HeLa Genome Data Access Working Group

Report to the Advisory Committee to the Director

December 8, 2016

Kathy Hudson, Ph.D.
Deputy Director for Science, Outreach, and Policy
National Institutes of Health

Clyde Yancy, M.D., MSc
Vice-Dean, Diversity & Inclusion
Professor in Medicine-Cardiology and Medical Social Sciences, Chief, Division of Medicine-Cardiology
Northwestern University
Feinberg School of Medicine

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

3

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

4

HeLa Genome Data Access Working Group Roster

Clyde Yancy, M.D., MSc (co-chair)
Professor in Medicine-Cardiology and Medical Social Sciences
Chief, Division of Medicine-Cardiology
Northwestern University
Feinberg School of Medicine

Kathy Hudson, Ph.D. (co-chair)
Deputy Director for Science, Outreach, and Policy
National Institutes of Health

Russ Altman, M.D., Ph.D.
Professor, Bioengineering, Genetics, & Medicine
Director, Biomedical Informatics Training Program
Stanford University

Lisa Cooper, M.D., M.P.H.
James F. Fries Professor of Medicine and Director, Johns Hopkins Center to Eliminate Cardiovascular Health Disparities
Johns Hopkins University School of Medicine

Ruth Faden, Ph.D., M.P.H.
Philip Franklin Wagley Professor in Biomedical Ethics
Director, Johns Hopkins Berman Institute of Bioethics
Johns Hopkins University

David Lacks Jr.
Representative, Henrietta Lacks Family

Richard Myers, Ph.D.
President, Director and Faculty Investigator
HudsonAlpha Institute

Robert Nussbaum, M.D.
Chief Medical Officer
Invitae Corporation

Veronica Spencer
Representative, Henrietta Lacks Family

5

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

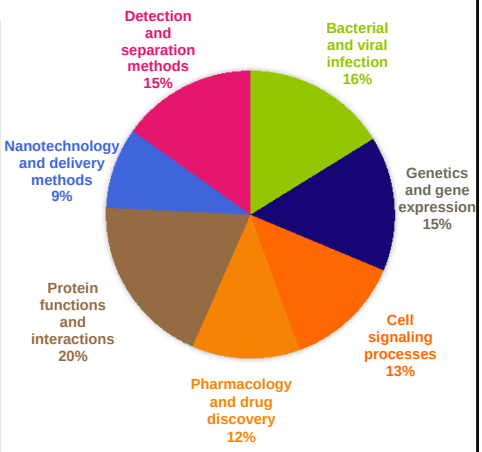
- Purpose:
  - Explore scientific and ethical questions about open-access HeLa genomic data
  - Discuss the pros and cons of prospectively applying the current HeLa whole genome sequence policy to other HeLa genomic data types
- Outcomes:
  - Scope of the HeLa Genome Data Policy to remain as is
  - NIH 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

6

Information Sharing: The Contribution of HeLa Cells to Science

Publications Resulting from Research Using HeLa Cells (January – July 2016):

- 1,500 publications identified
- 150 publications resulting from NIH-funded research, and that directly used HeLa cells and/or HeLa cell genome sequence



7

Special Session at a National Scientific Meeting: 2016 Annual Biomedical Research Conference for Minority Students (ABRCMS)

- Plenary Session: *An Update on the NIH-Lacks Family Partnership*
  - Jeri Lacks-Whye, granddaughter of Henrietta Lacks
  - David Lacks, Jr., grandson of Henrietta Lacks
  - Dina Paltoo, Office of Science Policy, NIH
- Scientific Session (Sponsored by American Association for the Advancement of Science): *HeLa Cells and the Future of Biomedical Science*
  - Erez Aiden, Assistant Professor, Baylor College of Medicine
  - Andrew Adey, Assistant Professor, Oregon Health and Science University

Both investigators have HeLa genome sequence data in dbGaP!!

8

HeLa Whole Genome Sequence Data in dbGaP

- Currently, five datasets in dbGaP:
  - The Genomic and Transcriptomic Landscape of a HeLa Cell Line
  - The Haplotype-Resolved Genome and Epigenome of The HeLa Cancer Cell Line
  - HeLa S3 (CCL-2.2) HiC Sequencing
  - Full-Length Single-Cell RNA-seq Applied to HeLa S3 Cells
  - High Resolution Maps of the HeLa 3D Genome Using Hi-C *(newest dataset)*

9

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?

10

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)

11

Status of Data Access Requests

Number of Requests	Status
61	Evaluated by the HeLa Genome Data Access Working Group
47	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

12

Working Group Finding: 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
Mobile Genetic Element Insertion Profiling	**Inserm Marseille**	• **The investigator previously conducted research that focused the mapping of locations within the HeLa cell genome at which mobile genetic elements (also known as "jumping genes") tend to insert themselves, in order to understand their contribution to biology and disease.** • **In the proposed research, the investigators will integrate their mobile genetic element insertion maps with data obtained by the HeLa Cell Genome Sequencing Studies and publicly available data from the ENCODE project to generate more precise maps that will help them understand why mobile genetic elements prefer to insert themselves into specific sites.**	**CONSISTENT WITH DATA USE AGREEMENT**
CHC22 Variation in Hela Cells (Collaborate project from three investigators = three separate data access requests)	**University of California, San Francisco** **University College, London**	• **The investigator previously identified genetic variants in a gene called CLTCL1, which produces a protein that controls how cells respond to insulin and remove glucose from the bloodstream.** • **The investigator proposes to characterize the role of this protein in glucose removal in HeLa cells, and, by analyzing the HeLa genome, learn how the HeLa specific variants within CLTCL1 may relate to the protein's function.**	**CONSISTENT WITH DATA USE AGREEMENT (all three requests)**

13

Working Group Finding: Evaluation of Access Requests

Project Title	Requestor's Affiliation	Project Overview	Working Group Findings
Genetic Modifications of HeLa cells for the Analysis of Cell Cycle Progression	University of Oxford	• The investigator uses HeLa cells to study how chromosomes properly segregate during the cell division process called mitosis. Aneuploidy, a condition where chromosomes are either lost or gained in cells during mitosis, is a consequence of faulty chromosome segregation and is considered a hallmark and driving force for the formation of tumors. • The investigator proposes to use the HeLa genome sequence to inform the design of genetic editing tools that will disable key proteins involved in the cell division process in order to understand their role in guiding appropriate chromosomal segregation during mitosis.	CONSISTENT WITH DATA USE AGREEMENT
Analysis of the HeLa Proteome Using Data Independent Acquisition Mass Spectrometry	University of Washington	• Analytical techniques used to identify and characterize the proteins present in a biological sample rely on having a corresponding high-quality genomic sequence as a reference, such as the HeLa genome sequence. • The investigator proposes to use the HeLa cell genome sequence data to improve the annotation of the host of proteins present in HeLa and advance development of a unique software that will be able to test new hypotheses.	CONSISTENT WITH DATA USE AGREEMENT

14

ACD Discussion, Vote, and Recommendations

15

Working Group Finding: Evaluation of Access Requests

Project Title	Requestor's Affiliation	Working Group Findings
Mobile Genetic Element Insertion Profiling	**Inserm Marseille**	**CONSISTENT WITH DATA USE AGREEMENT**
CHC22 Variation in Hela Cells (Collaboration between three investigators on this project = three separate requests)	**University of California, San Francisco** **University College London**	**CONSISTENT WITH DATA USE AGREEMENT** **(all three requests)**
Genetic Modifications of HeLa cells for the Analysis of Cell Cycle Progression	**University of Oxford**	**CONSISTENT WITH DATA USE AGREEMENT**
Analysis of the HeLa Proteome Using Data Independent Acquisition Mass Spectrometry	**University of Washington**	**CONSISTENT WITH DATA USE AGREEMENT**