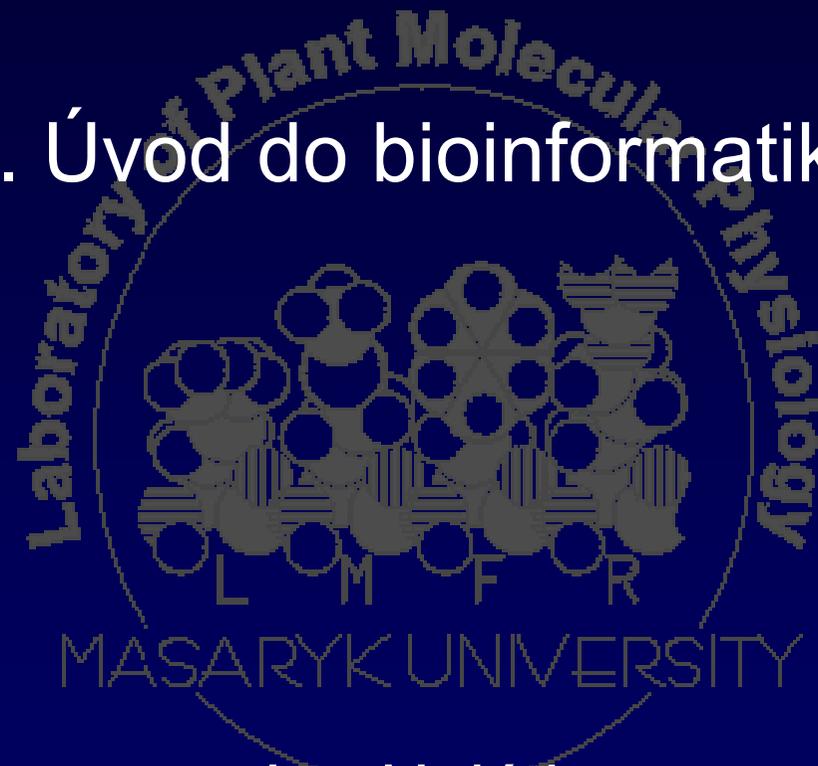


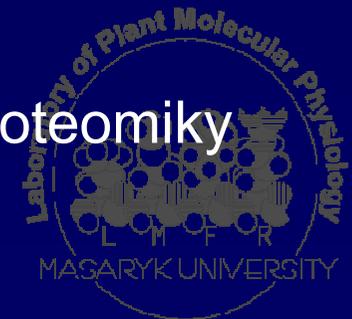
# Základy genomiky

## I. Úvod do bioinformatiky



Jan Hejátko

Masarykova univerzita, Laboratoř funkční genomiky a proteomiky  
Laboratoř molekulární fyziologie rostlin



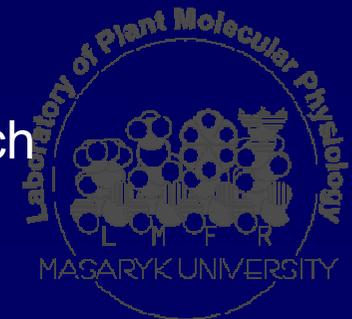
# Základy genomiky I.

- Zdrojová literatura ke kapitole I:
  - Plant Functional Genomics, ed. Erich Grotewold, 2003, Humana Press, Totowa, New Jersey



# Základy genomiky I.

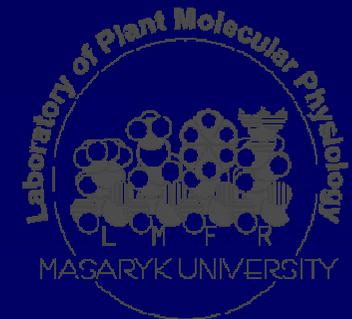
- Role BIOINFORMATIKY v současném pojetí FUNKČNÍ GENOMIKY
- Databáze
  - Spektrum „on-line“ zdrojů
  - PRIMÁRNÍ, SEKUNDÁRNÍ a STRUKTURÁLNÍ databáze
  - GENOMOVÉ zdroje
- Analytické nástroje
  - Vyhledávání homologií
  - Vyhledávání sekvenčních motivů, otevřených čtecích rámců, restrikčních míst....
  - Další www genomové nástroje



# Základy genomiky I.

## GENOMIKA-co to je?

- V širším pojetí-zkoumá **STRUKTURU** a **FUNKCI** genomů
  - Předpokladem je znalost genomu (sekvencí)-práce s databázemi
- V užším pojetí zkoumá **FUNKCI** jednotlivých genů - **FUNKČNÍ GENOMIKA**
  - používá zejména přístupy **REVERZNÍ GENETIKY**



# GENOMIKA-co to je?

## role BIOINFORMATIKY ve FUNKČNÍ GENOMICE

### Přístupy „klasické“ genetiky



3

:

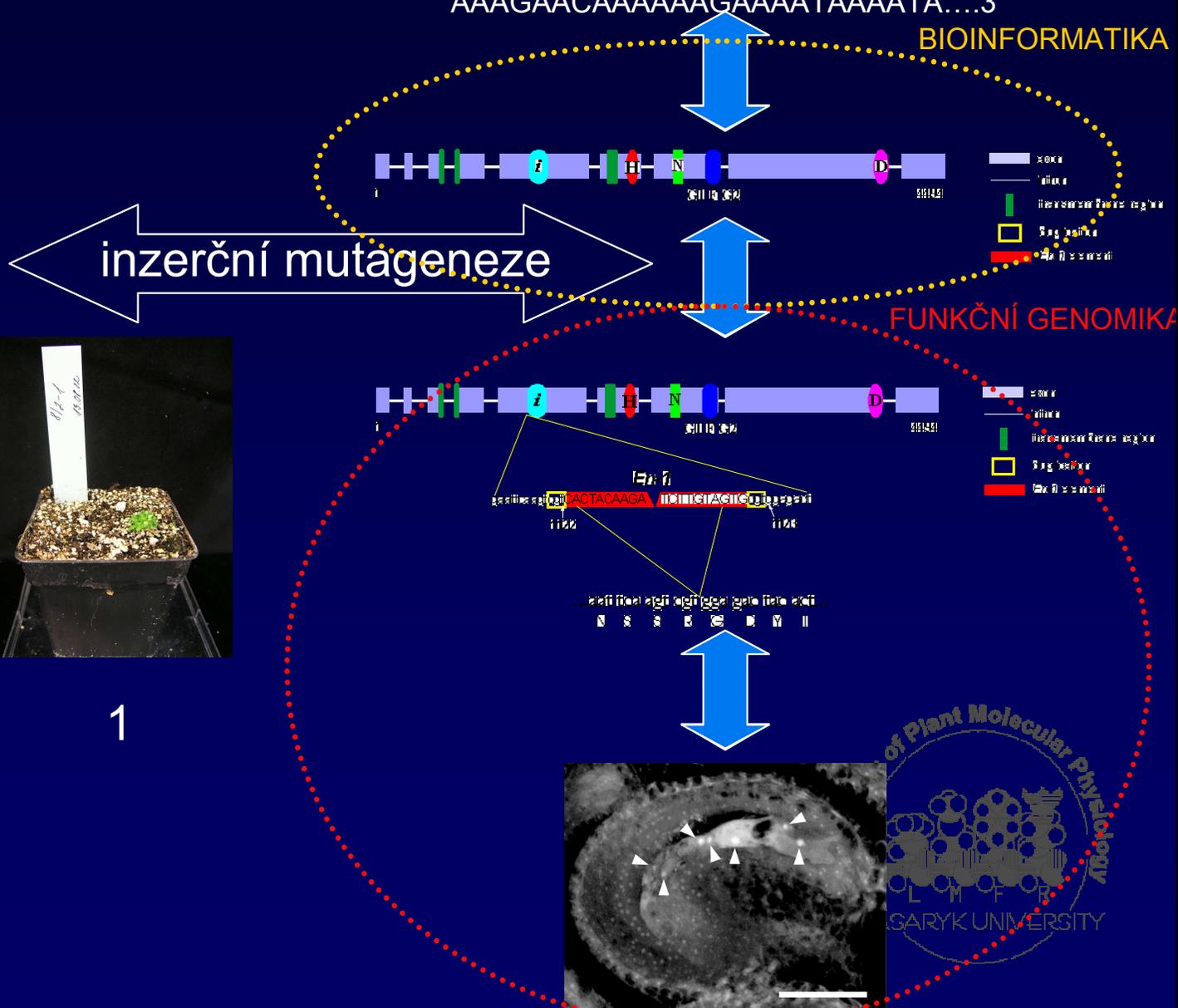
1



?

### „Reverzně genetický“ přístup

5'TTATATATATATATTTAAAAATAAAATA  
AAAGAACAAAAAGAAAATAAAATA...3'



# Základy genomiky I.

- Role BIOINFORMATIKY v současném pojetí FUNKČNÍ GENOMIKY
- Databáze
  - Spektrum „on-line“ zdrojů



# Databáze

## Spektrum on-line zdrojů

### *EMBNet National Nodes*

Vienna Biocenter	Austria	<a href="http://www.at.embnnet.org/">http://www.at.embnnet.org/</a>
BEN	Belgium	<a href="http://www.be.embnnet.org/">http://www.be.embnnet.org/</a>
BioBase	Denmark	<a href="http://biobase.dk/">http://biobase.dk/</a>
CSC	Finland	<a href="http://www.fi.embnnet.org/">http://www.fi.embnnet.org/</a>
INFOBIOGEN	France	<a href="http://www.infobiogen.fr/">http://www.infobiogen.fr/</a>
GENIUSnet	Germany	<a href="http://genome.dkfz-heidelberg.de/biounit/">http://genome.dkfz-heidelberg.de/biounit/</a>
IMBB	Greece	<a href="http://www.imbb.forth.gr/">http://www.imbb.forth.gr/</a>
HEN	Hungary	<a href="http://www.hu.embnnet.org/">http://www.hu.embnnet.org/</a>
INCBI	Ireland	<a href="http://acer.gen.tcd.ie/">http://acer.gen.tcd.ie/</a>
INN	Israel	<a href="http://dapsas.weizmann.ac.il/bcd/inn.html">http://dapsas.weizmann.ac.il/bcd/inn.html</a>
IEN-ADR	Italy	<a href="http://bio-www.ba.cnr.it:8000/BioWWW/Bio-WWW.htm">http://bio-www.ba.cnr.it:8000/BioWWW/Bio-WWW.htm</a>
CAOS/CAMM	Netherlands	<a href="http://www.caos.kun.nl/">http://www.caos.kun.nl/</a>
Bio	Norway	<a href="http://www.no.embnnet.org/">http://www.no.embnnet.org/</a>
IBB	Poland	<a href="http://www.ibb.waw.pl/">http://www.ibb.waw.pl/</a>
IGC	Portugal	<a href="http://www.igc.gulbenkian.pt/">http://www.igc.gulbenkian.pt/</a>
GeneBee	Russia	<a href="http://www.genebee.msu.su/">http://www.genebee.msu.su/</a>
CNB-CSIC	Spain	<a href="http://www.es.embnnet.org/">http://www.es.embnnet.org/</a>
BMC	Sweden	<a href="http://www.embnnet.se/">http://www.embnnet.se/</a>
SIB	Switzerland	<a href="http://www.ch.embnnet.org/">http://www.ch.embnnet.org/</a>
SEQNET	UK	<a href="http://www.seqnet.dl.ac.uk/">http://www.seqnet.dl.ac.uk/</a>

### *EMBNet Specialist Nodes*

MIPS	Germany	<a href="http://www.mips.biochem.mpg.de/">http://www.mips.biochem.mpg.de/</a>
ICGEB	Italy	<a href="http://www.icgeb.trieste.it/">http://www.icgeb.trieste.it/</a>
Pharmacia Upjohn	Sweden	<a href="http://www.pnu.com/">http://www.pnu.com/</a>
F.Hoffmann-La Roche	Switzerland	<a href="http://www.roche.com/">http://www.roche.com/</a>
EBT	UK	<a href="http://www.ebi.ac.uk/">http://www.ebi.ac.uk/</a>
HGMP-RC	UK	<a href="http://www.hgmp.mrc.ac.uk/">http://www.hgmp.mrc.ac.uk/</a>
Sanger	UK	<a href="http://www.sanger.ac.uk/">http://www.sanger.ac.uk/</a>
UMBER	UK	<a href="http://www.bioinf.man.ac.uk/dbbrowser">http://www.bioinf.man.ac.uk/dbbrowser</a>

### *EMBNet Associate Nodes*

IBBM	Argentina	<a href="http://sol.biol.unlp.edu.ar/embnnet">http://sol.biol.unlp.edu.ar/embnnet</a>
ANGIS	Australia	<a href="http://www.angis.su.oz.au/">http://www.angis.su.oz.au/</a>
CBI	China	<a href="http://www.cbi.pku.edu.cn/">http://www.cbi.pku.edu.cn/</a>
CIGB	Cuba	<a href="http://bio.cigb.edu.cu/">http://bio.cigb.edu.cu/</a>
CDFD	India	<a href="http://salarjung.embnnet.org.in/">http://salarjung.embnnet.org.in/</a>
SANBI	South Africa	<a href="http://www.sanbi.ac.za">http://www.sanbi.ac.za</a>

### *USA Information Providers*

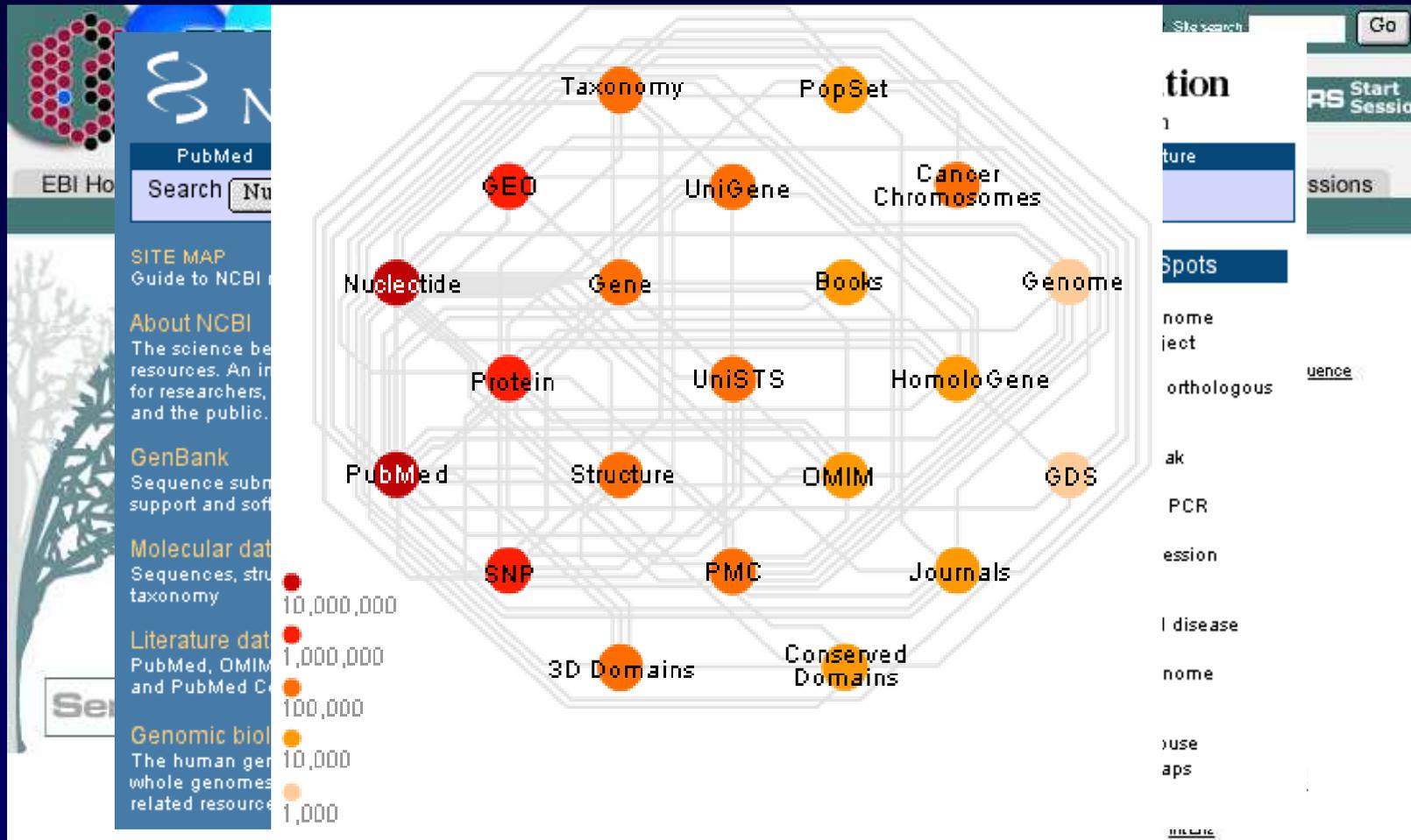
NCBI	USA	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>
NLM	USA	<a href="http://www.nlm.nih.gov/">http://www.nlm.nih.gov/</a>
NIH	USA	<a href="http://www.nih.gov/">http://www.nih.gov/</a>



# Databáze

## Spektrum on-line zdrojů

- EBI <http://www.ebi.ac.uk/services>
- NCBI <http://www.ncbi.nlm.nih.gov/>



# Základy genomiky I.

- Role BIOINFORMATIKY v současném pojetí FUNKČNÍ GENOMIKY
- Databáze
  - Spektrum „on-line“ zdrojů
  - PRIMÁRNÍ, SEKUNDÁRNÍ a STRUKTURÁLNÍ databáze



# Primární databáze

- zahrnují soubory primárních dat – sekvencí DNA a proteinů
  - DNA sekvence:
    - EMBL, <http://www.ebi.ac.uk/embl/>
    - GenBank, <http://www.ncbi.nih.gov/Genbank/GenbankSearch.html>
    - DDBJ, <http://www.ddbj.nig.ac.jp>
  - Proteinové sekvence:
    - PIR, <http://pir.georgetown.edu/>
    - MIPS, <http://www.mips.biochem.mpg.de>
    - SWISS-PROT, <http://www.expasy.org/sprot/>



# Primární databáze

- GenBank (NCBI) <http://www.ncbi.nlm.nih.gov/>

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

PubMed Entrez BLAST OMIM Books TaxBrowser Structure

Search  for

**SITE MAP**  
Guide to NCBI resources

**About NCBI**  
The science behind our resources. An introduction for researchers, educators and the public.

**GenBank**  
Sequence submission support and software

**Molecular databases**  
Sequences, structures and taxonomy

**Literature databases**  
PubMed, OMIM, Books and PubMed Central

**Genomic biology**  
The human genome, whole genomes and related resources

**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. [More...](#)

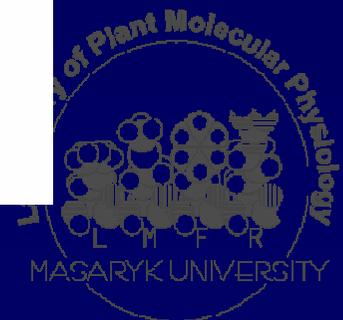
**Hot Spots**

- ▶ Cancer genome anatomy project
- ▶ Clusters of orthologous groups
- ▶ Coffee Break
- ▶ Electronic PCR
- ▶ Gene expression omnibus
- ▶ Genes and disease
- ▶ Human genome resources
- ▶ Human/mouse homology maps
- ▶ LocusLink

**Mouse Genome**  
*Resources: explore tools for manipulating the mouse genome.*

**Try these:** Map Viewer Sequencing Progress Human-Mouse Homology

**BLink** and get results fast!



# Primární databáze

```
/translation="MNGRYSFTRQDFKTGAKPWSILALIWAAMIFAFMAVASWQDNAT
TQAILLSQLRSINADSASLQRDVLRAHTGTVANVRFPIISRLGALRKNLEDLKQLFRQSH
IVSEBNAQLLRQLEVSLNSADAAVAAPGAQNVRIQDSLASFTRALSLSLPGKASTDQT
LEKPTTELASMMLOPLRQPSPAISFBI SLELERLQKQRGLDEAPVRILAREGPIILELL
POVKDLVNNIQTSDTAEIAENLQRECLEVYSLKNVVERSARIFLGSASVGLCLYITL
VYRLRKKTDWLARRLDYBELIKEIGVCFEGBAATSSAQAALRIIQRFADATCALAL
VDHRRWAVETFGAKHPKPVWDDSVLRRI VSRKADERATVFR IISKKIVHLPLEIP
GLEILLAHKSTDKLI AVCCLGYQSYRPRPCQGEIQLELATACLCHYIDVRRKQTECD
VLARLLEHAQREAVGTLAGGIAHEFNILGSELGHAELAQNVSRTSVTRRYIDYII
SSGDRAMLIIDQILTLERKQERMIKPPSEBELVTEIAPLLRMLPPNIELSFRPDQMQ
SVIEGSEPLELQQVLINICKNASQAMTANGQIDIIISQAFLPVKKILAHGVMPPGDVYL
LSISDNGGGIPEAVLPHIFEPFPTTRARNGGTGLGLASVHGHI SAFPAGYIDVSTVGH
GTRPD IYLPPESSKEPVNPDSPFGRNKAPRGNGBI VALVEPDDLREAYEDKIAALGYE
PVGPRTPNRI RDWISKNEADLVMDQASLPEQSPNSVDLVLTASIIIGGNDLKMT
LSREDVTRDLYLPKPISSRTMAHAILTKEIT"
```

## ORIGIN

```
1 atgaacggaa gatattcacc gacgcggcag gattttaaga caggcgcgaa gccttggtct
61 atattggccc ttatcgttgc tgeaatgatt ttocogttca tggcgggtgc gtcttggcag
121 gacaatgega ctaccccagg aatccctcage caactacgat cgat taacgc cgacagcgcc
181 tcaactgcage gogatgtact ccgcgcctcac acgggcaccgc tggcgaaacta ccgccccatt
241 atctccaggc tgggagctct gcggaagaat ctggaagatt tgaagcaatt atttagacaa
301 tctcatattg taagtgagag caatgctgct caactgctac gccagctaga agtgtctcta
361 aattcggtgc acgcgggcgt cgcgcctctt ggtgcgcaaa atgtacgcct gcaagattcg
421 ctggcccagtt tcaactcgtgc tttgagcagt cttccaggaa aagcctcaac cgatcagact
481 ttagaaaaac caacagaatt ggetagcagt atgctccaat ttcttcggca accaaagcccg
541 gctatttcat togagatcag ccttgaacta gagagggctcc aaaaaacaac cggtcttgat
601 gaagetcccg tgcgcatact tgcacgtgaa ggtcccaata tcttatcctt ttgcccacag
661 gtgaaaagtc tgggtgaacat gattcagacg tctgacaccg cagaaattgc ggagatgctg
721 cagcgcagat gtttggagggt ctatagcttg aaaaatgtag agggagcggag cgcacgctac
781 tttcttgggt ccgcttcagt gggctcttgc cctacatca tcaacttagt ctataggcta
841 cgcacaaaaa ccgattgggt agcgcggcgt ttagattacg aagagctaat caaagagatc
901 ggaagtatgt ttgaaggtag ggcggccacc acgtcgtccg cgcacagctgc acttcgtatt
961 attcagcgcct cctttgatgc cgatacgtgc cgttagctc tagtggacca tgaccgtaga
1021 tgggctgtcg aaacattcgg tgcgaaacac ccaaaacctg tgtgggacga cagcgtgcta
1081 cgcgaaatag tctctcgtac caaagcggac gaacgggcca cggctattccg catcatatcg
1141 tgcacaaaaa togtacattt gectctcga atccaggtc tctcgatact actggctcac
1201 aaatccacag ataaactaat tgcggtttgt tcaactgggt accaaagcta tgcctctoga
1261 ccttgcacaag gogaaaatca gcttcttgaa ctocccaccg cctgectctg tcaactatate
1321 gatgttcggc gtaagcagac cgaatgcgac gttttggcca gacgattgga gcatgcgcaa
1381 cgccttgagg cagttggtag acttgcgggc ggaatagcac atgaatttaa taacattttg
1441 ggcctcaatc togggcaocg agaattagca caaaactcgg tgtctcgaac atctgtcacc
1501 cgaagatata ttgactatat catctctgca ggcgacagag ccatgctcat tatcgatcag
1561 atcttgacgc tgagccgaaa acaggagcgc atgatcaagc catttagtgt ctcagagctt
1621 gtgaccgaaa togtccctct gctacgtatg gctcttcgcg caaacatcga gcttagtttc
1681 agatttgatc aatgcagag cgtgatcgaa ggaagccgcg ttgaactca acaggtacta
1741 attaacatct gcaagaatgc tcccaagcc atgactgca atggtcaaat cgacatctc
1801 atcagccaaag cttttttacc agttaagaaa attctggcgc atggtgttat gccacctggc
1861 gactatgttc tctatctat tagcgacaat ggtggaggca tcccgaggc tgtgttacc
1921 cacatttttg aacctctct tacgacacga gctcgcaacg gtggaacggg tctcggcctt
1981 gcttctgtgc atggtcatat cagcgcgttt ggggttaca togacgttag tcaactgtt
2041 gggcatggga cgcgctttga catctatctc cctcgtctt ctaaggaacc cgtaaatcca
2101 gacagttttt toggccgcaa taaggcaccg cgtggaaacg gggagattgt gccacttgtt
2161 gagcccgatg acctcctcgc ggaaggctat gaagacaaga togcgcctct aggatatgag
2221 ccggctcgggt tctgacctt taatgaaatt cgcgattgga tttcaaaag caatgaaagc
2281 gatctggtea tggctgacaa agcgtctctt cctgaagatc aaagtctcaa tccctggat
2341 ttagtctca agaccgcctc catcatcatt ggcggaaatg atctcaaaat gacctttca
```

# Proteinové sekundární databáze

- databáze funkčních nebo strukturálních *motivů* získaných srovnáním primárních dat (sekvencí)

□ PROSITE, <http://www.expasy.org/cgi-bin/browser/PRINTS/>

The screenshot displays the PROSITE database interface. At the top, there is a navigation bar with links: [ExpASY Home page](#), [Site Map](#), [Search ExpASY](#), [Contact us](#), [Swiss-Prot](#), [PROSITE](#), and [Proteomics tools](#). Below this, it indicates the site is hosted by SIB Switzerland and lists mirror sites for Australia, Bolivia, Canada, China, Korea, Taiwan, and USA.

The main content area shows two search results:

- >[PDOC50109 PS50109 HIS\\_KIN](#) Histidine kinase domain [profile].  
402 - 671 NASHDIRGALAGMKGLIDI CRDGVKPGSDVDTTLNQVNVCAKDLVALLNSVLDMSKIRSG  
KMQLVERDFNLSKLLLEDVIDFYHPVAMKKGVDVVLDPHDgavFKPSNVRGDSGRLLKQILN  
NLVSNVAVKPTVD--GHIAVRAWAQrggnsasvlasypkgvakFvkamFcknkeesatye  
teisnairnnanTMEFVFRVDDTGKGI FMEMRKSVPFNYVQVREtAQSHQGTGLGLGIVQ  
SLVRLMGGGIRITDKAMGekGTCPQFNVLIT
- >[PDOC50110 PS50110 RESPONSE\\_REGULATORY](#) Response regulatory domain [profile].  
987 - 1085 RVLVVDNPFISRKVATGKLLKMGVSeVEQCDSGKRALRLVTRGLtqreeggsvdklpFDY  
IFMDCQMPMDGYRATREIRkvekSYGVRTPIIAVSGHD-----  
-----

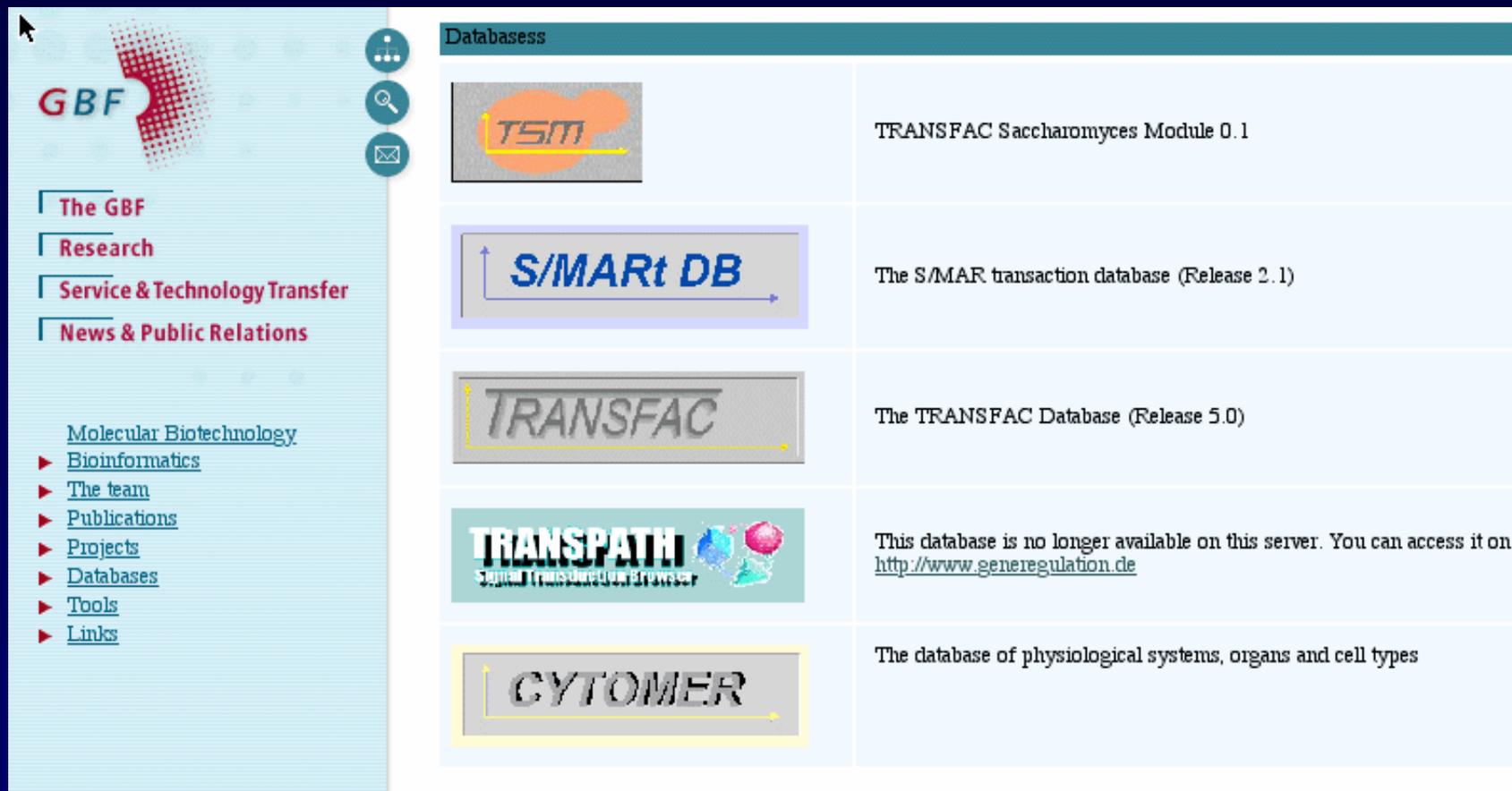
Below the results is a section titled "Graphical summary of hits (java applet)". It features a horizontal bar representing a protein sequence of 100 residues. Two red vertical cursors are positioned at the beginning and end of the sequence. A yellow bar highlights a region labeled "HIS\_KIN". Another yellow bar highlights a region labeled "RESPONSE\_REGULATORY". Below the bar are navigation controls: "Zoom", "Back", "Reset", and a text box containing "1123".

Underneath the graphical summary, it states "98 hits with 12 PROSITE entries".

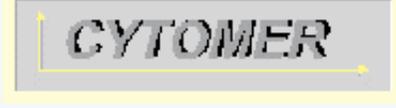
At the bottom of the interface, there is another navigation bar with the same links as the top: [ExpASY Home page](#), [Site Map](#), [Search ExpASY](#), [Contact us](#), [Swiss-Prot](#), [PROSITE](#), and [Proteomics tools](#). Below this bar, there are buttons for "START THE SCAN" and "RESET", and a search input field containing "LQVNVCAKDLVALLNSVLDMSKIRSG" with a "(to test a pattern, see help)" link.

# Sekundární databáze DNA

- TRANSFAC <http://www.gene-regulation.com/>



The screenshot shows the GBF website interface. On the left is a navigation menu with the GBF logo and links for 'The GBF', 'Research', 'Service & Technology Transfer', and 'News & Public Relations'. Below these are links for 'Molecular Biotechnology', 'Bioinformatics', 'The team', 'Publications', 'Projects', 'Databases', 'Tools', and 'Links'. The main content area is titled 'Databases' and contains a table of database entries.

Databases	
	TRANSFAC Saccharomyces Module 0.1
	The S/MAR transaction database (Release 2.1)
	The TRANSFAC Database (Release 5.0)
	This database is no longer available on this server. You can access it on <a href="http://www.generegulation.de">http://www.generegulation.de</a>
	The database of physiological systems, organs and cell types

# Strukturální databáze

- PDB <http://www.rcsb.org/pdb/>

Structure Explorer - 1P5Y Seite 1 von 2

 **Structure Explorer - 1P5Y** 

*Title* **The Structures Of Host Range Controlling Regions Of The Capsids Of Canine and Feline Parvoviruses and Mutants**

*Classification* **Virus/Viral Protein**

*Compound* **Mol. Id: 1; Molecule: Coat Protein Vp2; Chain: A; Fragment: Sequence Database Residues 190-737; Engineered: Yes; Mutation: Yes**

*Exp. Method* **X-ray Diffraction**

 **View Structure**   

[Summary Information](#)

[View Structure](#)

[Download/Display File](#)

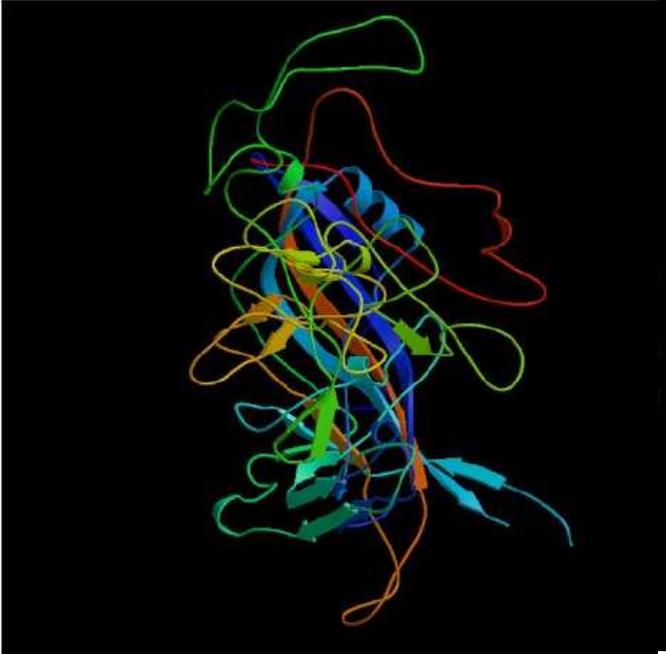
[Structural Neighbors](#)

[Geometry](#)

[Other Sources](#)

[Sequence Details](#)

[SearchLite](#) [SearchFields](#)





<http://www.rcsb.org/pdb/cgi/explore.cgi?job=graphics;pdbId=1P5Y;page=:pid=173561064349344&bio=1&opt=show&size=500> 12/29/2003



# Základy genomiky I.

- Role BIOINFORMATIKY v současném pojetí FUNKČNÍ GENOMIKY
- Databáze
  - Spektrum „on-line“ zdrojů
  - PRIMÁRNÍ, SEKUNDÁRNÍ a STRUKTURÁLNÍ databáze
  - GENOMOVÉ zdroje



# Primární data-genomové zdroje

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

PubMed Entrez BLAST OMIM Books TaxBrowser Structure

Search **Nucleotide** for

**SITE MAP**  
Guide to NCBI resources

**About NCBI**  
The science behind our resources. An introduction for researchers, educators and the public.

**GenBank**  
Sequence submission support and software

**Molecular databases**  
Sequences, structures and taxonomy

**Literature databases**  
PubMed, OMIM, Books and PubMed Central

**Genomic biology**  
The human genome, whole genomes and related resources

**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. [More...](#)

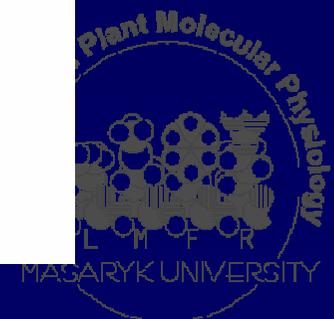
**Hot Spots**

- ▶ Cancer genome anatomy project
- ▶ Clusters of orthologous groups
- ▶ Coffee Break
- ▶ Electronic PCR
- ▶ Gene expression omnibus
- ▶ Genes and disease
- ▶ Human genome resources
- ▶ Human/mouse homology maps
- ▶ LocusLink

**Mouse Genome**  
*Resources: explore tools for manipulating the mouse genome.*

**Try these:** Map Viewer Sequencing Progress Human-Mouse Homology

**BLINK** and get results fast!



# Primární data-genomové zdroje

NCBI Entrez Genome

PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM Help

Search for  on chromosome(s) Find

Show linked entries Help FTP

Entrez Genomes  
MapViewer Home  
Prominent organisms  
FTP SITE  
Related Databases:  
TAIR  
TIGR  
MIPS  
KADS  
Sequencing Projects:  
SPP Consortium  
CSH / WashU  
TIGR  
Kazusa  
ESSA  
Genoscope

**Arabidopsis thaliana genome view** [BLAST search Arabidopsis genome](#)

**I** **II** **III** **IV** **V** **MT** **CHL**

**Lineage:** [Eukaryota](#); [Viridiplantae](#); [Streptophyta](#); [Embryophyta](#); [Tracheophyta](#); [Spermatophyta](#); [Magnoliophyta](#); [eudicotyledons](#); [core eudicots](#); [Rosidae](#); [eurosids II](#); [Brassicales](#); [Brassicaceae](#); [Arabidopsis](#)

*Arabidopsis thaliana* is a small flowering plant that is widely used as a model organism in plant biology. Arabidopsis is a member of the mustard (Brassicaceae) family, which includes cultivated species such as cabbage and radish. Arabidopsis is not of major agronomic significance, but it offers important advantages for basic research in genetics and molecular biology. Its genome has been sequenced by an international collaboration collectively termed the [Arabidopsis Genome Initiative \(AGI\)](#) ([The Arabidopsis Genome Initiative, 2000, Nature, 408:796-815](#)).

This sequence, map, and annotations are the result of a collaboration between [TIGR](#), [MIPS](#), and [TAIR](#). The non-redundant sequence of the chromosomes (pseudomolecules) and their annotations were provided to NCBI by TIGR on behalf of the collaborators.



# Primární data-genomové zdroje

**NCBI** Entrez Genome

Search  Find in This View

PubMed Entrez BLAST OMIM Taxonomy Structure

**Arabidopsis thaliana Map View**

Chromosome: [I](#) | [II](#) | [III](#) | [IV](#) | [V](#)

Master Map: Gene Maps & Options

Total Genes On Chromosome: 7173  
 Region Displayed: 0-30M bp [Download/View Sequence/Evidence](#)  
 Genes Labeled: 20 Total Genes in Region: 7173

Clone	Gene	Accession	Description
At1g05180	YUP8H12.21	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_100396</a> <a href="#">NP_172010</a>	auxin-resistance protein AXR1
At1g08210	T23G18.7	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_100695</a> <a href="#">NP_563808</a>	expressed protein
At1g11250	T28P6.10	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_100997</a> <a href="#">NP_172591</a>	syntaxin-related protein At-SYR1, putative
At1g14670	T5E21.15	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_101334</a> <a href="#">NP_172919</a>	endomembrane protein, putative
At1g17790	F2H15.2	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_101642</a> <a href="#">NP_564037</a>	expressed protein
At1g21050	T22I11.13	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_101958</a> <a href="#">NP_564130</a>	expressed protein
At1g24210	F3I6.14	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_102267</a> <a href="#">NP_564212</a>	expressed protein
At1g28370	F3M18.20	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_102603</a> <a href="#">NP_174159</a>	ethylene-responsive element binding factor, putative
At1g31885	F5M6.28	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_102926</a> <a href="#">NP_174472</a>	major intrinsic protein, putative
At1g35670	F15O4.8	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_103271</a> <a href="#">NP_174807</a>	calcium-dependent protein kinase
At1g48160	F21D18.11	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_103712</a> <a href="#">NP_175250</a>	signal recognition particle 19 kDa protein subunit, putative
At1g51980	F5F19.4	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_104079</a> <a href="#">NP_175610</a>	mitochondrial processing peptidase alpha subunit, putative
At1g55150	T7N22.9	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_104388</a> <a href="#">NP_175911</a>	ethylene-responsive RNA helicase, putative
At1g60140	T13D8.4	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_104705</a> <a href="#">NP_176221</a>	trehalose-6-phosphate synthase, putative
At1g63750	F24D7.6	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_105052</a> <a href="#">NP_176562</a>	putative disease resistance protein
At1g67090	F5A8.1	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_105379</a> <a href="#">NP_176880</a>	ribulose-bisphosphate carboxylase small unit, putative
At1g69870	T17F3.10	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_105655</a> <a href="#">NP_564979</a>	putative peptide transporter
At1g72970	F3N23.17	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_105955</a> <a href="#">NP_565050</a>	expressed protein
At1g75780	T4O12.1	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_106228</a> <a href="#">NP_177706</a>	tubulin beta-1 chain
At1g78700	F9K20.26	<a href="#">TIGR</a> <a href="#">MIPS</a> <a href="#">TAIR</a> <a href="#">NM_106517</a> <a href="#">NP_565187</a>	expressed protein

[Disclaimer](#) | [Write to the Help Desk](#)  
[NCBI](#) | [NLM](#) | [NIH](#)

# Primární data-genomové zdroje

NCBI Nucleotide

Search  for

History has expired due to inactivity.

Show:

1: NC\_003070. Arabidopsis thali...[gi:22330780] [Links](#)

**Reverse Complement Strand**  
[View on plus strand](#) [Protein coding genes](#) [Hide Toolbar](#)

Search for gene    CDS with gene and mRNA   
 Hide sequence

3002869 3' 20M 10M 1 5'

At1g05200 At1g05190 At1g05180 At1g05170

**Legend:**  
 - CDS (pink line)  
 - RNA (blue line)  
 - gene (grey line)  
 - sequence fragment shown (red line)

**Sequence:**

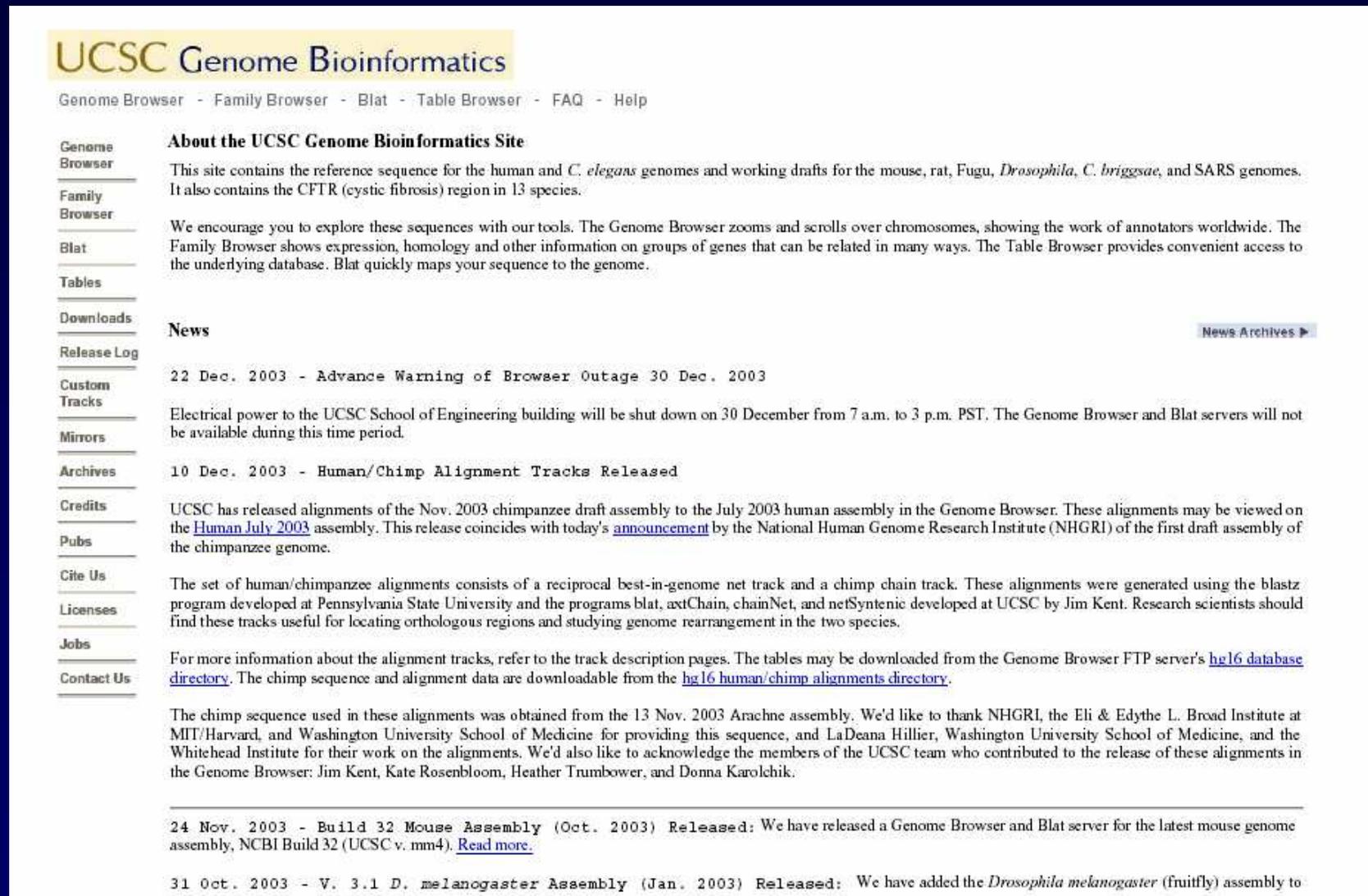
```

1501830 CTTTTTGT TT ATCAGTTCAC CCGAGGCCA AAATCGTCTCT CGCTTGAGCT CCGAAGATCC At1g05180
                                     M mRNA-protein id:
                                     CDS
1501770 AACGAGTAA AAGATCCAGG AGCATGTTG AACAGAGGCC AACAAATGGA CAACCTAAAA At1g05180
Q A V K R S R R H V E E E P T M V E P K mRNA-protein id:
                                     CDS
1501710 CCAAGTACGA TGCTCAGCTC AGGTATACAT ACTCTTTTTC CTAAACTCT ACTTCGAACT At1g05180
T K Y D R Q L R mRNA-protein id:
                                     CDS
1501650 ACTCTTGTCG GGAGGAAAT TGATGAGTA GTAATCGTTG GGTGACTGA TTAGGATTTG At1g05180
                                     I W mRNA-protein id:
                                     CDS
1501590 GGGGGAGSTA GGTCAAGCGG CCTTGGAAGA AGCGAGTATC TGTTTACTCA ATTGTGCCCC At1g05180
G E V G Q A A L E E A S I C L L N C G P mRNA-protein id:
                                     CDS
  
```



# Primární data-genomové zdroje

- Human Genome Browser <http://genome.ucsc.edu/>



**UCSC Genome Bioinformatics**

Genome Browser - Family Browser - Blat - Table Browser - FAQ - Help

**Genome Browser**  
**About the UCSC Genome Bioinformatics Site**

This site contains the reference sequence for the human and *C. elegans* genomes and working drafts for the mouse, rat, Fugu, *Drosophila*, *C. briggsae*, and SARS genomes. It also contains the CFTR (cystic fibrosis) region in 13 species.

**Family Browser**  
We encourage you to explore these sequences with our tools. The Genome Browser zooms and scrolls over chromosomes, showing the work of annotators worldwide. The Family Browser shows expression, homology and other information on groups of genes that can be related in many ways. The Table Browser provides convenient access to the underlying database. Blat quickly maps your sequence to the genome.

**Blat**

**Tables**

**Downloads**

**Release Log** News Archives ▶

**Custom Tracks**  
22 Dec. 2003 - Advance Warning of Browser Outage 30 Dec. 2003

**Mirrors**  
Electrical power to the UCSC School of Engineering building will be shut down on 30 December from 7 a.m. to 3 p.m. PST. The Genome Browser and Blat servers will not be available during this time period.

**Archives**  
10 Dec. 2003 - Human/Chimp Alignment Tracks Released

**Credits**  
UCSC has released alignments of the Nov. 2003 chimpanzee draft assembly to the July 2003 human assembly in the Genome Browser. These alignments may be viewed on the [Human July 2003](#) assembly. This release coincides with today's [announcement](#) by the National Human Genome Research Institute (NHGRI) of the first draft assembly of the chimpanzee genome.

**Pubs**

**Cite Us**  
The set of human/chimpanzee alignments consists of a reciprocal best-in-genome net track and a chimp chain track. These alignments were generated using the blastz program developed at Pennsylvania State University and the programs blat, actChain, chainNet, and netSyntenic developed at UCSC by Jim Kent. Research scientists should find these tracks useful for locating orthologous regions and studying genome rearrangement in the two species.

**Licenses**

**Jobs**

**Contact Us**  
For more information about the alignment tracks, refer to the track description pages. The tables may be downloaded from the Genome Browser FTP server's [hg16 database directory](#). The chimp sequence and alignment data are downloadable from the [hg16 human/chimp alignments directory](#).

The chimp sequence used in these alignments was obtained from the 13 Nov. 2003 Arachne assembly. We'd like to thank NHGRI, the Eli & Edythe L. Broad Institute at MIT/Harvard, and Washington University School of Medicine for providing this sequence, and LaDeana Hillier, Washington University School of Medicine, and the Whitehead Institute for their work on the alignments. We'd also like to acknowledge the members of the UCSC team who contributed to the release of these alignments in the Genome Browser: Jim Kent, Kate Rosenbloom, Heather Trumbower, and Donna Karolchik.

---

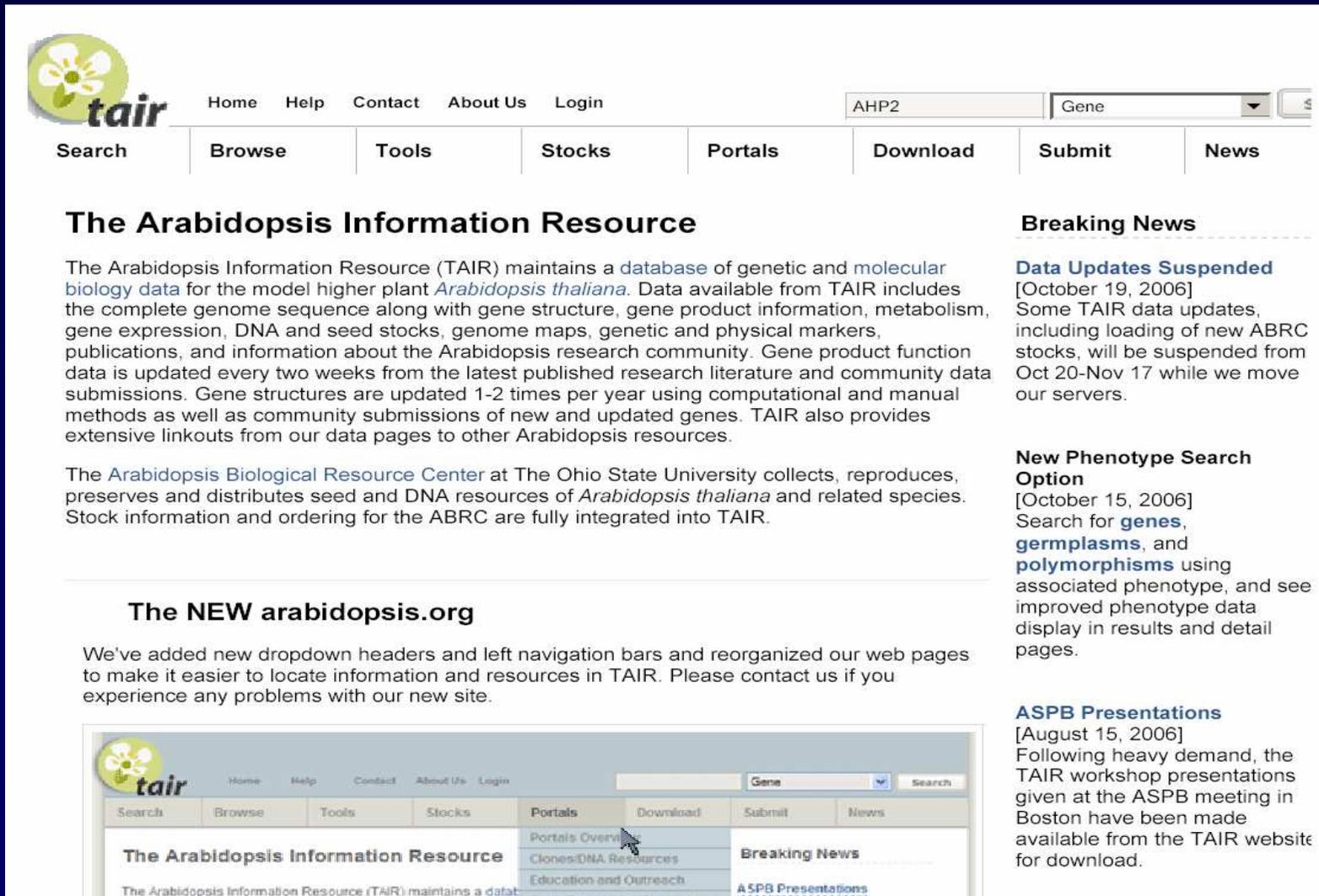
24 Nov. 2003 - Build 32 Mouse Assembly (Oct. 2003) Released: We have released a Genome Browser and Blat server for the latest mouse genome assembly, NCBI Build 32 (UCSC v. mm4). [Read more.](#)

31 Oct. 2003 - V. 3.1 *D. melanogaster* Assembly (Jan. 2003) Released: We have added the *Drosophila melanogaster* (fruitfly) assembly to



# Primární data-genomové zdroje

- TAIR, The Arabidopsis Information Resource, <http://www.arabidopsis.org>



The screenshot shows the TAIR website homepage. At the top left is the TAIR logo, a green flower. To its right are navigation links: Home, Help, Contact, About Us, Login. A search bar contains 'AHP2' and a dropdown menu is set to 'Gene'. Below the navigation is a horizontal menu with tabs: Search, Browse, Tools, Stocks, Portals, Download, Submit, News. The main content area is divided into two columns. The left column has a heading 'The Arabidopsis Information Resource' followed by a paragraph describing the database. Below that is another paragraph about the Ohio State University resource center. The right column has a heading 'Breaking News' followed by two news items: 'Data Updates Suspended' and 'New Phenotype Search Option'. At the bottom of the left column is a section 'The NEW arabidopsis.org' with a paragraph about website changes. At the bottom of the right column is a section 'ASPB Presentations' with a paragraph about workshop presentations. A watermark for 'Molecular Physiology OF OHIO STATE UNIVERSITY' is visible on the right side of the screenshot.

**The Arabidopsis Information Resource**

The Arabidopsis Information Resource (TAIR) maintains a database of genetic and molecular biology data for the model higher plant *Arabidopsis thaliana*. Data available from TAIR includes the complete genome sequence along with gene structure, gene product information, metabolism, gene expression, DNA and seed stocks, genome maps, genetic and physical markers, publications, and information about the Arabidopsis research community. Gene product function data is updated every two weeks from the latest published research literature and community data submissions. Gene structures are updated 1-2 times per year using computational and manual methods as well as community submissions of new and updated genes. TAIR also provides extensive linkouts from our data pages to other Arabidopsis resources.

The Arabidopsis Biological Resource Center at The Ohio State University collects, reproduces, preserves and distributes seed and DNA resources of *Arabidopsis thaliana* and related species. Stock information and ordering for the ABRC are fully integrated into TAIR.

**The NEW arabidopsis.org**

We've added new dropdown headers and left navigation bars and reorganized our web pages to make it easier to locate information and resources in TAIR. Please contact us if you experience any problems with our new site.

**Breaking News**

**Data Updates Suspended**  
[October 19, 2006]  
Some TAIR data updates, including loading of new ABRC stocks, will be suspended from Oct 20-Nov 17 while we move our servers.

**New Phenotype Search Option**  
[October 15, 2006]  
Search for **genes**, **germplasms**, and **polymorphisms** using associated phenotype, and see improved phenotype data display in results and detail pages.

**ASPB Presentations**  
[August 15, 2006]  
Following heavy demand, the TAIR workshop presentations given at the ASPB meeting in Boston have been made available from the TAIR website for download.



# Základy genomiky I.

- Role BIOINFORMATIKY v současném pojetí FUNKČNÍ GENOMIKY
- Databáze
  - Spektrum „on-line“ zdrojů
  - PRIMÁRNÍ, SEKUNDÁRNÍ a STRUKTURÁLNÍ databáze
  - GENOMOVÉ zdroje
- Analytické nástroje
  - Vyhledávání homologií



# Analytické nástroje

- BLAST <http://ncbi.nlm.nih.gov/BLAST/>

NCBI *nucleotide-nucleotide* **BLAST**

Nucleotide Protein Translations Retrieve results for an RID

[Search](#)

```
aaccaaccgc  
acaccatcat cattatcacc atcgttttgg ggcgatggtg tgtggttcca  
ggtattaat  
ataattaatt tattccacat gagatatgat atgatatact atgtatTTTT  
tgTTTTTTTT  
ttatttgtaa acctttaata taacaagaac tacaaaaaat gaaaa
```

[Set subsequence](#) From:  To:

[Choose database](#)

Now: **BLAST!** or

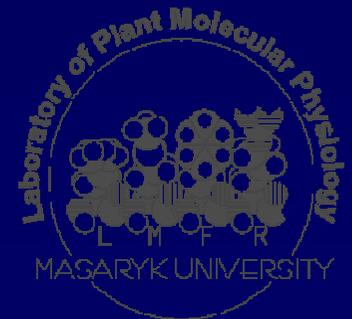






# Základy genomiky I.

- Role BIOINFORMATIKY v současném pojetí FUNKČNÍ GENOMIKY
- Databáze
  - Spektrum „on-line“ zdrojů
  - PRIMÁRNÍ, SEKUNDÁRNÍ a STRUKTURÁLNÍ databáze
  - GENOMOVÉ zdroje
- Analytické nástroje
  - Vyhledávání homologií
  - Vyhledávání sekvenčních motivů, otevřených čtecích rámců, restrikčních míst....



# Analytické nástroje

□ <http://workbench.sdsc.edu/>

**Biology WorkBench**  
click here to toggle between menus and buttons  
**WE Moved!** <http://workbench.sdsc.edu/>  
Version 3.2

Session Tools Protein Tools **Nucleic Tools** Alignment Tools Structure Tools (Alpha)

beta-glucosidase

GBPLN:804655 **Hordeum vulgare L. beta-glucosidase (BGQ60) gene, complete cds.**  
 GBPLN:170248 **Nicotiana tabacum glucan beta-1,3-glucosidase gene, complete cds.**

Select All Deselect All Ndjinn BATCH Add Edit Delete Copy View Download ViewRecords  
BL2SEQ BL2SEQX BLASTN BLASTX TBLASTX FASTA FASTX FASTY SSEARCH CLUSTALW  
CLUSTALWPROF ALIGN LALIGN LFASTA PATTERNMATCHDB PATTERNMATCH TACG PRIMER3  
NASTATS BESTSCOR PFSCAN PRIMERCHECK PRIMERTM SIXFRAME REVCOMP RANDSEQ

Copyright (C) 1999, Board of Trustees of the University of Illinois.



# Analytické nástroje

**View**  
View Nucleic Sequence(s)

---

Format  Case

[Download/view all sequences in text format](#)

[\[NEXT\]](#) [\[BOTTOM\]](#)

**Nicotiana tabacum glucan beta-1,3-glucosidase gene, complete cds.**  
**GBPLN:170248, 4699 bp**

>170248  
GAGCTCCCTTGGGGGGCAAGGGGCAAACTTTTGCTAAATGGAAAAATATTATACCAAGTGTGTTGTAATA  
GTTACTCAATTTGAATTAACAAAGGGGCAAATTTGACTATTTTGGCCCTTATATCTTTTGGTCAAAAAAC  
ATAAAATATCCCATCCGAAATTCCAAATGGTCCATTATCGGCAAGTAGCTTTCTTTAATTATAGTTAGTT  
GACAAAACACTATCAAGATATCATTATTATAATAAATAACTTCAAAGTCCATCATCTTAGCTGCCTCCTCA  
GTAGAGCCGCCAGTAAAAATAAGACCGATCAAATAAAGCCGCCATTAAAAATAATGAATTTTAGGACTCTC  
GATTGGCACGTAAGTGCCAAAACCTTTCCAATACTTTGTGCAACTTGGGGCTGCTAGGTTCTGAGCTTC  
CAGATATGGGATATTTCTAAGTTTATCTCCTAAITTTACATCTCAACTAATATTAAGAAITTAACAGGTA  
CAGCAAATCATAAAATTTTCTCTAAAGAAGACAAATGAATCCGGTTACTGATTCAATTGGCCTTTTTCAGAG  
TCTGCATGCCATATTCACTAAGGGGTCGTTTGGTAC AAGAAAATAATAATAAATTTCTGGGATAGAATTT  
GAGATTGCATTTATCTTGTGTTTAATTATAAGTATTAGCTAATTTT CAGAATAAAATTTTACACTAAAATAG  
TAAAATCAACTATCACATGTAGAAGGTGGAATGGAATAGCTAATCCCATAGCCACTCACATAGAATATCC  
TTATTTATCTCACTATTTTACCAAATGATCGGTTAGTCTTCATGAGAATCCAGTATCCTCAATAAATGCA  
GTAAGAAGTTAGAAAATTTTCATTAATCAAITCATATAATTTAAAAATATTAGATATGGAGCACTTAAG  
ATACAATAAAAAGATGTACCGTTAATAATAAAAAGATAAGATAGAGTTTTAAATAGGAAAAAAAAAACGGTT  
CGAGACACTCTTATGGAAGGCGTTGTCTTCAAAGTAGATTCTCATTCAATTGCTCTGGTGC AATAGCAAAA  
TGACATCTTACTCTTAAGATACAGCGAGCCACTCTACAATCTTCTATTGTATACTCAAATGAAAGTTTTA  
GAGAACTTTCAAATCTCTCAACTACTTTTTAAGGGAATTCAAAATACGACCAATATTTATTACTTACTTAC  
TTATAGTTAAATGATATGAATTTTTATTTTTAAATTTGAAATGAAAAATTTAAATTACTTGATTTAATATAA



# Analytické nástroje

## Regex pattern:

ctt. {1,32}ctt

0 sequences were searched

1 match was found

## Matches are indicated in blue

>170248

```
GAGCTCCCTTGGGGGGCAAGGGCAAACTTTTTGTCAAATGGAAAAATATTATACCAAGTGTTTTGTAAATA
GTTACTCAATTTGAATTAACAAAGGGGCAAAATTTGACTATTTTTGCCCTTATATCTTTTTGGTCCACAAAAAC
ATAAAATATCCCATCCGAAATTCCAAATGGTCCATTATCGGCAAGTAGCTTCTTTAAATATAGTTAGTT
GACAAACACTATCAAGATATCATTATTATAATAATACTTCAAAGTCCATCATCTTTAGCTGCCTCCTCA
GTAGAGCCGCCAGTAAAAATAAGACCGATCAAATAAAAGCCGCCATTAAAAATAATGAATTTTAGGACTCTC
GATTTGGCAGTAAGTGCCAAAACCTTCCAATACTTTTGCTGCAACTTTGGGGCTGCTAGGTTCTGAGCTTC
CAGATATGGGATATTTCTAAGTTTTATCTCCTAATTTACATCTCAACTAATATTAAAGAAATPAAACAGGTA
CAGCAAATCATAAAATTTTCTCTAAAGAAGACAATGAATCCGGTTACTGATTCATTTGGCCTTTTTAGAG
TCTGCATGCCATATTTACTAAGGGGTCGTTTTGGTACAAGAAATAATAATAATAATTTTGGGATAGAATTT
GAGATTGCATTTATCTTTGTGTTTTAATTATAAGTATPAGCTAATTTTCAGAATAAAATTTTACTACTAAAATAG
TAAATCAACTATCACATGTAGAAGGTGGAATGGAATAGCTAATCCCATAGCCACTCACATAGAATATCC
TTATTTATCTCACTATTTTACCAAATGATCGGTTAGTCTTCATGAGAATCCAGTATCCTCAATAAATGCA
GTAAGAAATPAGAAAAATTTTCATTAAATCAATTCATATAATTTAAAAATATTAGATATGGAGCACTTAAG
ATACAATAAAAGATGTACCGTTAATAATAAAAGATAAGATAGAGTTTTTAAATAGGAAAAAAAAAAAAACGGTT
CGAGACACTCTTATGGAAGGCGTTGCTTTCAAAGTAGATTCTCATTCAATTGCTCTGGTGCATAGCAAAA
TGACATCTTACTCTTAAGATACAGCGAGCCACTCTACAATCTTTCTATTGTATACTCAAATGAAAATTTTA
GAGAACTTTCAAATCTCTCAACTACTTTTTAAGGGAATTCAAAATACGACCAATATTTATTACTTACTTAC
TTATAGTTAAATGATATGAATTTTTATTTTTAAATTTGAATTTGAAAATATTTAAATTTACTTTGATTTAATATAA
ACAAATAGATATCGCTAAGTATTTTACCACAAACATGGAGATACTACAGAAGATTTTTATTATTITGTAACGAT
GATTAAGCAGCTATTCATCTGGTTGTGCAGGATGAAAAGAAAGTAACTAGCTATAATTTCTTTTTGTAAAGT
```



# Analytické nástroje

## Frame 1, 1 stop codon

Nicotiana tabacum glucan beta-1,3-glucosidase gene, complete cds. Tran

>170248 Translated - Frame 1  
ELPFGARAKLFAKWKNIIIPSV

```
1 gagctcccttggggggcaagggcaaaactttttgctaaatggaaaaatattataccaagt 60
  V C N S Y S I * I N K G A N L T I L P L
61 gtttgtaatagttactcaatttgaattaacaaaggggcaaatttgactattttgcctta 120
```

## Frame 2, 1 stop codon

Nicotiana tabacum glucan beta-1,3-glucosidase gene, complete cds. Tran

>170248 Translated - Frame 2  
SSLGGQGQNFLLNGKILYQV

```
2 agctcccttggggggcaagggcaaaactttttgctaaatggaaaaatattataccaagt 61
  F V I V T Q F E L T K G Q I * L F C P
62 tttgtaatagttactcaatttgaattaacaaaggggcaaatttgactattttgcctta 120
```



# Analytické nástroje

## = Linear Map of Sequence:

```

          StyI
          BsaJI
        CviJI
        AluI
          SacI
        EcoICRI
          Bsp1286I
          BsiHKAI
        BanII  BslI
          \ \ \ \ \
1 gagctcccttggggggcaggggcaaaactttttgctaaatggaaaaatattataccaagt 60
  ctcgagggaaccccccgttcccgttttgaaaaacgatttacctttttataaatatggttca
    ^ * ^ * ^ * ^ * ^ * ^ * ^ *
1  E L P W G A R A K L F A K W K N I I P S
2  S S L G G Q G Q N F L L N G K I L Y Q V
3  A P L G G K G K T F C * M E K Y Y T K C
4  L E R P P C P C F K K S F P F I N Y W T
5  S S G Q P A L A F S K A L H F F I I G L
6  L A G K P P L P L V K Q * I S F Y * V L

          Tsp509I          Tsp509I
        MaeIII Tsp509I  MseI          ApoI
          \ \ \ \ \
61 gtttgtaatagttactcaatttgaattaacaaaggggcaaatttgactattttgcctta 120
   caaacattatcaatgagttaaacttaattgtttccccgttttaaactgataaaaacgggaat
     ^ * ^ * ^ * ^ * ^ * ^ * ^ * ^ * ^ *
1  V C N S Y S I * I N K G A N L T I L P L
2  F V I V T Q F E L T K G Q I * L F C P *
3  L * * L L N L N * Q R G K F D Y F A L R
4  N T I T V * N S N V F P C I Q S N Q G *
5  T Q L L * E I Q I L L P A F K V I K G K
6  H K Y Y N S L K F * C L P L N S * K A R

```



# Analytické nástroje

## Selected Sequence(s)

- *Lycopersicon esculentum* beta-1,3-glucanase mRNA, complete cds.,
- Capsicum annuum* clone GC170 beta-1,3-glucanase-like protein gene.,
- Nicotiana tabacum* glucan beta-1,3-glucosidase gene, complete cds.,
- Nicotiana plumbaginifolia* beta-(1,3)-glucanase gene for a vacuolar,
- Hordeum vulgare* L. beta-glucosidase (BGQ60) gene, complete cds.

[Download a PostScript version of the output](#)

```

.....
2560 CTTTGCTTGGTCTCTGCTTGACAACTTCGAGTGGAGACTCGGCTACACTGCCCGTTTCGG 804655

                2650      2660      2670      2680      2690      2700
24 .....AAATAGGT. 170381
1 ..... 11321163
2430 .....GAACAATT. 170248
1743 CAGTCAAATGATTGACAGAACTGCCAAAAACAAGCGAAAAATGGTAAAAAAAAAAAAAATTC 196886
2620 GATCGTCTATGTGGACTTCAATACTGTGAAGAGGTACGGCAAGGAGTCAGGCTTCTGGTT 804655

                2710      2720      2730      2740      2750      2760
32 .....ATTATGTGCTTCTAGGATTACTTGTGGCA.GGACATTGACATAAG 170381
1 ..... 11321163
2438 .....AG.ATAATGATTTAGTTTCTAACAGTAATT.GTAATTCITATTGAGC 170248
1803 AGCATGTTTACAATTGTATGTGAAAGGGCGACTGACTATTTTCTAATTCATAATTGACC 196886
2680 GAAGAACATGCTGTCGGAAAGAAAGAGCAGCTAGGATCGAAACAGCATCTGGGAGGATE 804655

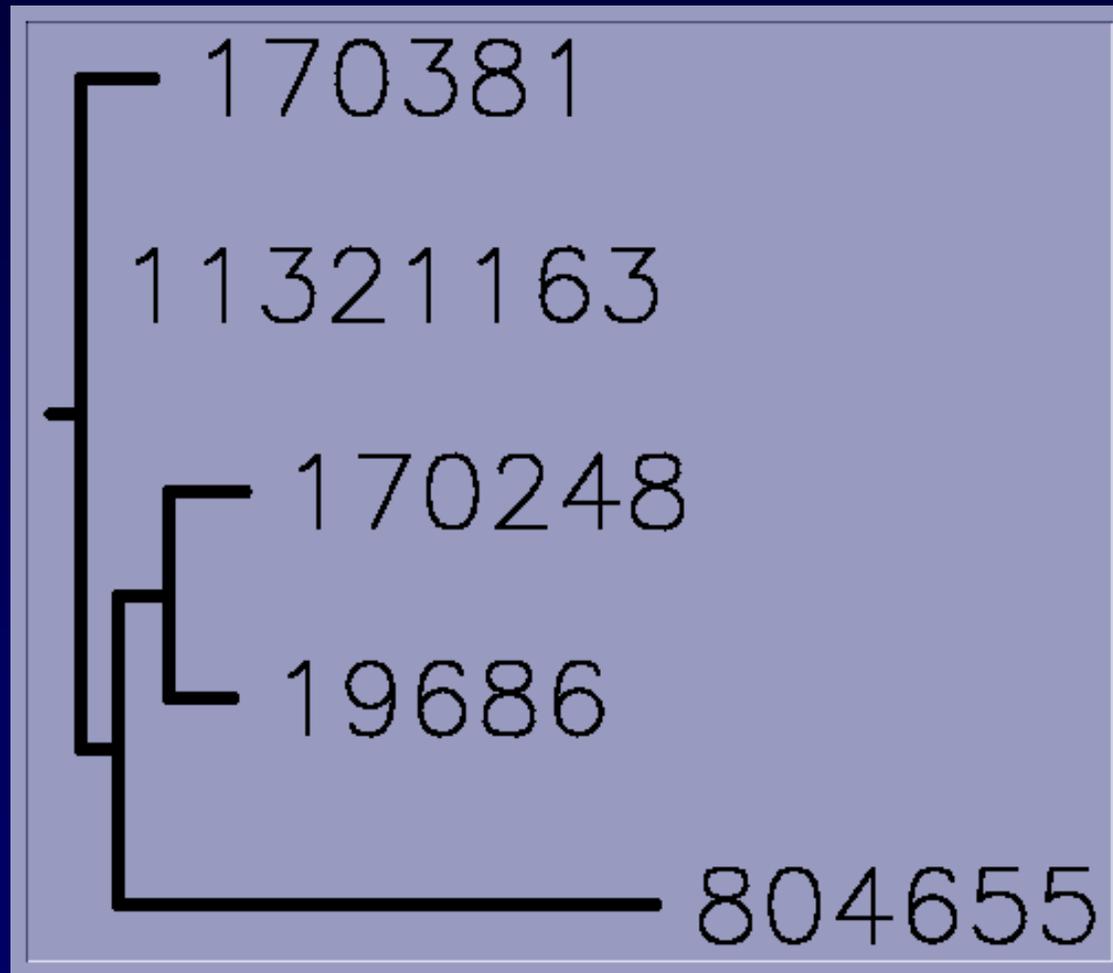
                2770      2780      2790      2800      2810      2820
79 AGAGGCTCAA...ATAGGTGT...TTGTTATGCAATCATGGCAACAAGCTTGGCATGAC 170381
1 T...ATCGGTGT...TTGGTATGCAATCATGGCAACAAGCTTGGCATGAC 11321163
2484 AGGGGCTCAATCAATAGGTGT...TTGCTATGCAATCATGGCAACAAGCTTGGCAATC 170248
1863 AGGGGCTCAATCAATAGGTGT...TTGCTATGCAATCATGGCAACAAGCTTGGCAATC 196886
2740 AGCTGCTTCAAGCTTTCAGAAATAAATTCATAATGTAATGTGTTTTATCATAGAACGTGAG 804655

                2830      2840      2850      2860      2870      2880
132 ATTCTGAAGTTATACAGC..TCTACAAGTCGAGAAACATTCGAAGACTGAGGCTTTATGA 170381
45 ATTGGGAAGTTATACAGC..TCTACAAGTCGAGAAACATTCGAAGACTGAGGCTTTATGA 11321163
2540 ATTGGGAAGTTATACAGC..TCTACAAGTCGAGAAACATTCGAAGACTGAGGCTTTATGA 170248
1919 ATTGGGAAGTTATACAGC..TCTACAAGTCGAGAAACATTCGAAGACTGAGGCTTTATGA 196886
2800 ACTTAGGGGCTGATTCGCTAAAGACGCACTGCGCAATATCCGCAGTCGGCAAAATGGCAG 804655

```



# Analytické nástroje



# Analytické nástroje

- VPCR <http://elanor.sci.muni.cz/LMFR/vpcr/>

**SEARCH**  [ABOUT](#) [DOWNLOAD](#) [LINKS](#)

VPCR 2.0 (WWW interface) - Please, enter nucleotide primer sequences ([IUB codes](#) allowed for degenerate primers). VPCR 2.0 searches the specified database for matches to the primers. If matches are found within 10000 bases, a PCR simulation model predicts amplification. Calculated PCR products are displayed within a minute.

NOTE: Abilities of VPCR 2.0 are still limited by BLAST capabilities and settings, as well as inability of our current software to deal with more than a couple thousand matches per primer. For example, using primers shorter or roughly equal to our 11-base word size misses most matches. Primers with overrepresented sequences cause problems as well. We are now busy solving most of these problems, please, be patient. If you have a minute, please, let us know what kind of expectations you have for VPCR 2.0 etc. Currently, this address is for testing VPCR 2.0, stable features will be installed on [VPCR 2.0 Homepage](#).

Search using  in the database for

Primer 1

Primer 2

Primer 3

Primer 4

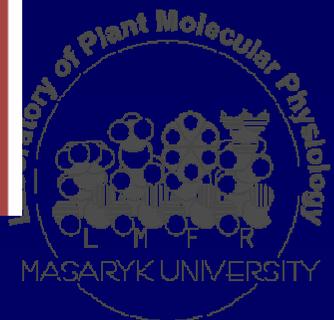
Primer 5

Primer 6

Primer 7

Primer 8

Annealing temperature

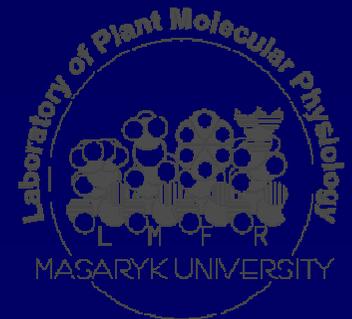


# Analytické nástroje



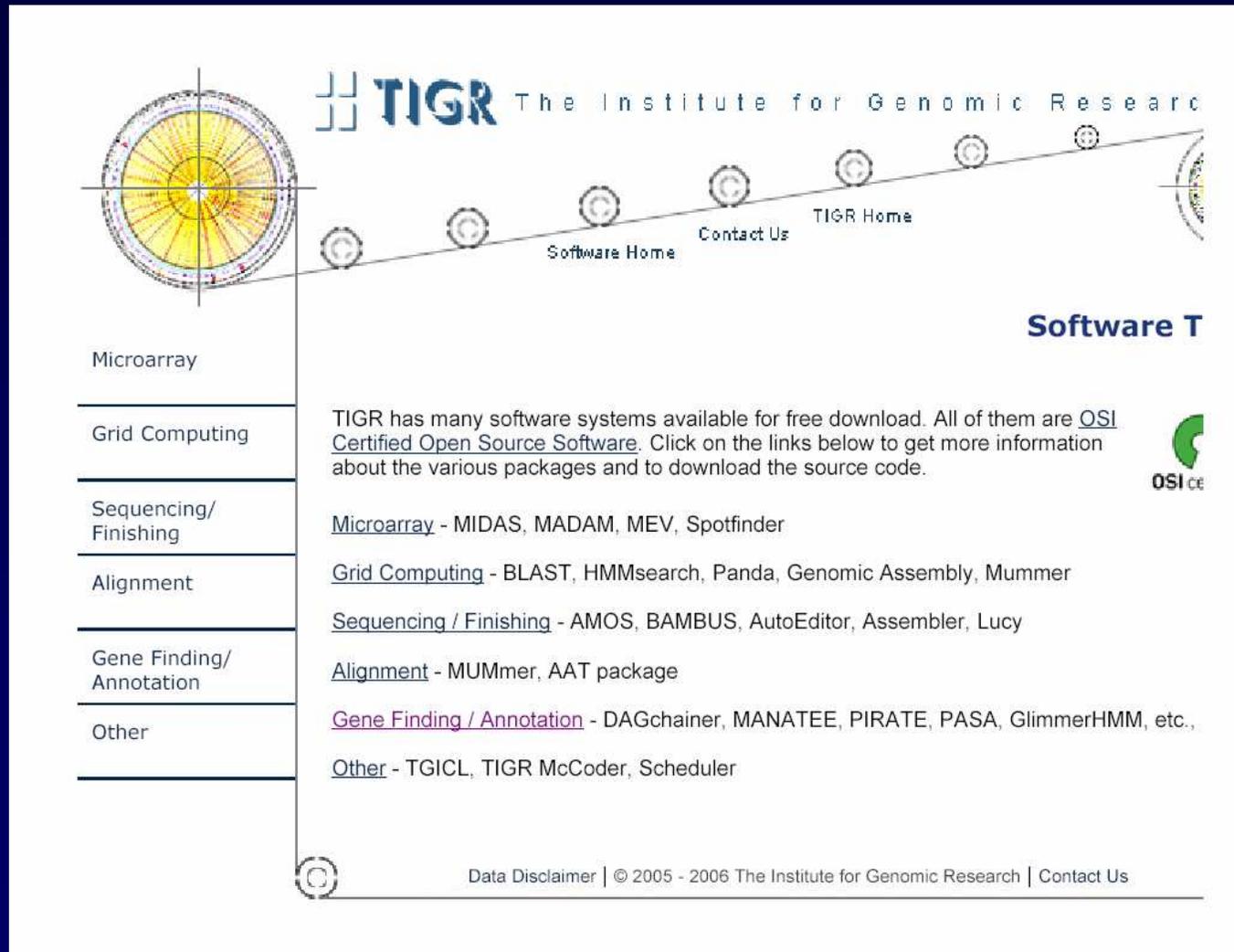
# Základy genomiky I.

- Role BIOINFORMATIKY v současném pojetí FUNKČNÍ GENOMIKY
- Databáze
  - Spektrum „on-line“ zdrojů
  - PRIMÁRNÍ, SEKUNDÁRNÍ a STRUKTURÁLNÍ databáze
  - GENOMOVÉ zdroje
- Analytické nástroje
  - Vyhledávání homologií
  - Vyhledávání sekvenčních motivů, otevřených čtecích rámců, restrikčních míst....
  - Další [www](#) genomové nástroje



# www analytické nástroje

- TIGR (The Institute for Genomic Research, <http://www.tigr.org/software/>)



The screenshot shows the TIGR website's software page. At the top left is a circular genomic map. To its right is the TIGR logo and the text 'The Institute for Genomic Research'. A navigation bar contains icons for 'Software Home', 'Contact Us', and 'TIGR Home'. The main heading is 'Software T'. Below this is a table with categories on the left and descriptions on the right. A paragraph of text explains that all software is OSI Certified Open Source Software. The OSI logo is present. At the bottom, there is a footer with a copyright notice and a 'Contact Us' link.

Category	Description
Microarray	
Grid Computing	TIGR has many software systems available for free download. All of them are <a href="#">OSI Certified Open Source Software</a> . Click on the links below to get more information about the various packages and to download the source code.
Sequencing/ Finishing	<a href="#">Microarray</a> - MIDAS, MADAM, MEV, Spotfinder
Alignment	<a href="#">Grid Computing</a> - BLAST, HMMsearch, Panda, Genomic Assembly, Mummer <a href="#">Sequencing / Finishing</a> - AMOS, BAMBUS, AutoEditor, Assembler, Lucy
Gene Finding/ Annotation	<a href="#">Alignment</a> - MUMmer, AAT package <a href="#">Gene Finding / Annotation</a> - DAGchainer, MANATEE, PIRATE, PASA, GlimmerHMM, etc.,
Other	<a href="#">Other</a> - TGICL, TIGR McCoder, Scheduler

Data Disclaimer | © 2005 - 2006 The Institute for Genomic Research | [Contact Us](#)



# Základy genomiky I.

## shrnutí

- Role BIOINFORMATIKY v současném pojetí FUNKČNÍ GENOMIKY
- Databáze
  - Spektrum „on-line“ zdrojů
  - PRIMÁRNÍ, SEKUNDÁRNÍ a STRUKTURÁLNÍ databáze
  - GENOMOVÉ zdroje
- Analytické nástroje
  - Vyhledávání homologií
  - Vyhledávání sekvenčních motivů, otevřených čtecích rámců, restrikčních míst....
  - Další www genomové nástroje



# Základy genomiky I.

diskuse

