NATIONAL INSTITUTES OF HEALTH

HELA GENOME DATA ACCESS

WORKING GROUP UPDATE TO THE ACD

Garth Graham, MD, MPH & Lyric Jorgenson, PhD
HeLa Working Group Co-Chairs

December 15, 2023

GOALS FOR DISCUSSION

- **Evaluate current HeLa genome access request**
  - Assess working group's findings
  - Approve/deny request
- **Discuss NIH efforts continuing to promote the NIH-Lacks Family Agreement**
  - Learn about NIH's efforts
  - Discuss proposal to expedite ACD reviews

CURRENT HELA GENOME DATA ACCESS REQUEST

REMINDER
THE HELA GENOME DATA USE AGREEMENT

Researchers are expected to:

- Apply for HeLa whole genome sequence access through dbGaP & ACD review
- Abide by HeLa Genome Data Use Agreement terms, including:
  - Using data only for biomedical research purposes
  - Disclosing whether the research could result in a patent or commercial product/service
  - Respecting the privacy of the family by safeguarding the data and not attempting to contact them
  - Recognizing the contributions of Henrietta Lacks and her family by including an acknowledgment in publications and presentations
- Deposit future whole genome sequence data into dbGaP

REMINDER

HELA GENOME DATA ACCESS WORKING GROUP ROSTER

- Garth Graham, MD, MPH (Co-Chair)
  Google/YouTube
- Lyric Jorgenson, PhD (Co-Chair)
  National Institutes of Health
- Russ Altman, MD, PhD
  Stanford University
- Ruth Faden, PhD, MPH
  Johns Hopkins University
- David Lacks, Jr.
  Representative, Henrietta Lacks Family
- Jeri Lacks-Whye
  Representative, Henrietta Lacks Family
- Richard Myers, PhD
  HudsonAlpha Institute for Biotechnology
- Veronica Robinson
  Representative, Henrietta Lacks Family
- Cheryl Jacobs, PhD (Executive Secretary)
  National Institutes of Health

REMINDER

ACCESS REQUEST EVALUATION CRITERIA

- **Is the proposed research focused on health, medical, or biomedical research objectives?**
  - Is the proposed research related to determining the ancestry or population origins of HeLa cells?
- **Are there any plans to develop intellectual property? Specifically:**
  - Does the requestor anticipate or foresee IP or developing commercial products or services from the proposed research?
  - Has the requestor agreed to notify NIH if their plans for IP or commercial products change?
- **Are there any plans to publish or present findings?**

UPDATE

SUBMITTED ACCESS REQUESTS

# Requests	Status/Outcome
97 (total)	Evaluated by the HeLa Genome Data Access Working Group
90	Approved by NIH Director
1	Disapproved by NIH Director
5	Rejected for insufficiency by NIH staff (i.e., requestors did not respond to requests for clarifications regarding publication plans, IP, and/or the non-technical summary)
1	To be reported to ACD today

UPDATE

CURRENT REQUEST: WG ASSESSMENT

Project Title	Requestor's Affiliation	Project Summary	Working Group Assessment
Using HeLa Genome Sequences for CRISPR sgRNA Design	Weizmann Institute of Science	• Requested dataset will be used to design sgRNAs that cut at specific locations in the HeLa genome to identify HeLa-specific CRISPR/Cas9 sequences • Access will allow designing molecular biology tools that could be helpful in studying how specific genomic perturbations affect cancer cells • Dataset use will be limited to applicants and not published, with exception of designed sgRNA sequences • No plans to develop a commercial product or service or to file IP but will notify NIH if their plans change	**CONSISTENT WITH DATA USE AGREEMENT**

UPDATE

CURRENT REQUEST: WG ASSESSMENT ACD VOTE

Project Title	Requestor's Affiliation	Project Summary	Working Group Assessment
Using HeLa Genome Sequences for CRISPR sgRNA Design	Weizmann Institute of Science	• Requested dataset will be used to design sgRNAs that cut at specific locations in the HeLa genome to identify HeLa-specific CRISPR/Cas9 sequences • Access will allow designing molecular biology tools that could be helpful in studying how specific genomic perturbations affect cancer cells • Dataset use will be limited to applicants and not published, with exception of designed sgRNA sequences • No plans to develop a commercial product or service or to file IP	**CONSISTENT WITH DATA USE AGREEMENT**

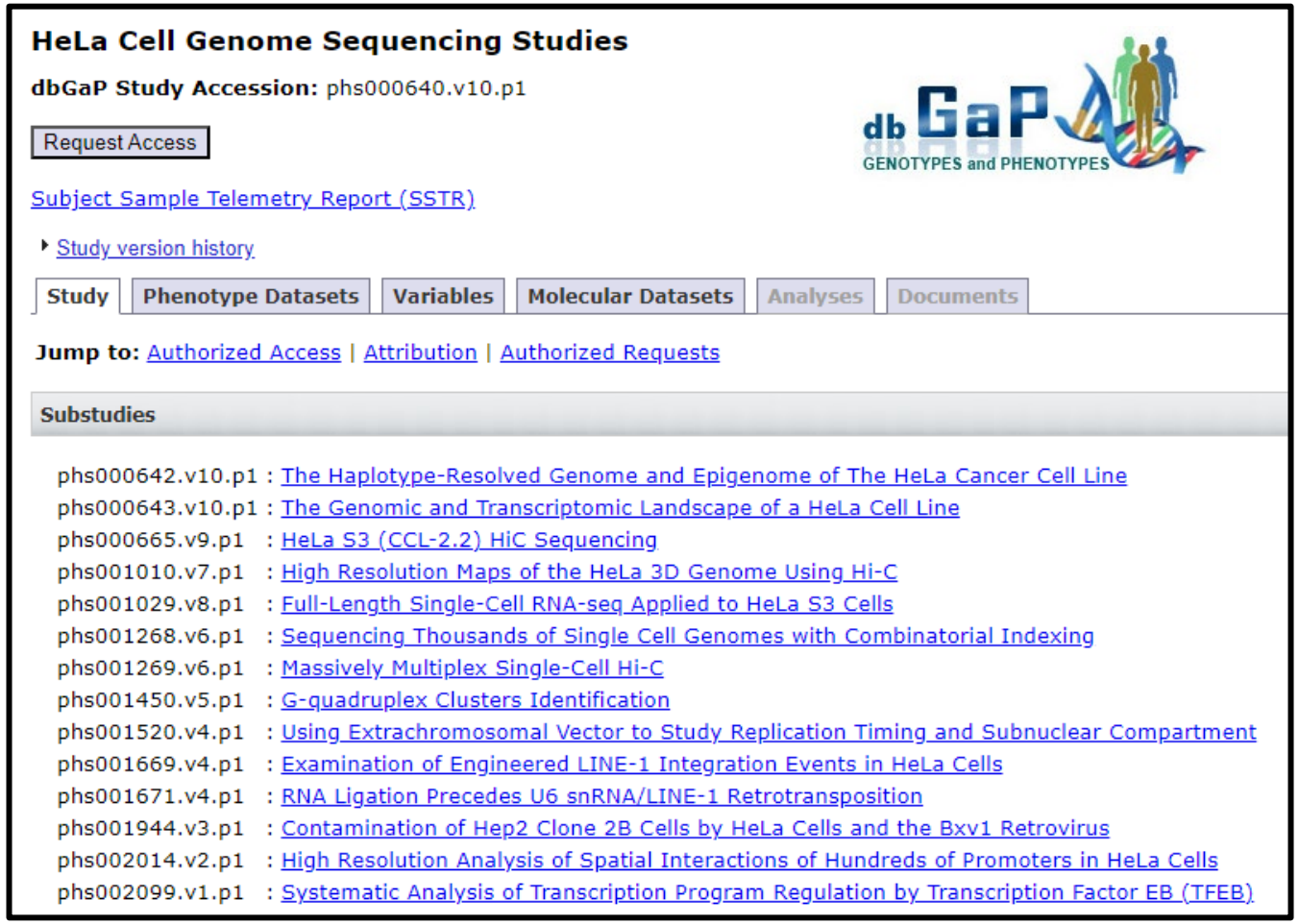
NIH EFFORTS TO CONTINUE TO PROMOTE THE NIH-LACKS FAMILY AGREEMENT

http://www.nature.com/nature/journal/v500/n7461/full/500141a.html

ANALYSES

DATA SUBMISSIONS UNDER THE AGREEMENT

- **14 datasets submitted to NIH under the Agreement since 2013**
- **Approximately 1/3 represent voluntary submissions**
- **Data have been submitted from 6 countries**



HeLa Cell Genome Sequencing Studies

dbGaP Study Accession: phs000640.v10.p1

Request Access

dbGaP GENOTYPES and PHENOTYPES

Subject Sample Telemetry Report (SSTR)

Study version history

Study | Phenotype Datasets | Variables | Molecular Datasets | Analyses | Documents

Jump to: Authorized Access | Attribution | Authorized Requests

Substudies

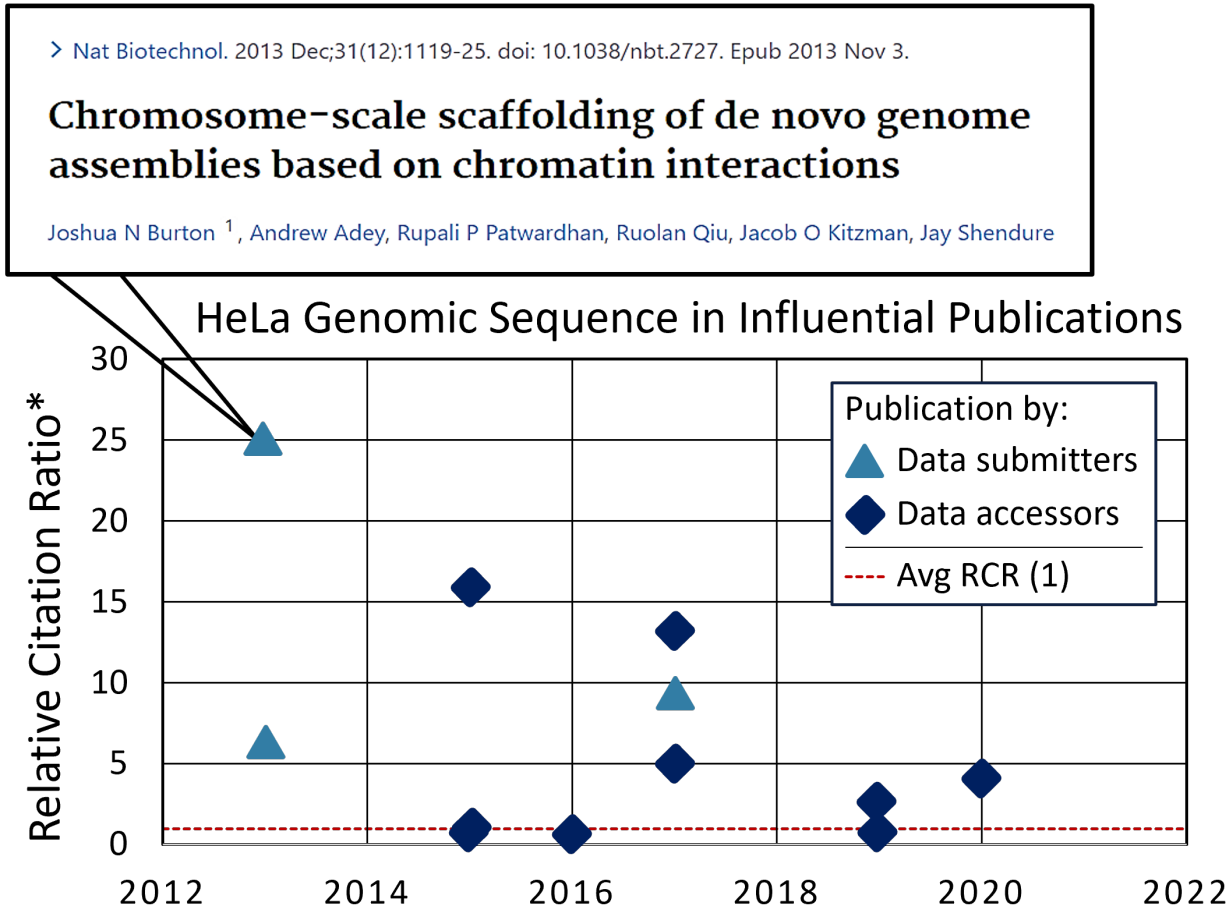
phs000642.v10.p1 : The Haplotype-Resolved Genome and Epigenome of The HeLa Cancer Cell Line
phs000643.v10.p1 : The Genomic and Transcriptomic Landscape of a HeLa Cell Line
phs000665.v9.p1 : HeLa S3 (CCL-2.2) HiC Sequencing
phs001010.v7.p1 : High Resolution Maps of the HeLa 3D Genome Using Hi-C
phs001029.v8.p1 : Full-Length Single-Cell RNA-seq Applied to HeLa S3 Cells
phs001268.v6.p1 : Sequencing Thousands of Single Cell Genomes with Combinatorial Indexing
phs001269.v6.p1 : Massively Multiplex Single-Cell Hi-C
phs001450.v5.p1 : G-quadruplex Clusters Identification
phs001520.v4.p1 : Using Extrachromosomal Vector to Study Replication Timing and Subnuclear Compartment
phs001669.v4.p1 : Examination of Engineered LINE-1 Integration Events in HeLa Cells
phs001671.v4.p1 : RNA Ligation Precedes U6 snRNA/LINE-1 Retrotransposition
phs001944.v3.p1 : Contamination of Hep2 Clone 2B Cells by HeLa Cells and the Bxv1 Retrovirus
phs002014.v2.p1 : High Resolution Analysis of Spatial Interactions of Hundreds of Promoters in HeLa Cells
phs002099.v1.p1 : Systematic Analysis of Transcription Program Regulation by Transcription Factor EB (TFEB)

ANALYSES

SUBMITTED DATA CATALYZING NEW DISCOVERIES

- **Since 2013, 13 citations of datasets**
  - 10 published by approved users
  - 3 published by data submitters

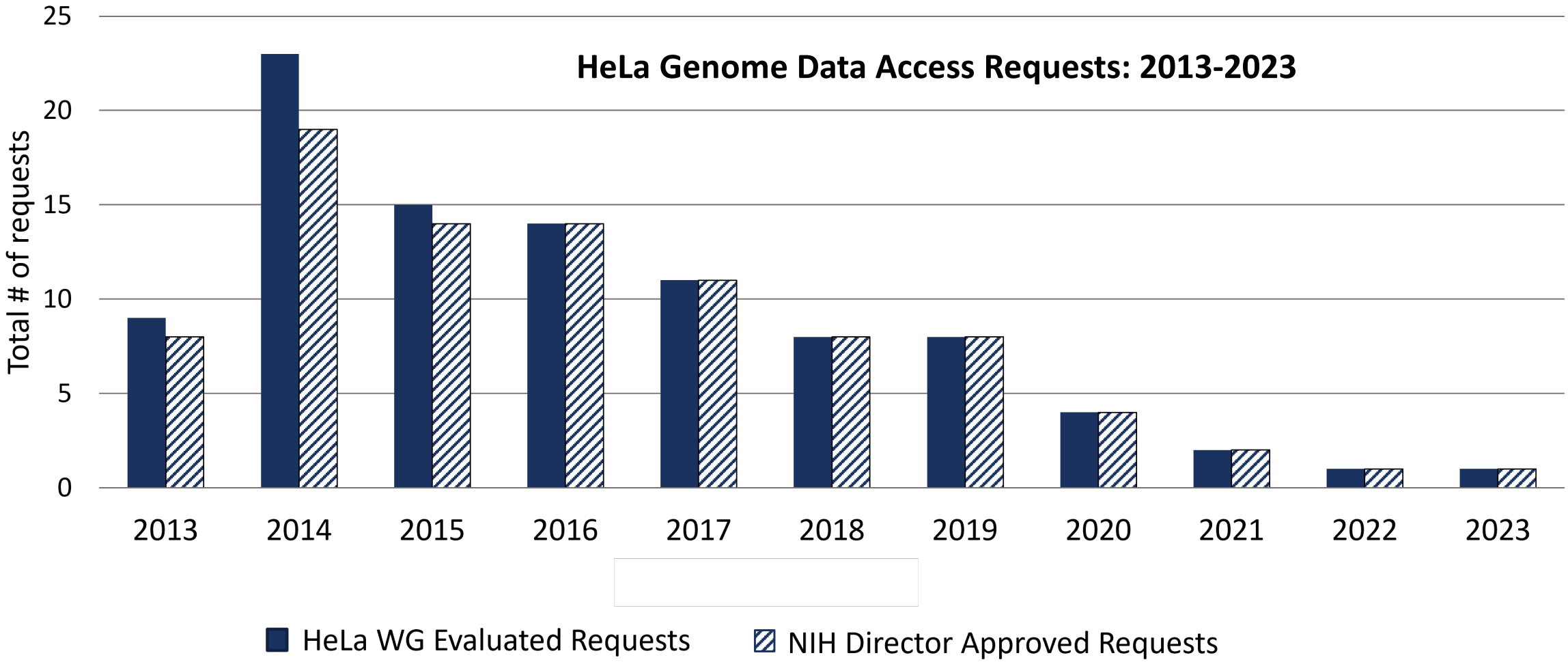
- **Publications citing datasets continue to be highly cited**



*RCR represents a citation-based measure of scientific influence calculated as citations/year normalized to the citations/year received by NIH-funded papers in the same field and published in the same year. https://dpcpsi.nih.gov/sites/default/files/iCite%20fact%20sheet_0.pdf

ANALYSES

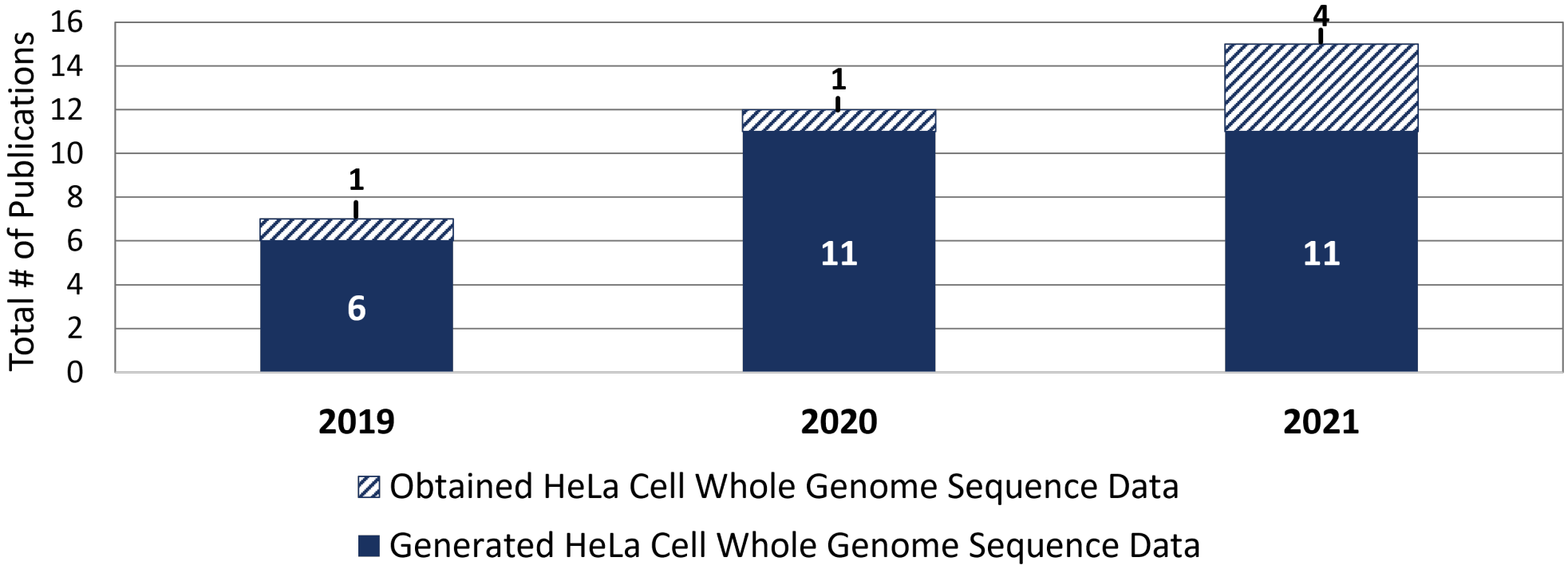
DIMINISHING DATA ACCESS REQUESTS TO NIH



ANALYSIS

HELA SEQUENCE DATA CONTINUE TO PROVIDE VALUE

- Publications citing NIH funding continue to report generation/use of HeLa cell whole sequence data
- **IMPLICATION: Decreasing awareness of Agreement and NIH expectations**



NEXT STEPS

NIH AND THE LACKS FAMILY REAFFIRM COMMITMENT

"We are committed to upholding our beloved Henrietta Lacks' legacy and continuing our partnership with the National Institutes of Health to assist respectfully in propelling science and technology forward."

- Descendants of Mrs. Lacks

August 24, 2023

10 years in, NIH-Lacks Family partnership holds strong

This August marks 10 years since the launch of a landmark partnership between the National Institutes of Health and descendants of Mrs. Henrietta Lacks. Through the 2013 NIH-Lacks Family agreement, NIH continues to honor the preferences of the Lacks family while responsibly advancing scientific discovery. Over the past decade, 90 researcher requests have been approved for controlled access to HeLa cell whole genome data and researchers have deposited 14 HeLa cell genome sequences in NIH's database of genotypes and phenotypes (dbGaP). To honor Mrs. Lacks' contribution to and her family's continued support of biomedical research, NIH also has a dedicated website that serves as a transparent, accessible resource to the public, scientific researchers, and the Lacks' family that is in keeping with the spirit of this historic agreement.

NEXT STEPS

OPPORTUNITY TO REINVIGORATE THE AGREEMENT

- **NIH is fully committed to upholding the NIH-Lacks Family Agreement**
  - Reinvigorating commitment to responsible data submission and use
  - Increased outreach efforts at scientific conferences and other venues to promote awareness
  - Updating NIH guidance and other resources such as sample Data Management and Sharing Plans for applicants planning to generate HeLa cell whole sequence data

- **As NEEDED: Convene ACD on an *ad hoc* basis (spring and fall) to minimize delays for researchers seeking access – FOR DISCUSSION**