Center for Research on
Genomics and Global Health

Missing heritability

Contributions from genomic studies in African ancestry populations

Charles N. Rotimi, PhD

Director: Center for Research on Genomics and Global Health

Chief: Metabolic, Cardiovascular and Inflammatory Disease Genomics Branch

May 1&2, 2018

1. Questions we are not asking

2. Disease influencing variables we are not measuring

3. Are these too difficult to measure?
    1. Neighborhoods
    2. Schools attended (strong correlation with educational attainment, income, etc)
    3. Tax policies
    4. Living under perpetual state of stress (epigenetics?)

The Seductive Allure of Behavioral Epigenetics

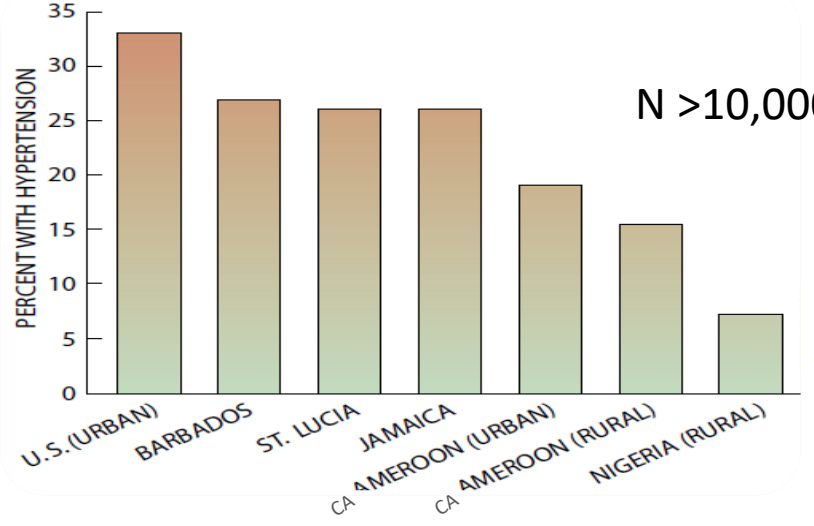
Different upbringings: Being raised by a nurturing (top left) or a lackadaisical (top right) mother can cause epigenetic differences that affect a rat pup's behavior later in life.

Whether similar differences occur in people raised in wealthy (bottom left) or impoverished (bottom right) neighborhoods remains an open question.

www.sciencemag.org SCIENCE VOL 329 2 JULY 2010

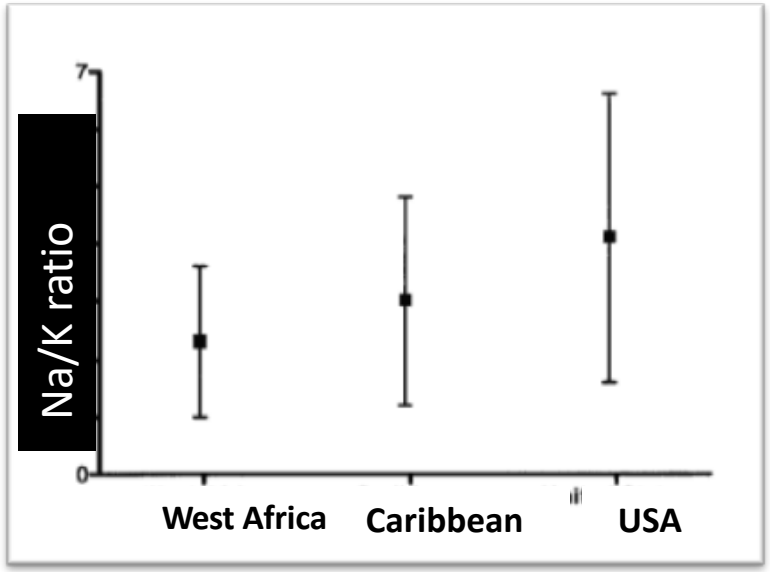
The Puzzle of Hypertension in African-Americans

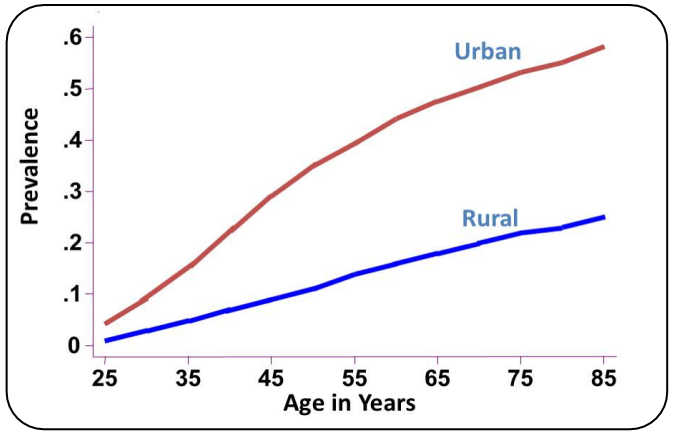
Prevalence of Hypertension in Seven Populations of West African Origin *Am J Public Health.* Feb1997



N >10,000

Urinary Na/K Ratio in Persons of West African Origin





Hypertension in Urban and Rural Nigerians

Genes are often invoked to account for why high blood pressure is so common among African-Americans. Yet rates are lower in rural Africans. This discrepancy demonstrates how genes and the environment interact

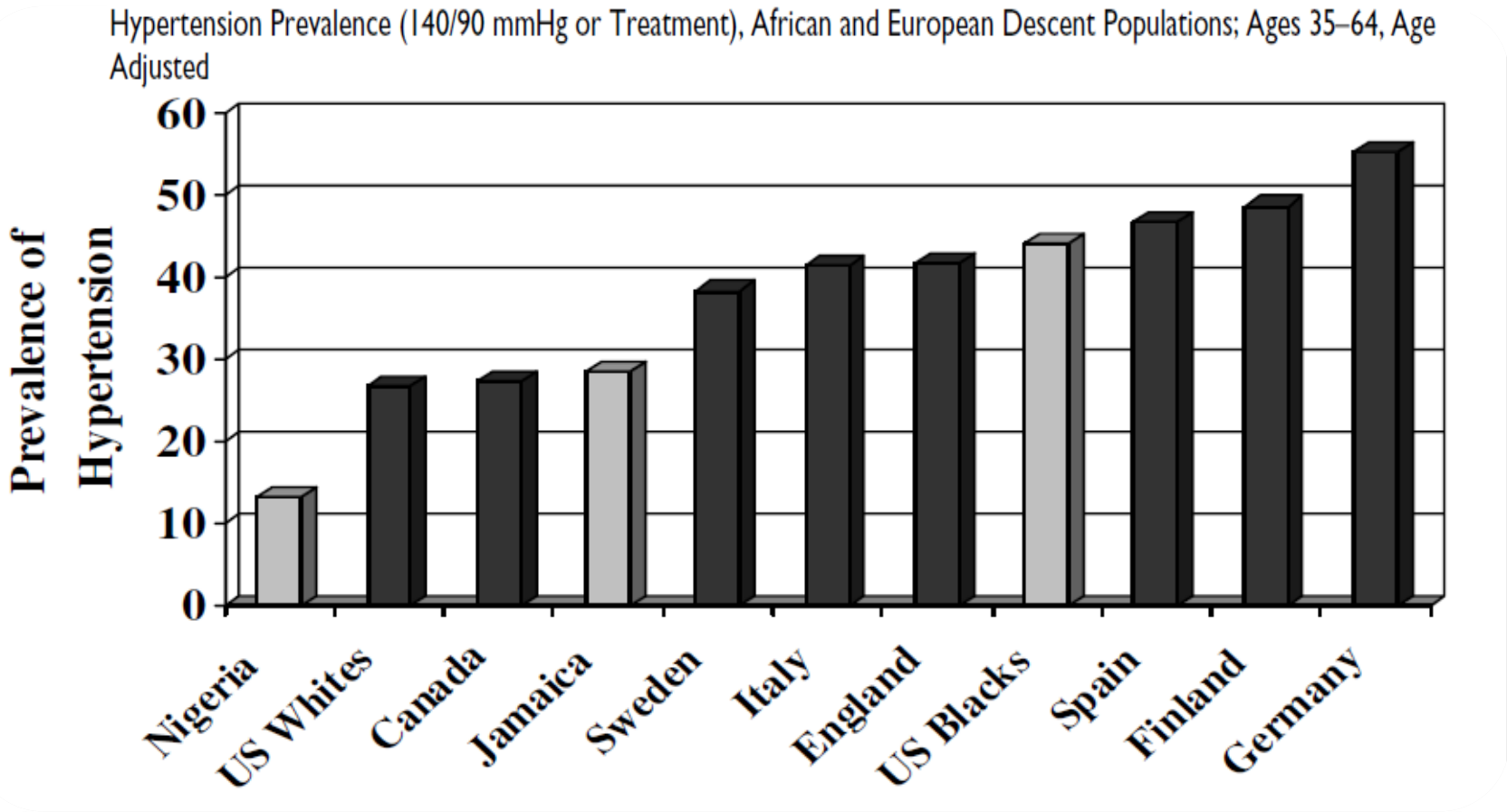
Cooper RS, Rotimi CN and Ward R. **The puzzle of hypertension in African-Americans**.

SCIENTIFIC AMERICAN 1999

Research article

An international comparative study of blood pressure in populations of European vs. African descent

Richard S Cooper*[1], Katharina Wolf-Maier[1], Amy Luke[1], *BMC Medicine* 2005,



Hypertension prevalence ranged

a. From 27 to 55% for European ancestry

b. From 14 to 44% for African ancestry populations

Hypertension
JOURNAL OF THE AMERICAN HEART ASSOCIATION

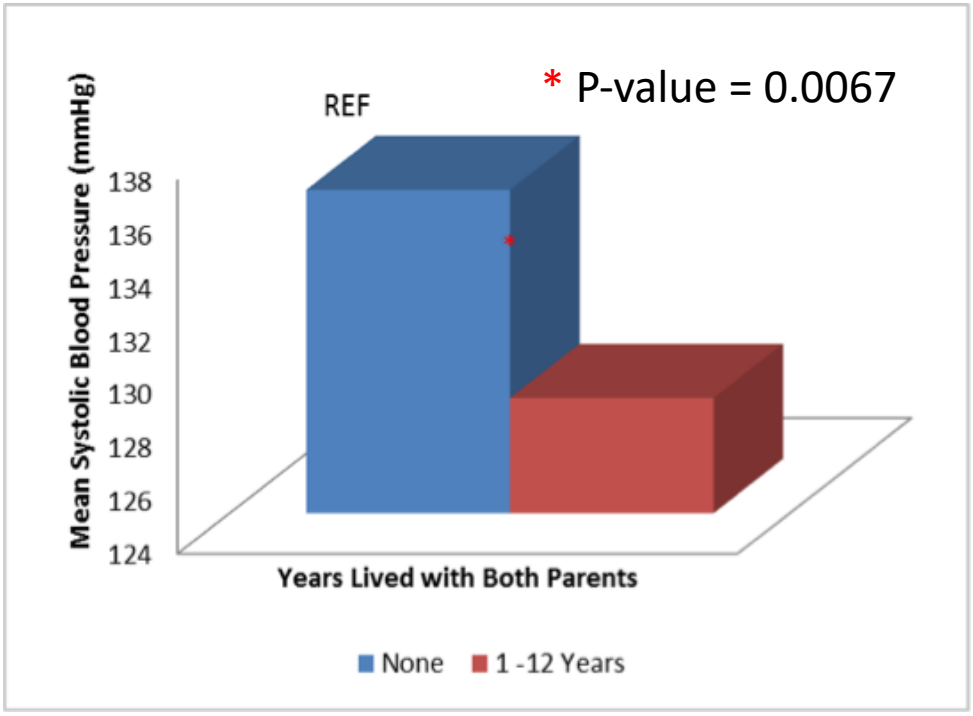
American Heart Association®

Childhood Family Living Arrangements and Blood Pressure in Black Men: The Howard University Family Study

Debbie S. Barrington, Adebowale A. Adeyemo and Charles N. Rotimi

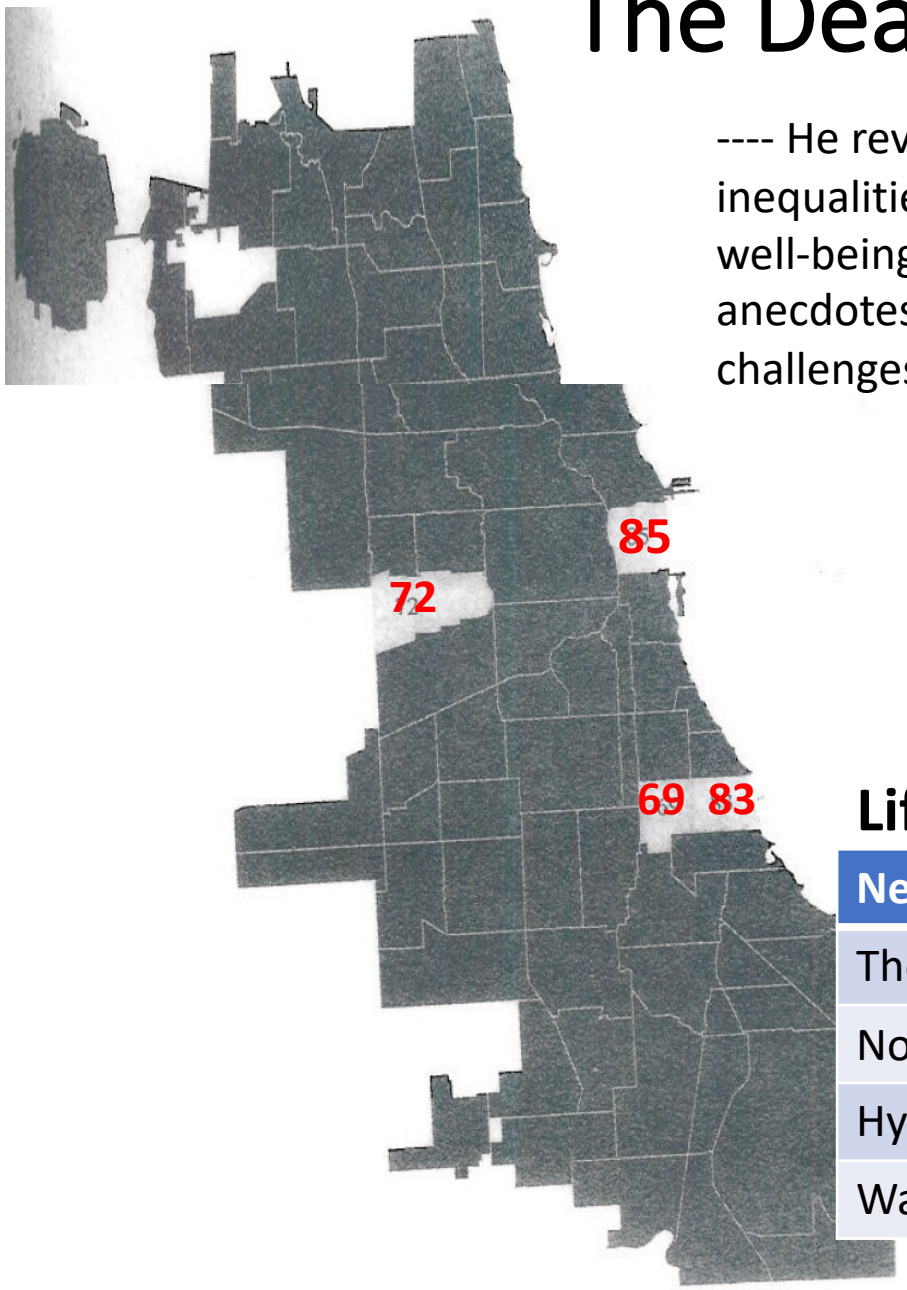
Men who lived with both parents between 1-12 years:

- 6.5 mmHg lower SBP
- 43% lower odds of HTN compared to never lived with both parents



Hypertension. 2014;63:48-53

The Death Gap – David A Ansell, MD



---- He reveals the profound inequalities, particularly racial inequalities, that generate tremendous differences in lifespan and well-being across neighborhoods, and he provides powerful patient anecdotes that provide a human face to otherwise abstract challenges." Harold Pollack, University of Chicago

Location, Location, Location – real estate
These are also the three most important words in understanding health and wealth inequality in America

Life Expectancy varies greatly by neighborhoods in Chicago

Neighborhoods	Income	Life expectancy	comments
The Loop	$80,000	85yrs	Affluent downtown
North Lawndale	$25,000	72yrs	91% AA (5miles from the Loop)
Hyde Park	$40,000	83yrs	Racially diverse
Washington Park	$22,000	69yrs	Almost entirely AA

AA – African Americans

http://www.press.uchicago.edu/ucp/books/book/chicago/D/bo25081418.html

In the US, states with greater inequality in the distribution of income

1. have higher rates of unemployment, incarceration, and higher proportion receiving income assistance

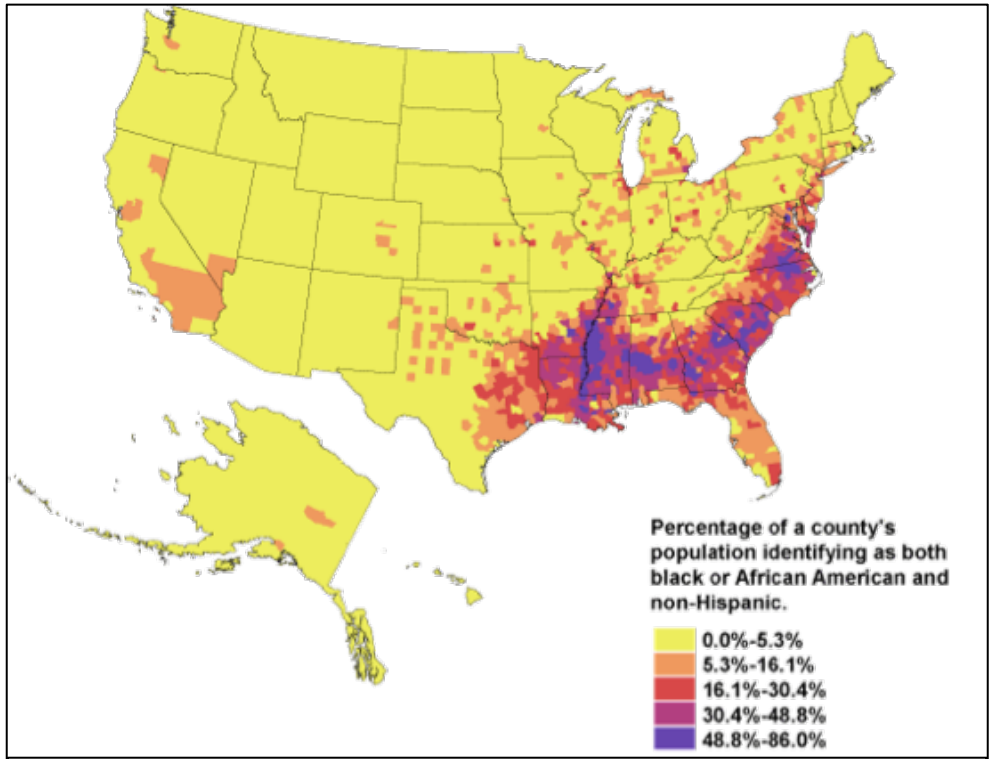
2. Greater % of people without medical insurance.

3. Spend less per person on education.

4. Greater proportion of babies born with low birth weight; higher rates of homicide, violent crime; persons unable to work because of disabilities

5. Have higher costs per-person for medical care

This population is concentrated largely in the southern states and urban areas.



http://www.censusscope.org/us/print_map_nhblack.html

http://www.huppi.com/kangaroo/Inequality&Health.htm

Phenotypic variance explained by local ancestry in admixed African Americans

*Daniel Shriner, Amy R. Bentley, Ayo P. Doumatey, Guanjie Chen, Jie Zhou, Adebowale Adeyemo and Charles N. Rotimi **

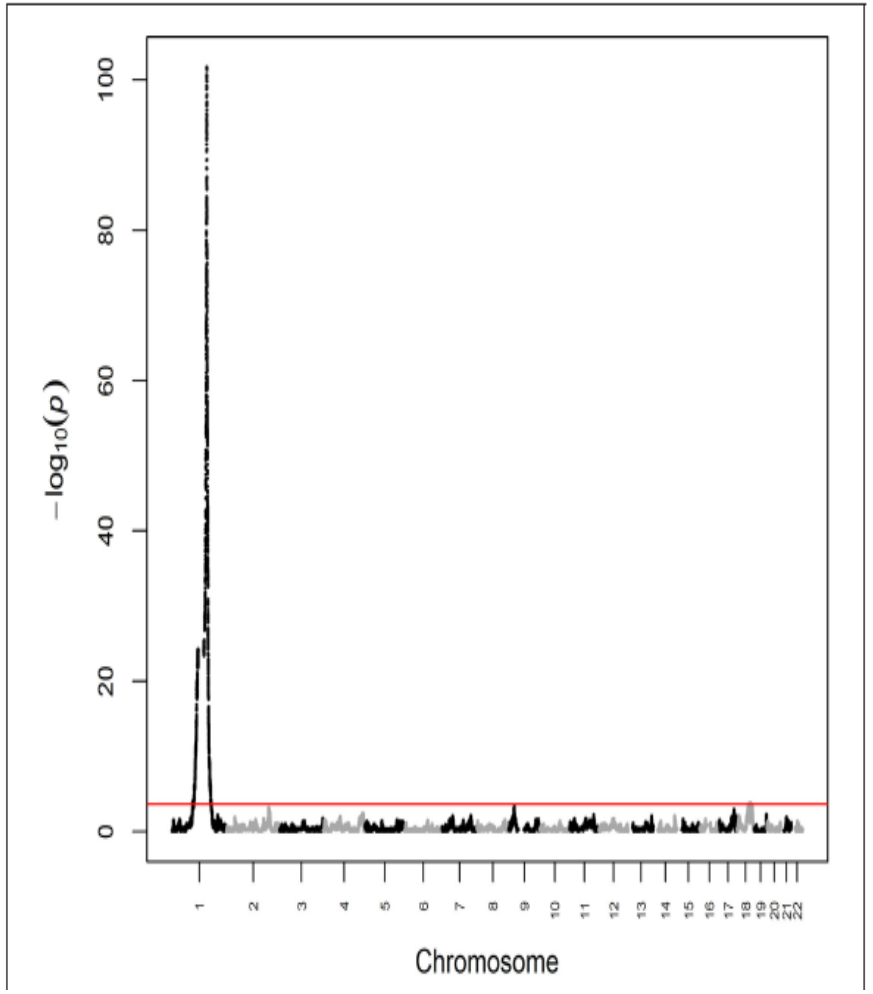
1. Surveyed 26 quantitative traits to estimated proportion of variance explained by local ancestry in African Americans.

2. Linear mixed models; genetic similarity defined in terms of local ancestry rather than genotype in ARIC and HUFS

3. **Local ancestry at major (WBC) and polygenic effect genes can explain up to 20 and 8% of phenotypic variance, respectively.**

4. WBC is remarkably different from other phenotypes in that the *DARC* promoter-null polymorphism is essentially fixed between West Africans and Europeans.

5. Put another way, a GWAS for WBC in Europeans would not identify *DARC*, nor would a GWAS in Africans. **The effect size is more in the realm of Mendelian genetics than complex disease genetics.**



Manhattan plot from admixture mapping for white blood cell count in ARIC. WBC regressed on local ancestry, adjusted for age, global ancestry, sex, and center.

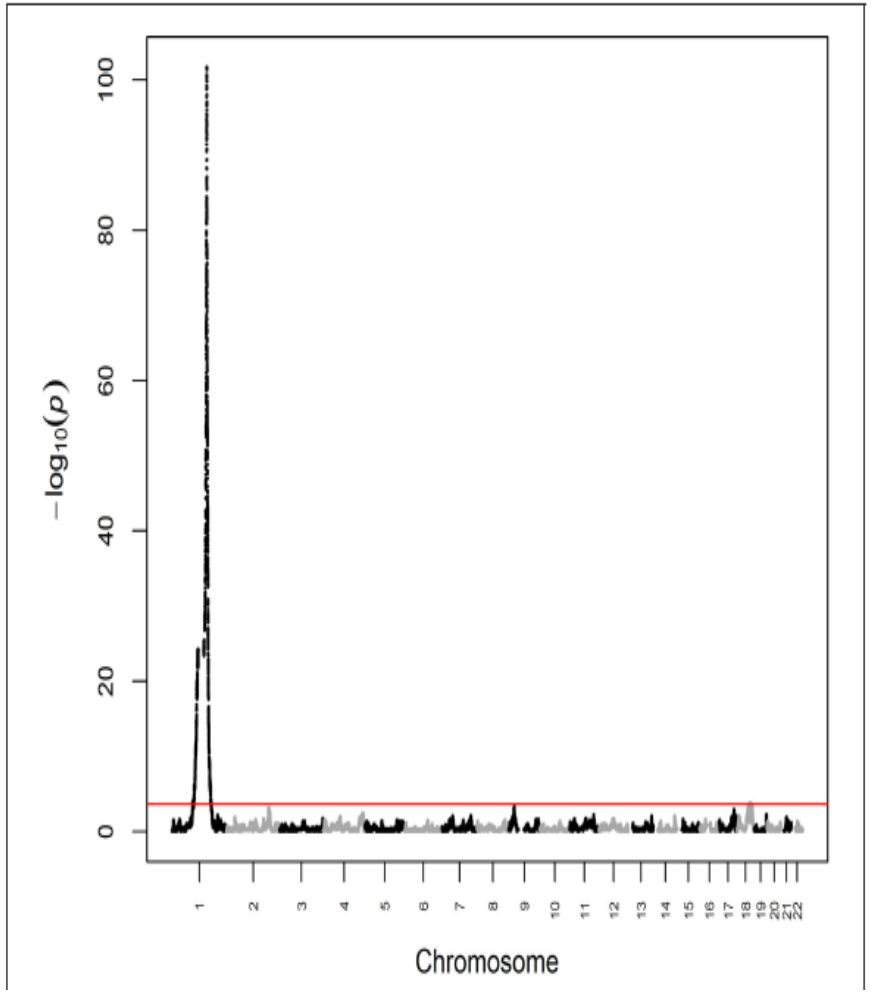
Front Genet. 2015 Oct 29;6:324.

Phenotypic variance explained by local ancestry in admixed African Americans

*Daniel Shriner, Amy R. Bentley, Ayo P. Doumatey, Guanjie Chen, Jie Zhou, Adebowale Adeyemo and Charles N. Rotimi**

Findings provide evidence that

a. most additive genetic variance is explained by genetic markers undifferentiated by ancestry
b. results inform proportion of health disparities due to genetic risk factors
c. and the magnitude of error in association studies not controlling for local ancestry.
d. Adjusting for global ancestry, *e.g.*, using PCA, does not control for the effects of local ancestry.
e. Local ancestry effects are not limited to African Americans or Latinos.
f. However, since genetic differentiation between West Africans and Europeans is at the high end of the scale (in terms of Fst), the amount of phenotypic variance explained by local ancestry in African Americans is at the high end



Manhattan plot from admixture mapping for white blood cell count in ARIC. WBC regressed on local ancestry, adjusted for age, global ancestry, sex, and center.

Front Genet. 2015 Oct 29;6:324.

Discovery and refinement of genome-wide associated loci for serum uric acid in Africans

1. Heritability is 35.4% in (HUFS - African Americans), 39.9% in (NHLBI Family Heart Study - European families), 34.0% in Sardinians, and 36% in Icelanders. For some unknown reason, the Framingham Heart Study yielded 63% (data from the 1970s). **Question – distortion of the range of heritability**

2. Performed first GWAS for uric acid in Africans (*n* = 4,126); replication in African Americans (*n* = 5,007)

3. Identified two signals in *SLC2A9* (evidence for allelic heterogeneity) and one signal in *SLC22A12*

4. Single candidate causal variant at *SLC22A12* - low frequency in AFR but monomorphic in EUR (must be other causal variants in Europeans, therefore also evidence for allelic heterogeneity)

5. The three signals account for 4.3% of variance of serum uric acid, or 12% of heritability

6. Compared to **7.0% in** GWAS of European-ancestry individuals (N > 140,000) that identified 26 loci (Kottgen A et al. *Nat Genetics 2013)*

Under review 2018

CHARGE Gene-Lifestyle Interactions Working Group: Gene × Smoking Status on Serum Lipids

Samples Model
Stage 1 + 2

Interaction Model (**1df**, Joint 2df): $Y = \beta_0 + \beta_C C + \beta_E E + \beta_G SNP + \beta_{GE} E \times SNP$

EUR	AFR	ASN	HISP	**TRANS[1]**
n=237,483	n=30,970	n=94,688	n=20,490	**N=387,283**

13 novel loci $p < 5 \times 10^{-8}$

Main Effect Model: $Y = \beta_0 + \beta_C C + \beta_G SNP$

Smoking-Adjusted Model: $Y = \beta_0 + \beta_C C + \beta_E E + \beta_G SNP$

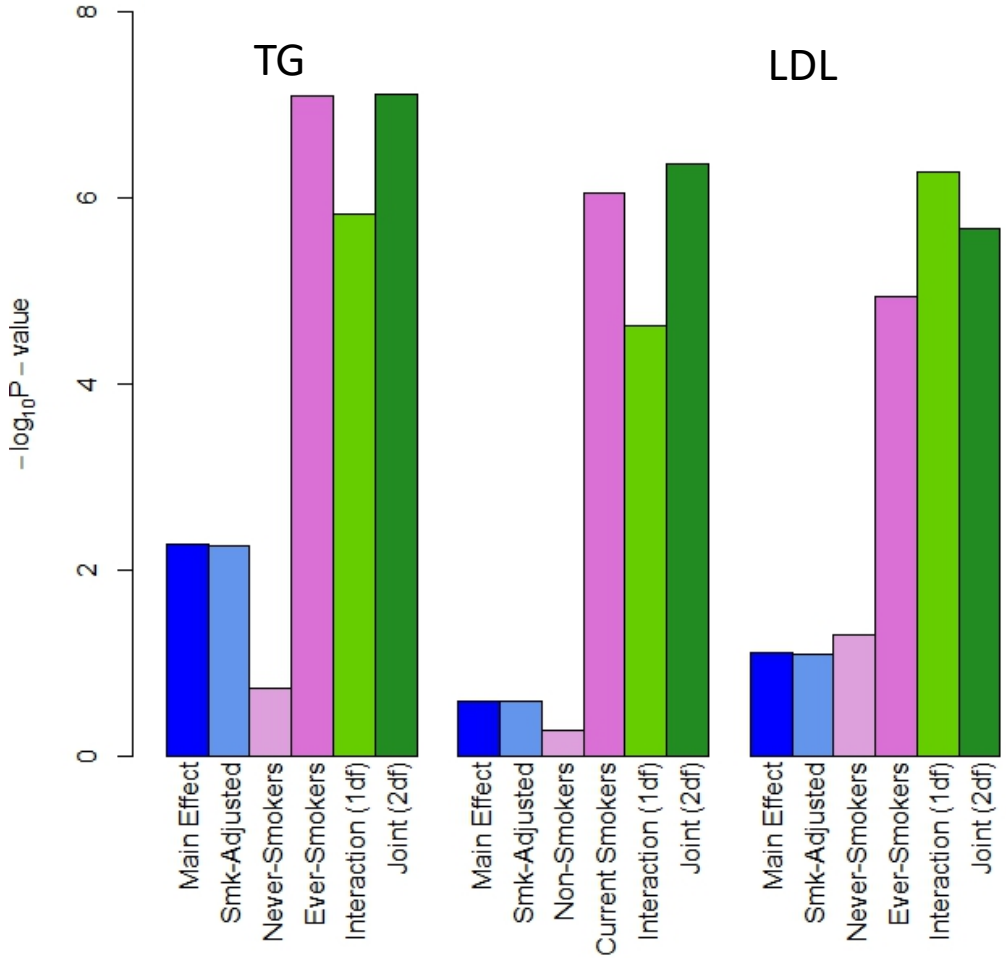
Main Effect Model (STRATIFIED E=0): $Y = \beta_0 + \beta_C C + \beta_G SNP$

Main Effect Model (STRATIFIED E=1): $Y = \beta_0 + \beta_C C + \beta_G SNP$

Interaction Model (1df): $Y = \beta_0 + \beta_C C + \beta_E E + \beta_G SNP + \beta_{GE} E \times SNP$

Interaction Model (Joint 2df): $Y = \beta_0 + \beta_C C + \beta_E E + \beta_G SNP + \beta_{GE} E \times SNP$

Identified loci which would not have been identified in a standard GWAS model or with adjustment for smoking.



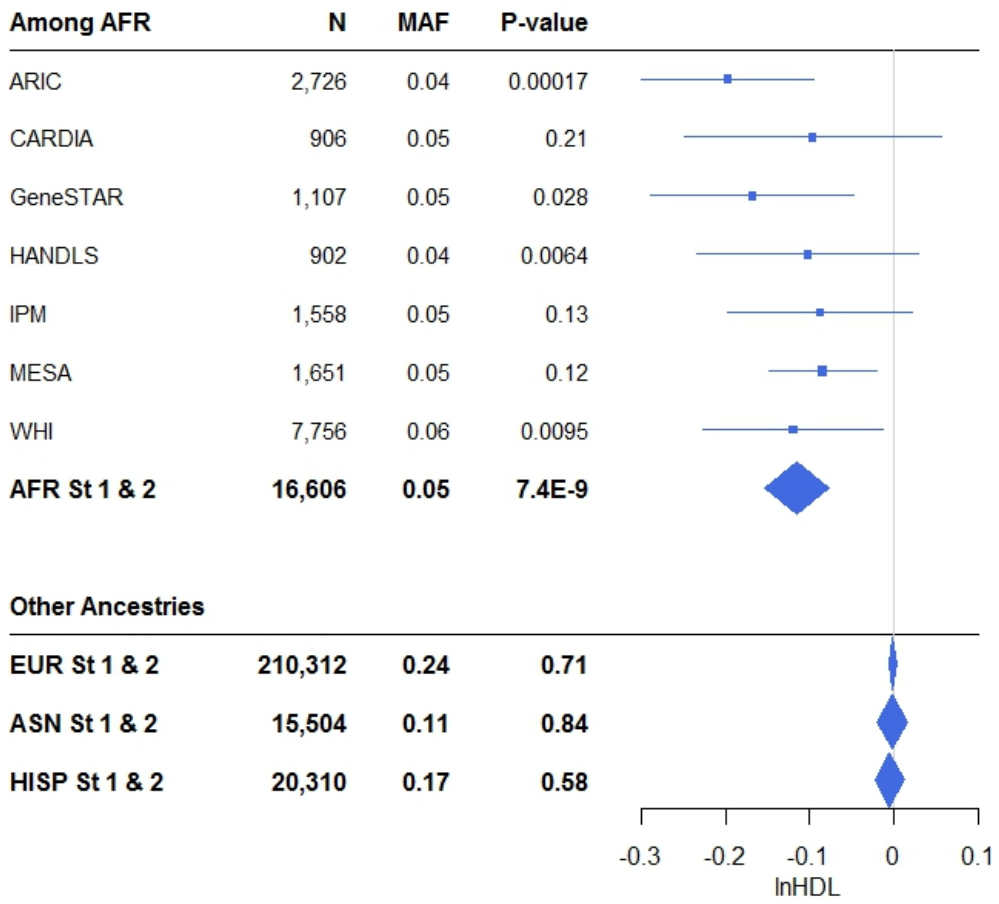
C: Covariates; E: Environmental Exposure (Smoking Status)

Bentley AR *et al*, *Nat Genet*, under review

Interethnic Differences in Interactions

- 6 of 14 novel signals only statistically significant in AFR or in Trans-ancestry meta-analyses included AFR
- Interethnic allele frequency differences contribute:
  - 4 of 6 of these are absent among EUR
  - But 2 of these variants are in higher frequency among other studied populations with no association
- Shown variant is a trans-eQTL for *TAS1R1*, which has been found to influence taste receptors, notably affecting cigarette smoking habits (Choi J-H, *Mol Nutr & Food Res*, 2016).

Interaction of chr 1 locus with Current Smoking Status

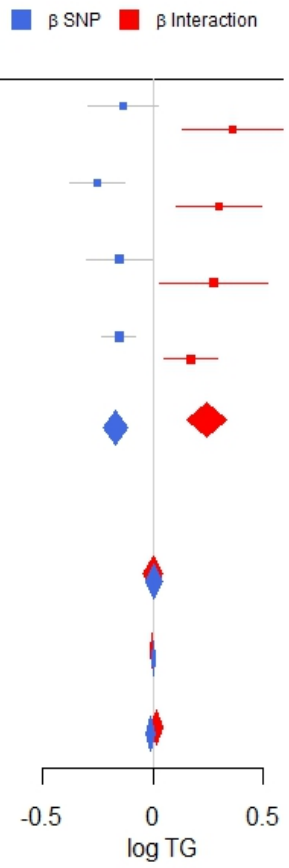


Bentley AR *et al*, *Nat Genet*, under review

Interethnic Differences in Interactions

A.) Among AFR

AFR Study	N	MAF	P-value
ARIC	2,533	0.02	0.005
HABC	1,095	0.04	0.0002
IPM	1,588	0.03	0.06
WHI	7,756	0.03	0.0004
AFR Meta-Analysis	**12,972**	**0.03**	**1.35 x 10-9**
B.) In Other Ancestries			
EA Meta-Analysis	**55,795**	**0.01**	**0.86**
AS Meta-Analysis	**112,887**	**0.34**	**0.61**
HA Meta-Analysis	**20,513**	**0.12**	**0.63**

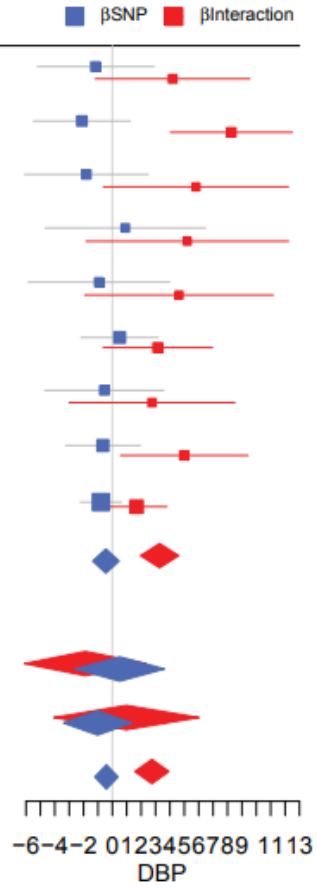


Another example using a 2 degree of freedom joint test of the interaction and main effect

Bentley AR *et al*, *Nat Genet*, under review

Effect of rs11931572 (T5−L10*) and its interaction with EverSmk on DBP

A. AFR Cohorts	N	EAF	2df P
ARIC	2862	0.979	0.203
FamHS	617	0.947	3.45e−05
GENOA	941	0.967	0.193
HABC	1136	0.967	0.02
HUFS	1686	0.97	0.159
JHS	2134	0.961	0.031
MESA	1594	0.968	0.542
BioMe	3101	0.971	0.054
WHI	7925	0.968	0.294
AFR meta−analysis	**23236**	**0.968**	**2.905e−08**
B. Other Ancestries			
EUR meta−analysis	**3133**	**0.987**	**0.6051**
ASN meta−analysis	**2466**	**0.962**	**0.6416**
Trans meta−analysis	**28835**	**0.969**	**9.235e−07**



Such findings also observed in similar meta-analysis of Gene × Smoking Status on **blood pressure measures**

Sung YJ *et al*, *AJHG*, March 2018

Interactions Minor Contribution to % Variance Explained

Trait	Ancestry	Study (N)	Known Lipids Variants	Novel Variants	
				From novel variants	From interactions
HDL	European	WGHS (23,141)	10.1	0.08	0.02
		ARIC (9,052)	9.8	0.05	0.1
		Airwave study (14,002)	11.9	0.1	0.03
	African	ARIC (2,479)	14.9	1.0	1.1
		JHS (1,988)	14.7	1.5	2.0
	Asian	GenSalt (1,813)	17.6	0.6	0.3
		SiMES (2,541)	14.3	0.3	0.2
		SINDI (2,496)	16.8	0.5	0.4
		SCES (1,853)	16.4	0.5	0.7
	Hispanic	SOL (10,091)	11.4	0.3	0.2

Bentley AR *et al*, *Nat Genet*, under review

Trait	Ancestry	Study (N)	Known BP Variants	Novel Variants	
				From novel variants	From interactions
SBP	European	WGHS (22,983)	3.1	1.2	0.4
		ARIC (9,645)	3.2	1.1	0.3
		UK Biobank (137,426)	1.9	0.5	0.1
		Airwave study (14,002)	1.6	0.7	0.5
	African	JHS (2,134)	2.5	1.2	1.4
	Asian	GenSalt (1,835)	2.6	0.4	0.4
		SiMES (2,531)	2.5	0.7	NA
		SINDI (2,491)	1.6	1.2	NA
		SCES (1,848)	2.8	1.1	NA
	Hispanic	SOL (10,294)	1.1	0.2	0.5

Sung YJ *et al*, *AJHG*, March 2018

Meta-Analysis of Genome-Wide Association Studies in African Americans Provides Insights into the Genetic Architecture of Type 2 Diabetes

Maggie C. Y. Ng[1,2¶], Daniel Shriner[3¶], Brian H. Chen[4,5], Jiang Li[2], Wei-Min Chen[6,7], Xiuqing Guo[8],
...
Charles N. Rotimi[3¶*], Donald W. Bowden[1,2,27¶*], for the MEta-analysis of type 2 DIabetes in African Americans (MEDIA) Consortium

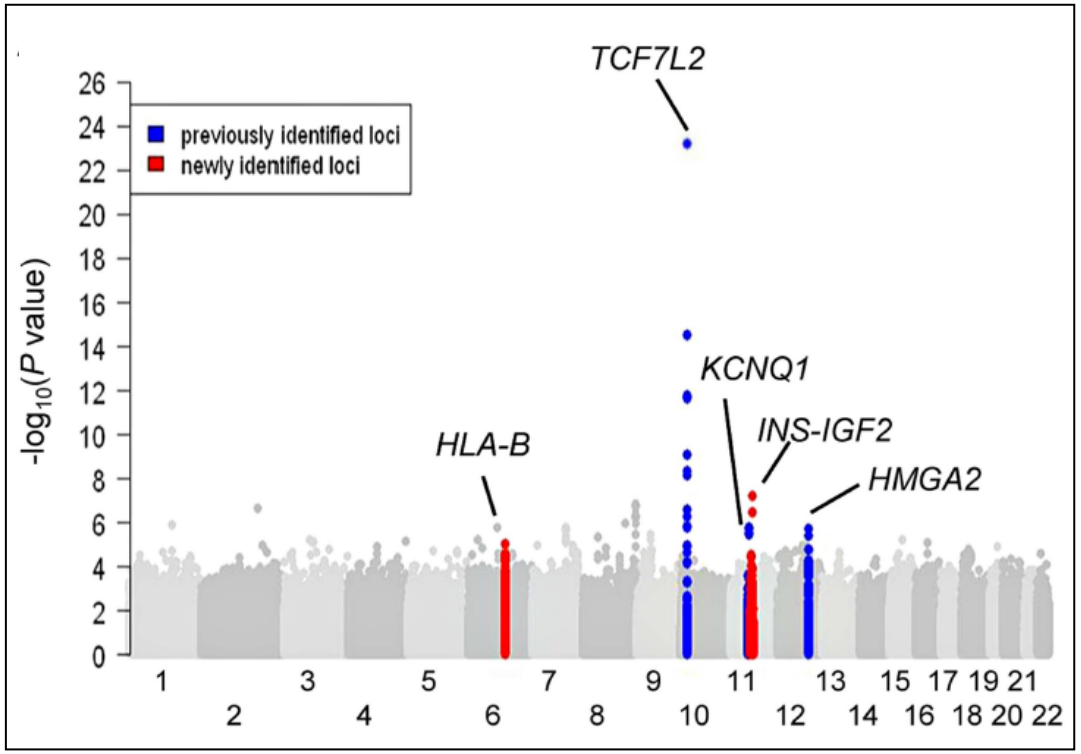
PLOS | GENETICS

August 2014 | Volume 10 | Issue 8 | e1004517

Meta-analysis of T2D in AA (8,284 cases & 15,543 controls). ~2.6M genotyped and imputed SNPs

Replication of 21 loci in AA (6,061 cases and 5,483 controls) and European ancestry (8,130 cases and 38,987 controls).

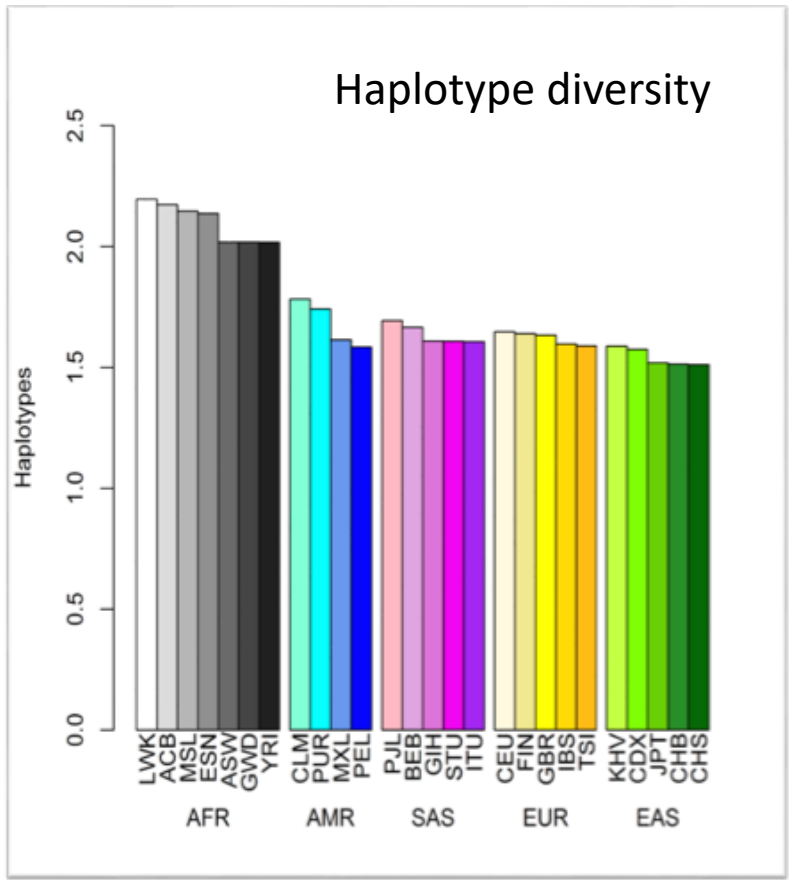
Identified five genome-wide significant loci - 3 known loci (*TCF7L2*, *HMGA2* and *KCNQ1*) & 2 novel loci (*HLA-B* and *INS-IGF2*).

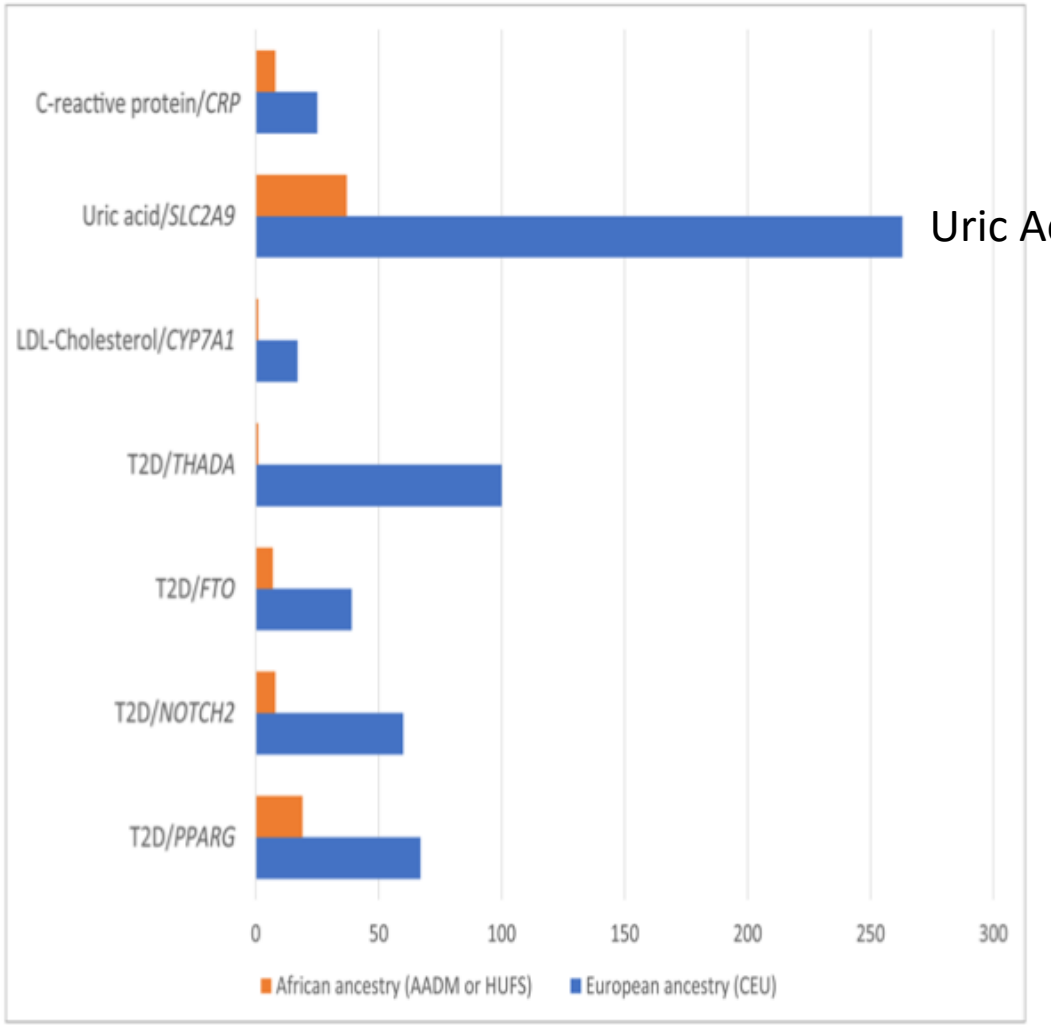


The genome landscape of Africa populations in health and disease

Rotimi CN, Bentley AR, Doumatey AP, Chen G, Shriner D, Adeyemo A

Hum Mol Genetics. 2017





Uric Acid/*SLC2A9*

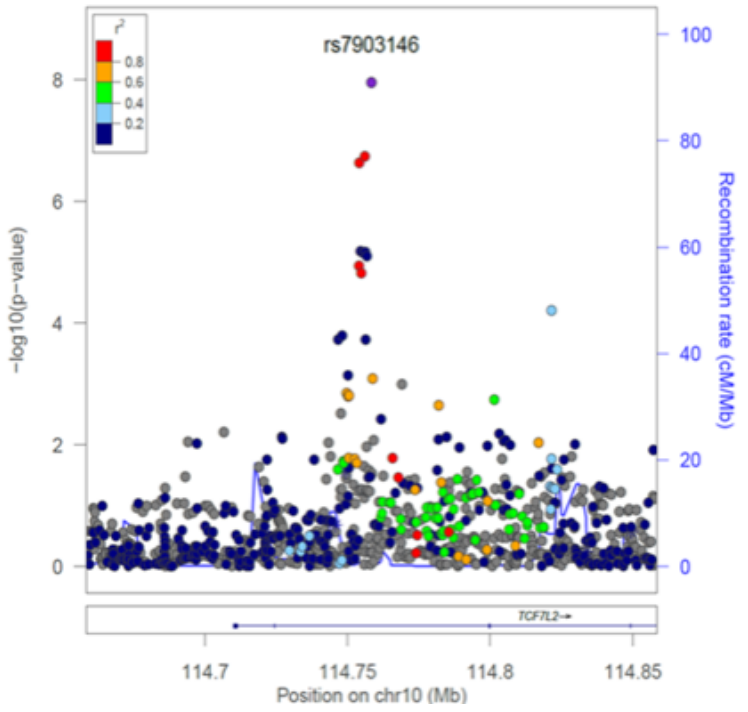
Fine-Mapping - smaller haplotype blocks in African ancestry populations refined genome-wide significant loci

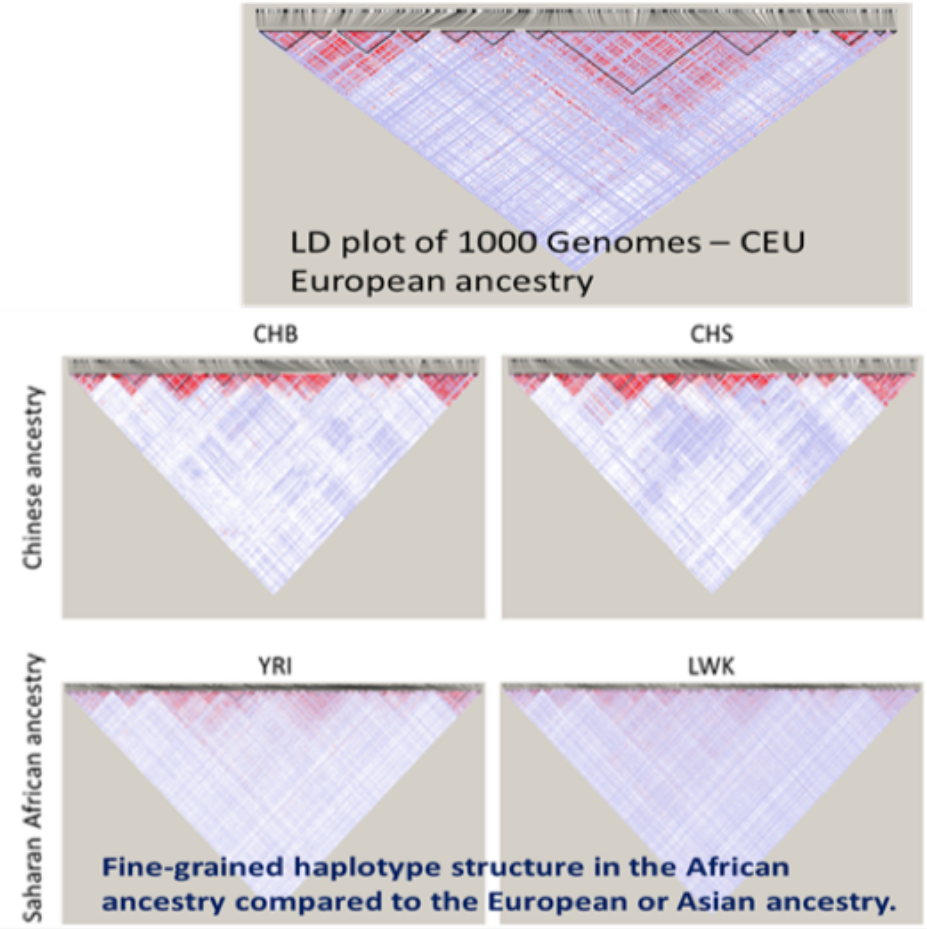
Refining the impact of *TCF7L2* gene variants on type 2 diabetes and adaptive evolution

Agnar Helgason[1], Snæbjörn Pálsson[1,2], Gudmar Thorleifsson[1], Struan F A Grant[1,13], Valur Emilsson[1], Steinunn Gunnarsdottir[1], Adebowale Adeyemo[3], Yuanxiu Chen[3], Guanjie Chen[3], Inga Reynisdottir[1], Rafn Benediktsson[4,5], Anke Hinney[6], Torben Hansen[7], Gitte Andersen[7], Knut Borch-Johnsen[7,8], Torben Jorgensen[9], Helmut Schäfer[10], Mezbah Faruque[3], Ayo Doumatey[3], Jie Zhou[3], Robert L Wilensky[11], Muredach P Reilly[11], Daniel J Rader[11], Yu Bagger[12], Claus Christiansen[12], Gunnar Sigurdsson[4,5], Johannes Hebebrand[6], Oluf Pedersen[7,8], Unnur Thorsteinsdottir[1], Jeffrey R Gulcher[1], Augustine Kong[1], Charles Rotimi[3] & Kári Stefánsson[1]

VOLUME 39 | NUMBER 2 | FEBRUARY 2007 NATURE GENETICS nature genetics

Fine-mapping of TCF7L2





New meta-analysis of two relatively large cohorts from West, East and South Africa

- Fine-mapping of the *TCF7L2* locus - T2D association signal shared between Europeans and Africans (indexed by rs7903146). **Detected a distinct, African-specific signal.**
- We detected one novel signal, MAF = 9.5%; monomorphic in Europeans and Asians.

Genome-wide analysis identifies an African- specific variant in *SEMA4D* that is associated with Body Mass Index

Guanjie Chen[#1], Ayo P. Doumatey[#1], Jie Zhou[1], Lin Lei[1], Amy R. Bentley[1], Olufemi Fasanmade[2], Godfrey Okafor[3], Benjamin Eghan Jr[4], Kofi Agyenim-Boateng[4], Albert Amoah[5], Clement Adebamowo[6], Joseph Acheampong[4], Thomas Johnson[2], Johnnie Oli[3], Daniel Shriner[1], Adebowale A. Adeyemo[1], and Charles N. Rotimi[1]

Obesity. 2017 Apr;25(4):794-800

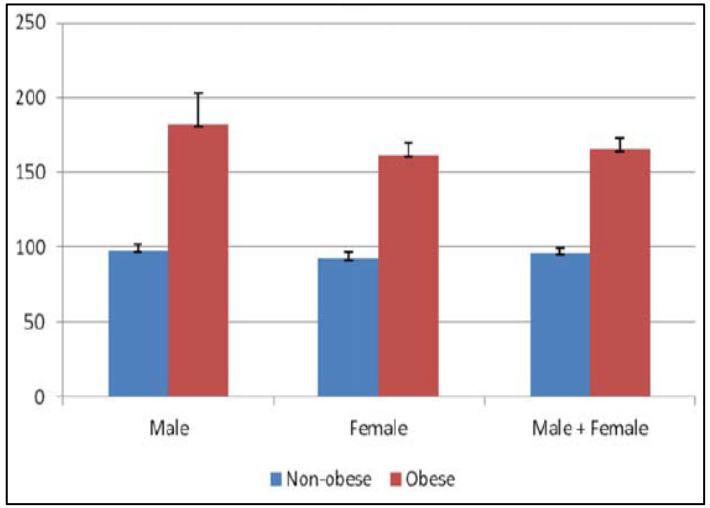
In this first GWAS of BMI in West Africans

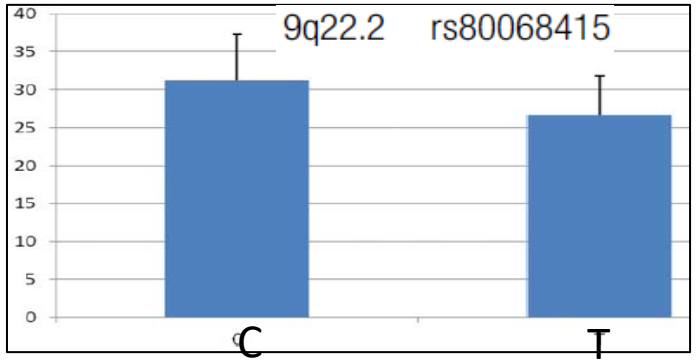
1. Discovered that ~ 1% of West Africans and African Americans carry a genetic variant that increases their risk of obesity

2. Carrier of the risk allele in the semaphorin-4D (SEMA4D) gene were ~6 lbs heavier than those without the genetic variant

3. Would not have been found in GWAS of Europeans and Asians: variant is monomorphic

Serum levels - obese vs non-obese ($p < 0.0001$)





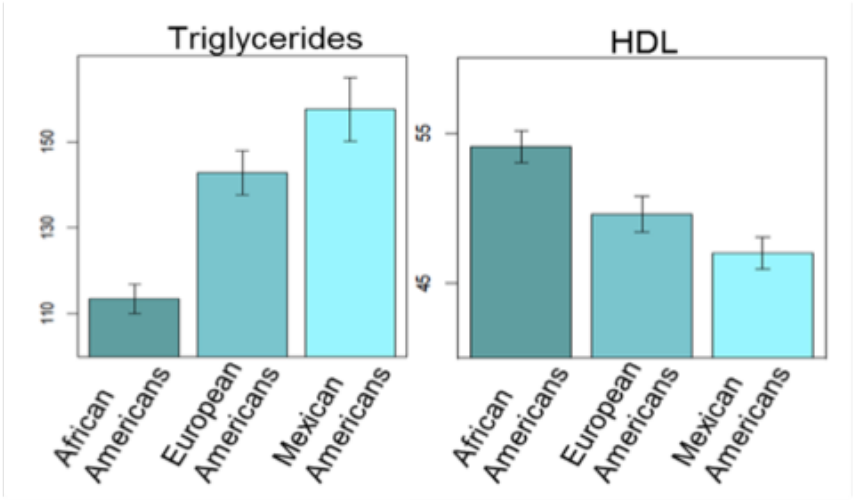
Mean BMI by alleles ($p = 0.0007$)

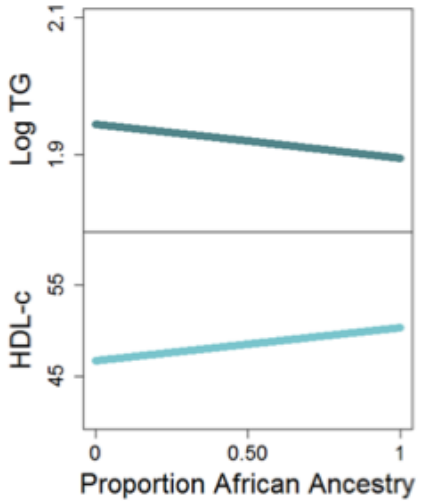
Gene-Based Sequencing Identifies Lipid-Influencing Variants with Ethnicity-Specific Effects in African Americans

PLOS | GENETICS

March 2014 | Volume 10 | Issue 3 | e1004190

Amy R. Bentley[1]*, Guanjie Chen[1], Daniel Shriner[1], Ayo P. Doumatey[1], Jie Zhou[1], Hanxia Huang[1], James C. Mullikin[2], Robert W. Blakesley[2], Nancy F. Hansen[3], Gerard G. Bouffard[2], Praveen F. Cherukuri[4], Baishali Maskeri[2], Alice C. Young[2], Adebowale Adeyemo[1], Charles N. Rotimi[1]*

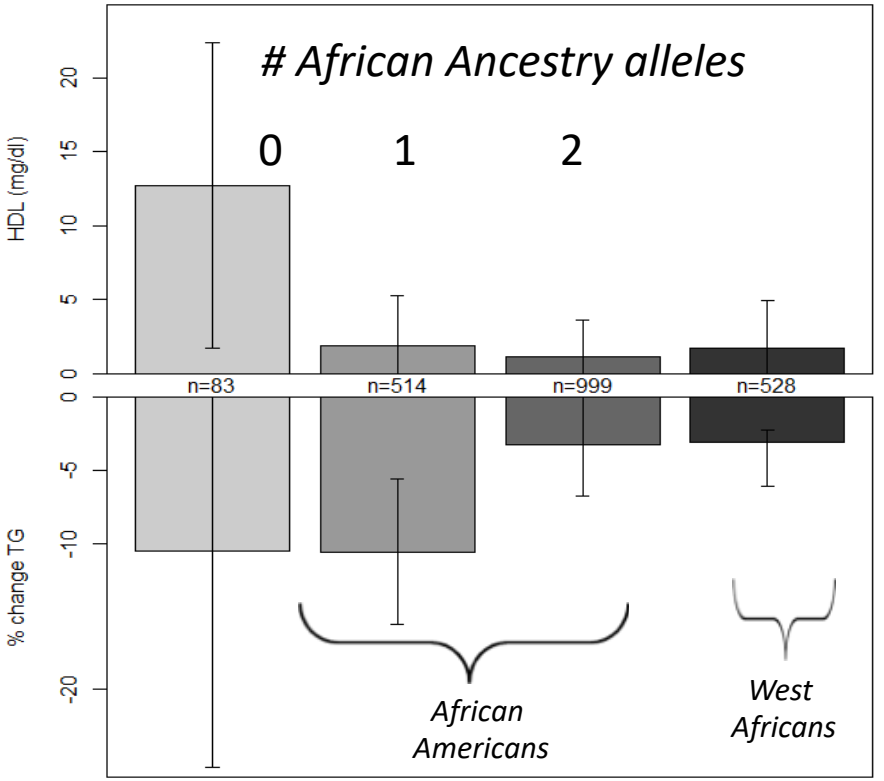




Sequenced 5 genes (*ABCA1*, *LCAT*, *LPL*, *PON1*, and *SERPINE1*) for variant discovery in African Americans with extreme lipids.

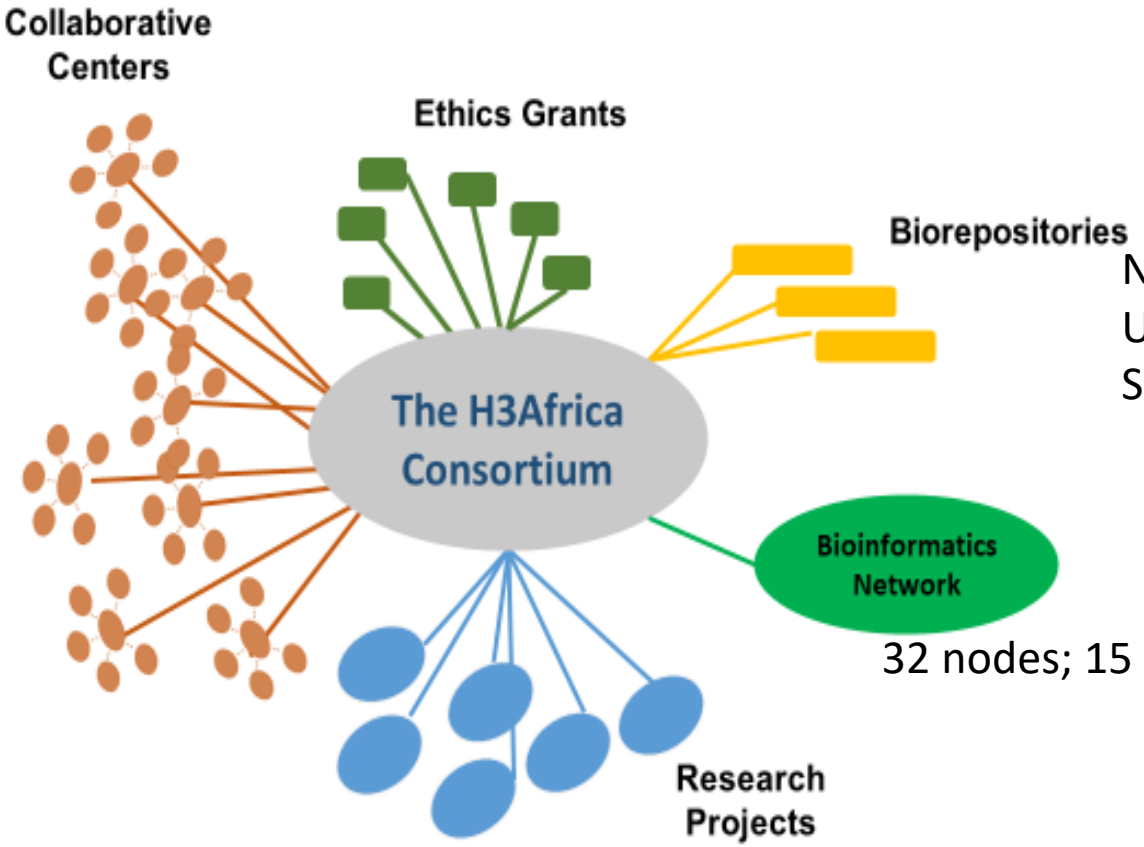
Follow-up in population-based sample of African Americans (n=1694)

Association of *LPL* variants (rs328) depended on ancestry at that locus.



H3Africa
Human Heredity & Health in Africa

Establishing a Consortium Structure



Nigeria
Uganda
S Africa

32 nodes; 15 countries

https://h3africa.org/

H3Africa To Date: By the Numbers

- >$170million of funding
- 48 research projects
- 29 African countries
- >500 investigators
- >54,000 research participants
- >130 publications
- 139 trainees/study coordinators supported
- 53 workshops/courses
- 2062 workshop/meeting attendees

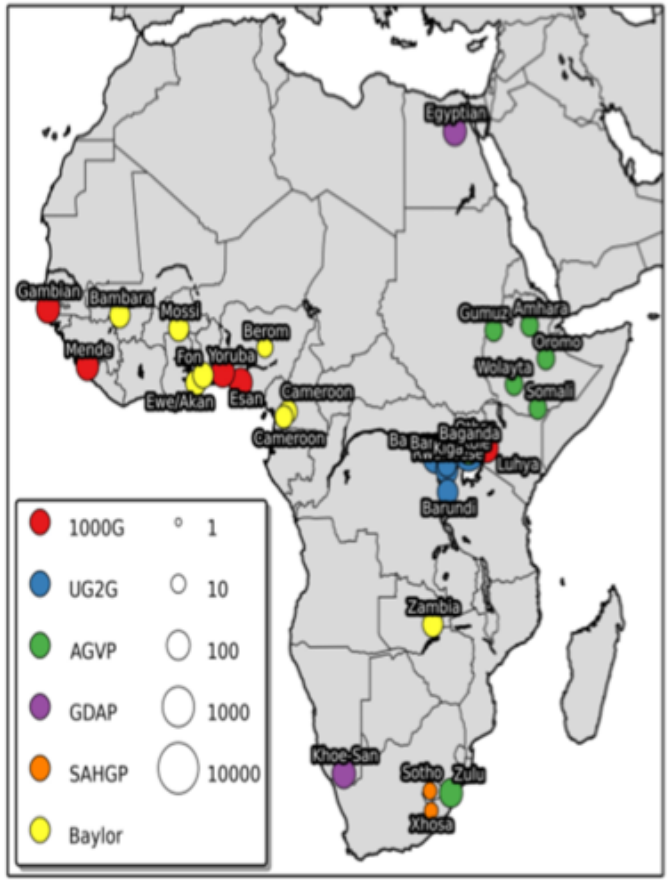
H3Africa Projects		
Collaborative Centers		
African Collaborative Center for Microbiome and Genomics Research (ACCME)	Clement Adebamowo	Nigeria
Eyes of Africa: The Genetics of Blindness	Adeyinka Ashaye	Nigeria
H3Africa Kidney Disease Research Network	Dwomoa Adu, Akinlolu Ojo, Salako Babatunde	Ghana
Genomic Characterization and Surveillance of Microbial Threats in West Africa	Christian Happi	Nigeria
TrypanoGEN: an integrated approach to the identification of genetic determinants of susceptibility to trypanosomiasis	Enock Matovu	Uganda
TrypanoGEN +: The Genetic Determinants of Two Neglected Tropical Diseases	Enock Matovu	Uganda
Collaborative African Genomics Network (CAfGEN)	Mogomotsi Matshaba, Moses Joloba, Adeodata Kekitiinwa,Graeme Mardon, Sununguko Mpoloka, Gabriel Anabwani	Botswana & Uganda
RHDGen: The genetics of rheumatic heart disease network	Bongani Mayosi	South Africa
Burden, spectrum and aetiology of type 2 diabetes in sub-Saharan Africa	Ayesha Motala	South Africa
Center for Research on the Respiratory Microbiota of African Children (ReMAC)	Mark Nicol	South Africa, The Gambia, Malawi & USA
SickleGenAfrica:Sickle Cell Disease Genomics Network of Africa	Dr. Solomon Ofori-Acquah, Dr. Gordon Awandare, Dr. Julie Makani	Ghana
Stroke Investigative Research and Educational Network (SIREN)	Mayowa Owolabi & Bruce Ovbiagele	Nigeria
AWI-Gen Phase 2: Genomic and environmental risk factors for cardiometabolic disease in Africans	Michele Ramsay	South Africa
Genomic and Environmental Risk Factors for Cardiometabolic Diseases in Africans (AWI-Gen)	Michele Ramsay & Osman Sankoh	South Africa & Ghana

Research Projects		
Breast Milk Microbiota Influence on Infant Immunity and Growth (BEAMING)	Alash'le Abimiku	Nigeria
African Female Breast Cancer Epidermiology (AFBRECANE) Study	Clement Adebamowo	Nigeria
Contribution of genetic variation to pharmacokinetic variability and toxicity in patients undergoing multi-drug	Dissou Affolabi	Benin
Systems Biology for Molecular Analysis of Tuberculosis in Ethiopia	Gobena Ameni	Ethiopia
TB Genetics Network in Africa (TBGENAfrica)	Abraham Aseffa	Ethiopia
Host and Microbial Genetic Determinants of Febrile Illness in West Africa	Christian Happi	Nigeria
Clinical and Genetic Studies of Hereditary Neurological Disorders in Mali	Guida Landoure	Mali
Deciphering Developmental Disorders in Africa (DDD-Africa) - Evaluating Clinical Exome Sequencing in an African Setting	Zane Lombard	South Africa
Immunoglobulin gene diversity in an African population and impact on antibody function in HIV infection	Lynn Morris	South Africa
Transgenerational Epigenomics of Trauma and PTSD in Rwanda	Leon Mutesa, Monica Uddin, Derek Wildman, Stefan Jansen	Rwanda
Genetic interactions between human populations and malaria parasites in different environmental settings a	Alfred Amambua Ngwa	The Gambia
The Nasopharyngeal Microbiome and Respiratory Disease in African Children	Mark Nicol	South Africa
Reprogramming of the Trypanosoma brucei Epigenome during Human Infection: Opportunities for New Therapies	Hugh-George Patterton	South Africa
Transgenerational Effects of Maternal Stressors: Investigating the Role of Infant Gene Expression	Dan Stein, Nastassja Koen, Kerry Ressler, Aliza Wingo	South Africa
The Genomics of Schizophrenia in the South African Xhosa People	Dan Stein & Raj Ramesar	South Africa
Hearing Impairment Genetics Studies in Africa (HI-GENES Africa)	Ambroise Wonkam	South Africa, Cameroon, Mali, Ghana

H3Africa Consortium Array

Whole genome datasets used

Source	Number of samples
GDAP	204
UG2G	2000
AGVP	320
1000G*	~2500
TrypanoGEN	168
Baylor	350
SAHGP	16
Total	**~5500**



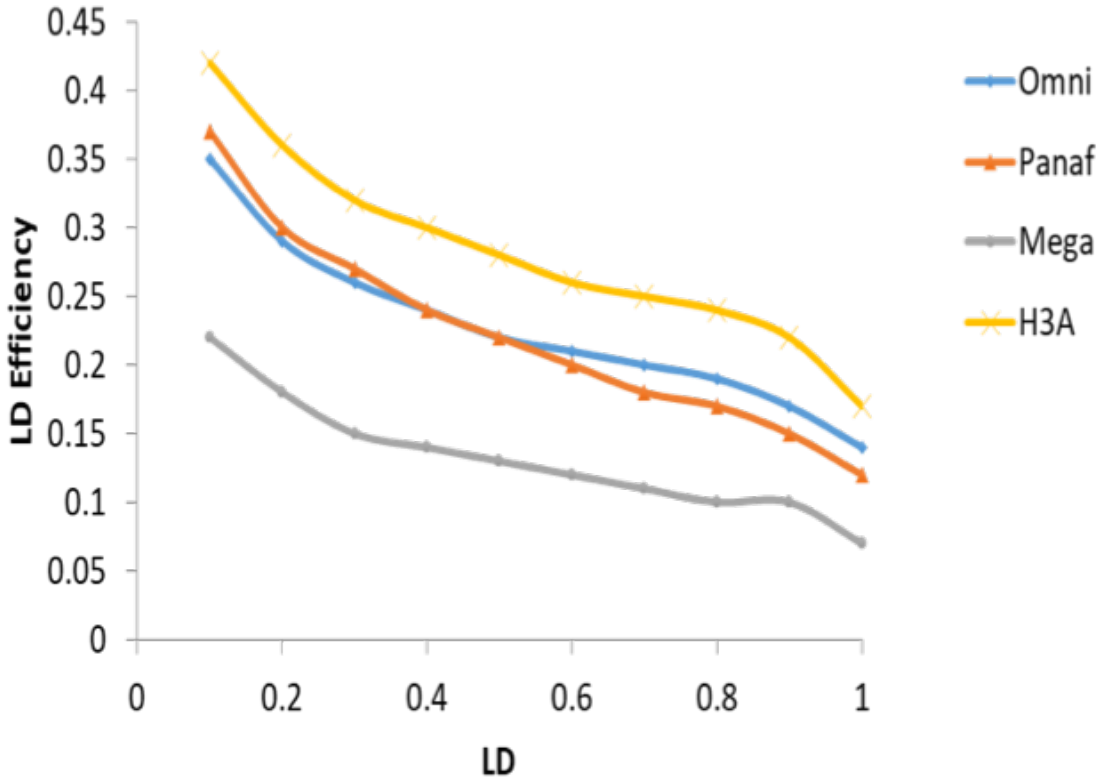
~3,500 African samples

Chip content

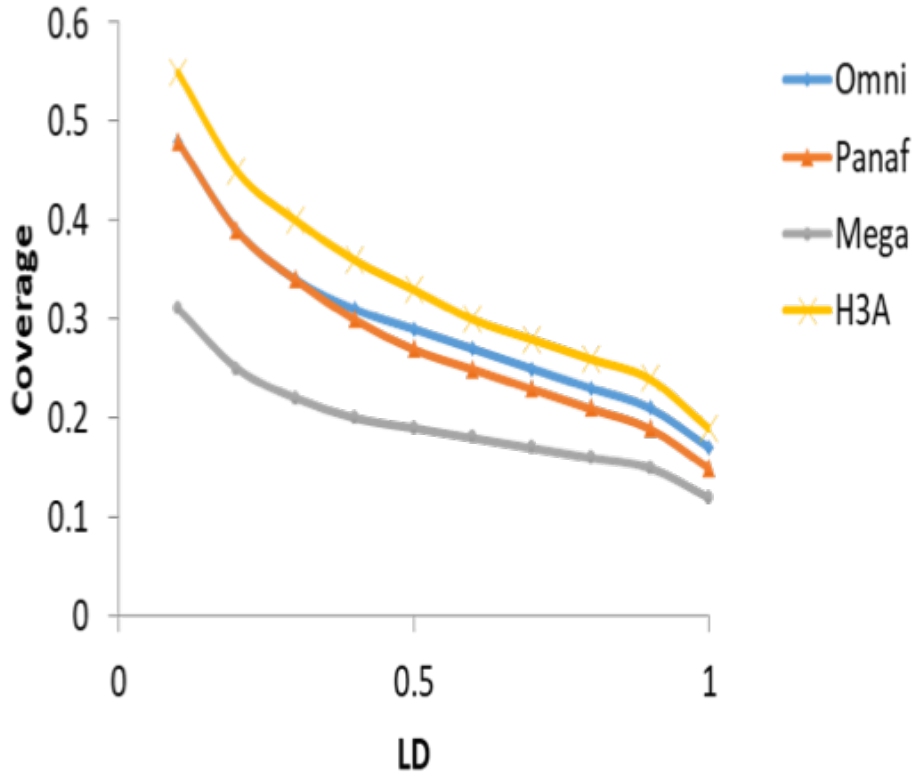
Total (unique SNPs)	2,397,996
Base content	1,561,404
Custom content	862,235 (8,200 replicated)

Exonic	155027
PI requests	9245
COSMIC	10538
ClinVar	26584
PharmGKB	2464
GWAS catalog	22641
MHC	17858
Y	2550
MT	229

Chip evaluation –LD efficiency



Chip evaluation –coverage



H3Africa Consortium Array

- Adebowale Adeyemo & Zane Lombard
- Data providers
- Sanger team:
  - Manj Sandhu
  - Tommy Carstensen
  - Deepti Gurdasani
  - Martin Pollard
- Univ Illinois:
  - Victor Jongeneel
  - Luidmila Sergeevna Mainzer
  - Gloria Rendon
- CHPC, UCT ICTS, Wits cluster
- Support from funders

CRGGH lab folks