Discoveries hidden in big data

One way of looking at the complexities of the human genome is to turn it into a computer print-out, an approach that produces what is known as a 'transcriptome'.

No two transcriptomes are alike, and analysing their similarities and differences—especially when something goes wrong—is the job of computational biologists such as Associate Professor Jessica Mar.

Jessica heads up the Mar Group at the Australian Institute for Bioengineering and Nanotechnology, part of the University of Queensland. Her research focuses on developing statistical methods for understanding how genetic regulation goes wrong, and how that affects human disease.

And while some other genetics researchers approach their tasks using pipettes, microscopes and Petri dishes, Jessica and her colleagues deploy keyboards and screens.

"We use statistical models to understand big data and to understand what's going on with stem cells," she says.

"Big data can be very messy. We use statistics and computers to try to separate the signals from the noise, and to really understand what is going on with different stem cells, and stem cells at different stages of development.

"The aim is to try to understand which genes and pathways are involved in these sorts of phenomena."

Bioinformatics, as Jessica's field is known, is an area of increasing importance across science, and for one starkly simple reason. As the amount of raw data grows ever larger—a product of increases in information processing capability—the insights that can be gained from its analysis grow ever more detailed.

"When I started my career, we were lucky to get stem cell lines that were from maybe 10 patients," says Jessica. "Nowadays we can get them from easily hundreds, if not thousands."

This explosion in available information has been of particular use to one of her colleagues, Dr Atefeh Taherian Fard.

Atefeh studies gene expression related to ageing, and diseases associated with getting older. Her work is focussed on identifying the genetic markers involved in both processes.

It is a daunting task, and one in which separating signal from noise is particularly challenging. Success, however, could transform not only medical practice, but also countless lives.

"If we can find these markers, we would potentially stop those diseases happening," she says, "and that might prolong healthy ageing in each and every member of the society."