Management and Analysis of Biomarker Data

Aiden Flynn, PhD

Director

Exploristics Limited

Introduction

There is great expectation in the Pharmaceutical Industry, by Regulatory Authorities, Medical Practitioners, Healthcare Providers and amongst patients that the era or Personalised Medicine will deliver safer and more effective treatments (Jørgensen & Winther, 2009). This promise is made possible by the emergence of new technologies and data acquisition platforms that can generate enormous amounts of patient-level biomarker data to characterise biological processes and explain response to treatment. The routine application and integration of genetics, genomics, transcriptomics, metabolic and imaging data is now possible within most clinical research programs.

Aside from the potential to develop safer and more effective drugs, Pharmaceutical companies are interested in biomarkers as they offer the potential to explain variability in drug response and therefore provide more options in development programs. There are now numerous examples of drugs that have failed to demonstrate a clinically meaningful effect in all-comers but have used biomarkers to identify subgroups of patients that derive benefit (Risner et al., 2006; Weber & McCormick, 2008). In addition, biomarkers have been successfully used to improve the safety profile of drugs (Hughes et al., 2004). Regulatory Authorities also recognise this potential and are increasingly requesting that biomarkers are an integral part of drug development (Lesko & Woodcock, 2004).

Whilst there is much interest in the use of biomarkers and the new tools make the data readily available, there remain many challenges to be overcome before the potential of biomarkers can be fully realised. A lot of biomarker research in recent years has been driven by the ability of biomarker platforms to generate increasingly large volumes of biomarker data. However, the capabilities to perform the best analysis of these data are still being developed. These analysis methods need to go beyond the traditional statistical approaches used and need to account for the biological relationships in the biomarker data as well as the statistical associations with other clinical variables. In addition, the increased demand for the analysis of biomarker data stretches the available analysis resources. Furthermore, there is the need to develop and implement standards, processes and infrastructure to enable the storage and integration of biomarker data in conjunction with other clinical data.

Exploristics, the innovative analysis company, offers a range of solutions to enable the integration and application of biomarker data in both an exploratory, hypothesis generation setting as well as in the regulated, hypothesis testing environment. We provide flexible resourcing options to support the analysis of biomarker studies. We facilitate the development of in-house capabilities through consultancy and training. We have also partnered with OmicSoft Corporation to offer state-of-the-art software that can be used by experienced statisticians and bioinformaticians as well as biologists. This case study describes the capabilities and functionality of two software packages, Array Studio and Array Server, offered through our partnership with OmicSoft.

Array Studio

Array Studio provides state of the art statistics and visualization for the analysis of high dimensional gene expression data and genotype data. It provides the fastest, easiest, and most powerful solution for -omic data analysis on the market. Array Studio supports numerous platforms and provides complete statistical analyses for most experimental designs. More than 400 features have been implemented based on feedback by industrial and academic users. Array Studio is highly interactive (Figure 1) and was designed so that biologists can function at the level of informaticians.

Figure 1: The highly interactive genome browser allows interrogation of results with quick access to information about genes, HapMap information, and other customizable views.

Array Studio also provides comprehensive support for project management (Figure 2), data manipulation, quality control, pathway analysis, gene ontology analysis, and power analysis. There is also an internal audit trail for workflow tracking.

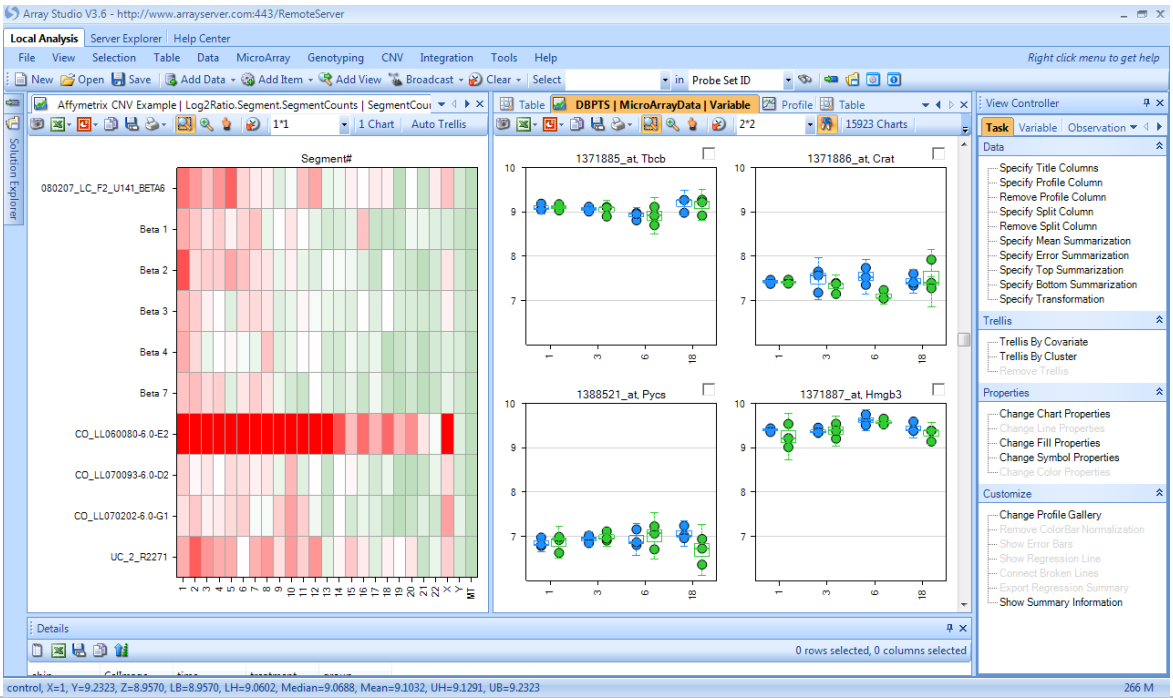

Figure 2: Multiple projects can be viewed at once. In this view, the results of a Copy Number Variation (CNV) segmentation are viewed alongside the Variable View of a gene expression project.

Array Server

Array Server is an enterprise solution, allowing users to store, share, search, and integrate their microarray/SNP/CNV projects and data. Results can be easily shared with clients and colleagues. Array Server enables cross project/platform data **storage,** data **search** and data **integration and** allows the generation of multi-project visualizations and charts, including analyzing SNP, CNV, and Microarray data together.

Array Server may be integrated alongside Array Studio, or accessed via the free Array Viewer software (for easy sharing of data with clients). There are over 5000 fully analyzed Affymetrix GEO/Array Express projects with every purchase of Array Server. All projects can be searched to find genes with similar properties (Figure 3 & 4).

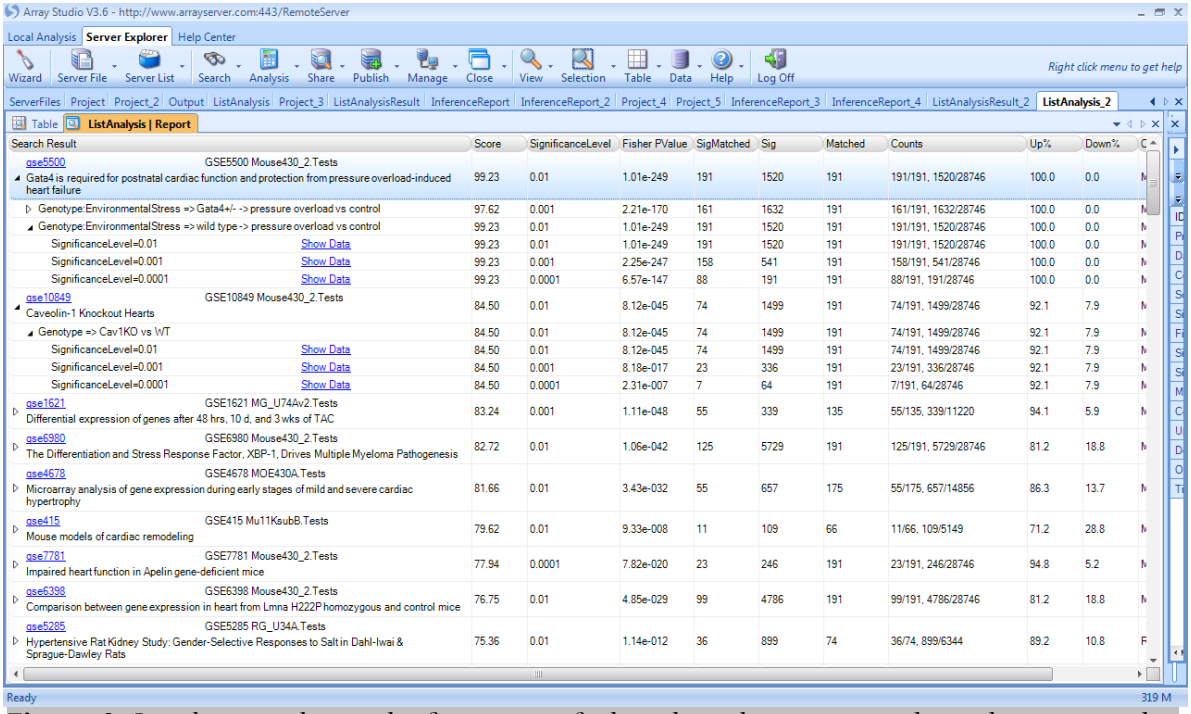

Figure 3: Search across thousands of projects to find similarity between input list and project results.

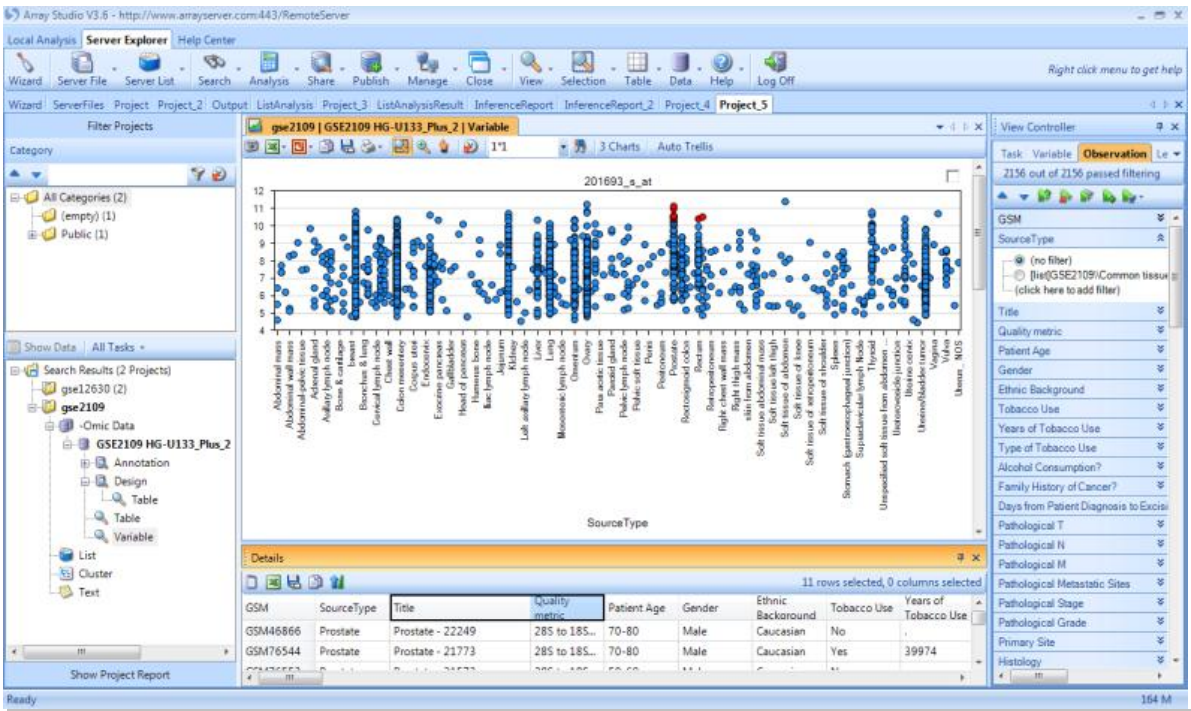

Figure 4: Query and filter by gene, disease, sample attributes (i.e sex, age, etc.) fold change, p-value, etc. and return interactive graphs. Outputs can be easily exported to PowerPoint.

Summary

The era of Personalised Medicine offers much promise for the development of safer and more effective drugs. However, there remain many challenges before this potential can be realised. Exploristics, the innovative analysis company, offers solutions to the management, storage, integration and analysis of large volumes of biomarker data. These solutions include the provision of analysis expertise, knowledge transfer through consultancy and training, and the development and implementation of standards, processes and strategies for the use of biomarker data in clinical studies. This case study describes the solutions provided through our partnership with OmicSoft: Array Studi o and Array Server provide easy-to-use, state-of-the-art capabilities for the storage, management, integration, analysis and visualisation of high-dimensional gene expression and genotype data. For more information about our services and Array Studio and Array Server, visit our website (www.exploristics.com) or contact us on info@exploristics.com.

References

Jørgensen J. and Winther H. The new era of personalized medicine: 10 years later (2009). Personalized Medicine, 6(4); 423-428

Lesko L., Woodcock, J. Translation of pharmacogenomics and pharmacogenetics: a regulatory prespective (2004). Nature Reviews, 3: 763-769.

Hughes A et al. Association of genetic variations in HLA-B region with hypersensitivity to abacavir in some, but not all, populations. Pharmacogenomics 2004;5:203-211

Risner, M et al. Rosiglitazone in Alzheimer's Disease Study Group. Efficacy of rosiglitazone in a genetically defined population with mild-to-moderate Alzheimer's disease. (2006). Pharmacogenomics Journal, *6: 246-54.*

Weber, J and McCormick, P. Panitumumab In Metastatic Colorectal Cancer with Wild-Type KRAS (2008). Biodrugs, 22 (6); 403-411.