A Web-Based Initiative To Accelerate Research On Genetics And Disease in African Americans

K.E. Barnholt[1], A.K. Kiefer[1], H.L. Gates, Jr.[2], M. Nelson[1], M. Mullins[1], E. Baker[3], J. Frank[1], C.D. Bustamante[4], T.W. Love[5], R.A. Kittles[6], N. Eriksson[1], J.L. Mountain[1]

[1]23andMe, Inc, Mountain View, CA; [2]W.E.B. Du Bois Institute for African and African American Studies, Harvard University, Cambridge, MA; [3]23andYou.com; [4]Department of Genetics, Stanford University School of Medicine, Stanford, CA; [5]Onyx Pharmaceuticals, Inc., South San Francisco, CA; [6]College of Medicine, University of Illinois at Chicago, Chicago, IL.

Goals of Study

Little is known about the connection between DNA and disease in African Americans. Genome wide association studies in European populations outnumber those conducted in all other populations combined by a ratio of ten to one[1]. This disparity stems from a tendency to limit studies to groups of more homogeneous ancestry for statistical reasons, and possibly from increased concerns expressed by African Americans about genetic testing[2]. As a result, gene-disease correlations discovered through previous research efforts are almost entirely limited to people of European descent. Since a better understanding of the connection between DNA and disease may lead to opportunities for more personalized health care, it is important to conduct research studies in all population groups[3].

In July 2011, 23andMe launched a research initiative called *Roots into the Future* to accelerate genetic research in populations of African ancestry.

The purpose of this study is to:

- Evaluate the feasibility of rapidly recruiting a cohort of African Americans,
- Assess the interest of this population in participating in a large-scale, web-based genetic study,
- Explore how genetic ancestry predicts the likelihood of having certain variants and reporting certain outcomes,
- Assess our ability to make discoveries by attempting to replicate associations found previously in African Americans, and
- Determine whether genetic associations previously identified in Europeans or Asians are relevant to African Americans.

The long-term goal of this study is to discover genetic markers linked to conditions of particular relevance to the African American community.

Methods

- Individuals who self-identified as African American, Black or of African descent were recruited through large events (NUL), online campaigns (TheRoot.com), print campaigns (Ebony), and word of mouth for the *Roots into the Future* research initiative.
- 10,182 African Americans recruited through the *Roots into the Future* initiative and otherwise provided IRB-approved consent, submitted a saliva sample for DNA extraction through the 23andMe at-home saliva collection kit, were genotyped across a set of between 500,000 and 1 million single nucleotide polymorphisms (SNPs) and were invited to complete medical history surveys.
- We evaluated effectiveness of recruitment through analysis of the response to invitations to take surveys.
- The full 23andMe African American cohort was included in genetic analyses. Cases and controls were selected based on responses to three health history surveys; height and weight were also assessed via web-based surveys.
- We attempted to replicate over 250 associations that had been discovered previously in a broad range of populations: body mass index (BMI) (34 associations discovered in Europeans), height (102, including 13 associated in African Americans), type 2 diabetes (39, including one associated in African Americans), lupus (31, in Europeans and Asians), osteoporosis (39, in Europeans and Asians), migraines (9, in Europeans) and uterine fibroids (4, in Asians).
- Sample sizes: type 2 diabetes: 580 cases, 6086 controls; lupus: 48 cases, 4912 controls; osteoporosis: 142 cases, 4713 controls; migraines: 734 cases, 6058 controls; uterine fibroids: 689 cases, 1638 controls.
- Significance of associations was assessed through regression analyses, taking into account sex, age, and proportion of African ancestry.

Results

Roots into the Future initiative recruitment

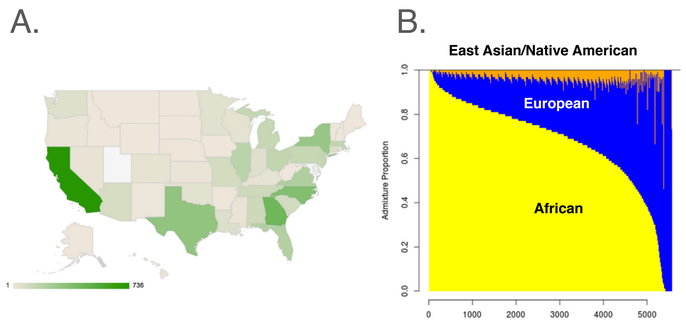

Figure 1. Rapid recruitment of African Americans through the *Roots into the Future* research initiative (>6000 individuals enrolled in eight months). (A) Participants reside in western, northeastern, southern United States. (B) Median estimates of ancestry for each participant: 73% African, 23% European, 4% uncertain.

Total African American cohort

Table 1. Enrollment for overall African American cohort.

	African American Cohort
Total Genotyped	11,431
Total Consented + Genotyped	10,182 (89%)
Of Consented:	
Mean Age	44
Male (%)	42
Taken One of Three Main Health Surveys (%)	65
Average # of Surveys Taken (If Taken Any)	11

Genetic ancestry as predictor

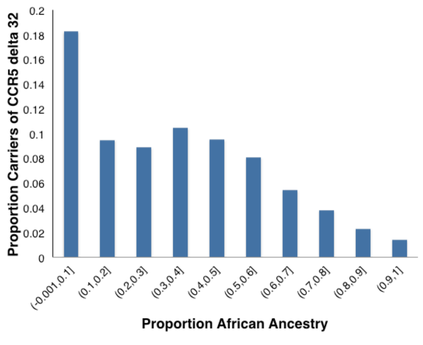

Figure 3. Higher proportion of European ancestry corresponds with a greater chance of carrying variants more common among Europeans than among Africans, such as CCR5 HIV-resistance variant (n = 616 carriers).

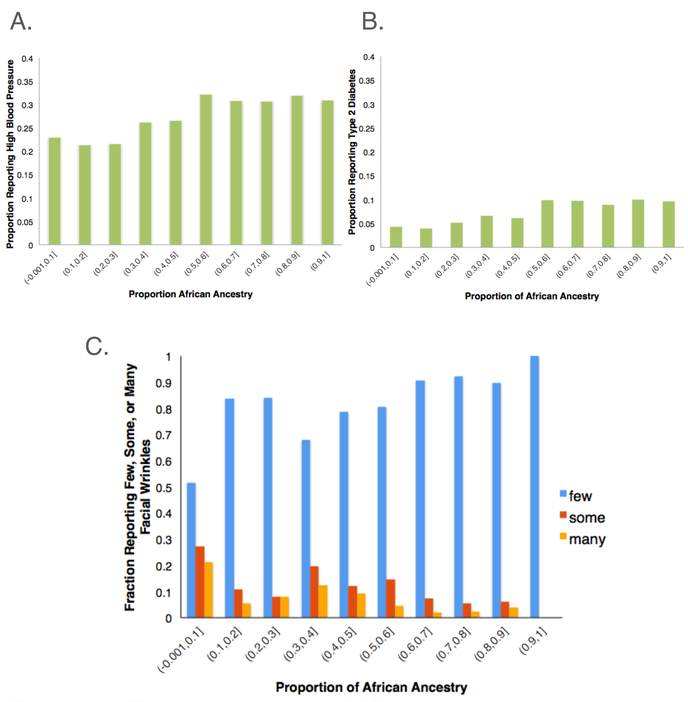

Figure 4. The higher the proportion of African ancestry, the more likely a person reported having (A) high blood pressure (n = 5286; OR per 10% African ancestry = 1.17) and (B) type 2 diabetes (n = 7226; OR per 10% African ancestry = 1.12), and the less likely a person reported having (C) facial wrinkles (n = 791; effect = -0.25 per 10% African ancestry, on scale of 0,1,2). All correlations highly significant: $p<1e-16$.

Replication analyses

Table 2. Replication in African American cohort of associations reported previously for BMI, height, type 2 diabetes, lupus, osteoporosis, migraines. No evidence was found for replication of associations with uterine fibroids.

trait	discovery PMID	orig pop	SNP	gene/location	p-value	effect (BMI units, inches, or OR)	alleles (effect is for 2nd)	same dir.?
BMI	19151714	Eur	rs1421085	[FTO]	3.64E-06	-0.85	C/T	y
BMI	20935630	Eur	rs1558902	[FTO]	7.08E-06	-0.86	A/T	y
BMI	20935630	Eur	rs2287019	[QPCTL]	2.25E-04	-0.67	C/T	y
BMI	20935630	Eur	rs543874	FAM5B---[]-SEC16B	9.64E-04	0.49	A/G	y
BMI	20935630	Eur	rs10968576	[LINGO2]	1.86E-03	0.50	A/G	y
BMI	19079261	Eur	rs10838738	[MTCH2]	4.37E-02	-0.37	A/G	n
BMI	19079261	Eur	rs9939609	[FTO]	4.53E-02	-0.24	A/T	y
BMI	19079260	Mult	rs6265	[BDNF]	3.95E-02	-0.53	C/T	y
BMI	19079260	Mult	rs6499640	[FTO]	4.79E-02	0.26	A/G	n
Height	21998595	Afr	rs9470004	[PPARD]	6.60E-06	-0.37	C/T	y
Height	22021425	Afr	rs606452	[SERPINH1]	1.18E-04	-0.24	A/C	y
Height	22021425	Afr	rs1787200	[DYM]	1.37E-04	-0.23	A/G	y
Height	22021425	Afr	rs1549519	MMD---[]--TMEM100	2.14E-03	-0.22	C/T	y
Height	21998595	Afr	rs7979673	[HMGA2]	5.42E-03	-0.18	C/T	y
Height	21998595	Afr	rs2351491	[ACAN]	1.21E-02	0.17	C/T	y
Height	22021425	Afr	rs7968682	HMGA2--[]---LLPH	1.84E-02	-0.14	G/T	y
Height	21998595	Afr	rs2589113	[CCDC88A]	4.58E-02	0.12	G/T	y
Height	18391952 &21998595	Afr/ Eur	rs6440003	[ZBTB38]	3.11E-03	-0.21	A/G	y
Height	20189936	Asi	rs7678436	FAM184B--[]-DCAF16	2.83E-04	0.21	A/G	y
Height	20189936	Asi	rs7571816	[DIS3L2]	2.18E-02	-0.23	A/G	y
Height	18391951	Eur	rs12198986	SNRNP48---[]-BMP6	1.42E-04	-0.25	A/G	y
Height	20397748	Eur	rs1812175	[HHIP]	5.42E-04	0.22	A/G	y
Height	20881960	Eur	rs6569648	[L3MBTL3]	1.04E-03	-0.39	C/T	y
Height	18391951	Eur	rs8756	[HMGA2]	2.45E-03	0.18	A/C	y
Height	18391950	Eur	rs724016	[ZBTB38]	3.70E-03	0.19	A/G	y
Height	20881960	Eur	rs7909670	CAMK1D--[]--CCDC3	3.97E-03	-0.18	C/T	y
Height	18391951	Eur	rs6763931	[ZBTB38]	5.28E-03	-0.19	A/G	y
Height	20881960	Eur	rs7112925	[RHOD]	6.73E-03	-0.17	C/T	y
Height	20881960	Eur	rs7864648	C9orf93---[]--BNC2	7.44E-03	-0.16	G/T	n
Height	18391951	Eur	rs7153027	FBLN5--[]-TRIP11	9.22E-03	-0.15	A/C	y
Height	18391950	Eur	rs314277	[LIN28B]	1.43E-02	-0.15	A/C	y
Height	20881960	Eur	rs10037512	MEF2C---[]	3.77E-02	0.14	C/T	y
Height	18391951	Eur	rs4743034	[ZNF462]	3.81E-02	-0.12	A/G	y
Height	20881960	Eur	rs7274811	[ZNF341]	4.42E-02	-0.14	G/T	y
Height	18391950	Eur	rs8007661	[TRIP11]	4.48E-02	-0.12	C/T	y
Height	18391952	Eur	rs6854783	[HHIP]	4.70E-02	-0.11	A/G	y
Height	20881960	Eur	rs2597513	HDAC11-[]--FBLN2	4.78E-02	-0.15	C/T	y
Lupus	19838193	Asi	rs1913517	[WDFY4,LRRC18]	1.94E-02	0.61	A/G	y
Lupus	20169177	Asi	rs1128334	[ETS1]	3.29E-02	1.89	C/T	y
Migraines	22683712	Asi/ Eur	rs10166942	hCG_1642047--[]TRPM8	2.61E-02	1.16	C/T	y
Osteop.	19249006	Eur/ Asi	rs11864477	[ADAMTS18]	1.02E-02	1.67	C/T	y
Osteop.	19801982	Eur	rs11898505	[SPTBN1]	4.33E-02	0.69	A/G	y
Type 2 diabetes	18711366	Eur/ Asi	rs2237897	[KCNQ1]	2.36E-03	0.65	C/T	y
Type 2 diabetes	22101970	Eur/ Afr/ Asi	rs7903146	[TCF7L2]	5.78E-03	1.21	C/T	y

*Two-sided *p*-values < 0.05). *FTO, HHIP, HMGA2, ZBTB38* genes are represented by more than one SNP.

Conclusions

- *Roots into the Future* initiative yielded rapid recruitment of a large genotyped cohort of individuals who self-identify as African American.
- A majority of participants chose to take at least one health history survey; those who took at least one took an average of eleven.
- Genetic ancestry is a predictor of several conditions, as demonstrated by the strong correlation between proportion of African ancestry and reporting of high blood pressure, type 2 diabetes, and facial wrinkles.
- Of 13 associations with height reported previously for African Americans, we see significance of $p<0.01$ for six and of $0.01<p<0.05$ for an additional three. In all nine cases the direction of the effect is the same as in the original report.
- We see evidence suggesting replication of 44 of 258 associations with BMI, height, lupus, osteoporosis, type 2 diabetes and migraines. Inferred direction of effect is the same as for the original discovery for all but three of the 44 SNPs.
- We expect continued growth of this African American cohort and consideration of local ancestry to lead to power sufficient for detecting novel associations.

Acknowledgments

We thank *Roots into the Future* participants and other 23andMe customers who consented to participate in research for their support and time dedicated to research. We also thank our project advisors and recruitment partners for enabling this study. Finally, we thank the employees of 23andMe who contributed to the development of the infrastructure that made this research possible.

References and Resources

1. Need and Goldstein, *Trends in Genetics*, 2011.
2. Halbert et al., *Cancer Epidemiol Biomarkers Prev*, 2006.
3. Daar and Singer, *Nature Reviews Genetics*, 2004.