Agilent GeneSpring GX 10: Gene Expression and Beyond

Pam Tangvoranuntakul
Product Manager, GeneSpring
October 1, 2008
GeneSpring GX 10 in the News

Agilent Technologies Introduces Next-Generation Bioinformatics Software Providing Systems-Level Microarray Analysis

GeneSpring GX 10.0 Offers Multiple Applications, Pathway Analysis for Biological Insight

SANTA CLARA, Calif., Sept. 22, 2008

Agilent Technologies Inc. (NYSE: A) today introduced Agilent GeneSpring GX 10.0, the next generation of Agilent's high-throughput gene expression bioinformatics platform. GeneSpring GX, considered the gold standard of desktop gene expression analysis, now offers tools for systems-level data interpretation and pathway analysis, enabling scientists to attain a new level of insight into the underlying mechanism of disease or biological process.

"We developed this new revision of GeneSpring GX to address the growth of systems-level studies in functional genomics research," said Bruce von Herrmann, Agilent vice president and general manager, Software and Informatics Division. "GeneSpring has always been excellent at gene expression data analysis. Now, it has evolved into an integrated software package that helps scientists discover the relationship between the data and the biological event occurring in their samples."

Beyond Gene Expression

GeneSpring GX has earned a solid reputation among gene-expression biologists, with more than 4,400 references in Google Scholar, including more than 1,600 in peer-reviewed publications. As systems-level studies become more prevalent in genomics research, GeneSpring GX 10.0 adds visualization and analysis tools for applications including alternative splicing, microRNA expression and real-time PCR. To get more insight into the underlying mechanism of disease, GeneSpring now has powerful pathway analysis capabilities. By providing a database of gene product interactions, scientists can build biological interaction networks from their genes of interest.

GeneSpring also provides statistical tools to detect significant changes in microRNA abundance, and integrates TargetScan microRNA gene-target information to determine the biological consequences of detected changes. Additionally, to facilitate validation of microarray experiments, GeneSpring GX 10.0 supports analysis of real-time PCR data.

Enterprise Solution for Systems-Level Research

To support enterprise functional genomics data analysis needs, Agilent offers GeneSpring Workgroup, a client-server application with GeneSpring GX as the data-analysis engine providing central and secure storage for GeneSpring data and analysis results.

"GeneSpring Workgroup has the flexibility to act as an enterprise ‘on-prem’ data warehouse while also allowing developers to build additional functionality using the extensive APIs," said Ketan Patel, Ph.D., associate research fellow, Computational Sciences, Pfizer Inc., an early-access user of GeneSpring GX 10.0. "This is the main reason why we favored this product, as it allows us to integrate the system with our existing databases and tools and also build out new functionality using scripting in languages such as R, Java and Python. The APIs allow us to extensively search the database using sample annotation attributes. This was very important to us and will enable broader use of omics data within the enterprise."

GeneSpring GX 10.0 Brochure

GeneSpring GX 10.0 User Guide
Our Goals for GeneSpring GX 10

**Goal 1: Bring back GeneSpring GX 7.3.1 features most requested by users**
- Update annotations directly from NCBI
- Normalization options: control genes, scaling, baseline transformation to multiple control groups
- Ability to plot raw intensity values
- And many more…

**Goal 2: Improve features in GeneSpring GX 9**
- Automate GX 7 to GX 9 Migration tool
- Hierarchical Clustering: improve visualization and navigation, color branches by parameters, mouse-over to show correlation coefficient
- Flexible GO annotation formats
- And many more…

**Goal 3: Add analysis tools for new applications**
- Analysis of miRNA expression using Agilent miRNA microarray data
- Analysis of alternative splicing using Affymetrix Exon Array data
- Analysis of ABI real-time PCR data
- Network analysis using interactions from Pathway Architect knowledge base or interactions extracted from text using NLP
New Applications and Analyses:
miRNA, alternative splicing, pathway analysis, and real-time PCR
MicroRNAs (miRNAs) may regulate up to 90% of the human transcriptome. miRNAs regulate expression through translation repression and mRNA degradation. Recently, evidence of miRNA inducing gene expression has been presented. Changes in miRNA expression can lead to changes in gene expression that mediate the underlying mechanism of a biological process or disease.
Agilent miRNA Data Analysis in GeneSpring GX 10

Support for Agilent miRNA data
- Download miRNA technologies in GeneSpring GX
- Full text files from FE automatically recognized

Guided Workflow for Agilent miRNA array data
- Normalization
- QC on Samples using Agilent Feature Extraction (FE) QC metrics
- QC on miRs using Agilent FE probe QC metric
- Statistical Analysis
- Fold change analysis
- GO Analysis on significant miR targets
- Pathway Analysis on miR targets
Biological Contextualization of miRNA Data

What are the targets of these differentially expressed miRs?

Is there an enrichment of these targets in any GO term?

Is there an enrichment of these targets in any pathways?

Biological Knowledge:
- Intracellular trafficking and cell motility are affected and may mediate disease process under study
- TGF-beta signaling and cell-cycle regulation pathways are affected and may mediate disease process under study
Gene targets of differentially expressed miRs can be translated into a corresponding gene expression experiment to view their expression levels.
Approximately 75 percent of human mRNAs undergo alternative splicing.

Splice variants can produce proteins with distinct and sometimes even antagonistic functions.

A number of genetic mutations associated with human disease map to changes in splicing signals or sequences that regulate splicing.

Thus, measuring changes in splicing patterns is integral to understanding the disease mechanism or biological process under study.
Alternative Splicing Analysis in GeneSpring GX 10

Support for Affymetrix GeneChip Exon Arrays

- Perform both gene-level (gene expression) and exon-level (alternative splicing) analysis
- Summarize Core, Extended, or Full probe sets using RMA, PLIER, or IterPLIER
- QC on samples and QC on exons using Affymetrix QC metrics
- Identify alternative splicing activities using the multivariate splicing ANOVA
Identification and Visualization of Differentially Spliced Transcripts

Transcripts with significantly different splicing events are identified using a multivariate splicing ANOVA. Significant transcripts can be filtered on splicing index value ("fold change" analysis at the exon level).

Each transcript can be plotted using:

- Variance Plot- profile represents average intensity values or gene normalized intensity values for each condition across probes of the chosen transcript (Red: Normal, Blue: Tumor)
  - Probesets on x-axis are ordered according to their chromosomal location
- Profile Plot- profile represents intensity values for each probeset in each condition
Real Time RT-PCR Data Analysis in GeneSpring GX 10

Support for ABI 7900 real-time PCR platform

• SDS 2.1, 2.2, and 2.3 RQ file formats supported
• RQ Manager 1.2 file format supported
• All analysis tools available to apply to data (statistical tests, fold change analysis, clustering etc)
• Results of real-time RT-PCR data can be directly compared to microarray gene expression data in GeneSpring GX 10
Improved Data Translation Tool

Expression data and list associated values for list also imported to experiment you are translating to.

Entity list associated values and raw or normalized intensity values from original experiment can be plotted in target experiment using Plot Entity List Associated Value tool.
Plot Entities by Entity List Values

Plot Entities from up to two Entity Lists in Scatter Plot, Profile Plot, and Histogram

All Entity List associated values and raw or normalized intensity values can be plotted for those Entities

If two Entity Lists are used, the union of those lists are plotted

Entity List translated from other experiments can also be plotted and intensity values are those from the other experiment

Color entities by any Entity List values
Compare Results From Maximum of Three Experiments Using Venn Diagram

Project up to three Entity Lists from different experiments (of different platforms or species) to the Venn Diagram.

Translation between platforms and species is performed automatically.

Allows comparison of results between different experiments.
Pathway Analysis in GeneSpring GX 10

Find Significant Pathway analysis has been added to each Guided Workflow

Full integration of tools in PathwayArchitect

• Import and view Biopax pathways and perform Find Similar Pathways analysis as in GX 9
• Create interaction network using entities of interest
  – Direct interaction network, expand network, shortest path network, network regulators, and network targets
• Create relevance interaction network using entities of interest
  – Generate an interaction network with an associated (high) confidence index
  – Is the number of interactions between my entities of interest greater than what would be expected by chance?
• Overlay expression data onto pathways
• Use Natural Language Processing (NLP) algorithm to mine full text files and Medline abstracts in PubMed to create new interactions and pathways
Superimpose Expression Data onto Pathways
Pathway Analysis in GeneSpring GX 10
Simple and Advanced Pathway Analysis
Ingenuity Pathway Analysis (IPA) Plug-in

Ingenuity Pathway Analysis (IPA) scripts can now be launched from a link in the Workflow browser

- Create Pathway in IPA
- Perform Data Analysis on Experiment
- Perform Data Analysis on Entity List
- List of genes from IPA can also be imported into GeneSpring GX
Data Migration
Automated GX 7 Migration Tool

**Step 1: Prepare for GS7 Migration** - tool automatically prepares data for migration

**Step 2: Select GS7 genome**
- GS10 - all experiments, samples, interpretation, gene lists, trees, parameter values, condition values, and classifications will be automatically migrated

**Step 3: Open Project with name corresponding to GX 7 genome to see the migrated data.** Note that if genome was assigned a project in GX 7, this name will be the name of the project in GX 10 instead of the name of GX 7 genome.
Automatic GX 7 Data Migration

When migration is finished running, a summary report is displayed

- Information for how many objects are supposed to be migrated versus how many were migrated is reported
- Each error is logged
Other New and Improved Features

- alternative splicing
- gene expression
- real-time PCR
- miRNA
- pathway analysis
Technology Annotation, Data Organization, Import, and Export

Options for creating and/or updating technology annotations:

- From Agilent Server
- From Agilent eArray
- From GeneSpring Update file
- From Biological Genome
- From text file

Support for .gpr files

Addition of a My Favorites folder

Copy/Paste export option in most result tables
Normalization and Filtering Options

Normalization options:

• Normalize to control genes
• Scaling
• Baseline transformation to multiple control groups

Filtering options:

• Filter on Volcano Plot (implemented as its own link in Workflow browser)
• Filter by Error (coefficient of variation and standard deviation)
• Filter on Data Files
• Filter by Expression (raw and normalized intensity values or percentile)
Data Visualization

Hierarchical clustering trees

- Navigator view of entire tree
- Zoom capability to make navigation easy
- Color branches by parameter and show coloring blocks
- Mouse over to see degree of similarity and gene annotation
- Ability to extract sub-tree

Mouse-over to get gene annotation in views

Addition of fold-change lines to Scatter plot

MvA plot
GeneSpring Workgroup 10
Changes in Workgroup 10

- Version number changed to align with GeneSpring version number
- Pathway interaction database incorporated into Workgroup
- My Favorites Folder added to data hierarchy of Workgroup
GeneSpring GX 10 e-Seminar Series

Register for live e-seminars at [http://genespring.com](http://genespring.com)

Topics for scheduled e-seminars:

- Introduction to GeneSpring GX 10 New Features (Oct 1st)
- Joint Analysis of miRNA and Gene Expression Data in GeneSpring GX (Oct 8th)
- Analysis of real-time PCR data in GeneSpring GX and Comparison to Microarray Results (Oct 22nd)
- Identifying Significant Differences in Alternative Splicing Events in GeneSpring GX (Oct 29th)
- Network Analysis in GeneSpring GX- Building Relevant Networks from Genes of Interest (Nov 5th)