Kinase interaction network

Introduction

Protein kinases are attractive objects of study in medicine, as understanding their interactions plays an important role in the fight against diseases such as cancer (Kanev et al. 2019, Kim et al. 2017). These Enzymes are most likely known to maintain post-translational control mechanisms in signal transduction. Furthermore, they play a major role in cellular energy balance (Carling, 2017). Primarily, these enzymes covalently add phosphoryl groups to other proteins via phosphorylation among different phosphorylation sites (Kim et al. 2017).

However, creating a network of these kind of molecules is specifically crucial for further investigations in medicine as well as any other natural sciences. The aim of this work therefore is to build such Kinase signaling a network using prior knowledge from the database SIGNOR

Methods

First all human phosphorylation data Kinase signaling data was downloaded from the SIGNOR database. The data has been processed with *pandas* library in Python and a new dataframe object has been created with four features (Target, Position, Residue, Kinase). To prevent the network from being to cluttered all proteins that are not protein kinases have not been included and only important nodes and edges have been drawn. Subsequently a directed graph with the *networkx* library in Python was computed. Furthermore, a degree distribution has been plotted and the Kinases have been ranked by degree from highest to lowest. Additionally, all self-phosphorylation loops and duplicate loops have been deleted to give more transparency.

Furthermore, the importance of nodes was calculated by using the eigenvector centrality and betweenness centrality values of each node. First, represents the influence of a node in a network by considering if connections to higher or lower scoring nodes were made and referring to how quick a node can communicate with others. Second is explained by how much effectively a node can communicate with other nodes and considering the number of

shortest paths between this and other nodes. Large nodes are more important, small nodes are less.

Moreover, an investigation in inhibition of the MAPK1_HUMAN enzyme and the resulting changings on the network was carried out. Additionally, a mutation in the 642nd residue of the WEE1_HUMAN protein from serine to glutamine was performed and analysed. In the end also an investigation in change of 474th residue of AKT1_HUMAN from tyrosine to phenylalanine has been conducted.

Results

The resulting kinase signaling network (Fig. 1) shows a directed graph of all the most important protein kinases circle-sized by importance. The amino acids Serine (blue), Threonine (green) and Tyrosine (orange) are the most likely to see residues on kinases visible on their over-representation in the network. The network consisted of 315 nodes and 814 edges after deleting self-loops only 642 edges were left. The importance of the kinases is characterised by the size of their nodes and identified on how quickly and effectively these nodes communicate inside the network. Non-circled Kinases have a lower importance on the network subsequently have not been colored. The five top ranked important kinases in this network (Fig. 1) are GSK3B, MAP2K1, RAF1, EEF2K and EGFR. Looking at the top ranked kinases in terms of degree after self-loop edges have been deleted AKT1 (37), SRC (36), AKT (35), PRKACA (25) and MAPK1 (25) were found in this order (Table 1). The general degree distribution shows that the majority of kinases are sparsely connected and only a few of them are covering large values of degrees. Summarizing a decreasing dependency on node quantity to degree quantity is visible (Fig. 2).

Furthermore, an investigation in inhibition of the MAPK1_HUMAN enzyme resulted in a decreasing number of total edges to 583. Moreover, a mutation in the 642nd residue of the WEE1_HUMAN protein from serine to glutamine changed the network by nothing considering it to be a phosphomimetic mutation. However, when investigating a change of 474th residue of AKT1_HUMAN from tyrosine to phenylalanine this residue has not been found in the network and no changes could have been explained.

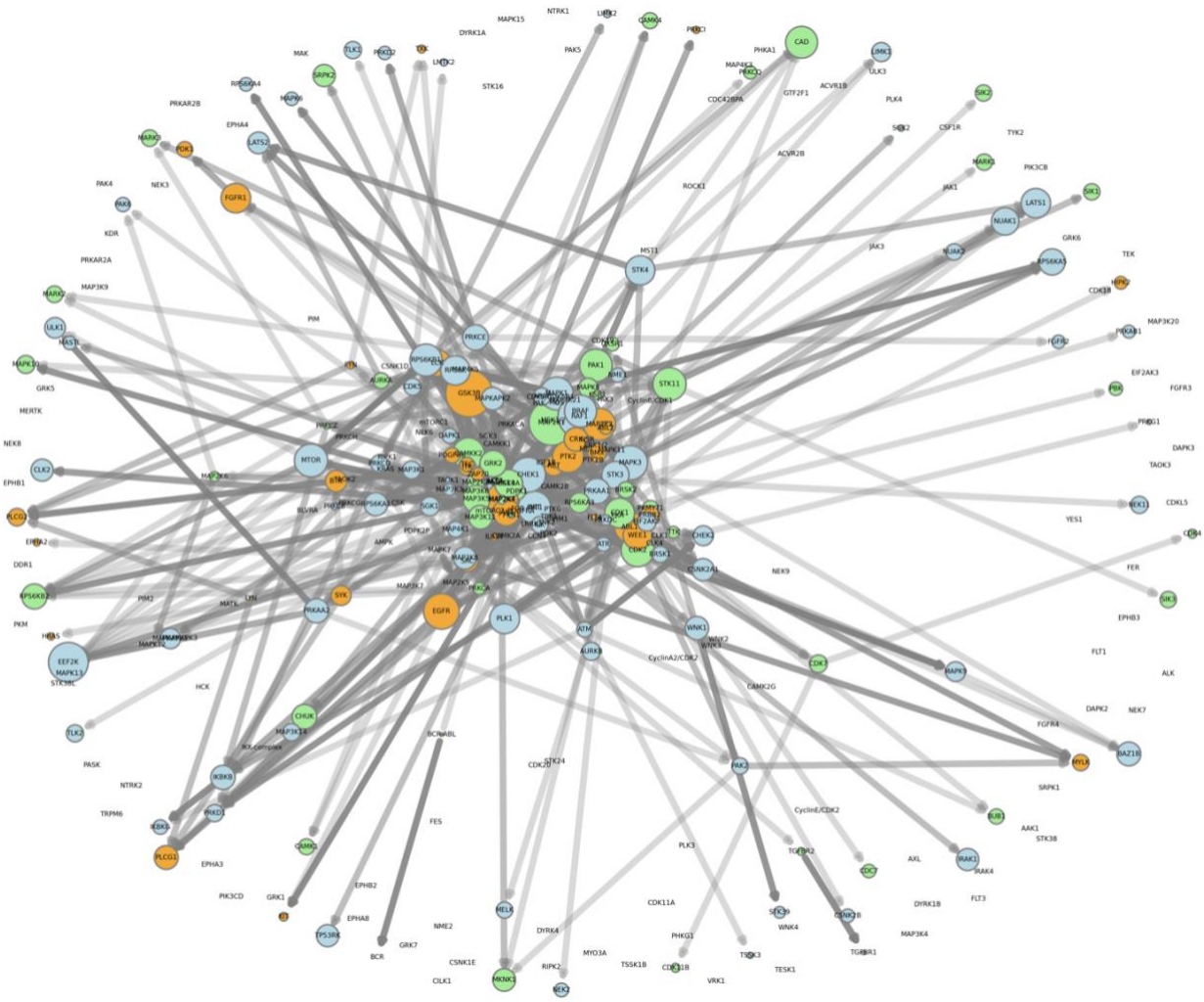

Figure 1 Kinase signaling network with no self-loops and only all important nodes and edges included. Colours of the nodes represent the specific aminoacid residue. H = red, Y = orange, T = green, S = blue. The arrows on the end of the edges describe the kinase and target relation. Non-circled Kinases have a lower importance on the network subsequently decided to be not coloured.

Table of top ranked protein degrees

Kinase	Degree
AKT1	37
AKT	36
SRC	35
PRKACA	25
MAPK1	25

Table 1 Top ranked protein degrees in kinase network

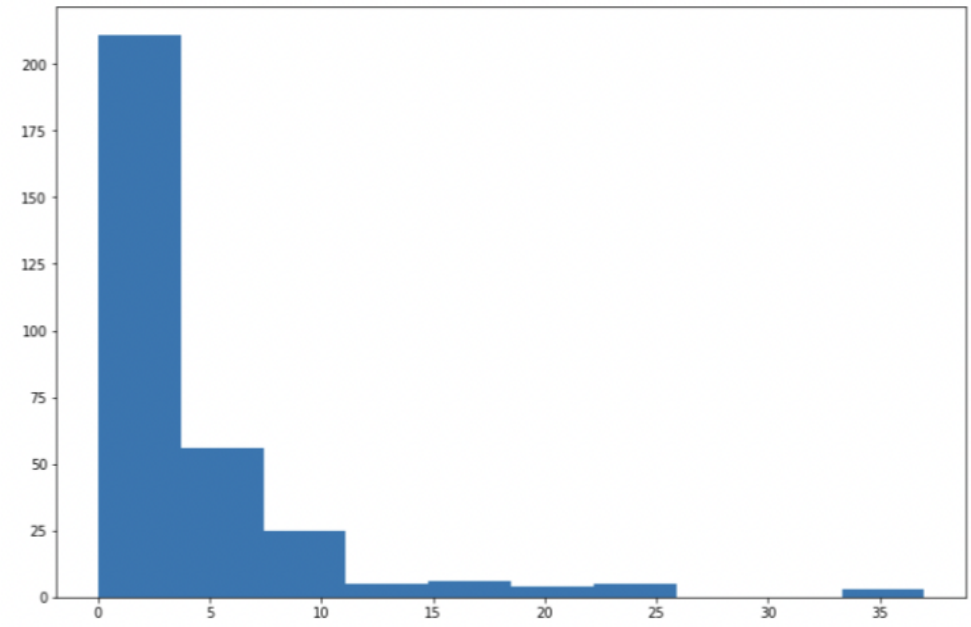

Figure 1 Degree distribution of the kinase signaling network

Discussion

Comparing the degree distribution of the drawn network to other networks it is mentionable that the given distribution was also found in a lot of other networks. But all the protein-protein interaction networks show a unique distribution pattern when considering the fact, that the main quantity of proteins are sparsely connected but still profoundly connected proteins were found as shown in this case. This is on one hand explainable by the gene duplication process merged with the condition of barely connection and it has also been proved that an increasing gene duplication process is increasing the number of profoundly connected proteins (Wu and Xu 2015).

The inhibition of MAPK1 results in the loss of 25 edges for the network since its degree is 25. An inhibition in this protein would affect broad majority of function of the network since it is one of the top 10 most important enzymes of the network with a score of 0.164. Considerably, it is also interacting with other of the topmost important proteins like MAP2K1, RAF1, EEF2K and EGFR which makes it activity crucial for the whole network communication system. In comparison to that the mutation in WEE1s 642nd residue has not changed the resulting analysis matrix since a mutation from serine to glutamine/glutamic acid considerably mimics phosphorylation due to their similar structures and thus did not have an effect as inhibition

factor to the whole system as the function keeps the same. Additionally, WEE1 is a less important protein of the system, and its downregulating would have minor effects in comparison to changes in proteins with more importance. When comparing the mutation of the 474th residue of AKT1 from tyrosine to phenylalanine it is suggestable that a decreased activation activation after stimulation would have been seen. However, it is also possible that the function will be completely downregulated. When seeing this in other proteins many studies some studies have found this to happen. For example, aniline triazole which has been investigated by Lippa et al., is structured like a phenylalanine subsequently will have equal functions and it was shown to inhibit a kinases function (Lippa et al. 2008).

A motif which definitely carried out a lot of representation in this network are self-loops since autophosphorylation is a type of post-translational modification of proteins (Petsko and Ringe 2009).

References

1. Aric A. Hagberg, Daniel A. Schult and Pieter J. Swart, “Exploring network structure, dynamics, and function using NetworkX”, in Proceedings of the 7th Python in Science Conference (SciPy2008), Gäel Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), pp. 11–15, Aug 2008.

2. Carling D. (2017). AMPK signalling in health and disease. *Current opinion in cell biology*, *45*, 31–37. https://doi.org/10.1016/j.ceb.2017.01.005

3. Kim, B. M., Yoon, W., Shim, J. H., Jung, H., Lim, J. H., Choi, H. J., Seo, M., Lee, T. H., & Min, S. H. (2017). Role of Protein Kinases and Their Inhibitors in Radiation Response of Tumor Cells. *Current pharmaceutical design*, *23*(29), 4259–4280. https://doi.org/10.2174/1381612823666170608084355

4. Kanev, G. K., de Graaf, C., de Esch, I., Leurs, R., Würdinger, T., Westerman, B. A., & Kooistra, A. J. (2019). The Landscape of Atypical and Eukaryotic Protein Kinases. *Trends in pharmacological sciences*, *40*(11), 818–832.

5. Lippa, B., Pan, G., Corbett, M., Li, C., Kauffman, G. S., Pandit, J., Robinson, S., Wei, L., Kozina, E., Marr, E. S., Borzillo, G., Knauth, E., Barbacci-Tobin, E. G., Vincent, P., Troutman, M., Baker, D., Rajamohan, F., Kakar, S., Clark, T., & Morris, J. (2008). Synthesis and structure based optimization of novel Akt inhibitors. *Bioorganic & medicinal chemistry letters*, *18*(11), 3359–3363. https://doi.org/10.1016/j.bmcl.2008.04.034

6. Z. Ma, X. Wu and H. Xu, "Understanding the degree distribution pattern of protein-protein interaction networks," *2015 International Symposium on Bioelectronics and Bioinformatics (ISBB)*, 2015, pp. 43-46, doi: 10.1109/ISBB.2015.7344919.

7. McKinney, W., & others. (2010). Data structures for statistical computing in python. In Proceedings of the 9th Python in Science Conference (Vol. 445, pp. 51–56).

8. Petsko, GA and Ringe, D 2009, 'Protein Structure and Function', Oxford University Press Inc., New York, U.S.A