Science and Technology

Genomic Privacy: Advocating for the Convergence of Legal and Technical Solutions

Can Kockan, Dov Greenbaum, Danielle Lee, and Mark Gerstein

The economy of DNA: consumer genomics and its risks

Genomic data collection and analysis is an incredibly valuable sector within the information economy, yet genomic information has never been more pervasive nor publicly accessible. Just this past spring, the National Institutes of Health's "All of Us" program announced that it would make nearly one hundred thousand unique and diverse whole genome sequences available for researchers, and several other public efforts have similarly set out to collect and disseminate enormous amounts of genomic data for research.[1] Beyond such institutions looking to innocuously advance their research and even beyond amateur genomic sleuths combing through our genetic information to try and crack unsolved crimes,[2] there also exists reason to be concerned about matters relating to genetic privacy on the corporate side.

Can Kockan is a postdoctoral associate at Yale University, Department of Molecular Biophysics and Biochemistry. His current research is focused on algorithm development for privacy-preserving and secure genome and metagenome analysis. He completed his PhD in Computer Science at Indiana University where he designed and implemented privacy-preserving algorithms for bioinformatics tasks such as Genome-Wide Association Studies (GWAS) and genotype imputation.

Professor Dov Greenbaum is the Director of the Zvi Meitar Institute for Legal Implications of Emerging Technologies and a professor of law at the Harry Radzyner Law School, Reichman University, Herzliya, Israel (IDC). Dov is also an affiliate researcher in the Department of Molecular Biophysics and Biochemistry at Yale University. Dov has degrees and postdoctoral fellowships from Yale, UC Berkeley, Stanford, and Eidgenössische Technische Hochschule Zürich (ETH Zürich) and is a practicing intellectual property attorney.

Danielle Lee is an undergraduate researcher at Yale University, Department of Molecular Biophysics and Biochemistry. She is jointly supervised by the Gerstein and Miranker labs, working on cell-free DNA and genome privacy.

Mark Gerstein is the Albert L. Williams Professor of Biomedical Informatics at Yale. He is associated with the Departments of Molecular Biophysics & Biochemistry, Computer Science and Statistics & Data Science. He is Co-Director (and founder) of the Computational Biology & Bioinformatics PhD Program and Co-Director (and founder) of the Yale Center for Biomedical Data Science. Prof. Gerstein completed his PhD training in Computational Chemistry and Biophysics at Cambridge University, followed by postdoctoral training at Stanford. Since then, he has published >600 manuscripts in total, including several in prominent venues, such as Science, Nature, and Cell, with an H-index of >175. He is a specialist in bioinformatics with a particular interest in large-scale data science, especially as it pertains to personal genome analyses. Current research foci in his lab include disease genomics (particularly neurogenomics and cancer genomics), human genome annotation, genomic privacy, network science (especially gene regulatory networks), wearable and molecular imaging data analysis, text mining of the biological science literature and macromolecular simulation. Prof. Gerstein has received awards such as being elected as a fellow of AAAS and the International Society of Computational Biology.

Companies such as MyHeritage and Ancestry.com collect client DNA in various degrees of quantity and quality, at times even using this data for commercial purposes and selling clients' genomic information to drug companies for hundreds of millions of dollars.[3]

In addition to the growing size of both public and private DNA databases, these databases have also grown increasingly more revealing as research continues to uncover additional correlations between genes and disease and hereditary conditions. For example, identical twins share 100 percent of their DNA, parents and siblings share up to 61 percent of their DNA, and even distant third cousins share up to 2.2 percent of their DNA.[4] Thus, if even one of these individuals chooses to disclose their genomic information, they are in effect disclosing not only theirs but also much of the genomes of their close relatives. One study showed how the genetic information of the majority of Americans of European ancestry can be identified through their distant relatives' DNA that already exists in various public and private current databases.[5] The result of such data is an ever-expanding genomic panopticon, as predicted by the late Supreme Court Justice Antonin Scalia.[6]

Following the COVID-19 pandemic, the genomic information consumed and produced around the world has only further contributed to expanding the realm of genomic data available to us. Genetic tests have become especially de rigueur when traversing borders, and it was even reported recently that various heads of state refused PCR testing when meeting the Russian president Putin out of fear that Russian intelligence might abuse and misuse the resulting genetic information.[7] Yet, this fear is not limited to politicians, celebrities, or sports personalities whose genetic information could contain exploitable knowledge.[8] In fact, given all the sunk costs for the infrastructure of pandemic genetic testing, some countries will likely continue to find excuses to collect and examine the DNA of visitors for a host of potentially actionable genetic data, including controversial correlations like propensity to violence or depression.[9] As such, now would be the best time to enhance genetic data protection and privacy laws before we become too complacent with giving away our genes at the border, or anywhere else.

A brief history of the technical and legal solutions for genomic privacy

U.S. legislators have long assumed the primary role of protecting our genetic data, from instituting federal regulations like the Federal Genetic Information Nondiscrimination Act of 2008 (GINA) to state laws including those that target the limited use of genetic information in areas of health insurance, employment, and even direct-to-consumer genetic testing companies like 23andMe.[10] However, laws are slow to change, and regulations are slow to implement. The Health Insurance Portability and Accountability Act (HIPAA) is exemplary of slowness in responding to outdated regulations. The 1996 law outlines limitations on the use of protected health information (PHI). Under HIPAA's Privacy Rule, individually identifiable health information is protected. Deidentified information is exempt from this rule, as is other data deemed unidentifiable.[11] This definition is anachronistic in our world of big data, and even heretofore perceived benign data has long been shown to be as revealing as PHI.[12]

Practically speaking, these regulations are also hard to enforce; oftentimes, genetic information is obtainable without needing the consent of the individual (e.g., from a drinking glass or a coffee cup), providing an abuser of genetic information plausible deniability regarding the misuse of the data—for example, in employment. Additionally, particularly with state laws, out-of-state and offshore labs can circumvent many regional rules. Recognizing that legislation has its limitations, researchers have also sought to uncover technical methods of genomic data protection. Specifically, bioinformatics researchers have been interested in quantifying the information leakage from genomic data sharing and applying different computational techniques for enabling privacy-preserving and secure genome analysis for almost two decades.[13]

Among the genetic encryption techniques that have been developed for genomic data

protection, homomorphic encryption schemes have shown significant potential. In very simple terms, homomorphic encryption is a technique that allows functions, such as addition and multiplication, to be computed over encrypted data. Therefore, it is even possible for external untrusted service providers to use cloud computation resources (e.g. Amazon AWS or Microsoft Azure) when analyzing client genomic information without the client fearing that their private data will be revealed—once the analysis is completed on the server side, the results, still encrypted, can be sent back to the client. Since only the client possesses the private key which allows the decryption of the data, the client can then decrypt the analysis results once they have received it.

. . . homomorphic encryption is a technique that allows functions, such as addition and multiplication, to be computed over encrypted data.

The possibility of such a fully homomorphic encryption scheme was first discussed by Ronald Rivest in 1978.[14] However, it was not until 2009 that the thirty-year-old open problem was finally resolved when Craig Gentry described the first plausible Fully Homomorphic Encryption (FHE) scheme in his doctoral thesis,[15] laying the foundation for a new wave of scientists to research homomorphic encryption and ways to improve its computational performance. Efforts have only intensified since then, alongside certain important developments in the fields of computer security and cryptography, such as the emergence of practical homomorphic encryption (HE) schemes like BFV and CKKS, secure multiparty computation (SMC), and trusted execution environments (TEE).[16] Differential privacy (DP) mechanisms for releasing summary statistics without violating privacy have also been rigorously studied to come up with ways to provide optimal utility versus privacy tradeoffs.[17] Resultantly, HE-based solutions have been developed for a number of tasks including Genome-Wide Association Studies (GWAS), genotype imputation, viral strain classification, edit distance computation, tumor classification, among others.[18]

Despite the advantages of FHE, there also exist a number of shortcomings to this technology. Most significantly, for certain complex applications such as deep convolutional and recurrent neural networks, FHE is notoriously impractical, prohibitively expensive, and sometimes even inaccurate. In these cases, alternative solutions are considered.

In 2015, another important development gained the attention of the genome privacy community. The introduction of Intel Software Guard Extensions (SGX), a TEE, made it possible for complex models to be analyzed with higher accuracy and at a reasonable price.[19] An SGX enclave can be understood as an area on the main processor whose contents cannot be tampered with even when an attacker takes control of the host operating system. The user's private data then remains encrypted during transit and is only decrypted once inside this secure enclave. Once the analysis is performed, the results are encrypted again and sent back to the client.

Although Intel SGX was neither the first nor the only available TEE, researchers have been inclined to utilize it over competitors such as ARM TrustZone and AMD Secure Encrypted Virtualization (AMD-SEV) because it exhibits greater practicality for evaluating certain privacy-preserving bioinformatics applications.[20] In terms of computational performance, TEEs generally have a significant advantage compared even to state-of-the-art FHE schemes. Moreover, while many instances of Intel SGX's vulnerability to side-channel attacks have led some to question the level of protection these platforms provide compared to FHE schemes, careful design and implementation choices can mitigate most, if not all, of these issues.[21]

Apart from HE and TEEs, recent efforts to protect genomic information also include the development of privacy-preserving file formats, which sanitize commonly utilized bioinformatics file formats such as the Sequence Alignment Map (SAM) and Binary Alignment Map (BAM) in order to both preserve the utility of

the genomic reads for downstream analysis and remove information that could potentially leak variants, as well as other SMC-based solutions.[22]

Given the wide array of techniques available for use in the field of genome privacy and security, one chief advantage of bioinformatics for genomic data protection is that it presents a multitude of approaches and protocols from which to draw when trying to attack security problems from multiple angles. Yet, underlying this advantage is an even greater challenge: there is no current technique that seems to be the clear solution to all genome privacy problems.

The need for standardization is not new—standardization efforts do exist and have been prevalent in technological circles for some time. For instance, the slow but steady transition of the internet from http to https set a strong precedent of one such standardization process. While the common user may not nor need not understand concepts like public-key cryptography, key generation and exchange, or Transport Layer Security (TLS), the entirety of modern secure web browsing and e-commerce depends on it all the same. Once HE, TEEs, or other such solutions rise to the same level of user experience, we might expect to observe a similar trend toward widespread community adoption of genomic data protection softwares. Still, most of the work done in this area so far has remained purely research-driven—its shift toward becoming a practical service has occurred extremely slowly, if at all.

Whether analyzing the legal or technical approaches to genetic privacy concerns, the ultimate goal of all development is to enhance and accelerate genome research. Both regulations and technology aim to promote the beneficial uses of data while preserving patient privacy and information security. Still, although both sides have made strides toward safeguarding our genomic information, these two approaches to genetic data protection have also long remained independent of one another. As such, we believe that the best approach to protecting genetic data privacy may be found through a convergence of legal and technical solutions.

The goal: symbiosis of technical and legal solutions

The intersection of law and technology is increasingly important in modern society. This is especially true in relation to biomedical research and genomics. The privacy implications of genomic data require legal regulation, but at the same time, we do not want the administrative burdens of law to hinder the rapid development of this technology and its benefits to precision medicine. Therefore, in order for genomic privacy research to be safe yet meaningful, it is crucial both for the researchers to understand and partake in the regulations that govern new technological development and also for the lawmakers to consider and carry out whatever legislation is most appropriate to support the rapidly evolving technological landscape of our modern day.

In particular, genetic privacy would be much better served if lawmakers would promote technical mechanisms designed to protect genetic information. Instead of locking in methods and technologies that may soon become obsolete, legislation ought to be designed specifically in deference to the concept of evolving technologies. To this end, federal and state governments should focus on both granting researchers and industry groups the power to guide genetic privacy protection through such evolving technologies as well as deputizing academic, industry, or mixed standard setting groups to set the necessary standards for technological protection.

With technologies such as HE, TEE, and SMC at the forefront of genetic privacy law, it would be easier to define an efficient legal framework for the secure sharing of genomic data across many institutions in different geographical locations. While many traditional laws and regulations are unable to keep up with the

Instead of locking in methods and technologies that may soon become obsolete, legislation ought to be designed specifically in deference to the concept of evolving technologies.

pace of scientific innovations in genomics, all of these technologies are tools that not only allow collaborative research across multiple national and international institutions but also provide the means for higher efficiency in research.

This technological advantage is especially apparent when considering time-sensitive cases such as rare diseases research. With rare diseases research, individual institutions might not possess sufficient data to perform statistical analyses or yield confident and actionable results. In contrast, the amount of data pooled across multiple institutions would significantly improve the statistical power of state-of-the-art analysis tools. With current access control limitations, the data sharing process required to begin the project might take months, assuming that access will eventually be granted, thereby inhibiting efficient collaborative analysis that could speed up the research on a potentially life-threatening condition by many orders of magnitude. Moreover, through the use of these technologies, policymakers and researchers alike would be able to identify parties exhibiting various levels of trust and subsequently regulate how much information should be shared at each level. Hence, we must also be precise in our definitions of threat and trust models as well as the acceptable levels of utility versus leakage tradeoffs.

The traditional multi-party computation (MPC) defines two primary threat models relevant to the types of analysis we are interested in. These models are the malicious (active) adversary, who might cause compromised or corrupted parties to deviate from the protocol in an attempt to violate security, and the honest-but-curious (semi-honest) adversary, who corrupts parties by passively attempting to learn as much as possible from the received data, all while still following the original protocol.[23] For example, an honest-but-curious adversary could be a compromised third-party cloud service provider who performs its promised functionality but also gathers and stores data from the user for future potential use in functions other than the original task. It should be noted that honest-but-curious adversaries do not include those individuals who we trust to safeguard our data while providing us utility, such as physicians and researchers.

In the past, individuals and institutions have primarily sought to counteract such threats through legal solutions. Indeed, the easiest and perhaps most ideal way to prevent misuse of genomic data would simply be for the engaged parties to sign contracts pertaining to how much of the client data may be used by the service provider outside the original analysis. However, enforcing these agreements becomes an entirely different matter due to the differences in the legal systems across states and countries. In addition, this solution assumes that the third-party cloud service provider will do everything in its power to take the necessary precautions and ensure its systems are secure from external attackers.

By combining the legal and technical approaches to genomic data protection, we propose an alternative way to prevent misuse of genomic data. Past data has shown that the time spent transforming FHE and TEE technologies from being virtually useless to completely viable was significantly more worthwhile than the time spent drafting and redrafting laws to keep up with each generation of consumers seeking privacy protection. Yet, beyond just conferring optimal utility for cases where a zero-trust solution is absolutely necessary, technology holds incredible potential for genomic privacy protection in light of its capacity for further development. Thus, it would not only be more realistic but also more effective for legislators to support and accelerate the development of technological solutions through launching joint programs among the government, industry, and academia.

One such program, "homomorphicencryption.org," has already begun to contribute toward research on a variety of HE schemes, having brought together representatives from the government, industry, and academia to advance secure computation and standardize homomorphic encryption.[24] The development of more programs like "homomorphicencryption.org" would serve to maximize the potential for rapid collaborative genome research in the next decade, a need which was clearly highlighted during the recent years of the pandemic.

Beyond these suggested technical solutions, lawmakers and researchers also have opportunities to share their knowledge and ideas through

conferences including RECOMB and ISMB.[25] These conferences have increasingly featured submissions related to genome privacy and security, and there have even been instances in which they have resulted in tangible solutions such as the Open Imputation Server in 2021.[26] The community has also formed workshops like the National Institutes of Health-sponsored iDASH Privacy and Security Workshop, calling for researchers to come up with efficient solutions to certain open problems of interest in genome privacy using state-of-the-art cryptographic technology.[27]

In the same vein, the institution of annual standards workshops would enable groups like "homomorphicencryption.org" to provide industry and academia with appropriate guidance on which technologies are necessary for promoting privacy protection given innovation in the field. However, deciding which groups should be empaneled to set standards is non-trivial. These decisions could likely be best directed by federal scientific agencies such as the National Institutes of Health (NIH) or the Office of Science and Technology Policy. Alternatively, Congress could create a council on bioethics, akin to past presidential commissions and councils, that in reflecting a bipartisan worldview, could be enlisted to either support or supplement the non-governmental standard-setting organizations in developing best privacy practices for genomics.

Evidently, investing time and resources to both promote and guide technological development is vastly more fruitful than trying to restrict its rapid advancement. Through improving technology, the goals set out for genomic privacy would be accomplished with significantly greater speed and efficacy. Yet, while regulations may be unable to keep up with the rapidly changing landscape of time-sensitive research, the legal landscape does play a valuable role in protecting the information acquired by genomic research from undesired parties. By focusing more on funding and supporting existing research than on seeking new ways to regulate this space, legislation would effectively work alongside technology for the greatest benefit, not only protecting individual privacy but also allowing researchers to continue accessing and learning from the massive amounts of collected DNA information.

Notes

1. "Research Roundup: Genomic Data Release Opens New Paths for Discovery," National Institutes of Health All of Us Research Program, March 17, 2022, https://allofus.nih.gov/news-events/announcements/research-roundup-genomic-data-release-opens-new-paths-discovery; Ting Wang *et al.*, "The Human Pangenome Project: A Global Resource to Map Genomic Diversity," *Nature* 604, no. 7906 (2022): 437–446, https://doi.org/10.1038/s41586-022-04601-8; "The International HapMap Project," *Nature* 426, no. 6968 (2003): 789–796, https://doi.org/10.1038/nature02168; Adam Auton *et al.*, "A Global Reference for Human Genetic Variation," *Nature* 526, no. 7571 (2015): 68–74, https://doi.org/10.1038/nature15393; "UK Biobank - UK Biobank," UK Biobank - UK Biobank, June 17, 2022, https://www.ukbiobank.ac.uk/; "ICGC Data Portal," dcc.icgc.org, https://dcc.icgc.org/.
2. Paige St. John, "The Untold Story of How the Golden State Killer Was Found: A Covert Operation and Private DNA." *Los Angeles Times*, December 8, 2020, https://www.latimes.com/california/story/2020-12-08/man-in-the-window.
3. Antonio Regalado, "More than 26 Million People Have Taken an at-Home Ancestry Test," MIT Technology Review, June 18, 2020, https://www.technologyreview.com/2019/02/11/103446/more-than-26-million-people-have-taken-an-at-home-ancestry-test/; Samual A. Garner and Jiyeon Kim, "The Privacy Risks of Direct-to-Consumer Genetic Testing: A Case Study of 23andMe and Ancestry," *Washington University Law Review* 96, no. 6, (2019): 1219–1265, https://openscholarship.wustl.edu/law_lawreview/vol96/iss6/6.
4. "Average Percent DNA Shared between Relatives - 23andme Customer Care," 23AndMe, accessed August 17, 2022, https://customercare.23andme.com/hc/en-us/articles/2121706

68-Average-percent-DNA-shared-between-relatives.
5. Yaniv Erlich *et al*., "Identity Inference of Genomic Data Using Long-Range Familial Searches," *Science* 362, no. 6415 (September 2018): 690–694, https://doi.org/10.1126/science.aau4832.
6. Maryland v. King, 569 U.S. 435, (2013).
7. Person and Michel Rose, "Macron Refused Russian Covid Test in Putin Trip over DNA Theft Fears," *Reuters*, February 11, 2022, https://www.reuters.com/world/europe/putin-kept-macron-distance-snubbing-covid-demands-sources-2022-02-10/.
8. Sonia M. Suter, "From Sweaty Towels to Genetic Stats: Stalking Athletes for Their Genetic Information," *Recent Patents on DNA &Amp; Gene Sequences* 6, no. 3 (January 2012): 189–192, https://doi.org/10.2174/187221512802717286.
9. Yiqiu Hu *et al*, "The Effect of Childhood Maltreatment on College Students' Depression Symptoms: The Mediating Role of Subjective Well-Being and the Moderating Role of MAOA Gene RS6323 Polymorphism," *European Journal of Developmental Psychology* 19, no. 3 (2021): 438–457, https://doi.org/10.1080/17405629.2021.1928491.
10. Pub. L. 110–233, enacted May 21, 2008; "More than 26 million people have taken an at-home ancestry test. MIT Technol Rev; "Genome Statute and Legislation Database," Genome.gov, accessed July 25, 2022, https://www.genome.gov/about-genomics/policy-issues/Genome-Statute-Legislation-Database.
11. Office for Civil Rights (OCR), "Summary of the HIPAA Privacy Rule," HHS.gov, July 27, 2021, https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html.
12. Michal Kosinski, David Stillwell, and Thore Graepel, "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior," *Proceedings of the National Academy of Sciences* 110, no. 15 (November 2013): 5802–5805, https://doi.org/10.1073/pnas.1218772110.
13. Dov Greenbaum *et al*., "Computer Security in Academia—a Potential Roadblock to Distributed Annotation of the Human Genome," *Nature Biotechnology* 22, no. 6 (2004): 771–772, https://doi.org/10.1038/nbt0604-771; Dov Greenbaum *et al*., "Genomics and Privacy: Implications of the New Reality of Closed Data for the Field," *PLoS Computational Biology* 7, no. 12 (January 2011): https://doi.org/10.1371/journal.pcbi.1002278; Arif Harmanci and Mark Gerstein, "Quantification of Private Information Leakage from Phenotype-Genotype Data: Linking Attacks," *Nature Methods* 13, no. 3 (January 2016): 251–256, https://doi.org/10.1038/nmeth.3746; Arif Harmanci and Mark Gerstein, "Analysis of Sensitive Information Leakage in Functional Genomics Signal Profiles through Genomic Deletions," *Nature Communications* 9, no. 1 (2018): https://doi.org/10.1038/s41467-018-04875-5.
14. Ronald L. Rivest et al., "On data banks and privacy homomorphisms." *Foundations of secure computation* 4.11 (1978): 169–180.
15. Craig Gentry, "Fully Homomorphic Encryption Using Ideal Lattices," *Proceedings of the 41st Annual ACM Symposium on Symposium on Theory of Computing - STOC '09*, 2009, https://doi.org/10.1145/1536414.1536440.
16. Junfeng Fan and Frederik Vercauteren. "Somewhat Practical Fully Homomorphic Encryption." *IACR Cryptol. ePrint Arch.* 2012 (2012): 144; Jung Hee Cheon *et al*., "Homomorphic Encryption for Arithmetic of Approximate Numbers," *Advances in Cryptology – ASIACRYPT 2017*, 2017, 409–437, https://doi.org/10.1007/978-3-319-70694-8_15.
17. Sean Simmons and Bonnie Berger, "Realizing Privacy Preserving Genome-Wide Association Studies," *Bioinformatics* 32, no. 9 (2016): 1293–1300, https://doi.org/10.1093/bioinformatics/btw009; Nour Almadhoun, Erman Ayday, and Özgür Ulusoy, "Differential Privacy under Dependent Tuples - the Case of Genomic Privacy," *Bioinformatics*, August 2019, https://doi.org/10.1093/bioinformatics/btz837; Abdullah Çağlar Öksüz, Erman Ayday, and Uğur Güdükbay, "Privacy-Preserving and Robust Watermarking on Sequential Genome Data Using Belief Propagation and

Local Differential Privacy," *Bioinformatics* 37, no. 17 (2021): 2668–2674, https://doi.org/10.1093/bioinformatics/btab128.

18. Marcelo Blatt et al., "Secure Large-Scale Genome-Wide Association Studies Using Homomorphic Encryption," *Proceedings of the National Academy of Sciences* 117, no. 21 (December 2020): 11608-11613, https://doi.org/10.1073/pnas.1918257117; "IDASH PRIVACY & SECURITY WORKSHOP 2021 - Secure Genome Analysis Competition," Competition tasks - IDASH privacy & security workshop 2021 - secure genome analysis competition, accessed July 25, 2022, http://www.humangenomeprivacy.org/2021/competition-tasks.html; Jung Hee Cheon, Miran Kim, and Kristin Lauter, "Homomorphic Computation of Edit Distance," *Financial Cryptography and Data Security* (2015): 194–212, https://doi.org/10.1007/978-3-662-48051-9_15; Seungwan Hong et al., "Secure Tumor Classification by Shallow Neural Network Using Homomorphic Encryption," *BMC Genomics* 23, no. 1 (September 2022): https://doi.org/10.1186/s12864-022-08469-w.
19. "Intel® Software Guard Extensions (Intel® SGX)," Intel, accessed July 25, 2022, https://www.intel.com/content/www/us/en/architecture-and-technology/software-guard-extensions.html.
20. Arm Ltd., "Trustzone for Cortex-M – ARM®," Arm, accessed July 25, 2022, https://www.arm.com/technologies/trustzone-for-cortex-m; "AMD Secure Encrypted Virtualization (SEV)," AMD, February 10, 2022, https://developer.amd.com/sev/.
21. Wenhao Wang *et al.*, "Leaky Cauldron on the Dark Land," *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, https://doi.org/10.1145/3133956.3134038; Johannes Götzfried et al., "Cache Attacks on Intel SGX," *Proceedings of the 10th European Workshop on Systems Security*, 2017, https://doi.org/10.1145/3065913.3065915; Jo Van Bulck *et al.*, "Foreshadow: Extracting the Keys to the Intel {SGX} Kingdom with Transient {Out-of-Order} Execution," *27th USENIX Security Symposium (USENIX Security 18)*, 2018; Natnatee Dokmai *et al.*, "Privacy-Preserving Genotype Imputation in a Trusted Execution Environment," March 2021, https://doi.org/10.1101/2021.02.02.429428.
22. Gamze Gürsoy *et al.*, "Data Sanitization to Reduce Private Information Leakage from Functional Genomics," *Cell* 183, no. 4 (2020), https://doi.org/10.1016/j.cell.2020.09.036; Hyunghoon Cho, David J Wu, and Bonnie Berger, "Secure Genome-Wide Association Analysis Using Multiparty Computation," *Nature Biotechnology* 36, no. 6 (July 2018): 547–551, https://doi.org/10.1038/nbt.4108.
23. "Defining multi-Party Computation - Securecomputation.org," accessed July 25, 2022, https://securecomputation.org/docs/ch2-definingmpc.pdf.
24. "Homomorphic Encryption," Homomorphic Encryption Standardization, accessed July 25, 2022, https://homomorphicencryption.org/.
25. "Home," Home, accessed July 25, 2022, https://www.recomb.org/home; "HOME - ISMB 2022," Home - ISMB 2022, accessed July 25, 2022, https://www.iscb.org/ismb2022.
26. Arif O. Harmanci *et al.*, "Open Imputation Server Provides Secure Imputation Services with Provable Genomic Privacy," January 2021, https://doi.org/10.1101/2021.09.30.462262.
27. "IDASH Privacy & Security Workshop 2021 - secure genome analysis competition - home," accessed July 25, 2022, http://www.humangenomeprivacy.org/2021/.