Agenda

DOE/NIH Genome Informatics Meeting

Dulles Hilton Hotel
April 2-3, 1998

Thursday, April 2

9:00 am	Aristides Patrinos, Associate Director, Office of Biological and Environmental Research, DOE
9:10 am	Francis Collins, Director, National Human Genome Research Institute, NIH

Moderator: Aravinda Chakravarti

9:20 am	David Thomassen, DOE
9:30 am	Aravinda Chakravarti, Chair of the NHGRI Planning Subcommittee
9:40 am	LaDeana Hillier, Large-Scale Sequencing
10:00 am	Takashi Gojobori, DNA Data Base of Japan
10:20 am	Anne Spence, Medical Genetics
10:40 am	Break
11:00 am	Deborah Nickerson, Genetic Variation
11:20 am	Roger Brent, Functional Genomics
11:40 am	Rainer Fuchs, Industry
12:00 noon	Bettie Graham, Training
12:10 pm	Lunch
1:30 pm	Breakout groups
5:00 pm	Adjourn for day

Friday, April 3

Moderator: Aravinda Chakravarti

9:00 am	Reports from the four breakout groups
11:00 am	Break
11:20 am	David Lipman, National Center for Biotechnology Information
11:40 am	Ed Uberbacher, Annotation Consortium
12:00 noon	Lunch

Moderators: Elbert Branscomb and Eric Green

1:15 pm	Discussion of goals and priorities
2:30 pm	Break
2:45 pm	Continue discussion
4:00 pm	Adjourn

DOE/NIH Genome Informatics Meeting

Dulles Hilton Hotel
April 2-3, 1998

Purpose: to discuss the types of queries that will be important in genome informatics, and what types of data, tools, and databases will be needed to address them. The emphasis is on setting priorities for current and future user needs. The results of this meeting will contribute to the five-year plan for the Human Genome Project that DOE and NIH are formulating. The results will also influence the agencies' plans for informatics projects and funding.

Questions to address in talks and breakout groups:

1) **Queries:** What scientific questions will you want to answer? What types of data will you need to answer these questions? Which of these data types are permanent, which are temporary but important, and which will need to be regularly updated? What uses will you have for genomic sequence data in the next 5 years?

2) **Tools:** What protocols and tools for data submission, viewing, analysis, annotation, curation, comparison, and manipulation will you need to make maximal use of the data? What sorts of links among datasets will be useful?

3) **Infrastructure:** What critical infrastructures will be needed to support the queries you want to perform and what attributes should these infrastructures have? In what ways should they be flexible, and how should they stay current? How should they be maintained?

4) **Standards:** What kind of community-agreed standards are needed, e.g. controlled vocabularies, datatypes, annotations, and structures? How should these be defined and established?

First afternoon breakout groups: the first name is the suggested moderator:

Sequencing, mapping for sequencing, gene maps:
Raju Kucherlapati, LaDeana Hillier, Eric Green, David Lipman, Takashi Gojobori, Peter Schad, Elbert Branscomb, Ray Gesteland, David Smith, Peter Cartwright, Rainer Fuchs, Peter Weinberger

Gene finding, OMIM, variation:
Ken Buetow, David Nelson, Anne Spence, Jim Ostell, Aravinda Chakravarti, David Valle, Bob Cottingham, Bruce Weir, Deborah Nickerson, Chuck Langley, Stan Letovsky

Annotation, function:
Brian Chait, Roger Brent, Martin Ringwald, Joanna Amberger, Mark Boguski, Manfred Zorn, Ed Uberbacher, Chris Overton, Temple Smith, Richard Mural, David Balaban, Dixon Butler, Nat Goodman, Barbara Wold, Randall Smith

Comparative genomics:
Carol Bult, Michael Cherry, Tony Kerlavage, Jean-Francois Tomb, Terry Gaasterland, Frederique Galisson, Reinhold Mann, Janan Eppig, Bill Gelbart, Katie Thompson, Paul Gilna