

Vanessa Hayes, PhD and Desiree Petersen, PhD Use Population Genetics to Dig into Mankind's Roots

Humans are said to have a 0.1% rate of diversity¹, seemingly small, but actually quite significant when considering the future of personalized medicine. Gaining a deeper understanding of this genetic variance can help determine susceptibility to disease as well as medicinal response and outcomes.

One-third of human genetic diversity can be found in a likely contender for the birthplace of mankind: Southern Africa. Africa as a whole boasts over 2,000 distinct languages and cultures.¹ The magnitude of ethnic diversity was a signal for Dr. Vanessa Hayes, Dr. Desiree Petersen, and their team at the J. Craig Venter Institute (www.jcvi.org). For that reason, it has been their mission to make sure that African genomes are examined and represented sufficiently in databases worldwide.

In 2010, Hayes and Petersen teamed with Dr. Stephan Schuster of Pennsylvania State University (www.schusterlab.com) and with Philip Venter, Professor in the Faculty of Medical Sciences at the University of Limpopo in South Africa (www.ul.ac.za) in a collaboration to put African genomes on the genetic map, so to speak, in an effort to expand the definition of human variation. The Southern African Genome Project (www.jcvi.org/cms/research/projects/southern-african-genome-diversity-study) stemmed from frustration regarding the limited representation of African genomes in existing databases, such as HapMap, 1000 Genomes, etc. Since Africa has the most diverse genetic data, they believe it should be included in genomic research to more fully understand genetic variance.

Genetic admixture introduced new genetic lineages when previously separated populations began interbreeding. The study of admixture patterns among the oldest lineages in the world not only

expands knowledge of genetic diversity, but also deepens the understanding of origins of mankind and cultural diversity. For example, within-population and between-population genetic diversity among the click-speaking foraging people of southern Africa was recently featured in an article by Petersen and Hayes in *PLoS ONE*: "Complex Patterns of Genomic Admixture within Southern Africa."²

But their research has not just put African genomes on the map; it has also given insight into TB, polio, and prostate cancer. Archbishop Desmond Tutu was the first African sequenced and gaining his acceptance was key in accessing the African people's genetic data. Tutu was a survivor of TB, polio, and prostate cancer, and his participation allowed the genetic markers potentially associated with his diseases to be correlated with his family and medical history.³

Hayes and Petersen have been involved in nearly every step of the study from collecting samples to genotyping, which has been very time consuming. Since time is so important, the team immediately

QUICK FACTS

Name: Dr. Vanessa Hayes and Dr. Desiree Petersen

Title: Principal Investigator and Staff Scientist, respectively

Institution: J. Craig Venter

Location: San Diego, CA

Studying: Genetic diversity among African populations

Data: Illumina HumanOmni1-Quad BeadChips

Number of Samples: 103 individuals representing 5 southern African populations

Number of SNPs: Almost 1 million

Analysis: Population genetics

SVS: DNA-Seq Analysis Package



A rendering of the new J. Craig Venter Institute building in San Diego, California

focused on areas that consumed too much of it.

At the onset of the project, the team was spending too much time merging data from different software programs and needed to find a single, optimal data management solution to ensure data quality and performance. Additionally, the team wanted a solution that did not require the user to be a code-writing statistician and instead, allowed them to focus on the science.

Having used Golden Helix's SNP & Variation Suite (SVS) at a prior job in Australia, Hayes decided once again to implement the platform for their work in population genetics, which SVS is not typically used for. The team found that SVS was just



Dr. Vanessa Hayes



Dr. Desiree Petersen

way to do what we need, even if it isn't what the software was made for," says Petersen.

what they were looking for: user-friendly with great visualizations and a support team that is responsive and willing to go the extra mile.

Petersen says, "Working with the Golden Helix Support Team is very helpful. If we need to do a specific analysis in SVS that isn't currently available, we can consult with Golden Helix, and they will create a script for us." Further, the team found that using a single program instead of shifting between several saved them time and manpower.

Hayes, Petersen, and team have been using SVS for three years now and are very pleased. "We always find a

And it is a good thing that Hayes and Petersen can continue their work because as population genetics constantly evolves, the team will be able to add even more insight on population admixture and help to identify novel genetic markers for disease.

¹ Interview with Dr. Vanessa Hayes at the 2011 Cancer Research Awards. Accessed on April 10, 2013. <http://www.youtube.com/watch?v=aS29LZMNQoI>

² Petersen, D et al. (2013) Complex Patterns of Genomic Admixture within Southern Africa. *PLoS Genetics*, 9(3):1003309, doi:10.1371/journal.pgen.1003309. <http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.1003309>

³ Sheahan, L (2010) "Scientist Spotlight: Meet Vanessa Hayes." JCVI Weblog. Retrieved April 18, 2013 from <http://blogs.jcvi.org/2010/11/scientist-spotlight-meet-vanessa-hayes/>

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.