

Data Analysis Services

Core Competency

WITH THE FLOOD OF DATA PRODUCED BY NEXT GENERATION SEQUENCING (NGS), MICROARRAYS, AND OTHER GENOMIC PLATFORMS, DATA ANALYSIS IS A BOTTLENECK FOR LIFE SCIENCE RESEARCHERS. FINDING AND HIRING A STAFF WITH THE RIGHT BACKGROUND IN BIOLOGY, STATISTICS AND DATA MINING IS BOTH TIME-CONSUMING AND EXPENSIVE. RESEARCHERS BRING IN 5AM'S DATA ANALYSIS TEAM – TOP-NOTCH BIOINFORMATICIANS – TO SOLVE THEIR MOST PRESSING ANALYSIS NEEDS QUICKLY.

5AM SOLUTIONS' DATA ANALYSIS SERVICES

Genomic technologies are evolving rapidly. Unfortunately, user-friendly tools and the training required to interpret the data have not been able to keep up. Let 5AM's bioinformatics team become your partner in making sense of your data. Our team has over 25 years of combined experience interpreting and analyzing biological data using tools such as R, Matlab and Spotfire, and fine-honed analytic skills. Having worked at world-renowned genome centers, pharmaceutical and biotechnology companies, 5AM bioinformaticians are expert users of most genomic tools and resources. 5AM engages in small to large analysis projects and will deliver one-time analysis results, tool assessments (or creation), analysis workflows and/or analysis training, depending on your requirements.

Examples of typical projects include:

BIOMARKER DISCOVERY

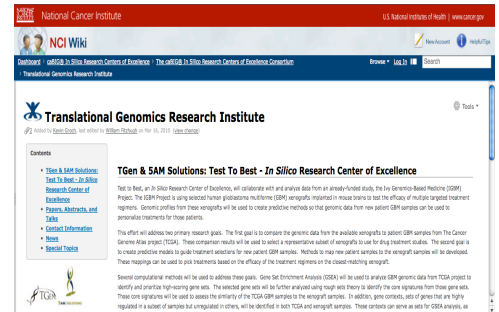
A client has gathered a collection of gene expression measurements acquired through RT-PCR assays run across cohorts of disease and control samples. The client wishes to identify biomarker signatures that can discern disease from control cases. 5AM works with the client to design an experiment that will meet their desired statistical power. 5AM performs data QC and cleaning to identify any abnormalities or unwanted experimental artifacts in the data. Our team then writes and executes analysis workflows to identify informative genes and combine gene expression results as diagnostics classifiers. Cross-validation or independent training and test sets are used to ensure meaningful results are found and to predict future diagnostic performance. These techniques are applicable to a range of other assay types, including quantitative protein assays and genetic variation data.

NEXT GENERATION SEQUENCING

A client has a next-generation sequencer and has bought commercial tools for running read mapping and downstream analysis. 5AM works with the client to assess results from the analysis tool and compares results to publicly-available tools. Our team works with the client and the scientists who receive the data to understand the results. 5AM staff has additional experience using next-generation sequencing data for copy number analysis and metagenomic studies.

MICROARRAY ANALYSIS

A client has Affymetrix gene expression data for a set of cancer related cell lines, but they lack the internal expertise needed to analyze the data themselves. 5AM works



VITAL STATS

GOT SCIENCE?

Biomedical research generates high volumes of complex and sophisticated data. Successful efforts to produce new therapeutics and diagnostics require *informed* use of this data from public and proprietary sources. To this end, our bioinformatics and science analysis team provides expertise in the tools and brings a deep knowledge of the data and data sources. We mentor, advise, and collaborate with our clients throughout the entire research and development process, from lab automation and assay design to data management, analysis and visualization.

EXAMPLE CLIENTS

- Top Ten Pharmaceutical Companies
- Diagnostics Start Ups
- Mid-Size Biotechs
- Translational Genomics Research Institute

TOOLS

- R & BioConductor
- Matlab
- Spotfire
- Perl, Python & Java

LINK

- <https://www.5amsolutions.com/science/>

(866) 526-6042 | www.5amsolutions.com

SOFTWARE FOR LIFE™



with the client to instruct them how to perform array QC, identify differentially expressed genes while accounting for multiple testing, and find relevant pathways through gene enrichment analysis. 5AM helps the client install necessary tools, teaches them how to use the tools and leaves them with workflows customized to their data and analysis needs. After the project, 5AM remains a dedicated partner, addressing questions and issues that arise as the client takes ownership over their analysis.

COMPUTATIONAL METHOD DEVELOPMENT

5AM is collaborating with the Translational Genomics Research Institute on a project to identify biomarkers for predicting treatment efficacy in brain cancer. The *In Silico* Research Center of Excellence is funded by the National Cancer Institute with the goal of developing algorithms and software tools for use in mining the vast amounts of data coming out of genomics studies of cancer, such as the Cancer Genome Atlas. 5AM works with researchers at TGen to develop visualization tools and analysis pipelines that make their algorithms more streamlined and the results more easily interpretable. 5AM is dedicated to publicly releasing data and software developed for this exciting project.

DATA INTEGRATION

A client wishes to integrate multiple public and internal genomic and proteomics resources for the purpose of discovering new potential drug targets. 5AM develops a custom software framework and database to organize the data. Working closely with the client, 5AM identifies relevant resources and develops tools to extract and normalize the data and sample annotations for integration. 5AM recruits a team of expert curators to read through scientific literature and parse out relevant data and confidence values. Utilizing R and Spotfire, 5AM runs meta-analyses and creates visualizations demonstrating targets that are marginal in individual studies but in aggregate are compelling.

DATA VISUALIZATION

Biological data is generally large and complex, and finding the right answers can often feel like hunting for a tiny needle in a vast haystack. Carefully-crafted visualizations can bring the right answers into focus quickly, saving time and effort. Visualizations can also help you communicate results and the value of your project to other stakeholders, or help make a point clearly for publication. 5AM has the expertise and tools to help you build powerful visualizations from your data analysis pipeline.

EXAMPLE KEY BIOS

Will FitzHugh, 5AM's Chief Science Officer, has over 17 years of experience working on software and data analysis in life sciences and biotechnology. He has worked on projects such as the Human Genome Project, model organism genome sequencing and annotation, cancer proteomics, and genome-wide association studies. Will is skilled in gathering requirements for analysis and software projects and is an ideal liaison between scientists and software/IT staff. He has extensive experience with a wide range of genomic and biomedical resources such as Ensembl, Genbank and PubMed. He has used such tools as Matlab and Perl to explore data and develop analysis methods.

James ("J") Ireland is a bioinformatician and data analyst with 15 years of professional experience in biotech, pharmaceutical and academic institutions. J led the bioinformatics team at The Whitehead Institute, Center for Genome Research (now the Broad Institute) for one of the first large scale SNP discovery and genotyping efforts in human genes. J has designed and analyzed genotyping and expression microarrays while at Affymetrix, Novartis and Avalon Pharmaceuticals. Among his accomplishments, J has designed commercial MIP genotype panels at ParAllele, methylation and somatic mutation assays at Affymetrix and resequencing PCR designs for SNP discovery at the Whitehead Institute, Center for Genome Research (now the Broad Institute). J is the author of the primer screening tool pgrep and is responsible for the assay design featured in more than ten journal articles and patents.

Andrew Evans brings over two decades of software engineering experience to 5AM. In addition, With his graduate-level coursework in molecular biology and bioinformatics, Andy is a valuable liaison between software and science teams. His specialization is in information architecture, user experience, and data visualization – skills he has applied in the Fortune 500 and on the founding teams at companies like CareerBuilder.com. Andy also serves as Technical Lead at the Translational Genomics Research Institute's *In Silico* Research Center of Excellence, and is the 2010 Visiting Lecturer in Bioinformatics at the West Virginia School of Osteopathic Medicine.

Viktoria Gontcharova, Ph.D., has worked as a bioinformatics scientist, software developer, and consultant. Through Viktoria's professional and academic careers, she has done research, analysis, and developed software in both Windows and Unix environments using languages such as Perl, Python, Matlab, R, and C#. She created a 454 post-processing pipeline, that included chimera detection and statistical analysis capabilities. Through the years, she wrote and coauthored a number of papers, often dealing with pyrosequencing topics and machine learning methods. Most recently, Viktoria has worked on microarray analysis and probe design.