Spotlight

Proteomics, Genomics & Microarrays

New Gene Expression Research Hopes to Replace the Need for Animal Testing

Developments in this area could provide viable alternatives to animal experiments for the cosmetics industry and beyond

"Because there are not yet any in vitro tests or test strategies available to test chemical compounds on their potential to induce allergies, the aim of the project is to develop alternatives to animal tests currently used for the risk assessment of potential skin or lung sensitisers."

Author Details:

Carl-Johan Ivarsson, MSc,
President of Qlucore

Researchers at Lund University in Sweden are using next-generation data analysis software from Qlucore to conduct important new studies that aim to reduce or replace animal testing within the cosmetics industry and beyond.

Until now, it has been impossible to determine whether cosmetics such as make-up, shampoo, or moisturisers are likely to cause an allergic reaction unless tested on a living creature. With this powerful new software, however, scientists can test their theories by conducting their sophisticated gene expression research on computers, instead of relying of living organisms, such as lab animals.

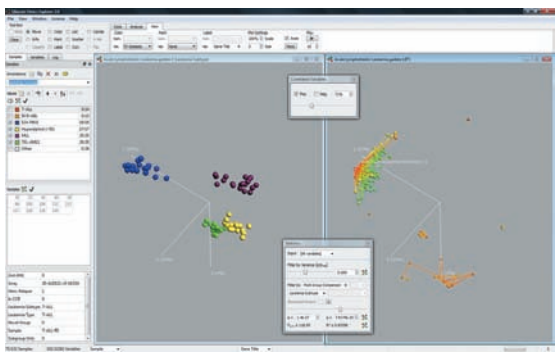
In 2005, a large EU-funded research project was launched to develop and optimise in vitro test strategies that could reduce or replace animal testing for sensitisation studies. By using a multi-disciplinary approach, this study is helping to address skin and lung sensitisation by focusing on the impact of compounds on cellular-molecular interactions, which play a central role in the development and elicitation of many allergies.

Because there are not yet any in vitro tests or test strategies available to test chemical compounds on their potential to induce allergies, the aim of the project is to develop alternatives to animal tests currently used for the risk assessment of potential skin or lung sensitisers.

The project, known as Sens-it-iv (http://www.sens-it-iv.eu/), combines both private and public research institutions, as well as several industrial and societal interest organisations. One of the key partners involved with the project is the European Centre for the Validation of Alternative Methods (ECVAM) at the Joint Research Centre. The presence of the ECVAM ensures a clear focus on tests and testing strategies that can be validated, which is a prerequisite for eventual regulatory acceptance.

Dr Ann-Sofie Albrekt is currently working in this exciting area, based at the Department of Immunotechnology, which is headed by Professor Carl Borrebaeck, a sub-coordinator of Sens-it-iv. As an internationally renowned centre for research and education, Lund University is highly regarded for its research in cutting-edge fields including nanotechnology, translational cancer research, and stem cell biology. Dr Albrekt is currently focusing on two key areas: looking for new biomarkers in cancer studies, as well as performing important research on allergens.

"Worldwide, more and more people are suffering from allergies, which means that this area has become an important health concern," said Dr Albrekt. "As a scientist, I am interested to find out why otherwise harmless compounds can often elicit an adverse immune response in humans."



When a number of interesting variables are found that responds to the effect the researcher are looking for it is natural to investigate further. Which variables are correlated to a specific variable? How do these relations match the common knowledge that the researcher holds about the subject. Are these findings new?

MAKING SENSE OF IMPORTANT DATA

Dr Albrekt is currently using sophisticated data analysis software, called Qlucore Omics Explorer, to help her to get the most value out of the data being produced by this research.

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

"Some data analysis applications can be very complicated and difficult to use, even for specialist statisticians, so it is very important to find software that has been developed by scientists, for scientists," Dr Albrekt said. "I am not a statistician, and yet I found the Qlucore software very easy to use, and without the need for any manuals or training. We started using the software straight away, and the fact that it is highly intuitive means that we were actually able to learn by using it."

ADVANCED DATA ANALYSIS: WHAT ROLE DOES SOFTWARE PLAY?

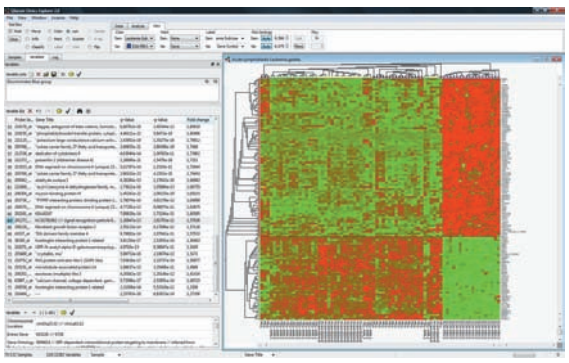
Sophisticated bioinformatics software now enables scientists to analyse very large data sets by a combination of statistical methods and visualisation techniques such as Heatmaps and Principal Component Analysis (PCA). With the benefit of instant user feedback on all actions, as well as an intuitive user interface that can present all data in 3D, scientists studying allergens and other aspects of human biology can now easily analyse their data in real-time, directly on their computer screen.

Modern data analysis software now enables researchers to use this approach with extremely large data sets – even those with more than 100 million data points – on a regular PC. This kind of specialist software can even take advantage of annotations and other links that are connected with the data being studied, as well as number statistical functions such as false discovery rates (FDR) and p-values.

As such, the research being conducted at the Department of Immunotechnology represents a significant breakthrough in how modern data analysis is being performed. Less than 10 years ago, researchers were only able to work with analysis methods that provided information about single genes. The number of information points per subject has grown to hundreds of thousands in recent years, however, thanks to important technological advances in this area.

RECENT TECHNOLOGICAL DEVELOPMENTS

Most recently, perhaps within the past two years, the overall performance of data analysis software has been optimised significantly. According to Dr Albrekt, modern data analysis software can be used to transform high dimensional data down to lower dimensions, which can then be plotted in 3-dimensions on a computer screen and rotated manually or automatically, so that they can be examined by the naked eye.



A second natural step would be to generate a list of a set of variables that can separate a group. In this example a list of the variables with highest fold change is generated. The p- and q-values are also listed. The researcher is provided with both the list and for instance a heatmap to be able to verify the results visually.

These instant visualisation techniques are combined with powerful statistical methods and filters, all of which are handled with only one mouse-click. Different colours can make this analysis even easier, as each sub-group can be labelled with its own unique colour. As such, the view of the data can be changed in an instant, so that researchers are only looking at the specific sub-group that interests them at any given moment. As a result, it is very easy to add and/or remove data as necessary, without having to start from the beginning and re-analyse the entire data set.

"When you are looking at such a large amount of genetic data, there is bound to be a number of confounding factors that distort the data," said Dr Albrekt. "The ability to remove this 'noise' is very important, in order for researchers to be sure that they are working with the most reliable data. Advanced data analysis software like Qlucore Omics Explorer makes it much easier to make a qualified judgment about the amount of noise present, so that researchers can see true patterns as they emerge."

In fact, with key actions and plots now displayed within a fraction of a second, scientists can increasingly perform the research they want and find the results they need instantly – without the wait. This approach has helped to open up new ways of working with the analysis and, as a consequence, has helped to bring the biologists back into the analysis phase, which means that bioinformaticians and biostatisticians are free to focus on their own areas of interest and expertise.

A STEP-BY-STEP APPROACH

According to Dr Albrekt, when performing her own research, she typically begins her workflow by coding any interesting factors (and confounding factors) into a single file. She then imports the data and looks at the pattern of samples in order to search for both anticipated and non-anticipated sub-patterns.

At this point, Dr Albrekt can begin to examine the sub-patterns using the coded factors that she had identified earlier. For example, by using the application's colour function and/or eliminating the factor function. Dr Albrekt can then look for any significant differences by using statistical tests.

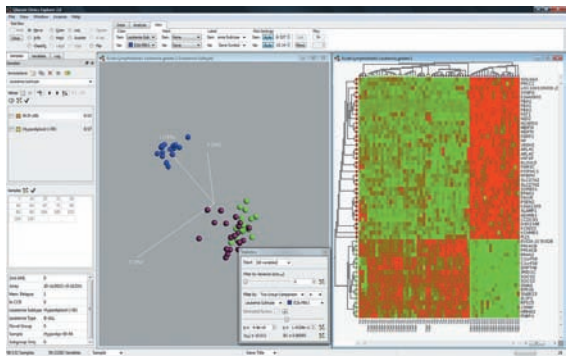
"We can test the robustness of these findings by using kNN visualisation, randomisation and permutation tools," Dr Albrekt explained. "That way, we can make a decision on which variables to trust, and then annotate any significant variables that we have found and export them for functional analysis using another software tool."

FLEXIBILITY HELPS BOOST CREATIVITY

With the freedom, speed and flexibility provided by this approach, it is now possible to evaluate and test a number of different scenarios and hypotheses in a very short time, and to fully understand the data being examined. This technique makes it possible for researchers to combine very large amounts of data, and therefore to conduct analysis in ways that were simply not possible before.

"In our studies, we are dealing with very large amounts of data, sometimes between 10 and 100 million data points, which we tend to view as graphics. With other software, these graphics would take a long time to appear, but with the latest data analysis tools, the information is presented instantly," Dr Albrekt said. "As a result, we can be much more creative with our theories, as we can easily test any number of hypotheses in rapid succession."

Although Dr Albrekt is currently using data analysis software to study gene expression micro array data, other researchers have used it to study protein array data, miRNA data, and RT-PCR data as part of their research studies. This kind of software has also been used to analyse protein data from 2-D gels, image analysis data, and in fact with any data set of multivariate data of sizes up to 1000 samples and 100,000 variables, or 1000 variables and 100,000 samples.



Here is Qlucore Omics Explorer being used to rapidly identify which which of the variables is responsible for separating the blue group (which can represent any type of grouping in a data set). The variables are identified and selected using a t-test. The results are then visualised both by a CA plot (Principal Component Analysis) and a heatmap plot. This gives the researcher a good basis for combining his/her own knowledge with the results of the study.

THE FUTURE

The latest technological advances in this area are making it much easier for researchers 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 the scientists (biologists, MD, and so on) to regain control of the analysis and to realise the true potential of gene expression profiling.

According to Dr Albrekt, her own research efforts will continue to focus on both the Sens-it-iv allergen studies, as well as on the ongoing cancer research within CREATE Health, which is a strategic centre for translational cancer research.

"In terms of the work we are doing to support Sens-it-iv, I feel confident that a successful project outcome will contribute to the reduction in the number of animals required for safety testing and the establishment of more accurate tools for product development," she said. "This project will therefore be of substantial benefit to all European citizens, and that goal continues to motivate me to make new discoveries in this area."

Diverse Molecules Engineered for High-Quality Fragment Screening-Based Programmes

Thermo Fisher Scientific introduced the latest addition to its Maybridge Fragment Range, the Maybridge Ro3 Diversity Fragment Library. The new library has been computationally engineered to build on the key customer-driven features of the original Maybridge Ro3 Library, such as 'Rule of Three' (Ro3) compliance, pharmacophoric enrichment and quality assurance of at least 95%.

The Maybridge Ro3 Diversity Library offers both an improved structural diversity profile and experimental solubility data for every one of the 1,500 member compounds. Full Ro3 compliance means that all 1,500 fragments have the physicochemical properties that increase the probability of successful 'hits'. The broad portfolio also provides access to analogues for fragment hopping, as well as reactive analogues for rapid hit evolution and derivatisation. Consequently, these Maybridge Fragments provide an ideal starting point for a lead-optimisation, structure-based drug discovery programme and the ultimate evolution of a new molecule in the drug discovery process.

Every fragment in the new Maybridge library has been experimentally triaged to assure solubility in both DMSO (200mM) and aqueous phosphate buffer (1mM). Aqueous solubility is a prerequisite for successful in-vitro testing, as poor solubility can compromise the robustness of the screening data through aggregation and promiscuous inhibition. Fragment hits with poor aqueous solubility are also more likely to produce evolved analogues with inferior ADME properties, such as plasma protein binding, poor systemic distribution and an increased likelihood of candidate attrition. The Maybridge Ro3 Diversity Fragment Library mitigates all such issues. When building the new Maybridge library, Maybridge scientists applied Daylight-based structural fingerprinting techniques (1024 bit length) to an Ro3 compliant set of 8,000 fragments and the fingerprinted set then clustered using industry standard clustering software (DBClus). The resulting representative set of 1,500 high-quality molecules consists of 819 cluster centroids and 690 singletons at a Tanimoto level of 0.66.