FEATURE: South African Start-up Electrifies Clustering Software Market

Claire Bisseker

CAPE TOWN, South Africa, Nov 22 - South Africa's first and only bioinformatics start-up, Electric Genetics, hopes to corner the market for clustering software with the release next month of updated versions of its core products, StackPack and Stackdb.

The products combine to form a gene clustering system that is part of the standard armory of many academic genomics researchers worldwide and has achieved growing commercial acceptance since its commercial release early last year.

The StackPack data assembly and transcript analysis system started out as a rather cumbersome academic research tool. Completely re-engineered by Electric Genetics, StackPack version 2.0 was released in July this year as a robust commercial product.

A peer-reviewed system, StackPack's data management and visualization tools enable the rapid clustering, alignment, and analysis of high volumes of ESTs and partial-length sequences. They highlight transcript variation and can accelerate gene discovery, gene function analysis, expression analysis, and drug development.

StackPack is the tool set and Stackdb the database it creates. The latter is regarded as one of the most thoroughly processed representations of expressed human genes in the world. Cambridge and Harvard Universities and the Pasteur Institute are among its 200-odd users.

The release of StackPack v2.1 is now on the horizon. It can accommodate data sets 10 times as large as version 2.0, is significantly faster, and has improved data visualization capabilities. Stackdb v3.0 has been constructed using StackPack v2.1 and will be released next month with free viewing software in relational database format.

“By having a technologically superior product and a strong service ethic we expect to capture the rest of the market and even some of our competitors’ customers,” said Electric Genetics managing director Tania Hide, sitting at the desk in her office at the University of the Western Cape in Cape Town.

“One thing that makes our product offering really unique is that we are the only group to offer both a database of clustered ESTs and the tools to create that database. These tools can then be applied to any set of data.”

Stackdb’s biggest competitors are the Institute for Genomic Research’s THC database and UniGene, produced by the US National Center for Biotechnology Information. However, UniGene offers only groups of sequences and does not provide alignments of clusters or detailed information about variation or alternative forms of transcripts represented by the clusters. THC focuses on making short, highly accurate consensus sequences while Stackdb strives to maximize consensus length.

StackPack competes with DoubleTwist's CAT system, which has over 20 users, all commercial. StackPack claims over 55 licensed users, both academic and commercial.

Electric Genetics provides its products and support to academics for free, in line with its aim of furthering genomics research. Academic feedback is then incorporated into fine-tuning the process.

Electric Genetics has been operating as a private company since January 1999 but its roots are firmly embedded in academic research, the company having grown out of the need to commercialize products developed by the South African National Bioinformatics Institute, a research group based at the University of the Western Cape, Cape Town.

The symbiotic relationship between SANBI and Electric Genetics is in no small part due to the fact that Tania Hide, who has 10 years of marketing experience in the bioinformatics industry, is married to SANBI director, Win Hide, a former director of genomics at MasPar Computers.

Lion Bioscience gave Electric Genetics a big thumbs up, with a deal to integrate StackPack into its SRS data integration platform, which allows for advanced querying across a number of biological databases.

"They [Electric Genetics] combine a thorough approach to tackling scientific problems with an excellent sense for bringing the final product to the market," said Clemens Suter-Crazzolara, Lion Bioscience's product manager.

The deal, Hide hopes, will also help to drive up sales.

"We believe that the more tightly integrated our product is with other scientifically respected products, the more value it will have for the customer," she said. "If the customer perceives more value, there will be more sales."

In order to help marketing efforts beyond South Africa, Electric Genetics has signed four non-exclusive distribution deals with Lion in Germany, eBioinformatics in Australia, Teijin Systems Technology in Japan, and AFS Informatics in California.

In addition, Electric Genetics' distributors tend to provide its products in a value-added way. For instance, eBioinformatics has integrated Stackdb into their BioNavigator online portal and is negotiating to integrate StackPack into BioNavigator, while Teijin has on occasion packaged both products with integration services to make them more useful.

With several big sales already under her belt and more in the pipeline, Hide expects earnings to increase several-fold by the year's end. She declined to offer current revenue figures. The company, which has received government funding, is currently seeking venture capital backing and is considering a share offering on the Nasdaq in two to three years time. Hide said she was also planning to open a US office in the near term.

Such steps could help to give Electric Genetics the money and access it needs to grow. Right now the company must find novel ways to deal with the paucity of bioinformaticists in Africa and the limitations of most biologists' computer programming skills.

Hide has sought out intellectually curious programmers with Internet and networking backgrounds who then work side by side with the biologists and bioinformaticists at SANBI. Each concentrates on what he or she does best with Hide acting as a bridge between the two disciplines.

“So far this approach has worked very well and we feel we are able to work more efficiently than groups that insist on hiring only bioinformaticists,” said Hide who points out that the total development time of StackPack v2.1 was a mere 10 weeks.

The company employs nine people, five of whom are computer programmers and none of whom are bioinformaticists.

Electric Genetics is also planning to expand its product offerings. The company is planning the first implementation of Lincoln Stein’s Distributed Annotation Server, which will allow researchers to publish their human genome sequence annotations over the Internet. Electric Genetics has undertaken to release the code for its initial DAS implementation freely to the scientific community before the end of November.