KENYA - GENOMIC AND ENVIRONMENTAL RISK FACTORS FOR CARDIOMETABOLIC DISEASE IN AFRICANS (KENYA) 2014, African Wits-INDEPTH Partnership for genetic studies on the role of the genome and environment in body composition, as a risk factor for cardiometabolic disease

Catherine Kyobutungi - African Population and Health Research Center (APHRC)
Prof Michèle Ramsay - University of the Witwatersrand – South Africa

Report generated on: August 5, 2020

Overview

Identification

ID NUMBER
DDI-KEN-APHRC-AWIGEN-2014-v01

Version

VERSION DESCRIPTION

PRODUCTION DATE
2020-07-06

Overview

ABSTRACT
The Genomic and environmental risk factor for cardiometabolic disease in Africans (AWI-Gen) project is a collaborative study between the University of the Witwatersrand (Wits) and the INDEPTH Network funded under the Human Heredity and Health in Africa (H3Africa) initiative. The H3Africa is a ground-breaking initiative to build institutional and individual capacity to undertake genetic and genomic studies in the African region. This collaboration, involves five INDEPTH sites i.e. 1) Navrongo - Ghana; 2) Nanoro - Burkina Faso; 3&4) Agincourt and Digkale - South Africa; and 5) Nairobi - Kenya) plus the Soweto-based birth-to-twenty cohort. The work in Kenya will be undertaken in the Nairobi Urban Health and Demographic Surveillance Site (NUHDSS) run by APHRC. The AWI-Gen project aims to understand the interplay between genetic, epigenetic and environmental risk factors for obesity and related cardiometabolic diseases (CMD) in sub-Saharan Africa. The project capitalizes on the unique strengths of existing longitudinal cohorts and well-established health and demographic surveillance systems (HDSS) run by the partner institutions. The six study sites represent geographic and social variability of African populations which are also at different stages of the demographic and epidemiological transitions. The study has two components: i) to understand the African population genetic structure and ii) to determine the association between genetic factors and obesity in determining cardiometabolic risk and the effect of environmental factors on this relationship.

In this application, we seek ethical approval for the Kenya study only. The other partners will seek approval from the appropriate ethics review authorities in their countries.

UNITS OF ANALYSIS
A survey with community-level questionnaire with the following units of analysis: individuals, households, and communities.

Scope

NOTES
The scope of the Survey includes: General information, Demographic information, Family composition, Pregnancy, Marital status, Education, Employment, Household attributes, Tobacco use, Alcohol use, General health, Infection history(Malaria, Tb), Cardio metabolic risk factors (Diabetes, Stroke, Hypertension, Angina, Heart attack, Congestive heart failure, High cholesterol), Thyroid disease, Kidney disease, Physical activity (Occupation-related physical activity (paid or unpaid work, Travel-related physical activity, Non-work related and leisure time physical activity, Sitting/resting activity), Sleep

Coverage

GEOGRAPHIC COVERAGE
The study was conducted in Viwandani and Korogocho informal settlements in Nairobi.

UNIVERSE
Component a): Adults (18+years) residents of Korogocho and Viwandani informal settlements in family trios (father, mother

and adult offspring). In selecting participants, emphasis will be placed on ethnic groupings that have not been included in previous and ongoing global genome studies (Kang et al., 2010, Patterson et al., 2006). To ensure that the project contributes new knowledge on the genetic diversity of African populations we will select participants whose self-reported ethnicity is either Luo or Kiisi or Luhya since the genomic structure of Kikuyu, Masaai and Kalenjin ethnicities has already been documented as part of completed and ongoing studies.

Component b): Adults (40-60 years) residents of Korogocho and Viwandani informal settlements registered in the NUHDSS.

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
Catherine Kyobutungi	African Population and Health Research Center (APHRC)
Prof Michèle Ramsay	University of the Witwatersrand – South Africa

OTHER PRODUCER(S)

Name	Affiliation	Role
Shukri Mohamed	African Population and Health Research Center (APHRC)	Kenya Site Project Coordinator
Christopher Khayeka Wandabwa	African Population and Health Research Center (APHRC)	Kenya site Research Officer
Tambo Ernest	University of the Witwatersrand – South Africa	Overall Project Manager

FUNDING

Name	Abbreviation	Role
National Institutes of Health (NIH) - (USA) - under the Human Heredity and Health in Africa (H3A) Initiative		

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
AFRICAN POPULATION AND HEALTH RESEARCH CENTER	APHRC		Documentation of the DDI

DATE OF METADATA PRODUCTION
2020-07-06

DDI DOCUMENT VERSION
Version 1.0 (July 2020)

DDI DOCUMENT ID
DDI-KEN-APHRC-AWIGEN-2014-v01

Sampling

Sampling Procedure

Study participants were drawn from the most recently updated NUHDSS database managed by the African Population and Health Research Center (APHRC). The NUHDSS follows about 74,000 individuals living in approximately 24,000 households in Korogocho and Viwandani.

Objective 1
The strategy for sampling for the population genome structure study is to collect 30 trios (mother, father and one child - all over 18 years of age) from each of the sites, and an additional 10 unrelated individuals to give a total sample of 100. Using the NUHDSS database, a sampling frame of participants who are related i.e. one adult child, father and mother based on information on household structure was drawn. Potential participant trios were then randomly selected and approached to participate. To minimise cases where biological paternity was in doubt, there was selection of the second or third adult child as the index participant in each trio and then both parents. First it was confirmed with the mother whether the other stated parent is the biological father. If either of the stated parents are not the biological parents, the trio was dropped and replaced. Given the possibly high rate of non-participation for various reasons (refusal to participate, non-biological relationship with either parent in the trios) there was randomly selection of as many as 60 trios aiming for a final sample size of 30. In addition to the 30 participating trios and 10 unrelated individuals, there was oversampling of about 10 more trios to compensate for any cases where the child turns out to be unrelated to the father/mother on genetic testing. Phenotype data was not collected on these participants.

Inclusion criteria
· Households with (trios) families consisting of mother, father and adult biological child in the NUHDSS area.
· Registered and resident in the most up-to-date NUHDSS database.
· All members should be over 18 years of age at sampling.
· Participants whose self-reported ethnicity is either Luo or Kiisi or Luhya

Exclusion criteria
· All visiting non-resident families in the NUHDSS frame.
· Household trios that do not have biological relations.
· Incapacitated adults who are unable to provide informed consent.
· Anyone under the age of 18.

Objectives 2-4
For the genetic association study; stratified sampling (based on sex) was applied to adults aged 40-60 years to ensure we had equal numbers of men and women. The sampling frame was for the current residents in the NUHDSS aged 40-60 years on the day of sampling. The aim was to collect phenotype and genotype data from 2000 participants in each of the study sites - hence we selected 1000 men and 1000 women in the site.

Inclusion criteria
· Adults (40-60 years) in Korogocho and Viwandani slums.
· Registered and resident in the most up-to-date NUHDSS database.

Exclusion criteria
· All visiting non-resident men and women in the NUHDSS frame.
· Incapacitated adults who are unable to provide informed consent.

Sample size
A sample size of 2000 per site (12000 in total) was based on power calculations and effect sizes. The power calculations show that we have power to detect realistic effect sizes, based on studies in other populations. Figure 2 illustrates the relationship between power and effect size for two different phenotypes, illustrating that the detectable effect size is realistic. Power analysis for a sample size of 12000 individuals based on proposed candidate gene study for BMI (shown on the left) and for DXA (total body fat) (shown on the right). Given a sample size of 12000 in the AWI-Gen study, this graph shows effect size (x) which could be detected at a given power (y) for different minor allele frequencies (ranging from 0.05-045). For example, with a minor allele frequency of 0.25, we will have 80% power to detect an effect size (Beta) of 0.20 per allele change in BMI, and an effect size of 0.25 per allele change in body fat percentage.
The Kenya study will thus contribute 2000 individuals (1000 males and 1000 females). In order to have sufficient power for

the genetic association studies, data from the Kenyan sample will be pooled with data from the other five sites.

Response Rate

(88.7%) accepted to participate in the study

Questionnaires

Overview

AWI_Gen Trait_Ass_Questionnaire: General information, Demographic information, Family composition, Pregnancy, Marital status, Education, Employment, Household attributes, Tobacco use, Alcohol use, General health, Infection history(Malaria, Tb), Cardio metabolic risk factors (Diabetes, Stroke, Hypertension, Angina, Heart attack, Congestive heart failure, High cholesterol), Thyroid disease, Kidney disease, Physical activity (Occupation-related physical activity (paid or unpaid work, Travel-related physical activity, Non-work related and leisure time physical activity, Sitting/resting activity), Sleep

Data Collection

Data Collection Dates

Start	End	Cycle
2014-10-01	2015-11-30	Phase 1

Data Collection Mode

Face-to-face [f2f]

Questionnaires

AWI_Gen Trait_Ass_Questionnaire: General information, Demographic information, Family composition, Pregnancy, Marital status, Education, Employment, Household attributes, Tobacco use, Alcohol use, General health, Infection history(Malaria, Tb), Cardio metabolic risk factors (Diabetes, Stroke, Hypertension, Angina, Heart attack, Congestive heart failure, High cholesterol), Thyroid disease, Kidney disease, Physical activity (Occupation-related physical activity (paid or unpaid work, Travel-related physical activity, Non-work related and leisure time physical activity, Sitting/resting activity), Sleep

Data Processing

Data Editing

a) Data management, storage and sharing
Individuals were assigned barcodes that was placed on all the samples collected, questionnaires, ultrasound images and sample aliquots. Data was captured on paper and entered in a secure online data capture system (RedCap®). Data transmitted to Wits for pooled analyses did not have any personal identifiers other than the barcodes.

The following key principles guided the data management, storage and sharing.
The long term storage of the DNA samples from this project was done in two locations, one in the country of origin (in this case Kenya) and one in the laboratory that was co-ordinating the genetic and epigenetic analyses. Since data collected from the Kenya site was analysed with those from five other sites, there was need for harmonization and standardization in the collection and analysis of samples for phenotyping and DNA extraction. The same applies for genomic and genetic analyses. For these reasons, the samples from this study and all those from the five other sites was shipped to (a yet-to-be determined) central place for DNA extraction and analysis. The choice of the central analysis laboratory depended on capacity to conduct the required analyses, quality standards and certifications, and cost. All sites retained samples in storage that can be used for future analyses and genetic studies.

Second, as part of the data sharing and access policy of the whole H3Africa consortium, all participating sites (not just those in the AWI-Gen project) were required to avail their data and a DNA sample in a publicly accessible repository after an appropriate period of time during which the data generators will exploit the data. These repositories included the H3A Bio-repository, funded through the same H3Africa initiative and the European Genome Phenome Archive (EGA). A draft H3Africa Data Access and Release Policy is provided in Appendix 5. Underpinning the H3Africa data access and sharing policy is the need to maximise the returns on investments in this pioneering initiative and to assure the availability of samples and data for future genetic studies, discoveries and innovations that are likely to benefit Africans as this field of research grows. It is therefore a distributed model which ensures that DNA is safeguarded against accident or disaster. This is important for capacity development in the different African countries and to give the Centres autonomy in the future use of the DNA samples. The access to the samples will be governed by a set of policies developed to ensure controlled, but wide access in an ethical manner with minimal risk to the research participants in terms of confidentiality and potential stigmatisation.

Data Appraisal

No content available