POLICY FORUM



ILLUSTRATION: MICROSTOCKHUB/ISTOCK.COM

Downloaded from http://science.sciencemag.org/ on February 4, 2021

HUMAN GENOME ANNIVERSARY

Complicated legacies: The human genome at 20

Millions of people today have access to their personal genomic information. Direct-to-consumer services and integration with other "big data" increasingly commoditize what was rightly celebrated as a singular achievement in February 2001 when the first draft human genomes were published. But such remarkable technical and scientific progress has not been without its share of missteps and growing pains. *Science* invited the experts below to help explore how we got here and where we should (or ought not) be going. **—Brad Wible**

An ethos of rapid data sharing, more relevant than ever

By **Kathryn Maxson Jones**[1,2] *and* **Robert Cook-Deegan**[3]

Sharing data can save lives. The "Bermuda Principles" for public data disclosure are a fundamental legacy of producing the first human reference DNA sequence during the Human Genome Project (HGP) (*1*). Since the 1990s, these principles have become a touchstone for open science.

In February 1996, the leaders of the HGP gathered in Bermuda to discuss how to scale up production for a human reference DNA sequence. With some caveats, the consortium agreed that all sequencing centers would release their data online within 24 hours. Other examples of sharing data before publication existed, but most—such as the Protein Data Bank—restricted sharing of prepublication data to a small community of users, sometimes withholding data even after the related papers were published (*2*). At the time, the Bermuda Principles were distinctive in their aspiration that all HGP-funded sequences be released to anyone online within a day. Yet implementing this policy was hardly simple; the challenges that the HGP faced inform data sharing today (*3*).

The Bermuda Principles required advocacy. This came from John Sulston and Robert Waterston, whose experiences with data sharing in *Caenorhabditis elegans* biology were the practical precedent for a radical idea. Context also mattered, and data release within 24 hours remained an aspirational ethos rather than a strict requirement. Flexibility allowed smaller centers to participate while also allowing the project to accommodate then-incompatible policies in Germany, France, Japan, and the United States.

Published by AAAS

Finally, the policy required enforcement. Administrators from the HGP's largest patrons sent stern letters intended to make funders' policies conform to the Bermuda Principles, threatening expulsion from the international sequencing consortium.

The Bermuda Principles have since been adapted to different communities and have served as an inspiration for many others (*4*). For example, rapid data sharing has been crucial in the current coronavirus crisis. The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) genome was identified quickly and its sequence released on 10 January 2020, starting the clock on the development of vaccines and diagnostic tests. The COVID-19 Host Genetics Initiative disseminated data rapidly and openly, building on precedents such as the Global Initiative on Sharing All Influenza Data (*5*, *6*).

Of course, unfettered data sharing is not, and should not be, universal. Identifiable individual medical data, for instance, cannot be treated the same way as samples contributed to build a reference genome sequence. Many communities have adopted prepublication sharing strategies with considerable success, such as the various consortia for Alzheimer's research, the "open science" experiments at the Montreal Neurological Institute and the Mario Negri Institute, and the advances enabled by the Structural Genomics Consortium.

The HGP set a high bar. Its core values of open science and rapid data flow persist, fomented by the urgency of rapid data sharing in biomedicine.

Lack of diversity hinders the promise of genome science

By **Charles N. Rotimi[4], Shawneequa L. Callier[4,5], Amy R. Bentley[4]**

The long-term global impact of human genomics will be compromised, and our understanding of human history and biology hindered, if we continue to focus predominantly on individuals of European ancestry (*7*). Although we all share a recent common origin in Africa, and the genetic difference between any two individuals is small (0.1%), this translates to about 3 million points where individual genomes can vary, and the distribution of these human genetic variants (HGVs) is not random. It has long been understood that genomes (and exposures to key nongenetic factors) differ across ancestral and geographical backgrounds; nonetheless, genomics has largely focused on European-ancestry genomes. Presumably this is attributable to the availability of large, well-characterized datasets of European-ancestry individuals, academic and research networks that exclude and disadvantage underrepresented scholars (*8*), and the absence of publishing or funding motivation for large-scale genomics of diverse individuals. But diversity and representation are now being elevated from the purview of specialized research to a broad awareness across genomics.

As this awareness develops, the field must grapple with understanding and communicating the implications: (i) Any two sub-Saharan Africans are more likely to be genetically different from each other than from an individual of European or Asian ancestry; (ii) a subset of HGVs can only be found in Africans because the small number of humans that left Africa about 100,000 years ago to populate the rest of the world carried a fraction of the variation that existed then; (iii) the African ecological environment has left its mark on human genomes (e.g., gene variants found to increase vulnerability to kidney failure) that are seen worldwide only in persons with ancestry from specific regions of Africa (*9*).

Similarly, there are HGVs of health and historical importance that are rare or absent in African populations. For example, genomic regions harboring ancient DNA—the result of interbreeding with archaic human relatives (such as Neanderthals) in Asia, Europe, and the Americas—have biological functions, such as susceptibility to diabetes and viruses (*10*). For genomics-driven technologies and clinical and public health approaches to be deployed globally without exacerbating health inequalities, we must include individuals from diverse ancestral and geographical backgrounds.

Growing prioritization of diverse populations in genomics research has begun to respond to these gaps. Programs such as TOPMed, All of Us, International Common Disease Alliance, Human Heredity and Health in Africa (H3Africa), Million Veteran Program, GenomeAsia, and the COVID global consortium contribute to advances in diversity and inclusion among research participants. The diversity of genomics researchers also merits continuing attention. The H3Africa initiative, for example, includes investments in training and infrastructure in each project, providing a blueprint for prioritizing capacity-building. The genomics community needs to value diverse samples in analyses and conclusions, as well as to focus resources on capacity-building and removing barriers to create a diverse workforce (*11*).

Algorithmic biology unleashed

By **Hallam Stevens[6]**

Over a few frenzied weeks in the middle of 2000, icing his wrists between coding sessions, Jim Kent, a graduate student at the University of California, Santa Cruz, created the first genome assembler software. GigAssembler pieced together the millions of fragments of DNA sequence generated at labs around the globe, literally making the human genome. At almost the same time, Celera Genomics acquired Paracel, a company that primarily designed software for intelligence gathering. Paracel owned specially designed text-matching hardware and software (the TRW Fast Data Finder) that was rapidly adapted for sniffing out genes within the vast spaces of the genome.

Untangling the jumble of genomic letters required rapidly and accurately searching for a specified sequence within a very large space. This demanded new forms of training and disciplinary expertise. Physicists, mathematicians, and computer scientists brought methods such as linear programming, hashing, and hidden Markov models into biology. Since 2005, the Moore's Law–like growth of next-generation sequencing has generated ever-increasing troves of data and required even faster algorithms for indexing and searching. Biology has borrowed "big data" methods from industry (e.g., Hadoop) but has also contributed to pushing the frontiers of computer science research (e.g., the Burrows-Wheeler transform) (*12*).

The coalescence of bioinformatics and computational biology around algorithms has also given rise to new institutional forms and new markets for biomedicine. Statistically powered "data-driven biology" has configured an emerging medical-industrial complex that promises personalized and "precision" forms of diagnosis and treatment. Algorithmic pipelines that compare an individual's genotype to reference data generate a range of predictions about future health and risk. Direct-to-consumer genomics companies such as 23andMe now promise us healthier, happier, and longer ways of living via algorithms.

This presents substantial challenges for privacy, data ownership, and algorithmic bias (*13*–*15*) that must be addressed if genomics is to avoid becoming a handmaiden of "surveillance capitalism" (*16*). Many tech companies have begun to look toward

Downloaded from http://science.sciencemag.org/ on February 4, 2021

Published by AAAS

using machine learning to combine more and more biological data with other forms of personal data—where we go, what we buy, whom we associate with, what we like. The hopes for genomics have long been tempered by fears that the genome could reveal too much about ourselves, exposing us to new forms of discrimination, social division, or control. Algorithmic biology is depicting and predicting our bodies with growing accuracy, but it is also drawing biomedicine more closely into the orbits of corporate tech giants that are aggregating and attempting to monetize data.

Value and affordability in precision medicine

By **Kathryn A. Phillips[7,8,9], Jeroen P. Jansen[7,9], Christopher F. Weyant[7]**

Debates about precision medicine (PM), which uses genetic information to target interventions, commonly focus on whether we can "afford" PM (*17*), but focusing only on affordability, not also value, risks rejecting technologies that might make health care more efficient. Affordability is a question of whether we can pay for an intervention given its impact on budgets, whereas value can be measured by the health outcomes achieved per dollar spent for an intervention. Ideally, a PM intervention both saves money and improves outcomes; however, most health care interventions produce better outcomes at higher cost, and PM is no exception. By better distinguishing affordability and value, and by considering how we can address both, we can further the agenda of achieving affordable and valuable PM.

The literature has generally not shown that PM is unaffordable or of low value; however, it has also not shown that PM is a panacea for reducing health care expenditures or always results in high-value care (*17*). Understanding PM affordability and value requires evidence on total costs and outcomes as well as potential cost offsets, but these data are difficult to capture because costs often occur up front while beneficial outcomes accrue over time (*18*). Also, PM could result in substantial downstream implications because of follow-up interventions, not only for patients but also for family members who may have inherited the same genetic condition. Emerging PM tests could be used for screening large populations and could include genome sequencing of all newborns, liquid biopsy testing to screen for cancers in routine primary care visits, and predictive testing for Alzheimer's disease in adults. These interventions may provide large benefits, but they are likely to require large up-front expenditures. Another complication is that many PM interventions measure multiple genes relevant to multiple conditions and provide myriad types of value, such as the personal value of this information to patients (*19*).

Various methods have been developed for integrating affordability and value, but cost-effectiveness analyses often do not examine the budget impact, which can result in incomplete or contradictory conclusions (*20*). However, assessments that consider affordability and value simultaneously, such as those by the Institute for Clinical and Economic Review, are becoming more accepted by decision-makers (*21*). The growing consideration of both affordability and value is less a result of methodological advances than of an increased focus on how to ensure sustainable and efficient health care (and the corresponding political will to do so). A positive consequence of this is an increase in research on how to best define and quantify affordability and value given the available data.

PM is here to stay. However, it can only achieve its potential if it is both affordable and of high value.

End the entanglement of race and genetics

By **Dorothy E. Roberts[10]**

In the aftermath of the first publication of the human genome, researchers confirmed what many scholars had recognized for decades: that race is a social construct, not a natural division of human beings written in our genes (*22*, *23*). Yet rather than hammer the final nail in the coffin, the human genome map sparked renewed interest in race-based genetic difference. The posting of recent genetic studies on white supremacist websites led the American Society of Human Genetics in 2018 to issue yet another statement denouncing genetics-based claims of racial purity as "scientifically meaningless," while many geneticists failed to see how the biological concept of race was itself invented to support racism. None of this history has restrained the search for genetic differences between races and genetic explanations for various racial disparities (e.g., in COVID-19 outcomes), which in turn generates persistent public confusion about race and genetics.

It is time to end the entanglement of race and genetics and to work toward a radically new understanding of human unity and diversity. There are two general approaches that can help guide innovative research questions and methods that no longer rely on invented racial classifications as if they were biological. First, genetic researchers should stop using race as a biological variable that can explain differences in health, disease, or responses to therapies (*24*). Treating race as a biological risk factor obscures how structural racism has biological effects and produces health disparities in racialized populations. Epigenetics offers promising models to investigate one pathway through which unequal social conditions get "embodied" or "under the skin" to generate disparate health outcomes. Still, researchers must use caution to avoid making deleterious epigenetic processes seem self-perpetuating and inevitable, taking attention away from structural inequities that caused the problem in the first place (*25*).

Second, genetic researchers should stop using a white, European standard for human genetics and instead study a fuller range of human genetic variation. Projects dedicated to expanding genetic databases with DNA from groups on the African continent, for example, have shown that these populations are the most genetically diverse on Earth and refute the myth that there is a genetically distinguishable Black race (*26*). The aim of diversifying biomedical research should not be to find innate genetic differences between racial groups; rather, it should be to give persons from racialized populations equal access to the benefits of participating in high-quality and ethical research (including clinical trials) and to give scientists a richer resource to understand human biology. In this way, genetic research can contribute to more individualized diagnoses and therapies that no longer rely on crude medical decisions based categorically on a patient's race.

Genetic privacy in the post-COVID world

By **Dina Zielinski[11,12]** *and* **Yaniv Erlich[13]**

In 2007, only two individuals had their full genome sequenced: Craig Venter and Jim Watson. Today, more than 30 million individuals have access to their detailed genomic datasets. This democratization of genomic data has helped to reunite families,

Downloaded from http://science.sciencemag.org/ on February 4, 2021

Published by AAAS

fight racism, and promote genetic literacy (*27*, *28*), but it has also enabled surveillance on a massive scale. The correlation of DNA variants between distant relatives means that relatively small databases can identify large parts of the population, including people who are not in the database (*29*). The high dimensionality of DNA data and linkage disequilibrium mean that efforts to obscure individual-level data, by pooling genomes or censoring parts of the genome, can fail unexpectedly (*30*). And with the advent of consumer genomics and third-party websites that allow participants to upload their genome data, it is increasingly easy to collect and access DNA data (*31*).

We envision that the COVID-19 pandemic will accelerate genetic surveillance. People will likely see infectious disease surveillance, swabbing upon arrival, at border crossings, including airports. Governments can harness pandemic control infrastructure to build a DNA database of all arrivals. Such databases can identify a substantial portion of the visitor's home-country population because genetic re-identification is magnified through familial connections. But massive surveillance will not be restricted to government efforts. With the growing size of third-party genetic databases, essentially everyone with the right technical skills will be able to identify individuals.

What are the implications of ubiquitous genetic surveillance? On the plus side, law enforcement agencies will be able to solve virtually all sexual assault cases. Screening at airports can help to reveal fraudulent identities, which is central in fighting human trafficking and espionage. However, the same technology can be used to target minorities or political opponents.

The convergence of these applications underscores the importance of treading lightly with these new forensic superpowers. On the technical side, one theoretical mitigation option to limit such re-identification could include creating a trail that leads a genealogical tracing attempt to a fake identity. But this and other methods have yet to be investigated in a principled approach. Beyond technological countermeasures, the field needs guidelines concerning the use of genetic surveillance technologies. An important step is the interim policy laid out by the U.S. Department of Justice restricting forensic investigators' usage of third-party genetic databases to investigations of violent crimes, and only with sites that receive informed consent from users for such searches (*32*). Open public discussion is vital to further shape policies and expectations so as to harness the power of the genomic revolution for the benefit of the public.



PHOTO: RENE BEGAY

Dr. Jessica Elm watches Alison Watson hold the pipette for a DNA extraction exercise during the Summer Internship for Indigenous Peoples in Genomics (SING) workshop in 2019.

Emerging ethics in Indigenous genomics

By **Nanibaa' A. Garrison**[14,15,16] *and* **Stephanie Russo Carroll**[17,18]

Despite considerable advances in genomics research over the past two decades, Indigenous Peoples are incredibly underrepresented. Biological materials from Indigenous Peoples have been collected to study diseases, medical traits, and the origins of human populations, yet many studies have not benefited the participants or their communities. Some research has even created harms such as exacerbation of derogatory and detrimental stereotypes or challenges to cultural beliefs. Without productive relationships, Indigenous communities may not benefit from research in areas such as precision medicine and pharmacogenomics, and health disparities may remain unaddressed. Thus, many Indigenous Peoples are hesitant to participate in genomics research without extensive discussions and agreements to ensure that the results have individual and collective benefits, as well as to learn what happens to samples and how they are used (*33*). Indigenous scholars are developing guidance to address concerns and pave pathways for more equitable and beneficial research that aligns with the rights and interests of Indigenous Peoples (*34*).

Culturally aligned research can increase Indigenous Peoples' participation in genomics research. The Summer Internship for Indigenous Peoples in Genomics (SING) trains and builds capacity for scientists and community members to shape research priorities of interest in their communities, and it has prompted the SING Consortium to develop a framework for ethical research engagement (*35*). The Center for the Ethics of Indigenous Genomic Research supports Indigenous-led research in biobanking and precision medicine that integrates sovereignty rights and Indigenous communities' ethical and cultural preferences. In Canada, Silent Genomes is creating an Indigenous Background Variant Library through close engagement with community and cultural advisors. Finally, in New Zealand, the Māori-developed Te Ara Tika framework integrates relationships, research design, cultural and social responsibility, justice, and equity as core interests for ethical genomic research with Māori people (*36*).

Recognizing the need to foster self-determination and collective rights within open science and secondary use, the Global

Downloaded from http://science.sciencemag.org/ on February 4, 2021

Published by AAAS

Indigenous Data Alliance's CARE Principles for Indigenous Data Governance (Collective Benefit, Authority to Control, Responsibility, and Ethics) complement the FAIR principles (Findability, Accessibility, Interoperability, and Reusability) that make data machine-readable and usable in multiple contexts (*37*, *38*). When operationalized together, CARE and FAIR enhance Indigenous leadership and innovation, leading to participatory governance and enabling opportunities for trust-building and accountability by incorporating Indigenous values and rights. For example, the creation of data standards and the use of Indigenous community-defined metadata can protect data while allowing them to be useful. The metadata become durable and persistent components of genomic information that provide guidance on future use, such as who has the authority to sanction that use, for what purposes, and to benefit whom (*34*, *37*).

An increased focus on rights and interests combined with enhanced engagement and capacity has the potential to reduce bias and produce more relevant and beneficial research for all.

N.A.G. is Diné, a citizen of the Navajo Nation. S.R.C. is Ahtna, a citizen of the Native Village of Kluti-Kaah.

Polygenic risk in a diverse world

By **Pilar N. Ossorio**[19]

Polygenic risk scores (PRSs) are a rapidly emerging technology for aggregating the small effects of multiple polymorphisms across a person's genome into a single score. A PRS can be calculated for any phenotype for which genome-wide association data are available, usually by summing the weighted effect sizes of alleles (*39*). In medicine and public health, PRSs could be used for selecting therapies, initiating additional risk screening, or motivating behavior change. Whether they will be used in medicine depends on factors such as the degree to which they provide actionable risk information beyond that provided by clinical algorithms, the availability of information technology for calculating PRSs in clinical settings, and the availability of decision support tools. To date, PRSs have demonstrated moderate utility for complex medical phenotypes, including blood pressure, obesity, diabetes, depression, schizophrenia, and coronary heart disease.

PRSs also highlight the complex intersection of race and ancestry in genomics. Substantiating and extending earlier work, a recent analysis showed that in 26 previous studies, PRSs performed significantly worse for people with predominantly African or South Asian ancestry than for people with predominantly European ancestry (*40*, *41*). There was not enough data to assess performance for many groups (e.g., South East Asians, Pacific Islanders). Researchers have attributed this result to underrepresentation of non-European individuals and racial/ethnic minorities in datasets used to develop PRSs. Relative to people who are included in most genomic datasets, racial/ethnic minorities tend to have a greater portion of recent ancestry from places other than Europe.

In response to the differential predictive power of PRSs, researchers have developed some PRSs specifically for people of predominantly African ancestry, and genome scientists are considering whether "ancestry-specific PRS are needed for every ethnic group…" (*42*). These developments occur as scholars of race call for an end to many uses of "race correction" in medicine (*43*). Appropriate attention to genetic ancestry's effects on PRSs can easily collapse into an ill-informed focus on race, without considering how social inequalities shape health and how race is an imperfect proxy for ancestry. Society needs a multidisciplinary approach for developing and implementing PRSs for diverse communities. Otherwise, ancestry-specific PRSs could reinvigorate people's misconceptions about human races as genetically distinct groups and encourage mistaken views that trait distribution between racial/ethnic groups is primarily caused by genetics (*39*). Such beliefs are central to white supremacy and racist medical practices. Injustice in science can occur because some groups of people are not included (*44*), but injustice can also result from inappropriate inclusion.

Risks of genomic surveillance and how to stop it

By **Yves Moreau**[20] *and* **Maya Wang**[21]

The use of DNA profiling for individual cases of law enforcement has helped to identify suspects and to exonerate the innocent. But retaining genetic materials in the form of national DNA databases, which have proliferated globally in the past two decades, raises important human rights questions. Landmark court decisions in Europe and in the United States set some limits on data collection and retention in DNA databases, such as restricting long-term retention of DNA profiles to people arrested for or convicted of a crime.

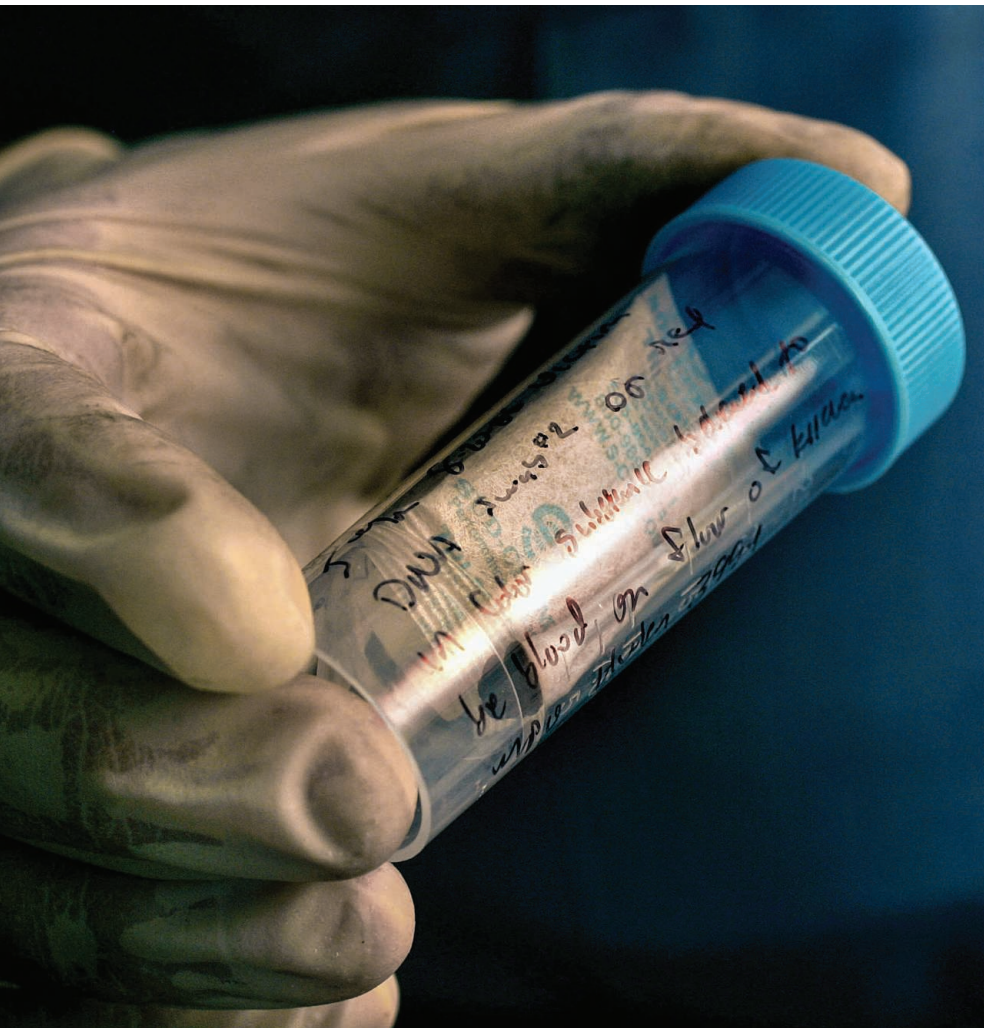
But these decisions are far from the comprehensive regulations we need. Privacy rights are fundamental human rights. Around the world, the unregulated collection, use, and retention of DNA has become a form of genomic surveillance. Kuwait passed a now-repealed law mandating the DNA profiling of the entire population. In China, the police systematically collected blood samples from the Xinjiang population under the guise of a health program, and the authorities are working to establish a Y-chromosome DNA database covering the country's male population. Thailand authorities are establishing a targeted genetic database of Muslim minorities (*45*). Under policies set by the previous administration, the U.S. government has been indiscriminately collecting the genetic materials of migrants, including refugees, at the Mexican border.

As the technology gets cheaper, and as the adoption of surveillance gets ever broader, there is an acute risk of pervasive genomic surveil-

[1]School of Life Sciences, Arizona State University, Tempe, AZ, USA. [2]Center for Medical Ethics and Health Policy, Baylor College of Medicine, Houston, TX, USA; kmaxson@alumni.princeton.edu. [3]School for the Future of Innovation in Society and Consortium for Science, Policy and Outcomes, Arizona State University, Washington, DC, USA. [4]Center for Research on Genomics and Global Health, National Human Genome Research Institute, Bethesda, MD, USA; rotimic@mail.nih.gov. [5]Department of Clinical Research and Leadership, George Washington University School of Medicine and Health Sciences, Washington, DC, USA. [6]School of Humanities, Nanyang Technological University, Singapore; hstevens@ntu.edu.sg. [7]Center for Translational and Policy Research on Personalized Medicine (TRANSPERS), Department of Clinical Pharmacy, University of California, San Francisco, CA, USA. [8]Philip R. Lee Institute for Health Policy Studies, University of California, San Francisco, CA, USA. [9]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA, USA; kathryn.phillips@ucsf.edu. [10]Department of Africana Studies, Department of Sociology, and Law School, University of Pennsylvania, Philadelphia, PA, USA; dorothyroberts@law.upenn.edu. [11]Université de Paris, INSERM 970, Paris Translational Research Centre for Organ Transplantation, Paris, France. [12]Doctoral School 515, Sorbonne Université, Paris, France. [13]Efi Arazi School of Computer Science, IDC Herzliya, Herzliya, Israel; erlichya@gmail.com. [14]Institute for Society & Genetics, University of California, Los Angeles, CA 90095, USA. [15]Institute for Precision Health, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA. [16]Division of General Internal Medicine & Health Services Research, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA; nanibaa@socgen.ucla.edu. [17]Native Nations Institute, Udall Center for Studies in Public Policy, University of Arizona, Tucson, AZ 85724, USA. [18]Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ 85724, USA. [19]Morgridge Institute for Research and University of Wisconsin Law School, Madison, WI, USA; pnossorio@wisc.edu. [20]University of Leuven (KU Leuven), Leuven, Belgium; yves.moreau@kuleuven.be. [21]Human Rights Watch, New York, NY, USA; wangm@hrw.org.

Published by AAAS

Downloaded from http://science.sciencemag.org/ on February 4, 2021



A police officer holds a DNA test swab of blood found at a crime scene in Kansas City, Missouri, 24 July 2008. DNA tests and genetic databases have been increasingly used, with some success, by law enforcement, but this also raises concerns about how to prevent potential overreach in the form of ubiquitous genomic surveillance.

PHOTO: FRED BLOCHER/KANSAS CITY STAR/TRIBUNE NEWS SERVICE/GETTY IMAGES

lance, not only by authoritarian regimes but also in democracies with weakening rights. But such a loss of autonomy and freedom is not inevitable. Governments should reform surveillance laws and draft comprehensive privacy protections that tightly regulate the collection, use, and retention of DNA and other biometric identifiers (*46*). They should ban such activities when they do not meet international human rights standards of lawfulness, proportionality, and necessity. They should develop a coordinated global regime of export control legislation, as well as sanctions akin to the U.S. Magnitsky Act, to hold businesses accountable that recklessly supply or market this technology for genomic surveillance.

Journal editors and publishers should reassess hundreds of ethically suspect DNA-profiling publications—for example, publications co-authored by police forces involved in the persecution of the minorities studied (*47*) or lacking proper consent or ethical approval (*48*). Although there have been a few retractions (*47*, *48*), such assessments should not be limited to the bureaucratic verification of informed consent and ethical approval documents; they also need to consider the basic ethical principles of beneficence, nonmaleficence, autonomy, justice, and faithfulness. The scientific community should also refuse to cooperate with law enforcement anywhere in the world that is proven to be violating human rights standards, in particular the Chinese police and military.

REFERENCES AND NOTES

1. E. S. Lander *et al.*, *Nature* **409**, 860 (2001).
2. B. J. Strasser, *Isis* **102**, 60 (2011).
3. K. M. Jones, R. A. Ankeny, R. Cook-Deegan, *J. Hist. Biol.* **51**, 693 (2018).
4. National Human Genome Research Institute, *NHGRI Genomic Data Sharing (GDS) Policy: Data Standards* (2020).
5. N. L. Yozwiak, S. F. Schaffner, P. C. Sabeti, *Nature* **518**, 477 (2015).
6. "Benefits of sharing" [editorial], *Nature* **530**, 129 (2016).
7. A. R. Bentley, S. Callier, C. N. Rotimi, *J. Community Genet.* **8**, 255 (2017).
8. T. A. Hoppe *et al.*, *Sci. Adv.* **5**, eaaw7238 (2019).
9. G. N. Nadkarni *et al.*, *N. Engl. J. Med.* **379**, 2571 (2018).
10. Y. Luo, *Nature* **587**, 552 (2020).
11. E. D. Green *et al.*, *Nature* **586**, 683 (2020).
12. H. Stevens, *Biosocieties* **11**, 352 (2016).
13. D. Grishin, K. Obbad, G. M. Church, *Nat. Biotechnol.* **37**, 1115 (2019).
14. L. Bonomi, Y. Huang, L. Ohno-Machado, *Nat. Genet.* **52**, 646 (2020).
15. K. Ferryman, M. Pitcan, *Fairness in Precision Medicine* (Data & Society, 2019).
16. S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier* (Profile, 2019).
17. D. M. Cutler, *JAMA* **323**, 109 (2020).
18. K. A. Phillips, M. P. Douglas, D. A. Marshall, *JAMA* **324**, 2029 (2020).
19. E. Faulkner *et al.*, *Value Health* **23**, 529 (2020).
20. A. Towse, J. A. Mauskopf, *Value Health* **21**, 249 (2018).
21. S. D. Pearson, *Value Health* **21**, 258 (2018).
22. J. L. Graves Jr., *The Emperor's New Clothes: Biological Theories of Race at the Millennium* (Rutgers Univ. Press, 2001).
23. D. Roberts, *Fatal Invention: How Science, Politics, and Big Business Re-Create Race in the Twenty-First Century* (New Press, 2012).
24. M. Yudell, D. Roberts, R. DeSalle, S. Tishkoff, *Science* **351**, 564 (2016).
25. D. E. Roberts, O. Rollins, *Annu. Rev. Sociol.* **46**, 195 (2020).
26. M. C. Campbell, S. A. Tishkoff, *Annu. Rev. Genomics Hum. Genet.* **9**, 403 (2008).
27. A. Panofsky, J. Donovan, *Soc. Stud. Sci.* **49**, 653 (2019).
28. J. S. Roberts *et al.*, *Publ. Health Genomics* **20**, 36 (2017).
29. Y. Erlich, T. Shor, I. Pe'er, S. Carmi, *Science* **362**, 690 (2018).
30. D. R. Nyholt, C.-E. Yu, P. M. Visscher, *Eur. J. Hum. Genet.* **17**, 147 (2009).
31. S. C. Nelson, D. J. Bowen, S. M. Fullerton, *Am. J. Hum. Genet.* **105**, 122 (2019).
32. J. Kaiser, *Science* 10.1126/science.aaz6336 (25 September 2019).
33. R. Taitingfong *et al.*, *J. Am. Med. Inform. Assoc.* **27**, 1987 (2020).
34. M. Hudson *et al.*, *Nat. Rev. Genet.* **21**, 377 (2020).
35. K. G. Claw *et al.*, *Nat. Commun.* **9**, 2957 (2018).
36. M. Hudson *et al.*, *He Tangata Kei Tua—Guidelines for Biobanking with Māori* (Te Mata Hautū Taketake—Māori & Indigenous Governance Centre, Univ. of Waikato, 2016).
37. S. R. Carroll *et al.*, *Data Sci. J.* **19**, 43 (2020).
38. M. D. Wilkinson *et al.*, *Sci. Data* **3**, 160018 (2016).
39. N. A. Rosenberg *et al.*, *Evol. Med. Public Health* **2019**, 26 (2019).
40. L. Duncan *et al.*, *Nat. Commun.* **10**, 3328 (2019).
41. A. R. Martin *et al.*, *Am. J. Hum. Genet.* **100**, 635 (2017).
42. National Human Genome Research Institute, *Genomic Medicine XII: Genomics and Risk Prediction*, Executive Summary (2019); www.genome.gov/sites/default/files/media/files/2019-07/GMXII_Executive_Summary.pdf.
43. D. E. Roberts, *Lancet* **397**, 17 (2021).
44. A. R. Martin *et al.*, *Nat. Genet.* **51**, 584 (2019).
45. UN Committee on the Elimination of Racial Discrimination, Letter to the Permanent Representative of Thailand to the United Nations Office (15 May 2015); https://bit.ly/39VxGJe.
46. Forensic Genetics Policy Initiative, *Establishing Best Practices for Forensic DNA Databases* (2017); https://bit.ly/3iasJzL.
47. D. Zhang *et al.*, *Int. J. Legal Med.* 10.1007/s00414-019-02049-6 [retracted] (2019).
48. X. Pan *et al.*, *Int. J. Legal Med.* **134**, 2079 [retracted] (2020).

ACKNOWLEDGMENTS

K.M.J. and R.C.-D. were funded by National Cancer Institute (NCI) grant R01 CA237118. C.N.R., S.L.C., and A.R.B. were supported in part by the Intramural Research Program of NIH through the Center for Research on Genomics and Global Health (CRGGH) at the National Human Genome Research Institute (NHGRI). The CRGGH is also supported by the National Institute of Diabetes and Digestive and Kidney Diseases. K.A.P. is a consultant to Illumina Inc. and was funded by NCI grant R01 CA221870 and NHGRI grant U01 HG009599. J.P.J. is a part-time salaried employee of the Precision Medicine Group. D.Z. and Y.E. hold equity in DNA.Land, a third-party genetic service. D.Z. is an employee of Cibiltech; Y.E. is an employee of Eleven Therapeutics, MyHeritage, and consultant to ArcBio. The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policies or positions of any of their employers.

10.1126/science.abg5266

Downloaded from http://science.sciencemag.org/ on February 4, 2021

Published by AAAS

Science

Complicated legacies: The human genome at 20

Kathryn Maxson Jones, Robert Cook-Deegan, Charles N. Rotimi, Shawneequa L. Callier, Amy R. Bentley, Hallam Stevens, Kathryn A. Phillips, Jeroen P. Jansen, Christopher F. Weyant, Dorothy E. Roberts, Dina Zielinski, Yaniv Erlich, Nanibaa' A. Garrison, Stephanie Russo Carroll, Pilar N. Ossorio, Yves Moreau and Maya Wang

Science **371** (6529), 564-569.
DOI: 10.1126/science.abg5266

ARTICLE TOOLS http://science.sciencemag.org/content/371/6529/564

RELATED CONTENT
http://science.sciencemag.org/content/sci/371/6529/545.full
http://science.sciencemag.org/content/sci/371/6529/556.full
http://science.sciencemag.org/content/sci/371/6529/560.full
http://science.sciencemag.org/content/sci/371/6529/570.full
http://science.sciencemag.org/content/sci/371/6529/578.full

REFERENCES This article cites 36 articles, 3 of which you can access for free
http://science.sciencemag.org/content/371/6529/564#BIBL

PERMISSIONS http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works

Downloaded from http://science.sciencemag.org/ on February 4, 2021