Sander van der Laan Implements SVS for Greater Efficiency and Ease-of-Use

Sander van der Laan, a doctoral candidate at the Interuniversity Cardiology Institute Netherlands (ICIN, *www.icin.nl*) based at the University Medical Center Utrecht (UMCU, *www.umcutrecht.nl*), performs genomic association studies with the goal of uncovering biomarkers and drug targets for atherosclerosis. Upon initiating his PhD work, van der Laan realized that he needed an analytical tool that was better suited to the manipulation of large-scale genomic data than open-source offerings. Van der Laan found SNP & Variation Suite suited his research needs exactly. "With Golden Helix, my frustration went out the door," he says.

Atherosclerosis is a chronic, often asymptomatic disease caused by the arterial plaque formation. Van der Laan conducts his research on atherogenesis and its markers within the university's experimental cardiology laboratory (*www.umcutrecht.nl/experimentalcardiology*). Using whole-genome SNP and CNV analyses coupled with proteomics data, van der Laan aims to identify and validate atherosclerotic plaque markers that are predictive for adverse cardiovascular events, such as stroke or myocardial infarction.

Van der Laan's studies are conducted in association with the Athero-Express Biobank (*www.atheroexpress.nl*), a collaborative, multi-institutional study involving hundreds of samples and millions of SNPs. Trained as a medical biologist, van der Laan came to this data-intensive work after spending time as an intern at the Erasmus Medical Center Rotterdam (*www.erasmusmc.nl*) in their bioinformatics department. After spending months using open-source tools to clean and export data, he realized that he needed a clean start.

"My internship was six months, and in that time, I was able to get from raw data to something you could work with," van der Laan says. "However, as soon as my internship finished, I knew the data wasn't good enough for reliable results. When I began my PhD, I knew I had to start over."

In the experimental cardiology lab at UMCU, there are sixty researchers who are primarily focused on proteomics; van der Laan has sole responsibility for "making sense" of the genomic data. Using Affymetrix' Genome-Wide Human SNP Array 5.0, the biobank collected blood-based DNA and plaque from 836 patients with atherosclerosis; cursory quality assurance yielded just more than 700 samples. Van der Laan planned to impute the Affymetrix 5.0 data using BEAGLE and BEAGLECALL, which would result in 2.5 million SNPs. He wanted a data analysis tool that would be both easy-to-use and compatible with multiple file formats.

At that point, van der Laan recognized the limitations of relying solely on command-

QUICK FACTS

Name: Sander W. van der Laan

Title: PhD Student

Institution: Interuniversity Cardiology Institute Netherlands at the University Medical Center Utrecht

Location: Netherlands

Studying: Atherosclerosis

Data: Affymetrix 5.0

Number of samples: Over 700

Number of SNPs: Imputed to 2.5 million

Analysis: SNP, CNV, and GWAS

SVS: Power Seat

Utrecht, Netherlands

line applications, such as PLINK and R, to complete his analyses. So he started searching for a more intuitive tool that would be capable of importing and exporting various file formats. However, most commercial offerings were better suited to expression pathway discovery and other proteomics applications. He considered one popular commercial tool, but it was neither user-friendly nor particularly suited to SNP analysis.

Sander W. van der Laan

After a brief evaluation period, he ultimately chose Golden Helix's SNP & Variation Suite (SVS) software. Within two months, van der Laan had clean data in a format he could work with.

This productivity gain exceeded van der Laan's initial expectations. Anticipating the full breadth of results required by the Athero-Express study, he chose to purchase the Power Seat of SVS. Without these tools, van der Laan imagines that he would still be working on quality assurance today—one year since he began his PhD work.

Van der Laan appreciates the software's extensive capabilities, in addition to it being user-friendly and capable of manipulating data in multiple formats. This is indispensable in large-scale genomic studies. Whereas free command-line tools crashed on large datasets, SVS accommodated van der Laan's data seamlessly. Learning how to use the software was easy, he says, thanks to Golden Helix's free training, tutorials, a live support team, and an informative website.

Moreover, van der Laan considers SVS ideal for data management, as it can convert and customize annotation library tracks as he explores his data. Because the software takes advantage of his computer's full capabilities, van der Laan's analyses are accelerated without compromising accuracy. Additionally, van der Laan found that the flexibilty of being able to export various file formats to work with programs such as BEAGLE was invaluable.

Since van der Laan began using SVS one year ago, several powerful quality assurance methods have been implemented through regular updates from Golden Helix. Thanks to intuitive and customizable graphical user interfaces, he can make his project's data accessible to colleagues without the need for prior training or coaching. Van der Laan has already written and integrated his own customized script for strand-flipping using SVS's Python script editor.

Currently, van der Laan plans to initiate both CNV and whole genome SNP analyses. Some exciting results regarding the predictive power of certain proteins on atherosclerosis have already been found. He hopes to finish his doctorate in the next three to four years, and then perhaps begin a career as a geneticist intermediate between academia and biotech.

In the meantime, his PhD supervisor continues to demand new genomic analyses as new results are uncovered daily. With the analytic power of SVS applied to his store of SNPs, van der Laan now feels like he can meet these expectations head on.

About Golden Helix

Founded in 1998, Golden Helix is known for helping genetic research groups working with large-scale DNA-sequencing or microarray data overcome the frustration and challenges of bioinformatic roadblocks: delayed projects, lack of quality findings, and low productivity. By empowering researchers with highly effective software tools, world-class support, and an array of complementary analytic services, we refute the notion that analysis has to be difficult or time consuming. Golden Helix's flagship software product, SNP & Variation Suite (SVS), is an integrated collection of powerful data management, quality assurance, visualization, and tertiary analysis tools for genetic data. SVS is delivered in a user-friendly, scalable platform and is supported by a team of highly trained bioinformaticians, statistical geneticists, and computer scientists that together make advanced statistical and bioinformatic methods accessible to scientists of all levels.