Workshop on Complete cDNA Sequencing

Gaithersburg, MD
May 19, 1997

Participants and Photo

In the fall of 1996, several research teams worldwide announced plans for complete sequencing of cDNAs, the sturdy representatives of the cell's fragile messenger RNAs (mRNAs) for gene expression. A workshop for participants in the international Integrated Molecular Analysis of Genome Expression (I.M.A.G.E.) consortium [http://image.llnl.gov/] [*HGN* **6**(6), 3] was called to extend the highly beneficial infrastructure I.M.A.G.E. has provided since 1994 to the challenges of complete cDNA sequencing. The meeting was organized and chaired by Greg Lennon of the Lawrence Livermore National Laboratory (LLNL), with Marvin Stodolsky coordinating for the meeting sponsor, the DOE Office of Biological and Environmental Research. Scientists attended from France, Germany, Italy, Japan, Sweden, United Kingdom, and the United States.

Several participants are members of the subgroup EURO-IMAGE, whose goals include generating and sequencing a master set of unique full-length cDNA clones (based on I.M.A.G.E. consortium resources) representing 3000 transcripts and 6 Mb of finished sequence, obtaining high resolution and comparative functional mapping in human and model organisms of 1000 genes from the master set, and developing the I.M.A.G.E. consortium data base for easy access to an integrated view of the sequence, map, and expression data generated.

Agencies providing funds for cDNA efforts were represented at the workshop and include DOE, NIH, and the recently established non-profit Merck Genome Research Institute (MGRI) [*HGN* **8** (3-4, 9)]. Selected highlights follow of technical progress in complete cDNA sequencing, as reported by several workshop participants.

Workshop Highlights
Attendees addressed a wide range of topics, including the following:

- Status of cDNA sequencing projects (number of cDNAs sequenced and R&D pilots underway),
- Targets for next year (how selected or obtained, sequencing strategy, coordination with others, funding status),
- Data- and I.M.A.G.E. clone-release policies (submitted to which databases, annotation, clone distribution),
- Quality criteria and assessment (coverage and strandedness, electronic reanalysis, and clone resequencing), and
- Mouse and other model organism cDNAs.

Speakers projected that, with adequate support from funding agencies, participating laboratories could generate up to 15,000 full-length cDNA sequences in the coming year. With average cDNA lengths of 2 kb, this represents some 30 Mb of total sequence.

Technical Progress
It has long been recognized that expression of a single gene may culminate in production of several different mRNA transcripts, depending both on the gene and the source tissue. Added to this biological complexity are the technical challenges of converting mRNAs to the sturdier cDNAs. Libraries with abundant truncated products are the common result, particularly for longer source mRNAs. Strategies devised for alleviating this truncation problem were described by Takao Isogai (Helix Research Institute, Japan), Nobuo Nomura (Kazusa DNA Research Institute, Japan), John Quackenbush (The Institute for Genomic Research, [TIGR, http://www.tigr.org]) and M. Bento Soares (University of Iowa).

A greater proportion of full-length cDNA products can be obtained through the use of two tactics. One protocol type takes advantage of the unusual nucleotide "cap" on the 5' end of mRNAs. This requires that the extension of the first cDNA strand be long enough to protect the cap, as a contingency for final cDNA clone production. However, Soares reported that about one-third of the cDNA transcripts begin within the mRNA, as contrasted with the desired starts at the

3' end of the mRNA, giving rise to 3' truncations. This problem can be alleviated substantially by size fractionating the mRNAs and later selecting out the cDNA products with lengths equal to the size-sorted mRNA templates.

Hans Lehrach (Max Planck Institut für Molekulare Genetik, Germany) related the value of massively parallel oligomer fingerprinting of cDNAs. This is an economical way to screen a library for novel and longer, potentially full-length cDNAs. Optimal candidate cDNAs chosen by the Lehrach team at the Resource Center of the German Genome Project are being sequenced in the laboratory of Annemarie Poustka, Deutsches Krebsforschungszentrum [http://www.dkfz.de/en/mga/].

More than one sequencing read is commonly necessary to display the complete sequence for cDNAs longer than a few hundred bases. Strategies for economical full-length sequencing were discussed by Greg Lennon and Richard Gibbs, Baylor College of Medicine [http://www.hgsc.bcm.tmc.edu/]. Sequence reads beyond 1000 bases now are being obtained with improvements to sequencing systems by a team led by Wilhelm Ansorge, European Molecular Biology Laboratory [http://www-db.embl.de/jss/EmblGroupsOrg/g_25]. Ansorge suggested that, for cDNAs shorter than 2 kb, good coverage could be achieved by two overlapping reads on complementary strands.

Giuseppe Borsani (Telethon Institute of Genetics and Medicine (TIGEM), Italy) reported on the benefits of the easily manipulated *Drosophila* model for studies of development and function to reveal roles represented by human cDNAs.

Mark Boguski (National Center for Biotechnology Information, NCBI) reported on the status of the dbEST cDNA sequence database [http://www.ncbi.nlm.nih.gov/dbEST/] and the Transcript Map, and made recommendations for the evolution needed under the impending new demands of complete DNA sequencing. He observed that each group will have its own selection criteria and sequencing priorities such as finding cancer genes, genes with *Drosophila* homologs, or genes that already have been mapped. Boguski coined the expression "the slicing problem" to describe the difficulties in avoiding undesirable duplication and redundancy due to overlapping choice categories. A possible solution would be to establish a registration and tracking database modeled after the successful European Bioinformatics Institute RHAlloc/RHdb approach used in the construction of the human transcript map [http://www.ncbi.nlm.nih.gov/SCIENCE96/]. Such a database would include an investigator or center name and contact information, identifiers for the physical cDNA clones being sequenced and associated EST accession numbers, and sequencing status. When participants registered a clone that they intended to sequence, the database would detect and report overlaps with clones selected by other groups.

Toward the close of the workshop, attendees agreed that I.M.A.G.E. meetings should take place every six months to maintain necessary coordination and efficiency. The next meeting, organized by John Quackenbush, was held in September in conjunction with the Ninth International Genome Sequencing and Analysis Conference in Hilton Head, South Carolina. Washington University I.M.A.G.E. participants will organize the next workshop, tentatively planned to be continuous with the May, 1998 Human Genome Workshop at the Cold Spring Harbor Laboratory [http://www.cshl.edu/].

Participants in the Workshop on Complete cDNA Sequencing

Gaithersburg, MD
May 19, 1997

Participants:
Front, left to right: Takao Isogai, Wilhelm Ansorge, Marvin Stodolsky, Giuseppe Borsani, Annemarie Poustka, Kirsten Timms, Marcelo BentoSoares, M.J. Finley Austin, Elise Feingold, Hans Lehrach.
Back, left to right: Greg Lennon, Charles Auffray, Nobuo Nomura, Richard Gibbs, Michael Metzker, Stephan Wiemann, John Quackenbush, Mark Yandell, Mark Boguski, Cleo Naranjo, Joakim Lundeberg, Chris Mundy.
Not pictured: Carol Dahl, Marvin Frazier, Aristides Patrinos, Robert Strausberg

Participant Roster

---

cDNAs represent mRNAs
When a gene is expressed, chromosomal DNA is transcribed into nuclear RNA molecules containing long non-coding introns separating the short protein-coding segments. A much shorter mRNA with a contiguous protein coding segment matures as the introns are naturally processed out within the nucleus. Processing of a single nuclear mRNA into a few different mature mRNAs is common, and depends on both the gene and the tissue or organ in which it is expressed. As a population, mRNAs range in length from a few hundred to a few thousand bases, depending on both the source gene and the nuclear RNA processing. mRNAs are exported to the ribosomal complexes of the cytoplasm, where they act as templates for protein synthesis. mRNAs can be short lived naturally and are also highly susceptible to degradative processes during harvesting from cells.

For analytical purposes, mRNAs can be worked up into the much sturdier libraries of double-stranded cDNAs. Using poly dT as a primer on the 3' poly A end of purified mRNAs, reverse transcriptase enzymes of viral origin polymerize the synthesis of a single-stranded DNA complement of the mRNA. These initial DNA transcripts often fail to extend to the 5' end of longer mRNAs. Using more routine biochemistries, the single-stranded DNA is converted into duplex DNA and combined with a DNA vector to support its propagation and maintenance as a DNA clone. Double-stranded DNAs are much more stable and less susceptible to degradative processes than their single-stranded mRNA predecessors.

All cDNA clones derived from a particular tissue constitute a library of clones representing the genes that were expressed when the source tissue was harvested. The analysis of libraries from many different tissues, and obtained under a variety of physiological conditions, will be necessary to decipher the organ-specific patterns of gene expression.

The need for coordination and infrastructure
In a cDNA library the numerical representation of particular cDNAs varies over a thousand-fold. The predominant members of cDNA libraries from all tissues are the genes for cellular maintenance functions. I.M.A.G.E.'s coordination minimizes the unwanted and expensive repetitive analysis of the already characterized cDNAs.

A single sequencing read of a few hundred cDNA bases is usually sufficient to serve as a distinguishing identifier (called an expressed sequence tag, or EST) of the predecessor mRNA. This approach was pioneered by Craig Venter, now director of TIGR [http://www.tigr.org], which made public a major EST data release in June 1997. High throughput production of ESTs from I.M.A.G.E. cDNAs has been predominantly funded by Merck & Co., with sequencing done at the Washington University (St. Louis) Genome Center Human EST Project. The ESTs are deposited in the public database dbEST (http://www.ncbi.nlm.nih.gov/dbEST/) which supports queries on the similarities of ESTs and cDNAs to cDNA molecules whose analysis is just beginning.

The I.M.A.G.E. consortium manages distribution of reference sets of cDNAs. In the U.S., libraries of cDNA clones representing many different tissues are donated to I.M.A.G.E. at Lawrence Livermore National Laboratory (LLNL) where the clones are placed into reference arrays and replicas are then provided to genome research centers and private sector resource distributors. Over 3 million clone replicas have been sent to over 1000 laboratories worldwide. End users return data on the analyzed clones. All data are entered into public databases, enabling researchers to compare this against their preliminary cDNA sequencing data and eliminate redundant efforts.

EST analyses of over 500,000 cDNA I.M.A.G.E. clones suggest that over 50,000 of the estimated 60,000 to 80,000 human genes are represented. I.M.A.G.E. researchers at LLNL are providing "subtracting cDNA reagents" to aid the production of new cDNA libraries (by M. Bento Soares) that preferentially contain clones not already represented in the current I.M.A.G.E. collection.

Support in the USA
Several funding agencies provide support for cDNA-related projects. To annotate developing chromosome maps, DOE in 1990 began dedicated support for improved cDNA library production, early EST generation by C. Venter's team, physical mapping of cDNAs onto chromosomes, and database support. High throughput correlations of cDNAs with the new BAC resources are also in progress. The sequencing of cDNAs corresponding to genes recognized during genomic

sequencing is often a component of major chromosome sequencing projects.

The NIH National Human Genome Research Institute [http://www.genome.gov/] is also supporting research and development in cDNA library improvement and mouse cDNA library production. In the NIH-supported chromosome map development using radiation hybrid methodologies, about one-third of the markers are derived from ESTs. The source genes are thus mapped onto the chromosomes. Recently the NIH National Cancer Institute (NCI) began providing substantial support for cDNA library production and analysis in a major effort to identify cancer-related genes [see article, *HGN* **8**(3-4), 8]. This effort, called the Cancer Genome Anatomy Project [http://www.ncbi.nlm.nih.gov/CGAP/], CGAP, was described at the workshop by Carol Dahl and Robert Strausberg of the NCI.

Merck (http://www.mgri.org/) grant administrator M.J. Finley Austin (mgri@merck.com) spoke of supporting programs to characterize cDNAs representing disease genes, which include full length cDNA cloning and sequencing.

The utility of the mouse model for studying human diseases is being advanced with diverse collaborative support, including NIH for library construction and DOE for I.M.A.G.E. efforts at LLNL to array mouse cDNA libraries [http://image.llnl.gov/image/html/muslib_info.shtml]. Washington University (St. Louis) generates mouse ESTs for the clone arrays with support from the Howard Hughes Medical Institute [http://www.hhmi.org/].

- I.M.A.G.E. at LLNL, *http://image.llnl.gov/*
- Wilhelm Ansorge Lab at EMBO
- Helix Research Institute, Japan
- Kazusa DNA Research Institute, Japan, *http://www.kazusa.or.jp/en/*
- The Institute for Genomic Research (TIGR), *http://www.tigr.org*
- M. Bento Soares (University of Iowa)
- Resource Center of the German Genome Project, *http://www.rzpd.de* of the German Genome Project, *http://www.dhgp.de*
- Laboratory of Annemarie Poustka, Deutsches Krebsforschungszentrum, *http://www.dkfz.de/en/mga/*
- Telethon Institute of Genetics and Medicine (TIGEM), Italy
- The dbEST database at the National Center for Biotechnology Information, *http://www.ncbi.nlm.nih.gov/dbEST/index.html*
- Human transcript map, *http://www.ncbi.nlm.nih.gov/SCIENCE96*
- Washington University (St. Louis) Genome Center Human EST Project, *http://genome.wustl.edu/est/index.php?human_merck=1*
- NIH National Human Genome Research Institute, *http://www.genome.gov/*
- NIH National Cancer Institute (NCI) Cancer Genome Anatomy Project, CGAP, *http://www.ncbi.nlm.nih.gov/ncicgap*
- Baylor College of Medicine, Human Genome Sequencing Center, *http://www.hgsc.bcm.tmc.edu/*
- Mouse cDNA resources at LLNL, *http://image.llnl.gov/image/html/muslib_info.shtml*
- Washington University Mouse EST Project, *http://genome.wustl.edu/est/index.php?mouse=1*
- Howard Hughes Medical Institute, *http://www.hhmi.org/*

Workshop on Complete cDNA Sequencing

Gaithersburg, MD
May 19, 1997

The European IMAGE Consortium

EURO-IMAGE is a shared-cost Research and Technology Development Program sponsored by the European Union BIOMED2 Program. The total budget for the program, which spans the period August 1997 to July 2000, is 5.6 million ECU, of which 2.8 million ECU are provided by the EU and the rest by the participating institutions.

General Objectives of the European IMAGE Consortium:

- To generate a miniminal set of non redundant cDNA clones for most human gene transcripts and a master set of unique full-length cDNA clones representing 3,000 transcripts based upon the IMAGE Consortium resources (arrays of cDNA and CpG islands clones).
- To characterize by DNA sequencing with high accuracy the complete sequence of the master set of 3,000 human gene transcripts (6 Mbases of finished sequence).
- To obtain high resolution and comparative functional mapping localization in man and model organisms of 1,000 genes represented in the master set.
- To develop the IMAGE Consortium Data Base to provide an easy access to an integrated view of the sequence, map and expression data generated.

Workplan
The European IMAGE Consortium will devote 20 percent of the resources to the assembly of the physical resources (cDNA and CpG island arrays characterized by end sequencing and fingerprinting, minimal sets of clones selected after comprehensive sequence, clone and functional clustering, master set of 3,000 full-length clones). These resources will be available throughout the European Union for the user community in both academy and industry. The Consortium will serve the future needs of the scientific community in the systematic identification of all human genes and their regulatory sequences by deciphering in an efficient and economic manner 6 Mb of complete, finished sequence for 3,000 transcripts of average size 2 kb: 50 percent of the resources will be devoted to the sequencing of full-length cDNAs and CpG islands selected by Consortium on the basis of their map position, similarity to known families or expression profiles. In order to ensure that advances in basic genetic knowledge is used to further enhance human health, the Consortium will seek to contribute to the identification of genes involved in human biology and diseases by correlating precise map location and phenotypic expression data, exploiting various comparative approaches for 1,000 of the genes represented in the master set: 10 percent of the resources will be devoted to high-resolution and comparative functional mapping in close interaction with the mapping consortia in order to obtain the most precise and evolutionary relevant map location. Last but not least, 20 percent of the resources will be devoted to the IMAGE Consortium Data Base. This will provide the community with up to date, integrated sequence, mapping and expression data related to the IMAGE consortium arrays, as they are collected by IMAGE Consortium members in Europe, the United States and Japan, and will help in sharing and harmonizing such data.

EURO-IMAGE Participants:

Partner 1- Coordinator -Charles Auffray
Centre National de la Recherche Scientifique (CNRS)
UPR 420
19, rue Guy Mquet
BP 8
94801 VILLEJUIF Cedex
FRANCE
Tel : 33 1 49 58 34 98 - Fax : 33 1 49 58 35 09

E-mail : auffray@infobiogen.fr
Secretary : obrasier@infobiogen.fr

Partner 2- Wilhelm Ansorge
European Molecular Biology Laboratories (EMBL)
Postfach 10.2209
Meyerhofstrasse 1
6900 HEIDELBERG
GERMANY
Tel : 49 6221 387 574 - Fax : 49 6221 387 306
E-mail : ansorge@embl-heidelberg.de
Hartmut Voss tel : 49 6221 387 474 E-mail : voss@embl-heidelberg.de

Partner 3- Andrea Ballabio
Telethon Institute of Genetics and Medicine (TIGEM)
Via Olgettina 58
20132 MILANO
ITALY
Tel : 39 221 560 206 - Fax : 39 221 560 220 E-mail : ballaba@dibit.hsr.it

Giuseppe Borsani
Tel : 39 221 560 1
E-mail : borsani@tigem.it

Partner 4 - Xavier Estivill
Institut de Recerca Oncologica (IRO)
Dept de Genetica Molecular
Hospital Duran i Reynals
Autovia de Castelldefels, km2.7
L'Hospitalet de Llobregat
BARCELONE 08907
SPAIN
Tel : 34 3 335 7152 - Fax : 34 3 263 22 51
E-mail : estivill@gm.iro.es

Melanie Pritchard
Tel : 34 3 335 7152
E-mail : pritchar@gm.iro.es

Partner 5 - Keith Gibson
Human Genome Mapping Project Resource Centre (HGMP)
Hinxton Hall
HINXTON, Cambridgeshire CB10 1RQ
UK
Tel : 44 1223 494 500 (or 514) - Fax : 44 1223 494 512
E-mail : kgibson@hgmp.mrc.ac.uk

Partner 6 - Hans Lehrach
Max-Planck-Institut fr Molekulare Genetik (MPI)
Ihnestrasse 73
Dahlem
14195 BERLIN
GERMANY
Tel : 49 30 84 131 221 - Fax : 49 30 84 131 380
E-mail : lehrach@mpimg-berlin-dahlem.mpg.de

John O'Brien
Tel : 49 30 84 13 12 37
E-mail : obrien@mpimg-berlin-dahlem.mpg.de

Mikael Wiles
E-mail : wiles@mpimg-berlin-dahlem.mpg.de

Partner 7 - Annemarie Poustka
Genomforschung und Bioinformatics
Deutsches Krebsforschungszentrum (DKFZ)
Im Neuenheimer Feld 280
69120 HEIDELBERG
GERMANY
Tel : 49 6221 424 742 (or 46 46) - Fax : 49 6221 423 454
E-mail : a.poustka@dkfz-heidelberg.de

Stefan Wiemann
Tel : 49 6221 423 416
E-mail : S.Wiemann@DKFZ-Heidelberg.de

Partner 8 - Mathias Uhlen
Kungl Tekniska Hgskolan (KTH)
Royal Institute of Technology
Department of Biochemistry & Biotechnology
Teknikringen 34
STOCKHOLM S 100 44
SWEDEN
Tel : 46 8 790 75 16 - Fax : 46 824 54 52
E-mail : mathias@biochem.kth.se

Joakim Lundeberg - Tel : 46 8 790 87 58 E-mail : joakiml@biochem.kth.se

Unite de Genetique Moleculaire
et Biologie du Developpement
CNRS UPR 420 - 7-19 rue Guy Moquet
BP 8 - 94801 VILLEJUIF CEDEX - FRANCE
Tel : 33 (0)1 49 58 34 98 - Fax : 33 (0)1 49 58 35 09
E-mail : auffray@infobiogen.fr

Workshop on Complete cDNA Sequencing

Gaithersburg, MD
May 19, 1997

IMAGE Participants	**Target/Yr (5/97-5/98)* (# of cDNA Molecules)**
Washington University (St. Louis)	1K-10K, full insert
Baylor College of Medicine	1K, full insert
The Institute for Genomic Research	500+, full length (1-5K/yr)
Kazusa Research Institute (Japan)	300, full length
EURO-IMAGE	1-1.5K, full length
Joint Genome Institute	250, full insert
Helix Research Institute (Japan)	500, full length
German Human Genome Project (DHGP)	1-1.5K, full length
TOTALS	2-11K full inserts 3-4K full lengths

*Projections contingent on adequate funding

Workshop on Complete cDNA Sequencing

Gaithersburg, MD
May 19, 1997

Helix Full Length cDNA Project for High Throughput Functional Analysis

In 1996 the *Helix Research Institute, Inc.* was established to develop effective technologies for identifying biologically important genes and evaluating their function. This six-year joint research project is cofunded by the Japanese government and industry, with a planned budget of 6.6 billion yen. About 70% of the money comes from the Ministry of International Trade and Industry (MITI) through the Japan Key Technology Center, and 30% from 10 private companies.

The institute consists of genomics, bioinformatics, and biological technology laboratories. Projects include high speed sequencing of full-length cDNAs, detection of evolutionary and functional relationships using bioinformatics, and use of novel experimental technologies (such as DNA chips) to elucidate the functions of genes considered to be of particular interest.

A recent major target of the worldwide genome project has been human gene cloning using full length cDNAs and their functional analysis as related to human health and disease. Our Genomics Department combines established methods for high-throughput cloning of complete full length cDNAs and identifying gene functions. Bioinformatics support includes performance of sequence analyses and functional predictions of new genes. The Biological Technology Department supports genomics efforts by developing experimental methods for evaluating gene function and gene functional analysis of biological mechanisms through gene expression profiles and other strategies. A unique set of full-length cDNA clones would be a powerful tool for exploring the functional analysis of genes. In collaboration with S. Sugano (University of Tokyo) we have developed an efficient method for cloning full-length cDNAs by the oligo-cap method (Maruyama, K. and S. Sugano, *Gene* **138**:171-174, 1994). Using this approach, full-length-enriched cDNA libraries are constructed from cultured cells and tissues. A mammalian expression vector is used for direct applications to functional analyses in biological assays. We are now analyzing the 5' and 3'ends of the cDNA clones and performing full cDNA sequencing as shown in the figure.

Helix Full-Length cDNA Project for High-Throughput Functional Analysis
Takao Isogai
isogai@hri.co.jp
Helix Research Institute, Inc.
1532-3 Yana, Kisarazu, Chiba 292, Japan
Tel +81-438-52-3951; FAX: -3952

Workshop on Complete cDNA Sequencing

Gaithersburg, MD
May 19, 1997

KASUZA DNA RESEARCH INSTITUTE

Located on the Boso Peninsula facing Tokyo Bay, the *Kazusa DNA Research Institute (KDRI) [http://www.kazusa.or.jp/]* supports sequencing and functional analysis of the genomes of biologically as well as biotechnologically important organisms.

KDRI is operated by the Kazusa DNA Research Foundation established by the Chiba Prefecture Government and authorized as a nonprofit organization by the Ministry of International Trade and Industry and the Science and Technology Agency. Annual budget, mostly supported by the Prefectural Government through the Kazusa DNA Research Foundation, was about 1.8 billion yen (16 million dollars). KDRI houses about 25 scientists and 50 support staff.

Projects include the following:

- Sequencing of the 3.6 Mb genome of a cyanobacterium (photosynthetic bacterium)
- Sequencing of chromosomes 3 and 5 from the genome of *Arabidopsis thaliana*, as members of an international consortium.
- Sequencing of full-length cDNAs from cultured human cells (including brain).
- Development of new technologies, particularly for large-scale DNA sequencing.
- Explorations of the molecular mechanisms of nuclear fusion in fission yeast.
- Determinations of the most efficicient ways to compile and transmit vast amounts of sequence information obtained from DNA databases
- KDRI also publishes the international bimonthly journal, DNA Research, which carries papers on genome sequence, DNA analysis technologies, and informatics topics.

Nobuo Nomura
Kazusa DNA Research Institute
1532-3 Yana, Kisarazu, Chiba 292 JAPAN
Tel +81-438-52-3930; Fax: -3931
nomura@kazusa.or.jp

Workshop on Complete cDNA Sequencing

Participant Roster

Wilhelm Ansorge
European Molecular Biology Laboratory
Postfach 10.2209
Meyerhofstrasse 1
W-6900 Heidelberg Germany
Tel: 49-6221-387-355
Fax: 49-6221-387-306
ansorge@embl-heidelberg.de

Charles Auffray
CNRS UPR 420 - Genetique Moleculaire
et Biologie du Developpement
7-19 rue Guy Moquet - Batiment
BP 8
94801 Villejuif Cedex
France
Tel: +33-1-49-58-3498
Fax: +33-1-49-58-3509
auffray@infobiogen.fr

M.J. Finley Austin
Administrative Director
Merck Genome Research Institute
Sumneytown Pike
West Point, PA 19486
Tel: 215-652-4180
Fax: 215-652-6913
finley_mgri@merck.com

Mark Boguski
National Center for Biotechnology Information
National Library of Medicine
Building 38A, Room 8S-810
8600 Rockville Pike
Bethesda, MD 20894

Joakim Lundeberg
Dept. of Biochemistry
KTH, Royal Institute of
Technology
S-100 44 Stockholm, Sweden
Tel: +46-8-790-87-58
Fax: +46-8-24-54-52
joakim.lundeberg@biochem.kth.se

Michael Metzker
Merck and Co.
Sumneytown Pike
PO Box 4, WP26A-3000
West Point, PA 19486
Tel: 215-652-1333
Fax: 215-652-2075
michael_metzker@merck.com

Chris Mundy
HGMP Resource Centre
Hinxton Hall, Cambridge
CB10 1SB United Kingdom
Tel: +44-1223-494-515
Fax: +44-1223-494-512
cmundy@hgmp.mrc.ac.uk

Cleo Naranjo
Los Alamos National Laboratory
LS-3 M888
Los Alamos, NM 87545 Tel:
505/6 67-2782
Fax: 505/665-6894
naranjo@telomere.lanl.gov

Nobuo Nomura

Tel:301/496-2475
Fax: 301/480-9241
boguski@ncbi.nlm.nih.gov

Giuseppe Borsani
Research Unit Coordinator
Tigem - Telethon Institute of Genetics and Medicine
Via Olgettina 58
20132 Milano - Italy
Tel: 39-2-215-601
Fax: 39-2-215-60220
borsani@tigem.it

Carol A. Dahl
Assistant to the Director
National Cancer Institute
7500 Wisconsin Avenue, Room 312
MSC 9010
Bethesda, MD 20892-9010
Tel: 301-496-1550
Fax: 301-496-7807
carol_dahl@nih.gov

Elise Feingold
Assistant Dir. for Program Coordination
NHGRI
Building 38A, Room 604
38 Library Drive, MSC 6050
NIH
Bethesda, MD 20892-6050
Tel: 301/402-5407
Fax: 301/480-2770
Elise_Feingold@nih.gov

Marvin Frazier
Director, OBER Life Sciences and
Life Sciences Research Div.
DOE
19901 Germantown Rd.
Germantown, MD 20874-1290
Tel: 301/903-5468
Fax: 301/903-5051
marvin_frazier@oer.doe.gov

Richard Gibbs
Kirsten Timms
Institute for Molecular Genetics
Baylor College of Medicine
Gibbs Tel: 713/798-6539
Timms Tel: 713/798-6538
Fax: 713/798-5741
agibbs@bcm.tmc.edu
ktimms@bcm.tmc.edu

Takao Isogai
Helix Research Institute

Kazusa DNA Research Institute
1532-3 Yana, Kisarazu, Chiba 292
JAPAN
Tel: +81-438-52-3930
Fax: +81-438-52-3931
nomura@kazusa.or.jp

Aristides Patrinos
Associate Director, OBER
Department of Energy
19901 Germantown Rd.
Germantown, MD 20874-1290
Tel: 301/903-3251
Fax: 301/903-5051
ari_patrinos@oer.doe.gov

Annemarie Poustka
Stefan Wiemann
Molekulare Genomanalyse
Deutsches
Krebsforschungszentrum (DKFZ)
Im Neuenheimer Feld 280/560
D-69120 Heidelberg, Germany
Tel: +44-131-332-2471
Fax: +44-131-343-2620
A.Poustka@DKFZ-Heidelberg.DE

S.Wiemann@DKFZ-
Heidelberg.DE

John Quackenbush
The Institute for Genomic
Research
9712 Medical Center Drive
Rockville, MD 20850
Tel: 301/838-3528
Fax: 301/838-0208
johnq@tigr.org

Marcelo Bento Soares
Associate Professor of Pediatrics,
Physiology and Biophysics
University of Iowa
451 Eckstein Med. Research Bldg.

Iowa City, IA 52242
msoares@blue.weeg.uiowa.edu

Marvin Stodolsky
US Dept of Energy, ER-72
Human Genome Program
19901 Germantown Rd.
Germantown, MD 20874-1290
Tel: 301/903-4475
Fax: 301/903-8521
Marvin.Stodolsky@oer.doe.gov

1532-3 Yana, Kisarazu, Chiba 292, Japan
Tel: +81-438-52-3969
Fax: +81-438-52-3952
isogai@hri.co.jp

Hans Lehrach
Max-Planck Institut fur Molekulare Genetik
D-14195 Berlin (Dahlem)
Ibnestrasse 73
Germany
Tel: +49-30-8413-1221
Fax: +49-30-8413-1380
lehrach@mpimg-berlin-dahlem.mpg.de

Greg Lennon
Gene Logic, Inc.
708 Quince Orchard Road
Gaithersburg, MD 20878
Tel: (301) 987-1700
Fax: (301) 987-1701
glennon@genelogic.com

Note: Greg Lennon has recently moved to Gene Logic from LLNL. For I.M.A.G.E. issues, he is retained as a consultant to LLNL with the below email address still active.
greg@mendel.llnl.gov

Robert Strausberg
Assistant to the Director
Strategic Technologies
National Cancer Institute
7500 Wisconsin Ave., Room 312, MSC 9010
Bethesda, MD 20892-9010
Tel: 301/496-1550
Fax: 301/496-7807
strausbb@od.nci.nih.gov

Mark Yandell
Genome Sequencing Center
Campus Box 8501
Washington University
4444 Forest Park Blvd.
St. Louis, MO 63108
Tel: 314/286-1844
Fax: 314/286-1810
myandell@watson.wustl.edu