PLEASE FILL OUT AND RETURN THIS FORM TO: *Center for Public Genomics, Duke University; c/o Susan Brooks; Center for Genome Ethics, Law, and Policy; 304 Research Drive, Box 90141; Durham, NC, 27708.* ***OR:*** *You may fax it to us at (U.S.) 1-919-668-0799.*

Interviewee Information. *Please list an address where we can contact you.*

Full name: LaDeana W. Hillier Date of interview: April 5, 2012

Current institutional affiliation: University of Washington Genome Sciences

Street Address: Foege Building S-250, Box 355065 3720 15th Ave NE, Seattle WA 98195-5065

Phone: (206) 221-7377 Email address: lhillier@uw.edu

Interviewer Information.

Full name(s): Robert Cook-Deegan and Kathryn Maxson

Affiliations(s): Duke University

I, the undersigned, have read the above, and I **AGREE** to release my interview materials, subject to any restrictions listed below:

(A) ☑ I place **no restrictions** on my interview materials.

OR

(B) ☐ My interview materials may be reviewed, used, and quoted by the researchers affiliated with the Center for Public Genomics, Duke University; *and in addition* (check all that apply):

☐ Researchers unaffiliated with the Center for Public Genomics may **read** the interview transcript and any related documents only after obtaining my permission.

☐ Researchers unaffiliated with the Center for Public Genomics may **quote** from the interview only after obtaining my permission.

☐ Researchers unaffiliated with the Center for Public Genomics **DO NOT HAVE** my permission to **read or quote** from the interview.

Posting interview materials to public digital archives: In spite of any restrictions listed above, I give permission for my interview materials to be made publicly available on the Internet by deposit in an institutionally affiliated archive:

☐ 1 year from the date of this form

☐ 5 years from the date of this form

☐ 10 years from the date of this form

☐ 25 years from the date of this form

☐ After my death

☐ Other: _____ (please specify a date or condition)

☐ NEVER: MAY NOT BE DEPOSITED IN A PUBLIC ARCHIVE

Please specify any further restrictions in the space below:

_____

Signature: LaDeana W. Hillier Date: 10/15/2012

Robert Cook-Deegan <bob.cd@duke.edu> April 9, 2012 2:24 PM
To: Kathryn Maxson <kat.maxson@duke.edu>
Cc: LaDeana Hillier <lhillier@watson.wustl.edu>, Rachel Ankeny <rachel.ankeny@adelaide.edu.au>
Reply-To: bob.cd@duke.edu
Re: analysis meeting notes

LaDeana,

The meeting can't have been conducted under the formal Chatham House Rule (see wikipedia explanation: http://en.wikipedia.org/wiki/Chatham_House_Rule), since there is a public formal list of who was there. (The formal rule is to make it so no one can deduce who was there and no one can say, in addition to attribution of quotes to individuals.)

But folks often refer to CHR informally as a set of rules in which you can cite *what* is said, but not identify *who* said it unless you have permission.

In this case, I think you are free to share notes and your own thoughts freely, but you can't attribute quotes or positions to others unless you (or we) get subsequent approval. FWIW, many others have shared notes, and we got agenda and list of attendees and notes through a Freedom of Information Act request. Many individuals have also given us their notes and photos and such.

We're hoping to pool what we have with Wellcome itself at some point, and either give the public materials to them to post, or post them here with links. They're still trying to free up the transcript. The Wellcome Library wants to share it, but at this point, it's property of the Wellcome Trust operational side, so far as we understand (Michael Morgan is trying to manage this process).

Since there's a bit of ambiguity about the rules, I think it's mainly a matter of common sense and judgment. I hope that helps.

BCD

On Fri, Apr 6, 2012 at 1:39 PM, Kathryn Maxson <kat.maxson@duke.edu> wrote:

> Dear LaDeana,
> Wow, thanks so much! It was a pleasure talking to you too, and this project has indeed been a fascinating ride. Your story is particularly interesting to us, as you were involved at the crossroads of so many different genome projects.
>
> I'll let Bob answer your question about Chatham House Rules, as I simply don't know well enough to advise you one way or another, so I'll refer you to the expert!
>
> All best, and thanks again! You should have your transcript in a week or so.
>
> Kathryn
>
> On Apr 6, 2012, at 12:59 PM, LaDeana Hillier wrote:
>
> > Greetings -
> >
> > Thanks for taking time to talk yesterday. It must be so interesting
> > to hear how accounts of the meeting vary between people and to get a glimpse
> > at how memories are affected by time and by the perspective we've
> > gained based on the things that have happened since that time.
> >
> > This was certainly an absolutely amazing time to be alive and
> > a spectacular project to be a part of. I am incredibly fortunate.
> > As I mentioned, getting to work with the people in this project,
> > going to meetings with them, racing for deadlines with them, it has
> > all been a total privilege.
> >
> > Here are my notes from the January, 1996, analysis meeting we had at
> > the Sanger.
> >
> > <analysis.jan1996.txt.gz>
> >
> >
> > I have my outlines for my talks from the 1996 and 1998 Bermuda meetings.
> > Because of the Chatham House rules, does that mean I'm not supposed to share
> > those with you? Or because they're my outlines, then I'm allowed to give them to you?
> >
> > Thanks,
> > LaDeana

Kathryn Maxson, B.S.
Research Aide, Duke University
Genome Ethics, Law, & Policy in the IGSP
http://www.linkedin.com/pub/kathryn-maxson/23/371/b04

--
BCD
==========
Robert Cook-Deegan, MD
Genome Ethics, Law & Policy
Institute for Genome Sciences & Policy and
Sanford School of Public Policy
Duke University, Box 90141
304 Research Drive
Durham, NC 27708-0141
919.668.0790
gelp@duke.edu
http://www.genome.duke.edu/directory/faculty/cook-deegan/

SEQUENCE HANDLING (Part I of INFORMATICS)

==========

GOAL: Reduce personnel time by reducing human decision

making without sacrificing data quality

I. TOOLS FOR SEQUENCE HANDLING

A. Data collection

1. image analysis

GETLANES

RETRAK

TPP

PHRED (base quality / error distribution estimates from phred /phrap he worked on 14 of our cosmids to get them)

This table shows the quality / error distribution in 14 automatically

assembled cosmids. Overall the error rate is 1 per 11kb, with the bulk of

the errors being in the phrap low-quality regions. These datasets include

finishing reads but are completely unedited.

Restriction digests used as a check of clone fidelity and then of the resulting assembly and each clone represents a single haplotype

QUALITY/ERROR DISTRIBUTION

IN 14 COSMIDS

(6 mammalian, 8 C. elegans)

Phrap

Quality Bases Errors Error rate

0-10 886 (.17%) 29 1/31 bp

11-20 1679 (.32%) 8 1/210 bp

21-30 5101 (.97%) 4 1/1.3 kb

31-40 17106 (3.20%) 4 1/4.3 kb

> 40 502030 (95.30%) 3 1/167 kb

==========

Total 526802 (100.0%) 48 1/10.9 kb

Current estimate of error rate based on C. elegans comparisons to mRNA sequence in GenBank:

TOTAL BASES: 175123 Discrepancies: 251 Of those: 12 errors

ERROR RATE: 1/14593 bp

Estimate based on yeast comparisons:
In 176,521 bp there were 13 differrences. I judged we're wrong for 3 of them; I'm certain of our sequence for the others. I think the Europeans agreed that they are wrong for 6 of them.

2. data handling

OTTO/ASP

3. prefinishing

FINISH

B. ASSEMBLY

1. engine - PHRAP

2. interface - CONSED

XGAP

C. DATA TRACKING

1. DACE

a. nightly cron jobs

b. UNIX spreadsheets

c. graphical interfaces (FREEZER, ORDERFIN)

2. STLACE

all analysis-related data

II. PLANNED AREAS OF DEVELOPMENT

A. Data Collection

1. reduce human manual scheduling

2. improved "finish" program

B. Assembly

1. better handling of larger inserts with complicated repeats

2. automated use of external mapping data

C. Data Tracking

1. bar codes

2. improved/specialized user interfaces

3. automation of workflow and scheduling

DATA RELEASE (Part II of INFORMATICS)

==========

GOAL: Immediate data release with local annotation

of the sequence

I. UNITS OF DATA RELEASED

A. units sequenced

BACs, PACs, cosmids, etc. (according to clone origin)

B. can be updated into larger contiguous pieces

C. large contiguous chunks of data should be made available,

with tools to pull out any size chunk through links

II. FORM OF DATA

A. represent sized gaps

B. include per base confidence measures in submissions

III. ANNOTATION

A. All features identified in a reasonably automatic fashion

1. Alu's and other defined repeats

2. simple repeats >80% conserved

3. EST identities

4. polymorphisms

5. internal repeats

6. signficiant blast similarities

7. exon based >95% confidence

B. Work towards making things more and more automatic and

more and more complete

IV. MECHANISM FOR ASSESSING ERROR RATES

A. local analysis and annotation important for error rate assessment

B. comparison against existing databank entries

C. other assemblies + auto-editing

D. confirmation of assembly using mapping data

1. STS data
2. restriction data
3. fingerprints
4. overlapping clones

E. public availability of raw trace data