

GS01 0163

Analysis of Microarray Data

Keith Baggerly and Kevin Coombes
Section of Bioinformatics

Department of Biostatistics and Applied Mathematics
UT M. D. Anderson Cancer Center

kabagg@mdanderson.org

kcoombes@mdanderson.org

12 October 2004

Lecture 13: Biological Interpretation

- Introduction
- Primary probe identifiers
 - Sigma-Genosys
 - Agilent
 - Affymetrix
 - IMAGE
- GenBank
- UniGene
- LocusLink
- Batch Resources

Introduction

After analyzing a microarray experiment, you typically end up with a list of “interesting” genes. Today’s lecture deals with how to make biological sense of that list.

Keep in mind that the list may be quite long. For example, in our analysis of the prostate cancer data set, we found about 3500 genes that were differentially expressed (along with 3500 potential biomarkers, with an overlap of about 2500).

There are numerous databases (GenBank, UniGene, LocusLink, etc.) that include gene-related information. It can be difficult to keep track of exactly what kind of information each identifier describes. Each of the three kinds of microarrays has a different primary descriptor that tells you what biological material was placed on the array.

Primary Identifiers

synthesized oligonucleotide array : These are the Affymetrix arrays. The primary identifier is an **Affymetrix probe set ID**, which refers to the collection of 25-mers that form a probe set.

spotted cDNA array : These are the most common glass microarrays. The primary identifier is an **IMAGE clone id**, which refers to the actual cDNA clone attached to a vector and propagated in bacteria

spotted oligonucleotide array : Newer glass arrays often spot commercially synthesized 60- or 70-mers instead of cDNA clones. The primary identifier is usually the commercial identifier that corresponds to the actual synthesized sequence.

http://www.sigma-genosys.com

Custom Oligos | **Gene Expression Analysis** | **Custom Peptides** | **Peptide Antisera** | **Molecular Biology**

SIGMA GENOSYS Gene Expression Analysis

Tue, 10/12/2004 10:54:54 AM CST

ORDER ONLINE
ORDER NOW

SEARCH
Enter Criteria **GO**

SIGN IN
E-mail Address
Password
SUBMIT

[Forgot Password](#)
 [Register Now](#)
 [Learn More](#)

ORDERING INFO
 [Oligos](#)
 [Peptides](#)

Oligonucleotide Libraries

 An accurate transcriptome should be the starting point for gene expression probe design. Sigma-Genosys and Compugen together offer Oligonucleotide Libraries designed using a superior transcriptome database generated by Compugen's LEADS™ technology. These genome-wide probe sets:

- ▶ Capture the maximum number of splice variants
- ▶ Avoid SNPs, repeats, chimeras, and intron contamination
- ▶ Are logically organized by Gene Ontology™ functional assignment
- ▶ Maximize sensitivity and specificity

Each library probe is a 65-mer (except human library probes, which are 60-mers) with a 5'-C6 amino modifier for covalent attachment. Oligos in this length range can be designed

UPDATES

New
E. coli Oligonucleotide Library
This novel three-strains-in-one library contains 6048 oligonucleotide probes representing 15,185 genes from three *E. coli* strains.
[more ▶](#)

OLIGO LIBRARIES

- ▶ [Human Release 2](#)
- ▶ [Human Release 1](#)
- ▶ [Mouse](#)
- ▶ [Rat](#)
- ▶ [Zebrafish](#)
- ▶ [Bacillus subtilis](#)
- ▶ [E. coli](#)

Technical Information

Sigma-Genosys Human Oligo

Sigma-Genosys Human Library Release 2 - Netscape

File Edit View Go Bookmarks Tools Window Help

http://www.sigma-genosys.com/oligonucleotide_human2.asp

Search

Home Radio Bookmarks Google Bioinformatics ...

Custom Oligos | Gene Expression Analysis | Custom Peptides | Peptide Antisera | Molecular Biology

SIGMA GENOSYS Gene Expression Analysis

Tue, 10/12/2004 10:55:52 AM CST

NEW 29,000 - Gene Human Oligonucleotide Library

Designed using newly available genomic and expressed sequence data and an enhanced LEADS platform for transcriptome analysis, the **Human Oligonucleotide Library Release 2.0** sets new standards for coverage and quality.

This new library represents an advance over Release 1.0 in the following ways:

- Wider genomic coverage
- New representation of predicted genes
- Better splice variant representation
- Lower cross homology
- Better sequence quality and fewer polymorphisms
- Fewer bases in repeat alignments
- Improved identification of chimeric sequences and antisense genes

HUMAN LIBRARY 2.0

- Product Info
- Gene List
- Plate selection by GO assignment
- Evaluation Plate
- For our customers

OLIGO LIBRARIES

- Human Release 2
- Human Release 1
- Mouse
- Rat
- Zebrafish
- Bacillus subtilis
- E. coli

Technical Information

- Oligonucleotide FAQ

ORDER ONLINE

ORDER NOW

SEARCH

Enter Criteria **GO**

SIGN IN

E-mail Address

Password

SUBMIT

- ▣ Forgot Password
- ▣ Register Now
- ▣ Learn More

ORDERING INFO

- Oligos
- Peptides

http://www.sigma-genosys.com/oligo.asp

Sigma-Genosys Product List

LabOnWeb
by Corrugon

Home | OligoLibraries | Custom Oligos | Science & Technology | FAQs | Ordering | Contact Us

Genome-wide OligoLibrary Product List

OligoLibraries are comprehensive oligo collections designed to optimally represent each gene in the genome. OligoLibraries are available for human, mouse, rat, zebrafish, and *B. subtilis*. From this page, you can download a gene list, a product information sheet, and a product profile for each OligoLibrary. Evaluation plates and library subsets are also available.

Catalog #	Description	Downloads	
		Gene list	Product info (PDF)
BACLIB96	B. Subtilis OligoLibrary: 96 well format, 43 plates (4,128 oligos)	Excel, 2 MB Zip, 262 KB	Info sheet Product profile
BACLIB384	B. Subtilis OligoLibrary: 384 well format, 11 plates (4,128 oligos)	Excel, 2 MB Zip, 285 KB	Info sheet Product profile
HUBLIB96	Human OligoLibrary Release 2.0: 96 well format, 304 plates (29,134 oligos)	Excel, 14 MB Zip, 2 MB	Info sheet Product profile

OLIGOLIBRARIES

- [About OligoLibraries](#)
- [View presentation](#)
- [Reasons to choose our oligos](#)
- [Gene lists](#)
- [Plate selection by GO assignment](#)
- [Evaluation plates](#)
- [FAQs](#)
- [Ordering](#)
- [For our customers](#)

Sigma-Genosys Oligo List

Microsoft Excel - HUBLIB96_ps.xls

File Edit View Insert Format Tools Data Window Help

100% Arial 16 B I U

C7 GenBank_accession

Copyright (c) 2002 CompuGen Incorporated. All rights reserved.
 This file was retrieved after accepting LabOnWeb Oligo Sequence
 Any usage of any data contained in this file is subject to LabOnWeb
 Terms of Use.

Library: HUBLIB96 Version: 1 created at

Plate	Location	GenBank_accession	UniGene_ID	Serial_number	
1	A1	M33197	Hs.169476	CGEN HUMCTRL 10001 1 0	Human gly
1	A2	NM_005856	Hs.25691	CGEN HUMAN LIB2 6010300008 0	Homo sap
1	A3	NM_012076	Hs.169745	CGEN HUMAN LIB2 6010300016 0	Homo sap
1	A4	NM_001243	Hs.1314	CGEN HUMAN LIB2 6010300024 0	Homo sap
1	A5	NM_006028	Hs.241377	CGEN HUMAN LIB2 6010300032 0	Homo sap
1	A6	NM_000810	Hs.24969	CGEN HUMAN LIB2 6010300040 0	Homo sap
1	A7	NM_004857	Hs.48714	CGEN HUMAN LIB2 6010300048 0	Homo sap
1	A8	S76942	Hs.99922	CGEN HUMAN LIB2 6010300056 0	dopamine
1	A9	NM_000871	Hs.22180	CGEN HUMAN LIB2 6010300064 0	Homo sap
1	A10	NM_014286	Hs.301760	CGEN HUMAN LIB2 6010300072 0	Homo sap
1	A11	NM_002531	Hs.110642	CGEN HUMAN LIB2 6010300080 0	Homo sap
1	A12	U37028	Hs.381264	CGEN HUMAN LIB2 6010300088 0	Human leu
1	B1	NM_002924	Hs.79348	CGEN HUMAN LIB2 6010300001 0	Homo sap
1	B2	NM_001458	Hs.58414	CGEN HUMAN LIB2 6010300009 0	Homo sap
1	B3	NM_017589	Hs.128180	CGEN HUMAN LIB2 6010300017 0	Homo sap
1	B4	NM_014207	Hs.58685	CGEN HUMAN LIB2 6010300025 0	Homo sap
1	B5	NM_001702	Hs.194654	CGEN HUMAN LIB2 6010300033 0	Homo sap
1	B6	NM_002246	Hs.24040	CGEN HUMAN LIB2 6010300041 0	Homo sap
1	B7	NM_001302	Hs.374846	CGEN HUMAN LIB2 6010300049 0	Homo sap
1	B8	NM_152219	Hs.135211	CGEN HUMAN LIB2 6010300057 0	Homo sap
1	B9	NM_019074	Hs.3736	CGEN HUMAN LIB2 6010300065 0	Homo sap

Ready

http://www.chem.agilent.com

Agilent Technologies

Search in
All of Life Sciences/Chemical

Register · Login Products & Services Industries Customer Center About Agilent

[Registration Benefits](#)

[Select a Country](#)
[Contact Us](#)

Home > Products & Services > Life Sciences/Chemical Analysis > DNA Microarrays
> Catalog Microarray Kits

Catalog Microarray Kits

- [Library](#)
- [Technical Support](#)
- [Online Store](#)
- [Education](#)
- [Events](#)
- [News Releases](#)
- [Alliances](#)


Other Links

- [Stay Current with e-Notes](#)
- [Attend an e-Seminar](#)

Agilent Catalog Oligo Microarray Kits

When consistency counts, count on our catalog microarrays!

Agilent's committed to enabling the 1" x 3" microarray market with its high-quality, affordable Catalog Microarray Products featuring pre-printed microarrays with content you can trust. These microarrays are packaged with a comprehensive user guide and CD containing annotation and design files. And because they are printed with Agilent SurePrint technology, you'll enjoy unparalleled performance from microarray-to-microarray.



Key Information

- [Pricing & Availability](#)

Library Information

- [Brochures](#)
- [Application Notes](#)
- [Technical Notes](#)
- [Specifications](#)
- [Reagent Compatibility Chart](#)

Products - Microarray Kits

Animal

Agilent Microarray

The screenshot shows a Netscape browser window displaying the Agilent Technologies website. The browser's address bar shows the URL: <http://www.chem.agilent.com/Scripts/PCol.asp?IPage=494>. The website header includes the Agilent logo, a search bar, and navigation links: Register, Login, Products & Services, Industries, Customer Center, and About Agilent. Below the header, there are links for Registration Benefits, Select a Country, and Contact Us. The main content area is titled "DNA Microarrays & Scanner" and features a sidebar with links to Library, Technical Support, Online Store, Education, Events, News Releases, and Alliances. The main text describes Agilent Printed Microarray Solutions, highlighting their integrated and flexible approach for gene expression analysis. A "Special Offers" section is also present. The footer includes a link to Cancer Research and Agilent Microarrays and a mention of MOgene, LC - Agilent's Newest Certified Microarray Service.

Agilent Technologies

Search in All of Life Sciences/Chemical

[Register](#) [Login](#) [Products & Services](#) [Industries](#) [Customer Center](#) [About Agilent](#)

[Registration Benefits](#) [Select a Country](#) [Contact Us](#)

Home > [Products & Services](#) > [Life Sciences/Chemical Analysis](#) > DNA Microarrays & Scanner

DNA Microarrays & Scanner

Agilent Printed Microarray Solutions

Agilent Printed Microarray Solutions are based on an integrated, flexible, open approach for successful gene expression analysis. Each component is designed to work with others in our system, or with your existing set-up. Allowing you to express it your way. Buy just what you need or purchase the entire system.

Special Offers

NEW MICROARRAY: Whole Mouse Genome Oligo Microarray Kit
Designed to truly represent all known genes in the mouse genome and their resulting transcripts
[Review Product Information](#)

Now Available: Search Routine for Oligo Microarray Gene Lists
Search the Agilent oligo microarrays gene lists and retrieve available information to help in your research
[View Search Page](#)

MOgene, LC - Agilent's Newest Certified Microarray Service

Cancer Research and Agilent Microarrays
Welcome to a world of new cancer research possibilities - Made possible

Other Links

- [Library](#)
- [Technical Support](#)
- [Online Store](#)
- [Education](#)
- [Events](#)
- [News Releases](#)
- [Alliances](#)
- [Stay Current with e-Notes](#)
- [Attend an e-Seminar](#)

Agilent Human Microarrays

The screenshot shows a Netscape browser window displaying the Agilent website. The browser's address bar shows the URL: <http://www.chem.agilent.com/Scripts/PDS.asp?Page=13064>. The website header includes the Agilent Technologies logo, a search bar, and navigation links: Register, Login, Products & Services, Industries, Customer Center, and About Agilent. Below the header, there are links for Registration Benefits, Select a Country, and Contact Us. The main content area is titled "Whole Human Genome Oligo Microarray Kit" and features a sidebar with links to Library, Technical Support, Online Store, Education, Events, News Releases, and Alliances. The main text describes the kit as a "Double Your Data!" product, providing a compiled view of the human genome. It mentions that the sequence information was derived from a broad survey of well-known sources such as RefSeq, Goldenpath, Ensembl, Unigene, and others. The resulting view covers 41K unique genes and transcripts, which have been verified and optimized by alignment to the human genome assembly and by Agilent's Empirical Validation process. A note states that the product is intended for use in research only and requires agreement to specific licenses. To the right of the main text, there is a circular image of a human figure and a "Key Information" section with links for Pricing & Availability and Library Information. The Library Information section includes links for Brochure (0.4 MBytes), Technical Notes, Sample Amplification/Labeling Protocols, and Microarray Hybridization.

Agilent Technologies

Search in All of Life Sciences/Chemical

Register Login Products & Services Industries Customer Center About Agilent

Registration Benefits

Home > Products & Services > Life Sciences/Chemical Analysis > DNA Microarrays > Whole Human Genome Oligo Microarray Kit

Select a Country Contact Us

Whole Human Genome Oligo Microarray Kit

- Library
- Technical Support
- Online Store
- Education
- Events
- News Releases
- Alliances

Other Links


- Stay Current with e-Notes
- Attend an e-Seminar

Whole Human Genome Oligo Microarray Kit

Double Your Data!

This single microarray represents a compiled view of the human genome as it is understood today. The sequence information used to design this product was derived from a broad survey of well known sources such as RefSeq, Goldenpath, Ensembl, Unigene and others. The resulting view of the human genome covers 41K unique genes and transcripts which have been verified and optimized by alignment to the human genome assembly and by [Agilent's Empirical Validation process](#).

This product is intended for use in research only. Use of the product requires agreement to specific [licenses](#).



Key Information

- Pricing & Availability

Library Information

- Brochure (0.4 MBytes)
- Technical Notes
- Sample Amplification/Labeling Protocols
- Microarray Hybridization

Agilent Product List


Microarray Kit	Gene List			ref Guide	
	Standard file		Compressed File		
	html format	Excel format	html format	Excel format	
Oligo					
Human 1A (V2) G4110B		.xls 4.8 Mbytes		xls.zip 1.3 Mbytes	xls.zip 4.2 MBytes
Human 1B G4111A		.xls 3.2 Mbytes		xls.zip 0.8 Mbytes	xls.zip 2.5 MBytes
Human Genome, Whole G4112A		.xls 8.4 Mbytes		xls.zip 2.2 Mbytes	xls.zip 5.1 MBytes
Mouse (Dev) G4120A	.htm 10.4 Mbytes	.xls 3.8 Mbytes	htm.zip 1.0 Mbytes	xls.zip 1.8 Mbytes	xls.zip 1.9 Mbytes
Mouse G4121A	.htm 8.3 Mbytes	.xls 3.6 Mbytes	htm.zip 0.9 Mbytes	xls.zip 1.0 Mbytes	xls.zip 2.1 MBytes
NEW! Mouse Genome, Whole		.xls 10.8 Mbytes		xls.zip 2.8 Mbytes	

Agilent Oligo List

Agilent Technologies				
Whole Human Genome Oligo Microarray G4112A				
Systematic Name	Probe Name	GeneName	Additional Accessions	Description
NM_000015	A_23_P31798	NAT2	ref NM_000015.1 gb D90042 ens ENST00000286479 gb	Hom
NM_000016	A_23_P96761	ACADM	ref NM_000016.2 gb NM_000016 gb BC005377 ens ENS	Hom
NM_000017	A_23_P65022	ACADS	ref NM_000017.1 gb BC025963 ens ENST00000242592	Hom
NM_000018	A_23_P207650	ACADVL	ref NM_000018.1 gb BC000399 ens ENST00000322910	Hom
NM_000019	A_23_P24515	ACAT1	ref NM_000019.2 gb D90228 ens ENST00000265838 th	Hom
NM_000020	A_24_P324783	ACVRL1	ref NM_000020.1 gb NM_000020 gb BC042637 ens ENS	Hom
NM_000021	A_23_P205686	PSEN1	ref NM_000021.2 ref NM_007318.1 gb L76517 gb AK122	Hom
NM_000021	A_23_P340728	PSEN1	ref NM_000021.2 ref NM_007319.1 ref NM_007318.1 gb	Hom
NM_000022	A_23_P210482	ADA	ref NM_000022.1 gb BC040226 gb K00509 gb AK12398	Unkn
NM_000023	A_23_P66637	SGCA	ref NM_000023.1 ens ENST00000262018 gb U08895 gb	Hom
NM_000024	A_23_P145024	ADRB2	ref NM_000024.3 gb BC063486 gb BC012481 ens ENST	Hom
NM_000025	A_23_P168993	ADRB3	ref NM_000025.1 gb NM_000025 gb X70811 thc THC20	(Hom
NM_000026	A_23_P155103	ADSL	ref NM_000026.1 gb AF067853 ens ENST00000216194	Hom
NM_000027	A_23_P7099	AGA	ref NM_000027.2 gb NM_000027 gb X55330 ens ENST0	Hom
NM_000028	A_23_P200298	AGL	ref NM_000028.1 gb NM_000028 ref NM_000642.1 ref N	Hom
NM_000029	A_23_P115261	AGT	ref NM_000029.1 gb NM_000029 gb K02215 ens ENST0	Hom
NM_000030	A_23_P83590	AGXT	ref NM_000030.1 gb NM_000030 gb AF348451 gb X534	Hom
NM_000031	A_23_P324278	ALAD	ref NM_000031.3 gb NM_000031 gb BC000977 ens ENS	Hom
NM_000032	A_32_P385587	ALAS2	ref NM_000032.1 gb X60364 gb AF130113 gb BC03023	(Hom
NM_000033	A_23_P148556	ABCD1	ref NM_000033.2 gb BC025358 ens ENST00000218104	Hom
NM_000034	A_23_P88963	ALDOA	ref NM_000034.2 ref NM_184043.1 ref NM_184041.1 gb	Hom
NM_000035	A_23_P415974	ALDOB	ref NM_000035.2 gb X02747 ens ENST00000259398 gb	Hom

http://www.affymetrix.com

home | login | register | your profile | contact search site

AFFYMETRIX


PRODUCTS | ANALYSIS | SUPPORT | TECHNOLOGY | RESEARCH COMMUNITY | CORPORATE

アフィメトリクス・ジャパンへはこちら
 Web Talk
 October 20, 2004
 10:00am PDT

**Tools to take you
as far as your vision.™**

== RESOURCES ==
 → customers
 → investors
 → media
 → educators
 → developers
 → careers

== WHAT'S NEW ==
 → **Affymetrix Recognized with R&D 100 Award**
 → EPA Signs Research Agreement with Affymetrix to Develop Improved Tests for Identifying Hazardous Chemicals
 → Affymetrix and Broad Institute Collaborate on Three Landmark Research Projects
 → Affymetrix Technology to be Used in Groundbreaking Studies of Neurological Disease
 → Affymetrix Recognized with Outstanding Corporate Innovator Award by Product Development and Management Association

== EVENTS ==
 → **2004 Affymetrix European Seminar Tour**
 10/12/2004
 Cambridge, United Kingdom
 → All Affymetrix Events
 → **1st Affymetrix Asia Pacific User Group Meeting**
 Scientific Meeting
 November 15 - 16

== THE WAY AHEAD™ ==
 → **On The Trail of a Killer**
Phytophthora ramorum is responsible for Sudden Oak Death as well as for millions of dollars of damage to soybean crops every year. However, not much is known about this deadly microorganism. Affymetrix and the soybean genome community included *Phytophthora* on the GeneChip® Soybean Genome Array to help researchers learn more about the

Affymetrix login page

The screenshot shows a Netscape browser window titled "Affymetrix - Login - Netscape". The address bar contains "https://www.affymetrix.com/site/login/login.affx". The page features a navigation menu with links for "home", "login", "register", "your profile", and "contact", along with a "search site" input field. The Affymetrix logo is visible on the left, and a purple banner for the "Affymetrix European Seminar Series 2004" is at the top. A horizontal menu includes "PRODUCTS", "ANALYSIS", "SUPPORT", "TECHNOLOGY", "RESEARCH COMMUNITY", and "CORPORATE". The main content area is split into two columns: "AFFYMETRIX.COM MEMBERS" with a sign-in form (e-mail and password fields, "login" button, and "reset" link) and "NOT A MEMBER?" with a "register for free" link and "register" button. At the bottom, there are links for "Login Problems?" and "Forgot your password?". The footer contains contact information: "888-DNA-CHIP (888-362-2447) | +44 (0) 1628 552550" and links for "feedback", "e-mail support", "terms of use", and "privacy policy".

home | login | register | your profile | contact search site

AFFYMETRIX
Affymetrix European Seminar Series 2004

PRODUCTS | ANALYSIS | SUPPORT | TECHNOLOGY | RESEARCH COMMUNITY | CORPORATE

AFFYMETRIX.COM MEMBERS
Sign in here. Fields are case sensitive.
e-mail:
password:
login reset

NOT A MEMBER?
Register for free.
register

[Login Problems?](#) | [Forgot your password?](#)

888-DNA-CHIP (888-362-2447) | +44 (0) 1628 552550 feedback | e-mail support | terms of use | privacy policy

Return to the home page

home | login | register | your profile | contact search site

AFFYMETRIX

PRODUCTS | ANALYSIS | SUPPORT | TECHNOLOGY | RESEARCH COMMUNITY | CORPORATE

Tools to take you
as far as your vision.™

Web Talk
October 20, 2004
10:00am PDT

アフィメトリス・ジャパンへはこちら

≡ RESOURCES

- customers
- investors
- media
- educators
- developers
- careers

≡ WHAT'S NEW

- Affymetrix Recognized with R&D 100 Award
- EPA Signs Research Agreement with Affymetrix to Develop Improved Tests for Identifying Hazardous Chemicals
- Affymetrix and Broad Institute Collaborate on Three Landmark Research Projects
- Affymetrix Technology to be Used in Groundbreaking Studies of Neurological Disease
- Affymetrix Recognized with Outstanding Corporate Innovator Award by Product Development and Management Association

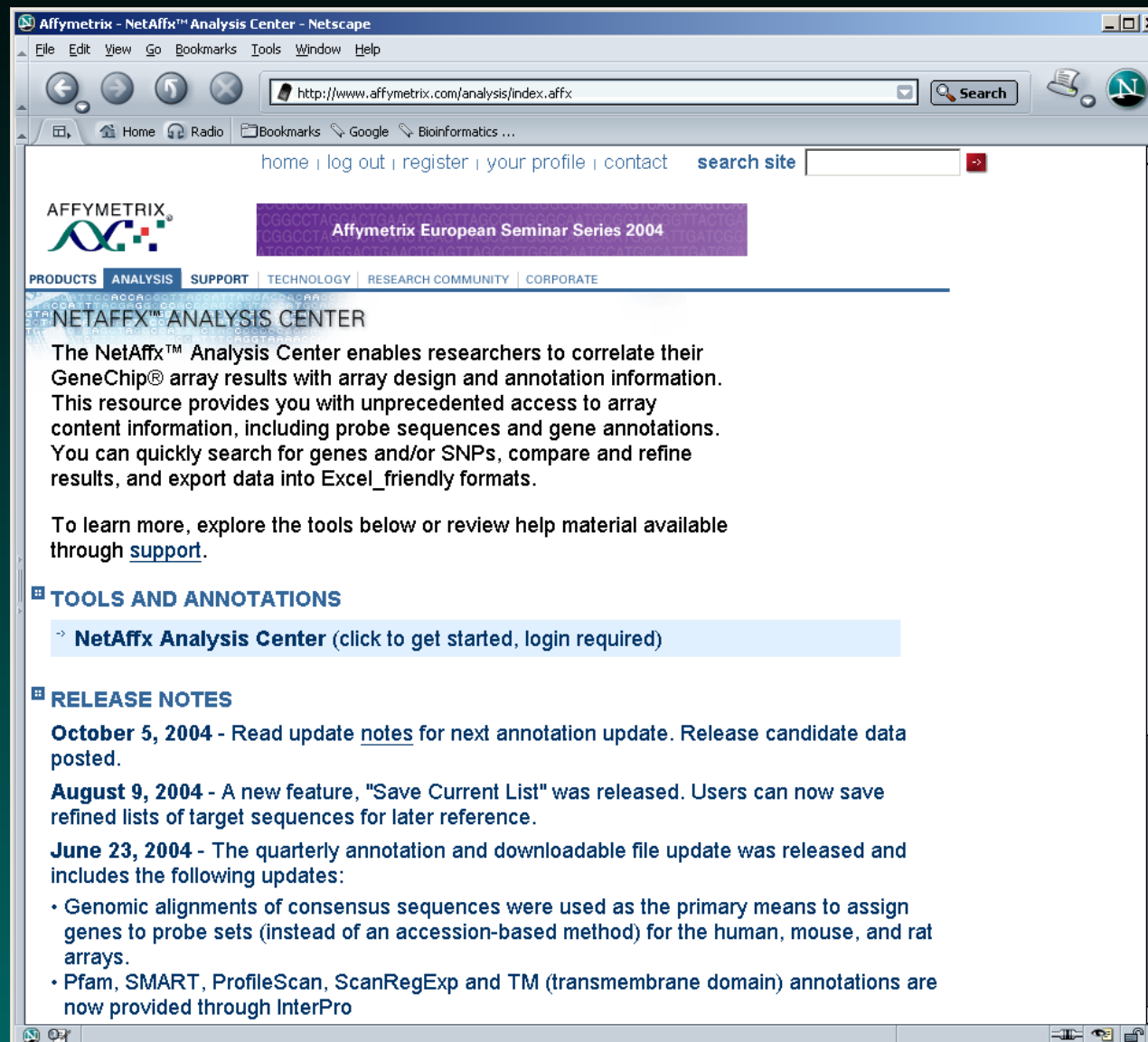
≡ EVENTS

- **2004 Affymetrix European Seminar Tour**
10/12/2004
Cambridge, United Kingdom
- All Affymetrix Events
- **1st Affymetrix Asia Pacific User Group Meeting Scientific Meeting**
November 15 - 16

≡ THE WAY AHEAD™

- **On The Trail of a Killer**
Phytophthora ramorum is responsible for Sudden Oak Death as well as for millions of dollars of damage to soybean crops every year. However, not much is known about this deadly microorganism. Affymetrix and the soybean genome community included *Phytophthora* on the GeneChip® Soybean Genome Array to help researchers learn more about the

Main page for NetAffx



A screenshot of a Netscape browser window displaying the main page of the NetAffx Analysis Center. The browser's address bar shows the URL <http://www.affymetrix.com/analysis/index.affx>. The page features the Affymetrix logo and a navigation menu with links for PRODUCTS, ANALYSIS, SUPPORT, TECHNOLOGY, RESEARCH COMMUNITY, and CORPORATE. A prominent purple banner advertises the "Affymetrix European Seminar Series 2004". Below the navigation, the page is titled "NETAFFX™ ANALYSIS CENTER" and provides a detailed description of the tool's capabilities, such as correlating GeneChip array results with design and annotation information. A section titled "TOOLS AND ANNOTATIONS" includes a link to the "NetAffx Analysis Center" (noting that login is required). A "RELEASE NOTES" section lists updates from October 5, 2004, August 9, 2004, and June 23, 2004, detailing changes in annotation and file updates.

home | log out | register | your profile | contact

AFFYMETRIX

Affymetrix European Seminar Series 2004

PRODUCTS | ANALYSIS | SUPPORT | TECHNOLOGY | RESEARCH COMMUNITY | CORPORATE

NETAFFX™ ANALYSIS CENTER

The NetAffx™ Analysis Center enables researchers to correlate their GeneChip® array results with array design and annotation information. This resource provides you with unprecedented access to array content information, including probe sequences and gene annotations. You can quickly search for genes and/or SNPs, compare and refine results, and export data into Excel_friendly formats.

To learn more, explore the tools below or review help material available through [support](#).

TOOLS AND ANNOTATIONS

→ [NetAffx Analysis Center](#) (click to get started, login required)

RELEASE NOTES

October 5, 2004 - Read update [notes](#) for next annotation update. Release candidate data posted.

August 9, 2004 - A new feature, "Save Current List" was released. Users can now save refined lists of target sequences for later reference.

June 23, 2004 - The quarterly annotation and downloadable file update was released and includes the following updates:

- Genomic alignments of consensus sequences were used as the primary means to assign genes to probe sets (instead of an accession-based method) for the human, mouse, and rat arrays.
- Pfam, SMART, ProfileScan, ScanRegExp and TM (transmembrane domain) annotations are now provided through InterPro

Affymetrix Analysis Central

home | log out | register | your profile | contact search site

AFFYMETRIX
Custom SNP Genotyping. Flexibility and Performance. Find out more.

PRODUCTS ANALYSIS SUPPORT TECHNOLOGY RESEARCH COMMUNITY CORPORATE

QUERY
Getting Started

Expression

- Search all available information in the database for a particular term or identifier. This is recommended as a starting point for your searches. [Quick Query]
- Search specific fields in the database for a term or identifier [Standard Query]
- Retrieve annotations for a probe list [Batch Query]
- Find probe sets that align to your sequence(s) through BLAST [BLAST]
- Find probes that identically match your sequence(s) [Probe Match]
- Query the UCSC Browser for genomic alignment [UCSC Query]

Genotyping

- Search all available information in the database for a particular term or identifier. This is recommended as a starting point for your searches. [Quick Query]
- Search specific fields in the database for a term or identifier [Standard Query]
- Retrieve annotations for a probe list [Batch Query]
- Query the UCSC public genome by position [UCSC Query]
- Search for SNPs between microsatellites [SNP Finder]

Begin

GETTING STARTED
-> Wizard

QUERY
Expression
-> Quick Query
-> Standard Query
-> Batch Query
-> BLAST
-> Probe Match
-> UCSC Query
Genotyping
-> Quick Query
-> Standard Query
-> Batch Query
-> UCSC Query
-> SNP Finder

CURRENT QUERY
88 probe sets
-> Annotations
-> Show Orthologs
-> GO Browser
-> Export

QUERY HISTORY
Annotation Views

Affymetrix Quick Search

Affymetrix - Quick Query - Netscape

File Edit View Go Bookmarks Tools Window Help

https://www.affymetrix.com/analysis/netaffx/quickquery.affx

home | log out | register | your profile | contact search site

AFFYMETRIX

Web Talk
October 20, 2004. 10:00am PDT

PRODUCTS ANALYSIS SUPPORT TECHNOLOGY RESEARCH COMMUNITY CORPORATE

GETTING STARTED
-> Wizard

QUERY
Expression
-> Quick Query
-> Standard Query
-> Batch Query
-> BLAST
-> Probe Match
-> UCSC Query
Genotyping
-> Quick Query
-> Standard Query
-> Batch Query
-> UCSC Query
-> SNP Finder

QUERY HISTORY

Annotation Views
-> Expression
-> Genotyping

BLAST
-> Status
-> New Folder

Expression Queries
-> All Descriptions

QUERY
Quick Query

1. Select a GeneChip Array:
(Use control-select to search up to three arrays simultaneously.)

Human Genome U133 Plus 2.0 Array
Mouse Genome 430 2.0 Array
Mouse Genome 430A 2.0 Array
Human Genome U133 Set
Human Genome U95 Set

2. Enter search terms:
Query All Descriptions

Leave empty to get all probe sets for the selected array(s).
Use & for AND, | for OR, and ! for NOT between terms. See Query Examples
Quick query adds wildcards(%) at the front and end of each term.

search

http://www.affymetrix.com/technology/index.affx

Affymetrix Batch Search

AFFYMETRIX
GeneChip[®] Yeast Genome 2.0 Array

PRODUCTS | ANALYSIS | SUPPORT | TECHNOLOGY | RESEARCH COMMUNITY | CORPORATE

GETTING STARTED
-> Wizard

QUERY
Expression
-> Quick Query
-> Standard Query
-> Batch Query
-> BLAST
-> Probe Match
-> UCSC Query
Genotyping
-> Quick Query
-> Standard Query
-> Batch Query
-> UCSC Query
-> SNP Finder

CURRENT QUERY
88 probe sets
-> Annotations
-> Show Orthologs
-> GO Browser
-> Export

QUERY HISTORY
Annotation Views
-> Expression

Batch Query
Rapidly obtain results for up to 3,000 Affymetrix[®] probe set accession numbers, gene names, or sequences ids at once.

1. Select a GeneChip Array:
(Use control-select to search up to three arrays simultaneously.)
Human Genome U133 Plus 2.0 Array
Mouse Genome 430 2.0 Array
Mouse Genome 430A 2.0 Array
Human Genome U133 Set
Human Genome U95 Set

2. Select a search option:
Probe Set ID

3. Upload a text (*.txt) file with the appropriate query values: [Example](#)
Browse...

Note: Query values should be separated by a new line (\n) character. A Microsoft Excel file should be saved as "text only" file before uploading.

4. Select a view:
* Annotation List *

5. Name Query:(Optional)
(95 Character max)

search ->

Affymetrix Search Results

home | log out | register | your profile | contact

AFFYMETRIX
2 organisms, one array

PRODUCTS | **ANALYSIS** | SUPPORT | TECHNOLOGY | RESEARCH COMMUNITY | CORPORATE

GETTING STARTED
 -> Wizard
QUERY
Expression
 -> Quick Query
 -> Standard Query
 -> Batch Query
 -> BLAST
 -> Probe Match
 -> UCSC Query
Genotyping
 -> Quick Query
 -> Standard Query
 -> Batch Query
 -> UCSC Query
 -> SNP Finder
CURRENT QUERY
76 probe sets
 -> Annotations
 -> Show Orthologs
 -> GO Browser
 -> Export
QUERY HISTORY
 Annotation

QUERY Results

Results 1-50 of 76.
[Export](#) | [GO Browser](#) | [Show Orthologs](#)
[Full Screen](#) | [First Page](#) | [Preceding Page](#) | [Next Page](#) | [Last Page](#) | [* Annotation List *](#)

Details	Probe Set ID	Gene Title	Gene Symbol	GO Biological Process Description	GO Molecular Function Description
<input type="checkbox"/> Details	1569449_a_at	cofactor required for Sp1 transcriptional activation, subunit 2, 150kDa	CRSP2	regulation of transcription from Pol II promoter transcription initiation from Pol II promoter androgen receptor signaling pathway	receptor activity RNA polymerase II transcription mediator activity transcriptional activator activity ligand-dependent nuclear receptor transcription coactivator activity vitamin D receptor binding thyroid hormone receptor binding
<input type="checkbox"/> Annotation	1570188_at	nuclear receptor	NR113	regulation of	transcription factor

Affymetrix Detailed Search Results

NETAFFX™ ANALYSIS CENTER

Full Record

Details for HG-U133_PLUS_2:211621_AT
Full Screen

NetAffx Links	Cluster Members Consensus/Exemplar Probe Set Display
---------------	--

GeneChip Array Information

Probe Set ID	211621_at
GeneChip Array	Human Genome U133 Plus 2.0 Array
Organism Common Name	Human

Probe Design Information

Transcript ID	g178655
Sequence Type	Exemplar sequence
Representative Public ID	M73069 NCBI
Target Description	gb:M73069.1 /DB_XREF=gi:178655 /GEN=AR /FEA=FLmRNA /CNT=1 /TID=HsAffx.900362.564 /TIER=FL /STK=0 /DEF=Human androgen receptor mutant gene, mRNA, complete cds. /PROD=androgen receptor /FL=gb:M73069.1

Annotation Method Description

Annotation Description	This probe set was annotated using the Genome based pipeline to a Locus Link identifier using 2 transcripts.
Annotation Transcript	

Navigation Sidebar:

- GETTING STARTED
 - Wizard
- QUERY
 - Expression
 - Quick Query
 - Standard Query
 - Batch Query
 - BLAST
 - Probe Match
 - UCSC Query
 - Genotyping
 - Quick Query
 - Standard Query
 - Batch Query
 - UCSC Query
 - SNP Finder
- CURRENT QUERY
 - 76 probe sets
 - Annotations
 - Show Orthologs
 - GO Browser
 - Export
- QUERY HISTORY
 - Annotation Views
 - Expression
 - Genotyping
 - BLAST
 - Status
 - New Folder

Affymetrix Sequence Information

Sequence

>HG-U133_PLUS_2:211110_S_AT
 tgcaactccaggatgctctacttcgcccctgatctggttttcaatgagtaccgcatgca
 caagtcccggatgtacagccagtggtgcccgaatgaggcacctctctcaagagtttgatg
 gctccaaatcacccccaggaattcctgtgcatgaaagccatgctactcttcagcattat
 tccagtggatgggctgaaaaatcaaaaattctttgatgaacttcgaatgaactacatcaa
 ggaactcgatcgtatcattgcatgcaaaagaaaaatcccacatcctgctcaagacgctt
 ctaccagctcaccagctcctggactccgtgcacacctattgagagagctgcacagtt
 cacttttgacctgctaataagtcacacatggtagcgtggactttccggaaatgatggc
 agagatcactctctgtgcaagtgcccagatcctttctgggaaagtcagcccactctattt
 ccacaccagtgaaagcattggaacctatttccccaccccagctcatgcccccttccag
 atgtcttctgctgttataactctg

Target Sequence

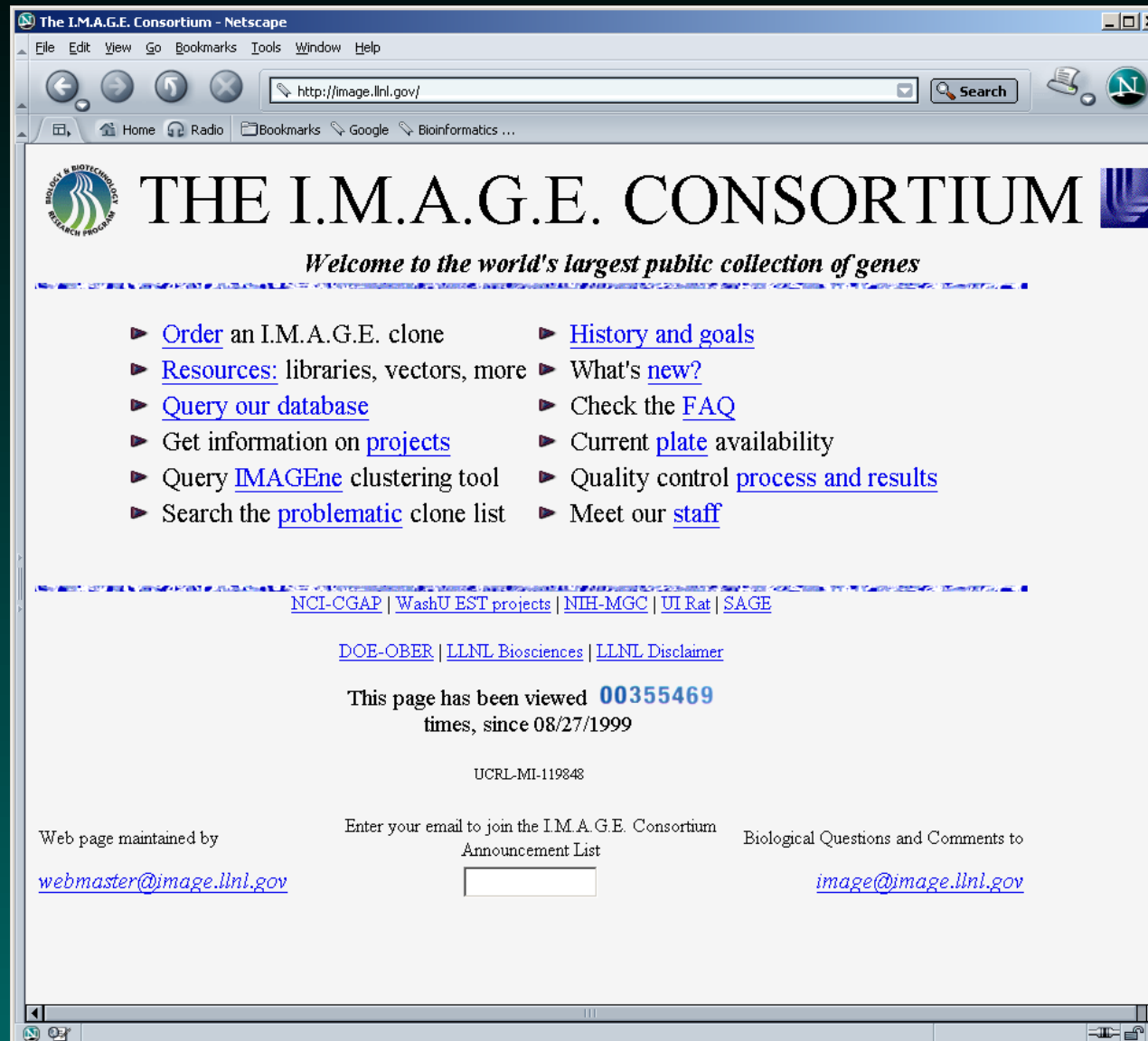
BLASTn GenBank NR

Probe Sequence(5'-3')	Probe X	Probe Y	Probe Interrogation Position	Strandedness
TGTCAACTCCAGGATGCTCTACTTC	344	925	2244	Antisense
CGCATGCACAAGTCCCGGATGTACA	571	417	2296	Antisense
GGATGTACAGCCAGTGTGCCGAAT	17	839	2312	Antisense
TGTCCGAATGAGGCACCTCTCTCAA	817	921	2328	Antisense
ACCCCCCAGGAATTCCTGTGCATGA	589	159	2374	Antisense
GGAACTCGATCGTATCATTGCATGC	488	863	2484	Antisense
GAGCTGCATCAGTTCACTTTTGACC	399	629	2590	Antisense
ATGGTGAGCGTGGACTTTCCGGAAA	446	67	2632	Antisense
AGAGATCATCTCTGTGCAAGTGCCC	134	115	2664	Antisense
GAAAGTCAAGCCCATCTATTTCCAC	1158	591	2703	Antisense
ATGTCTTCTGCCTGTTATAACTCTG	135	65	2784	Antisense

Probe Info

Sequence Source GenBank

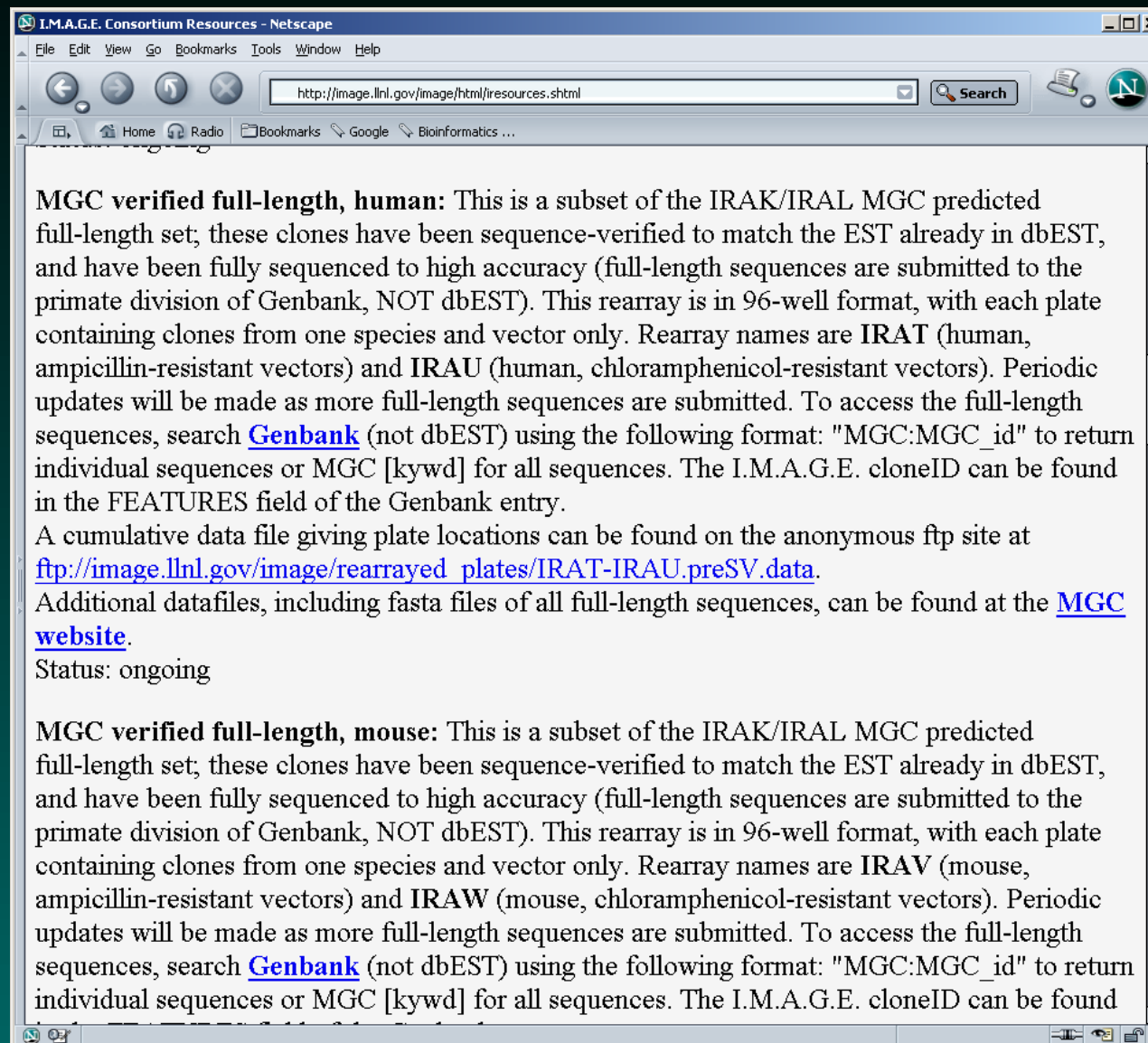
http://image.llnl.gov



The screenshot shows a Netscape browser window titled "The I.M.A.G.E. Consortium - Netscape". The address bar contains "http://image.llnl.gov/". The website content includes the following elements:

- Logo:** A circular logo with the text "BIOLOGY & BIOTECHNOLOGY RESEARCH PROGRAM" around a central graphic.
- Section Header:** "THE I.M.A.G.E. CONSORTIUM" in large, bold, black letters.
- Text:** "Welcome to the world's largest public collection of genes" in italics.
- Navigation Links:** A list of links with blue arrowheads:
 - Order an I.M.A.G.E. clone
 - Resources: libraries, vectors, more
 - Query our database
 - Get information on projects
 - Query IMAGEne clustering tool
 - Search the problematic clone list
 - History and goals
 - What's new?
 - Check the FAQ
 - Current plate availability
 - Quality control process and results
 - Meet our staff
- Footer Links:**
 - NCI-CGAP | WashU EST projects | NIH-MGC | UI Rat | SAGE
 - DOE-OBER | LLNL Biosciences | LLNL Disclaimer
- View Counter:** "This page has been viewed 00355469 times, since 08/27/1999"
- UCRL-MI-119848**
- Contact Information:**
 - Web page maintained by webmaster@image.llnl.gov
 - Enter your email to join the I.M.A.G.E. Consortium Announcement List:
 - Biological Questions and Comments to image@image.llnl.gov

IMAGE is integrated into GenBank



MGC verified full-length, human: This is a subset of the IRAK/IRAL MGC predicted full-length set; these clones have been sequence-verified to match the EST already in dbEST, and have been fully sequenced to high accuracy (full-length sequences are submitted to the primate division of Genbank, NOT dbEST). This rearray is in 96-well format, with each plate containing clones from one species and vector only. Rearray names are IRAT (human, ampicillin-resistant vectors) and IRAU (human, chloramphenicol-resistant vectors). Periodic updates will be made as more full-length sequences are submitted. To access the full-length sequences, search [Genbank](#) (not dbEST) using the following format: "MGC:MGC_id" to return individual sequences or MGC [kywd] for all sequences. The I.M.A.G.E. cloneID can be found in the FEATURES field of the Genbank entry.

A cumulative data file giving plate locations can be found on the anonymous ftp site at ftp://image.llnl.gov/image/rearrayed_plates/IRAT-IRAU.preSV.data.

Additional datafiles, including fasta files of all full-length sequences, can be found at the [MGC website](#).

Status: ongoing

MGC verified full-length, mouse: This is a subset of the IRAK/IRAL MGC predicted full-length set; these clones have been sequence-verified to match the EST already in dbEST, and have been fully sequenced to high accuracy (full-length sequences are submitted to the primate division of Genbank, NOT dbEST). This rearray is in 96-well format, with each plate containing clones from one species and vector only. Rearray names are IRAV (mouse, ampicillin-resistant vectors) and IRAW (mouse, chloramphenicol-resistant vectors). Periodic updates will be made as more full-length sequences are submitted. To access the full-length sequences, search [Genbank](#) (not dbEST) using the following format: "MGC:MGC_id" to return individual sequences or MGC [kywd] for all sequences. The I.M.A.G.E. cloneID can be found

GenBank: the common denominator

You may have noticed that all the commercial products map their proprietary identifiers to GenBank. GenBank is the primary repository for sequence information.

While the IMAGE ID refers to the actual clone, the corresponding GenBank entry (or entries) describes the actual mRNA sequence.

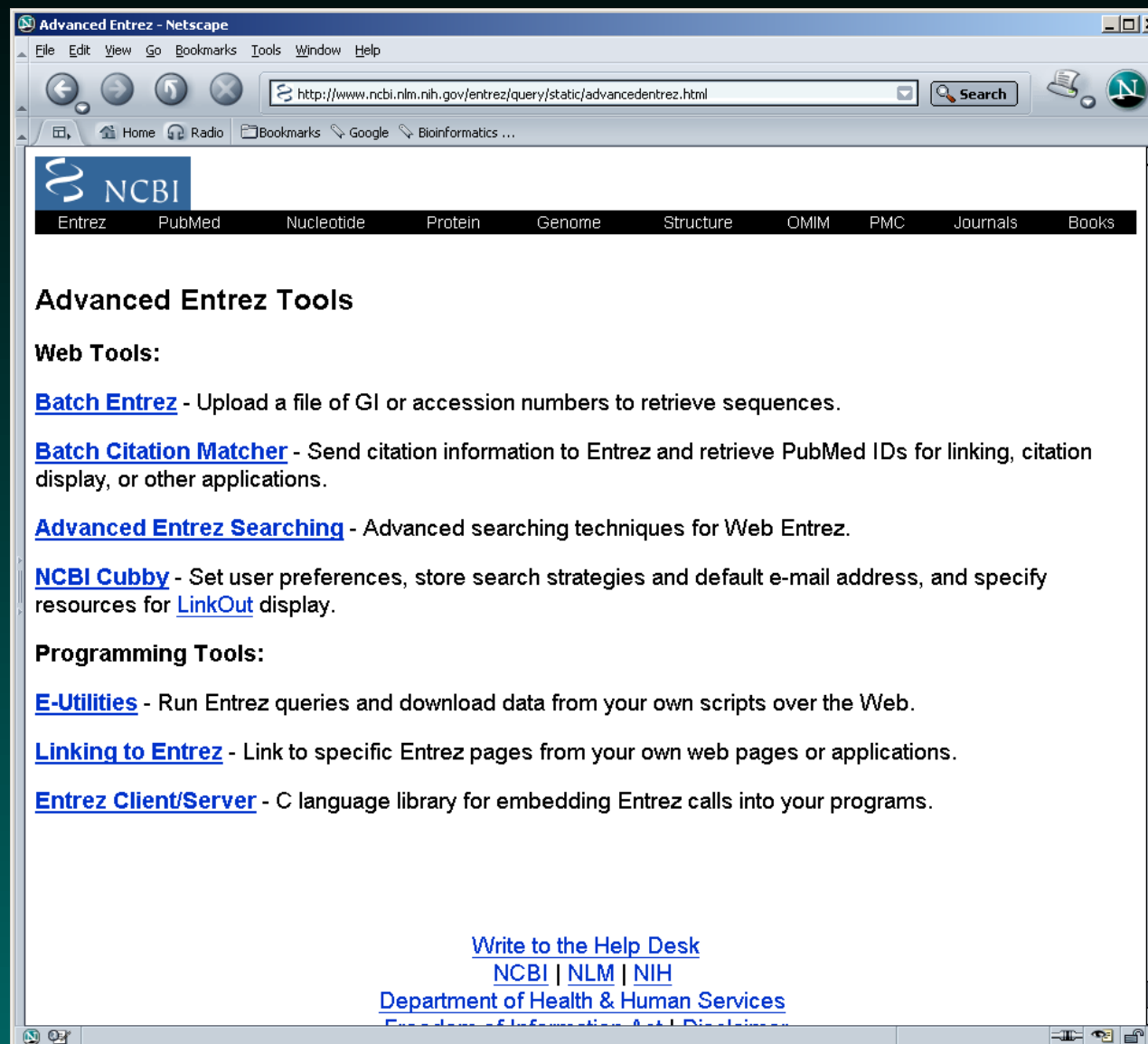
For Affymetrix, the GenBank entry describes the sequence from which the probe set was constructed.

For Sigma-Genosys or Agilent, the GenBank entry describes the sequence from which the long oligo was selected.

Entrez Nucleotide

The screenshot shows the Entrez Nucleotide database homepage in a Netscape browser window. The browser's address bar displays the URL `http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide`. The page features the NCBI logo and a navigation menu with tabs for Entrez, PubMed, Nucleotide, Protein, Genome, Structure, PMC, Taxonomy, and Books. A search bar is present with the text "Search Nucleotide for" and "Go" and "Clear" buttons. Below the search bar, there are tabs for "Limits", "Preview/Index", "History", "Clipboard", and "Details". The main content area includes a yellow highlighted text block stating: "The Entrez Nucleotides database is a collection of sequences from several sources, including GenBank, RefSeq, and PDB. The number of bases grows at an exponential rate. As of April 2004, there are over 38,989,342,565 bases." Below this is a blue-bordered box titled "Human Genome" with the text: "Explore [human genome resources](#) or browse the human genome sequence using the [Map Viewer](#)." At the bottom, there is a purple header for "Building the human genome" followed by text: "The Human Genome Reference DNA Sequence was completed in April 2003. The current version is listed as a build number on the [Genome View](#) page and includes an accompanying set of [statistics](#) and [release notes](#)." A left sidebar contains links for "About Entrez", "Entrez Nucleotide Help | FAQ", "Entrez Tools", "Check sequence revision history", "LinkOut", "Related resources BLAST", and "Reference sequence project".

Entrez Tools



The screenshot shows a Netscape browser window titled "Advanced Entrez - Netscape". The address bar contains the URL "http://www.ncbi.nlm.nih.gov/entrez/query/static/advancedentrez.html". The browser's menu bar includes "File", "Edit", "View", "Go", "Bookmarks", "Tools", "Window", and "Help". The browser's toolbar includes navigation buttons (back, forward, home, stop), a search box, and a "Search" button. The browser's status bar shows "Home", "Radio", "Bookmarks", "Google", and "Bioinformatics ...".

The main content area of the browser displays the NCBI logo and a navigation menu with the following items: Entrez, PubMed, Nucleotide, Protein, Genome, Structure, OMIM, PMC, Journals, and Books. Below the navigation menu, the page title is "Advanced Entrez Tools".

The page content is organized into two main sections:

- Web Tools:**
 - [Batch Entrez](#) - Upload a file of GI or accession numbers to retrieve sequences.
 - [Batch Citation Matcher](#) - Send citation information to Entrez and retrieve PubMed IDs for linking, citation display, or other applications.
 - [Advanced Entrez Searching](#) - Advanced searching techniques for Web Entrez.
 - [NCBI Cubby](#) - Set user preferences, store search strategies and default e-mail address, and specify resources for [LinkOut](#) display.
- Programming Tools:**
 - [E-Utilities](#) - Run Entrez queries and download data from your own scripts over the Web.
 - [Linking to Entrez](#) - Link to specific Entrez pages from your own web pages or applications.
 - [Entrez Client/Server](#) - C language library for embedding Entrez calls into your programs.

At the bottom of the page, there are links for "Write to the Help Desk", "NCBI | NLM | NIH", and "Department of Health & Human Services".

Batch Entrez

Entrez Nucleotide - Netscape

http://www.ncbi.nlm.nih.gov/entrez/batchentrez.cgi?db=Nucleotide

NCBI Nucleotide

Entrez PubMed Nucleotide Protein Genome Structure PMC Taxonomy Books

Database: Nucleotide File: Browse... Retrieve

Batch Entrez

Batch Entrez has changed!

You will need Batch Entrez now only to upload a file of GI or accession numbers for an Entrez search. *You can do all other large searches directly within Entrez.*

There are now two ways to retrieve large amounts of sequence data using Entrez:

1. Begin at one of the [Entrez](#) pages and perform a large search just like any other Entrez search:
 - Enter a query
 - Designate the database as Nucleotide or Protein, as appropriate
 - Press Go; you will see a list of document summaries

Getting a Gene List

```
> extreme <- abs(t.statistics)>6 & tr.stats > 20
> extreme.genes <- gene.info[extreme,]
> dim(extreme.genes)
[1] 1005      6
> extreme.genes[1:8,]
      Clone.ID Accession Gene.Symbol Cluster.ID
X24     IMAGE:34849   R20379      EEF2     Hs.75309
X27     IMAGE:45525   H08440      RFP      Hs.440382
X63     IMAGE:295831  N74602     CGI-26    Hs.24332
X82     IMAGE:143322  R74357      Hs.367688
X100    IMAGE:753381  AA406332    SEC23A    Hs.272927
X132    IMAGE:341083  W58562     C6orf56   Hs.102471
X205    IMAGE:78946   T61792     PDK4      Hs.8364
X225    IMAGE:586831  AA130866    TMLHE     Hs.133321
```

Just GenBank

```
> extreme.gb <- extreme.genes$Accession
> write.table(extreme.gb, 'extreme.txt',
+ quote=FALSE, row.names=FALSE, col.names=FALSE)
> write.table(extreme.genes, 'extreme-genes.txt',
+ quote=FALSE, row.names=FALSE, col.names=TRUE,
+ sep='\t')
```

We need the one-column list of GenBank identifiers for a batch search of GenBank. Other tools can use the more general list.

Batch Entrez

Entrez Nucleotide - Netscape

http://www.ncbi.nlm.nih.gov/entrez/batchentrez.cgi?db=Nucleotide

NCBI Nucleotide

Entrez PubMed Nucleotide Protein Genome Structure PMC Taxonomy Books

Database: Nucleotide File: Browse... Retrieve

Batch Entrez

Batch Entrez has changed!

You will need Batch Entrez now only to upload a file of GI or accession numbers for an Entrez search. *You can do all other large searches directly within Entrez.*

There are now two ways to retrieve large amounts of sequence data using Entrez:

1. Begin at one of the [Entrez](#) pages and perform a large search just like any other Entrez search:
 - Enter a query
 - Designate the database as Nucleotide or Protein, as appropriate
 - Press Go; you will see a list of document summaries

Entrez Nucleotide Help | FAQ Entrez Tools Check sequence revision history LinkOut Related resources BLAST Reference sequence

Entrez Results

The screenshot shows a Netscape browser window displaying the Entrez Nucleotide search results. The browser's address bar shows the URL <http://www.ncbi.nlm.nih.gov/entrez/batchentrez.cgi>. The page features the NCBI logo and a navigation menu with tabs for Entrez, PubMed, Nucleotide, Protein, Genome, Structure, PMC, Taxonomy, and Books. The search criteria are set to 'Nucleotide' with 'Go' and 'Clear' buttons. Below the search bar, there are options for 'Limits', 'Preview/Index', 'History', 'Clipboard', and 'Details'. The 'Display' is set to 'Summary', 'Show' is set to '20', and 'Send to' is set to 'Text'. The results are displayed as 'Items 1 - 20 of 1003' on 'Page 1 of 51 Next'. Three results are visible:

- 1: [R20379](#)** Links
 yg40f10.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
 IMAGE:34849 5' similar to gb:X51466 ELONGATION FACTOR 2
 (HUMAN);, MRNA sequence
 gi|775013|gb|R20379.1|[775013]
- 2: [H08440](#)** Links
 yl89c06.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
 IMAGE:45525 5' similar to contains Alu repetitive element;contains MSR1
 repetitive element ;, MRNA sequence
 gi|873262|gb|H08440.1|[873262]
- 3: [N74602](#)** Links
 za48h04.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
 IMAGE:295831 3', MRNA sequence
 gi|1231887|gb|N74602.1|[1231887]

The left sidebar contains links for 'About Entrez', 'Entrez Nucleotide Help | FAQ', 'Entrez Tools', 'Check sequence revision history', 'LinkOut', 'Related resources BLAST', and 'Reference sequence project'. The bottom of the browser window shows the status bar with 'Search for'.

Entrez Results

Why did we only get 1003 results from a batch query with 1005 entries? Probably because two of the GenBank identifiers have been “retired” since the microarray was constructed. This typically happens when someone finds out there was a problem (often vector contamination) with the original sequence entry.

Note: To save the results, click the “Send to” button after first changing the destination to “File”.

We can use the results of the search to compare the IMAGE clone IDs supplied with the original array with the current record in GenBank. It is certainly possible that someone (in the distant past) typed one of the numbers incorrectly.

UniGene

GenBank only refers to individual sequences. Because lots of people have sequenced lots of fragments of RNA and entered them into GenBank, a sequence entry is not the same thing as a gene.

UniGene is the NCBI's attempt to organize sequences into coherent clusters that should represent genes.

Critical Fact: UniGene changes regularly. The current sequence data is reclustered about once a month. The presence of new sequence information can change the clusters. As of two or three years ago, as many as 25% of the UniGene cluster assignments changed over the course of a year. One hopes that the rate of change is decreasing.

UniGene Home Page

The screenshot shows the UniGene home page in a Netscape browser window. The browser's address bar displays the URL `http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene`. The page header includes the NCBI logo and the UniGene logo with the tagline "ORGANISED VIEW OF THE TRANSCRIPTOME". A navigation menu at the top lists various databases: Entrez, PubMed, Nucleotide, Protein, Genome, Structure, PMC, Taxonomy, and Books. Below this is a search bar containing the text "UniGene" and buttons for "Go" and "Clear".

The main content area is divided into two columns. The left column contains a sidebar with links for "NCBI", "UniGene", "Homepage", "Query Tips", "FAQ", "DDD", "Download", "UniGene", "Related Resources", "LocusLink", "HomoloGene", "dbEST", "Trace Archive", "BLAST", "CGAP", "MGC cDNA", and "clones". The right column contains a descriptive paragraph about UniGene and a table of UniGene entries.

UniGene is an experimental system for automatically partitioning GenBank sequences into a non-redundant set of gene-oriented clusters. Each UniGene cluster contains sequences that represent a unique gene, as well as related information such as the tissue types in which the gene has been expressed and map location.

Species	UniGene Entries
Chordata	
Mammalia	
Bos taurus	25,713
Canis familiaris	15,694
Homo sapiens	54,560
Mus musculus	46,544
Ovis aries	3,154
Rattus norvegicus	40,329
Sus scrofa	24,028
Aves	
Gallus gallus	21,371
Amphibia	

UniGene Results

The screenshot shows the UniGene web interface in a Netscape browser window. The browser's address bar displays the URL: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=unigene>. The page features the NCBI logo and the UniGene logo with the tagline "ORGANISED VIEW OF THE TRANSCRIPTOME". A navigation bar includes links for Entrez, PubMed, Nucleotide, Protein, Genome, Structure, PMC, Taxonomy, and Books. The search bar contains the text "UniGene" and "for R20379", with "Go" and "Clear" buttons. Below the search bar are tabs for "Limits", "Preview/Index", "History", "Clipboard", and "Details". A "Display" dropdown is set to "Summary", "Show" is set to "20", and "Sort" is set to "Text". The results section shows "Items 1 - 2 of 2" and "One page".

Item 1: [Hs.75309](#) MGC cDNA clone, Li
EEF2: Eukaryotic translation elongation factor 2
Homo sapiens, 1493 sequence(s)

Item 2: [At.47255](#) Li
Glycine-rich protein (GRP16)
Arabidopsis thaliana, 8 sequence(s)

The left sidebar contains a menu with the following items: NCBI, UniGene, Homepage, Query Tips, FAQ, DDD, Download, UniGene, Related Resources, LocusLink, HomoloGene, dbEST, Trace Archive, BLAST, CGAP, MGC cDNA clones.

UniGene Details

The screenshot shows the UniGene website interface in a Netscape browser window. The address bar displays the URL: <http://www.ncbi.nlm.nih.gov/UniGene/cluster.cgi?ORG=Hs&CID=75309>. The page features the NCBI logo and the UniGene logo with the tagline "ORGANISED VIEW OF THE TRANSCRIPTOME". Navigation tabs include PubMed, Nucleotide, Protein, Genome, Structure, Popset, and Taxonomy. A search bar contains the text "UniGene" and buttons for "Go" and "Clear". Below the search bar are links for "Limits", "Preview/Index", "History", "Clipboard", and "Details".

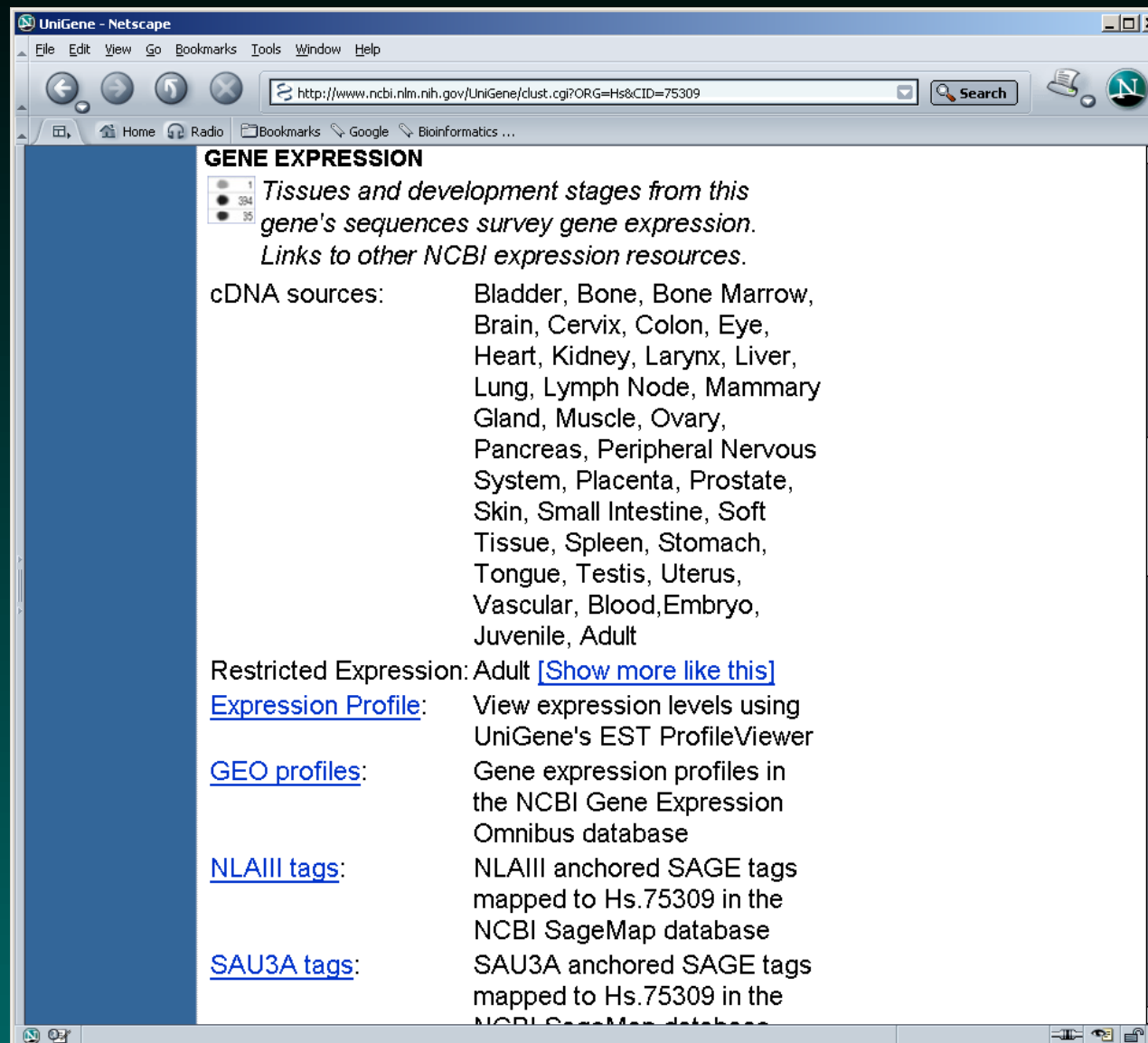
The main content area displays the following information:

- NCBI** (left sidebar menu)
- UniGene** (left sidebar menu)
- UniGene Cluster Hs.75309 Homo sapiens Eukaryotic translation elongation factor 2 (EEF2)** (main title)
- SELECTED PROTEIN SIMILARITIES** (section header)
- Comparison of sequences in UniGene with proteins supported by a complete genome. The alignments can suggest function of a gene.* (description)
- A. thaliana** [ref:NP_172112.1](#) 39.59 % / 846 aa - unknown (see [ProtEST](#)) protein [Arabidopsis thaliana]
- C. elegans** [ref:NP_492457.1](#) 78.74 % / 858 aa - Elongation factor Tu family (see [ProtEST](#))
- D. melanogaster** [pir:S05988](#) - S05988 78.42 % / 858 aa (see [ProtEST](#)) translation elongation factor

The left sidebar menu includes the following items:

- NCBI
- UniGene
- Homepage
- Query Tips
- FAQs
- DDD
- Download
- UniGene
- Related Resources
- LocusLink
- HomoloGene
- dbEST
- Trace Archive
- BLAST
- CGAP
- MGC cDNA clones

UniGene Expression Information



The screenshot shows a Netscape browser window with the address bar containing the URL <http://www.ncbi.nlm.nih.gov/UniGene/clust.cgi?ORG=Hs&CID=75309>. The page content is titled "GENE EXPRESSION" and includes the following information:

GENE EXPRESSION
1
394
35
*Tissues and development stages from this gene's sequences survey gene expression.
Links to other NCBI expression resources.*

cDNA sources: Bladder, Bone, Bone Marrow, Brain, Cervix, Colon, Eye, Heart, Kidney, Larynx, Liver, Lung, Lymph Node, Mammary Gland, Muscle, Ovary, Pancreas, Peripheral Nervous System, Placenta, Prostate, Skin, Small Intestine, Soft Tissue, Spleen, Stomach, Tongue, Testis, Uterus, Vascular, Blood, Embryo, Juvenile, Adult

Restricted Expression: Adult [\[Show more like this\]](#)

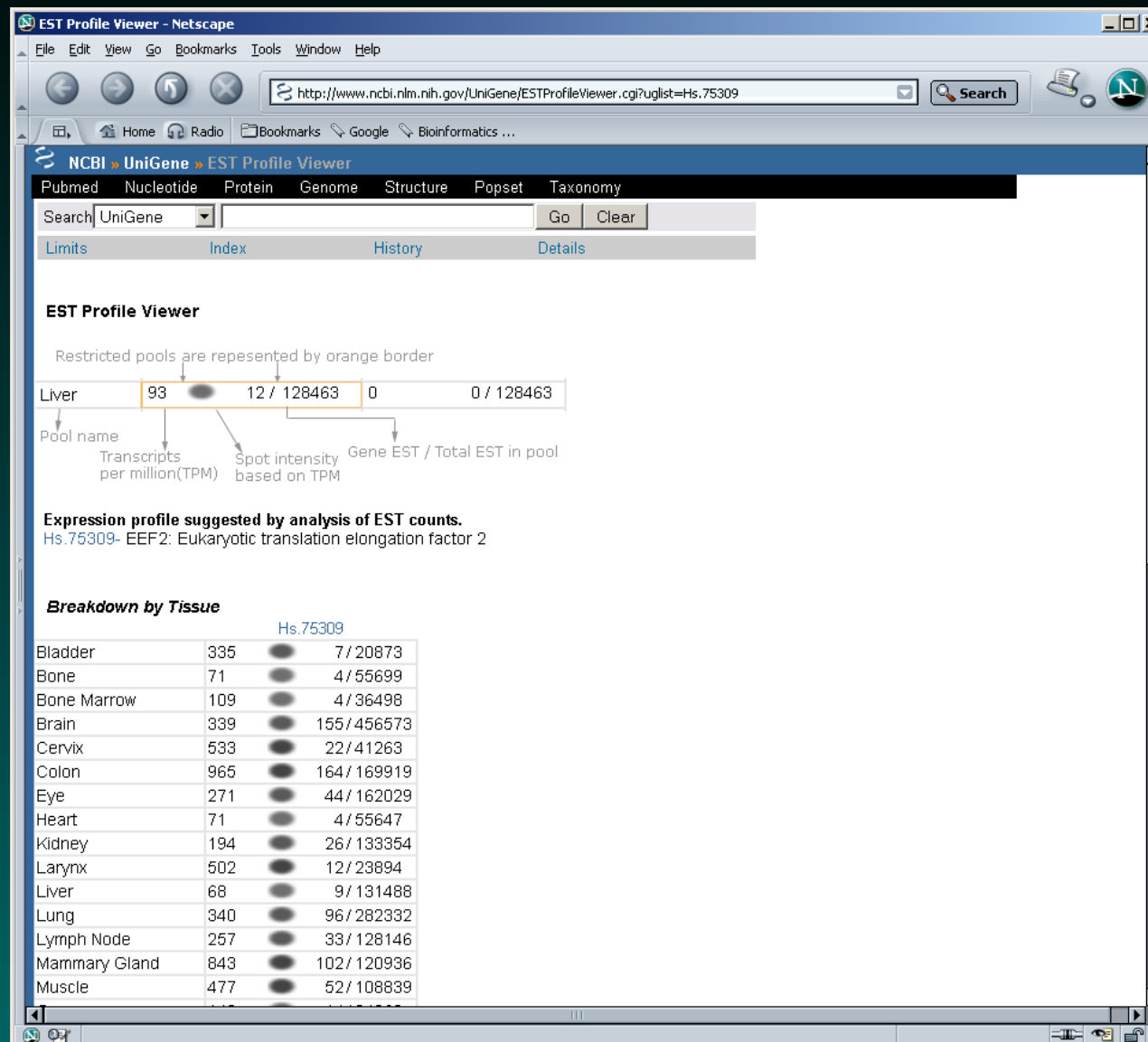
[Expression Profile:](#) View expression levels using UniGene's EST ProfileViewer

[GEO profiles:](#) Gene expression profiles in the NCBI Gene Expression Omnibus database

[NLAll tags:](#) NLAll anchored SAGE tags mapped to Hs.75309 in the NCBI SageMap database

[SAU3A tags:](#) SAU3A anchored SAGE tags mapped to Hs.75309 in the NCBI SageMap database

Expression Profile



UniGene Mapping Information

MAPPING POSITION

Genomic location specified by transcript mapping, radiation hybrid mapping, genetic mapping or cytogenetic mapping.

Genome View: Multiple Mappings [MapView](#)

CytoGenetic map: 19pter-q12

UniSTS entry: Chr 19 [RH80123](#)

UniSTS entry: Chr 20 [EST11A2](#)

UniSTS entry: Chr 17 [RH11500](#)

SEQUENCES

Sequences representing this gene; mRNAs, ESTs, and gene predictions supported by transcribed sequences.

mRNA sequences (12)

M19997.1	Human elongation factor 2 (EF-2) mRNA, 3' end.	P
NM_001961.2	Homo sapiens eukaryotic translation elongation factor 2 (EEF2), mRNA	P
BC068002.1	Homo sapiens cDNA clone MGC:70402 IMAGE:6052873, complete cds	PA
CR616093.1	full-length cDNA clone CS0DF036YA20 of Fetal brain of Homo sapiens (human)	P
CR608809.1	full-length cDNA clone CS0CAP008YN19 of Thymus of Homo sapiens (human)	P
CR602709.1	full-length cDNA clone CS0DG006YD23 of B cells (Ramos cell line) of Homo sapiens (human)	P
CR608809.1	full-length cDNA clone CS0D1002YK05 of Human	P

Chromosome Mapping

Entrez Genome view - Netscape

http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?query=Hs.75309&taxid=9606

NCBI

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM Help

Search for on chromosome(s) assembly Find

Show related entries

Homo sapiens genome view BLAST search the human genome

build 35 version 1 statistics

Hits: 1 2 3 4 5 6 7 8 9 10 11 12 13

Hits: 14 15 16 17 18 19 20 21 22 X Y MT

Search results for query "Hs.75309": 8 hits

Chr	Match	Map element	Type	Maps
5	171076	171076	Hs_EST_Cl	HsUniG
11	13500	13500	Hs_EST_Cl	HsUniG
19	all matches			
	82751	82751	Hs_EST_Cl	HsUniG
	82749	82749	Hs_EST_Cl	HsUniG
	82747	82747	Hs_EST_Cl	HsUniG
	82746	82746	Hs_EST_Cl	HsUniG
	82742	82742	Hs_EST_Cl	HsUniG
	78954	78954	Hs_EST_Cl	HsUniG

UniGene Links

UniGene - Netscape

http://www.ncbi.nlm.nih.gov/UniGene/cluster.cgi?ORG=Hs&CID=75309

NCBI **UniGene** ORGANISED VIEW OF THE TRANSCRIPTOME

PubMed Nucleotide Protein Genome Structure Popset Taxonomy

Search UniGene

Limits Preview/Index History Clipboard Details

NCBI

UniGene

Homepage

Query Tips

FAQs

DDD

Download

UniGene

Related Resources

LocusLink

HomoloGene

dbEST

Trace Archive

BLAST

CGAP

MGC cDNA clones

UniGene Cluster Hs.75309 *Homo sapiens* Eukaryotic translation elongation factor 2 (EEF2)

SELECTED PROTEIN SIMILARITIES

Comparison of sequences in UniGene with proteins supported by a complete genome. alignments can suggest function of a gene.

<i>A. thaliana</i>	ref:NP_172112.1	39.59 % / 846
	- unknown protein [Arabidopsis thaliana]	(see ProtEST)
<i>C. elegans</i>	ref:NP_492457.1	78.74 % / 858 aa
	- Elongation factor Tu family	(see ProtEST)
<i>D. melanogaster</i>	pir:S05988	78.42 % / 858 aa
	S05988 translation elongation factor	(see ProtEST)

Links

- ▶ LocusLink
- ▶ Gene
- ▶ GEO
- ▶ HomoloGene
- ▶ Nucleotide
- ▶ Nucleotide_MGC
- ▶ OMIM
- ▶ Protein
- ▶ SNP
- ▶ UniSTS
- ▶ LinkOut

http://www.ncbi.nlm.nih.gov/UniGene/cluster.cgi?ORG=Hs&CID=75309#MenuLinks

LocusLink

The first link out from UniGene is to LocusLink. LocusLink provides a single query interface to curated sequence information and descriptive information about genetic loci. This includes

- official nomenclature (symbol, name)
- aliases
- sequence accession numbers
- phenotypes
- MIM numbers
- UniGene clusters
- homology
- map locations

LocusLink Home Page

The screenshot shows the LocusLink Home Page in a Netscape browser window. The browser title is "LocusLink Introduction - Netscape" and the address bar shows "http://www.ncbi.nlm.nih.gov/LocusLink/index.html". The page features the NCBI logo and the LocusLink logo, which is a stylized 'S' with arrows pointing to the right. Below the logos is a navigation menu with links to PubMed, Entrez, BLAST, OMIM, Map Viewer, Taxonomy, and Structure. A search bar is present with a dropdown menu set to "LocusLink", a "Display" dropdown set to "Brief", and an "Organism:" dropdown set to "All". A "Query:" input field is followed by "Go" and "Clear" buttons. A yellow banner message states: "LocusLink will be replaced by [Entrez Gene](#). Check [Gene FAQ](#) for current information." Below this is a navigation bar with letters A through Z. The main content area has a purple header for "Introduction" and a paragraph of text: "LocusLink provides a single query interface to curated sequence and descriptive information about genetic loci. It presents information on official nomenclature, aliases, sequence accessions, phenotypes, EC numbers, MIM numbers, UniGene clusters, homology, map locations, and related web sites." A second paragraph follows: "Sequence accessions include a subset of GenBank accessions for a locus, as well as a new type, the NCBI Reference Sequence (RefSeq). RefSeq records are built according to the process [detailed here](#). See the [About](#) and [FAQ](#) pages for more information." A left sidebar contains a list of links: LocusLink Home, Build Process, Collaborators, Download, FAQ, GeneRIF, Add/Correct, Data, Help, Statistics, Genomic Biology, and NCBI Genome Guides.

LocusLink Results

The screenshot shows the LocusLink web interface in a Netscape browser window. The browser title is "LocusLink Report - Netscape" and the address bar shows "http://www.ncbi.nlm.nih.gov/LocusLink/LocRpt.cgi?l=1938". The page features the NCBI logo and a navigation menu with options like PubMed, Entrez, BLAST, OMIM, Map Viewer, Taxonomy, and Structure. A search bar contains "LocusLink" and "Display Brief". The main content area displays "View Hs EEF2" and "One of 1 Loci". A yellow banner states "LocusLink will be replaced by Entrez Gene. Check Gene FAQ for current information." Below this, there is a link to "Click to Display mRNA-Genomic Alignments (spanning 9391 bps)". A table of database links is shown, including Gene, PUB, OMIM, ACEVIEW, UNIGENE, MAP, VAR, HOMOL, GDB, and UCSC. The gene symbol "el" is listed. The official gene symbol is "Homo sapiens Official Gene Symbol and Name (HGNC)" and the description is "EEF2: eukaryotic translation elongation factor 2". The LocusID is 1938. An "Overview" link is present with a question mark. The RefSeq Summary states: "This gene encodes a member of the GTP-binding translation elongation factor family. This protein is an".

LocusLink Results

LocusLink Report - Netscape
 http://www.ncbi.nlm.nih.gov/LocusLink/LocRpt.cgi?l=1938

LocusLink Home
EEF2
Index:
 Top of Page
 Nomenclature
 Overview
 Function
 Relationships
 Map
 RefSeq
 Related
 Seqs
 Links

LocusLink:
 Collaborators
 Download
 FAQ
 Help
 Statistics

RefSeq:

Overview ?

RefSeq Summary: This gene encodes a member of the GTP-binding translation elongation factor family. This protein is an essential factor for protein synthesis. It promotes the GTP-dependent translocation of the nascent protein chain from the A-site to the P-site of the ribosome. This protein is completely inactivated by EF-2 kinase phosphorylation.

Locus Type: gene with protein product, function known or inferred

Product: eukaryotic translation elongation factor 2

Alternate Symbols: EF2, EEF-2

Alias: polypeptidyl-tRNA translocase

Function [Submit GeneRIF](#) ?

Please note: As a consequence of our transition to [Entrez Gene](#), GeneRIFs will only be displayed in the Entrez Gene record for this locus. [Gene](#)

Gene Ontology™:

LocusLink Results

LocusLink Report - Netscape
 http://www.ncbi.nlm.nih.gov/LocusLink/LocRpt.cgi?l=1938

angriment
 with mRNA

mRNA: [NM 001961](#)
Protein: [NP 001952](#) **BL**

Related Sequences ?

Nucleotide	Type	Protein	
BC006547	m	AAH06547	BL
BC024689	m	AAH24689	BL
M19997	m	AAA50388	BL
X51466	m	CAA35829	BL
Z11692	m	CAA77750	BL
None	p	P13639	BL

Additional Links ?

- **OMIM:** [130610](#)
- **UniGene:** [Hs.75309](#)

[To Top](#)

Questions or Comments?
 Write to the [NCBI Service Desk](#)

[Disclaimer](#) [Privacy statement](#)

LocusLink:
 Home
 EEF2
 Index:
 Top of Page
 Nomenclature
 Overview
 Function
 Relationships
 Map
 RefSeq
 Related
 Seqs
 Links

RefSeq:
 Collaborators
 Download
 FAQ
 Help
 Statistics

Online Mendelian Inheritance in Man

Both UniGene and LocusLink create automatic links to OMIM.

OMIM is a curated database of human genes and genetic disorders. It typically includes information about which diseases appear to be linked to specific genes, along with primary references that explain how the gene was sequenced and mapped to specific chromosomal regions.

OMIM Main Page

OMIM - Online Mendelian Inheritance in Man - Netscape

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM

NCBI OMIM Online Mendelian Inheritance in Man Johns Hopkins University

PubMed Nucleotide Protein Genome Structure PMC Taxonomy OMIM

Search OMIM for Go Clear

Limits Preview/Index History Clipboard Details

Entrez

- OMIM
- Search OMIM
- Search Gene
- Map
- Search Morbid
- Map

Help

- OMIM Help
- How to Link

FAQ

- Numbering System
- Symbols
- How to Print
- Citing OMIM
- Download

OMIM Facts

- Statistics
- Update Log
- Restrictions on Use

- Enter one or more search terms.
- Use **Limits** to restrict your search by search field, chromosome, and other criteria.
- Use **Index** to browse terms found in OMIM records.
- Use **History** to retrieve records from previous searches, or to combine searches.

OMIM™ - Online Mendelian Inheritance in Man™

Welcome to OMIM, Online Mendelian Inheritance in Man. This database is a catalog of human genes and genetic disorders authored and edited by Dr. Victor A. McKusick and his colleagues at Johns Hopkins and elsewhere, and developed for the World Wide Web by NCBI, the National Center for Biotechnology Information. The database contains textual information and references. It also contains copious links to MEDLINE and sequence records in the Entrez system, and links to additional related resources at NCBI and elsewhere.

You can do a search by entering one or more terms in the text box above.

OMIM Results

OMIM - EUKARYOTIC TRANSLATION ELONGATION FACTOR 2; EEF2 - Netscape

http://www.ncbi.nlm.nih.gov/entrez/dispmim.cgi?id=130610

NCBI

OMIM
Online Mendelian Inheritance in Man

PubMed Nucleotide Protein Genome Structure PMC Taxonomy OMIM

Search OMIM for Go

Clear

Limits Preview/Index History Clipboard Details

Display Detailed Show: 20 Send to Text

*130610 [Links](#)

**EUKARYOTIC TRANSLATION ELONGATION FACTOR 2;
EEF2**

Alternative titles; symbols

ELONGATION FACTOR 2; EF2
POLYPEPTIDYL-tRNA TRANSLOCASE

Gene map locus [19pter-q12](#)

TEXT

Diphtheria toxin and Pseudomonas exotoxin A (PA toxin) inhibit protein synthesis by catalyzing covalent binding of the ADP-ribose moiety of NAD to elongation factor-2 (EF2). EF2 is required for the translocation step in protein

Entrez Gene
Nomenclature
R RefSeq
G GenBank
P Protein
U UniGene

LinkOut

OMIM References

OMIM - EUKARYOTIC TRANSLATION ELONGATION FACTOR 2; EF2 - Netscape

http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=130610

By analysis of human-mouse hybrid cells, [Kaneda et al. \(1987\)](#) narrowed the assignment of EF2 to 19pter-q12.

REFERENCES

1. Kaneda, Y.; Hayes, H.; Uchida, T.; Yoshida, M. C.; Okada, Y. :
Regional assignment of five genes on human chromosome 19.
Chromosoma 95: 8-12, 1987.
PubMed ID : [3034518](#)
2. Kaneda, Y.; Yoshida, M. C.; Kohno, K.; Uchida, T.; Okada, Y. :
Chromosomal assignment of the gene for human elongation factor-2.
Proc. Nat. Acad. Sci. 81: 3158-3162, 1984.
PubMed ID : [6427766](#)
3. Rapp, G.; Klaudiny, J.; Hagendorff, G.; Luck, M. R.; Scheit, K. H. :
Complete sequence of the coding region of human elongation factor 2 (EF-2) by enzymatic amplification of cDNA from human ovarian granulosa cells. *Biol. Chem. Hoppe-Seyler* 370: 1071-1075, 1989.
PubMed ID : [2610926](#)

CONTRIBUTORS

Rebekah S. Rasooly - updated : 2/10/1999

CREATION DATE

NCBI

MIM *130610

Text

References

Contributors

Creation Date

Edit History

• Gene map

Entrez Gene

Nomenclature

RefSeq

GenBank

Protein

UniGene

LinkOut

Batch Resources

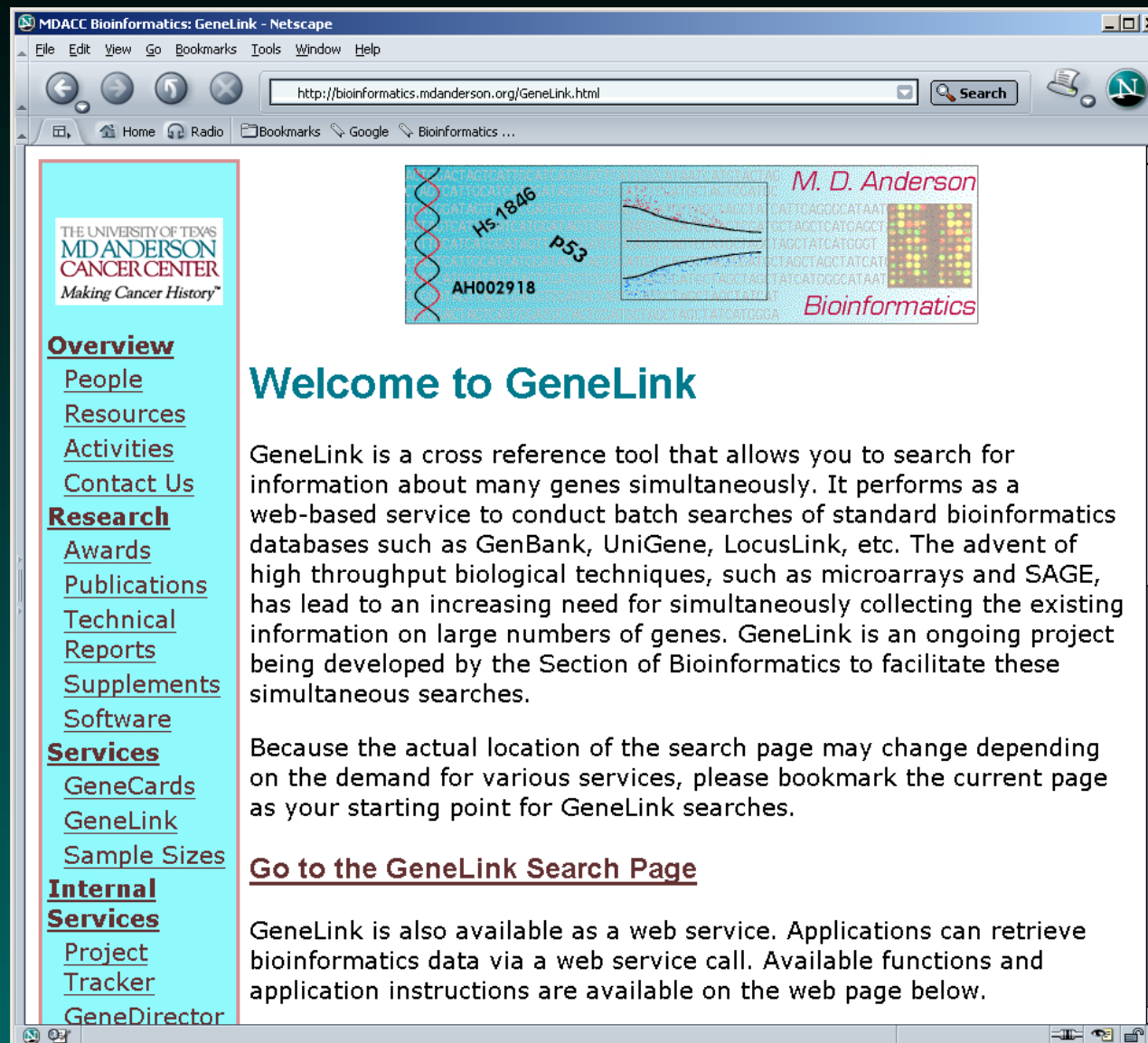
You may have noticed that only GenBank (Entrez-Nucleotide) provided a “batch” option that allowed us to search for an entire list of genes at once. Every other example we have presented works with one gene at a time. That’s probably not a good way to deal with 1000 genes.

M.D. Anderson maintains a service called GeneLink to deal with this problem. You can try it at [http:](http://bioinformatics.mdanderson.org/GeneLink.html)

[//bioinformatics.mdanderson.org/GeneLink.html](http://bioinformatics.mdanderson.org/GeneLink.html)

Stanford has a similar tool (that right now is faster) called Source. Check out <http://source.stanford.edu>.

GeneLink Description



MDACC Bioinformatics: GeneLink - Netscape

File Edit View Go Bookmarks Tools Window Help

http://bioinformatics.mdanderson.org/GeneLink.html

Home Radio Bookmarks Google Bioinformatics ...

THE UNIVERSITY OF TEXAS
MD ANDERSON
CANCER CENTER
Making Cancer History™

Overview
[People](#)
[Resources](#)
[Activities](#)
[Contact Us](#)

Research
[Awards](#)
[Publications](#)
[Technical Reports](#)
[Supplements](#)
[Software](#)

Services
[GeneCards](#)
[GeneLink](#)
[Sample Sizes](#)

Internal Services
[Project Tracker](#)
[GeneDirector](#)

Welcome to GeneLink

GeneLink is a cross reference tool that allows you to search for information about many genes simultaneously. It performs as a web-based service to conduct batch searches of standard bioinformatics databases such as GenBank, UniGene, LocusLink, etc. The advent of high throughput biological techniques, such as microarrays and SAGE, has lead to an increasing need for simultaneously collecting the existing information on large numbers of genes. GeneLink is an ongoing project being developed by the Section of Bioinformatics to facilitate these simultaneous searches.

Because the actual location of the search page may change depending on the demand for various services, please bookmark the current page as your starting point for GeneLink searches.

[Go to the GeneLink Search Page](#)

GeneLink is also available as a web service. Applications can retrieve bioinformatics data via a web service call. Available functions and application instructions are available on the web page below.

GeneLink Search Page

GeneLink links gene information across databases in spreadsheet format that can be pasted into an existing spreadsheet. [More](#) information is available below.

GeneLink is currently using Homo Sapiens Unigene **Build 175**.

Select the desired output columns. Check the box next to "Include input" if you want your search terms included as the first column of the result table.

Search for that the items listed either in the file or on separate lines of the following list:

Rows to display per page(default is 25)
Note: Enter 'ALL' to display all rows on one page

Output file delimiter (default is tab '\t')

Include links in output file

Output as an html table:

Output Fields

Basic output

Include Search Term

Accession Number

Unigene Cluster

Gene Symbol

Chromosome

Cytoband

Gene Name

LocusLink

Advanced output

Gene Ontology

Biocarta

GeneLink Search Page

Gene Link - Netscape

File Edit View Go Bookmarks Tools Window Help

http://bioinformatics.mdanderson.org/GeneLink/ Search

Home Radio Bookmarks Google Bioinformatics ...

Search for that the items listed either in the file or on separate lines of the following list:

Rows to display per page(default is 25)
Note: Enter 'ALL' to display all rows on one page

Output file delimiter (default is tab '\t')

Include links in output file

Output as an html table:

Output to Email Address:

(A tab-delimited text file will be emailed to you when it is ready. This option is good for large queries such as those that contain more than 1000 search items.)

Output Fields

Basic output

Include Search Term

Accession Number

Unigene Cluster

Gene Symbol

Chromosome

Cytoband

Gene Name

LocusLink

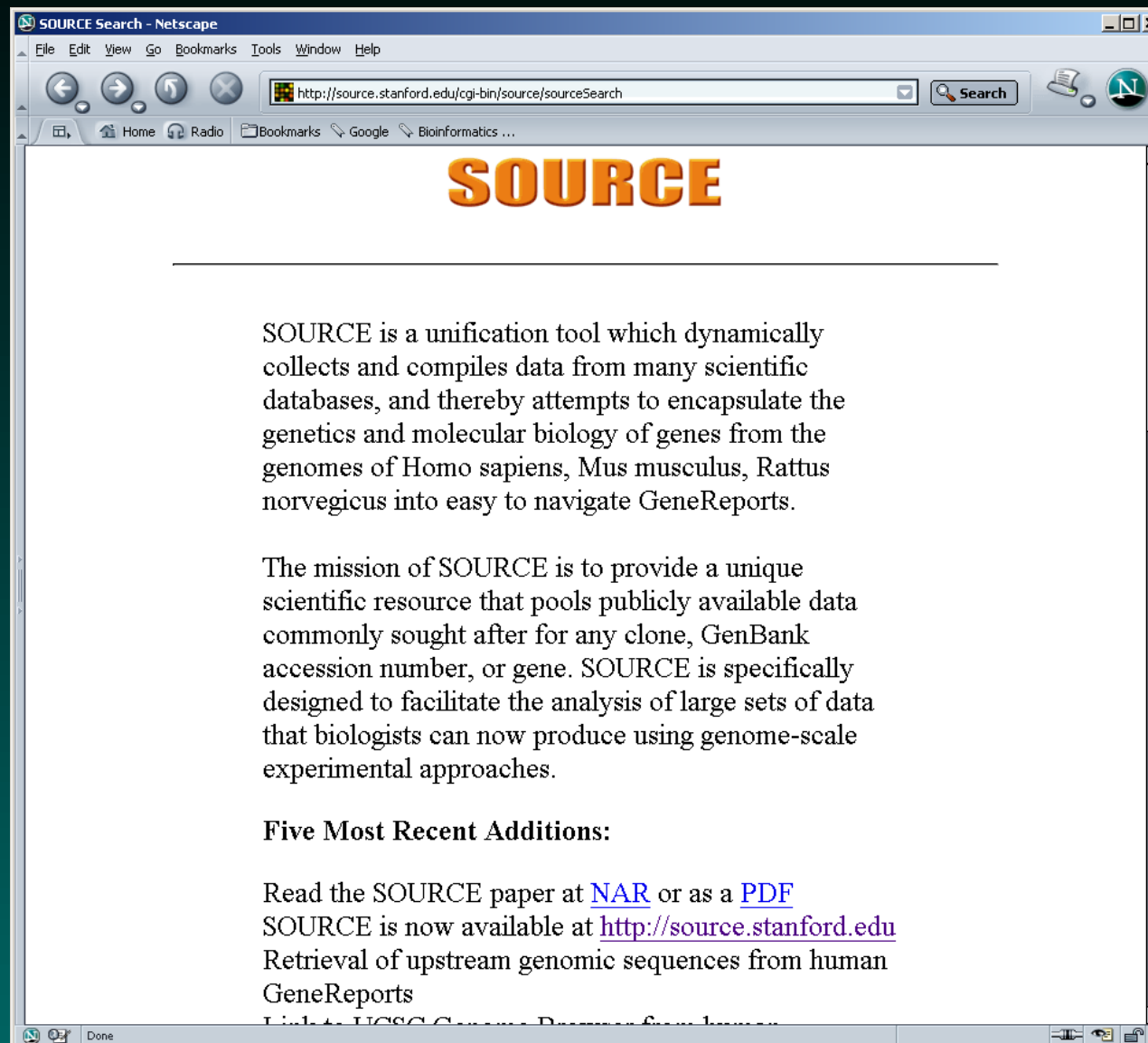
Advanced output

Gene Ontology

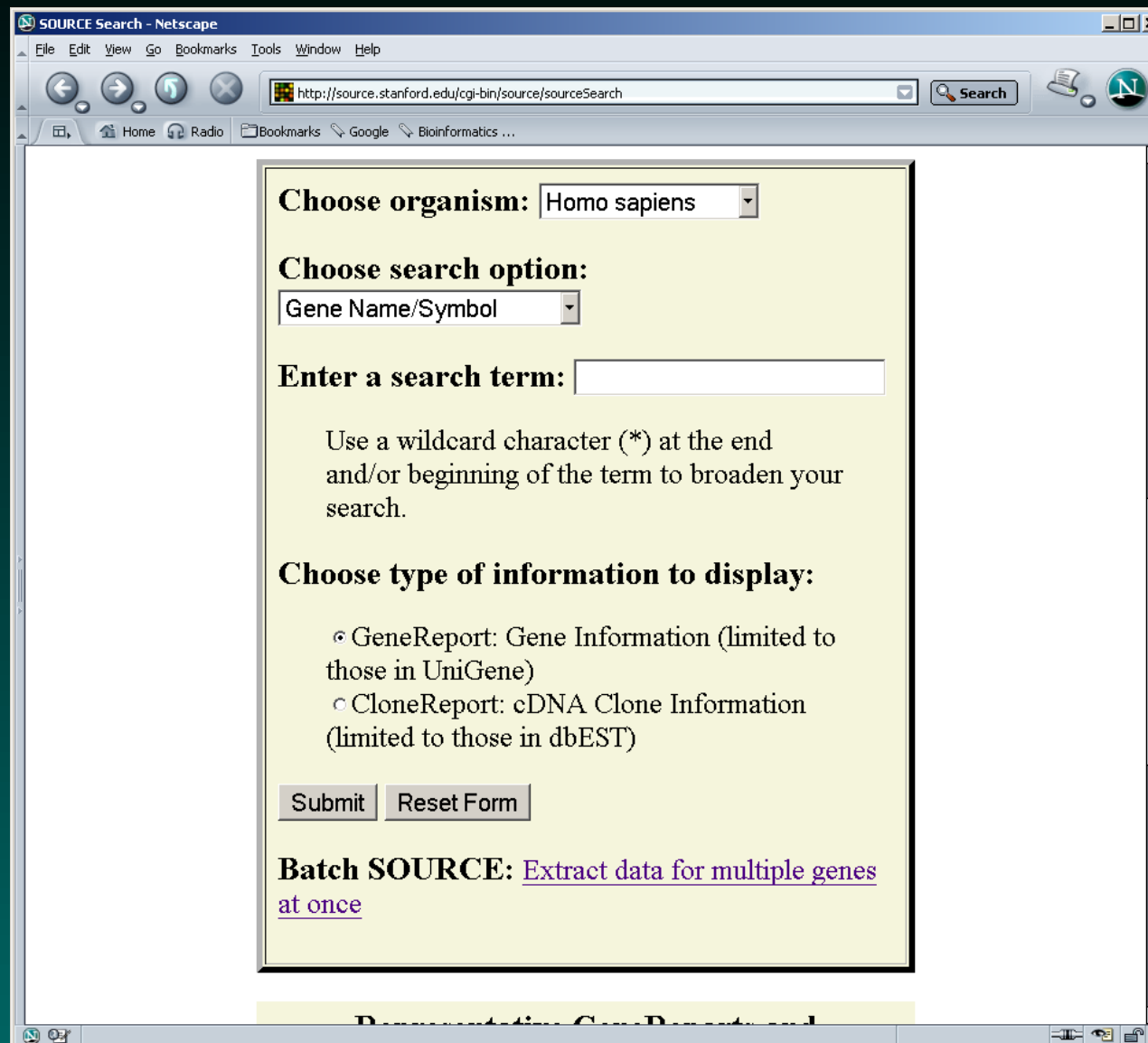
Biocarta

Kegg

Source Description



Source Search Page



The screenshot shows a Netscape browser window titled "SOURCE Search - Netscape". The address bar contains the URL "http://source.stanford.edu/cgi-bin/source/sourceSearch". The main content area features a search form with the following elements:

- Choose organism:** A dropdown menu with "Homo sapiens" selected.
- Choose search option:** A dropdown menu with "Gene Name/Symbol" selected.
- Enter a search term:** An empty text input field.
- Instructions:** "Use a wildcard character (*) at the end and/or beginning of the term to broaden your search."
- Choose type of information to display:** Two radio buttons: "GeneReport: Gene Information (limited to those in UniGene)" (selected) and "CloneReport: cDNA Clone Information (limited to those in dbEST)".
- Buttons:** "Submit" and "Reset Form".
- Batch SOURCE:** A link: [Extract data for multiple genes at once](#).

At the bottom of the form area, the text "Representative Gene Reports and" is partially visible.

Source Batch Search Page

The screenshot shows a Netscape browser window titled "SOURCE Batch Search - Netscape". The address bar contains the URL "http://genome-www5.stanford.edu/cgi-bin/source/sourceBatchSearch". The page content is as follows:

Input File:
C:\R\projects\tailrank\extreme.txt

Or enter a list of identifiers:

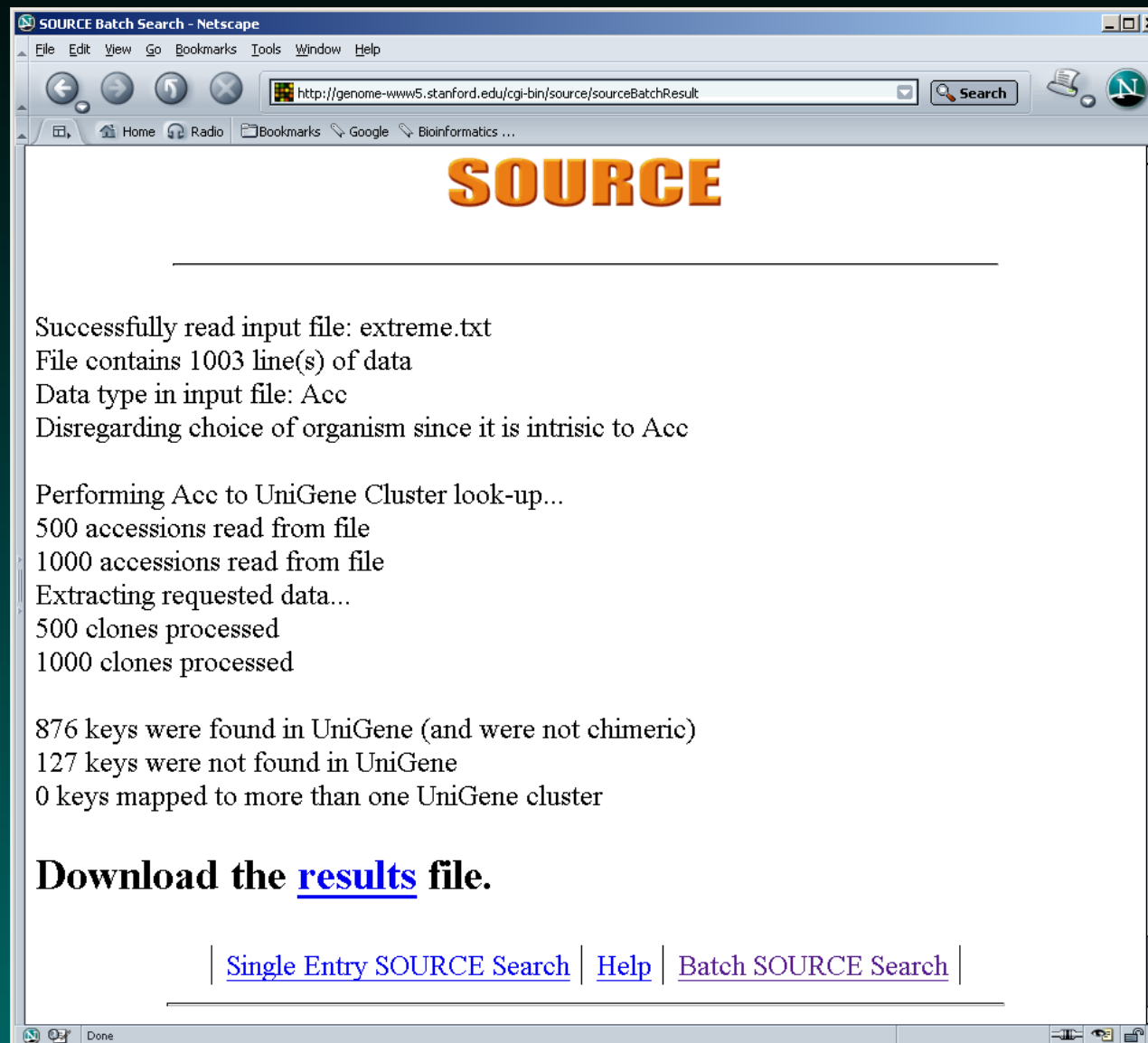
Select the type of input identifier: GenBank Accession

Choose organism: Homo sapiens

2) Choose field(s) for extraction:

<input checked="" type="checkbox"/> UniGene Cluster ID	<input type="checkbox"/> Representative Protein Acc.	<input type="checkbox"/> Other Ontology Annotations (full)
<input checked="" type="checkbox"/> UniGene Name	<input checked="" type="checkbox"/> Chromosome Location	<input type="checkbox"/> Other Ontology Annotations (short)
<input checked="" type="checkbox"/> UniGene Symbol	<input checked="" type="checkbox"/> Cytoband	<input type="checkbox"/> Enzymatic Function

Source Results



The screenshot shows a Netscape browser window titled "SOURCE Batch Search - Netscape". The address bar contains the URL "http://genome-www5.stanford.edu/cgi-bin/source/sourceBatchResult". The main content area displays the following text:

SOURCE

Successfully read input file: extreme.txt
File contains 1003 line(s) of data
Data type in input file: Acc
Disregarding choice of organism since it is intrinsic to Acc

Performing Acc to UniGene Cluster look-up...
500 accessions read from file
1000 accessions read from file
Extracting requested data...
500 clones processed
1000 clones processed

876 keys were found in UniGene (and were not chimeric)
127 keys were not found in UniGene
0 keys mapped to more than one UniGene cluster

Download the [results](#) file.

[Single Entry SOURCE Search](#) | [Help](#) | [Batch SOURCE Search](#)

Source Details

Acc	UGCluster	Name	Symbol	Chromosome	Cytoband
R20379	Hs.501443	Transcribed sequence with strong similarity to protein sp:P1			
H08440	Hs.440382	ret finger protein	RFP	6	6p22
N74602	Hs.3727	serine/threonine kinase receptor associated protein	STRAP	12	
R74357	Hs.367688	Clone IMAGE:4794726, mRNA		7	
AA406332	Hs.528298	Sec23 homolog A (<i>S. cerevisiae</i>)	SEC23A	14	14q11
AA406332	Hs.528298	Sec23 homolog A (<i>S. cerevisiae</i>)	SEC23A	14	14q11
W58562	Hs.102471	phosphatase and actin regulator 2	PHACTR2	6	6q24
T61792	Hs.8364	pyruvate dehydrogenase kinase, isoenzyme 4	PDK4	7	7q21
AA130866	Hs.133321	trimethyllysine hydroxylase, epsilon	TMLHE	X	
W79123	Hs.58561	G protein-coupled receptor 87	GPR87	3	3q24
AA047778	Hs.121576	myosin IB	MYO1B	2	2q12-q34
AA133243	Hs.35380	bobby sox homolog (<i>Drosophila</i>)	BBX	3	3q13
AA400247	Hs.42173	chromosome 6 open reading frame 107	C6orf107		
AA126673	Hs.212787	KIAA0303 protein	KIAA0303	5	5q13
AA478854	Hs.190477	zinc finger CCCH type domain containing 6			ZC3H1
N26658	Hs.342874	transforming growth factor, beta receptor III (betaglycan, 30			
AA701491	Hs.257111	hypothetical protein MGC3265	MGC3265	5	5q33
AA988746	Hs.459470	WD repeat and SOCS box-containing 2	WSB2	12	
AA779843	Hs.404741	nuclear factor (erythroid-derived 2)-like 3	NFE2L3		
AA918601	Hs.407934	neuron navigator 2	NAV2	11	11p15.1
AA931102	Hs.419149	junctional adhesion molecule 3	JAM3	11	11q24
AA985587	Hs.452398	CDNA FLJ30740 fis, clone FEBRA2000319			2
R34875	Data not found				
AI346326	Hs.157211	C1q and tumor necrosis factor related protein 5	C1QTNF5	5	5q11
AI657057	Hs.159642	glucosaminyl (N-acetyl) transferase 1, core 2 (beta-1			
AI263076	Hs.398124	dynein, axonemal, heavy polypeptide 5	DNAH5	5	
AA916784	Data not found				
W58562	Hs.102471	phosphatase and actin regulator 2	PHACTR2	6	6q24
AA495790	Hs.406064	ras homolog gene family, member B	RHOB	2	
R76613	Hs.102541	netrin 4	NTN4	12	12q22-q23
W63789	Hs.90797	hypothetical protein BC016005	LOC129642	2	2p25
AA453616	Hs.381089	AKAP-associated sperm protein	ASP	5	5p15
AA043501	Hs.134859	v-maf musculoaponeurotic fibrosarcoma oncogene homol			

Reprise

- Primary identifiers describe the biological material used as probes
- GenBank accession numbers are a common denominator to start updating gene annotations
- UniGene clusters sequences that represent the same gene
- LocusLink contains curated sequence information about names, mappings, etc
- OMIM contains curated connections between genes and diseases, with references into the literature
- There are batch processing tools to go from a list of GenBank accession numbers through UniGene to LocusLink or OMIM.

Main Question: How do we learn about gene functions, networks, and pathways?

Return to LocusLink

completely inactivated by EF-2 kinase phosphorylation.

Locus Type: gene with protein product, function known or inferred

Product: eukaryotic translation elongation factor 2

Alternate Symbols: EF2, EEF-2

Alias: polypeptidyl-tRNA translocase

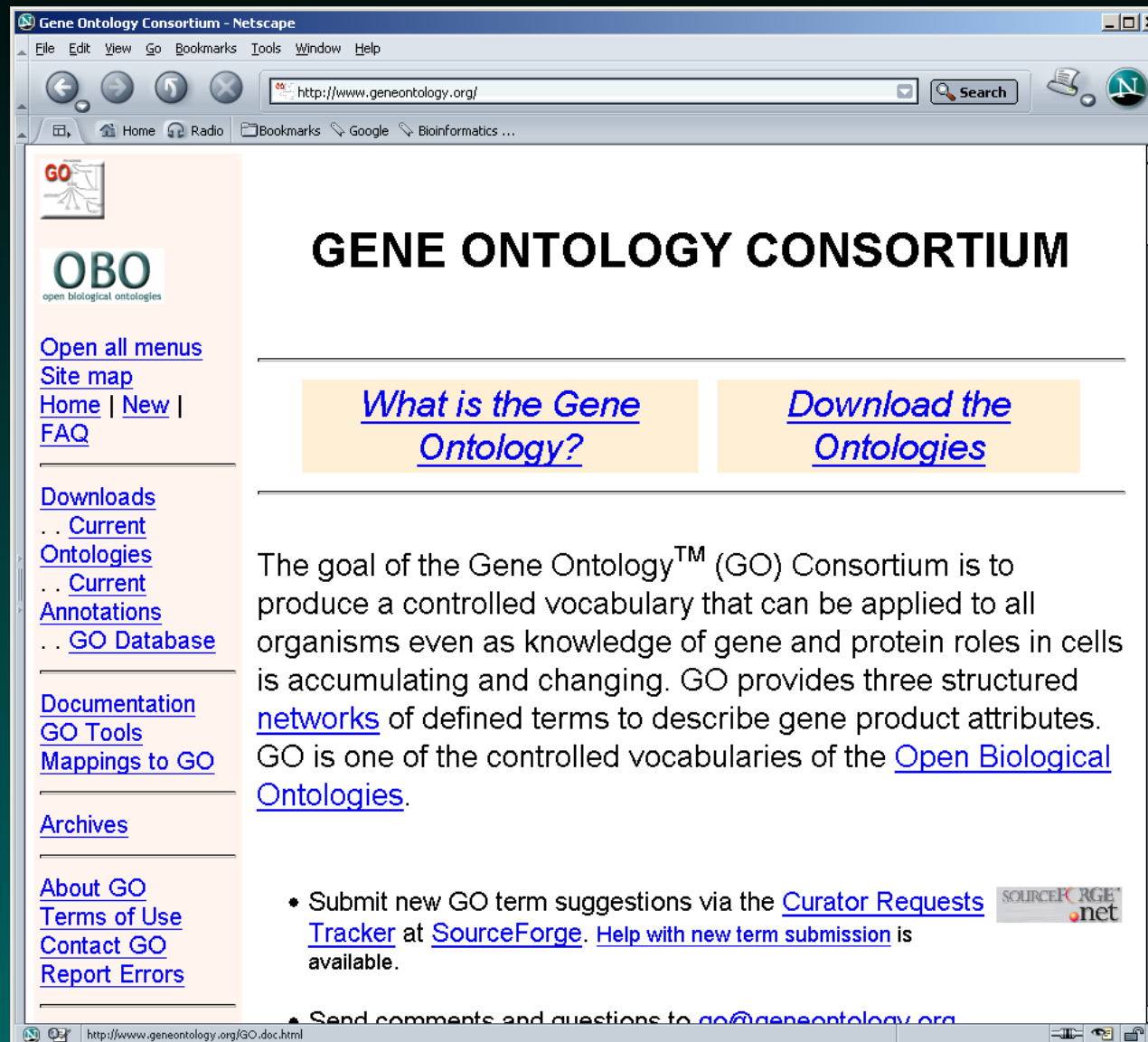
Function [Submit GeneRIF](#) ?

Please note: As a consequence of our transition to [Entrez Gene](#), GeneRIFs will only be displayed in the Entrez Gene record for this locus. [Gene](#)

Gene Ontology™:

Term	Evidence	Source	Pub
• GTP binding	IEA		GOA
• protein biosynthesis	IEA		GOA
• translation elongation factor activity	NR		GOA
• translational elongation	IEA		GOA

http://www.geneontology.org



The screenshot shows a Netscape browser window displaying the Gene Ontology Consortium website. The browser's address bar shows the URL <http://www.geneontology.org/>. The website features a navigation menu on the left with links for 'Open all menus', 'Site map', 'Home', 'New', 'FAQ', 'Downloads', 'Current Ontologies', 'Current Annotations', 'GO Database', 'Documentation', 'GO Tools', 'Mappings to GO', 'Archives', 'About GO', 'Terms of Use', 'Contact GO', and 'Report Errors'. The main content area is titled 'GENE ONTOLOGY CONSORTIUM' and includes two prominent yellow buttons: 'What is the Gene Ontology?' and 'Download the Ontologies'. Below these buttons, a paragraph explains the consortium's goal: 'The goal of the Gene Ontology™ (GO) Consortium is to produce a controlled vocabulary that can be applied to all organisms even as knowledge of gene and protein roles in cells is accumulating and changing. GO provides three structured networks of defined terms to describe gene product attributes. GO is one of the controlled vocabularies of the Open Biological Ontologies.' A bulleted list at the bottom provides instructions on how to submit new GO term suggestions via the Curator Requests Tracker at SourceForge and how to contact the consortium via email at go@geneontology.org. A SourceForge.net logo is visible in the bottom right corner of the page content.

GeneOntology

GeneOntology uses controlled vocabularies to create a directed acyclic graph (a generalized tree) that describes the kinds of functions or properties that a gene might have.

The properties are divided into three categories:

1. Biological process (what)
2. Molecular function (how)
3. Cellular component (where)