The Genomics of Schizophrenia in the South African Xhosa population

Dan Stein, Department of Psychiatry

Raj Ramesar, Human Genetics, University of Cape Town

Ezra Susser, Psychiatry/Epidemiology, Columbia University, New York

Mary- Claire King, Genetics, Washington University, Seattle, USA

- Background: originally granted by NIMH as RO1 (regrettably not in format of typical H3A project)

- Preliminary/pilot studies
  - HT SNP Chip studies handful of pts/controls
  - Preliminary Study Identification of CNVs
  - Exome sequencing in Xhosa persons with schizophrenia and Xhosa controls
  - Identification of Private Truncating Mutations in Schizophrenia Candidate Genes

- Current status

Schizophrenia

- Common neuropsychiatric condition
- Occurs across all populations
- High heritability

Genomics of schizophrenia research to date

- High number of people with rare CNVs
- 8-fold increase risk of de novo CNVs
- Rare CNVs significantly more common
- Many genes of small effect
- Focus thus far mainly on CNVs >100kb

Significance

- Closing the gap
- **Capacity building**
- Change genetics research worldwide – new candidate genes
- Address health disparities in Africa – Africa currently under-represented in genetics research

The amaXhosa people

Notable Xhosa: Nelson Mandela, Thabo Mbeki, Desmond Tutu

Total population

7.9 million (2001 estimate)

Regions with significant populations

Eastern Cape: 5.4 million

Western Cape: 1.1 million

Gauteng: 0.7 million

Free State: 0.25 million

Kwazulu-Natal: 0.22 million

(2001 estimates[1])

Languages

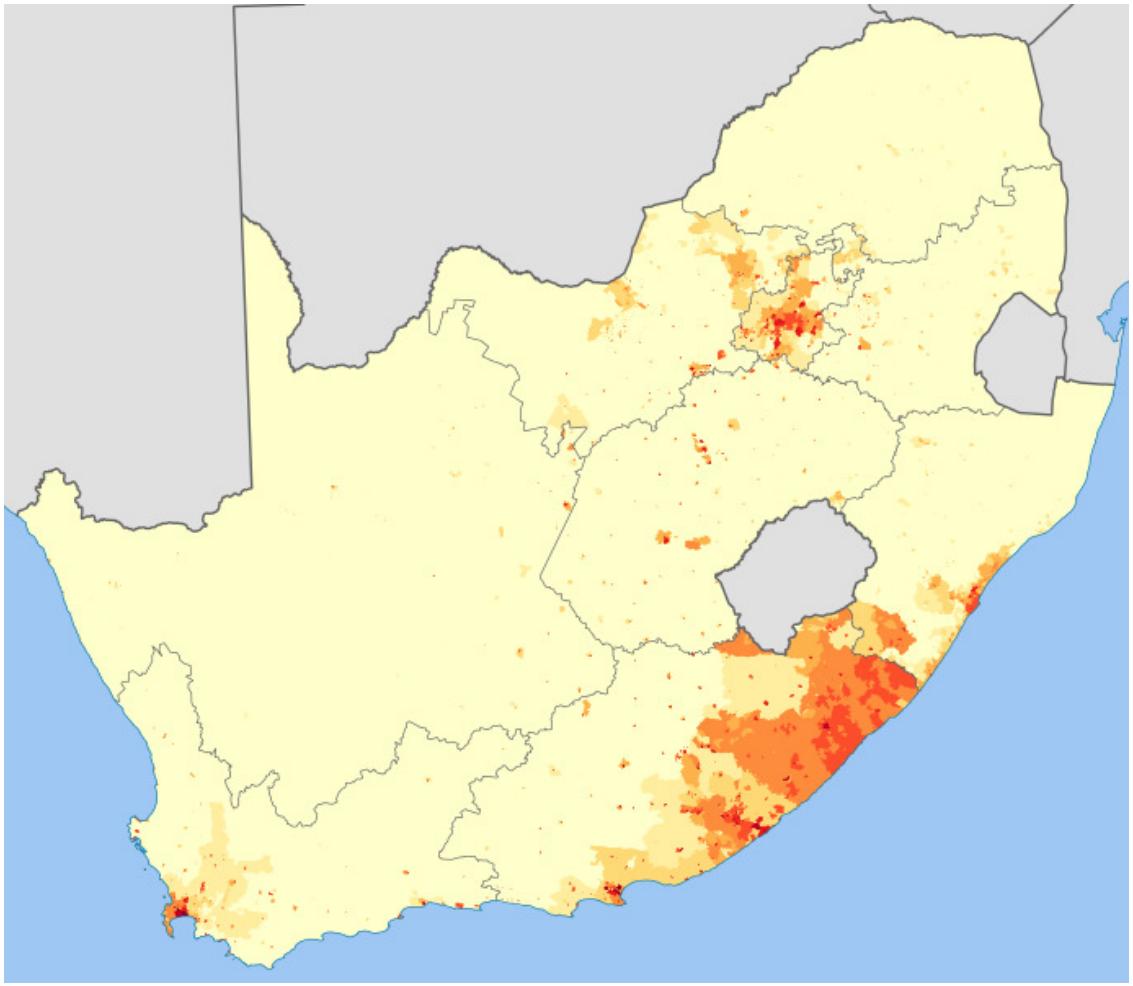
Xhosa (many also speak Zulu, English, and/or Afrikaans)

Religion

African Traditional Religion, Christianity

Related ethnic groups

Nguni, Basotho, Zulu, Khoisan



Language: isiXhosa

Aims

- Aim 1

  Ascertain, assess and enroll 1100 matched Xhosa cases and controls

- Aim 2

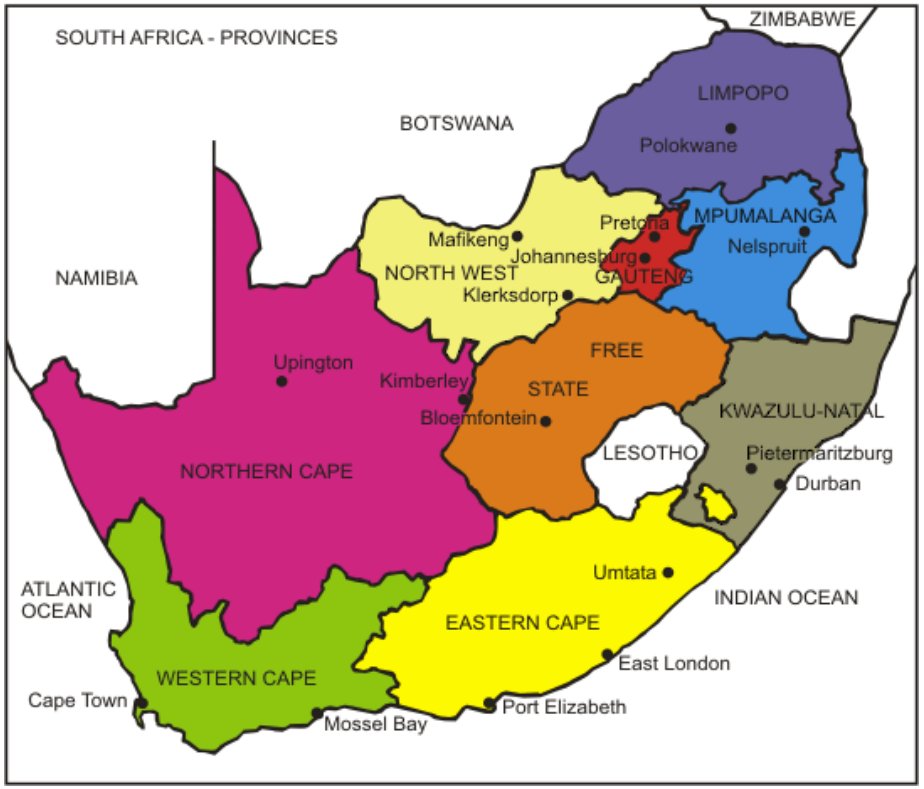
  Utilise Exome (200 pts + 200 controls) and
  GWAS (1100 cases + 1100 controls) strategies to identify a range of variants that impact coding regions

- Aim 3

  Identify genes enriched for damaging mutations in Xhosa individuals with schizophrenia as compared to controls

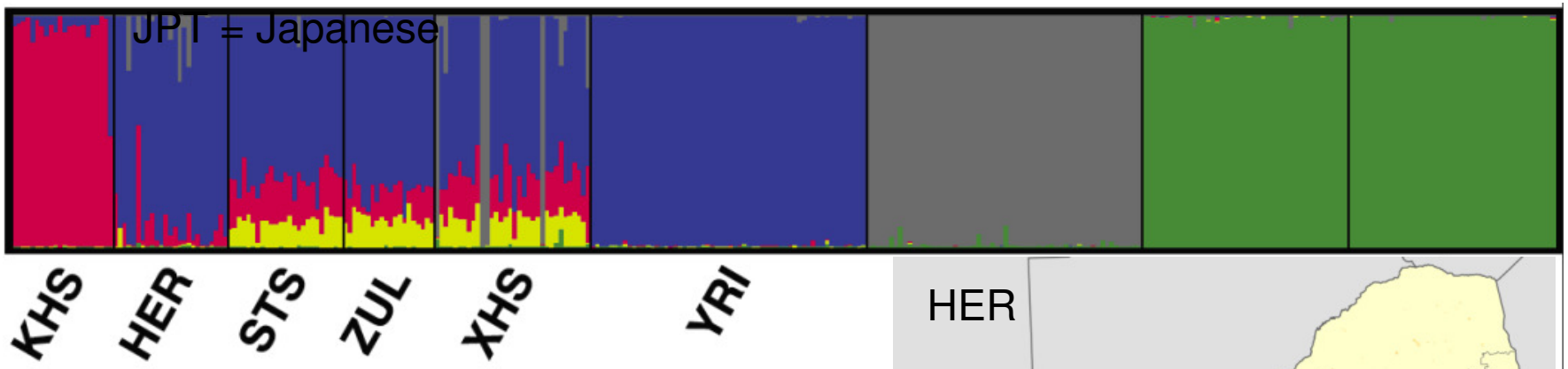
Recruitment

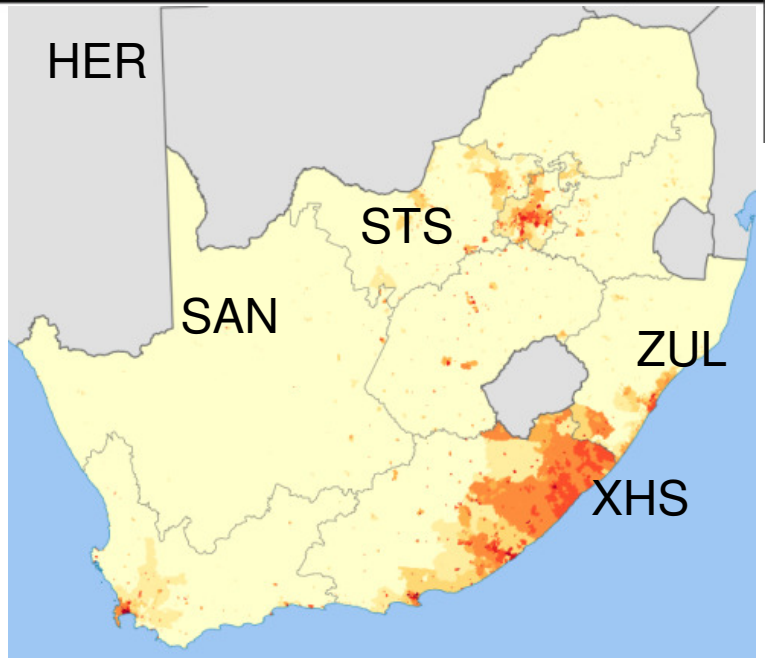
- 1100 Cases + 1100 matched Controls
  - Half from Cape Town,
  - Half from Eastern Cape
- 100 Trios
- Assessed using structured clinical interview (SCID)
- Neurocognitive assessment



Subjects

STRUCTURE analysis of the genetic relationship between South African Xhosa people and other population groups
KHS = Khoisan, HER = Herero, STS = Sotho-Tswana, ZUL = Zulu, XHS =Xhosa, YRI = Yorubans, CEU = Caucasians, CHB = Chinese, JPT = Japanese





CPGR and CBIO @ UCT

Preliminary Study/Logistics: Genome Wide SNPs, and Exome sequencing in Xhosa persons with schizophrenia and Xhosa controls

- cases / controls
- DNA extraction at UCT
- Aliquots shared
- Affy SNP6 (UCT)/Nimblegen HD2 Arrays (WU)
- HiSeq platform (WU)
  - Sequenced exomes, flanking splice sites and non-coding RNA
  - 42MB Nimblegen exome v2 pool

Pilot Study – Results of CNV Identification

- Novel, gene-impacting CNVs
- Good parallel analysis capacity
- Compatible technology

Pilot Study – Identification of CNVs

- Screen 10 cases and 86 controls (including the above) @ UCT (Afffymetrix6.0) and WU (Nimblegen HD2)
- Each lab called CNVs independently
- Cases filtered against controls
- CNVs retained if
  - Impacted on gene
  - Did not overlap >20%

Table 2. CNVs in Xhosa cases not present in Xhosa controls or the DGV

Chr	Type	HD2 size (kb)	HD2 probes	Affy6.0 size (kb)	Affy6.0 markers	Gene	Case ID
1 q25	gain	29	27	57	34	CEP350	5
1 q32.2	loss	56	29	28	17	KCNH1	16
1 q42.2	loss	28	25	19	24	KCNK1	19
2 p16.3	loss	18	17	25	15	NRXN1	12
2 q14.2	gain	335	280	309	254	GLI2	16
2 q37.1	loss	52	36	13	12	UGT1A	18
18 q22.1	loss	243	193	238	136	CDH19	16

Pilot Study – Identification of CNVs

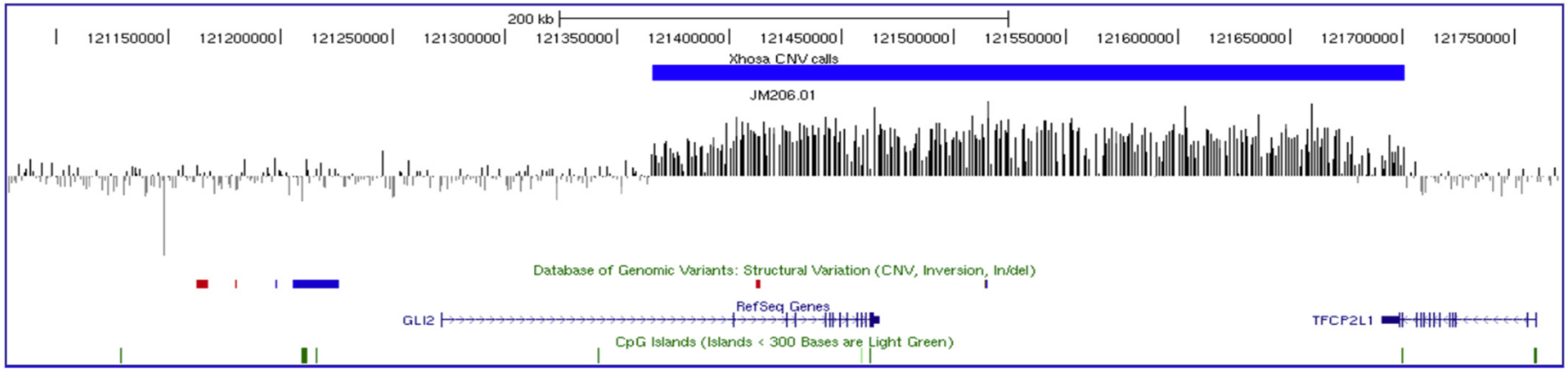


Figure 3. Duplication at *GLI2* in Xhosa patient 16. GLI2 is a C2H2 transcription factor that mediates signaling by sonic hedgehog (SHH) and activates patch. GLI2 is critical to forebrain development. Heterozygous loss-of-function mutations in *GLI2* are associated with holoprosencephaly[52]. Over-expression of *GLI2* leads to uncontrolled expression of multiple SHH target genes[53]. The duplication was detected as 335 kb by HD2 analysis and 309kb by Affymetrix6.0 analysis. By both analyses, *GLI2* exons 2-13 are duplicated. This event was not found in 8314 persons with schizophrenia and other serious mental disorders or in 7547 controls of various ancestries. (Thanks to S McCarthy and J Sebat for screening for this variant in their cohort.)

- 7 CNVs
  - 5 deletions
  - 2 duplications
  - Found by UCT and UW

Preliminary study: Exome sequencing in Xhosa persons with schizophrenia and Xhosa controls

- TruSeq paired end libraries
- Shearing germline DNA to peak of 250bp CovarisS2
- Repair, a-tailed, ligated to TruSeq adaptor
- 7 cycle amplification
- Hybridization to DNA oligos
- Purification with MyOne Streptavidin T1 dynabeads
- Amplification on dynabeads for 10 cycles

Pilot Study – Identification of Private Truncating Mutations in Schizophrenia Candidate Genes

- Tiled 20mer cRNA oligos at 3x coverage across all gene regions (excluding repetitive regions)
- Did not exclude segmentally duplicated genes 3.8Mb total
- Same process as above, but unique indices added to library during post capture amplification
- Median coverage of ~200x

Pilot Study – Identification of Private Truncating Mutations in Schizophrenia Candidate Genes

Results

- 3 novel truncation mutations identified in target regions

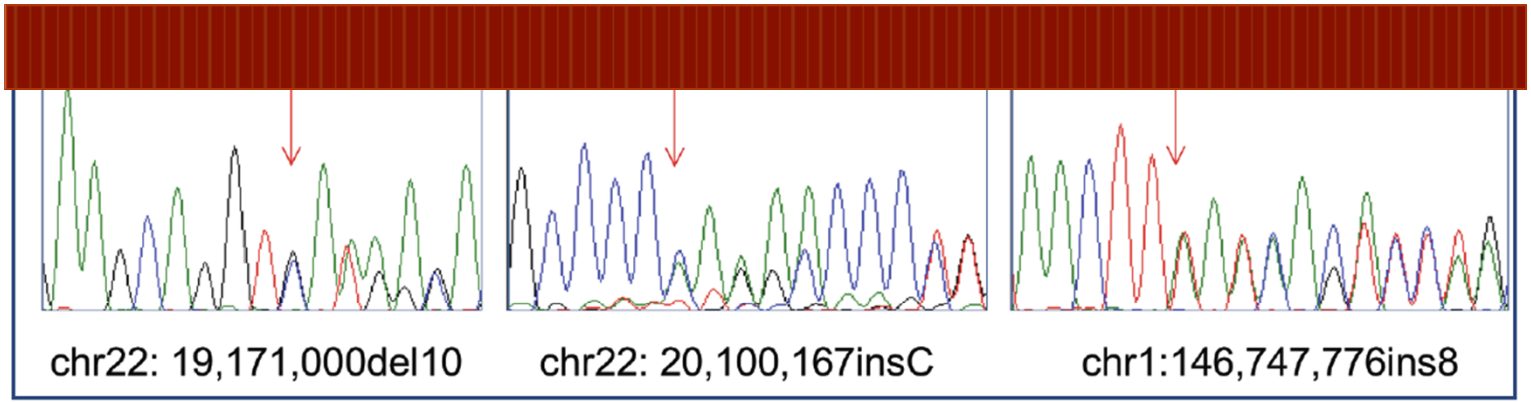


Figure 4: Validation of private truncating mutations in patients with schizophrenia. Mutations lie in genomic hotspots and were originally detected by targeted capture and massively parallel sequencing.

Achievements to date

- IRB approval at UW, Columbia and UCT
- Development and implementation of recruitment and and assessment procedures was initiated
- Optimisation of sequencing strategies (UW): Development of protocols at Human Genome Sequencing Centre (HGSC) at Baylor College of Medicine

  - planning addressed library preparation, coordination of batch sampling and data analysis
- Sample shipping procedures (UCT): Planning logistics of obtaining and shipping samples from Eastern Cape and Western Cape

  - Current model involves batch - freezing and storing material for lymphoblast cell lines locally (H3Africa Stellenbosch University Laboratory)

Achievements to date (Continued)

- Development of protocols for training at all sites.

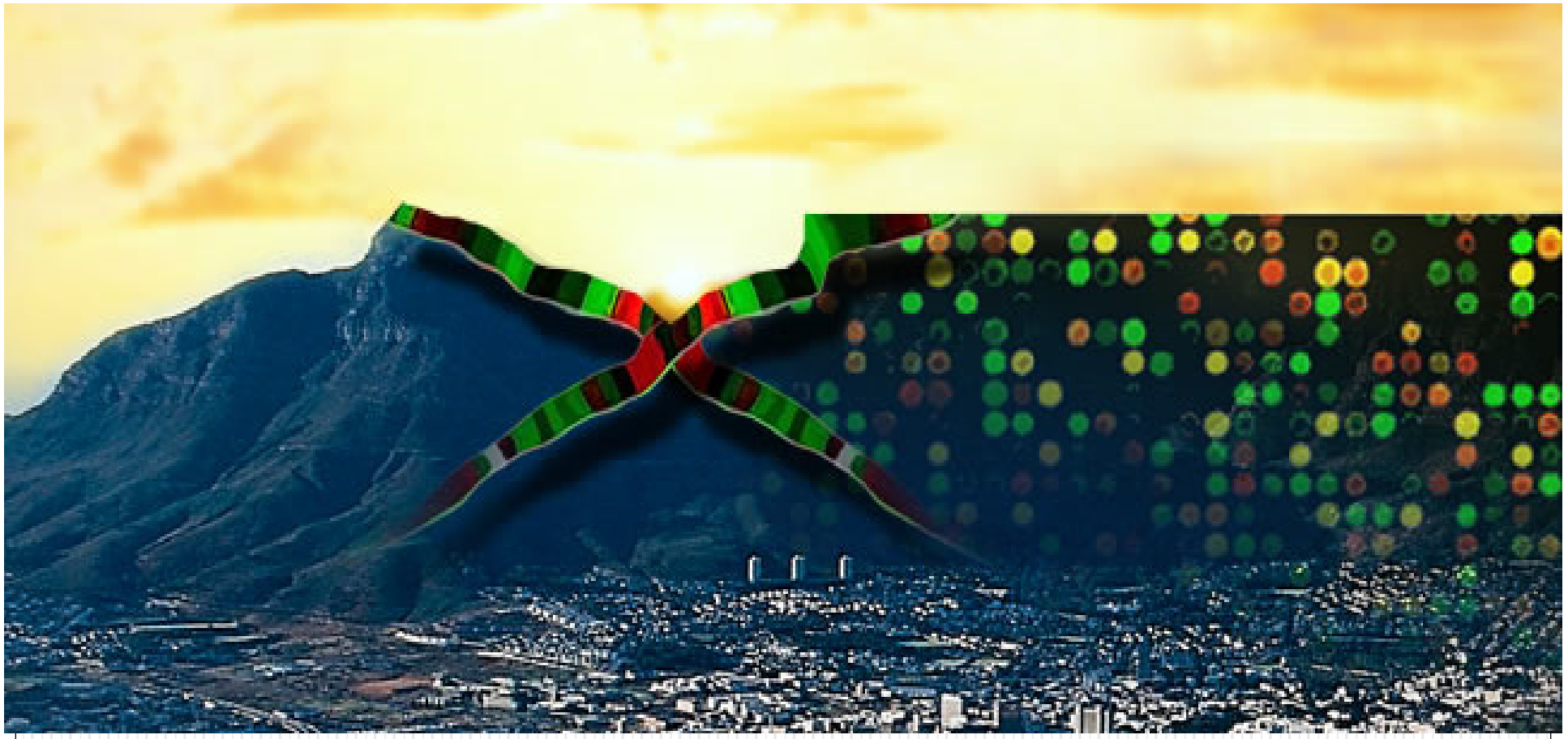
  - addressed aspects such as training goals, data sharing and ownership,

  publication guidelines and opportunities for cross-site training

- Study initiation.

  - the recruitment and assessment of individuals suitable for participation in the study has commenced (UCT)

  - in conjunction with this the preparation of DNA has therefore also commenced (UCT)



Thank you! Enkosi!