"A network of protein domains that are present in chimeric tyrosine-kinase proteins in cancer"

Introduction

Chromosomal translocations resulting in fusion genes that code for fusion chimeric proteins is a very common feature of cancer. The presence of fusion proteins with tyrosine-kinase (TK) activity, in particular, has been implicated in many different types of tumors. These fusion TKs are generally composed by an oligomerization domain together with a TK domain from a different protein, resulting in deregulated and constitutive TK activity in specific cell types.

We have recently created TICdb, a collection of gene-mapped translocation breakpoints in human tumors. Using the information contained in that database, a network of genes translocated in cancer was also created. In the present work, we have focused on the part of this network that includes genes coding for TKs (what we call the TK subnetwork). For all gene fusions in this subnetwork, we obtained the PFAM domains present in chimeric TK proteins, with a view to create a network of protein domains that are present in the same fusion protein.

Methods

For each chromosomal translocation involving a gene in the TK subnet we obtained from Ensembl the accession number of PFAM motifs coded by the exons included in the fusion gene. This enabled us to see which PFAM motifs are present in the same fusion protein.
We also recorded the reading frame for each exon, in order to check that the fusion gene keeps an intact reading frame.
This information was then processed using Cytoscape 2.4.0 in order to construct a network of interacting PFAM domains.

Results and Discussion

The TK subnetwork of genes fused in cancer (Figure 1) is comprised by 58 genes (58 nodes and 59 edges). There are 5 hubs (5 or more edges per node), four of which correspond to known TK genes and the other to ETV6, a gene frequently rearranged in cancer. Only 9 fusions (6.5%) do not seem to keep an intact reading frame, supporting the hypothesis that a complete fusion protein is necessary for cancer development.
PFAM motifs were obtained for all these gene fusions and represented as a network of protein domains that are present in fusion TK proteins in cancer. In most cases the protein motif is completely preserved in the fusion protein, with a few exceptions in which the breakpoint fell within the motif.
The network of PFAM domains is represented in Figure 2 and contains 43 nodes and 57 edges. The topology is clearly different to the corresponding network of fused genes shown in Figure 1. In the protein domain network we see two clear hubs with more than 10 edges, which correspond to the PNT domain of the ETV6 gene and to the TK domain present in all the TK genes of the network. Thus, transforming the gene fusion network into the protein domain network has not only reduced the complexity, but also identified potential mechanisms by which these translocations drive the oncogenic process.

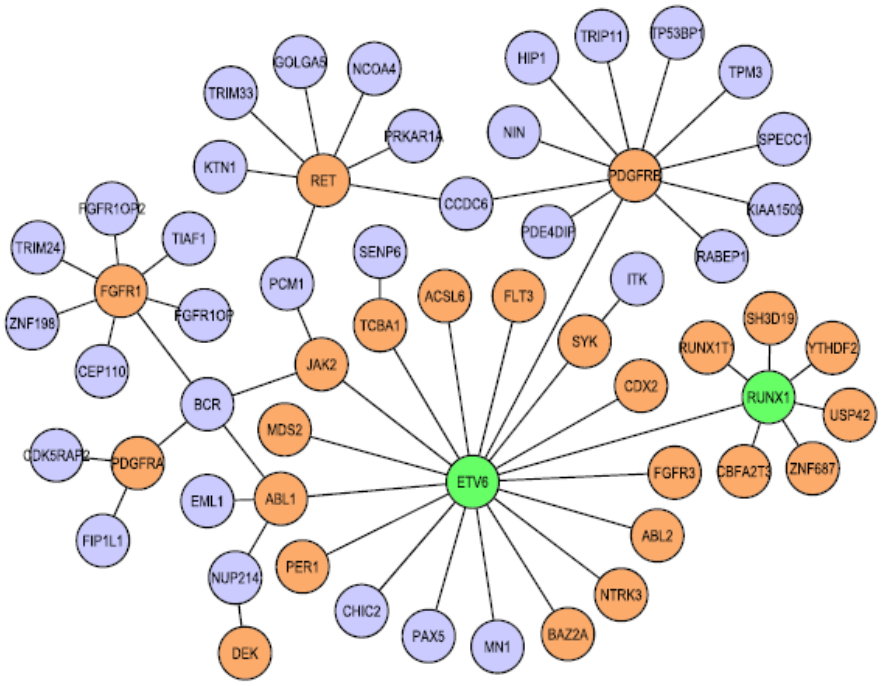

Figure 1: A network of TK genes translocated in cancer

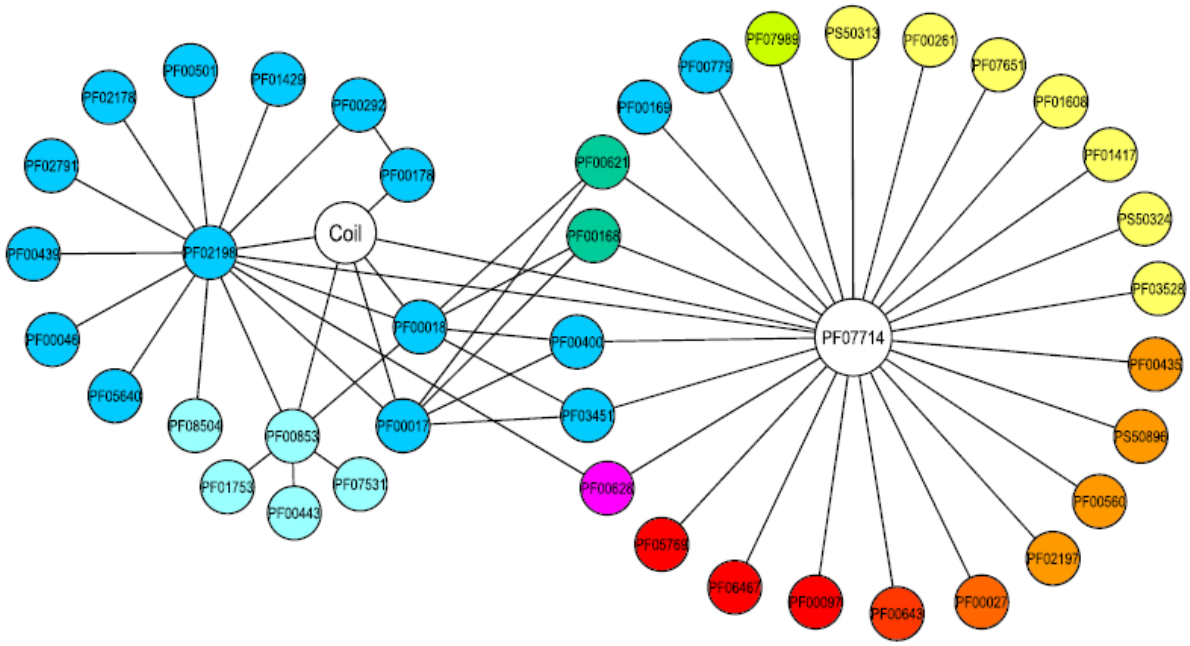

Figure 2: Network of PFAM domains present in the genes shown in Figure 1.