H3Africa Consortium Meeting
Genome Analysis Working group

Co-Chairs:
Zané Lombard (Wits University) & Adebowale Adeyemo (NHGRI/NIH)

November 9, 2014
Tasks of the WG

• Recommendations for standardization & harmonization of genomic data

• Assess desirability & feasibility of using standardized array/platform across H3

• Analysis of population structure & phenotypes across H3A through meta-analysis

• Design of fingerprint SNP panel
Genome Analysis WG

• Meeting by teleconference; WG meetings at H3A consortium meetings

• Regular reports to the Steering Committee

• Organized workshops in conjunction with H3A consortium meetings

• Convened custom chip task group

• Monitor advances in the field & needs of H3A projects; update plans/recommendations/guidelines
National Institutes of Health
Wellcome Trust H3Africa Research Network

Other collaborating institutions are in Belgium, Canada, France, the United Kingdom and the United States of America.
Project level data

- Working with H3ABioNet, collect minimum information needed from all projects (study-level data):
  - Study design
  - Genotyping platform
  - WES and/or WGS and/or targeted sequencing
  - WES/WGS depth
  - N samples; Ethnic groups and N samples/ethnic group
  - Phenotype(s) collected
  - Estimated date when samples will be ready for genomic analysis
Meetings and Workshops

The WG has organized two meetings/workshops/special sessions in conjunction with H3A consortium meetings

• African genomic variation, Johannesburg, SA – October 2013
• African genomic variation and disease gene discovery, Stellenbosch, SA – June 2014
DEVELOPMENTS IN THE FIELD
Release of 1000 Genomes Phase 3

Phase 1 samples

Samples in the final phase

Bubble size = sample size
Completion of AGVP

Dense genotyping data on 1,481 individuals from 18 ethno-linguistic groups

Low coverage whole genome sequencing data on 320 individuals from 7 ethno-linguistic groups

Manuscript accepted for publication in Nature

Figure 2: ‘Allele sharing between sequenced populations in the AGVP’

New genotyping arrays

- Illumina African Power Chip

Content mainly from African-Americans/Afro-Caribbeans/other African diaspora populations

~700k SNPs, meant to be used with another chip, e.g. Omni Express
New publications on African genomic diversity

• Shriner et al, Scientific Reports 2014
• Gurdasani et al, Nature 2014/2015

Genome-wide genotype and sequence-based reconstruction of the 140,000 year history of modern human ancestry

Daniel Shriner, Fasil Tekelo-Ayate, Adobowale Adeyemo & Charles N. Rotimi
Why an African custom array?

African-derived populations have greater genetic diversity and lower levels of LD, requiring a greater density of SNPs to provide genome-wide coverage of common variation.

Existing GWAS arrays still have less than optimal coverage for African populations (see Nelson et al G3(Bethesda) 2013)

Arrays optimized for African populations derive content from HapMap/1000 Genomes African populations and African-Americans/Afro-Caribbeans/other African diaspora populations.
An African custom array is needed

H3Africa projects cover many African ethnic groups, few of which have dense genotype/sequence data and for whom coverage of existing products is often suboptimal

Ethno-linguistic maps

Library of Congress 1996
Work on African custom array

• Genome analysis WG set up a custom array task group

• Custom array task group:
  – outlined the scope of the issue
  – identified and contacted people and entities to form a SNP consortium
  – identified sources of data for content
  – drafted parameters for the array
  – Identified two analysis groups
  – provided periodic updates to WG
Work on African custom array

• Identified need for more sequencing data; working with multiple partners – H3A groups, NHGRI, others
• canvassing of research groups and PIs for datasets
• activities of analysis groups: teleconferences, meeting in Cape Town November 20, 2014
Draft parameters for African custom array

• A GWAS chip that adequately captures common variation in African populations, especially those in H3A projects

• Desired features:
  – “good” coverage of common variation in H3A populations;
  – good scaffold for imputation;
  – useful specific content - e.g. reported GWAS hits, PGx variants, HLA, fingerprint SNPs,…

• Consideration for the option to add custom content

• Anticipated density: ~ 2M to 2.5M SNPs

• Timeline: ready before H3A projects samples ready for genotyping; possibility of more than one version
NHGRI supplement for WGS

- NHGRI supplement for WGS towards African custom array
- ~325 samples at 30x (@ Baylor)
- Ethnic groups selection to fill gaps in existing data
- Multi-way discussion to develop action plan
Criteria for selecting ethnic groups/samples

- Preference for ethnic group in H3Africa project on which such data does not exist
- Avoid overlap with existing data (e.g. 1000 Genomes, AGVP) and ongoing/in process (e.g. GDAP)
- Must have appropriate informed consent
- Site must be able to execute an appropriate MTA
- Must have extracted, QC’ed and quantified DNA
Update from WG meeting

• High costs of reagents/maintenance for doing genome analysis on the African continent. Possibility of negotiating discounts?

• Are we at the stage at which the WGS recommend WGS rather than WES for most indications?

• Should the WG make recommendation on WES capture protocol, especially the capture kit?

• State of each project’s budget for genotyping/sequencing?
WG meeting: African custom array

• Chip based on existing data will be much better than any existing array. Proceed with design but with option for custom bundle (Nicky hosting analysis meeting in 2 weeks in Cape Town)

• Contact manufacturers about range of prices for set quantities

• Identify H3Africa groups with African samples that might use the chip

• Engage CIDR
African custom array: what we need from you

• *What specific content (e.g. specific genes, variants) do you want to see on the array?*

• *Which ethnic groups are you studying for which there is no reference high throughput genotype/sequence data?*
Thank you