

The vast amount of data that is produced by gene expression studies poses a serious challenge for scientists studying human biology, says Carl-Johan Ivarsson from Qlucore. However, new software in this area is now helping researchers to perform sophisticated data analysis on proteomic, genomic and microarray data, without the need for specialist biostatisticians

“Modern data analysis software now enables researchers to analyse and explore extremely large data sets on a regular PC.”

Author Details:



Carl-Johan Ivarsson, President from Qlucore. Scientists involved in the project from the University of Nottingham; Genomics Centre at Kings College London; Department of Pathology at QMC; and the division of Ophthalmology and Visual Sciences, Nottingham.

New Approach to Genomic Data Analysis Helps to Unravel the Mystery of the Human Eye

Although gene expression studies are proving invaluable to the study of human biology, the amount of data that is produced by these experiments is enormous. As a result, it is impossible to derive any real biological meaning from these findings unless sophisticated data algorithms are used to help interpret the data effectively.

For this reason, most of the software that has been designed for use in this area has mainly focused on the ability to handle increasingly vast amounts of data, which means that the role of the scientist/researcher has been largely set aside. As a result, a lot of data analysis has been passed on to bioinformaticians and biostatisticians. However, in most cases, this model has several drawbacks, since it is typically the scientists themselves who know the most about biology.

Fortunately, the latest technological advances in this area are making it much easier for scientists to compare the vast quantity of genomic data generated, to test different hypotheses, and to explore alternative scenarios within seconds. Not only that, but the latest generation of data analysis software is also helping scientists to regain control of the analysis and to realise the true potential of gene expression profiling, which means that bioinformaticians and biostatisticians are free to focus on their own areas of interest and expertise.

USING GENE EXPRESSION ANALYSIS TO STUDY THE HUMAN EYE

Scientists who are currently studying the human eye at the Division of Ophthalmology and Visual Sciences at Queen's Medical Centre (QMC), part of the University of Nottingham, are already benefiting from this paradigm shift.

"Compared to what is possible with modern data analysis software, our analysis of microarray data from previous studies was more complex to analyse and difficult for biologists to understand," according to Dr Kulkarni, one of the researchers working at QMC. "Before, our analysis had to be performed on software for spot analysis and generation of raw data set files (which were then filtered and normalised), so that a range of filters and Interarray normalisation could be performed on the data set prior to the application of statistical analysis."

"However, performing microarray analysis is now easier, as the latest software provides instant graphical visualisation of the statistical tests in the form of Heat maps, as well as variable and sample PCA plots, which helps us to understand the analysis and the changes in gene expression patterns across different samples," Dr Kulkarni adds. "Not only is this software highly interactive, but it can also be easily understood by biologists, even if they had little or no previous knowledge of bioinformatics."

Professor H S Dua is leading QMC team that is studying the Limbal Stem Cells on the ocular surface of the eye. There is a large body of clinical and scientific evidence to suggest that corneal epithelial stem cells are located at the limbus. However, despite multiple studies in this area, definitive markers for limbal stem cells have remained elusive.

Professor H S Dua and his team had discovered a unique anatomical structure at the limbus; termed the Limbal Epithelial Crypt (LEC) [1]. The LEC is a solid cord of epithelial cells (between 50-100 microns long) that extends from the base of palisade of Vogt into the subepithelial tissue of the limbus, either peripherally or circumferentially.

Preliminary studies on the microanatomy and immunohistochemistry of the LEC strongly suggest that it has the potential to serve as the limbal stem cell niche [2], [3]. As such, the aim of the QMC study is to generate a gene expression profile of LEC in comparison with rest of the ocular surface epithelia with Human Gene ST 1.0 Array in order to find the putative stem cell marker.

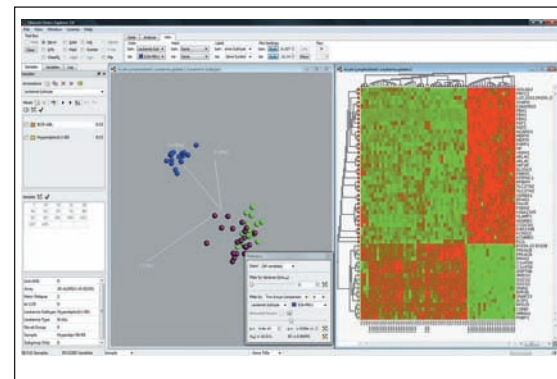
A team of six scientists are involved in this project, including

Professor Dua (supervisor), Dr Kulkarni (research student) at the University of Nottingham, as well as scientists from the division of Ophthalmology and Visual Sciences, Nottingham, the Genomics Centre at Kings College London (Dr Matthew Arno), and the Department of Pathology at QMC.

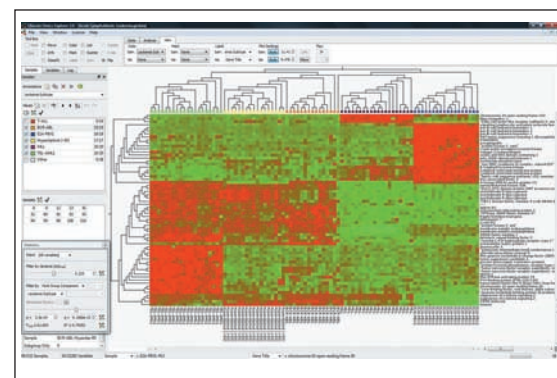
STEM CELL RESEARCH: HOW CAN TECHNOLOGY HELP?

The latest breakthrough in bioinformatics software enables scientists to analyse proteomic, genomic and microarray data with a combination of statistical methods and visualisation techniques such as Heat maps and Principal Component Analysis (PCA). With the benefit of instant user feedback on all actions, as well as an intuitive user interface that can present all data in 3D, scientists studying the structure of the human eye and other organs can now easily analyse all of this important information in real-time, directly on their computer screen.

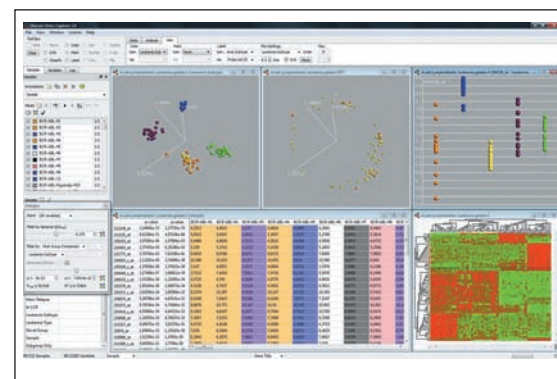
For all of these reasons, the latest software in this area is increasingly playing a key role in unveiling important new discoveries, as it allows the actual researchers involved – the people with the most biological insight – to study the data and to look for patterns and structures, without having to be a statistics or computer expert.



Software can be used to highlight the variables that discriminate the blue group from the rest



Plot types such as Heat map (pictured here) and Principal Component Analysis (PCA) are supported by the latest data analysis software



With the right software, the data contained in five synchronised plots can be viewed simultaneously

COPING WITH VAST AMOUNT OF DATA

Modern data analysis software now enables researchers to analyse and explore extremely large data sets on a regular PC, and to perform statistical tests to determine significant differentially expressed genes from the data set by eliminating background noise. This approach allows for linking the annotation to the genes and the calculation of values such as fold change false discovery rates (FDR) and p-values.

According to Professor H S Dua, this approach opens up many new possibilities, as previous research on the LEC was based on morphological and histological studies in which antibodies of known stem cell markers (identified from other studies) were used to identify limbal stem cells.

However, with these methods, it was not possible to perform quantitative protein analysis on the identified antibodies due to the small sample size of the tissue.

Recent advances in research techniques (such as laser microdissection) and microarrays from amplified RNA samples, however, have made it possible to perform gene expression study on a small amount of tissues like the LEC.

RECENT IMPROVEMENTS IN TECHNOLOGY

Most recently, perhaps within the past two years, the overall performance of data analysis software has been optimised significantly. In fact, with key actions and plots now displayed within a fraction of a second, researchers can increasingly perform the research they want and find the results they need instantly.

The programme being used by the team at QMC, called Qlucore Omics Explorer, is unique in that it is able to

generate PCA-plots between various sample data interactively and in real time directly on the computer screen, and to work with all annotations and other links in a fully integrated way, all at the same time. This approach has helped to open up new ways of working with the microarray analysis and, as a consequence, has helped the biologists to be involved actively in the analysis process.

MAKING SENSE OF IMPORTANT DATA

The Qlucore software graphically represents the high dimensional data in form of 3-dimensions plots on the computer screen. This instant visualisation technique is then combined with powerful statistical methods and filters, all of which are handled with just a single mouse-click. The points on the graph belonging to different subgroups or samples can then be coloured differently for identification purposes, so that these subgroups can be further analysed separately.

Qlucore's software has also been used to study protein array data, miRNA data, and RT-PCR data by other researchers. The software has also been used to analyse protein data from 2-D gels, image analysis data, and multivariate data sets of different sizes could be used.

THE FUTURE

The research project being conducted at QMC started with the discovery of the LEC, followed by its anatomical and immunohistochemical characterisation. This was followed by gene expression profiling of the LEC in comparison with surrounding ocular surface epithelia, such as the cornea and the limbus. This project will also serve to complement the existing morphological studies of the LEC, and also to

highlight geographical differences in gene expression of the LEC with the epithelium and the rest of the ocular surface epithelium regions.

The genes identified from the microarray analysis would be further evaluated for their potential for isolation and growing the limbal stem cells on suitable substrate for transplantation purposes in patients with limbal stem cell failure, a condition which occurs due to causes such as alkaline burns and autoimmune disorders like Steven Johnson syndrome.

Although current treatment modalities for limbal stem cell failure have proven benefits for a limited time period, patients often need to be kept on immunosuppressive drugs afterwards, and the failure rates in long-term periods are still quite high.

The ongoing research project at Queen's Medical Centre therefore has the real potential to progress from bench to bedside. The successful outcome of the microarray project on the LEC will pave the way for further research, which would contribute to enhancing the efficiency and the survival of the corneal grafts in limbal stem cell failure patients.

References:

- [1] Dua, H. S, V. A. Shanmuganathan, et al. (2005). *Br J Ophthalmol* 89(5): 529-32.
- [2] Shanmuganathan VA, Foster T, B Kulkarni, et al. (2007). *Br J Ophthalmol* 91(4): 514-9
- [3] Yeung AM, Schlötzer-Schrehardt U, Kulkarni B, Tint NL, Hopkinson A, Dua HS. *Arch Ophthalmol*. 2008 May; 126(5):665-9.