H3Africa Consortium
Phenotype Harmonization WG

Co-chairs
– Alia Benkala
– your name here
  (Jeff Struewing)

Accra, Ghana, May 16-18, 2013
H3Africa Phenotype Harmonization

Remit

1) Survey each grantee to identify phenotypes being studied at more than one site
2) Develop phenotyping standard (common) terms whenever possible, [REDCap]
3) Determine which phenotypes could be studied across the consortium (e.g., height, weight).
H3Africa Phenotype Harmonization

Remit (2)

4) Consider developing a “recommended” set of common phenotypes for use across the entire consortium

- “required” not appropriate (data already collected; not covered by existing consent; scientifically not justifiable; cost; etc.)
H3Africa Phenotype Harmonization

Remit (3)

5) Develop recommendations for which phenotypes are shared within Consortium and through European Genome-Phenome Archive (EGA)
H3Africa Phenotype Harmonization

Remit (4)

6) Consult with Ethics WG to develop consent language that allows data to be shared widely for other studies.
H3Africa Phenotype Harmonization

- Cross-consortium analyses – more statistically powerful and informative

- Opportunity to make the whole greater than the sum of its part – synergy
H3Africa Phenotype Harmonization

- A model to follow:

Phenotype harmonization and cross-study collaboration in GWAS consortia: the GENEVA experience

Siiri N. Bennett\textsuperscript{1,*}, Neil Caporaso\textsuperscript{2}, Annette L. Fitzpatrick\textsuperscript{1,3}, Arpana Agrawal\textsuperscript{4}, Kathleen Barnes\textsuperscript{5}, Heather A. Boyd\textsuperscript{6}, Marilyn C. Cornelis\textsuperscript{7}, Nadia N. Hansel\textsuperscript{5}, Gerardo Heiss\textsuperscript{8}, John A. Heit\textsuperscript{9}, Jae Hee Kang\textsuperscript{10}, Steven J. Kittner\textsuperscript{11}, Peter Kraft\textsuperscript{12}, William Lowe\textsuperscript{13}, Mary L. Marazita\textsuperscript{14}, Kristine R. Monroe\textsuperscript{15}, Louis R. Pasquale\textsuperscript{10}, Erin M. Ramos\textsuperscript{16}, Rob M. van Dam\textsuperscript{17}, Jenna Udren\textsuperscript{1}, and Kayleen Williams\textsuperscript{1} for the GENEVA Consortium

*Correspondence to: siiribennett@gmail.com

**TABLE II. Examples of possible (a) smoking-related questions and (b) new variables for cross-study analyses**

<table>
<thead>
<tr>
<th>Study (N)</th>
<th>Smoking-related questions</th>
<th>Possible responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study 1 (2,500)</td>
<td>1. Do you currently smoke cigarettes?</td>
<td>Y/N</td>
</tr>
<tr>
<td></td>
<td>2. If yes, how many cigarettes per day?</td>
<td>###</td>
</tr>
<tr>
<td>Study 2 (1,200)</td>
<td>1. Have you smoked more than 100 cigarettes in your lifetime?</td>
<td>Y/N</td>
</tr>
<tr>
<td></td>
<td>2. If yes, do you currently smoke?</td>
<td>Y/N</td>
</tr>
<tr>
<td></td>
<td>3. If yes, how many packs per day do you smoke?</td>
<td>###</td>
</tr>
<tr>
<td>Study 3 (8,500)</td>
<td>1. Have you ever smoked?</td>
<td>Y/N</td>
</tr>
<tr>
<td>Study 4 (1,250)</td>
<td>1. Do you currently smoke?</td>
<td>Y/N</td>
</tr>
<tr>
<td>Study 5 (4,200)</td>
<td>1. Do you smoke?</td>
<td>Y/N</td>
</tr>
<tr>
<td></td>
<td>2. When did you first start smoking regularly?</td>
<td>Past year; 1–5 years ago; &gt; 5 years ago</td>
</tr>
<tr>
<td>Study 6 (6,600)</td>
<td>1. Have you smoked tobacco in the past month?</td>
<td>Y/N</td>
</tr>
<tr>
<td>Study 7 (800)</td>
<td>1. Have you ever smoked regularly?</td>
<td>Y/N</td>
</tr>
<tr>
<td></td>
<td>2. If yes, do you still smoke?</td>
<td>Y/N</td>
</tr>
<tr>
<td></td>
<td>3. If yes, how much do you smoke a day?</td>
<td>1–10 cigarettes, 11–20 cigarettes, 21–30 cigarettes, &gt; 30 cigarettes</td>
</tr>
</tbody>
</table>
What phenotypes might be studied in 2 or more studies?

- 2 questionnaires received (recently)
  - Adu & Ojo, Rasmsay & Sankoh
  - Both are large, complex questionnaires, likely to be significant “natural” overlap [both use REDCap]

- Being analyzed by Alia & Seydou (H3ABionet)
  - Populate a table that will grow over time
Examples

- Pilot questionnaires - opportunity for modification
- Baseline & Follow-up questionnaire – add module
- Questionnaires in development – opportunity to modify questions, add new ones
- Subjects from one study can serve as controls for another
- “Replication” cohorts
- Cost/benefit trade-off
H3Africa Phenotype Harmonization

- Questions, Suggestions?
H3Africa Phenotype Harmonization

Working Group Membership

- Eyitayo Fakunle, Catherine Rossouw (Abayomi)
- Tay Croxton, Nicaise Ndemi, Alash'le Abimiku (Abimiku)
- Fatiu Arogundade, David Burke, Charlotte Osafo (Adu)
- Branwen Hennig, Corinne Merle, Aurel C. Allabi (Afollabi)
- Naomi Levitt, David Adeyemi, Ayesha Motala, Branwen Hennig (Amoah)
- Bruno Bucheton, Christianne Hertzfowler (Matovu)
- Mark Engel (Mayosi)
- James Brandful, Seydou Doumbia, Oyekanmi Nash, Yasmina Jaufeerally Fakim, Ezekiel Adebiyi, Dean Everett, Alia Benkahla, Alan Christoffels (Mulder)
- Alisha Wade, Nigel Crowther (Ramsay)
- Deborah Colantuoni (NHGRI)
<table>
<thead>
<tr>
<th>Project PI</th>
<th>Main Phenotype</th>
<th>Approx Sample Size*</th>
<th>Questionnaire Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adu &amp; Ojo</td>
<td>kidney disease</td>
<td>8,000</td>
<td>Finalized!</td>
</tr>
<tr>
<td>Affolabi</td>
<td>TB pharmacogenomics</td>
<td>630</td>
<td>nested</td>
</tr>
<tr>
<td>Amoah</td>
<td>Type II diabetes</td>
<td>24,000</td>
<td>Finalizing</td>
</tr>
<tr>
<td>Matovu</td>
<td>trypanosomiasis</td>
<td>2,400</td>
<td></td>
</tr>
<tr>
<td>Mayosi</td>
<td>rheumatic heart disease</td>
<td>6,000</td>
<td></td>
</tr>
<tr>
<td>Ramsay &amp; Sankoh</td>
<td>cardiometabolic disease</td>
<td>12,000</td>
<td>Near final</td>
</tr>
<tr>
<td>Stein &amp; Ramesar</td>
<td>schizophrenia</td>
<td>2,000</td>
<td></td>
</tr>
</tbody>
</table>

* Total study size - not all subjects genotyped in each study.

- **What phenotypes might be studied in 2 or more studies?**
  - 2 questionnaires received (recently) – being analyzed by Alia & Seydou