EAHIL Conference, Köln, 21st September 2002

Wisdom world-wide?

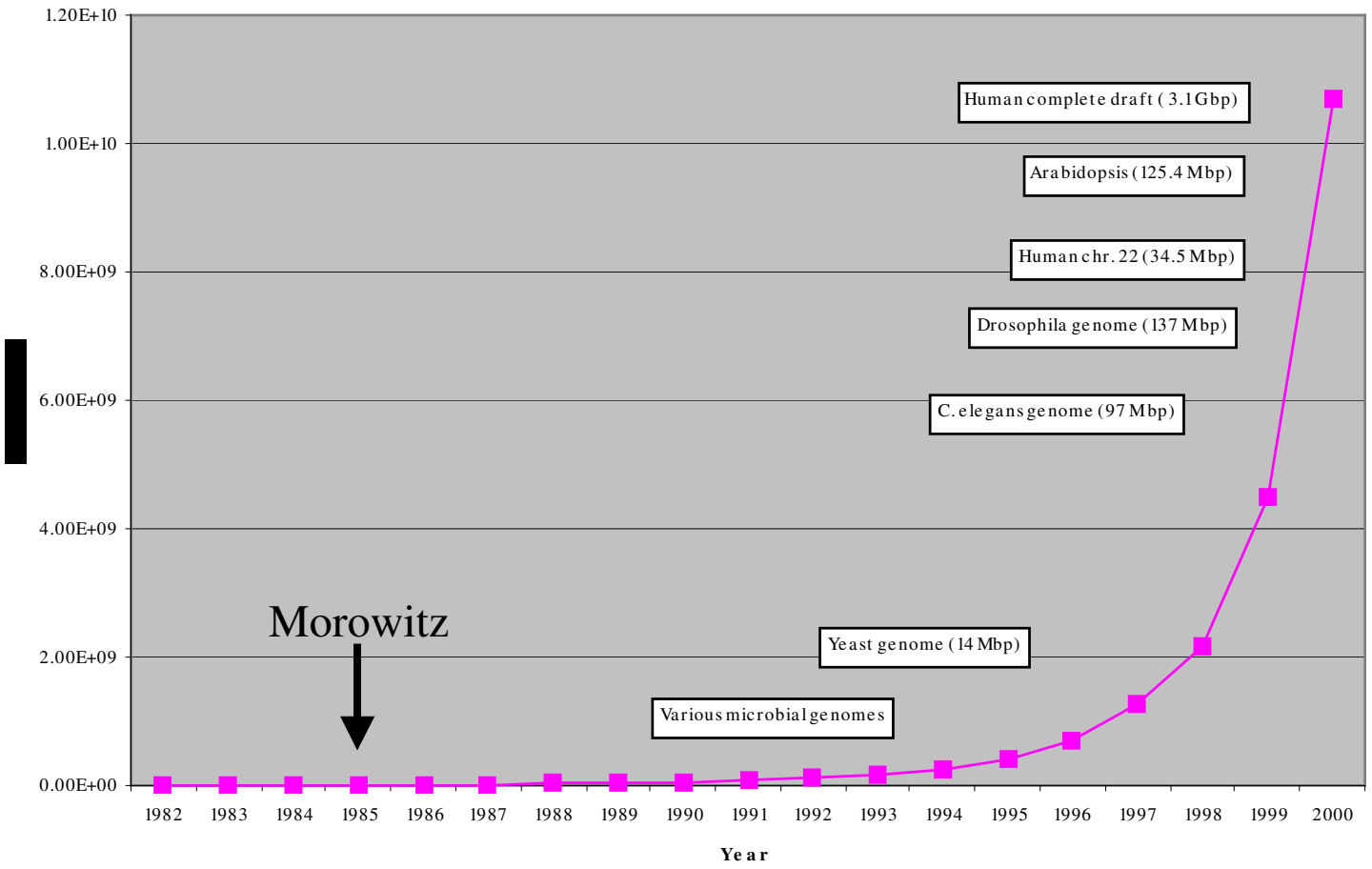
Portals to bio-medical information viewed from a user perspective

Les Grivell
European Molecular Biology Organization

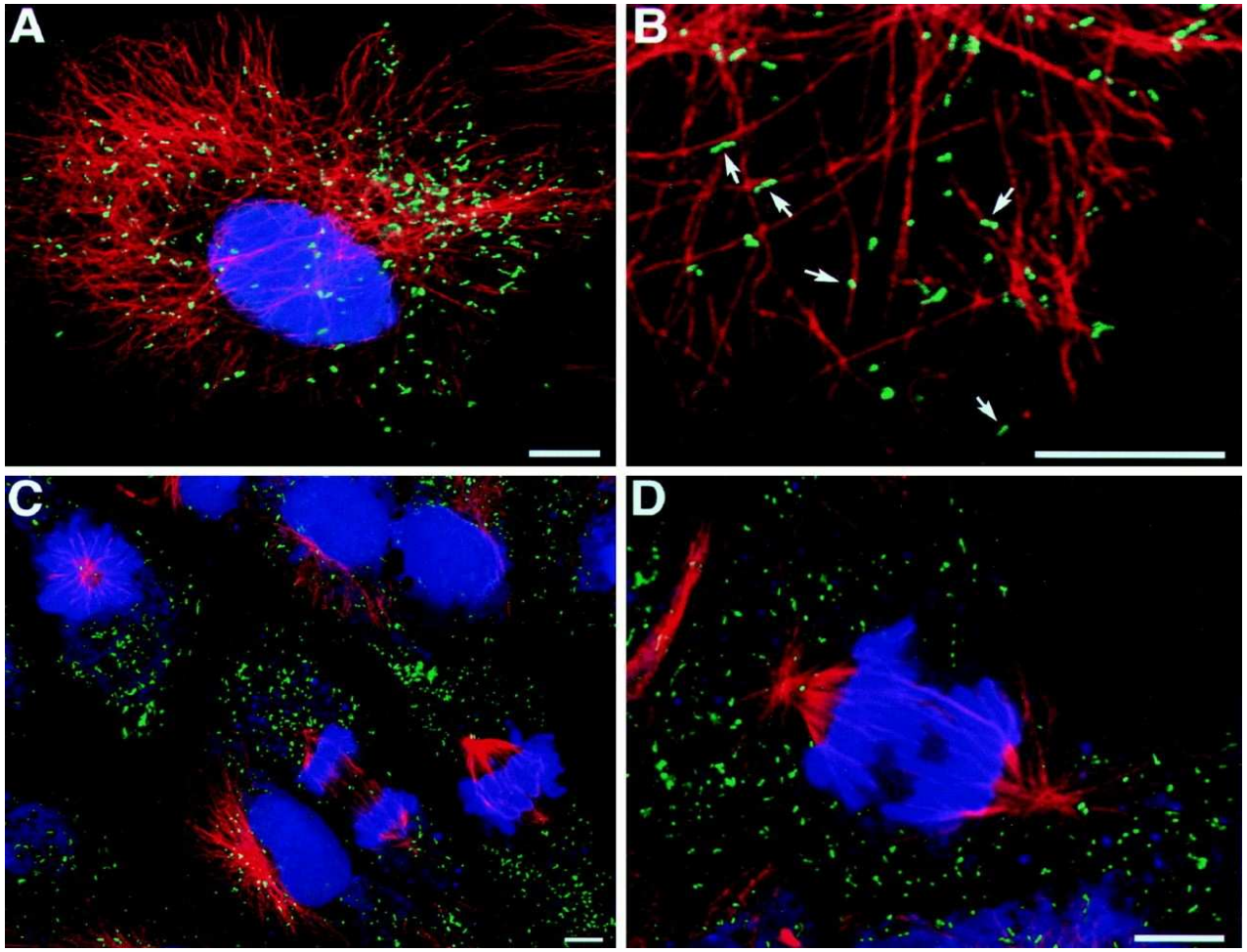
"Biological research has reached a point where new generalizations and higher order biological laws are being approached, but may be obscured by the simple mass of data"

Harold Morowitz, 1985
Report to the U.S. National Academy of Sciences

One part of the information explosion ….



New technologies also generate digital information



The problem ….

…and the user perspective

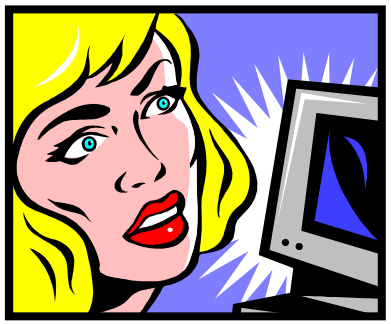
Biological information: current reality …..

- Knowledge scattered across the literature in many thousands of non-computer readable journal articles
- Hundreds of different databases, many in flat-file format
- Non-uniform or lack of external identifiers
- Lack of interoperability at the level of syntax and semantics

What do most users want?

- Intuitive interfaces
- To navigate smoothly from one resource to another
- Immediate availability of authoritative information; free of charge at the point-of-use
- Information that can be integrated / manipulated / visualized /output in another form

A sample scenario……



Friedreich ataxia is a disease characterized by neurodegeneration and cardiomyopathy. The gene responsible for the disease is frataxin. It encodes a mitochondrial protein required for iron homeostasis and is evolutionarily conserved from yeast to man.

- Critically assess the experimental evidence that frataxin functions in iron-sulphur cluster synthesis.
- Review the effectiveness of anti-oxidants as treatment for the disease.
- Assemble as far as possible all other known information on the frataxin gene and the action of its protein

Now choose a portal …

Part 1: Molecular Biology Search and Analysis

- 3D-ALI - Database relating Proteins Structures and Sequences (see Documentation) at EMBL-Heidelberg W3
- AA Analysis - Protein Identification in SwissProt and PIR using Amino Acid Composition at EMBL-Heidelberg W3
- AA CompIdent - Protein Identification in SwissProt using Amino Acid Composition at ExPASy, Switzerland W3
- AA CompSim - Compare the Amino Acid Composition of a SwissProt Entry with Other Entries at ExPASy, Switzerland W3
- AA Sequence Analysis - Multiple Analysis of a Native or Modified Amino Acid Sequence at Rockefeller U. W3
- AbCheck - Test an Antibody Sequence Against the Kabat Database at U. College-London, UK W3
- ALIGN - Optimal Global Alignment of Two Sequences with No Short-cuts at EERIE-Nimes, France and Align at MIPS, Gern
- AllAll - Relationships in a Set of Related Peptides at ETH-Zürich, Switzerland W3
- AllAllDB - Search the All-Against-All Database of SwissProt at ETH-Zürich, Switzerland W3
- AMAS - Analysis of Protein Multiple Sequence Alignments at U. Oxford, UK W3
- ASC - Analytic Surface Calculation of PDB Protein Structures at EMBL-Heidelberg W3
- ATLAS - Search DNA and Protein Sequence Databases at MIPS, Germany W3
- BCM Search Launcher at Baylor College of Medicine W3
- BEAUTY - BLAST Enhanced Alignment Utility (see Documentation) at Baylor College of Medicine W3
- BERLIN - RNA Databank of 5S rRNA and 5S rRNA Gene Sequences at CAOS/CAMM, The Netherlands W3
- BIOACCELERATOR at Weizmann Institute of Science, Israel W3
- bioCUSI - Multiple Access to Biological Servers at OCMS-Oxford W3
- Biological Units - Expanded PDB Entries with Full Biological Units at Brookhaven National Laboratory FTP
- Biologist's Search Palette - Collection of Search Engines for Biological Databases at Oxford U., UK W3
- BLAST - Basic Local Alignment Research Tool at NCBI, BLAST (hiv, vector, etc) at Stanford U., BLAST at GenomeNet, Ja

France BLAST (with Entrez and SRS Links) at Baylor College of Medicine and BLAST at EPFL, Switzerland W3

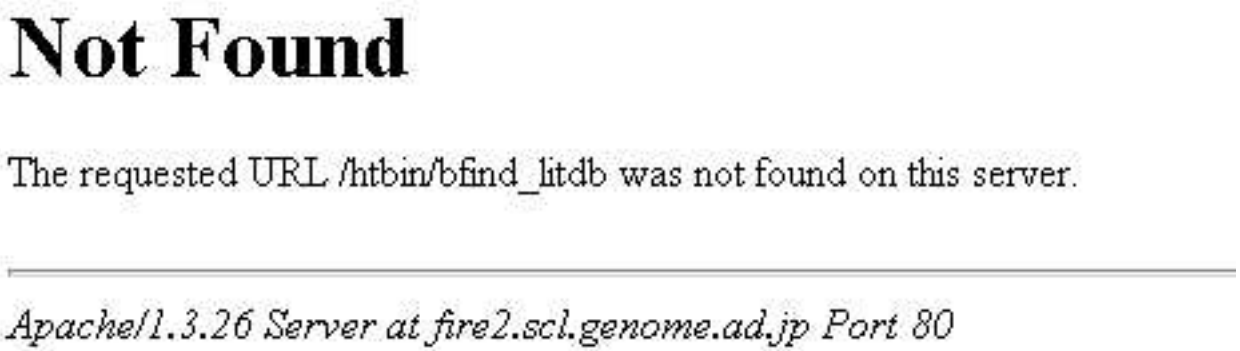
- BLASTPAT - BLAST-based Pattern Database Search (see Documentation) at Baylor College of Medicine W3
- BLITZ - Fast comparison of protein sequences against SwissProt (see Documentation) at EBI, UK W3
- BLOCKS - Database of Highly Conserved Regions in Proteins W3 and BLOCKS-WAIS Gopher at Fred Hutchinson Cancer
- BioMagResBank - A Database of NMR-Derived Protein Structures at BIMAS-NIH W3
- BioSCAN - Biological Sequence Comparative Analysis at U. North Caroline W3
- BMCD - Biological Macromolecule Crystallization Database at CARB W3

There's plenty of choice!

Part 2: Bibliographic, Text and WWW Searches

- **AA-Index** - Amino Acid Index Database at GenomeNet, Japan and **NAKAI** at CAOS/CAMM, The Netherlands W3
- **AGRICOLA** ✔ - Bibliographic References (Plant Genome Subset) at USDA-NAL Gopher
- **Annual Reviews** ✔ - Bibliographic Searches at LMMB-National Cancer Institute W3
- **ArchiePlexForm** - Search Files Available by Anonymous FTP at LeRC-NASA W3
- **ATCG** ✔ - Anderson's Timesaving Comparative Guides for Restriction Enzymes, Modifying Enzymes, Clone Libraries and Fi
- **Beckman-Bioresearch** ✔ - Information on Protein and DNA Applications, Meetings and Product Information at Beckman I
- **BIBLIO** - Molbio Journal Searches at Genethon, France Gopher and **Biology Journal Contents** at CERN, Switzerland W3
- **BIBSYS** - Interuniversitary Bibliographic System in Norway W3
- **BIOBIB** - Biocomputing Bibliography at ICGEB-Trieste, Italy Gopher and **CompAlgo** at CERN, Switzerland W3
- **BioCatalog** - Catalog of Molecular Biology Programs at EBI, **Bio-catalogue** at Genethon, France and **BIOCAT** at CAOS/
- **Bio-Email** - List of Biomedical E-Mail Servers and **Bio-Email-Search** at Washington U. W3
- **Bioline** - Bio-Publications at Base de Dados Tropical, Brazil W3
- **BioMOO** - The Biologists' Virtual Meeting Place at Weizmann Institute of Science, Israel W3
- **Bionet Journal** Articles and **Bionet News** Gopher
- **BIOSCI** Archives at EMBnet-UK Gopher and **BIOSCI** at CERN, Switzerland W3
- **BioSpace** ✔ - Find People, Companies and Resources in Biotech Industry at Synergistic Designs W3
- **Biotech Company WWW Registry** ✔ - Martin Leach's List of Commercial Net sites/Email addresses for Molecular Biolog
- **Biotechnology Abstracts DataBase** at Stack-Serpukhov, Russia Gopher
- **Biotechnology Bibliographies** at BIC-U. Maryland/College Park W3

Although some may be difficult to find …



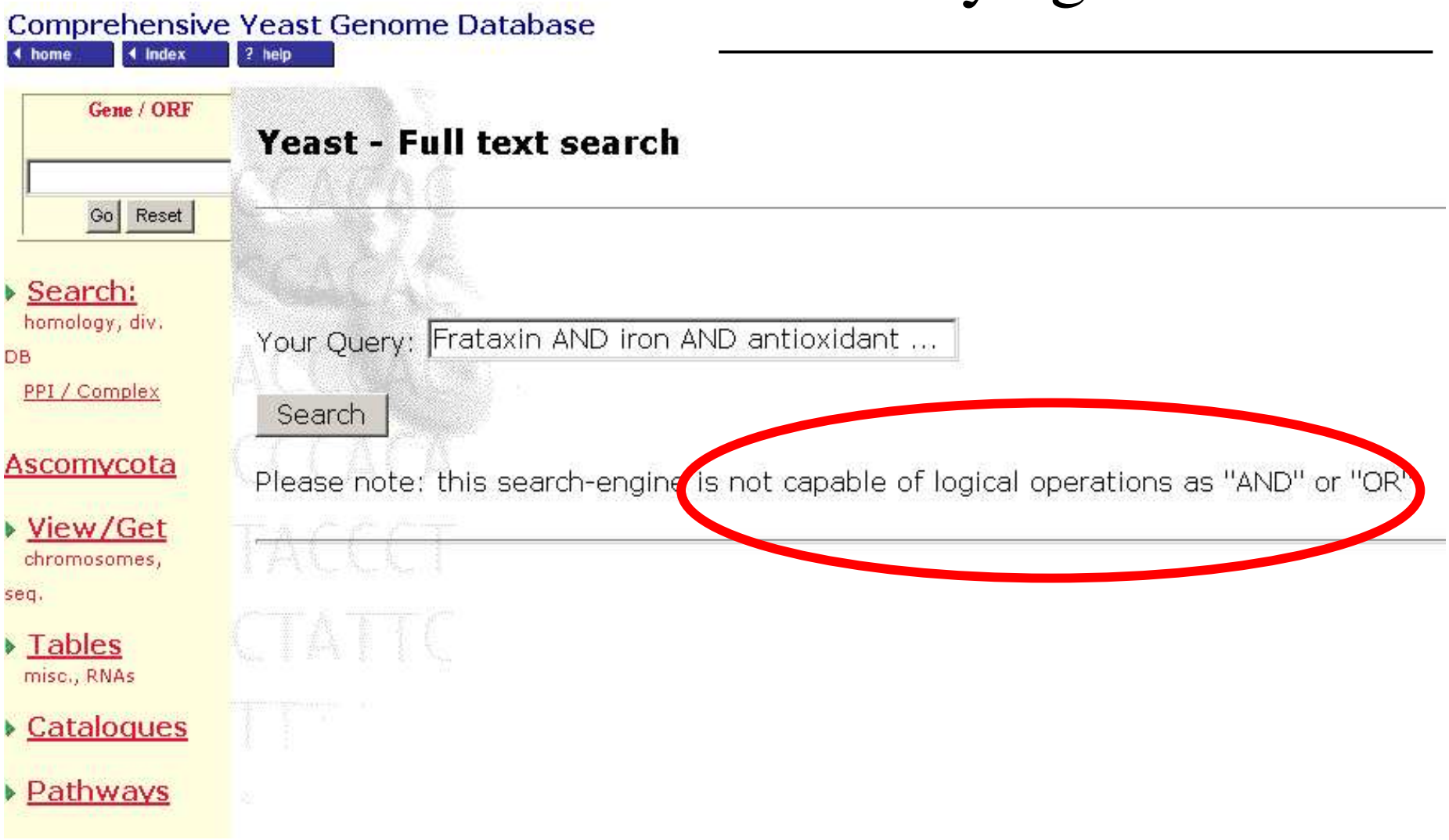
Not Found

The requested URL /htbin/bfind_litdb was not found on this server.

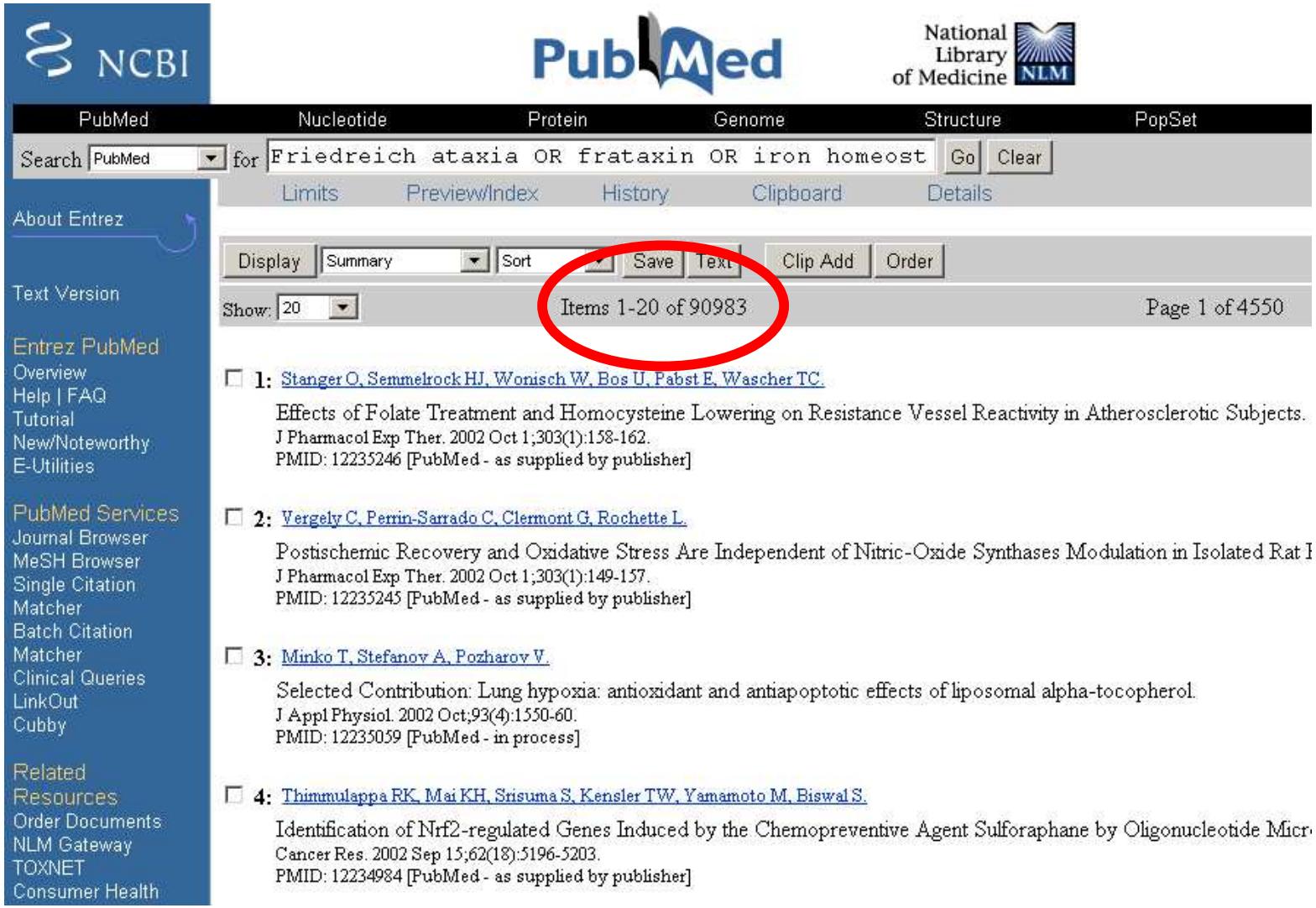
Apache/1.3.26 Server at fire2.scl.genome.ad.jp Port 80

(23% web-links invalid after 1 yr; rising to 53% after 6 years Lawrence et al., 2001)

Try again …


Comprehensive Yeast Genome Database
home
Index
? help
Gene / ORF
Go
Reset
Search:
homology, div.
DB
PPI / Complex
Ascomycota
View/Get
chromosomes,
seq.
Tables
misc., RNAs
Catalogues
Pathways
Yeast - Full text search
Your Query:
Frataxin AND iron AND antioxidant ...
Search
Please note: this search-engine is not capable of logical operations as "AND" or "OR"

Another attempt ….



and another …..


Bookmarks
Location: http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?CMD=search&DB=PubMed
Instant Message
The Net
Members
Marketplace
RealPlayer
NCBI
PubMed
National Library of Medicine
NLM
PubMed
Nucleotide
Protein
Genome
Structure
Search PubMed
for
Friedreich AND ataxia AND Frataxin AND iron A
Go
Clear
Limits
Preview/Index
History
Clipboard
Details
About Entrez
No items found.
Text Version
See Details.
Entrez PubMed
Overview
Help | FAQ
Tutorial
New/Noteworthy
E-Utilities
PubMed Services
Journal Browser
MeSH Browser
Single Citation
Matcher
Batch Citation

And again ? …


CiteSeer
Scientific Literature Digital Library


Google™

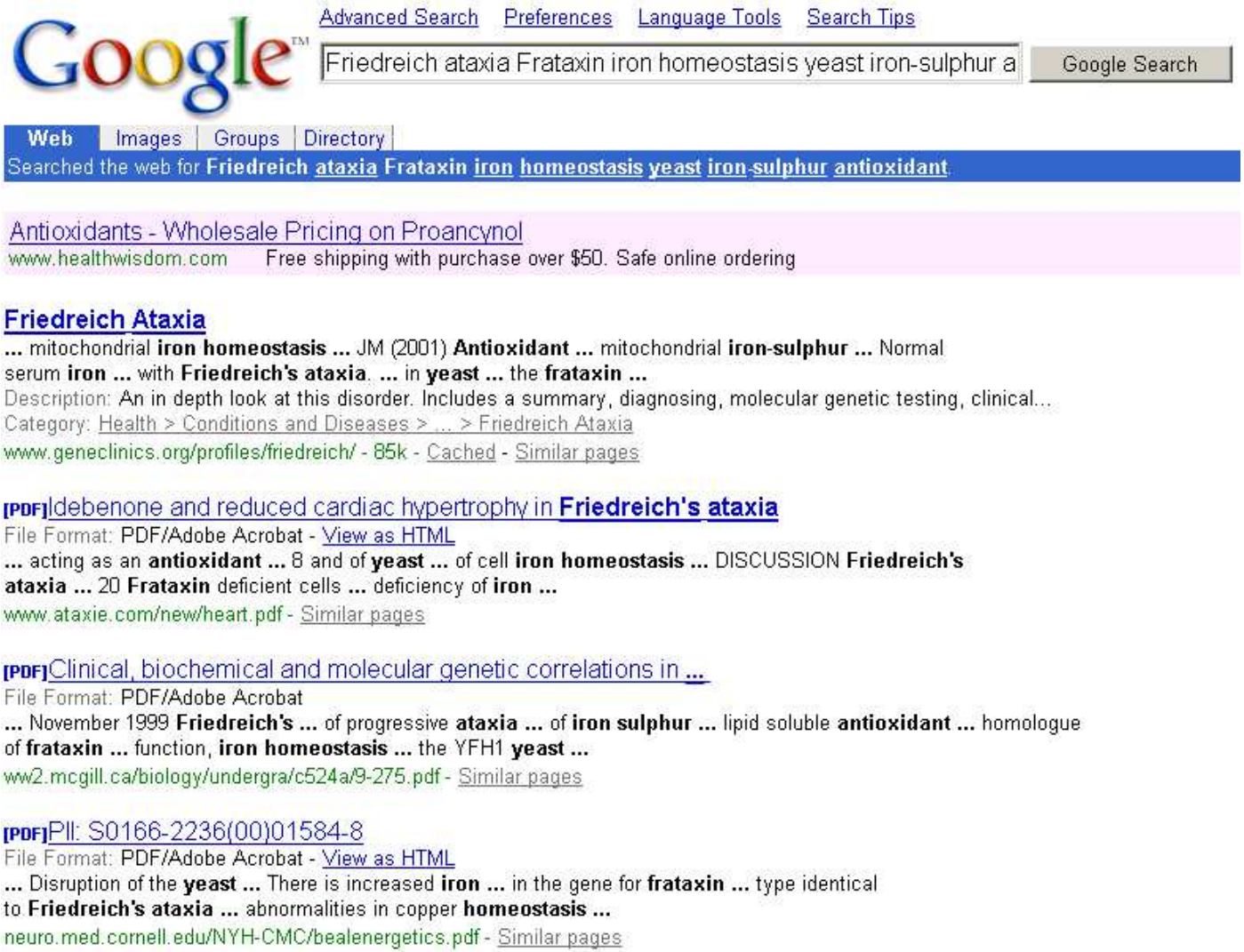

Ask Jeeves®
Ask.com


DOGPILE®
All results, no mess.


altavista
THE SEARCH COMPANY


metacrawler®
Search the Search Engines!

Success? ….


Google
Advanced Search
Preferences
Language Tools
Search Tips
Friedreich ataxia Frataxin iron homeostasis yeast iron-sulphur a
Google Search
Web
Images
Groups
Directory
Searched the web for Friedreich ataxia Frataxin iron homeostasis yeast iron-sulphur antioxidant
Antioxidants - Wholesale Pricing on Proancynol
www.healthwisdom.com Free shipping with purchase over $50. Safe online ordering
Friedreich Ataxia
... mitochondrial iron homeostasis ... JM (2001) Antioxidant ... mitochondrial iron-sulphur ... Normal serum iron ... with Friedreich's ataxia. ... in yeast ... the frataxin ...
Description: An in depth look at this disorder. Includes a summary, diagnosing, molecular genetic testing, clinical...
Category: Health > Conditions and Diseases > ... > Friedreich Ataxia
www.geneclinics.org/profiles/friedreich/ - 85k - Cached - Similar pages
[PDF]Idebenone and reduced cardiac hypertrophy in Friedreich's ataxia
File Format: PDF/Adobe Acrobat - View as HTML
... acting as an antioxidant ... 8 and of yeast ... of cell iron homeostasis ... DISCUSSION Friedreich's ataxia ... 20 Frataxin deficient cells ... deficiency of iron ...
www.ataxie.com/new/heart.pdf - Similar pages
[PDF]Clinical, biochemical and molecular genetic correlations in ...
File Format: PDF/Adobe Acrobat
... November 1999 Friedreich's ... of progressive ataxia ... of iron sulphur ... lipid soluble antioxidant ... homologue of frataxin ... function, iron homeostasis ... the YFH1 yeast ...
ww2.mcgill.ca/biology/undergra/c524a/9-275.pdf - Similar pages
[PDF]PII: S0166-2236(00)01584-8
File Format: PDF/Adobe Acrobat - View as HTML
... Disruption of the yeast ... There is increased iron ... in the gene for frataxin ... type identical to Friedreich's ataxia ... abnormalities in copper homeostasis ...
neuro.med.cornell.edu/NYH-CMC/bealenergetics.pdf - Similar pages

A next generation, scalable information service for the life sciences that will interlink factual and image data repositories with the research literature

EU Quality of Life research infrastructure:
platform under construction

The current E-BioSci partnership

- Distributed network of information resources
- Europe-based; world-wide role (Act locally - think globally!)

The E-BioSci platform

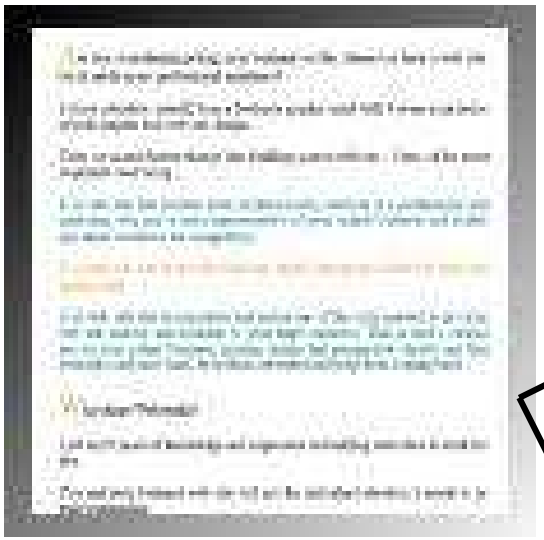
- Free access to the academic research community
- Set of *distributed* resources (literature, sequence- and image-databases)
- Integrated access to bibliographic abstracts, full text, factual and image databases
- Full-text search implemented via conceptual fingerprinting
  - Cross-repository
  - Cross-language
  - Implementation of effective 2-way links between databases and literature via gene symbol recognition

A discovery tool

powered by Collexis®

Conceptual fingerprints

powered by Collexis®

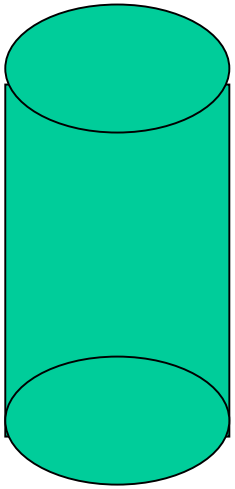


Full text document

Linkage of index terms to thesaurus

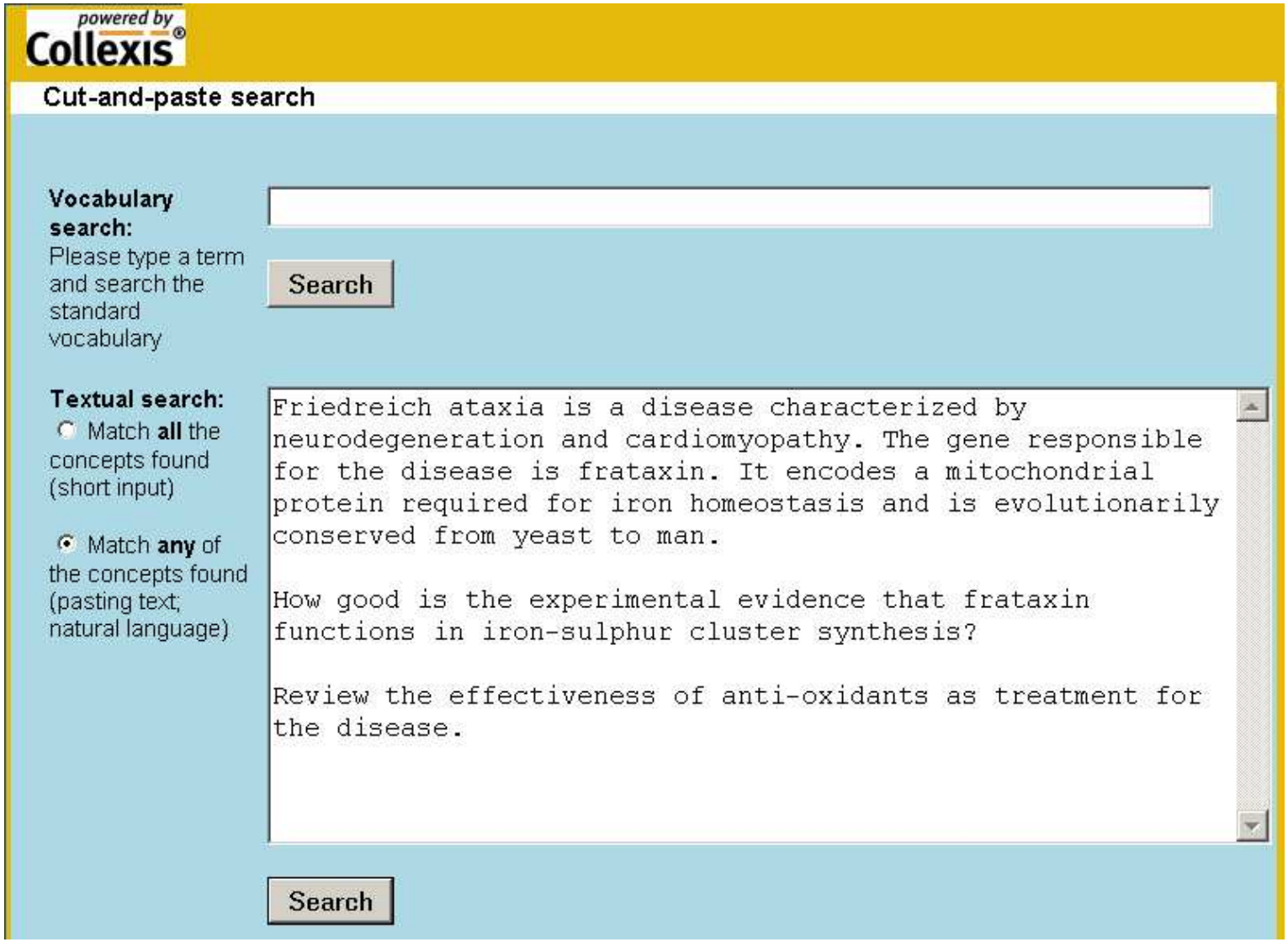
C19881 0.99
C92992 0.67
C02002 0.66
C99229 0.44
C00392 0.33
C93939 0.21

Fingerprint database

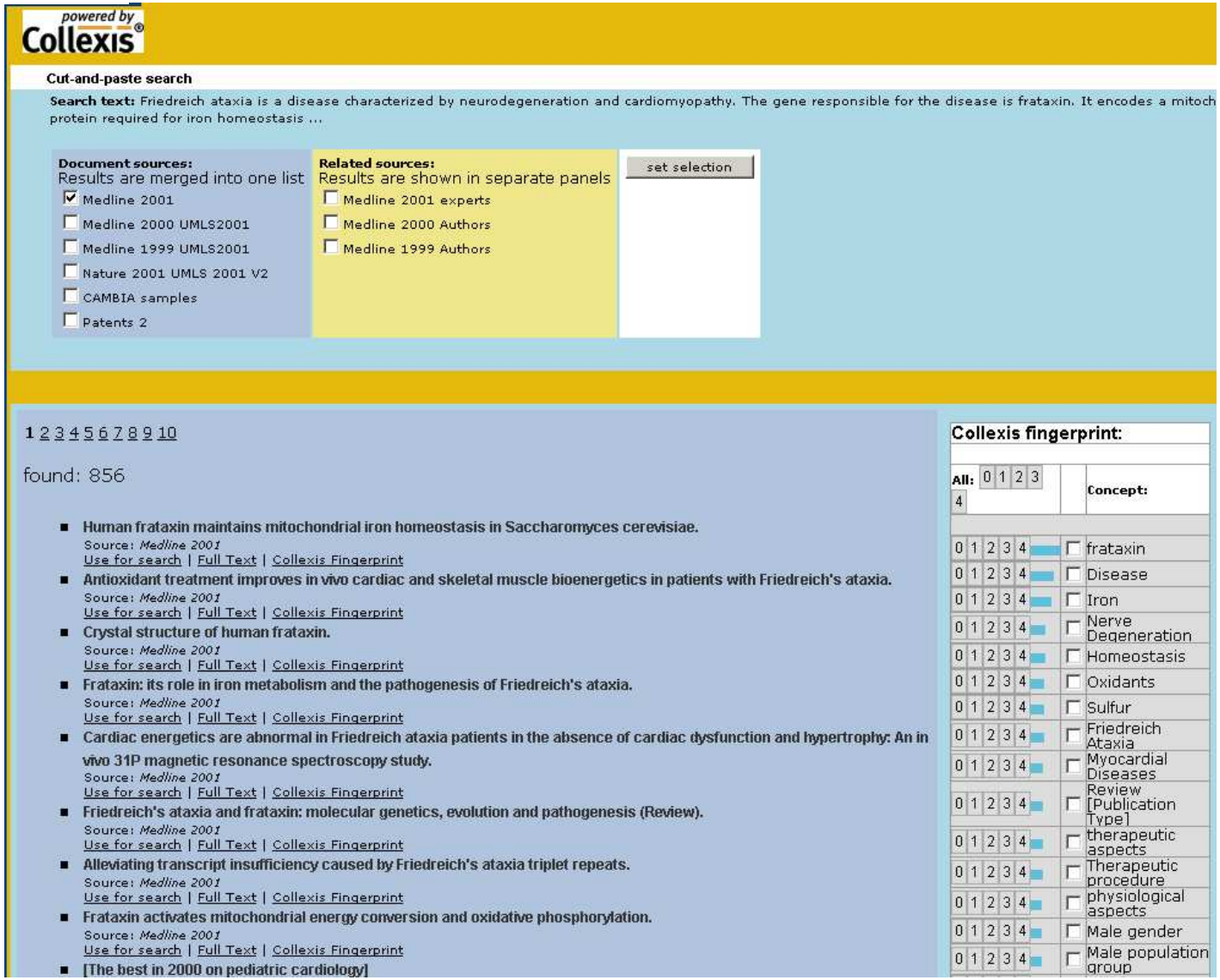


- 1 CFP = 400 bytes
- Abstraction: 250.000 pages/PC/day
- Matching: 500.000 CFP's: 40 millisec.

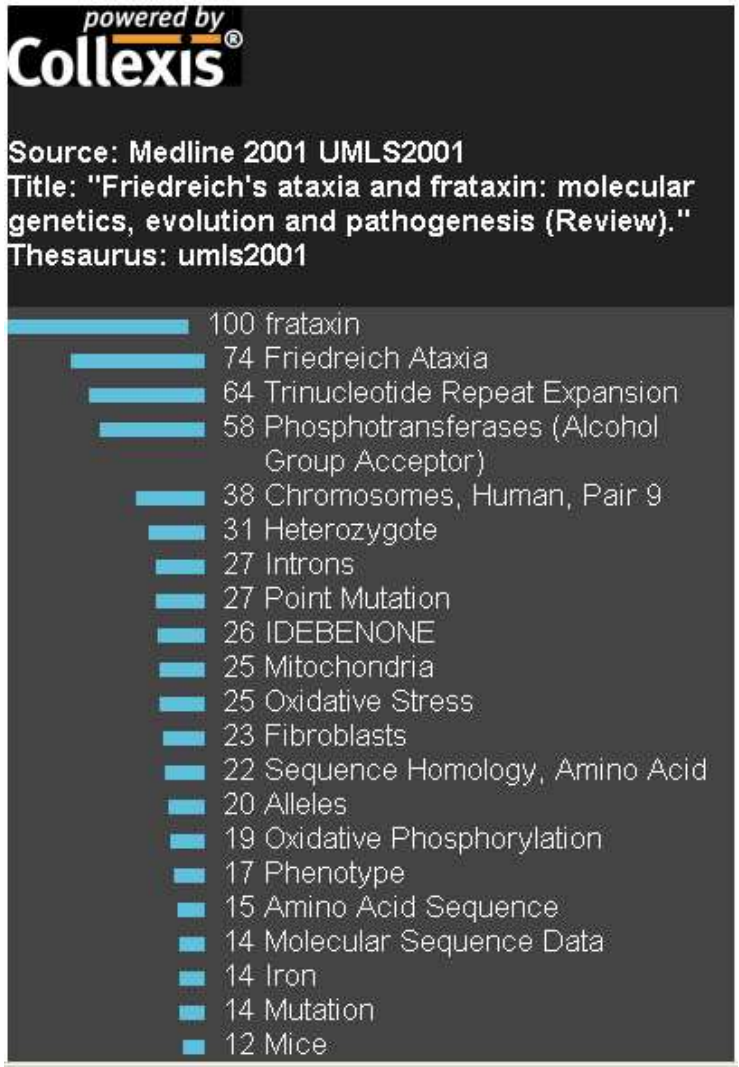
E-BioSci Our literature search once more….

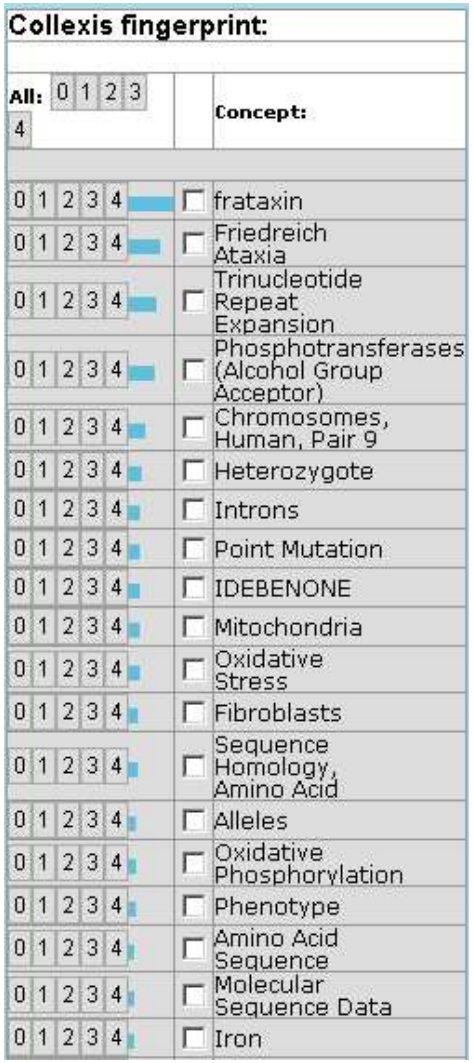


First search results …

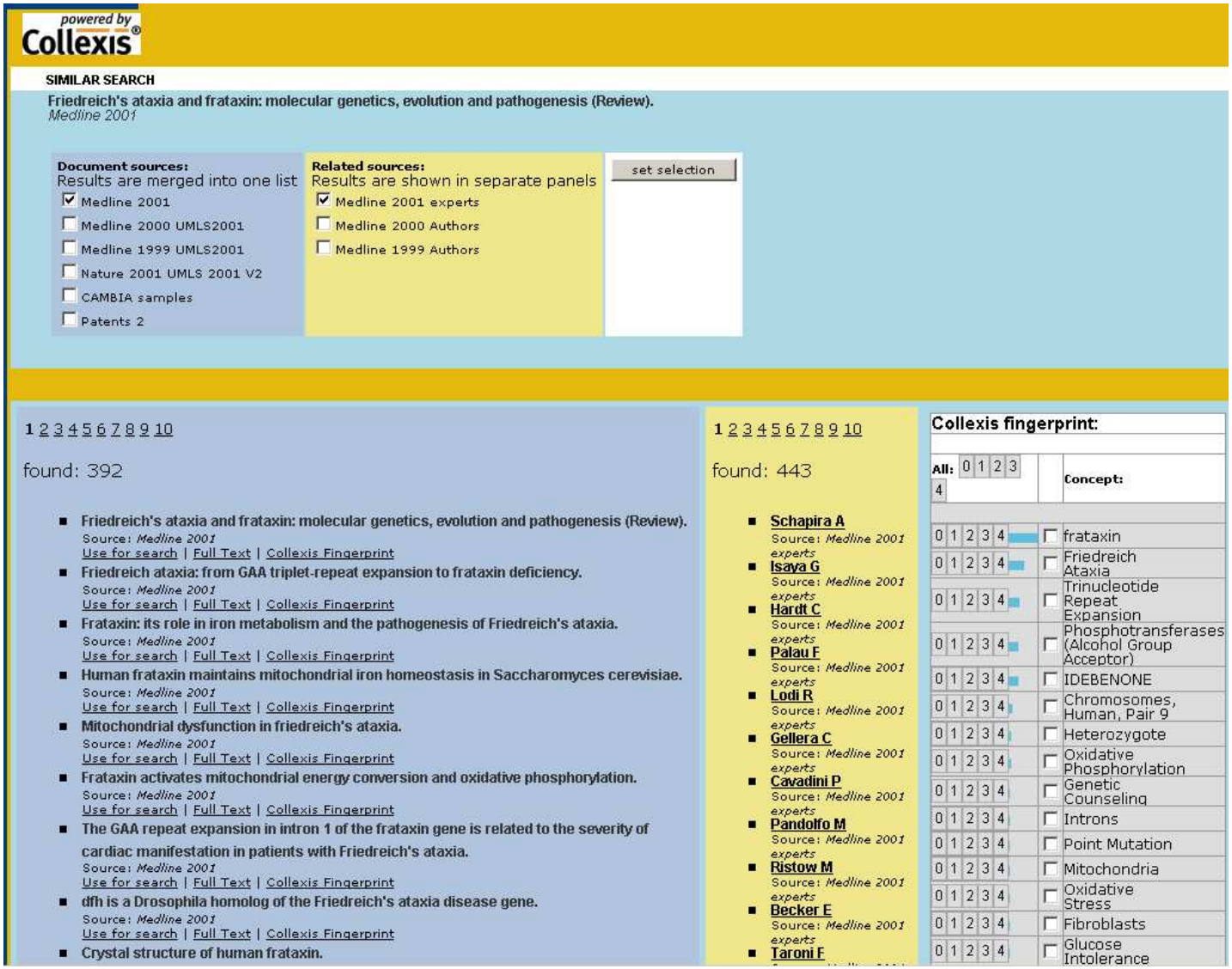


Refining the search …..

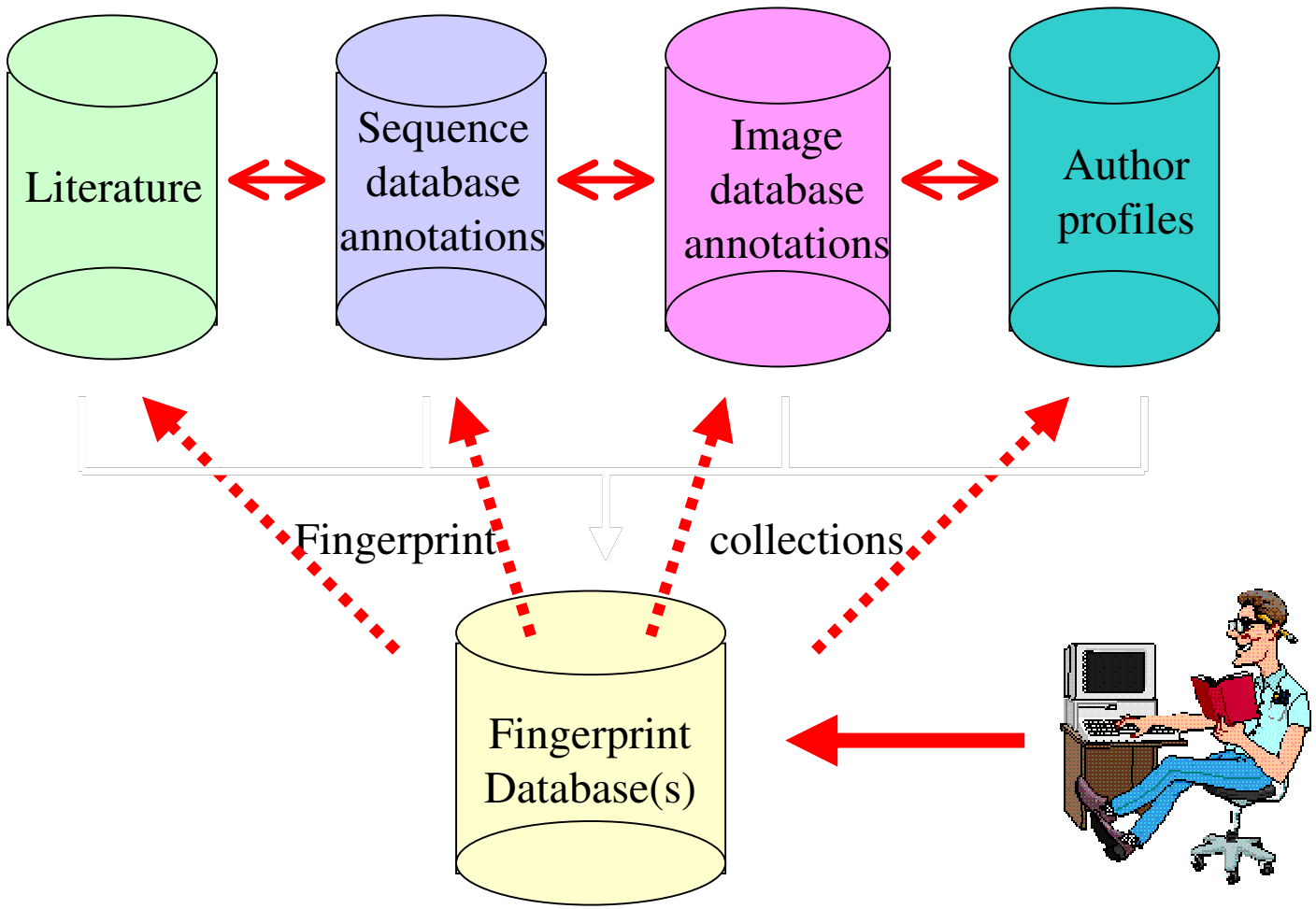




The result ….


powered by Collexis
SIMILAR SEARCH
Friedreich's ataxia and frataxin: molecular genetics, evolution and pathogenesis (Review).
Medline 2001
Document sources:
Results are merged into one list
Medline 2001
Medline 2000 UMLS2001
Medline 1999 UMLS2001
Nature 2001 UMLS 2001 V2
CAMBIA samples
Patents 2
Related sources:
Results are shown in separate panels
Medline 2001 experts
Medline 2000 Authors
Medline 1999 Authors
set selection
1 2 3 4 5 6 7 8 9 10
found: 392
Friedreich's ataxia and frataxin: molecular genetics, evolution and pathogenesis (Review).
Source: Medline 2001
Use for search | Full Text | Collexis Fingerprint
Friedreich ataxia: from GAA triplet-repeat expansion to frataxin deficiency.
Source: Medline 2001
Use for search | Full Text | Collexis Fingerprint
Frataxin: its role in iron metabolism and the pathogenesis of Friedreich's ataxia.
Source: Medline 2001
Use for search | Full Text | Collexis Fingerprint
Human frataxin maintains mitochondrial iron homeostasis in Saccharomyces cerevisiae.
Source: Medline 2001
Use for search | Full Text | Collexis Fingerprint
Mitochondrial dysfunction in friedreich's ataxia.
Source: Medline 2001
Use for search | Full Text | Collexis Fingerprint
Frataxin activates mitochondrial energy conversion and oxidative phosphorylation.
Source: Medline 2001
Use for search | Full Text | Collexis Fingerprint
The GAA repeat expansion in intron 1 of the frataxin gene is related to the severity of cardiac manifestation in patients with Friedreich's ataxia.
Source: Medline 2001
Use for search | Full Text | Collexis Fingerprint
dfh is a Drosophila homolog of the Friedreich's ataxia disease gene.
Source: Medline 2001
Use for search | Full Text | Collexis Fingerprint
Crystal structure of human frataxin.
1 2 3 4 5 6 7 8 9 10
found: 443
Schapira A
Source: Medline 2001 experts
Isaya G
Source: Medline 2001 experts
Hardt C
Source: Medline 2001 experts
Palau F
Source: Medline 2001 experts
Lodi R
Source: Medline 2001 experts
Gellera C
Source: Medline 2001 experts
Cavadini P
Source: Medline 2001 experts
Pandolfo M
Source: Medline 2001 experts
Ristow M
Source: Medline 2001 experts
Becker E
Source: Medline 2001 experts
Taroni F
Collexis fingerprint:
All: 0 1 2 3 4
Concept:
frataxin
Friedreich Ataxia
Trinucleotide Repeat Expansion
Phosphotransferases (Alcohol Group Acceptor)
IDEBENONE
Chromosomes, Human, Pair 9
Heterozygote
Oxidative Phosphorylation
Genetic Counseling
Introns
Point Mutation
Mitochondria
Oxidative Stress
Fibroblasts
Glucose Intolerance

Interconnecting searchable resources



E-BioSci

The prototype platform:

Currently under construction,
pre-testing, evaluation ….

See http://www.e-biosci.org for further information

Acknowledgements

- Frank Gannon, Executive Director EMBO
- … and many others who contributed ideas to the concept of E-BioSci
- The E-BioSci partners
- European Commission
  (contract no QLRI-2001-30266)