# Project brief

We as a company are looking at developing a new focus area in drug discovery around the protein GSTM1.

We wish to make connections with researchers looking at this protein. We therefore need you to determine who the key active researchers are in this field. In order to avoid trying to work with researchers who are in competition with one another, we also need to build up a picture of who is working with who.

Our database team has assembled CSV files of papers which seem to relate to the gene of interest, including title, abstract, journal, year and information about the first author. A second CSV file provides the full author list for each paper. Can you give us some graphics to present to the board on who are the most active researchers in this field, who they collaborate with, and how research in this area has changed over time? If there is anything else you can find out from the data supplied that would also be helpful.