April 22, 2011

Robert Cook-Deegan, M.D.
Duke University
Institute for Genome Sciences & Policy
Durham, NC 27708

Re: FOIA Case Number: 11-FOI-00165-NHGRI - 38639

Dear Dr. Cook-Deegan:

This is our final response to your March 3, 2011, Freedom of Information Act (FOIA) request addressed to the National Human Genome Research Institute (NHGRI), National Institutes of Health (NIH). You requested: 1) any notes (hand written or otherwise) taken at the 1996 International Strategy Meeting for Human Genome sequencing, and 2) any email or memo summaries of this meeting distributed within National Institutes of Health afterwards.

We searched the files of the NHGRI Office of the Director for records responsive to your request. That search produced 48 pages responsive to your request. Enclosed are: 1) notes taken at the 1996 International Strategy Meeting for Human Genome sequencing (45 pages), and 2) email/memo summaries of this meeting distributed within National Institutes of Health following the meeting (3 pages). A total of 48 pages are being released with this response.

In certain circumstances provisions of the FOIA and Department of Health and Human Services FOIA Regulations allow us to recover part of the cost of responding to your request. Because the cost is below the $25 minimum, there is no charge for the enclosed materials.

Thank you for your interest in the National Human Genome Research Institute.

Sincerely,

Christy Cecil
Freedom of Information Specialist, NHGRI

Enclosures – 48 pages
1. Software
   - Comptd. more action.
   - Retained.

2. Waterston
   - 10th anniv. 1983, 19th anniv. 1986
   - Compliments, technology & costs, not duplicating systems.

3. Presentations
   1. Waterston
      - Regional 575s. in line next BK-line (10x)
      - local very competitive & stable.
      - minimal delay.
      - Set, equipped, seduced.
      - Amendment, edit.
      - And mine the sheen for high quality sequence.
   2. Canada
      - Definite, 1 wk. beginning (17/9/76)
      - Goals: Some in next year for Waterston/Robin team.
      - Ann. at chr. 22 completion
      - Also x, 7, 6.

May direct targeted effort for.

2. London
   - Anchored. OKs. (c. 575s)
   - Suggest 3 seq. directed demol (underneath).

Distinctive Owens:
   - Fort esch m. (c. 575, mg. of BKs, demonstrated) select. 5 seq. Squadron.
   - Hotel to continue.
3. Vadder
   Cell line BACs → 29
   BAC end sequencing:
   BACs: 180k to 250k
   Fragments sequenced: 5% (99.96% average) 120,000
   Total sequencing: 100,000 fragments (10%)
   Partial and clone sequence <1 yr to 5x ABI
   → Fast next run in taking path

4. Roed
   Alloc NC3D
   C.C. done
   BAC's + 2x BACS sequenced
   Ream 100% results in pair sequenced 35 BACs
   500 by read
   Replication error also an issue
   LINEs an issue
   11% of reads more than
   Sticking also gets into BACs
   Plans multiple participations

5. Wateri - ashic. Gajarean
   Japanese effort
   Process will fail in 5-yr effort "brain drain" 2 new labs
   STA - Ren for 5x6/yr. Nakani, Nakada, Sano, Kuma, Iaka
6. Castano
   - Chromosome 19
   - 17.0 Mb up + 5.0 Mb down
   - 80.0 Mb karyotyping to 1.40 Mb
   - Started sequencing a 200.0 Mb region
   - Shipped edited clones

   UNLV recognized $ sequence as the driver for in repeat pieces

7. Caibbe - X12
   - 6 notebooks, 25 people
   - Solves redundancy in insert mapping
   - Bars up Xp28, Xq22, Bar 1q26-40

8. Mayers
   - 88% of chr 11 in tandem 1B9E5
   - 100% size up
   - chr. 5 - 100% size up (Hain + 1ST)
   - PAC found a 1X SAGE
   - 0.8% only 20% sequence
     (98% clone coverage)
   - 100% up 2.0 Mb of 1bp
     Use of 1ST 15

9. Lahoud - German genome program
   - Will be assessed around functional info
   - By conference in second semester center
   - 70% map of X
     Working on 21
10. 
Averages
2 - day, 2 lanes
include on prior label
200 clone debaters
Dig and E0.6 bp - expect 1 - 2.5 kb
Infant genome is hard with the clone.
Plan to quote machine.

11. Wefan - whole genome sequencing
Our boy (2.5 kb) runs odd a whole.
Quality varies in each nucleotide.
Run for 10x coverage (25 kb of raw sq.) by 2000.
Plot multiple sources of DNA 3 gel polymorphism.
May be less computer simulations on assembly.
$0.005/bp raw sequencer.
× 10 = $0.05/bp (5xassembly)

5. More on sequencing - Really maps.
A. Selection
[clipper] YG complete to 91,000.

Wasting in clean of 6000 × 3000 years.

Finished 33,900 6000 years (1/5th average, 1/3rd x).

45% of gene in these data base match.
53% of seq. is nonsense, 21% coding.
Total gene count est. 14,000.

Whole BRL-40 strain from using the data before.

Improving efficiency.
Quality in worldwide.

* Gene rats in mids '92! All are low - relatively unsuitable.
Selon les informations, il devrait être réalisé un essai pour l'eau finale.

Gaps - O log-0.5
- e moléculaire (forme, diff. points)
- délire - 10% 4200 coulées (4x13, point 3)
-鞄 形状 will be asymptotic

B. Craig, Victor

- MTC 355, 600 in TB2, Web Site
- 19112 (11-79)
- 1
- 100
- 127 7.69
- 17.67
- KTC 3516

- 25 000 req genes 85-70% sequence much
- 172 7 genes
- 82 8 genes
- 82 8 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
- 172 7 genes
Date release:
- Publish when completed
- Think ftp site is a good idea
- Strain: prembriae - being double E. coli
- Transform: peblam (NAM10)
- S$3 =
- What to be part of NHLDRS?
- Hemocline virus & telephoto?
- Existing resources, how are they being used?
- C. myxus - Transform: breed evening of research
- I'm S$3
- Clu. 4: 20MB - not www yet, will be cancelled
- Bason 2: C3 = 700 STSs. NMR
- TNG: WIN available
- New 02/03 120K
- C54
- S$3: 4000 new
- 3000 integrated from others 8-10K/yr. to get 30K (700/ak)
- 1/2 DBHAE
- TNG's loaded on 1/2 region
- Will finish 69 at 1K
- Then on TNG (will do some 1K, plus add 2K more)
- Southern "There exist a single published
- Validation of chips for sequencing or sequencing validation"
D. Bentley 22, 6
5 Mb x 4^12
2 Mb x 4^25
3 Mb x 4^26

While chromosome Flow chart 47.47% mids (introns) Flank by Pacs
Every point Centrosome

Ch. 22 (With U/Sigma donor linked to Hpa2)

linked at 4 x 10^2

60% from within colonies at 8-250 Kbp
3 Mb ready for mapping

Ch. 6

Not marked by YAC MPs
1200 markers public domain
700 unlinked 679 RH mapped

Tied 1G on PACs 10-20 pm
Golde/573
26 deletion 4 deletion 573
E. Lester - Resources

Page: 16,500

- 5,000 AR addresses in Canada
- 1500 YAC addresses

May go to 20,000, ending 6/18

- 14x YAC 43
- 2x BAC 14
- 10% on the way to RG

- Rapid SRS content mapping

Funding to support collaboration, not for

- Initial development

F. de Jong

PACS pCYPAC2 11

PAC4 (17.5 kb) EBV, 11 kb, 5500 paiz/plot

BACs pBAC3.4 (11.4 kb) To 7 do call him, schedule

- PACS at set time?

- PACS 10x 35% empty wells

46 x 3

- PACS 3, 4, 5 aliquots 3x, 3x, 4x, 6x

New wood to BAC, expect 10x by CSH

Will make 1 plate with 1 wood.

- Send list for PACS - might want BAC's listed

- Sensibility?

- FAD BAC 6 9x1 in 12x6 at Greenup

- 22x6 in their hands

- Seq BACs 5 under "Telemer" form PDB

- gel purification of product

- Use 60% to determine Robin labelled grnt
G. Some
ACB: 215 x k 4/8
Some [hears]: 1 [word]
CH: 2 ACB cytop [4] 16 300x changes.
Level 05 GAC plan

Infused concept?

H. Even
14, 132 [words]: 5.5x coding
14, 777 T T and say...
- ELSANX, 4000 molecules
- VISI broad view [AC] pole (221)
12,138 class [sympathy]
10 once + 1 PAC required at high accuracy

M2 region sequences
- Likely unburst
- Rin for 2006 by June
- Did 2 classes already by former biology
- "and artistic..."- very high density 46-67.
- Gomu build 192 channel oligo again.
- PMAP & first claim
- Cost $4/batch - $1/oligo

LTI
III.

[Introduction] - Bob & John

1. Taming of Data Release - Immediate (1kb embed)
2. No potential of generic sequence & handled to
   80kb only and available for research work (today)
3. Things need speed

Finding agencies as needed to act policy

Our Father asks for

& Faster to NESR from R & R... - several suit data release policy
A larger game than before

A. Hood - Phil Green's

PhDQ (PhD = quality, more related) -> 1/5kb

Phil Q = ~10x?

Factors on for PhD, inadequate data editing
Highest quality case: 1/167kb

B. Wilson

Box calling will be able to do 4k on a 377

GETLINES (instruction AB5) now in max, available

FAX: 91

TPR - can degression break subatomic, ability needed
AB5 - upgrade available 6/96?
As of 41 - will be at 379M yr limited 70k/yr

If you don't receive this email, let helpful
first thing is for emails)
C. Subter (3-5 Papers)
2mb of 85 (1.5 mb actually) "Signet 99"
Cost ~ 20k (Street)

MCE - tracking system
Chem. 6 - Sangam Bank will start their own
Preliminary X20

Mean anomaly v.2
thread 0.2
substit. 2.7

7.6 public

\[ x_0 \cdot 10^3 \approx 3.6 \text{ km} \]

\[ r_7/98 \approx 190 \text{ mbs/yr. lower} \]

(That's sweet flying)

D. Gabbi

"Goby 11.7 - structural similarity to platista"
How will they be
2 days/priime energy, after
More emphasis, 1.5x less template

Terminators have been constructed
An earlier assay added to Moth siever, primer chemistry

E. Hawkins

10,000 columns/1d to go, do 3 Gs in 100,000
Sedlab phone, reasonable miniaturization

Sequence - DNAampl, set up, load cycling

\[ 6 \text{ hrs} / 20 \text{ plates} / 1 \text{ FTE} = 5000 \text{ samples/24h} \]

No user intervention
Flexible - M.5, M.7, phone

Call p. 6 12,000
CRS activated and commercially available. What
4 in each
120 plates caesal

KE
8 - 10 plates/day

M13 stocks (2K) - 1 X Dye Primer
PUC (58) - 1.2 X P/K

Finishing strategy = FINISHER, hands off automated
Run code everywhere
LabPure data management

GRACE - Gel analyzer, Morning P/Seed equal. Score 0-3
5 mm +/- 5 cm yld

Rainfall

F. # harvested - June
36 people (15 techs, 7 Ph.D., 12 techs, 2 managers)
12 ABI, 377
1 Lico

Strategy - Gel from 2nd or 3rd
100% dye sensitivity - better GC read

Cleanse - try 1 second, undo start over
Dental protractor

PCR amplification of product = clone from DNA pool (6-10 km)
3 yds/week/plot

Amp for 10-15 mb finished Sep/yr.
Goal: 3 mb = 2/21, 7 mb 3rd batch (and 1st/102)
7 mb = Xp11.2 x2
1.2 mb = ABI 9 x

Send 1-2 mb projects in 7, 11, 17 (02 form)
Ch. 21 Seq. Launch
10 groups to do 30-35 MB
in Spain
No lab at main & Fijian
5 ea.
5 ea.

Not targeting the 4 MB Dave vision (II)

PM I

III 5 MB
IV 0.8 MB PNE region

Fijian - Max line 200 KB
L1 - same as line & size

Report to Dr. 1
660 smail - how a change of 5 -> on change?!

Last - 1 May 69 year (m.41)
Why so many customers - just need to have
wells
Lots of PABs are there already
Expect to dump out the new PABs

C. Fox - ch. 22 & PNE 2
Strategy - ds vendors, FS, terminators
red tape for shipment

Economic - day 12 to read
Customer - excluding out 1 BACs, 2 BACs done
Puting them into, on Fijian
Dr. George - new line coming, ready for

Dry reg. 6 month
Marine related region  

failed 2 BACS

Staph n 3mb some  

(0.25 mb in Gen Bank) will be  

cloned now 80% OK

Special thanks by NASA to do S. pyogenes  

E. N. group for  

Load to get  

SVVA  

10,000 bp plasmid added  

2-3K sequence

H. Welker

Cohort output 1.2 m/year in 1996

Cost = $1/gene by end of 1996 (mmpth?)

Total genome shotgun  

4.5 x 10^7 fws in 99%

0.55 $/kb

Would like lower Tog price?  

We could use more

I. Kattar:

A1 21 = STA

Mut1 deletion  

1500 pb (2x)

\[ \text{Depot, insertion \rightarrow mutd deletion \rightarrow seg, cut, gel}\]

\[ \text{run pl controls \rightarrow make map} \]

\[ \text{software grants} \]

\[ \text{mill bp sequent} \]

\[ \text{1500 bp of AP} \]

Use family to do 1.1 mb of AP

\[ \text{Finish by end of May} \]
IV. Inferences

A. Adams – Annotation

GRAIL Handling

*Nice to integrate all of them*  

Who updates annotation?

- GenBank is willing to

**GSDB emphasizes community annotation**

B. Killier

Data not required (read from

Report phylogenies

Access error rates

1) Local analysis & annotation locally is helpful
2) Compare to existing entries
3) Alternative assembly & new sequence
4) Annotation using – STS, restriction,

Repeat, overlapping clones
5) Public availability

Assign a GSDB # to every clone?

C. Wilson – Large clones + small clones

Simulation of whole genome shotgun versus good biology

Different levels of confidence output from

ED: 4.6 x 10^4 if one has 10 repeats + ref genome should

yield 4 x 10^7 if one x 10^4 phylogenies
D. Bulletin
   Coding for accuracy
   Prepare/circulate future position
   Y/31 by
   Y/167Kb

   What about the preliminary data? Should also
   include
   Need to retain original CD file to have
   kept record of comparisons.
   Sending stories? - where to archive?
   Currently considered I can cut file

IV. Final Session
A. Data Release/ISR
   Release - automatic release of sequence >1kb
   preferably daily
   mutation/advances of finished
   annotated sequence
   available immediately
   appropriate access
   available for both research and development
   in order to maximize the benefit to society

   Primary focus on the cloning of
   additional experiment information that
   function or diagnostic utility.
There is the principle that genome sequences, containing any functional information about functional or diagnostic value, is not an appropriate subject for patent protection.
B. Coordination

Head (Seattle) - ch. 7

TCKx/β, FNC/2

Sierra/Adams - 30mb of 16p / 3yrs

Reactor - 30mb of 21p (including 15, 85, FNC) / 3yrs

+ XPG (6p & 6fms)

+ Xp11.23-24 + PRAP

+ 6 mm memo repair of 1-2 mb

Hydrophobic他也修

Meyers - 16 or 8p

20mb

16 p fault not being pursued by TCK/xp

SASE plans - ? amount

Total plan run 100mb (UPL 10)

20mb of food

MIT - Quite vague

105mb / 3yrs. winst ch. 17

Amidst from more recent

Calkins - XPG 29, XPG 27, 12.13

30mb / 3yrs

Roe - ch. 22, NFL - ch.

6 mb

Japan - not sure

Chm - 4mb of X

Carthage - ch. 19

50mb

Rite under 4 - 10mb / qr.

ch. 2 Latia

St. Louis - 50-100mb / 3yrs. (until 30)

ch. 7, 22, X
Melanie - ch. 12/18, 5 MB (of funded)
End may - present

Reps 5
550 MB/7 yrs, 150-250 MB/3 yrs.
22, 8, 6, 20, 1, 16

Water
A few MB of whole genome

Events
11/15, 5, 11/12, 11/23
2 MB in each -> 6 MB/2 yrs
+
Saradhan

Kudrange
?25 MB of ch. 12

LBL - Ch. 5 12 MB/3 yrs

Sahpur - Ch. 4

C. Fastig

1. NGAGE
2. Welcome - Space is enough to do all of it
End of June - occupancy
CJ 5 position end of 74
Endurred 6/77

Orthogas sequencing
Sanger - TB
Molecular chimerism

GETSFUN (can do molecules/fishes)
M fries in 9/6
Is UK well placed?

Beware of molecular biology, all biology
Protein expression, transcription

Budget
Sanger £910k/7 yrs.
£ 8-9 m... the biggie
£65k/yr... 3 yrs
3. EC

5 yr financial programme

Nov 94 - Apr 98 (FP4)
20 projects; $178 M US dollar
15 member states
+ new Israel - Norway - Iceland

Budget - Human Genome Research

$111 M/yr.

5 schemes -

- Human
- Human
- Human
- Human
- Human

Current spending $5.3 M on seq. related
Advisory group meets Thursday to decide on 1996-2002 (FP5)

4. France

"Serious discussion" of large scale sequencing effort

5. Germany - England

Next week 5 Sci Adm. Committee will meet
Some progress can be

Diagram:

[Diagram of DNA structure with labels and measurements]
6. DF (Smith)
   $10M/yr
   About $10M is going towards him
   Expect to see $40M/yr by the end of 3 yrs.

7. MRC
   Finds C. cheaper - got an extra
   $10M to bring it along

D. Plans
   Another meeting of core group in 1 year
   HUB has a web site on 6/18
   List current plans

E. Notification of plans to be sequenced
   Put on a Web site
   All of 22 will be on Super
   HUB Web page will point to where info is

F. Accuracy
   What info?
   do coverage?
   rule of thumb: on the diff. decisions

Also - do them at some frequency? RC said yes, monthly!
No real concern - done by next year? Don't want much
Gaps - how many?

- Technical gaps
- Biological gap - not represented in database

Is it right and then

Need more records in GC beginning?

Goal is no gaps
- If all the facts, must type, map

Quality check?

- Sequencing will require > 100,000 bp
- GC content will affect

Try reasonably?

- Issue is 1) about read depth
- Quality of probe

Reasonably - problem for non-shipper?

Should un blocker agent be required? 28ng

- E. coli gives yes
- Not a big extra cost

Fad III

Ben RZ
Princess Hotels

RFA - Tech. Development $60,000
Large Scale Pilot - process for review (see note)
Analysis by 3/1 (review)
Review 3/14-93 (51 recommended)
Submit 3/15
Completion planned 9/93

Other issues -

Database submission?
Timing of final submission - will be based on GSDP standards?

RFA organization?
Analysis updating?

Represent summary database, reviewing - need what is to make entering more user-friendly.

Quality check?
So it is not irresponsible to proceed.
Do we include companies of other years?
Are we to use commercial software?

Timing review?
Independent review?
How do we handle differences?

Yest RFA

[Handwritten notes]
<table>
<thead>
<tr>
<th></th>
<th>14 m Tech. Dec. + 5 = 21</th>
</tr>
</thead>
<tbody>
<tr>
<td>$14.5 m$ Production $+ 15 = 29.50$</td>
<td></td>
</tr>
</tbody>
</table>

96

T R C

$40 m$ $60$

C6.
RESOURCES AVAILABLE

Name of participant: Mark Adams/Craig Venter - TIGR

Nature of resources available (software, maps, clones etc.)

**Software**
- TIGR Assembler - sequence assembly
- HRACM - human-based computer tool
- gkna - GenBank extraction software
- TIGR sybase schema
- Human cDNA Database - >355,000 ESTs, >50,000 THC assemblies

Available via:
- Email request to arkerb@tigr.org (Tony Keravage)
- DBH clone through TIGR/ATCC and WWW

Any conditions attached:

None.
RESOURCES AVAILABLE

Name of participant: **ANSORENA MB CL**

Nature of resources available (software, maps, clones etc.):
- **SEQUENCING TECHNOLOGY:** 100 kb per run
- **GENESKIPPER:** ASSEMBLY PROGRAM
  - **SEQUENCE ANALYSIS**

**RAN-DI (Random-Directed) Strategy:**
- Assembling first 50-100 clones randomly sequenced
+ all EcoR1 fragments

→ No cloning gaps observed

Available via: **FAX**

Any conditions attached:
# RESOURCES AVAILABLE

Name of participant:

Tony Carrano  
Lawrence Livermore National Laboratory

Nature of resources available (software, maps, clones etc.)

<table>
<thead>
<tr>
<th>Resource</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-resolution, metric map of chromosome 19</td>
<td>Published version available in Dec issue of Nature Genetics. Detailed version available by collaboration</td>
</tr>
<tr>
<td>Arrayed cosmid libraries of human chromosomes</td>
<td>Through major genome centers. Soon to be available through the UK and German resource centers.</td>
</tr>
<tr>
<td>IMAGE collection of cDNAs</td>
<td>Available through industry and resource centers.</td>
</tr>
<tr>
<td>DNA sequence sample tracking software</td>
<td>Contact Tom Slezak @ LLNL</td>
</tr>
<tr>
<td>Clone fingerprinting assembly and database software</td>
<td>Contact Tom Slezak @ LLNL</td>
</tr>
<tr>
<td>Mapping infrastructure resource</td>
<td>Contact Tony Carrano @ LLNL</td>
</tr>
<tr>
<td>(creating high-resolution sequence ready maps in cosmids and BACs)</td>
<td></td>
</tr>
</tbody>
</table>

Available via:

see above

Any conditions attached:

Creating maps as part of the mapping infrastructure resource would require funding.
RESOURCES AVAILABLE

Name of participant: RICHARD DURBIN

Nature of resources available (software, maps, clones etc.)

SOFTWARE: ACEDE Database system
SPOT (Can software) rather assembly/secondary
FPC (Can) support soon
DOUTENT - sequence assembler editor for 2D-526
(Richard Hult)
V5PCRUNCH/BELRO/RUSSER - sequence analysis/aligning code

Available via: ANONYMOUS FTP (ftp.SANGER.AC.UK)
Email: RD@SANGER.AC.UK

Any conditions attached:

NO COMMERCIALIZATION (use by companies OK)
Name of participant:
Glen A. Evans

Nature of Resources:

1. Chromosome 11 Sequencing Databses
   - YAC/STS coordinates database
   - cosmid end sequence database
   - YAC-cosmid coordinate database
   - Primers (new STSs)
   - Homology/Identities listed by match significance
   - Chromosome 11 sequencing data (complete cosmid/PAC sequences)
     11p15 project, 11p12 project
   - WWW http://mcdermott.swmed.edu/
   - Genbank

2. Clone libraries
   - chromosome 11 cosmid 5X, arrayed
   - chromosome 11 YAC 7X, arrayed (T. Shows/N. Nowak, RP)
   - chromosome 11 and 15 PAC set in preparation
     (can be made available on request to G. Evans)

3. Software
   - Mermade driver software for 192 channel synthesizer
   - Primer prediction software for primer directed walking
   - SUMU Lab sample tracking software
   - Robotics control software for Biomek
   - Data Inspector software for sequence quality control
   - WWW http://mcdermott.swmed.edu/

4. Hardware specifications and construction plans
   - Prepper III miniprep robot
   - Mermade 192 channel oligonucleotidesynthesizer
   - Lab workstations
   - TREC multigel controller
Lab workstation plans and ordering information

WWW http://mcdermott.swmed.edu/

Available via:

WWW http://mcdermott.swmed.edu/

Any conditions attached:

Data resources are made available within 6 months after generation.

Hardware and software are supplied without warranty and without support other than helpful hints when needed. Hardware specifications and plans are available to all non-commercial users.
RESOURCES AVAILABLE

Name of participant: Chris Fields

Nature of resources available (software, maps, clones etc.)

Chris Fields

GISDB (complete, genomic, scale relational DB) scheduled for operational mid-summer.

GisDB "Annotator" multiplatform client interface (view/edit) available free mid-summer.

Available via: http://www.nrgx.org

Any conditions attached: none
RESOURCES AVAILABLE

Name of participant: Richard A. Gibbs

Nature of resources available (software, maps, clones etc.)

- X chromosome mapped reagents - including binned cosmids (>2,000)
- Sequences, cosmids and the shotgun libraries from >1Mb of human DNA from X;12;17 available
- molxed cosmids/clone years available for X-chromosome from C.C. Lep

Available via:

- X chromosome + ch12 resources are described in their respective web pages.

Any conditions attached:

No
RESOURCES AVAILABLE

Name of participant: Trevor Huchins

Nature of resources available (software, maps, clones etc.)

> 15,000 Human mapped SSRs
> 7,000 Mouse mapped SSRs

GRACE (BASS Gel analysis and basecalling software, UNIX based).

Primer Picking Software (PRIMER 2.2)

LabBase database system

Available via: http://www-genome.sr.usmit.edu

Any conditions attached: None
RESOURCES AVAILABLE

Name of participant: LaDeana Hillier

Nature of resources available (software, maps, clones etc.)

SOFTWARE:
- GETLANES (tracking gel images)
- RETRAK (UNIX interface for editing lane tracking)
- TPP (trace processing software)
- PHRED (base calling)
- PHRAP (sequence assembly)
- FINISH (following shotgun completion, finish picks needs to configure & improve sequence quality)
- DACE (implementation of a laboratory notebook tracking system in ACEDB, other software tools are also available)

Available via: HTTP://genome.wustl.edu/gschmp.html

PHRED & PHRAP available: phg@u.washington.edu
ACEDB code available: ncbi.nlm.nih.gov/pub/pcrkey/acdb

Any conditions attached:

retrak & tpp are still under intensive development.
RESOURCES AVAILABLE

Name of participant: Pieter de Jong

Nature of resources available (software, maps, clones etc.)

- Human PAC library (16-fold redundant)
  - Male donor, DNA from blood
  - ~1100 384-well dishes
- Human PAC library (15-fold redundant)
  - Not yet arrayed 150kb
  - Female donor, DNA from blood
- Human BAC library: in progress
  - Expect to deliver 10-fold redundant
    by May 96 and 20-fold by Summer 96

Available via: PdJ, Roswell Park Cancer Institute

Any conditions attached:

- No secondary distribution of library,
  no problems to distribute individual
  clones (no ties attached).
- Cost-recovery of labor/plasticware/
  mailing costs for library replicates.
RESOURCES AVAILABLE

Name of participant: Dr. Hans Lehrach

Nature of resources available (software, maps, clones etc.)

The Resource Centre distributes high-density gridded filters of genomic libraries, cultures of individual library clones, or (in the future) PCR pools.

The table below gives details of those genomic libraries for which this service is now available, in the near future this will be supplemented with libraries from the I.M.A.G.E. consortium:

<table>
<thead>
<tr>
<th>Cosmid (Human)</th>
<th>Chromosome specific cosmid library</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>L4/FS1</td>
<td>Chromosome 1 specific cosmid library</td>
<td>112</td>
</tr>
<tr>
<td>L4/FS6</td>
<td>Chromosome 6 specific cosmid library</td>
<td>109</td>
</tr>
<tr>
<td>L4/FS7</td>
<td>Chromosome 7 specific cosmid library</td>
<td>113</td>
</tr>
<tr>
<td>L4/FS11</td>
<td>Chromosome 11 specific cosmid library</td>
<td>107</td>
</tr>
<tr>
<td>L4/FS13</td>
<td>Chromosome 13 specific cosmid library</td>
<td>108</td>
</tr>
<tr>
<td>L4/FS17</td>
<td>Chromosome 17 specific cosmid library</td>
<td>105</td>
</tr>
<tr>
<td>L4/FS18</td>
<td>Chromosome 18 specific cosmid library</td>
<td>111</td>
</tr>
<tr>
<td>L4/FS21</td>
<td>Chromosome 21 specific cosmid library</td>
<td>102</td>
</tr>
<tr>
<td>L4/PE22</td>
<td>Chromosome 22 specific cosmid library</td>
<td>106</td>
</tr>
<tr>
<td>L4/FSC X/LA</td>
<td>Chromosome X specific cosmid library</td>
<td>101</td>
</tr>
<tr>
<td>L4/FSC X</td>
<td>Chromosome X specific cosmid library</td>
<td>104</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cosmid (other)</th>
<th>Specific cosmid library</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>L4/S.Pombe</td>
<td>S. pombe specific cosmid library</td>
<td>60</td>
</tr>
<tr>
<td>L4/Bi/S.Pombe</td>
<td>S. pombe specific cosmid library</td>
<td>61</td>
</tr>
<tr>
<td>Fugu-Cosmid</td>
<td>Fugu DNA partial cut with MboI in Lawrist4 and DH10B</td>
<td>65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P1 Human</th>
<th>Total Genomic P1 Human Library</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP1 Mouse P1 library pomP1</td>
<td>Total Genomic Mouse C57Bl/6J P1 Library</td>
<td>700</td>
</tr>
<tr>
<td></td>
<td>Schizosaccharomyces pombe (wt 972 h−) P1 library</td>
<td>703</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PAC</th>
<th>Human PAC library brought by Peter de Jong</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human PAC</td>
<td>Human PAC library brought by Peter de Jong</td>
<td>704</td>
</tr>
</tbody>
</table>
Name of participant: Dr. Hans Lehrach

Nature of resources available (software, maps, clones etc.)

(Continued from previous page)

<table>
<thead>
<tr>
<th>Resources Available</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>YAC (Human)</strong></td>
<td></td>
</tr>
<tr>
<td>4X YAC</td>
<td>Human YAC library</td>
</tr>
<tr>
<td>4Y YAC</td>
<td>Human YAC library</td>
</tr>
<tr>
<td>CEPH YAC</td>
<td>Human CEPH YAC library</td>
</tr>
<tr>
<td>LSXY</td>
<td>Human YAC library</td>
</tr>
<tr>
<td>CHH YAC</td>
<td>Mouse YAC library</td>
</tr>
<tr>
<td><strong>YAC (other)</strong></td>
<td></td>
</tr>
<tr>
<td>St. Mary's Mouse YAC</td>
<td>Mouse YAC library from female C57BL/10 in host strain which is recombination deficient due to mutation in RAD52</td>
</tr>
<tr>
<td>YAC</td>
<td>Large insert Mouse YAC library constructed at the Whitehead Institute for Biomedical Research/MIT Center for Genome Research</td>
</tr>
<tr>
<td>YAC</td>
<td>Schizosaccharomyces pombe (wt 972 b) YAC library</td>
</tr>
<tr>
<td>YAC</td>
<td>Pig YAC library</td>
</tr>
<tr>
<td><strong>cDNA (Human)</strong></td>
<td></td>
</tr>
<tr>
<td>Human fetal brain cDNA</td>
<td>Human fetal brain cDNA made from 17 week embryo polyA+RNA</td>
</tr>
<tr>
<td>HFL cDNA</td>
<td>cDNA using dT primed polyA+ purified RNA from 21 weeks old human fetal liver</td>
</tr>
<tr>
<td>HTE cDNA</td>
<td>cDNA using dT primed polyA+ purified RNA from 21 weeks old human fetal thymus</td>
</tr>
<tr>
<td>HPO cDNA</td>
<td>cDNA from 21 weeks human fetal lung, poly dT primed, directionally cloned, excise enzyme MspI</td>
</tr>
<tr>
<td><strong>cDNA (other)</strong></td>
<td></td>
</tr>
<tr>
<td>MBR cDNA</td>
<td>Mouse adult brain cDNA, synth: oligo dT primed, directionally cloned, cloning site: NotI/SalI; 1.5kb average insert size</td>
</tr>
</tbody>
</table>
Name of participant: Dr. Hans Lehrach

Nature of resources available (software, maps, clones etc.)
(see previous pages)

Available via:
The Resource Centre/Primary Database of the German Human Genome Project,
Max-Planck-Institut für Molekulare Genetik,
(Abteilung Lehrach),
Innestrade 73,
14195 Berlin (Dahlem)
GERMANY
Tel: +49 30 8413 1627
Fax: +49 30 8413 1395
WWW: http://rznd.rz-berlin.mrz.de/

Any conditions attached:
Distribution of these resources will be free of charge to all participants in the German Human Genome Project, otherwise charges will be made to cover manufacturing expenses and postage costs.

In the case of some libraries additional conditions governing usage and distribution have been imposed by the owners.
RESOURCES AVAILABLE

Name of participant: DAVID J. LIPMAN

Nature of resources available (software, maps, clones etc.):

Databases & Software


Available via: WWW, FTP, CORUM

Any conditions attached: None
RESOURCES AVAILABLE

Name of participant:
Dr. Robert K. Moyzis
Ph: 505-667-3912
Center for Human Genome Studies
FAX: 505-667-2891
Los Alamos National Laboratory
email: moyzis@telomere.lanl.gov
Los Alamos, New Mexico 87545

Nature of resources available (software, maps, clones, etc.)

A) Complete digest libraries for each human chromosome
B) Partial digest phage and cosmid libraries for approximately half of the human karyotype (phage: 4, 5, 6, 8, 11, 13, 16, 17, X; cosmid: 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 20, X, Y)
C) YAC libraries for human chromosomes 9, 12, 16 and 21
D) M13/STS libraries (can be constructed for any human chromosome)
E) High-resolution YAC/STS/cosmid maps of human chromosomes 5 and 16

Available via:

A) American Type Culture Collection
B) Request from Los Alamos. Will also be available from commercial sources
C) Request from Los Alamos
D) Collaboration with Los Alamos
E) http://www-ls.lanl.gov; GDB and GSDB; request materials from Los Alamos

Any conditions attached:

A) Small fee; agreement to acknowledge Los Alamos in publications
B) Must sign Material Transfer Agreement with University of California limiting use to scientific purposes, limiting further distribution and agreeing to a limited collaboration with Los Alamos investigators
C) Collaboration with Los Alamos
D) Collaboration with Los Alamos
E) Sequencing coordinated with Los Alamos
RESOURCES AVAILABLE

Name of participant: Richard Myers + David Cox

Nature of resources available (software, maps, clones etc.)
- two panels of whole genome radiation hybrid DNA
  (Stanford 63 panel - 400 kb resolution)
  (Stanford TNG panel - 100 kb resolution)
  available from Research Genetics
- map positions of 7300 STSs on the 63 radiation hybrid
- an email server allowing anonymous STS radiation hybrid scores to be integrated with our mapping data on the 63 hybrids

Available via: https://www-slhgsc.stanford.edu

Any conditions attached:
- none
RESOURCES AVAILABLE

Name of participant: Bruce Roe

Nature of resources available (software, maps, clones etc.)

Laboratory Protocols
CosmidD, PI and BAC sequence data (20 pages)

Available via: HTTP://dna1.chem.walnor.edu

Any conditions attached:

let us know if you find something cool that we missed
RESOURCES AVAILABLE

Name of participant: Melvin I. Simon

Nature of resources available (software, maps, clones etc.)

1. Mouse 129ES Cell - BAC Library A 235,000 clones (n=10x coverage)
2. Human Fibroblast - BAC Library B 75,000 clones (n=3x coverage)
3. Human Sperm - BAC Library C 100,000 clones (n=3x coverage)
4. Human Primary Fibroblast - BAC Library D 75,000 clones
5. 6'-G4 - Ch 22 Specific Mapped BAC clones

Available via:
1, 2 and 3 now available - Research Genetics Inc. (Huntsville)
4 available - Research Genetics Inc. (April 1996)
5 available for screening via Haruko Sugasawa - Biology Division Caltech - Pasadena - FAX (818) 746-7066

Also see:
http://www.tree.caltech.edu

No conditions or restrictions are attached to this material.
RESOURCES AVAILABLE

Name of participant:
Jim Weber

Nature of resources available (software, maps, clones etc.):
- Crude, but comprehensive human linkage maps
- STRP information
- Methods
- Image analysis software
- Construction information for water bath thermal cycler and some SCAFUD components
- Sequence assembly simulation program (from Gene Myers at University of Arizona)

Available via:
Website: http://genetics.mftdcin.edu
Email: gene@cs.arizona.edu

Any conditions attached:
Software is not supported.
RESOURCES AVAILABLE

Name of participant:

Jean Weissenbach

Nature of resources available (software, maps, clones etc.)

The Généthon Human Linkage Map
(5,264 microsatellite markers)

Map + description of reagents
(sequences, primers, alleles, frequencies, etc.)

Available via:

http://www.genethon.fr

Any conditions attached:

freely available
Here's the summary that we gave to the staff:

Summary of program staff meeting -- 2/23/96

Report on the International Strategy Meeting on Human Genome Sequencing

Mark and Jane both have copies of the full agenda and attendance list, if anyone wants to see them. The major groups that were represented were: Sulston, Waterston, Lander, Myers, Venter, Ansorge, Carrano, Molyzis, Evans, Caskey, Chen, Gibbs, Hood, Lehich, Rosenthal, Weissenbach, McCombie, Roe, Weber, Hattori, Simon, de Jong, as well as Lipman, Fields, Ashburner. There was more than one person from several of the groups, a total of 50 people altogether, including agency types from Wellcome, NCHGR, DOE, MRC, Germany, Japan, and HUGO.

The major topics discussed were:

- the sequencing plans/strategies/accomplishments of each of the groups;
- the sequencing resources each group has and will make available (a list of these should have been distributed to staff);
- data release;
- data quality;
- coordination among large sequencing groups.

Some of the key conclusions were:

most sequencing groups seem to be converging on a general strategy of using BACs selected by STS screening and a combination of shotgun and directed sequencing strategies; other strategies, including BAC end sequencing across the genome, shotgun sequencing of the entire genome, and sample sequencing across a complete chromosome were discussed, but in none of these cases was there group consensus that the strategy was superior to the generally-accepted paradigm;

data should be released regularly and very quickly from large-scale sequencing projects, perhaps as frequently as daily but maybe weekly would do; this refers to preliminary data (i.e. contigs > 1 kb, not finished to database
submission quality) which would be put up locally automatically; it was also agreed that finished, annotated sequence would be immediately submitted to databases;

the attendees unanimously agreed to the following statement: for primary genomic sequence data from large-scale sequencing projects, the aim is to have all sequence freely available and in the public domain for both research and development, in order to maximize its benefit to society. This was intended to mean that the primary producers of the sequence from the Human Genome Project would not attempt to patent the sequence they generate. This statement was understood to be the sense of the attendees and that different organizations/agencies/countries might be under different constraints that might or might not allow them to adopt this as policy. The agencies were, however, urged to foster such policies.

data quality issues — representation of data quality is becoming possible and should be reported along with sequence data, particularly in the case of preliminary sequence; the group seemed to be moving toward agreement that the goal is 99.99% accuracy

International Coordination:

As a first step, each of the groups present discussed its goals for the next few years:

Seattle: 25-30 Mb in the next 3 years; primarily centered around the T cell alpha and delta regions on chromosome 7; also the human and mouse MHC regions

TIGR/Cal Tech: 30 Mb in 3 years on chromosome 16p

German consortium (administered by A. Rosenthal at Jena): 30 Mb of chromosome 21 (excluding the minimal Downs and PME regions); 1-2 megabase regions of X (Xq28, Xp11.2, the PAR1 region), 7, 11, 17

LANL: Moyzis original statement was that he intended to do one-pass sequencing across all of chromosome 16; by the time the meeting ended, he was reconsidering that and discussed producing finished sequence of regions of chromosome 16p not being pursued by Sanger or TIGR (a total of about 20 Mb) plus a region near the 5p telomere around the Cri du Chat locus

Whitehead: 105 Mb in 3 yr; human chromosome 17 and mouse syntenic regions plus a few random megabases here and there

Baylor: 30 Mb in 3 yr in Xq28, Xp22, 12p1.3
Oklahoma: I am funded to do 6 Mb in 3 years; working on the region of 22q between NF2 and the centromere; this is being coordinated with Sanger and Wash U.

Japan: did not say, will report by correspondence after the meeting

Chen/ABO: 4 Mb in various regions on X (this is being coordinated through the X chromosome workshops)

LLNL: 50 Mb in 3-5 years on chromosome 19 and mouse syntenic regions

Wash U: 100 Mb in 3 years; regions on chromosomes 7, 22, and X to begin with

CSHL: 5 Mb from chromosomes 13 and 18

EMBL (Ansorge): cDNAs from chromosomes 21 and X (a total of 2 Mb)

Sanger: 150 Mb in 3 years (actually funded for 7 years to do 250 Mb); beginning with 22 and regions of X, then have all of 6 and 20 targeted, followed by chromosome 1

Marshfield: whole genomic shotgun sequencing will be pursued, hoping for a level of 2-4 Mb (raw?) per year

Dallas: 11p15.5, 11p12, 11q23 (2 mb in each region in 2 years); also pursuing sample sequencing across the whole chromosome.

Agencies: Each of the agencies made a short presentation about its plans. The gist of the information discussed was that the major funders for production human DNA sequencing will be NCHGR and the Wellcome Trust. There is a possibility that DOE and the German genome program will make some significant contributions. There was a brief allusion to developing French plans that could not be discussed in public at present. The European Union and the U.K., MRC will not be spending significant amounts of money on production sequencing. And the Japanese were silent.

There was general agreement that this had been quite a useful meeting and that it would be worthwhile reconvening (perhaps a smaller group) about a year from now.