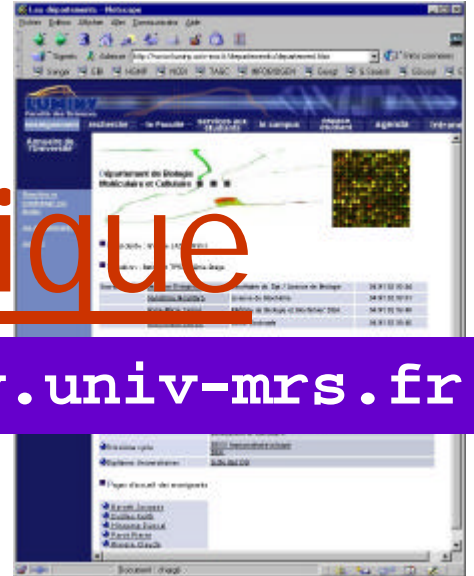


Bioinformatique

<http://www.luminy.univ-mrs.fr>



- Qu 'est ce?

Ce que ce n 'est pas...

Définitions!

Exemples

Science / Technologie ?

(créateurs / utilisateurs?)

Bref historique

Catégories

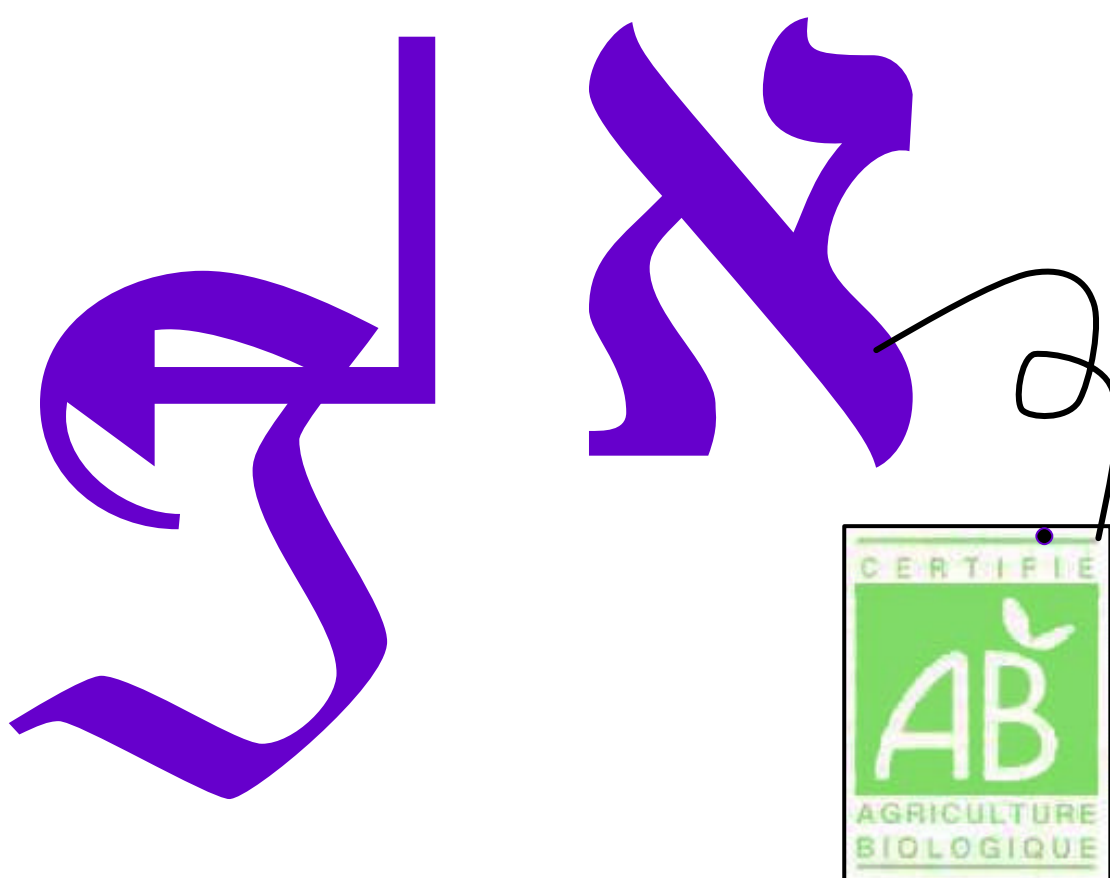
- A quoi ça ressemble?

lignes de commande

logiciels interactifs

interfaces Internet

Informatique «Bio»





Enki Bilal (1977) Mémoires d'Outre Espace
éditions *Les humanoïdes Associés*

Bioinformatiques

Bioinfo de labo

support au clonage,
séquençage,
& PCR...

Analyse séquence

identification gènes,
comparaisons de
séquences,
prédiction motifs...

Phylogénie

évolution à
l'échelle
moléculaire...

Structure des

protéines

calcul,
visualisation,
prédiction...

Liaison génétique

gènes candidats
de maladies
génétiques...

Génomique

fonctionnelle

transcriptome,
protéome,
interactome...

Bref historique

- 1970** Needleman & Wunsch
Alignement optimal de 2 séquences
- 1974** Chou & Fasman
Prédiction structure II protéines
- 1978** Dayhoff *et. al.*
Matrice de substitution (PAM)
- 1980** EMBL data library
Banque internationale de séquence
- 1981** Smith & Waterman
Similarités de séquences dans les banques
- 1981** Felsenstein *et. al.*
Arbres phylogénét. par comparaison de séq. (phylip)
- 1984** Devereux *et. al.*
Logiciel d'analyse de séquence (UW GCG)
- 1987** Feng & Doolittle
Alignement multiple progressif de séquences
- 1988** Bairoch *et. al.*
Banque de signatures protéiques (PROSITE)
- 1988** Pearson & Lipman
Sim. de séq. dans les banques (FASTA)

Bref historique (suite)

1990 Atschul *et. al.*

Sim. de séq. dans les banques (BLAST)

1991 Bowie *et. al.*

Prédiction struct. III protéines

1993 Borodovsky *et. al.*

Prédiction gènes génomes bactériens (GeneMark)

1995 Venter *et. al.*

Séquençage 'shotgun' génome Haemophilus 1.8Mb

1997 Sonnhammer *et. al.*

Banque de domaines protéiques (PFam)

1997 Burge *et. al.*

Prédiction gènes génomes eucaryotes (GENSCAN)

1998 Eisen *et. al.*

Analyse des profils d'expression (CLUSTER)

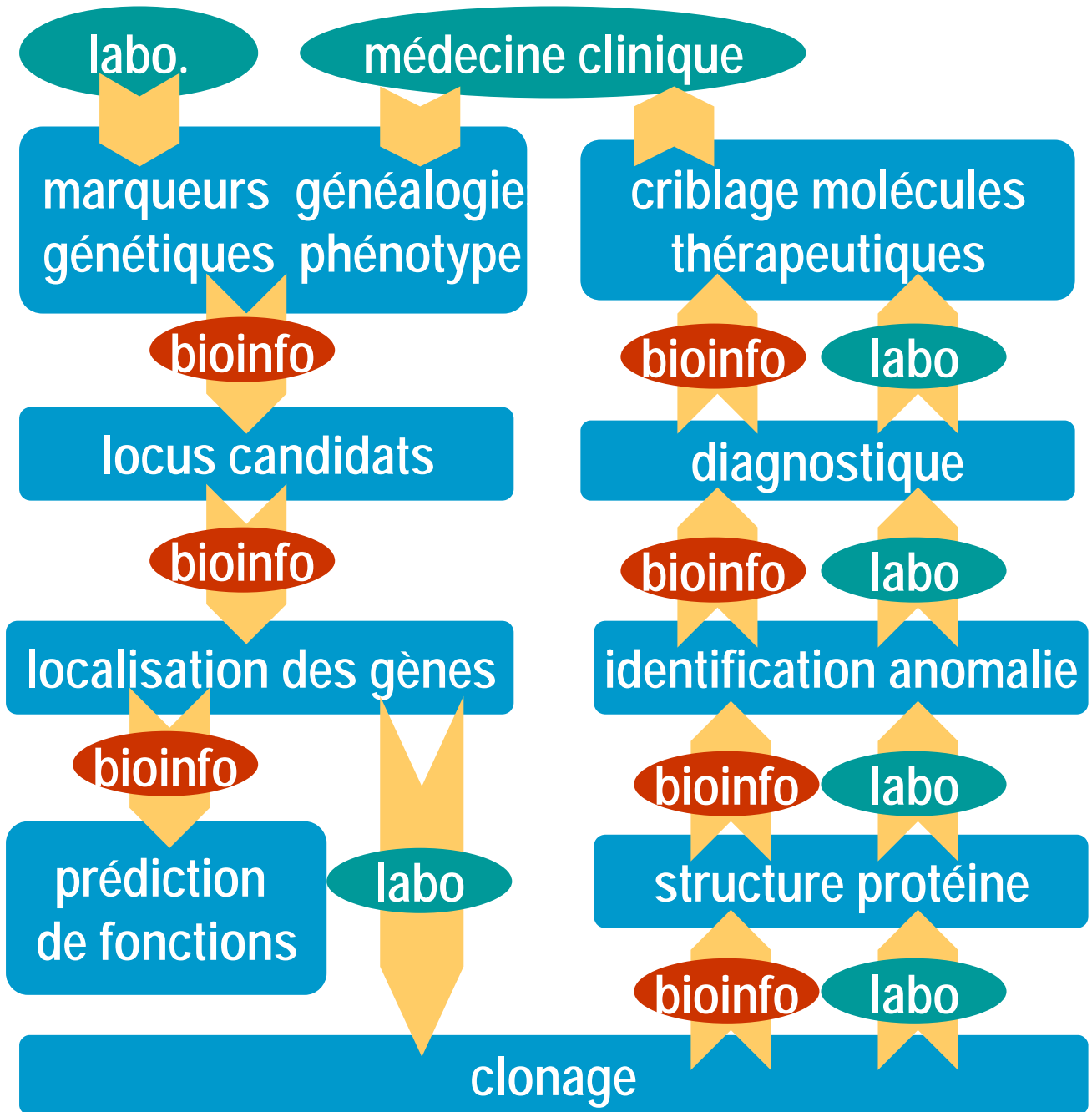
2000 Venter *et. al.*

'Shotgun' génome complet Drosophile 120 Mb

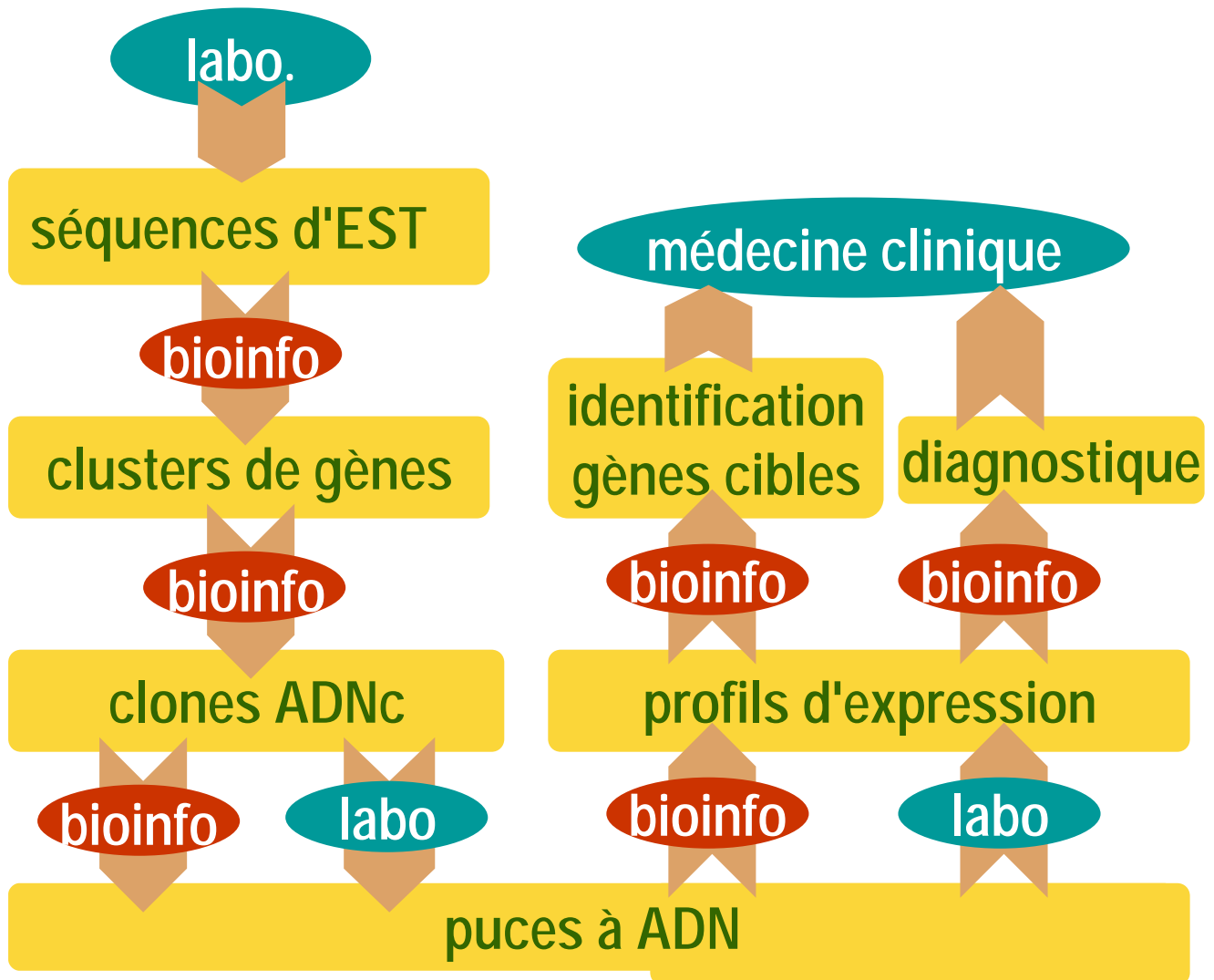
2001 Venter *et. al.*

'Shotgun' génome complet Humain 3000 Mb

Exemple d 'application 1



Exemple d 'application 2



Ligne de commande

```
Unix % genefinder -corrfactor 0.8 -tablenamefile
/packages/genefinder_washu/genefinder/human/humtables.hex -cpenalty 6.35 -penaltyfactor
0.85 -maxintronlength 10500 -minintronlength 70 -penaltycluster 50 -seqfile
pax6.wild.type.seq

Genefinder Version 084, release 960118
Parameters:
-seqfile pax6.wild.type.seq, -genefile "", -debug 0, -syncodon FALSE,
-normalize FALSE, -exoncutoff 3, -genecutoff 0, -atgcutoff 0,
-intron3cutoff 0, -intron5cutoff 0, -orfcutoff 5, -minexonlength 33,

Sequence: hsa1280.embl,.C
Length: 22253, nuc counts (N,A,C,G,T): 0 6645 4848 4692 6068

394 3' sites w/ scores > 0.00
310 5' sites w/ scores > 0.00
232 atg sites w/ scores > 0.00

Sequence: hsa1280.embl,
Length: 22253, nuc counts (N,A,C,G,T): 0 6068 4692 4848 6645

408 3' sites w/ scores > 0.00
429 5' sites w/ scores > 0.00
249 atg sites w/ scores > 0.00

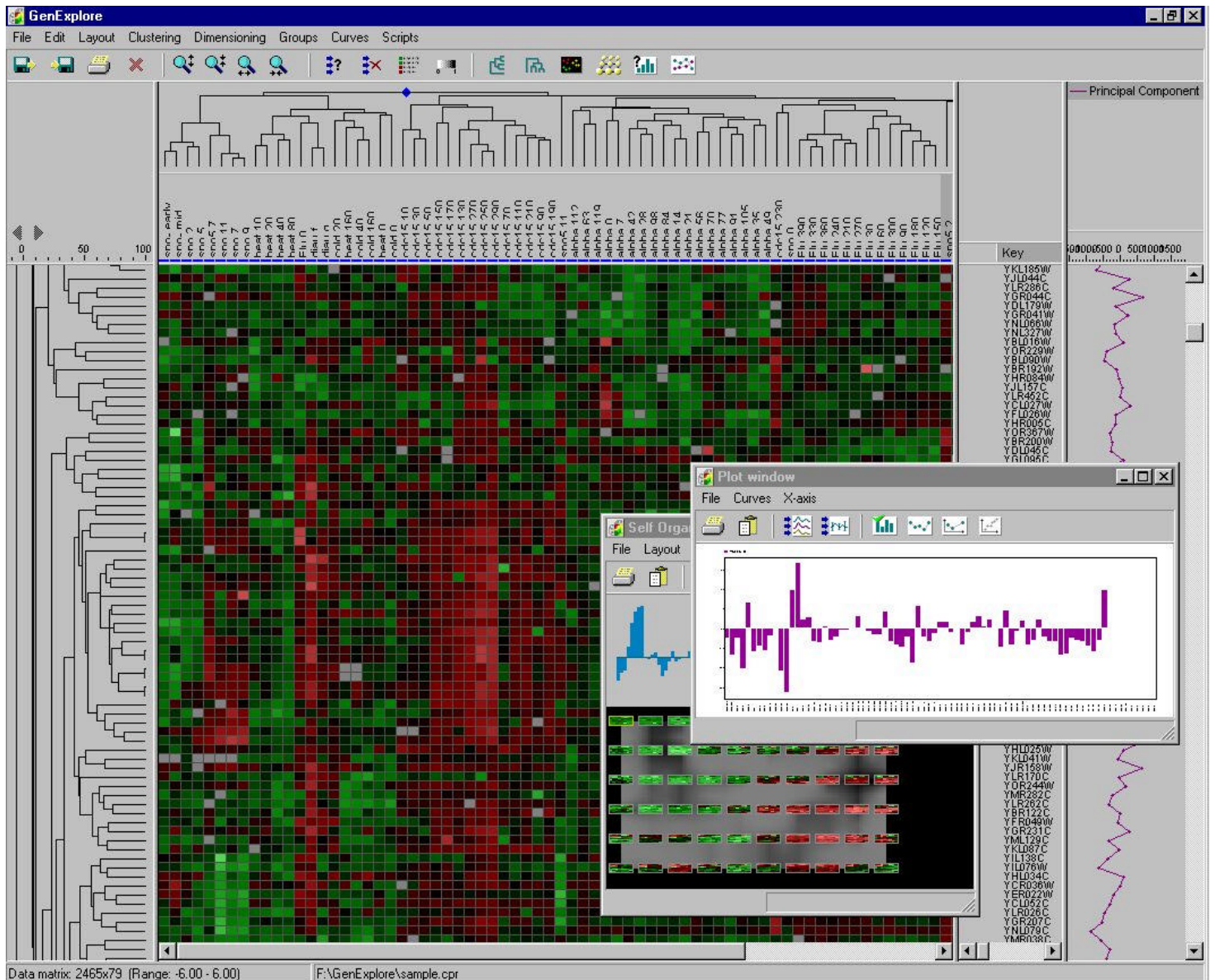
No predicted genes in hsa1280.embl,.C.

ORFs:

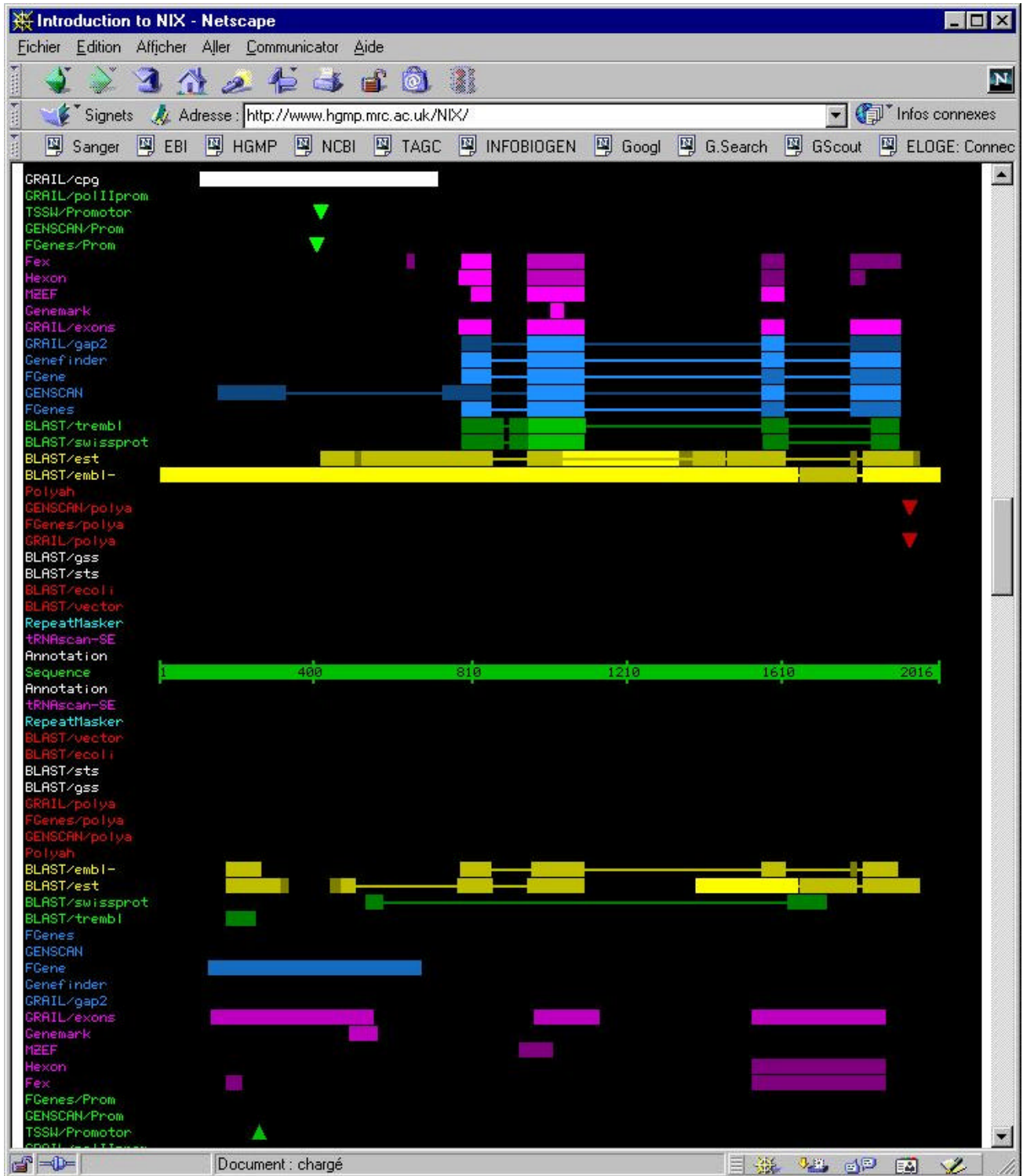
ORF: 17183 16701 483 4.16 5.79 6.00
*TKIIISTMSKRPTPRHCYHTEPAMLSPRECLELCLLEGQLMSLGCQKARQFPFSCRTGRLLQLSSRFFRLYQTVLLCPGNPDLQSTHIRA
QVHTYRPSAPTSLLFSSHLLHLACDIRVSKVSAVYTTEPWAGTGCLTVEITASEKEEVSQPAAVACTPT*
17144 17109 36 1.12
1 3' 17122 1.62
atg 17114 2.07
2 3' 17100 0.99
atg 17066 1.10
2 5' 17056 2.54
1 3' 17056 0.92
1 5' 17054 1.24
2 3' 17010 1.55
17009 16983 27 2.10
0 3' 16997 0.85
2 3' 16989 0.67
2 3' 16977 0.94
0 3' 16967 1.24
0 3' 16931 1.53
16853 16719 135 5.79
1 3' 16798 1.52
2 5' 16795 1.03
0 5' 16737 2.49

ORF: 5582 4998 585 5.66 4.06 6.00
*SQSPKNTLCGWGGDCQLTQNFLATRRVKVSPWTSSTSGPNHRRRLADLTLTSRQPQTAGDENPPESGATSGGSCCPRLLPNSKKRKG
GVSCAPSSWEPTHREGPGVEALGSALPGLGHRALSPSPRLLGLSGPDSCQGRDHQDLRLLLGAEGWTLVSAKFGCHLGVSSVYKWEVYGFK
SQAGMGLPTLGGVNR*
5507 5457 51 1.27
2 3' 5445 2.97
5408 5256 153 4.06
0 3' 5405 3.67
1 5' 5363 1.70
0 3' 5324 1.78
```

Logiciels interactifs



Interfaces Internet



Centres de ressources

Infobiogen

Page d'accueil INFOBIOGEN - Netscape

Centre de Ressources INFOBIOGEN
Le GIS INFOBIOGEN devient le Centre de Ressources INFOBIOGEN

Accueil
Activités Accès au site R&D
Services
SRS
PDB
GDB
GoldenPath
SEQWEB
Analyse de Séquences
Liens
DEAMBULUM
ORPHANET
AGORA

Présentation d'INFOBIOGEN
Education-Enseignement

Services en Accès Anonyme
Banques de données de biomolécules, génomes, maladies génétiques

Formations

Documentations
Informatique - Bioinformatique - Médecine
Destinés aux utilisateurs détenteurs d'un compte

Informations sur les Services en Accès Authentifié

Nouveautés
Liens sur l'ailleurs...

EBI

EMBL Outstation
European Bioinformatics Institute

The EMBL Outstation - European Bioinformatics Institute (EBI) is a centre for research and services in bioinformatics. The Institute manages databases of biological data including nucleic acid, protein sequences and macromolecular structures.

EMBL Webin-Align	Ethics 2001 Workshop	Ensembl
January 12th 2001 Web based alignment tool for all types of alignment data from the analysis and annotation analysis - (GDB)	January 15th 2001 The EBI organizes a workshop on ethics, legal and social implications of Human Genetic Research (GDB)	January 16th 2001 Ensembl 2.5 released with many improvements - (GDB)

Information
The EBI main information pages about staff, the information, travel, jobs, conferences and seminars - visitors programme

Databases
Access to EBI and the EMBL, GDB, GOLD, GOLDEN, SWISS-PROT, TrEMBL, MSD databases, etc.

Tools
Fast sequence database searches and interactive analysis tools e.g. FASTA, VAL-PROT, ClustalW, InterProScan etc.

Submissions
Information about how to submit data to databases maintained at the EBI, including EMBL using dEEMBL

Groups
Index of Groups and Team Leaders at the EBI and their projects e.g. Proteome Analysis Collaborative, Complex Genomes, EMBL's GenBank, Sequence Variation Database, Microarray, HEPATITIS, etc.

Publications
Lists of publications from the EBI

EIP
The EBI databases and software public ftp archives

Search EBI
The searchable collection of EBI's web pages, can be used to look up a person or a certain topic.

EMBL-European Bioinformatics Institute
Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SD, UK

NCBI

NCBI HomePage - Netscape

National Center for Biotechnology Information
National Library of Medicine National Institutes of Health

PubMed Entrez BLAST OMM Taxonomy Structures

Search: GenBank for

SITE MAP
About NCBI
GenBank
Molecular databases
Literature databases
Genomic biology
Tools for data mining
Research at NCBI
Education

What does NCBI do?
Established in 1988 as a national resource for molecular biology information, NCBI creates public databases, conducts research in computational biology, develops software tools for analyzing genome data, and disseminates biomedical information - all for the better understanding of molecular processes affecting human health and disease.

OMM in Entrez
Online Mendelian Inheritance in Man (OMIM) has now been incorporated into NCBI's Entrez system, as you can now search OMM using the same approach as for PubMed. Make the most of OMM's great content by use of Links, History, Links, etc!

NCBI in the News
The Genes and Disease site is described as: "Cutting-edge scientific knowledge delivered in a remarkably clear, concise, and accessible style that is free of jargon. This is an excellent..."

Hot Spots
Cancer genome anatomy project
Clusters of orthologous groups
Coffee Break
Electronic PCR
Gene expression credits
Genes and disease
Human genome resources
Human/mouse homolog maps
LocalLink
Mammal genetics & genomics
ORF Finder
Reference sequence project
RefSeq database

Expasy

Expasy Molecular Biology Server

This is the Expasy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (SIB). This server is dedicated to the analysis of protein sequences and structures as well as 2-D PAGE (Discipline).

Essential! Databases | Index Site!

Databases	Tools and Software Packages
<ul style="list-style-type: none"> SWISS-PROT and TrEMBL - Protein sequences PROSITE - Protein families and domains SWISS-2DPAGE - Two-dimensional polyacrylamide gel electrophoresis SWISS-3DIMAGE - 3D images of proteins and other biological macromolecules SWISS-MODEL Repository - Automatically generated protein models CDLBASE - CD40 ligand database ENZYME - Enzyme nomenclature SeqAnalRef - Sequence analysis bibliographic references Links to many other molecular biology databases 	<ul style="list-style-type: none"> Proteomics tools <ul style="list-style-type: none"> Identification and characterization <ul style="list-style-type: none"> OMA - O-Protein Peptide sequencing 0-Scan and protein 2D-probe Electrostatic interaction prediction Protein structure analysis Secondary structure prediction Tertiary structure Transmembrane regions detection Classification MolProbis - Software for 2-D PAGE analysis SWISS-MODEL - Automated knowledge-based protein modeling center Swiss-ProtServer - Molsoft/PC tool for structure display and analysis Databases: Manohar's Biochemical Pathways

Education and services	Documentation
<ul style="list-style-type: none"> The Expasy FTP server Swiss-Check - automatically obtain (by email) new sequence entries relevant to your field(s) of interest 2-D PAGE training - attend a one-week course in Geneva SWISS-2DSEARCH - get your 2-D Gels performed according to Swiss standards 	<ul style="list-style-type: none"> Who's New in Expasy SWISS-EL-ASP electronic bulletins SWISS-PROT documents How to create HTML links to Expasy Complete table of available documents

Links to lists of molecular biology resources	Links to good quality molecular biology servers
<ul style="list-style-type: none"> Access WWW links - The Expasy list of Biochemical servers Entrez - Search the Internet for molecular biology information NCBI - 2DPAGE - Link to 2-D PAGE database servers and 2-D PAGE related servers and services 	<ul style="list-style-type: none"> European Bioinformatics Institute (EBI) National Center for Biotechnology Information (NCBI) Japanese Genbanknet Australian National Genomic Information Service

Plan du cours

1. Introduction
2. Banques de données
3. Matrices de substitutions
4. Alignement de 2 séquences
5. Recherche dans les banques par similarité de séquence
6. Alignements multiples
7. Motifs & signatures protéiques
8. Profils & domaines protéiques

Bases de données

- Modes de constitution

Soumission directe:
archives **primaires**

redondantes

Sélection & expertise:
banques de **référence**

incomplètes

- Catégories de données

base de données moléculaires

banques de connaissances

- Accès

Numéros d'accession

Annotation

Formats

Banques de données moléculaires

Séquences nucléiques

EMBL/GENBANK

dbEST

*UNIGENE

*ENSEMBL

*RefSeq...

Séquences protéiques

TrEMBL

*SWISSPROT

*PIR

Structures protéines

PDB, * CATH, * SCOP..

** banques de référence expertisées (non redondantes)*

Banques de connaissances

Fonct/relat-ionnelles

LocusLink

GeneCards

PROSITE, PFam

KEGG...

Ontologies

Taxonomie

Symboles gènes

Enzymes

Tissus...

Organisme centrique

GDB

MGI

FlyBase

AceDB...

Bibliographiques

MEDLINE

Journaux Online

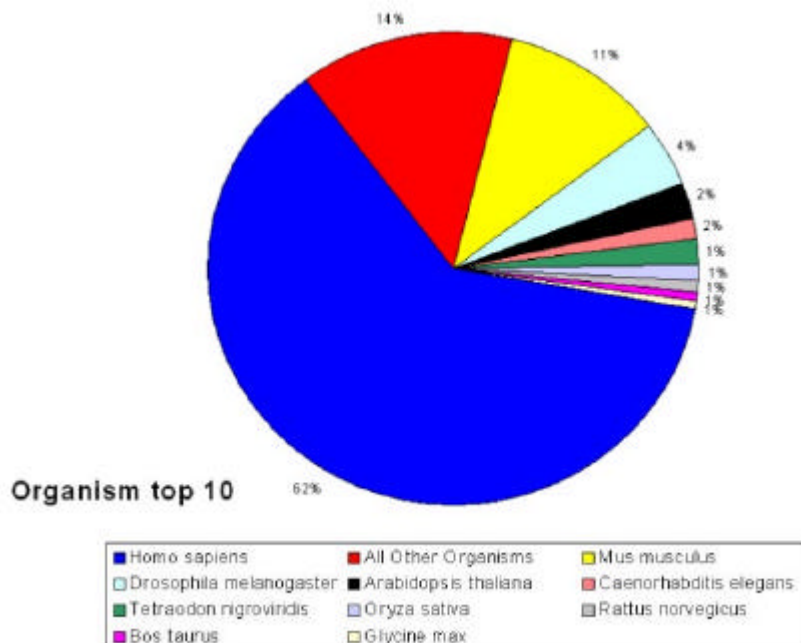
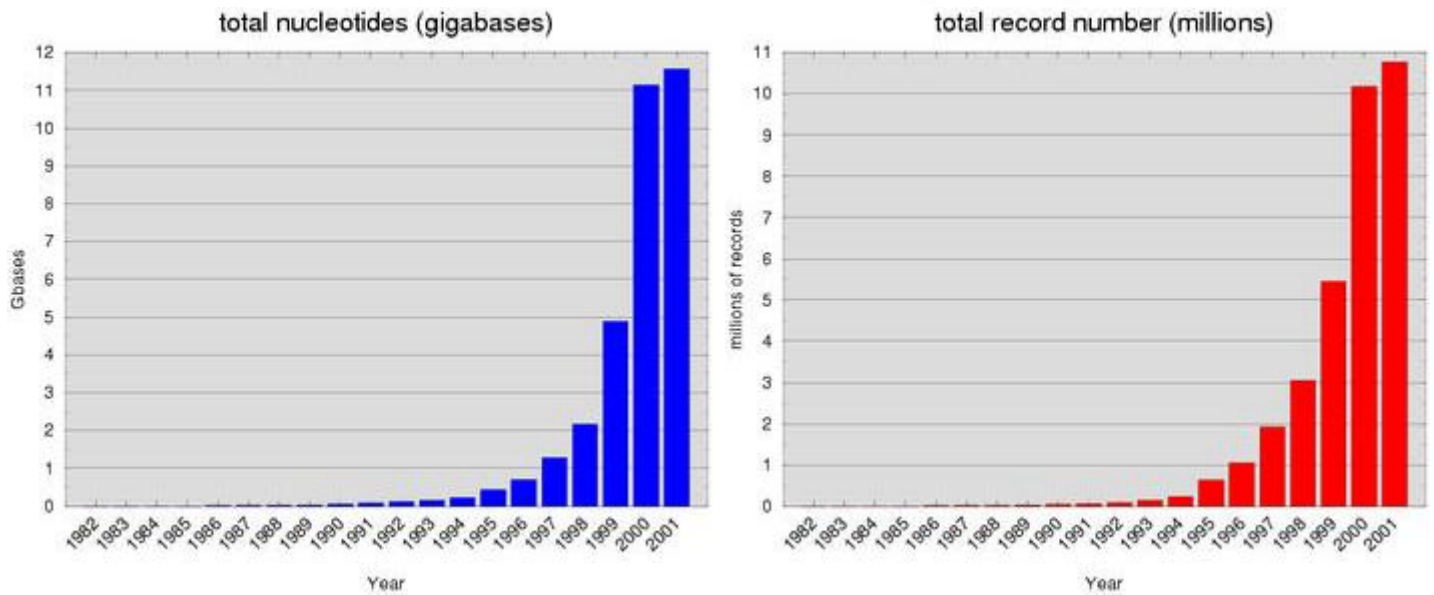
OMIM



EMBL Nucleotide Sequence Database Statistics

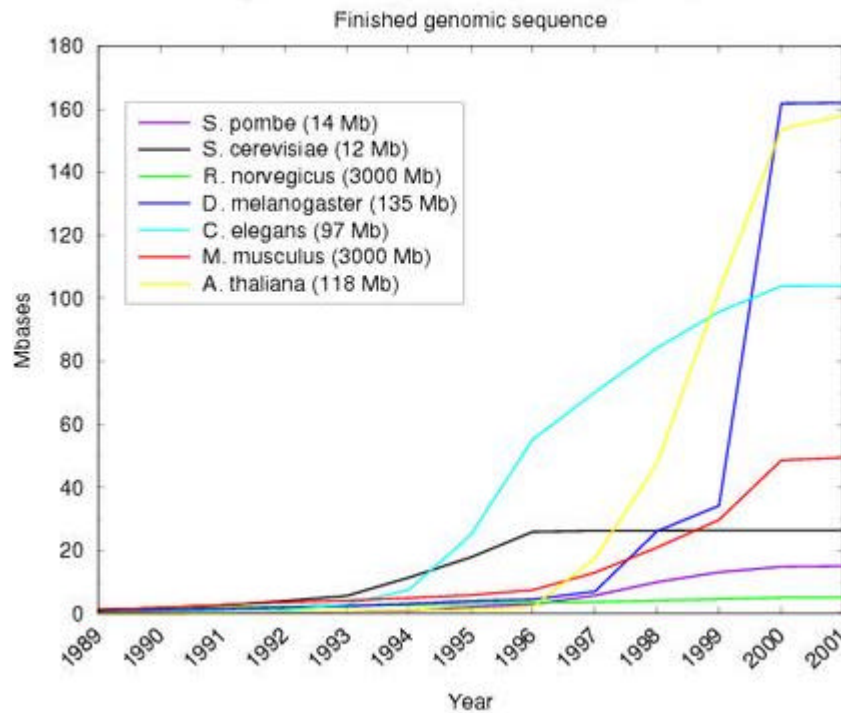
EMBL Database growth

This morning, the EMBL Database contained **11,571,085,945** bases in **10,770,275** records.

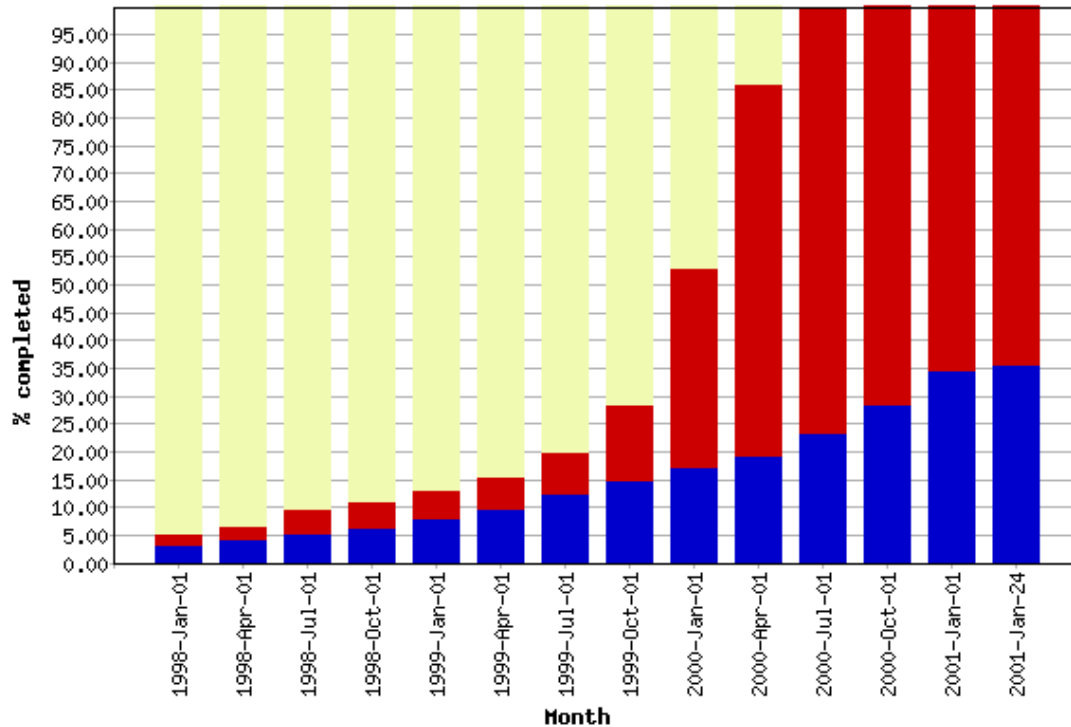


EMBL: génomes

Progress of Major Sequencing Projects



Progress of Human Genome Project (01 Jan 1998 - today)



Séquence brute

```
GGGTTTTATACGGATCTCCCTTCTCGTTTGATAATTATGCCATTAAAGGTTTTACCAGTTCATAAATTTA
GTAAAAATGAACCCCATAAAAAACAAAAGAGGTTTCATCTACTTTTTAAGGAATTAACCAAGGAATTTAAT
TCATATTAATAGCCATGGTTTCCAGTTTTTTACTGGCAGAGTACAAAAACCTAATAGTGAATCCCTCTG
AGCATTTCAAAATCTCAGTGAATGAAGATAATTTGACTGAATGGGATGTCATCTTAAAAGGCCACCTGA
CACTCTTTATGAGGGAGGCTTATTTCAAAGCAAAGATTGTCTTTCCCTCCAAAATACCCATATGAACCACC
AGATTAACATTCACCTCTGAAATGTGGCATCCCAATATCTACTCTGATGGGAAATTTATGTATTTCTATCT
TGCATGGAGACAATGCTGAAGAACAGGGAATGACTTGGTCTCCGGCTCAAAAGATTGATACCGTACTTCT
TAGTGTAATTTCTCTGCTCAATGAGCCAAATCCAGATTCTCCAGCAAATGTAGATGCAGCTAAAAGCTAC
CGTAAATATCTATATAAAGAGGATTTAGAATCATACCCCATGGAAGTTAAAAAGACTGTCAAAAAATCAT
TGGATGAGTGTTCAGCGGAAGACATAGAATATTTTAAAAATGTTCCAGTGAATGTTCTACCAGTACCCAG
TGATGATTATGAAGATGAAGAAATGGAGGATGGCACCTATATCTTAAACCTATGATGATGAGGATGAAGAA
GAGGATGAAGAGATGGATGATGAGTAGTGCTGATTTTTAATGCATAACATATTAGTTACTTACACTTTAGT
GCTTAGATTTTAGTGTTTAACTTTAGTGATTAGATTTTAGTGCTTAGATTTTAGTGTTTAACTTTAGT
GATTAGATTTTAGTGCTTAGATTTTAAATGTTTAACTTTAGTGATTAGATTTTAGTGCTTAGATTTTAA
```

Format FASTA

```
>X62440 African swine fever virus DNA for ubiquitin conjugating enzyme
GGGTTTTATACGGATCTCCCTTCTCGTTTGATAATTATGCCATTAAAGGTTTTACCAGTTCATAAATTTA
GTAAAAATGAACCCCATAAAAAACAAAAGAGGTTTCATCTACTTTTTAAGGAATTAACCAAGGAATTTAAT
TCATATTAATAGCCATGGTTTCCAGTTTTTTACTGGCAGAGTACAAAAACCTAATAGTGAATCCCTCTG
AGCATTTCAAAATCTCAGTGAATGAAGATAATTTGACTGAATGGGATGTCATCTTAAAAGGCCACCTGA
CACTCTTTATGAGGGAGGCTTATTTCAAAGCAAAGATTGTCTTTCCCTCCAAAATACCCATATGAACCACC
AGATTAACATTCACCTCTGAAATGTGGCATCCCAATATCTACTCTGATGGGAAATTTATGTATTTCTATCT
TGCATGGAGACAATGCTGAAGAACAGGGAATGACTTGGTCTCCGGCTCAAAAGATTGATACCGTACTTCT
TAGTGTAATTTCTCTGCTCAATGAGCCAAATCCAGATTCTCCAGCAAATGTAGATGCAGCTAAAAGCTAC
CGTAAATATCTATATAAAGAGGATTTAGAATCATACCCCATGGAAGTTAAAAAGACTGTCAAAAAATCAT
TGGATGAGTGTTCAGCGGAAGACATAGAATATTTTAAAAATGTTCCAGTGAATGTTCTACCAGTACCCAG
TGATGATTATGAAGATGAAGAAATGGAGGATGGCACCTATATCTTAAACCTATGATGATGAGGATGAAGAA
GAGGATGAAGAGATGGATGATGAGTAGTGCTGATTTTTAATGCATAACATATTAGTTACTTACACTTTAGT
GCTTAGATTTTAGTGTTTAACTTTAGTGATTAGATTTTAGTGCTTAGATTTTAGTGTTTAACTTTAGT
GATTAGATTTTAGTGCTTAGATTTTAAATGTTTAACTTTAGTGATTAGATTTTAGTGCTTAGATTTTAA
```

Fiche EMBL

```
ID   ASFVUCE   standard; DNA; VRL; 979 BP.
XX
AC   X62440;
XX
SV   X62440.1
XX
DT   15-JAN-1992 (Rel. 30, Created)
DT   27-APR-1992 (Rel. 31, Last updated, Version 2)
XX
DE   African swine fever virus DNA for ubiquitin conjugating enzyme
XX
KW   ubiquitin; ubiquitin-conjugating enzyme.
XX
OS   African swine fever virus
OC   Viruses; dsDNA viruses, no RNA stage; Asfarviridae;
OC   African swine fever-like viruses.
XX
RN   [1]
RX   MEDLINE; 92155177.
RA   Hingamp P.M., Arnold J.E., Mayer R.J., Dixon L.K.;
RT   "A Ubiquitin Conjugating Enzyme Encoded by African Swine Fever Virus";
RL   EMBO J. 11:361-366(1992).
XX
RN   [2]
RP   1-979
RA   Dixon L.K.;
RT   ;
RL   Submitted (04-OCT-1991) to the EMBL/GenBank/DDBJ databases.
RL   L.K. Dixon, Institute of Animal Health, Pirbright Laboratory, Ash Road,
RL   Pirbright, Woking, Surrey GU24 0NF, UK
XX
DR   SWISS-PROT; P25869; UBC_ASFM2.
XX
FH   Key          Location/Qualifiers
FH
FT   source          1..979
FT                   /db_xref="taxon:10497"
FT                   /organism="African swine fever virus"
FT                   /cell_line="pig blood"
FT                   /clone_lib="lambda EMBL1 Lmw"
FT                   /clone="Lmw22"
FT                   /isolate="Malawi LIL20/1"
FT   5'UTR          1..155
FT   CDS            156..797
FT                   /db_xref="SWISS-PROT:P25869"
FT                   /gene="ASFVUBC"
FT                   /product="ubiquitin conjugating enzyme"
FT                   /protein_id="CAA44305.1"
FT                   /translation="MVSSFLLAEYKNLIVNPSEHFKISVNEDNLTEWDVILKGGPPDTLY
FT                   EGGLFKAKIVFPPKYPYEPRLTFTSEMWHPNISDGKLCISILHGDNAEEQGMTWSPA
FT                   QKIDTVLLSVISLLNEPNPDSPANVDAAKSYRKYLYKEDLESYPMEVKKTVKSLDECS
FT                   AEDIEYFKNVPVNVLPVPSDDYEDEEMEDGTYILTYDDEDEEDEEMDDE"
FT   misc_feature   408..410
FT                   /note="active site cysteine"
FT   3'UTR          798..979
XX
SQ   Sequence 979 BP; 329 A; 151 C; 182 G; 317 T; 0 other;
      gggttttata cggatctccc ttctcgtttg ataattatgc cattaaaggt tttaccagtt          60
      cataaattta gtaaaaatga accccataaa aaacaaaaga ggttcatcta cttttaagga          120
      [...]
      aaactttagt gattagattt tagtgcttag attttaatgt ttaaacttta gtgattagat          960
      tttagtgctt agattttaa          979
//
```

Fiche GENBANK

```
LOCUS      ASFVUCE          979 bp      DNA          VRL          27-APR-1992
DEFINITION African swine fever virus DNA for ubiquitin conjugating enzyme.
ACCESSION  X62440
VERSION    X62440.1  GI:58649
KEYWORDS   ubiquitin; ubiquitin-conjugating enzyme.
SOURCE     African swine fever virus.
  ORGANISM African swine fever virus
            Viruses; dsDNA viruses, no RNA stage; African swine fever-like
            viruses.
REFERENCE  1 (bases 1 to 979)
AUTHORS    Hingamp,P.M., Arnold,J.E., Mayer,R.J. and Dixon,L.K.
TITLE      A ubiquitin conjugating enzyme encoded by African swine fever virus
JOURNAL    EMBO J. 11 (1), 361-366 (1992)
MEDLINE    92155177
REFERENCE  2 (bases 1 to 979)
AUTHORS    Dixon,L.K.
TITLE      Direct Submission
JOURNAL    Submitted (04-OCT-1991) L.K. Dixon, Institute of Animal Health,
            Pirbright Laboratory, Ash Road, Pirbright, Woking, Surrey GU24 0NF,
            UK
FEATURES   Location/Qualifiers
  source   1..979
            /organism="African swine fever virus"
            /isolate="Malawi LIL20/1"
            /db_xref="taxon:10497"
            /clone="Lmw22"
            /cell_line="pig blood"
            /clone_lib="lambda EMBL1 Lmw"
  5'UTR    1..155
  gene     156..797
            /gene="ASFVUBC"
  CDS      156..797
            /gene="ASFVUBC"
            /codon_start=1
            /product="ubiquitin conjugating enzyme"
            /protein_id="CAA44305.1"
            /db_xref="GI:58650"
            /db_xref="SWISS-PROT:P25869"
            /translation="MVSSFLLAEYKNLIVNPSEHFKISVNEEDNLTEWDVILKGPPDTL
            YEGGLFKAKIVFPPKYPYEPRLTFTSEMWHPNISYDGLKLCISILHGDNAEEQGMTWS
            PAQKIDTVLLSVISLLNEPNPDSPANVDAAKSYRKYLYKEDLESYPMEVKKTVKSLD
            ECSAEDIEYFKNVPVNVLPVPSDDYEDEEMEDGTIILTYDDEDEEDEEMDDE "
  misc_feature 408..410
            /gene="ASFVUBC"
            /note="active site"
  3'UTR    798..979
BASE COUNT 329 a    151 c    182 g    317 t
ORIGIN
  1  gggttttata  cggatctccc  ttctcgtttg  ataattatgc  cattaaaggt  tttaccagtt
 61  cataaattta  gtaaaaatga  accccataaa  aaacaaaaga  ggttcatcta  cttttaagga
      [...]
 901 aaactttagt  gattagattt  tagtgcttag  attttaatgt  ttaaacttta  gtgattagat
 961 tttagtgctt  agattttaa

//
```

Fiche SWI SSPROT

```
ID   UBC_ASFM2          STANDARD;          PRT;    213 AA.
AC   P25869;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   UBIQUITIN-CONJUGATING ENZYME E2-21 KDA (EC 6.3.2.19)
DE   (UBIQUITIN-PROTEIN LIGASE) (UBIQUITIN CARRIER PROTEIN).
GN   ASFV-UBC.
OS   African swine fever virus (isolate Malawi Lil 20/1) (ASFV).
OC   Viruses; dsDNA viruses, no RNA stage; Asfarviridae;
OC   African swine fever-like viruses.
OX   NCBI_TaxID=10500;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92155177; PubMed=1310934;
RA   Hingamp P.M., Arnold J.E., Mayer R.J., Dixon L.K.;
RT   "A ubiquitin conjugating enzyme encoded by African swine fever
RT   virus.";
RL   EMBO J. 11:361-366(1992).
CC   -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC   OTHER PROTEINS.
CC   -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
CC   PYROPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC   -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC   -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC   UBIQUITIN-THIOLESTER FORMATION.
CC   -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
DR   EMBL; X62440; CAA44305.1; -.
DR   EMBL; X71982; CAA50851.1; -.
DR   PIR; S19158; S19158.
DR   HSSP; Q02159; 2UCZ.
DR   InterPro; IPR000608; -.
DR   Pfam; PF00179; UQ_con; 1.
DR   PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR   PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
KW   Ubiquitin conjugation; Ligase.
FT   BINDING          85      85      UBIQUITIN (BY SIMILARITY).
FT   DOMAIN          183     213     ASP/GLU-RICH (BASIC).
SQ   SEQUENCE        213 AA;  24468 MW;  F9671BC7385D6DCE CRC64;
      MVSSFLLAEY KNLIVNPSEH FKISVNEDNL TEWDVILKGP PDTLYEGGLF KAKIVFPPKY
      PYEPPRLTFT SEMWHPNIYS DGKLCISILH GDNAEEQGMT WSPAQKIDTV LLSVISLLNE
      PNPDSPANVD AAKSYRKYLY KEDLESYPME VKKTVKKS LD ECSAEDIEYF KNPVNVLPV
      PSDDYEDEEM EDGTYILTYD DEDEEEDEEM DDE
```

//

Abstracts PubMed

The screenshot shows the Entrez-PubMed website in a Netscape browser window. The browser's address bar contains the URL: `http://www.ncbi.nlm.nih.gov/80/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=10812478&dopt=Abstract`. The page header features the NCBI logo and the text "National Library of Medicine PubMed". Below the header is a navigation menu with tabs for "PubMed", "Nucleotide", "Protein", "Genome", "Structure", "PopSet", "Taxonomy", and "OMIM". The "PubMed" tab is selected, and a search bar is visible with the text "Search PubMed for".

On the left side, there is a sidebar with various links under "About Entrez", "Entrez PubMed", "PubMed Services", "Related Resources", and "Privacy Policy".

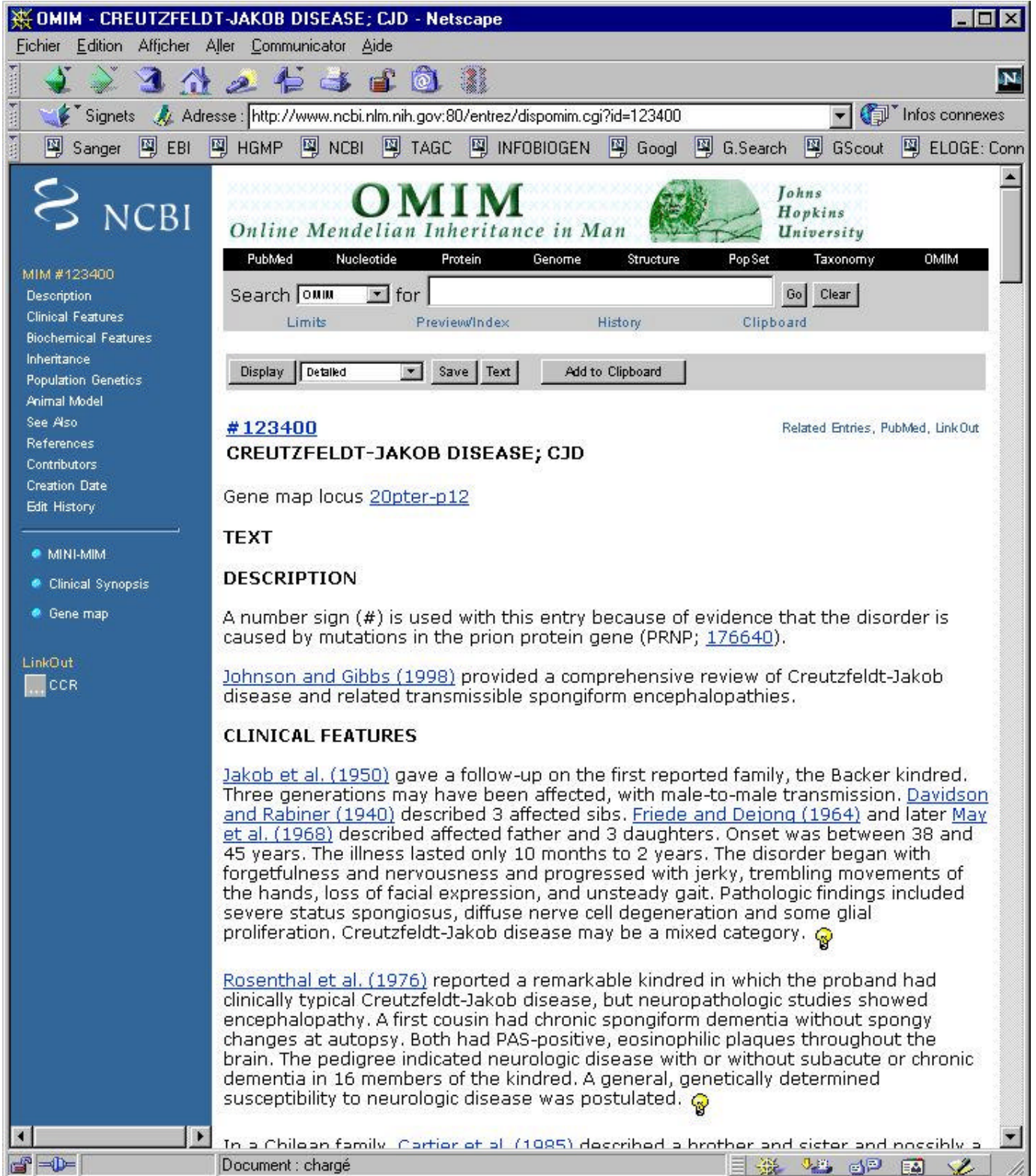
The main content area displays a search result for the article "A brief history of genome research and bioinformatics in France." by Danchin A. The article is listed as "1: *Bioinformatics* 2000 Jan;16(1):65-75". A button labeled "Full text article at bioinformatics.eupjournals.org" is provided. The abstract text reads: "Regulation of Gene Expression, Institut Pasteur, Paris, France. adanchin@pasteur.fr. The development of in silico genomics has progressed slowly in France for a number of political reasons. Two administrative organizations, the Groupement de Recherche sur les Genomes (GREG) and the Groupement de Recherche 1029 (GDR 1029) of the Centre National de la Recherche Scientifique (CNRS) have been established. These organizations have created the dynamics that hopefully will place France (which coordinated consortia that completed several of the first large microbial genomes) among the developed nations that support Large-Scale Biology." The publication type is listed as "Historical article" and the PMID is 10812478.

At the bottom of the page, there are controls for displaying the abstract, saving it, and adding it to a clipboard. The status bar at the very bottom indicates "Document : chargé".

Journaux 'Online'

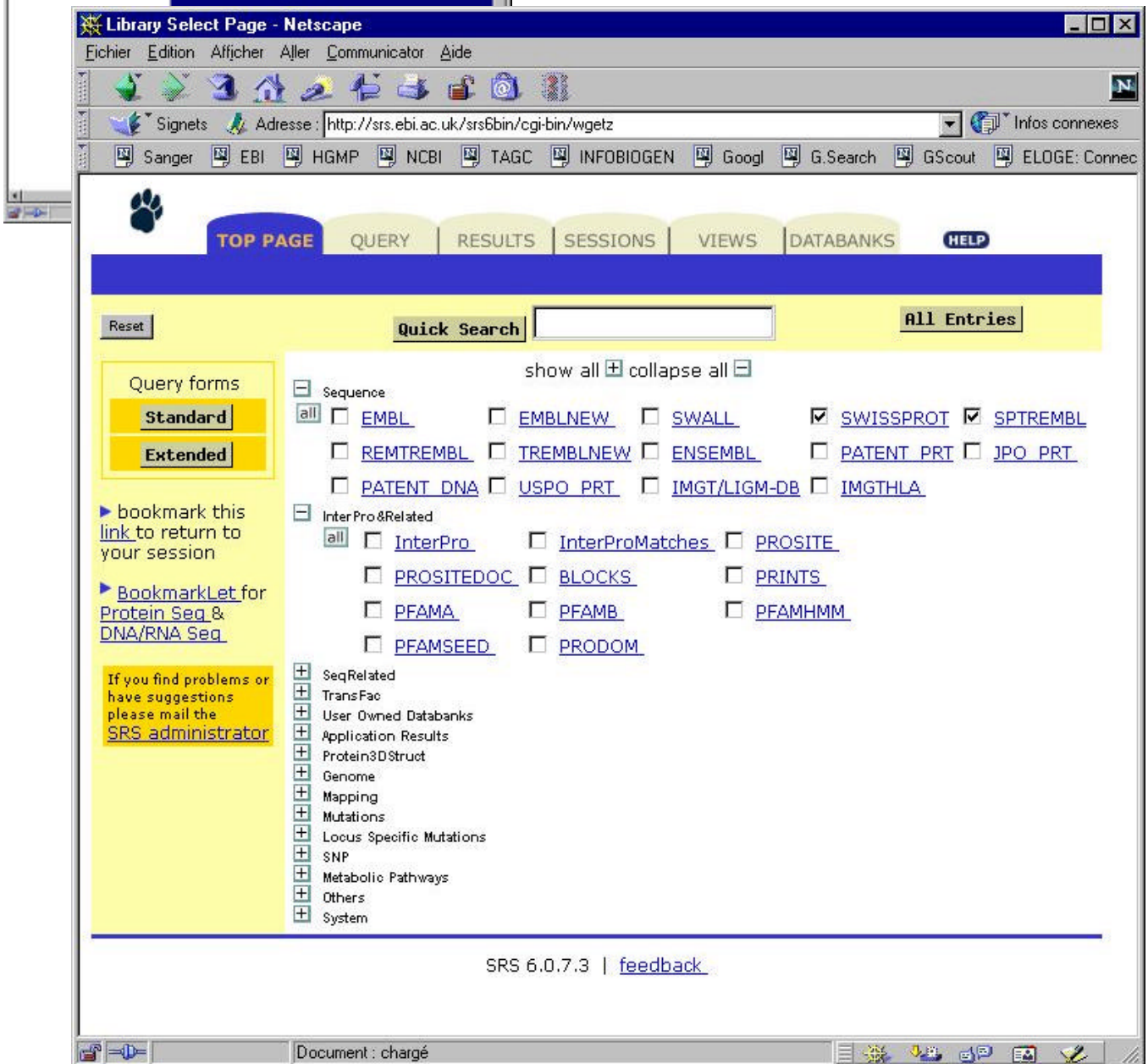


OMIM



The screenshot shows a Netscape browser window displaying the OMIM (Online Mendelian Inheritance in Man) website. The browser title is "OMIM - CREUTZFELDT-JAKOB DISEASE; CJD - Netscape". The address bar shows the URL: <http://www.ncbi.nlm.nih.gov:80/entrez/dispomim.cgi?id=123400>. The page content includes the NCBI logo, the OMIM logo, and the Johns Hopkins University logo. A navigation bar at the top lists various databases: Pubmed, Nucleotide, Protein, Genome, Structure, Pop Set, Taxonomy, and OMIM. A search bar contains the text "OMIM" and "for". Below the search bar are buttons for "Limits", "Preview/Index", "History", and "Clipboard". A "Display" dropdown menu is set to "Detailed", with "Save" and "Text" buttons. The main content area displays the entry for "#123400 CREUTZFELDT-JAKOB DISEASE; CJD". It includes a "Gene map locus" link to [20pter-p12](#) and a "TEXT" section. The "DESCRIPTION" section states: "A number sign (#) is used with this entry because of evidence that the disorder is caused by mutations in the prion protein gene (PRNP; [176640](#)). Johnson and Gibbs (1998) provided a comprehensive review of Creutzfeldt-Jakob disease and related transmissible spongiform encephalopathies." The "CLINICAL FEATURES" section describes the disorder based on Jakob et al. (1950), Davidson and Rabiner (1940), Friede and Dejong (1964), and May et al. (1968). It mentions symptoms like forgetfulness, nervousness, jerky movements, and pathologic findings like status spongiosus. It also references Rosenthal et al. (1976) and Cartier et al. (1985). The browser's status bar at the bottom shows "Document : chargé".

Recherche par l'annotation



Recherche par l'annotation

The image displays two screenshots of the SRS (Swiss-Prot/TrEMBL) database query interface. The top window, titled "Query Form - Netscape", shows the search criteria: Description: ubiquitin, Organism: Homo sapiens, and GeneName: Ubc. The bottom window, titled "Query Result - Netscape", shows the results of the query. The query string is: `query "[!libs=(swissprot sptrembl)]-Description: ubiquitin*" & [!libs-Organism: Homo* & sapiens* | Homo sapiens*] & [!libs-GeneName: Ubc*]" found 10 entries`. The results table lists the following entries:

SWISSPROT SPTREMBL	Accession	Description	Seq.Length
SWISSPROT:UBC1_HUMAN	P50550	UBIQUITIN-CONJUGATING ENZYME E2-18 KDA (EC 6.3.2.19) (UBIQUITIN-PROTEIN LIGASE) (UBIQUITIN CARRIER PROTEIN) (P18).	158
SWISSPROT:UB5A_HUMAN	P51668	UBIQUITIN-CONJUGATING ENZYME E2-17 KDA 1 (EC 6.3.2.19) (UBIQUITIN-PROTEIN LIGASE) (UBIQUITIN CARRIER PROTEIN) (E2(17)KB 1).	147
SWISSPROT:UB5B_HUMAN	P51669	UBIQUITIN-CONJUGATING ENZYME E2-17 KDA 2 (EC 6.3.2.19) (UBIQUITIN-PROTEIN LIGASE) (UBIQUITIN CARRIER PROTEIN) (E2(17)KB 2).	147
SWISSPROT:UB5C_HUMAN	P47986	UBIQUITIN-CONJUGATING ENZYME E2-17 KDA 3 (EC 6.3.2.19) (UBIQUITIN-PROTEIN LIGASE) (UBIQUITIN CARRIER PROTEIN) (E2(17)KB 3).	147
SWISSPROT:UBC6_HUMAN	P51965	UBIQUITIN-CONJUGATING ENZYME E2-21 KDA UBCH6 (EC 6.3.2.19) (UBIQUITIN-PROTEIN LIGASE) (UBIQUITIN CARRIER PROTEIN).	193
SWISSPROT:UBC7_HUMAN		UBIQUITIN-CONJUGATING	