

# Ethical considerations when co-analyzing ancient DNA and data from private genetic databases

Éadaoin Harney,<sup>1,2,\*</sup> Kendra Sirak,<sup>2,3</sup> Jakob Sedig,<sup>2,3</sup> Steven Micheletti,<sup>1</sup> Roslyn Curry,<sup>1,2</sup> Samantha Ancona Esselmann,<sup>1</sup> and David Reich<sup>2,3,4,5</sup>

## Summary

Ancient DNA studies have begun to explore the possibility of identifying identical DNA segments shared between historical and living people. This research requires access to large genetic datasets to maximize the likelihood of identifying previously unknown, close genetic connections. Direct-to-consumer genetic testing companies, such as 23andMe, Inc., manage by far the largest and most diverse genetic databases that can be used for this purpose. It is therefore important to think carefully about guidelines for carrying out collaborations between researchers and such companies. Such collaborations require consideration of ethical issues, including policies for sharing ancient DNA datasets, and ensuring reproducibility of research findings when access to privately controlled genetic datasets is limited. At the same time, they introduce unique possibilities for returning results to the research participants whose data are analyzed, including those who are identified as close genetic relatives of historical individuals, thereby enabling ancient DNA research to contribute to the restoration of information about ancestral connections that were lost over time, which can be particularly meaningful for families and groups where such history has not been well documented. We explore these issues by describing our experience designing and carrying out a study searching for genetic connections between 18th- and 19th-century enslaved and free African Americans who labored at Catoctin Furnace, Maryland, and 23andMe research participants. We share our experience in the hope of helping future researchers navigate similar ethical considerations, recognizing that our perspective is part of a larger conversation about best ethical practices.

We recently identified genetic relationships between 18th- and 19th-century African Americans from Catoctin Furnace, Maryland, and over forty thousand consenting research participants in 23andMe, Inc.'s genetic database.<sup>1</sup> These connections include distant relatives and possible direct descendants of the Catoctin individuals. This represents one of the first studies, to our knowledge, where identical-by-descent (IBD) segments of the genome (segments of DNA that are shared by two or more people because they have been inherited from a recent common ancestor) have been identified between historical individuals and their previously unknown living relatives by leveraging a dataset held by a private company.<sup>2</sup> In breaking this new ground, this study encountered ethical issues that had not been addressed in the existing literature on ethics in ancient DNA (aDNA),<sup>3–10</sup> although related concerns surrounding the rise of direct-to-consumer genetic ancestry testing have been discussed broadly in the social sciences literature.<sup>11–14</sup> Here, we discuss the approach we took in response to these issues with the hope that our experiences help inform future work by researchers in the field of aDNA, genetic ancestry companies interested in contributing to research collaborations or creating product features involving aDNA data, and anthropologists and curators who serve as stewards of human remains.

The genetic analyses of the remains that are reported in Harney et al. were carried out at the request of the Catoctin Furnace Historical Society (CFHS) (led by current president Elizabeth Comer), which advocated for further study of the Catoctin remains in partnership with the African American Resources and Cultural Heritage (AARCH) Society of Frederick, Maryland, later joined by individuals identified by researchers at CFHS as descendants of two African Americans who labored at Catoctin Furnace. The remains of the Catoctin individuals have been under the stewardship of the Smithsonian Institution since 1980 following their excavation from the African American Cemetery at Catoctin Furnace in advance of planned highway construction. Preliminary anthropological analyses of the historical Catoctin individuals were performed following their excavation to determine the identity of those buried in the undocumented cemetery,<sup>15,16</sup> but these analyses were limited in scope. They did not address all of the key goals of the CFHS and AARCH, including to “identify a descendant community for the Catoctin African American workers (enslaved and free); to connect the individuals within the cemetery to their ancestral roots in Africa; and to share the discovery process and its results with the public.”<sup>17</sup>

In 2016, CFHS received a Maryland Heritage Areas Authority grant to undertake a forensic re-analysis of the

<sup>1</sup>23andMe, Inc., Sunnyvale, CA 94043, USA; <sup>2</sup>Department of Human Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA; <sup>3</sup>Department of Genetics, Harvard Medical School, Boston, MA 02115, USA; <sup>4</sup>Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02115, USA; <sup>5</sup>Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

\*Correspondence: [eadaoinh@23andme.com](mailto:eadaoinh@23andme.com)  
<https://doi.org/10.1016/j.ajhg.2023.06.011>

© 2023 The Authors. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Catoctin individuals,<sup>18</sup> initiating a new investigation into Catoctin using a variety of anthropological and biomolecular tools. This work has contributed to community and national-level conversations about the role that scientific approaches can play in restoring information about the lives of enslaved people that would otherwise be lost to history.<sup>19,20</sup> In its first phase, researchers at the Smithsonian Institution partnered with the CFHS to study the demography and life history of the Catoctin individuals using osteological and isotopic approaches.<sup>21</sup>

In the next phase of the project, this research team sought to investigate the ancestral origins, familial relationships, and legacy of the Catoctin individuals through genetic analysis. Members of the Reich Laboratory at Harvard University were invited to join the study team based on their expertise in aDNA recovery and population genetic analyses. Sampling for aDNA was carried out with the approval of the Anthropology Collections Advisory Committee at the Smithsonian Institution, and the sequenced aDNA were reported in a technical note<sup>22</sup> and made fully publicly available following open science best practices<sup>23</sup> and the Smithsonian Institution's curatorial requirements.<sup>6</sup> However, without access to a diverse genetic database, it would have been impossible to effectively search for a Catoctin-descendant community using genetic approaches. Access to a database that was larger and more diverse than any that is publicly available provided an extremely powerful and otherwise unattainable tool<sup>13</sup> for addressing the request of the CFHS to search for a descendant community because the number of connections that can be found between past and present-day people is directly proportional to the number of people in the database and also dependent on the diversity of people in that database. The research team identified 23andMe as an appropriate partner due to the size of their database, which contains genome-wide data from ~12.8 million people (more than an order-of-magnitude more individuals than any publicly available dataset),<sup>24</sup> the high diversity of people represented in the 23andMe dataset, 23andMe's experience developing population genetic tools to identify genetic connections through IBD analysis,<sup>25–27</sup> and 23andMe's track record of publishing open-access research by leveraging its database.<sup>28,29</sup> Critically, over 80% of 23andMe customers have provided consent for their de-identified survey answers and genetic data to be used for research purposes and can be recontacted based on research findings for a variety of purposes, including to answer follow-up questions or to be invited to participate in additional, in-depth studies.<sup>12,30</sup> This also means that, in theory, it could be possible to return results to descendants identified over the course of the study (although how best to do this, if at all, is still under consideration, and no recontacting based on the results of Harney et al. has yet occurred). Therefore, with support from CFHS and AARCH, a partnership was explored between 23andMe and the research team in which ethics remained at the forefront of discussions.

We addressed these ethical considerations in a robust way that respects core academic principles through a formal written agreement between Harvard University and the Howard Hughes Medical Institute (HHMI), which were responsible for generating and sharing the Catoctin dataset, and 23andMe, which was responsible for comparing the Catoctin data with the 23andMe database. During the course of the study design and contract negotiations, we identified three key areas that warranted special consideration: data sharing, study reproducibility, and return of results.

## Sharing of aDNA datasets

To date, the vast majority of aDNA data have been made fully publicly available in accordance with the principles of open science.<sup>23,31</sup> However, for datasets that are closely associated with living descendants and other stakeholder communities, some researchers have advocated for restricting access.<sup>3,4,32,33</sup> Restrictions on data access that have been applied have included that they can only be used for non-commercial purposes, or in some cases, only for replication purposes, and that explicit permission from community stakeholders may be required before sharing data for additional research purposes (e.g., Fleskes et al.,<sup>34</sup> Molthe et al.,<sup>35</sup> and Severson et al.<sup>36</sup>).

While the best data-sharing policies for aDNA data may be context specific, all authors agree that a scenario in which a for-profit company—or any research or commercial entity that may benefit financially, either directly (e.g., by charging fees to access research results) or indirectly (e.g., career advancement), by restricting access to data—retains exclusive, unrestricted access to an aDNA dataset following study completion would be inappropriate. Instead, clear access guidelines should be created that apply to all parties who would like to access the data. In some for-profit companies, like 23andMe, there is a distinction between that organization's research and their commercial activity. Thus, even if data are available for non-commercial purposes only, this need not prevent research-focused collaborations with for-profit organizations. Regardless of the data-sharing policy that is implemented, it is critical to ensure that the details and implications of this policy are thoroughly communicated to all stakeholders involved in the study, as was done as a part of Harney et al.

The collaboration agreement between Harvard, HHMI, and 23andMe ensures that any non-public aDNA dataset shared with 23andMe can only be used for research and publication purposes. The data can only be used by 23andMe for other purposes after it is made fully publicly available. In the case of the Catoctin project, the terms of the sampling agreement set forth by the Smithsonian Institution required the Catoctin aDNA dataset to be made fully publicly available upon publication, or within three years of sampling. To meet the terms of this

agreement, we released a technical report describing the Catoctin data in June 2022,<sup>22</sup> over a year prior to publication of the joint analyses of the historical genomes and 23andMe research participant data in Harney et al.

### Challenges in ensuring reproducibility in studies involving private genetic databases

The 23andMe genetic database, like many other private genetic databases, is subject to restrictions on individual-level data sharing that make publishing fully reproducible studies challenging.<sup>30</sup> Some areas of population genetics, particularly genome-wide association studies, have developed standards for sharing summary statistics that do not contain any individual-level information, thereby respecting the privacy restrictions required of these private genetic databases while also enabling some degree of reproducibility (e.g., Buniello et al.<sup>37</sup> and Shi et al.<sup>38</sup>). However, for IBD-based studies that require direct comparisons between individual genomes, there is no corresponding approach that would enable a similar level of reproducibility of study results through the release of summary statistics. There is no perfect solution for this problem, but weighed against this is the fact that leveraging a genetic database as large as 23andMe's aDNA powers research that otherwise would be impossible. In fact, it is possible that 23andMe's high research participation rate can be attributed in part to the strong privacy protections that are offered to research participants, including the ability to withdraw consent at any time and restrictions on individual-level data sharing without additional explicit consent.

Through a series of discussions between researchers and legal advisors at Harvard, HHMI, 23andMe, and later editors at the journal where Harney et al. was published, we developed a two-phase approach to maximize the ability of researchers to independently replicate the results of our study: (1) we included genetic data drawn from several publicly available datasets, including the 1000 Genomes Project<sup>39</sup> and 23andMe's African American Imputation Panel,<sup>40</sup> in our study dataset. Whenever possible, we report the results of each analysis for these subsets of individuals. This approach makes it possible for researchers who wish to investigate the accuracy of the results to reproduce fully a portion of the study using these publicly available datasets. (2) For all analyses that cannot be fully reproduced using the above method, 23andMe committed to rerun comparisons upon request by academic and non-profit researchers on reasonable terms for a period of up to 7 years from the date of publication or for as long as the study co-authors are employed at 23andMe, whichever is shorter. This commitment is time limited in recognition that changes to data storage formats and technology (and the roles held by researchers) over time mean that no studies are likely to be fully reproducible indefinitely. This is similar to the common academic institutional requirement

that laboratory notebooks be retained to support reproducible research for a number of years, but not indefinitely.<sup>41</sup>

### Return of results to community stakeholders and research participants

In recognition of the importance of involving community stakeholders in the research process, the formal collaboration agreement enabled sharing of preliminary 23andMe research results not only with the external research partners at Harvard and HHMI who signed the agreement but also with other research partners, including those at the Smithsonian, and members of relevant stakeholder communities, including CFHS, AARCH, and previously identified descendants of the Catoctin individuals.

We also recognize that 23andMe's re-contactable research cohort presents an opportunity to explore options to return results not only to stakeholder communities with known ties to the Catoctin individuals but also to the research participants whose genetic information was used in the study—an option that is not typically available for studies performed using publicly available datasets. Although the research consent process makes it clear to 23andMe research participants that they should not expect to have individual-level results returned to them as part of the research process, it leaves open the possibility that research could aid in the development of future 23andMe features through which results could be returned.<sup>30</sup>

### Ethical considerations for future feature development by genetic ancestry companies

Although the aims of Harney et al. were purely research focused, over the course of the study, we considered how a feature or report in the 23andMe user interface could be used to help realize one of the original goals of CFHS and AARCH: to identify and foster a Catoctin-descendant community. In what follows, we discuss some issues that we believe would be critical for genetic ancestry companies to address when designing a product feature that reports IBD connections shared between customers and ancient or historical individuals.

#### *Providing opportunities to opt in to receive results*

Although use of re-contactable research cohorts in studies like Harney et al. provide a unique opportunity to return results to research participants, participants may not wish to learn about these connections for a variety of reasons. For instance, in the case of Catoctin, learning about genetic connections to a direct ancestor or distant relative who was enslaved could be a painful experience that evokes known family trauma or reveals an unknown personal connection to slavery in the United States. In the case of Harney et al., while community stakeholders expressed an interest in learning about any potential genetic connections they might share with the Catoctin individuals, they were also clear that they would like to be given

the opportunity to choose to receive this information rather than have it shared with them without warning. When returning results to research participants or other customers in their database, genetic ancestry companies should give customers the opportunity to actively opt in to receive this information and provide adequate information to help customers anticipate what they might learn by doing so. In making this recommendation, we acknowledge that users who opt not to receive this information from genetic ancestry companies directly may still learn about their likely genetic connections to historical individuals via relatives who do choose to opt in or explore their connections using other means. Our recommendation to include an opt-in step before returning results to customers mirrors the consent that customers at 23andMe and other genetic ancestry companies give before receiving information about living genetic relatives.<sup>42</sup>

#### ***Avoiding biological gatekeeping to community membership***

While identifying a Catoctin-descendant community using a genetic approach was one of the goals of Harney et al., genetic connections are only one of a variety of ways through which stakeholders can be connected to historical communities and may not even be the most important form of connection. Other forms of connection include non-genetic familial relationships through shared culture or heritage or through a collective kinship bond that does not depend on recent shared ancestry. For instance, key community stakeholders associated with the Catoctin individuals self-identify as members of a “collective” descendant community that was intentionally created by the CFHS and AARCH to honor and protect the Catoctin individuals.<sup>1</sup> Although this collective descent community was created because there was no other known community that claimed to have a connection to the Catoctin individuals, any newly identified descendant community (connected via genetics, genealogy, or any other means) would not diminish the importance of this established community.

When returning results to research participants who share a genetic connection to a historical population, it could be tempting for genetic ancestry companies to emphasize these genetic connections over the many other ways through which stakeholders can be connected to a historical group. However, when returning results to research participants or other customers, genetic ancestry companies should be explicit that genetic connections are not necessarily more important than other types of connections. Research on genetic connections to historical individuals has the potential to contribute to the biologization of identity, whether intentionally or not. Thus, when returning results, genetic ancestry companies should take care to identify and describe the presence of existing stakeholder communities. In the case of Catoctin, existing stakeholders have been explicit that they are eager to engage with genetic relatives of the Catoctin individuals. In other cases, genetic relatives who lack any other

connection to a historical group may not be welcomed by the existing stakeholder community on the basis of their genetic results alone.

#### ***Prioritizing accuracy when returning results***

When designing product features that reveal genetic connections to historical people, it is critical to use approaches that are optimized to work on aDNA data. Although many population genetic tools are capable of generating output when applied to aDNA data, it should not be assumed that these results are meaningful unless the performance of those tools on aDNA data has been rigorously assessed—paying particular attention to the impact of missing data, C-to-T misincorporations, low-confidence genotype calls, and, when applicable, imputation errors.<sup>43</sup> Approaches that do not account for the unique characteristics of aDNA data have the potential to be highly inaccurate, thereby failing to identify true genetic connections (“false negatives”) and incorrectly identifying genetic connections in cases where none exist (“false positives”). It is therefore critical to identify optimal approaches for analyzing aDNA data and minimum quality standards for the type of aDNA data that can be analyzed. For instance, in the Catoctin study, we optimized an approach for identifying IBD connections that can be confidently applied to aDNA data with over 1× coverage. We recommended caution when interpreting the results of data with less than 1× coverage and did not report results for data with less than 0.5× coverage.

In many cases, the identification of a genetic connection to a historical individual may impact customers in profound ways—particularly in cases where customers may have sought out genetic ancestry testing in order to fill gaps in their family and community history. For instance, many descendants of enslaved Africans in the United States have turned to genetic ancestry testing to learn about their family history.<sup>11,44–46</sup> Discovering a shared genetic connection to a historical African American from Catoctin or another site could therefore be extremely impactful, while returning inaccurate or misleading results has the potential to cause confusion and other harm.<sup>14,46</sup> We therefore believe that methods for comparing aDNA with customers should be rigorous to avoid false positives, and if an analysis identifies only weak affiliations—for example, due to sharing a common ancestor hundreds or thousands of years ago—genetic ancestry companies should make this clear and not overstate the likelihood (or in the case of more ancient individuals, the uniqueness) of direct genealogical connections. A methodological tradeoff associated with reducing the false-positive rate for any well-optimized analysis is an increase in the frequency of false negatives; therefore, genetic ancestry companies should also be sure to communicate that the lack of IBD sharing cannot be interpreted as conclusive evidence that there is no shared genetic relationship.

### **Taking responsibility for returning results in a comprehensible way**

Genetic ancestry companies should also recognize that returning inaccurate results or returning results in a confusing way may actively harm community stakeholders, people or organizations who serve as stewards for the human remains, those who serve in other roles that involve educating the public about specific historical individuals or sites, or broader public understanding. In the case of Catoctin, while groups like CFHS, AARCH, and researchers at the Smithsonian Institution are committed to helping potential descendants and other stakeholders learn about Catoctin, they should not have an added responsibility of helping customers of genetic ancestry companies interpret their genetic results. Reports generated by genetic ancestry companies should be created with enough educational content and detail so that customers with shared genetic connections can interpret their results without the need for guidance from outside organizations. In fact, genetic ancestry companies may wish to consult with members of these organizations to ensure that any product features that involve these historical populations are represented accurately and clearly. But in doing so, they should again be careful not to place the burden of responsibility for providing comprehensible results on these organizations.

### **Conclusion**

We believe that joint analysis of historical genomes and data in genetic databases controlled by for-profit genetic ancestry companies represents an exciting new direction for aDNA research. As aDNA research continues to leverage the power of ever-growing datasets, all parties involved in this work must make a concerted effort to predict ethical issues that may arise and proactively design an ethically sound strategy to ensure that sensitive genetic data remains protected. The continued application of aDNA technology to new contexts is inevitable, and we are encouraged by the robust conversation about best ethical practices that provided a foundation for this work.<sup>3–14</sup> The present paper aims to contribute to this body of literature by highlighting the specific issues we identified as ethical challenges in carrying out Harney et al. in the hope that it might help future researchers navigate similar ethical considerations in projects like this one going forward. Future work will certainly add to and amend the guidance we provide, which is part of moving forward the ethics of the rapidly growing field of aDNA research.

### **Acknowledgments**

We thank Kathryn G. Barca, Karin S. Bruwelheide, Arnab Chowdry, Elizabeth Comer, William A. Freyman, David Hinds, Fatimah Jackson, Joanna Mountain, Douglas W. Owsley, and Joyce Tung for providing critical comments on this and earlier versions of the manuscript. We would like to thank the research participants and employees of 23andMe for making this work possible. David

Reich acknowledges support from the Howard Hughes Medical Institute (HHMI), the John Templeton Foundation (grant number 61220), and NIH grant GM100233. This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

23andMe research participants provided informed consent and volunteered to participate in the research online under a protocol approved by the external AAHRPP-accredited IRB, Ethical & Independent (E&I) Review Services. As of 2022, E&I Review Services is part of Salus IRB (<https://www.versiticlinicaltrials.org/salusirb>).

### **Declaration of interests**

E.H., S.M., and S.A.E. are employees and shareholders of 23andMe, Inc. During the summer of 2021, R.C. was employed as an intern at 23andMe, Inc.

### **References**

1. Harney, É., Micheletti, S., Bruwelheide, K.S., Freyman, W.A., Bryc, K., Akbari, A., Jewett, E., Comer, E., Gates Jr, H.L., Heywood, L., et al. (2023). The Genetic Legacy of African Americans from Catoctin Furnace. *Science* 381, 6657. <https://doi.org/10.1126/science.ade4995>.
2. Begg, T.J.A., Schmidt, A., Kocher, A., Larmuseau, M.H.D., Runfeldt, G., Maier, P.A., Wilson, J.D., Barquera, R., Maj, C., Szolek, A., et al. (2023). Genomic analyses of hair from Ludwig van Beethoven. *Curr. Biol.* 33, 1431–1447.e22. <https://doi.org/10.1016/j.cub.2023.02.041>.
3. Wagner, J.K., Colwell, C., Claw, K.G., Stone, A.C., Bolnick, D.A., Hawks, J., Brothers, K.B., and Garrison, N.A. (2020). Fostering Responsible Research on Ancient DNA. *Am. J. Hum. Genet.* 107, 183–195. <https://doi.org/10.1016/j.ajhg.2020.06.017>.
4. Kowal, E., Weyrich, L.S., Argüelles, J.M., Bader, A.C., Colwell, C., Cortez, A.D., Davis, J.L., Figueiro, G., Fox, K., Malhi, R.S., et al. (2023). Community partnerships are fundamental to ethical ancient DNA research. *HGG Adv.* 4, 100161. <https://doi.org/10.1016/j.xhgg.2022.100161>.
5. Alpaslan-Roodenberg, S., Anthony, D., Babiker, H., Bánffy, E., Booth, T., Capone, P., Deshpande-Mukherjee, A., Eisenmann, S., Fehren-Schmitz, L., Frachetti, M., et al. (2021). Ethics of DNA research on human remains: five globally applicable guidelines. *Nature* 599, 41–46. <https://doi.org/10.1038/s41586-021-04008-x>.
6. Austin, R.M., Sholts, S.B., Williams, L., Kistler, L., and Hofman, C.A. (2019). Opinion: To curate the molecular past, museums need a carefully considered set of best practices. *Proc. Natl. Acad. Sci. USA* 116, 1471–1474. <https://doi.org/10.1073/pnas.1822038116>.
7. Bardill, J., Bader, A.C., Garrison, N.A., Bolnick, D.A., Raff, J.A., Walker, A., Malhi, R.S.; and Summer internship for Indigenous peoples in Genomics SING Consortium (2018). internship for In. peoples in G. (SING) (2018). Advancing the ethics of paleogenomics. *Science* 360, 384–385. <https://doi.org/10.1126/science.aaq1131>.

8. Claw, K.G., Anderson, M.Z., Begay, R.L., Tsosie, K.S., Fox, K., Garrison, N.A.; and Summer internship for Indigenous peoples in Genomics SING Consortium (2018). A framework for enhancing ethical genomic research with Indigenous communities. *Nat. Commun.* 9, 2957. <https://doi.org/10.1038/s41467-018-05188-3>.
9. Gibbon, V.E. (2020). African ancient DNA research requires robust ethics and permission protocols. *Nat. Rev. Genet.* 21, 645–647. <https://doi.org/10.1038/s41576-020-00285-w>.
10. Prendergast, M.E., and Sawchuk, E. (2018). Boots on the ground in Africa's ancient DNA 'revolution': archaeological perspectives on ethics and best practices. *Antiquity* 92, 803–815. <https://doi.org/10.15184/aqy.2018.70>.
11. Abel, S., and Schroeder, H. (2020). From Country Marks to DNA Markers: The Genomic Turn in the Reconstruction of African Identities. *Curr. Anthropol.* 61, S198–S209. <https://doi.org/10.1086/709550>.
12. Harris, A., Wyatt, S., and Kelly, S.E. (2013). The Gift of Spit (and the Obligation to Return It). *Inf. Commun. Soc.* 16, 236–257. <https://doi.org/10.1080/1369118X.2012.701656>.
13. Sharon, T. (2016). The Googlization of health research: from disruptive innovation to disruptive ethics. *Per. Med.* 13, 563–574. <https://doi.org/10.2217/pme-2016-0057>.
14. Nelson, S.C., Bowen, D.J., and Fullerton, S.M. (2019). Third-Party Genetic Interpretation Tools: A Mixed-Methods Study of Consumer Motivation and Behavior. *Am. J. Hum. Genet.* 105, 122–131. <https://doi.org/10.1016/j.ajhg.2019.05.014>.
15. Kelley, J.O., and Angel, J.L. (1983). The workers of Catoctin furnace. *Maryland Archeology* 19, 2–17.
16. Angel, J.L., and Kelley, J.O. (1983). Health-status of colonial iron-worker slaves. *Am. J. Phys. Anthropol.* 60, 170–171.
17. Hyman, E.B. (2020). Catoctin Slave Speak (Catoctin Furnace Historical Society, Inc.).
18. Comer, E.A. (2016). Catoctin Furnace: Academic Research Informing Heritage Tourism. In Conference Presentation. 49th Annual Conference on Historical and Underwater Archaeology.
19. Lepore, J. (2021). When Black History Is Unearthed, Who Gets to Speak for the Dead? *The New Yorker*. <https://www.newyorker.com/magazine/2021/10/04/when-black-history-is-unearthed-who-gets-to-speak-for-the-dead>.
20. Anderson, E. (2015). Buried Secrets: Study of Skeletal Remains May Hold Key to Slave History (The Frederick News-Post). [https://www.fredericknews-post.com/news/human\\_interest/buried-secrets-study-of-skeletal-remains-may-hold-key-to-slave-history/article\\_587f3e81-55c3-52ef-b2e0-96ab11990f76.html](https://www.fredericknews-post.com/news/human_interest/buried-secrets-study-of-skeletal-remains-may-hold-key-to-slave-history/article_587f3e81-55c3-52ef-b2e0-96ab11990f76.html).
21. Bruwelheide, K.S., Owsley, D.W., Barca, K.G., France, C.A.M., Little, N.C., and Comer, E.A. (2020). Restoring Identity to People and Place: Reanalysis of Human Skeletal Remains from a Cemetery at Catoctin Furnace, Maryland. *Hist Arch* 54, 110–137.
22. Harney, É., Olalde, I., Bruwelheide, K., Barca, K.G., Curry, R., Comer, E., Rohland, N., Owsley, D., and Reich, D.E. (2022). Technical Report on Ancient DNA Analysis of 27 African Americans from Catoctin Furnace. Preprint at bioRxiv. <https://doi.org/10.1101/2022.06.12.495320>.
23. Anagnostou, P., Capocasa, M., Milia, N., Sanna, E., Battaglia, C., Luzi, D., and Destro Bisol, G. (2015). When Data Sharing Gets Close to 100%: What Human Paleogenetics Can Teach the Open Science Movement. *PLoS One* 10, e0121409. <https://doi.org/10.1371/journal.pone.0121409>.
24. Heilbron, K., Mozaffari, S.V., Vacic, V., Yue, P., Wang, W., Shi, J., Jubb, A.M., Pitts, S.J., and Wang, X. (2021). Advancing drug discovery using the power of the human genome. *J. Pathol.* 254, 418–429. <https://doi.org/10.1002/path.5664>.
25. Freyman, W.A., McManus, K.F., Shringarpure, S.S., Jewett, E.M., Bryc, K., the 23andMe Research Team, and Auton, A. (2020). Fast and robust identity-by-descent inference with the templated positional Burrows-Wheeler transform. *Mol. Biol. Evol.* <https://doi.org/10.1093/molbev/msaa328>.
26. Jewett, E.M., McManus, K.F., Freyman, W.A., 23andMe Research Team, and Auton, A. (2021). Bonsai: An efficient method for inferring large human pedigrees from genotype data. *Am. J. Hum. Genet.* 108, 2052–2070. <https://doi.org/10.1016/j.ajhg.2021.09.013>.
27. Henn, B.M., Gignoux, C.R., Jobin, M., Granka, J.M., Macpherson, J.M., Kidd, J.M., Rodríguez-Botigué, L., Ramachandran, S., Hon, L., Brisbin, A., et al. (2011). Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proc. Natl. Acad. Sci. USA* 108, 5154–5162.
28. Bryc, K., Durand, E.Y., Macpherson, J.M., Reich, D., and Mountain, J.L. (2015). The Genetic Ancestry of African Americans, Latinos, and European Americans across the United States. *Am. J. Hum. Genet.* 96, 37–53. <https://doi.org/10.1016/j.ajhg.2014.11.010>.
29. Micheletti, S.J., Bryc, K., Ancona Esselmann, S.G., Freyman, W.A., Moreno, M.E., Poznik, G.D., Shastri, A.J., 23andMe Research Team, Beleza, S., Mountain, J.L., et al. (2020). Genetic Consequences of the Transatlantic Slave Trade in the Americas. *Am. J. Hum. Genet.* 107, 265–277. <https://doi.org/10.1016/j.ajhg.2020.06.012>.
30. 23andMe DNA Genetic Testing & Analysis - 23andMe. <https://www.23andme.com/about/consent/>.
31. Kaye, J., Heeney, C., Hawkins, N., de Vries, J., and Boddington, P. (2009). Data sharing in genomics — re-shaping scientific practice. *Nat. Rev. Genet.* 10, 331–335. <https://doi.org/10.1038/nrg2573>.
32. Hudson, M., Garrison, N.A., Sterling, R., Caron, N.R., Fox, K., Yracheta, J., Anderson, J., Wilcox, P., Arbour, L., Brown, A., et al. (2020). Rights, interests and expectations: Indigenous perspectives on unrestricted access to genomic data. *Nat. Rev. Genet.* 21, 377–384. <https://doi.org/10.1038/s41576-020-0228-x>.
33. Tsosie, K.S., Begay, R.L., Fox, K., and Garrison, N.A. (2020). Generations of genomes: advances in paleogenomics technology and engagement for Indigenous people of the Americas. *Curr. Opin. Genet. Dev.* 62, 91–96. <https://doi.org/10.1016/j.gde.2020.06.010>.
34. Fleskes, R.E., Cabana, G.S., Gilmore, J.K., Juarez, C., Karcher, E., Oubré, L., Mishoe, G., Ofunniyin, A.A., and Schurr, T.G. (2023). Community-engaged ancient DNA project reveals diverse origins of 18th-century African descendants in Charleston, South Carolina. *Proc. Natl. Acad. Sci. USA* 120, e2201620120. <https://doi.org/10.1073/pnas.2201620120>.
35. Moltke, I., Korneliussen, T.S., Seguin-Orlando, A., Moreno-Mayar, J.V., LaPointe, E., Billeck, W., and Willerslev, E. (2021). Identifying a living great-grandson of the Lakota Sioux leader Tatanka Iyotake (Sitting Bull). *Sci. Adv.* 7, eabh2013. <https://doi.org/10.1126/sciadv.abh2013>.
36. Severson, A.L., Byrd, B.F., Mallott, E.K., Owings, A.C., DeGiorgio, M., de Flamingh, A., Nijmeh, C., Arellano, M.V.,

- Leventhal, A., Rosenberg, N.A., and Malhi, R.S. (2022). Ancient and modern genomics of the Ohlone Indigenous population of California. *Proc. Natl. Acad. Sci. USA* *119*, e2111533119. <https://doi.org/10.1073/pnas.2111533119>.
37. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* *47*, D1005–D1012. <https://doi.org/10.1093/nar/gky1120>.
38. Shi, Z., Zhan, J., Wei, J., Ladson-Gary, S., Wang, C.-H., Hulick, P.J., Zheng, S.L., Cooney, K.A., Isaacs, W.B., Helfand, B.T., et al. (2022). Reliability of Ancestry-specific Prostate Cancer Genetic Risk Score in Four Racial and Ethnic Populations. *Eur. Urol. Open Sci.* *45*, 23–30. <https://doi.org/10.1016/j.euros.2022.09.001>.
39. 1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. *Nature* *526*, 68–74. <https://doi.org/10.1038/nature15393>.
40. O’Connell, J., Yun, T., Moreno, M., Li, H., Litterman, N., Kolesnikov, A., Noblin, E., Chang, P.-C., Shastri, A., Dorfman, E.H., et al. (2021). A population-specific reference panel for improved genotype imputation in African Americans. *Commun. Biol.* *4*, 1269. <https://doi.org/10.1038/s42003-021-02777-9>.
41. Burroughs Wellcome Fund, and Howard Hughes Medical Institute Making the Right Moves: A Practical Guide to Scientific Management for Postdocs and New Faculty Second Edition.
42. DNA Relatives: The Genetic Relative Basics 23andMe Customer Care. <https://customercare.23andme.com/hc/en-us/articles/115004659068-DNA-Relatives-The-Genetic-Relative-Basics>.
43. Orlando, L., Allaby, R., Skoglund, P., Der Sarkissian, C., Stockhammer, P.W., Ávila-Arcos, M.C., Fu, Q., Krause, J., Willerslev, E., Stone, A.C., and Warinner, C. (2021). Ancient DNA analysis. *Nat. Rev. Methods Primers* *1*, 14–26. <https://doi.org/10.1038/s43586-020-00011-0>.
44. David, L.T. (2023). Addressing the feasibility of people of African descent finding living African relatives using direct-to-consumer genetic testing. *Am. J. Biol. Anthropol.* *181*, 163–165. <https://doi.org/10.1002/ajpa.24705>.
45. Nelson, A. (2016). *The Social Life of DNA: Race, Reparations, and Reconciliation after the Genome* (Beacon Press).
46. Nelson, A. (2008). Bio Science: Genetic Genealogy Testing and the Pursuit of African Ancestry. *Soc. Stud. Sci.* *38*, 759–783. <https://doi.org/10.1177/0306312708091929>.