HHS Public Access

Author manuscript

Genet Med. Author manuscript; available in PMC 2021 October 01.

Published in final edited form as:

Genet Med. 2021 April ; 23(4): 777–781. doi:10.1038/s41436-020-01034-4.

A State-Based Approach to Genomics for Rare Disease and Population Screening

Kelly M. East, MS[1], **Whitley V. Kelley, MS**[1], **Ashley Cannon, PhD, MS**[2], **Meagan E. Cochran, MS**[1], **Irene P. Moss, MA**[2], **Thomas May, PhD**[1,3], **Mariko Nakano-Okuno, PhD**[2], **Stephen O. Sodeke, PhD, MA**[4], **Jeffrey C. Edberg, PhD**[2], **James J. Cimino, MD**[2], **Mona Fouad, MD, MPH**[2], **William A. Curry, MD**[2], **Anna C.E. Hurst, MD, MS**[2], **Kevin M. Bowling, PhD**[1], **Michelle L. Thompson, PhD**[1], **E. Martina Bebin, MD, MPA**[2], **Robert D. Johnson, MS**[2], **AGHI Consortium**[*], **Gregory M. Cooper, PhD**[1], **Matthew Might, PhD**[2], **Gregory S. Barsh, MD, PhD**[1], **Bruce R. Korf, MD, PhD**[2]

[1]HudsonAlpha Institute for Biotechnology, Huntsville, AL

[2]University of Alabama at Birmingham, Birmingham, AL

[3]Washington State University, Pullman, WA

[4]Tuskegee University, Tuskegee, AL.

Abstract

Corresponding Author: Kelly East, MS, (256) 327-0461, keast@hudsonalpha.org.
[*]A list of authors and their affiliations appears at the end of the paper.

CRediT Author Statement

Conceptualization: K.M.E., W.V.K., A.C., T.M., G.M.C., M.M., G.S.B., B.R.K.; Data curation: K.M.E., W.V.K., A.C., M.E.C., I.P.M., K.M.B., M.L.T., R.D.J.; Formal Analysis: K.M.E, W.V.K., A.C., M.E.C, I.P.M., K.M.B., M.L.T., R.D.J.; Funding acquisition: I.P.M., G.M.C., M.M., G.S.B., B.R.K.; Investigation: W.V.K., M.E.C., I.P.M., A.C.E.H., K.M.B., M.L.T., E.M.B.; Methodology: K.M.E., W.V.K., A.C., I.P.M., T.M., M.N.O., S.O.S., J.C.E., J.J.C., M.F., W.A.C., A.C.E.H., G.M.C., M.M., G.S.B., B.R.K.; Project Administration: I.P.M., Supervision: G.M.C., M.M., G.S.B., B.R.K.; Writing – original draft: K.M.E., W.V.K., A.C., M.E.C.; Writing – review & editing: K.M.E, W.V.K., A.C., M.E.C, I.P.M, T.M., M.N.O., S.O.S., J.C.E., J.J.C, M.F., W.A.C., A.C.E.H., K.M.B., M.L.T., E.M.B., R.D.J., G.M.C, M.M., G.S.B., B.R.K.

Ethics Declaration

The Alabama Genomic Health Initiative was reviewed and approved by the University of Alabama at Birmingham (UAB) IRB (protocol number F170303004). Informed consent was obtained from all participants.

Data Sharing

The data that support the findings of this study are available from the corresponding author, [KE]. Global screening array identified and Sanger confirmed P/LP variants have been submitted to ClinVar (study ID: AGHI_GT).

AGHI Consortium: Non-author contributors

Aras Acemgil, MBA[2]
David K. Crossman, PhD[2]
Candice R. Finnila, PhD[1]
David E. Gray, MS[1]
Veronica Greve, MS[1]
Sharonda Hardy, MAEd[2]
Susan M. Hiatt, PhD[1]
Donald R. Latner, PhD[1]
James M. J. Lawlor, MS[1]
Edrika L. Miskell, MPH[2]
Whitney Narmore, BS[2]
Julie H. Schach, BS[2]

[1]HudsonAlpha Institute for Biotechnology, Huntsville, AL; [2]University of Alabama at Birmingham, Birmingham, AL; [3]Washington State University, Pullman, WA; [4]Tuskegee University, Tuskegee, AL

Purpose: The Alabama Genomic Health Initiative (AGHI) is a state-funded effort to provide genomic testing. AGHI engages two distinct cohorts across the state of Alabama. One cohort includes children and adults with undiagnosed rare disease; a second includes an unselected adult population. Here we describe findings from the first 176 rare disease and 5369 population cohort AGHI participants.

Methods: AGHI participants enroll in one of two arms of a research protocol that provides access to genomic testing results and biobank participation. Rare disease cohort participants receive genome sequencing to identify primary and secondary findings. Population cohort participants receive genotyping to identify pathogenic and likely pathogenic variants for actionable conditions.

Results: Within the rare disease cohort, genome sequencing identified likely pathogenic or pathogenic variation in 20% of affected individuals. Within the population cohort, 1.5% of individuals received a positive genotyping result. The rate of genotyping results corroborated by reported personal or family history varied by gene.

Conclusion: AGHI demonstrates the ability to provide useful health information in two contexts: rare undiagnosed disease and population screening. This utility should motivate continued exploration of ways in which emerging genomic technologies might benefit broad populations.

Keywords

population screening; genotyping; state-wide; genome sequencing; family history; penetrance

INTRODUCTION

Emerging genomic technologies have begun to transition from “promise” to practical use. Genomic sequencing has been well-established as a diagnostic tool for patient populations with rare disease.[1–2] However, few broad population-wide programs of enrollment and testing with a primary goal of returning medically-actionable results have been conducted, though enrollment, testing, and result return have occurred or are occurring in the setting of other large-scale research projects.[3,4] Studies have suggested that approximately one percent of individuals in the general population may harbor pathogenic or likely pathogenic variation in one or more highly-penetrant, medically-actionable genes.[5–7]

The ACMG published recommendations regarding a list of medically actionable genes for which pathogenic and likely pathogenic variation should be reported in clinical sequencing, regardless of the primary indication for testing.[8] However, these recommendations were intended to apply to clinically-indicated sequencing, rather than population screening.[9] Limited guidelines exist for determining which, if any, genetic variation should be reported on a population level. In addition, recent work has shown that the actual penetrance of medically-actionable genes may be lower than previously expected at a population level.[10] Further research is required to determine if penetrance, management guidelines, and other considerations would be altered outside of a clinically-indicated population.

In addition to a paucity of population penetrance data, the underrepresentation of racial and ethnic minorities is another area of concern for genomic research. Most studies include an overrepresentation of individuals of European descent compared to the general population.

This discrepancy has significant implications for health disparities, access to care, and variant interpretation.[11]

The Alabama Genomic Health Initiative (AGHI; www.uabmedicine.org/aghi) was sponsored by the state of Alabama to test the feasibility and utility of genomic platforms to implement and realize these benefits at a broad, state-wide population level. AGHI seeks both to address racial disparity in genomic research recruitment and representation and investigate personal and clinical outcomes within its two distinct cohorts of participants: rare disease and population. Engaging these two cohorts allows for the implementation and evaluation of both genome sequencing and a more cost-effective genotyping assay. AGHI began in June of 2017 and is led and implemented through a collaboration between UAB Medicine and the HudsonAlpha Institute for Biotechnology. Leaders at each institution co-direct the initiative, and working groups have been established to carry key aspects of the project forward. Working groups include Bioethics, Data & Biobank, Education & Genetic Counseling, Genomics, Recruitment, and Engagement. While recruitment and analysis is ongoing, this paper describes data from the initiative to date, including 386 participants (176 probands) in the rare disease cohort and 5369 participants in the population cohort.

MATERIALS AND METHODS

Rare Disease Cohort Pipeline

Participation in the rare disease cohort is available to all Alabama residents (children and adults) for whom there is a strong suspicion of an undiagnosed, genetic condition. Potential participants are referred by their healthcare providers. Referrals are evaluated by AGHI clinicians to determine the utility of genome sequencing. Rare disease participants are enrolled in-person at one of three clinical research sites. Medical and family history is gathered, study participation is described, informed consent is obtained, and blood samples are collected from the proband and, when possible, biological parents. If applicable, samples are also collected from similarly-affected siblings. Participants provide optional consent for secondary finding disclosure, biobank submission, and research recontact.

Genome sequencing is used to identify a genetic cause for symptoms (primary findings) as well as other unrelated, actionable genetic changes (secondary findings).

Probands and affected siblings are sequenced with an average depth of 30X. Pathogenic (P), likely pathogenic (LP), and variants of uncertain significance (VUS) are reported for primary findings, while only P and LP variants are reported for secondary findings. Secondary findings include variants in genes on the ACMG SF v2.0 gene list[8] as well as P/LP variants in other genes identified incidentally through the primary analysis pipeline.

All reportable variants are confirmed via targeted, clinical Sanger sequencing in a CLIA certified lab. Participants receive results, regardless of whether there are findings, via in-person consultation or phone conversation with a genetic counselor and/or geneticist. A report summarizing the findings is written by genetic counselors and provided to the participant in-person or via mail.

Population Cohort Pipeline

Participation in the population cohort is available to all adult Alabama residents. Permanent and "pop-up" enrollment sites are located in medical and non-medical community settings to provide access to a broad, representative population.[12] "Pop-up" sites were chosen based on identifying areas of the state where participation rates were low and/or at a distance from permanent recruitment locations. Participants are recruited via media, social media, and word of mouth. During enrollment, participants meet with a recruitment team member to discuss study benefits, limitations, and logistics and provide informed consent. Participants can elect to participate in the biobank, have results shared with a healthcare provider, and/or be contacted about future research. Basic information is collected from participants including demographic data, contact information, and a targeted personal and family health history.

A health history questionnaire was developed to identify participants with a personal and/or family history relevant to the medically actionable gene list (see Supplementary Appendix). This information is triaged to identify participants who have a strong personal or family history suggestive of a genetic risk factor. Family history triage criteria were developed based on current clinical guidelines and published literature.[13–15] This triage is performed prior to and independent of any genotyping results.

A SNP genotyping assay, the Illumina Global Screening Array (GSA-24, v1.0 and GSA-24, v2.0), is used to detect rare, damaging variants in highly penetrant, medically actionable genes derived from the ACMG SF v2.0 gene list.[8] These variants are classified as P/LP in accordance with ACMG recommendations and reside in ClinVar.[5] Reportable variants are confirmed by targeted, clinical Sanger sequencing in a CLIA certified lab.

Participants with no genotyping findings receive a report via mail explaining the limitations of a genotyping test. If a participant has a family history that was flagged as having an elevated risk of hereditary disease, the mailed report is modified to highlight the suggestive history and include a recommendation for genetics follow-up (see Supplementary Appendix). In contrast, participants with positive genotyping results receive a phone call from a genetic counselor to describe the results, implications, and next steps. An individualized research result report is written by genetic counselors and sent to the participant following the phone disclosure (see Supplementary Appendix). With consent, results are sent to the participant's healthcare provider.

RESULTS

Rare Disease Cohort Findings

To date, 176 probands and affected siblings have consented to be part of the rare disease cohort, representing 154 families. Because informative relatives were also enrolled when available and appropriate, the total number of enrolled individuals was 386 (12 of whom only received Sanger sequencing for a variant of interest). 77.3% of 154 families had at least one biological parent available, with 70.6% of those trios and 29.4% duos. See Table 1 for detailed demographic information. 91.5% (n=353) of participants consented to receive

secondary findings, when available, and 72.3% (n=279) of participants consented to future recontact and to allow their sample to become part of a research biobank.

Of the 176 affected probands and siblings, 19 individuals (10.8%) received a pathogenic result, 16 (9.1%) received a likely pathogenic result, and 42 (23.9%) received a result of a variant of uncertain significance. When counting by family unit, proportions were largely similar (12.3% pathogenic, 9.1% likely pathogenic, and 24.7% variant of uncertain significance). Thirteen individuals (3.5%, including both affected and unaffected participants) from nine families received a medically-actionable secondary finding. A list of genes in which variation was reported for participants in the rare disease cohort is available in supplementary material.

Population Cohort Findings

To date, 5369 individuals have consented to be a part of the population cohort. This cohort includes individuals residing in each of the 67 Alabama counties, with a mean age of 51 years. A majority of population participants were female (75%) and Caucasian (74%). Additional demographics of this cohort and a comparison to the overall Alabama population are summarized in Table 1.

The vast majority of participants consented to the biobank and to research recontact (92%, n=4926). A smaller percentage (44%, n=2359) of participants consented to have their AGHI participation and subsequent results shared with their healthcare provider.

Eighty-one positive genotyping results among 80 individuals (1.5%) were identified in the population cohort. These results include risk-increasing variants for hereditary cancer, cardiomyopathy, malignant hyperthermia, and hypercholesterolemia. A summary of genotyping findings among the population cohort is found in Table 2. Genetic counselors were successful in reaching 76 positive participants (95%) to discuss the results via phone. In addition, result reports were delivered by mail to all positive result participants.

Nearly half (46%, n=2414) of the 5305 participants for whom a health history questionnaire was completed reported a personal or family history suggestive of an elevated disease risk. Of participants receiving a positive genotyping result, 73% (n=58) had a personal or family history considered at elevated risk. Notably, in approximately half of cases where a participant had both an elevated *a priori* risk of hereditary disease and a P/LP genotyping result, the reason for the flagged history was not relevant to the genotyping result found. When only considering histories relevant to the genomic result(s) identified, the proportion of individuals with a corroborating history dropped to 36% (n=29). The proportion of positive results corroborated by reported personal and family history varied by gene (Table 2).

DISCUSSION

AGHI has engaged a diverse group of individuals from across the state in both the rare disease-focused genome sequencing cohort and a population-based genotyping cohort for highly penetrant, medically actionable conditions.

The diagnostic rate within our rare disease cohort is consistent with previously published rates in similar populations.[1–2] Given the expense of genome sequencing and current lack of widespread insurance coverage, it is possible that AGHI has provided a path to diagnosis in some participants that would not have been available through standard clinical evaluations, but further research will be needed to substantiate this. Of note, evaluation of the rare disease cohort is more resource-intensive and available at fewer, exclusively urban enrollment sites, and this may represent a barrier to access that is not as applicable to the population cohort.

While a considerable proportion of individuals in the population cohort with positive results reported a personal or family history potentially related to an actionable genetic condition, only a minority (36%) of individuals with a P/LP variant found by genotyping reported a history relevant to their genomic result. This indicates the potential utility of genomic screening in addition to traditional risk assessment tools such as detailed family history collection. The lack of concordance between genomic test results and reported personal and family history may also suggest that the penetrance of some of these variants, particularly those related to cardiac conditions, may not be as high as previously expected in an unselected population. This is consistent with other recent studies.[16,17]

The advent of large scale population testing ("All of Us,"[3] eMERGE,[4] etc.) has motivated increased interest in the potential value of genetic testing for screening purposes.[18] This is because, despite considerable value, current screening through imaging, testing of cholesterol and blood pressure, etc., and traditional family health history intake all have recognized deficiencies in identification of health risk.[19] A large fraction (44%) of AGHI participants had an increased *a priori* risk based on reported personal and family history. The combination of history triaging and customized non-informative result reports may help to mitigate the risks of limited sensitivity genetic screening. We have found this to be a sustainable use of genetic counseling resources and a potential model for other large-scale population genetic screening programs.

Major limitations to overall AGHI recruitment include ascertainment bias and, in the case of the rare disease cohort, its recruitment concentration in urban areas. The population cohort reflects an enrollment bias toward those with family histories of inherited disease, this is not uncommon for population-wide genetic screening programs[20] and reflects the phenomenon that any voluntary population-wide intervention will have uptake that over-represents those with strong personal interest in the intervention's target. Detection of disease-causing variants is limited by the use of a genotyping test unable to detect all P/LP variants in the genes tested. This limitation is more pronounced in individuals from populations not adequately represented in genomic databases.

A future research interest is to study clinical and personal outcomes in both cohorts. Investigators would also like to further explore participants' decision-making at the time of consent (e.g. decision to/not to share results with healthcare providers) and result understanding. Of particular interest is gaining additional information regarding the clinical history and follow-up actions in participants receiving a result discordant from their reported personal and family history, to further explore the reason(s) for discordance.

In conclusion, AGHI demonstrates a feasible method of engaging a diverse population in genomics via genome sequencing and genotyping. Reported personal and family history among population participants receiving positive results suggests utility of genomic screening as a potential risk assessment tool. It also highlights opportunities for further investigation of genomic variation in an unselected population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the AGHI participants for their contributions to this study. This study was supported by funding from the State of Alabama. This study was also supported by the UAB Center for Clinical and Translational Science, funded by a NIH CTSA grant (UL1TR003096).

References

1. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. Nature 2015;519:223–228. [PubMed: 25533962]
2. Ramoni RB, Mulvihill JJ, Adams DR, et al. The Undiagnosed Diseases Network: Accelerating Discovery about Health and Disease. Am J Hum Genet 2017;100:185–192. [PubMed: 28157539]
3. All of Us Research Program Investigators, Denny JC, Rutter JL, et al. The “All of Us” Research Program. N Engl J Med. 2019;381(7):668–676. doi:10.1056/NEJMsr1809937 [PubMed: 31412182]
4. Wiesner GL, Kulchak Rahm A, Appelbaum P, et al. Returning Results in the Genomic Era: Initial Experiences of the eMERGE Network. J Pers Med. 2020;10(2):30. Published 2020 Apr 27.
5. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17:405–424. [PubMed: 25741868]
6. Amendola LM, Dorschner MO, Robertson PD, et al. Actionable exomic incidental findings in 6503 participants: challenges of variant classification. Genome Res 2015;25:305–315. [PubMed: 25637381]
7. Natarajan P, Gold NB, Bick AG, et al. Aggregate penetrance of genomic variants for actionable disorders in European and African Americans. Sci Transl Med 2016;8:364ra151.
8. Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. Genet Med 2017;19:249–255. [PubMed: 27854360]
9. ACMG Board of Directors. The use of ACMG secondary findings recommendations for general population screening: a policy statement of the American College of Medical Genetics and Genomics (ACMG). Genet Med 2019;21:1467–1468. [PubMed: 31019278]
10. Zoltick ES, Linderman MD, McGinniss MA, et al. Predispositional genome sequencing in healthy adults: design, participant characteristics, and early outcomes of the PeopleSeq Consortium. Genome Med 2019;11:10. [PubMed: 30808425]
11. Spratt DE, Chan T, Waldron L, et al. Racial/Ethnic Disparities in Genomic Sequencing. JAMA Oncol 2016;2:1070–1074. [PubMed: 27366979]
12. May T, Cannon A, Moss IP, et al. Recruiting diversity where it exists: The Alabama Genomic Health Initiative. J Genet Couns 2020;29:471–478. [PubMed: 32220047]
13. Adam MP, Ardinger HH, Pagon RA, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993–2020. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1116/.

14. National Comprehensive Cancer Network. Genetic/Familial High-Risk Assessment: Breast, Ovarian, and Pancreatic (Version 1.2020). https://www.nccn.org/professionals/physician_gls/pdf/genetics_bop.pdf. Accessed May 1, 2020.
15. National Comprehensive Cancer Network. Genetic/Familial High-Risk Assessment: Colorectal (Version 1.2020). https://www.nccn.org/professionals/physician_gls/pdf/genetics_colon.pdf. Accessed May 1, 2020.
16. Bick AG, Flannick J, Ito K, et al. Burden of rare sarcomere gene variants in the Framingham and Jackson Heart Study cohorts. Am J Hum Genet 2012;91:513–519. [PubMed: 22958901]
17. Peloso GM, Rader DJ, Gabriel S, Kathiresan S, Daly MJ, Neale BM. Phenotypic extremes in rare variant study designs. Eur J Hum Genet 2016;24:924–930. [PubMed: 26350511]
18. US Preventive Services Task Force, Owens DK, Davidson KW, et al. Risk Assessment, Genetic Counseling, and Genetic Testing for BRCA-Related Cancer: US Preventive Services Task Force Recommendation Statement [published correction appears in JAMA. 2019 Nov 12;322(18):1830]. JAMA. 2019;322(7):652–665. [PubMed: 31429903]
19. Murray MF, Evans JP, Khoury MJ. DNA-Based Population Screening: Potential Suitability and Important Knowledge Gaps [published online ahead of print, 2019 Dec 6]. JAMA. 2019;10.1001/jama.2019.18640.
20. Manickam K, Buchanan AH, Schwartz MLB, et al. Exome Sequencing-Based Screening for BRCA1/2 Expected Pathogenic Variants Among Adult Biobank Participants. JAMA Netw Open. 2018;1(5):e182140. Published 2018 Sep 7. [PubMed: 30646163]

Table 1.

AGHI participant demographics by rare disease and population cohort

Demographics	Rare disease cohort	Population cohort	State of Alabama[a]
Sex			
Female	82 (47%)	4010 (75%)	52%
Male	94 (53%)	1359 (25%)	48%
Mean Age (yrs)	5 (Pediatric) 35 (Adult)	51	--
Race			
Caucasian	125 (71%)	3965 (74%)	69%
African American	38 (22%)	1063 (20%)	27%
Asian	6 (3%)	128 (2%)	2%
Other/Unknown	7 (4%)	213 (4%)	2%
Ethnicity			
Hispanic	18 (10%)	176 (3%)	4%

[a]State of Alabama 2010 Census of Population

Table 2.

Population cohort variants and rate of corroborating personal or family history

Disease	Gene[a]	Number of variants identified	Corroborated history # (%)[b]
Hereditary breast and ovarian cancer	*BRCA1*	9	7 (78%)
	BRCA2	11	6 (55%)
Lynch syndrome	*MLH1*	3	1 (33%)
	MSH2	1	1 (100%)
	MSH6	3	0 (0%)
	PMS2	2	2 (100%)
MYH-associated polyposis	*MUTYH*	5 (4 heterozygous, 1 homozygous)	0 (%)
Multiple endocrine neoplasia type 2, familial medullary thyroid cancer	*RET*	2	1 (50%)
Hereditary paraganglioma-pheochromocytoma syndrome	*SDHB*	1	0 (0%)
Hypertrophic cardiomyopathy, dilated cardiomyopathy	*MYBPC3*	9	2 (22%)
	MYH7	5	1 (20%)
	GLA	2	0 (0%)
Arrhythmogenic right ventricular cardiomyopathy	*PKP2*	3	0 (0%)
Romano-Ward long-QT syndrome types 1, 2, and 3, Brugada syndrome	*KCNQ1*	1	0 (0%)
	KCNH2	2	0 (0%)
	SCN5A	2	1 (50%)
Familial hypercholesterolemia	*LDLR*	3	2 (67%)
	APOB	6	5 (85%)
Malignant hyperthermia	*RYR1*	9	1 (11%)

[a] Reportable variation in the following genes have not yet been identified in any AGHI population cohort participants to date: *TP53, STK11, APC, BMPR1A, SMAD4, VHL, MEN1, PTEN, RB1, SDHD, SDHAF2, SDHC, TSC1, TSC2, WT1, NF2, COL3A1, FBN1, TGFBR1, TGFBR2, SMAD3, ACTA2, MYH11, TNNT2, TNNI3, TPM1, MYL3, ACTC1, PRKAG2, MYL2, LMNA, RYR2, DSP, DSC2, TMEM43, DSG2, PCSK9, ATP7B, OTC, CACNA1S*.

[b] Corroborated history defined as having a relevant reported personal or family history that was flagged by AGHI criteria.