

Complete southern African genomes available to everyone

On the 18th February 2010 the first two complete genomes of southern African individuals were published and became immediately accessible to the world. In addition to the genomes of one Bantu-speaker (Xhosa/Tswana) (Archbishop Desmond Tutu) and one San (!Gubi) from Namibia, the exomes of a further three San octogenarians were compared to the reference human genome and all documented human genetic variation. An astounding 1.3 million novel genomic variants were detected, confirming the extensive untapped source of genetic variation in African populations. The project was led by Stephan Schuster (Pennsylvania State University, USA) and Vanessa Hayes (Children's Cancer Institute, Australia) and involved a further 46 coauthors, the majority of who were involved with data analysis. Some observational gems include the facts that: all four San are homozygous for the allele that is associated with lactose intolerance; the Duffy null allele which has almost gone to fixation in black African populations (proposed to be due to selection for protection against *Plasmodium vivax* infection) was absent in the four San; and the taste receptor gene (TAS2R38) bitter taste haplotype was present in all San tested, suggesting selection for taste discernment, important in hunter-gatherers. Perhaps one of the most surprising findings has been the apparent distant relationship between the four San, possibly reflecting their long term geographic isolation despite their small population sizes.

It is important to be reminded that a handful of individuals does not represent a population and these studies are, therefore, the prelude to genomic science on the sub-continent and an opportunity to start exploring the relevance of the rich genetic variation of Africa's peoples in the context of health and disease. Genotype phenotype association studies in the context of extensive collaborations between molecular scientists and clinical researchers would be crucial for the success of such studies. This should be viewed as an opportunity to build local scientific capacity and develop locally relevant diagnostic and prognostic indicators of health relevance.

In South Africa there is a paucity of skills in many disciplines essential for conducting genomic, transcriptomic and proteomic research and these include molecular genetics, population genetics, high throughput biology, data analysis, biostatistics and bioinformatics. However the vast diaspora of Africans working in the world's top institutions could play a crucial role in developing greater capacity on the sub-continent and ensuring that the rich biological resources are tapped together with those who work among the African peoples for their benefit. In this context we wish to encourage members of the Southern African Society of Human Genetics to get involved in large population-based studies and we encourage you to consider working with your local peers in addition to international collaborators, in an effort to build further capacity on the sub-continent.

Publication: Complete Khoisan and Bantu genomes from southern Africa

Schuster SC, Miller W, Ratan A, Tomsho LP, Giardine B, Kasson LR, Harris RS, Petersen DC, Zhao F, Qi J, Alkan C, Kidd JM, Sun Y, Drautz DI, Bouffard P, Muzny DM, Reid JG, Nazareth LV, Wang Q, Burhans R, Riemer C, Wittekindt NE, Moorjani P, Tindall EA, Danko CG, Teo WS, Buboltz AM, Zhang Z, Ma Q, Oosthuysen A, Steenkamp AW, Oostuisen H, Venter P, Gajewski J, Zhang Y, Pugh BF, Makova KD, Nekrutenko A, Mardis ER, Patterson N, Pringle TH, Chiaromonte F, Mullikin JC, Eichler EE, Hardison RC, Gibbs RA, Harkins TT, Hayes VM.

Nature. 2010 Feb 18;463(7283):943-7

Web based access to data: Galaxy site - <http://main.g2.bx.psu.edu/bushman>

SASHG Committee
February 2010