

# HeLa Genome Data Access Working Group

## Report to the Advisory Committee to the Director

June 14, 2018

### **Carrie D. Wolinetz, Ph.D.**

Acting Chief of Staff  
Associate Director for Science Policy  
National Institutes of Health

### **Spero M. Manson, Ph.D.**

Distinguished Professor of Public Health and Psychiatry  
Director, Centers for American Indian and Alaska Native Health  
The Colorado Trust Chair in American Indian Health  
Associate Dean for Research at the Colorado School of Public  
Health  
University of Colorado

# The HeLa Genome Data Use Agreement

Per the agreement between NIH and the Lacks family, NIH is requesting that ***all researchers***:

- Apply for access to HeLa whole genome sequence in the database of Genotype and Phenotype (dbGaP)
- Abide by terms outlined in the HeLa Genome Data Use Agreement, such as:
  - Data can only be used for biomedical research only; this does not include the study of population origins or ancestry
  - Requestors are not to make contact with the Lacks family
  - Requestors are to disclose any commercial plans
  - Requestors are to include an acknowledgment in publications and presentations
- Deposit future whole genome sequence data into dbGaP

# Role of HeLa Genome Data Access Working Group

- Evaluate requests to access HeLa cell genome data in dbGaP for consistency with the terms of the HeLa Genome Data Use Agreement
  - Evaluation not based on scientific merit
- Report findings to the Advisory Committee to the Director
- Make recommendations to the ACD on changes to the terms specified in the HeLa Genome Data Use Agreement

# HeLa Genome Data Access Working Group Roster

## **Spero M. Mason, Ph.D. (Co-Chair)**

Distinguished Professor of Public Health and Psychiatry  
Director, Centers for American Indian and Alaska Native Health  
The Colorado Trust Chair in American Indian Health  
Associate Dean for Research at the Colorado School of Public Health  
University of Colorado

## **Carrie D. Wolinetz, Ph.D. (Co-Chair)**

Acting Chief of Staff  
Associate Director for Science Policy  
National Institutes of Health

## **Russ B. Altman, M.D., Ph.D.**

Professor, Bioengineering, Genetics, & Medicine  
Director, Biomedical Informatics Training Program  
Stanford University

## **Ruth Faden, Ph.D., M.P.H.**

Philip Franklin Wagley Professor in Biomedical Ethics  
Director, Johns Hopkins Berman Institute of Bioethics  
Bloomberg School of Public Health  
Johns Hopkins University



## **David Lacks Jr.**

Representative, Henrietta Lacks Family

## **Jeri Lacks-Whye**

Representative, Henrietta Lacks Family

## **Richard M. Myers, Ph.D.**

President, Director and Faculty Investigator  
HudsonAlpha Institute for Biotechnology

## **Robert L. Nussbaum, M.D.**

Chief Medical Officer  
Invitae Corporation

## **Veronica Spencer**

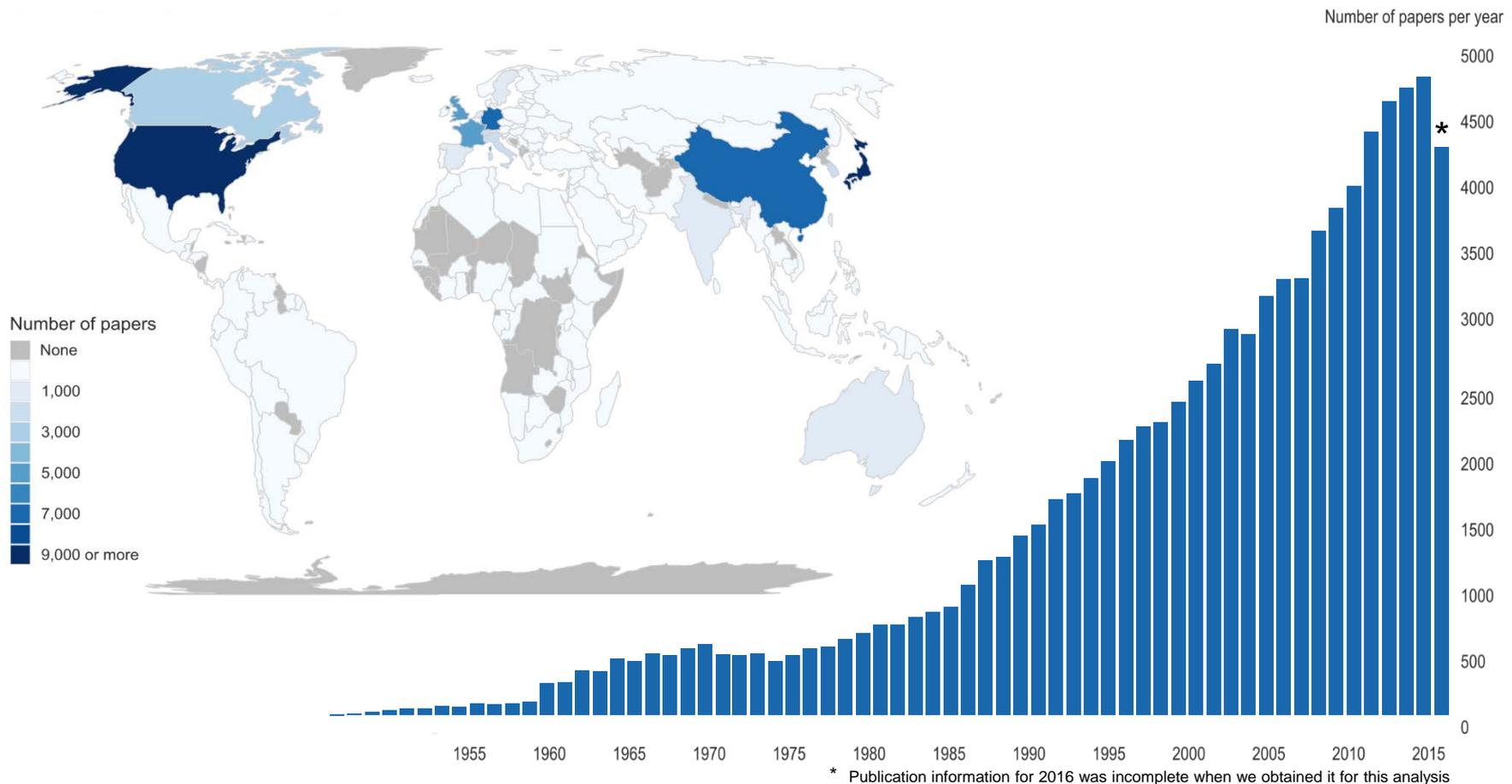
Representative, Henrietta Lacks Family

# 2014 NIH Workshop: *Scientific and Ethical Issues Related to Open-Access HeLa Genomic Data*

- Purpose:
  - Explore scientific and ethical questions about open-access HeLa genomic data
  - Discuss the pros and cons of prospectively applying the current HeLa whole genome sequence policy to other HeLa genomic data types
- Outcomes:
  - Scope of the HeLa Genome Data Policy to remain as is
  - **NIH to disseminate information on the state of the science using HeLa cells**
- September 2014 ACD recommendations to the NIH Director:
  - NIH should not change the HeLa genome data policy
  - NIH should hold a special session at a national scientific meeting that would focus on revolutionary research using HeLa cells

# The Contribution of HeLa Cell Research to Biomedical Science

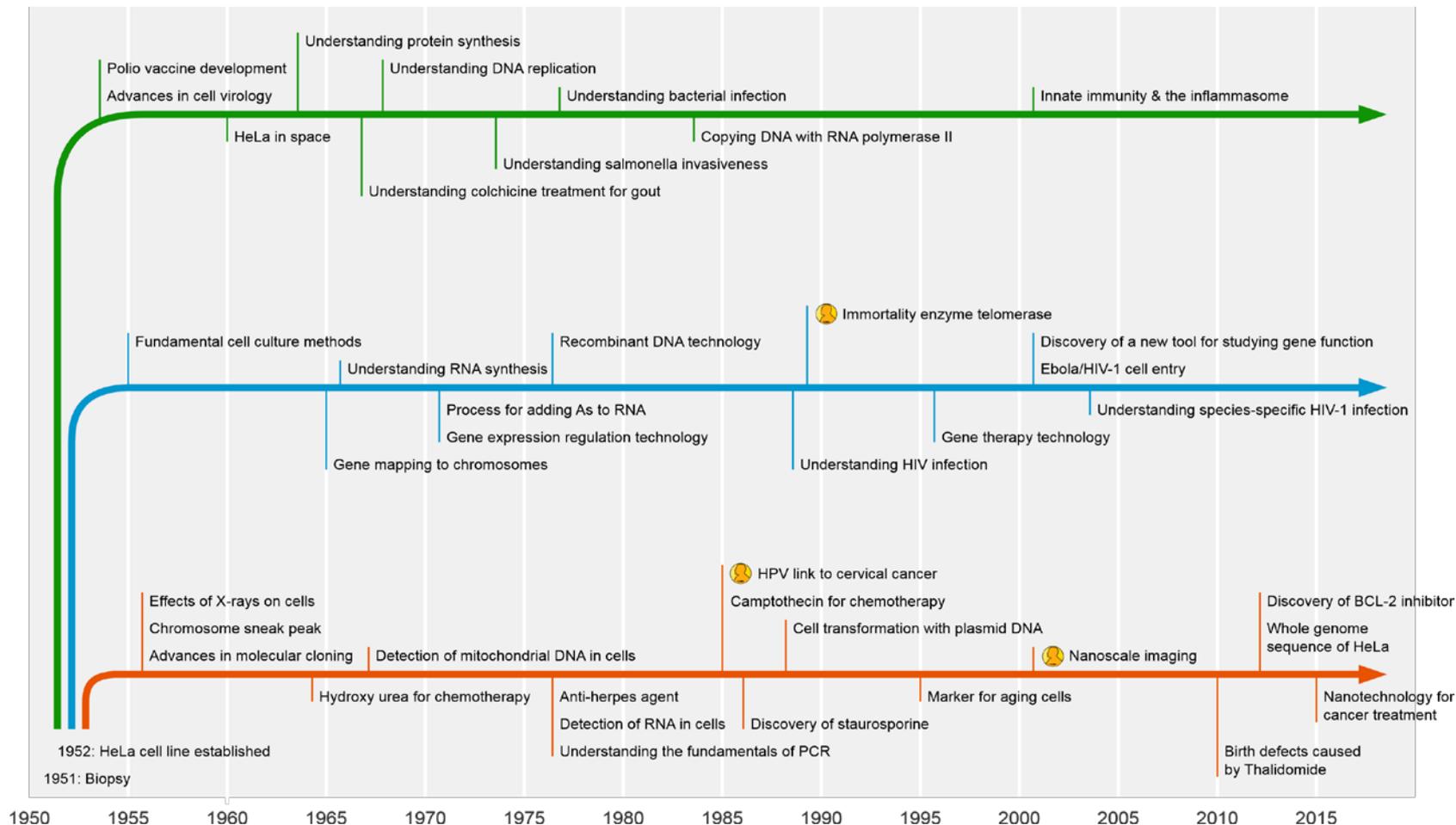
Over 96,000 publications resulting from research using HeLa Cells representing 142 countries (1953 -2016)





# Major Discoveries Using HeLa Cells

- █ Cell physiology and disease
- █ Molecular biology and genetics
- █ Research methods and cancer research
-  Nobel prize-winning discovery



# HeLa Whole Genome Sequence Data in dbGaP

Principal Investigator	Institution	Project Title	Submission Year
Wataru Yoshida	Tokyo University of Technology, Japan	Identification of G-quadruplex clusters by high-throughput sequencing of whole genome amplified products with G-quadruplex ligand	2018
Andrew Adey	Oregon Health and Science University, USA	Construction of thousands of single cell genome sequencing libraries using combinatorial indexing	2017
Jay Shendure	University of Washington, USA	Massively multiplex single-cell Hi-C	2016
Xun Xu	BGI-Shenzhen, China	Full-length single-cell RNA-seq applied to a viral human cancer: Applications to HPV expression and splicing analysis in HeLa S3 cells	2016
Erez Aiden	Baylor College of Medicine, USA	A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping	2016
Jay Shendure	University of Washington, USA	Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions	2014
Jay Shendure	University of Washington, USA	The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line	2013
Lars Steinmetz	European Molecular Biology Laboratory	The Genomic and Transcriptomic Landscape of a HeLa Cell Line	2013

# Working Group 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?

# Types of Findings Reported by the Working Group

In evaluating a Data Access Request, the Working Group will report a finding as:

- **Consistent** with the Data Use Agreement
- **Inconsistent** with the Data Use Agreement
- **Conditional** (will be consistent with the Data Use Agreement if NIH staff find that additional information obtained from the Requestor is satisfactory)
- **Pending** (will require a re-evaluation from the Working group once additional information is obtained from the Requestor)

# Status of Data Access Requests

Number of Requests	Status
80	Evaluated by the HeLa Genome Data Access Working Group
66	Approved by NIH Director
1	Disapproved by NIH Director
7	Disapproved by NIH staff (requestors did not respond to requests for clarifications regarding publication plans, IP, and/or the non-technical summary)
Number of New Requests	Status
6	Being reported to ACD today

# Working Group Findings: Evaluation of Access Requests

Since the last ACD meeting, the Working Group has found 6 requests to be consistent with the HeLa Genome Data Use Agreement:

Project Title	Requestor's Affiliation	Project Overview	Working Group Findings
<b>Analysis of Chromatin Organization in HeLa Cells</b>	<b>University of North Carolina-Chapel Hill</b>	<ul style="list-style-type: none"> <li>• In the cell nucleus, DNA is tightly folded around proteins called histones and together form a structure called chromatin. The physical properties of chromatin (e.g. the tightness of compaction and proximity of different sections of chromatin to one another) has been shown to exert control over gene expression implicating chromatin organization in normal biological processes and disease.</li> <li>• The investigator proposes to use the HeLa cell genome sequence to improve upon existing software to detect the physical properties of chromatin (e.g. mark the frequency by which discrete sections of chromatin interact in proximity with one another) to provide a resource tool to the scientific community that will allow investigators to query the impact certain regions of chromatin have on one another in normal biological processes and disease.</li> </ul>	<b>CONSISTENT WITH DATA USE AGREEMENT</b>
<b>Genome-Editon of HeLa Cells for the Study of Genes Involved in Endocytosis</b>	<b>Catholic University of Louvain</b>	<ul style="list-style-type: none"> <li>• Endocytosis is an essential cellular process required for the uptake of nutrients and proteins from the cell environment into the cell. HeLa cells have been used to investigate the biological properties of the endocytic pathway.</li> <li>• The investigator proposes to use the HeLa cell genome to design molecular editing tools, such as the CRISPR-Cas9 editing protein complex, to investigate how endocytic pathways dynamically operate in HeLa cells.</li> </ul>	<b>CONSISTENT WITH DATA USE AGREEMENT</b>

# Working Group Findings: Evaluation of Access Requests

Project Title	Requestor's Affiliation	Project Overview	Working Group Findings
<b>Finessing Predictors of Cognitive Development (part 2)</b>	<b>University of Auckland</b>	<ul style="list-style-type: none"> <li>• In the cell, DNA is compacted into organized structures called chromatin, which form 3-D architectural structures within the cell. Previous studies have shown that manipulation of the 3-D genomic architecture can influence normal development as well as disease.</li> <li>• How genomic architecture influences cognitive development is poorly understood. The investigator proposes to characterize the regulatory networks associated with cognitive development by comparing the HeLa cell genome sequence that dictates the 3-D architecture of the HeLa genome ("Hi-C sequence") to brain cell genomic architectural data to identify regions of the genome that play a role in cognitive development.</li> </ul>	<b>CONSISTENT WITH DATA USE AGREEMENT</b>
<b>The Role of Architectural Proteins in Shaping the Promoter Interactome</b>	<b>Babraham Institute</b>	<ul style="list-style-type: none"> <li>• Previous work has shown that specific DNA conformations influence transcriptional activity, an early step in gene expression, such as topologically associated domains (TADs) and specific DNA looping interactions between DNA fragments and promoters, the start site of transcription.</li> <li>• The investigator proposes to use the HeLa cell genome sequence, in particular the sequence that describes the 3D chromatin architecture of the HeLa genome ("Hi-C sequence"), to use as a reference to validate their approach that seeks to identify critical regions that control DNA conformation (i.e. DNA looping and TADs), transcriptional activity, and gene expression in HeLa cells.</li> </ul>	<b>CONSISTENT WITH DATA USE AGREEMENT</b>

# Working Group Findings: Evaluation of Access Requests

Project Title	Requestor's Affiliation	Project Overview	Working Group Findings
<p><b>RNA-editing Analysis in Hela S3 Cells by Single-cell RNA-seq and Machine Learning</b></p>	<p><b>Shenzhen BGI Technology Company</b></p>	<ul style="list-style-type: none"> <li>• RNA editing is a molecular process that can make discrete changes to specific nucleotide sequences within a RNA molecule after it has been generated from DNA.</li> <li>• The investigator requests access to HeLa cell genome sequence to develop a computational tool that discovers locations within the genome that have undergone RNA editing. Additionally, the investigator proposes to use single-cell RNA sequencing technologies to identify unique RNA editing sites in HeLa cell populations and compare them to HeLa cell genome sequence.</li> </ul>	<p><b>CONSISTENT WITH DATA USE AGREEMENT</b></p>
<p><b>Role of Nuclear Structure in Transcription Regulation</b></p>	<p><b>Duke University</b></p>	<ul style="list-style-type: none"> <li>• In the cell, DNA is tightly compacted around proteins called histones. Together, the compacted DNA and histones form a structure called chromatin. Special nuclear proteins associate with the chromatin to modify its overall architecture, which influences gene expression by controlling how regions of chromatin interact with each other and the genes in those regions.</li> <li>• The investigator proposes to use the HeLa cell genome sequence, in particular the sequence that dictates the 3D architecture of the HeLa genome (“Hi-C sequence”), to evaluate the role of nuclear architectural proteins that associate with chromatin and change gene expression.</li> </ul>	<p><b>CONSISTENT WITH DATA USE AGREEMENT</b></p>

# ACD Discussion, Vote, and Recommendations

---

# Working Group Findings: Evaluation of Access Requests

<b>Project Title</b>	<b>Requestor's Affiliation</b>	<b>Working Group Findings</b>
<b>Analysis of Chromatin Organization in HeLa Cells</b>	<b>University of North Carolina-Chapel Hill</b>	<b>CONSISTENT WITH DATA USE AGREEMENT</b>
<b>Genome-Editing of HeLa Cells for the Study of Genes Involved in Endocytosis</b>	<b>Catholic University of Louvain</b>	<b>CONSISTENT WITH DATA USE AGREEMENT</b>
<b>Finessing Predictors of Cognitive Development (part 2)</b>	<b>University of Auckland</b>	<b>CONSISTENT WITH DATA USE AGREEMENT</b>
<b>The Role of Architectural Proteins in Shaping the Promoter Interactome</b>	<b>Babraham Institute</b>	<b>CONSISTENT WITH DATA USE AGREEMENT</b>
<b>RNA-editing Analysis in Hela S3 Cells by Single-cell RNA-seq and Machine Learning</b>	<b>Shenzhen BGI Technology Company</b>	<b>CONSISTENT WITH DATA USE AGREEMENT</b>
<b>Role of Nuclear Structure in Transcription Regulation</b>	<b>Duke University</b>	<b>CONSISTENT WITH DATA USE AGREEMENT</b>