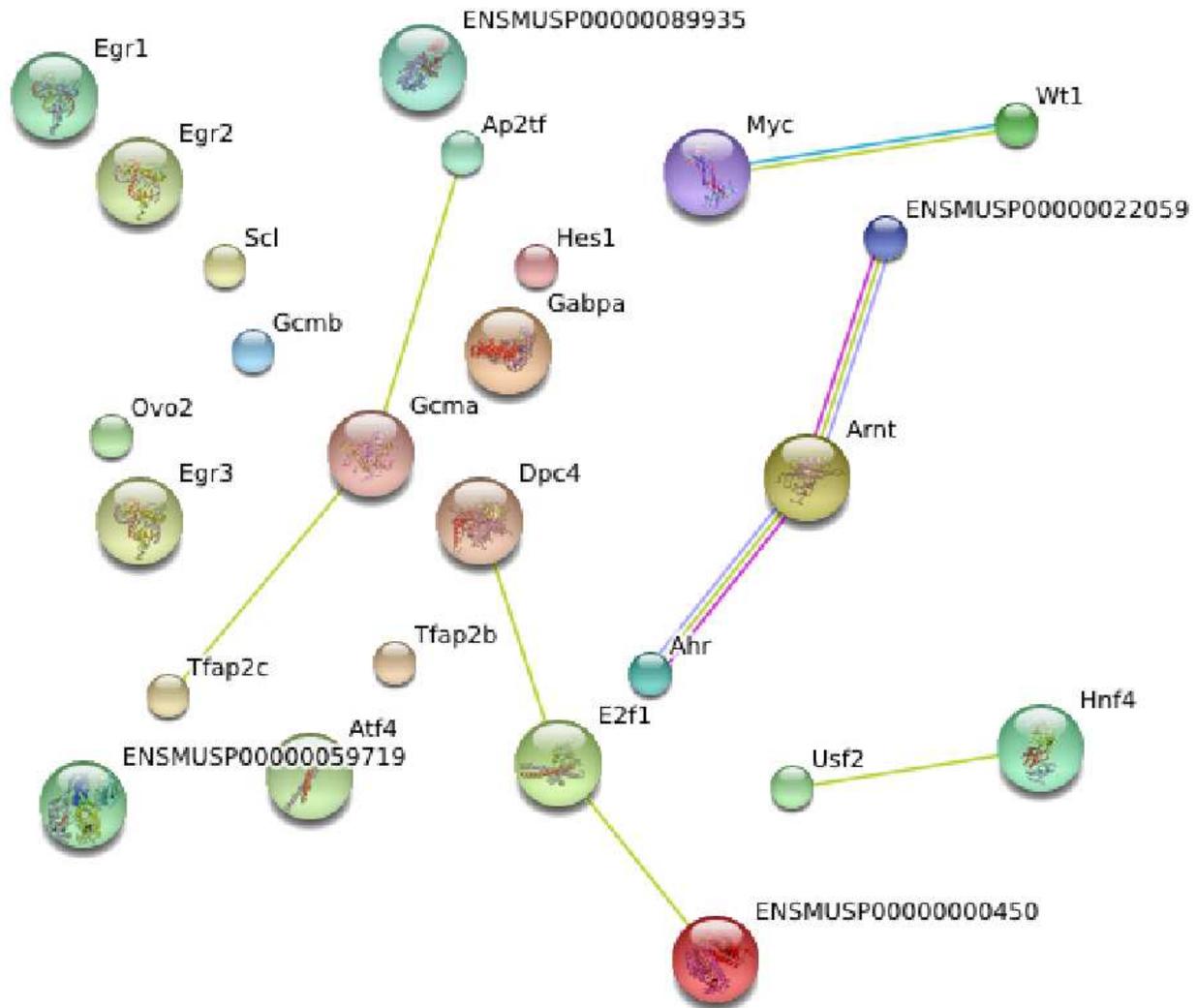


Figure 6: STRING output for the 8th bicluster ($\alpha = 0.5, \beta = 0.6$). Transcription factors are represented by nodes, and a variety of associations between the nodes are represented by the colored edges. The following associations between nodes are shown in the graph: 1) 'neighborhood' indicating frequent occurrence of genes in the same genomic neighborhood (dark green); 2) 'co-occurrence' of linked orthologous groups across species (dark blue); 3) 'gene fusion' indicating whether fusion occurred (red); 4) 'co-expression' representing evidence of mRNA co-expression of the associated genes (black); 5) 'experiments' representing experimental evidence for an interaction (purple); 6) 'databases' representing evidence from pathway databases for association (turquoise) and 7) 'text mining' representing co-publication information (light green).

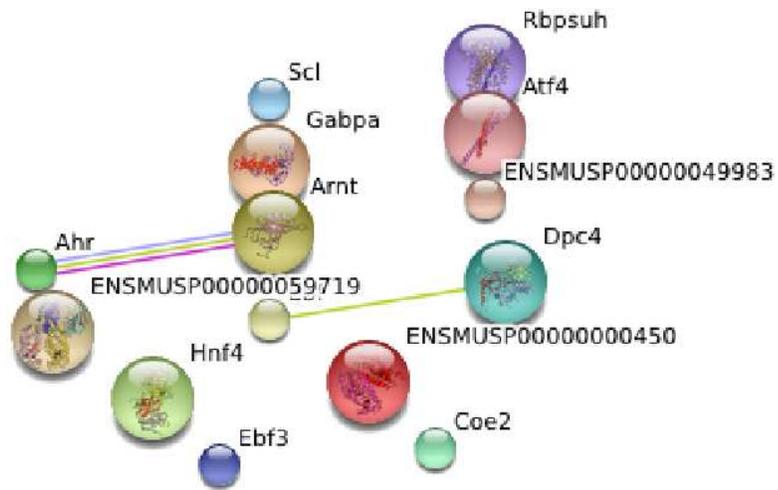


Figure 8: STRING output for the 10th bicluster ($\alpha = 0.5$, $\beta = 0.6$). Transcription factors are represented by nodes, and a variety of associations between the nodes are represented by the colored edges. The following associations between nodes are shown in the graph: 1) 'neighborhood' indicating frequent occurrence of genes in the same genomic neighborhood (dark green); 2) 'co-occurrence' of linked orthologous groups across species (dark blue); 3) 'gene fusion' indicating whether fusion occurred (red); 4) 'co-expression' representing evidence of mRNA co-expression of the associated genes (black); 5) 'experiments' representing experimental evidence for an interaction (purple); 6) 'databases' representing evidence from pathway databases for association (turquoise) and 7) 'text mining' representing co-publication information (light green).

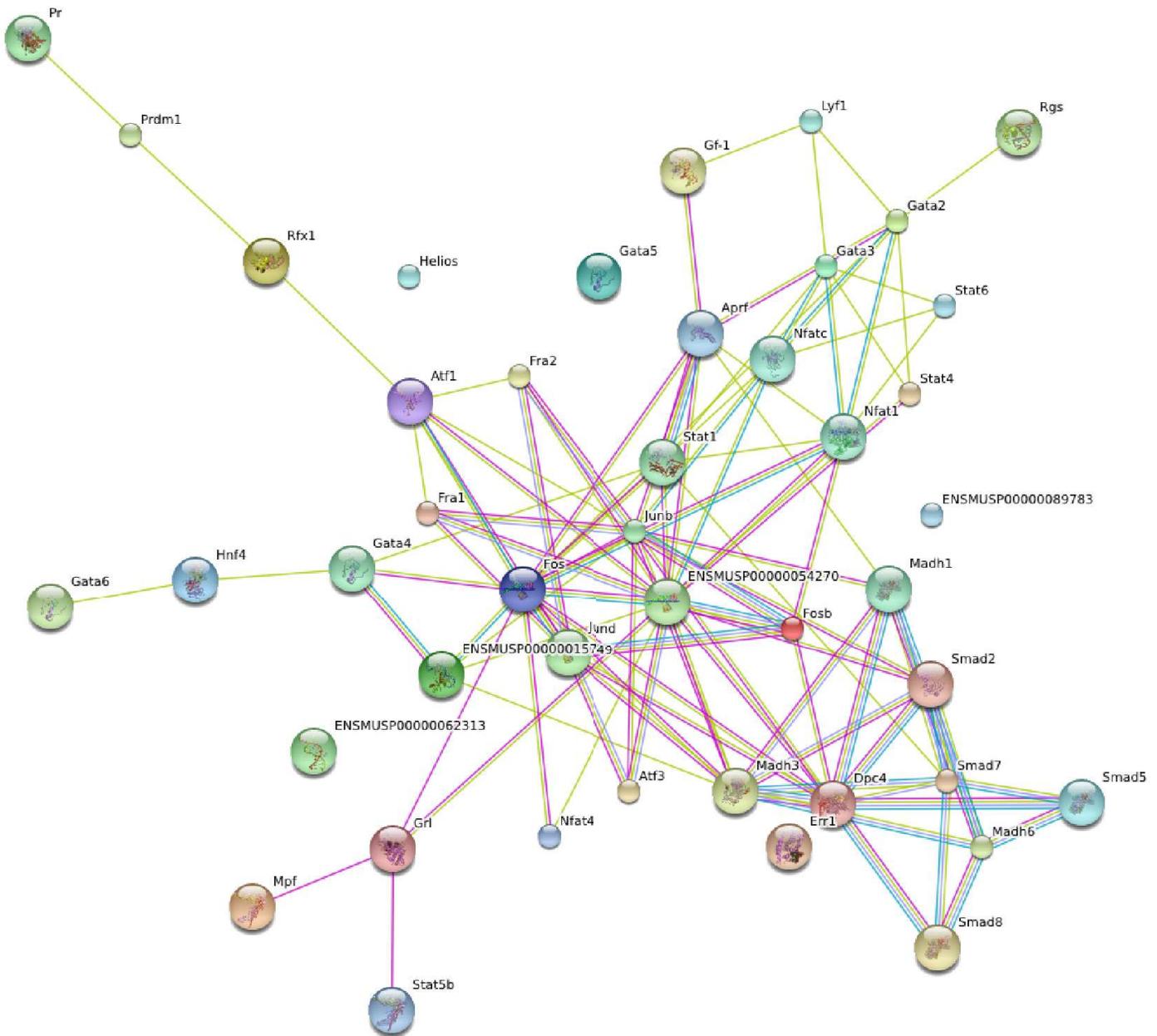


Figure 9: STRING output for the 13th bicluster ($\alpha = 0.5$, $\beta = 0.6$). Transcription factors are represented by nodes, and a variety of associations between the nodes are represented by the colored edges. The following associations between nodes are shown in the graph: 1) 'neighborhood' indicating frequent occurrence of genes in the same genomic neighborhood (dark green); 2) 'co-occurrence' of linked orthologous groups across species (dark blue); 3) 'gene fusion' indicating whether fusion occurred (red); 4) 'co-expression' representing evidence of mRNA co-expression of the associated genes (black); 5) 'experiments' representing experimental evidence for an interaction (purple); 6) 'databases' representing evidence from pathway databases for association (turquoise) and 7) 'text mining' representing co-publication information (light green).

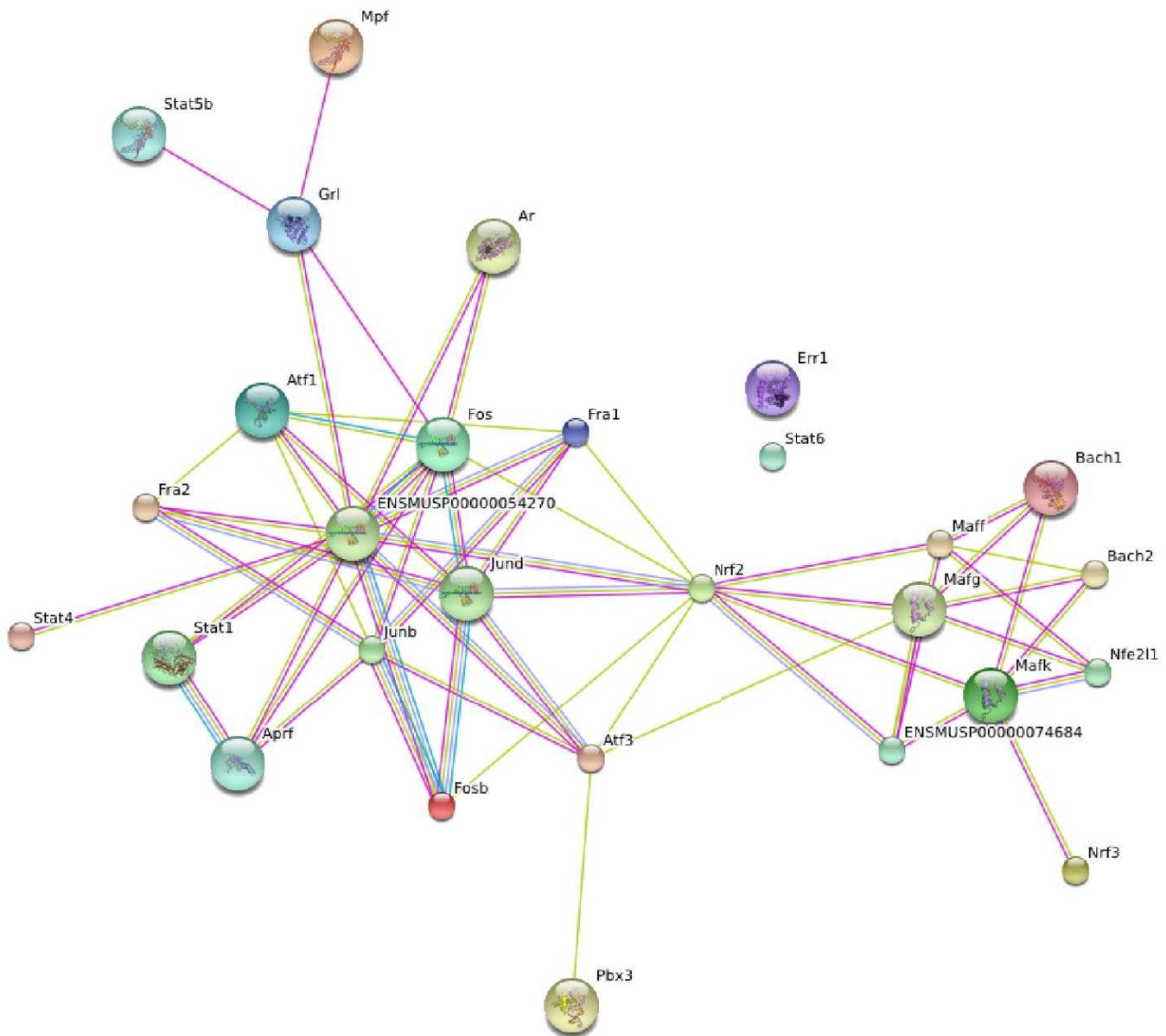


Figure 10: STRING output for the 20th bicluster ($\alpha = 0.5$, $\beta = 0.6$). Transcription factors are represented by nodes, and a variety of associations between the nodes are represented by the colored edges. The following associations between nodes are shown in the graph: 1) 'neighborhood' indicating frequent occurrence of genes in the same genomic neighborhood (dark green); 2) 'co-occurrence' of linked orthologous groups across species (dark blue); 3) 'gene fusion' indicating whether fusion occurred (red); 4) 'co-expression' representing evidence of mRNA co-expression of the associated genes (black); 5) 'experiments' representing experimental evidence for an interaction (purple); 6) 'databases' representing evidence from pathway databases for association (turquoise) and 7) 'text mining' representing co-publication information (light green).

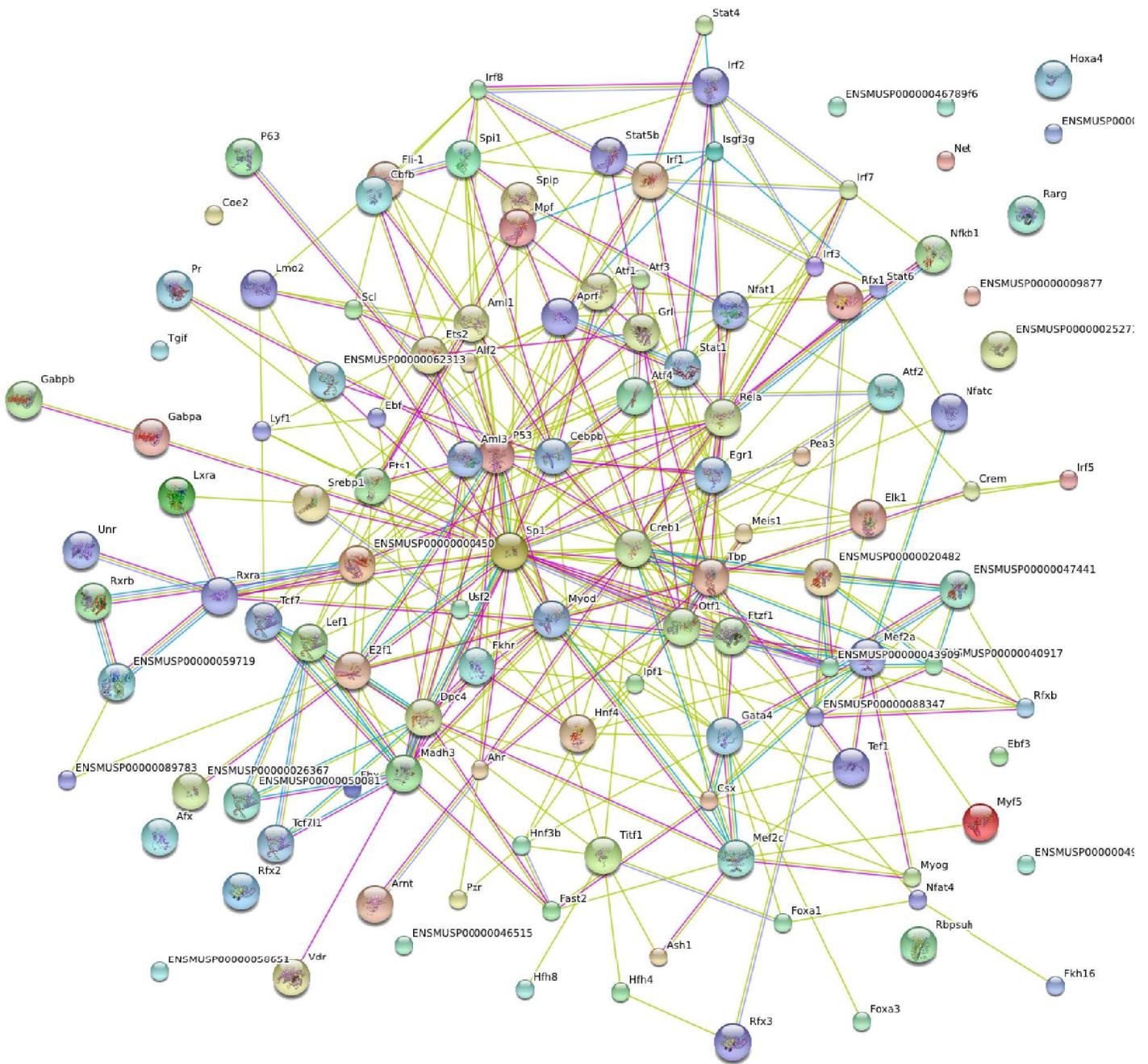


Figure 11: STRING output for the first bicluster ($\alpha = 0.6, \beta = 0.7$). Transcription factors are represented by nodes, and a variety of associations between the nodes are represented by the colored edges. The following associations between nodes are shown in the graph: 1) 'neighborhood' indicating frequent occurrence of genes in the same genomic neighborhood (dark green); 2) 'co-occurrence' of linked orthologous groups across species (dark blue); 3) 'gene fusion' indicating whether fusion occurred (red); 4) 'co-expression' representing evidence of mRNA co-expression of the associated genes (black); 5) 'experiments' representing experimental evidence for an interaction (purple); 6) 'databases' representing evidence from pathway databases for association (turquoise) and 7) 'text mining' representing co-publication information (light green).

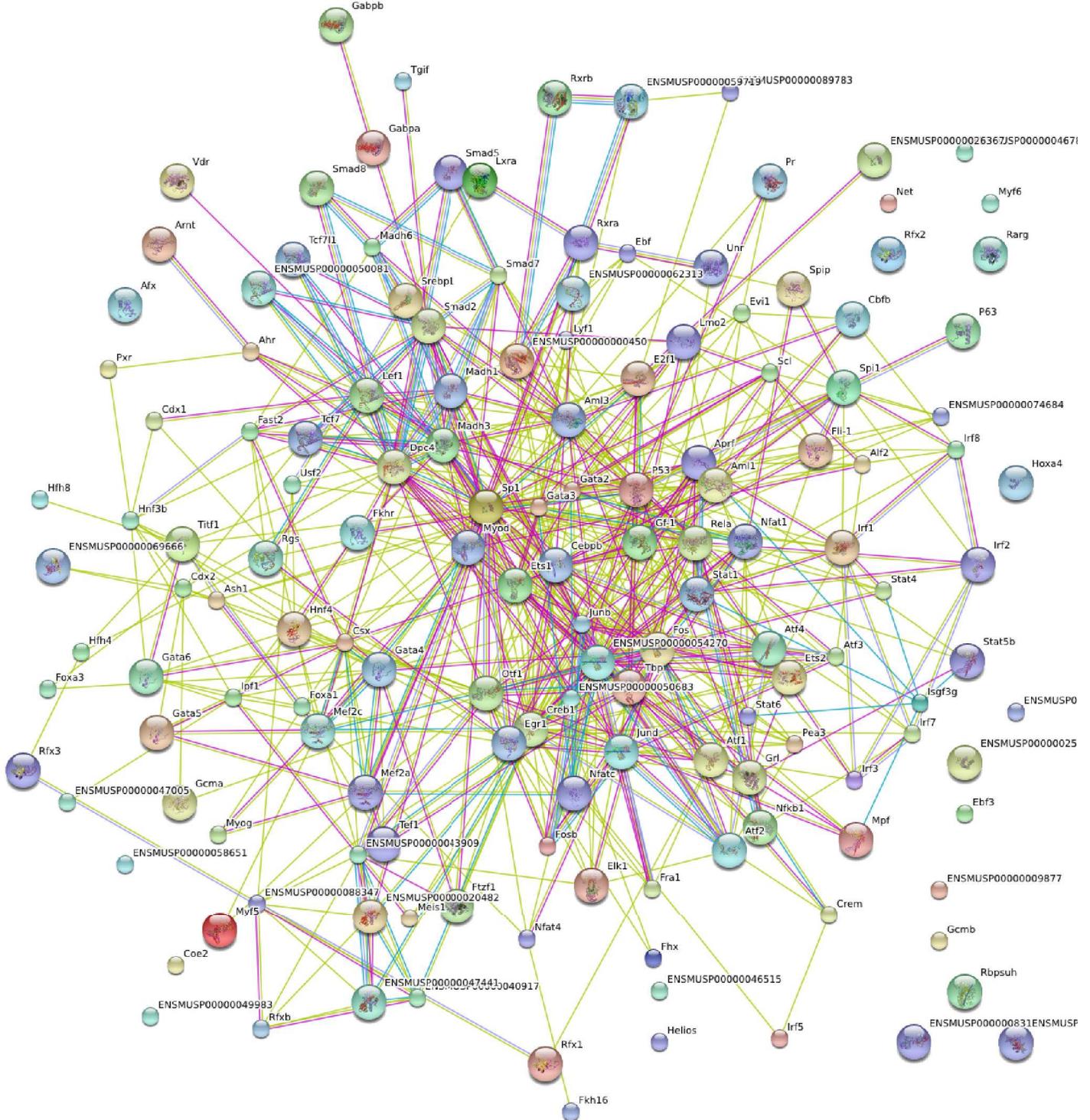


Figure 12: STRING output for the first bicluster ($\alpha = 0.5$, $\beta = 0.7$). Transcription factors are represented by nodes, and a variety of associations between the nodes are represented by the colored edges. The following associations between nodes are shown in the graph: 1) 'neighborhood' indicating frequent occurrence of genes in the same genomic neighborhood (dark green); 2) 'co-occurrence' of linked orthologous groups across species (dark blue); 3) 'gene fusion' indicating whether fusion occurred (red); 4) 'co-expression' representing evidence of mRNA co-expression of the associated genes (black); 5) 'experiments' representing experimental evidence for an interaction (purple); 6) 'databases' representing evidence from pathway databases for association (turquoise) and 7) 'text mining' representing co-publication information (light green).

