Collaborative interdisciplinary work

You are facing the challenge of embarking on an interdisciplinary project where you are creating a collaborative project as an interactive process. It might not be clear how a project can be shaped as an interactive process rather than as something where the tasks are already pre-assigned form the very start.

We thought it might be helpful if we simply recount how we started. The project started with a biological question; Jacqui thought it might be interesting to look at caloric restriction and extended lifespan and did a bit of literature searching. This relates to concepts like intermittent fasting. It turned out that some of the early studies in this area used yeast as a model organism and found that SRT, which had already been seen to be associated with longer lifespans in yeast, appeared to be involved. There are several homologous proteins known in humans (and other species) which are the SIRT-proteins. (This is a bit backwards from your projects because although the question was about lifespan in humans we hit on the yeast protein before the human one, but we definitely had yeast and human homologs). The yeast species we are using has a paralog (protein that has arisen through gene duplication in the same organism), HST1 and we decided to use this protein as our starting point but it was hard to know what direction to take because HST1 makes nearly 500 PPIs – if Jacqui tried to manually trace these all it would be almost impossible (or at least take a very long time) and would be really prone to bias and easy to miss something interesting because it was a few steps out. So Jacqui really needed an unbiased network analysis to cope with all of the data that needed to be incorporated.

Georg then initially analysed a smaller network where he looked at the immediate 200 neighbours of HST1 and the next shell of 200 neighbours. He then ran a couple of centrality measures on this reduced network and presented his results to Jacqui who had a look at the protein being identified. When she looked at what had been identified she could see that the central nodes were actually parts of large protein complexes that we already knew HST1 could be part of, and/or it was highlighting central nodes – and we know that knocking out many of these would be lethal to the cell. So although the approach was coming up with the correct answers as far as the biology was concerned it wasn't really telling us anything new yet. Moreover, Jacqui questioned Georg's approach of restricting the analysis to only the immediate shells surrounding HST1 – this would bias the analysis to HST1 and would not allow us to find important nodes that come from other pathways of interactions that are further removed.

Georg then downloaded the whole network. He then did some analysis looking at the most central nodes. Since then the analysis of just looking at the most central nodes wasn't giving interesting new leads, he ventured into community detection algorithms. Some of the algorithms that he knew from books were terribly slow so Georg needed to find more efficient algorithms (by reading lots of papers). Georg then run a couple of algorithms and some gave back a partition into 4 or 5 communities only. Jacqui then told him that she estimated that there are 100's of functional protein communities or pathways with the yeast PPI network, so his communities were probably still too large and that we were hoping to find more subtle connections rather than just seeing connections between essential proteins which is what these initial analyses were showing (and when we thought

about it that made sense – centrality was probably always going to bias towards identifying the essential nodes) – was there a way to do this? Then Georg needed to think of how to further subdivide these communities.

In the mean-time, we were discussing what proteins/nodes one should focus on – the central hubs, or maybe those which connect important nodes? We also spent a lot of time talking to one another - what we thought was possible to test/what might yield more interesting leads and trying to explain our methods in laymen terms to the other one. We also spent a lot of time reading papers, trying to understand a bit more about what is possible in terms of computation and what types of networks we were really looking at. The first few papers on applying network theory to biology made no sense to Georg at all, but the more he read the more he got into the topic (some review papers he found helped a lot). From a Biochemistry perspective we realised that we are interested in trying to find some pathways of interactions that didn't take the shortest possible pathway from a biologically essential central node to another. The first approach would probably yield PPI leads that could be modulated to have an effect, the second would be more likely to kill cells because we'd be disrupting essential cellular processes.

Whenever Georg thought that he had found a good partition, he asked Jacqui if she could identify any biologically relevant proteins (which required him to write the data in a format which Jacqui could use for her analysis). Jacqui liked to look at the sets of proteins coming up overall, looking for functionalities that seemed to be heading in the right direction and trace through the series of PPIs connecting HST1 to other activities, and looking at the type of evidence for these interactions, initially in STRING. As we progressed, including steps like reanalysing a partitioned set, pre-screening by selecting higher confidence interactions we found quite a few interesting pathways and connections that linked HST1 to the types of protein communities that could potentially bring about changes in lifespan. For example, we found some connections to proteins that regulate telomeres, which attracted Jacqui's attention because telomers are parts of your chromosomes that get shorter as you age and are well known to be associated with lifespan. But we found several other leads too, to other types of cellular activities that were relevant and that we could follow up on – some were particularly relevant to yeast and not humans. As we got closer to interesting leads it became more important to look for the actual evidence for the interactions because STRING does bundle up a lot of information in a few terms, and it was important to distinguish cause versus effect.

We learned that this project asked us to try a range of different approaches – there was plenty of room for playing; the playground was constantly changing in a discourse of exchange of results, knowledge and ideas. It wasn't as simple as handing over a problem and the answer being handed back, we had to spend a lot of time understanding and reframing what we were asking of each other and limitations from both perspectives. We both found this a very rewarding and fun experience. Start playing!