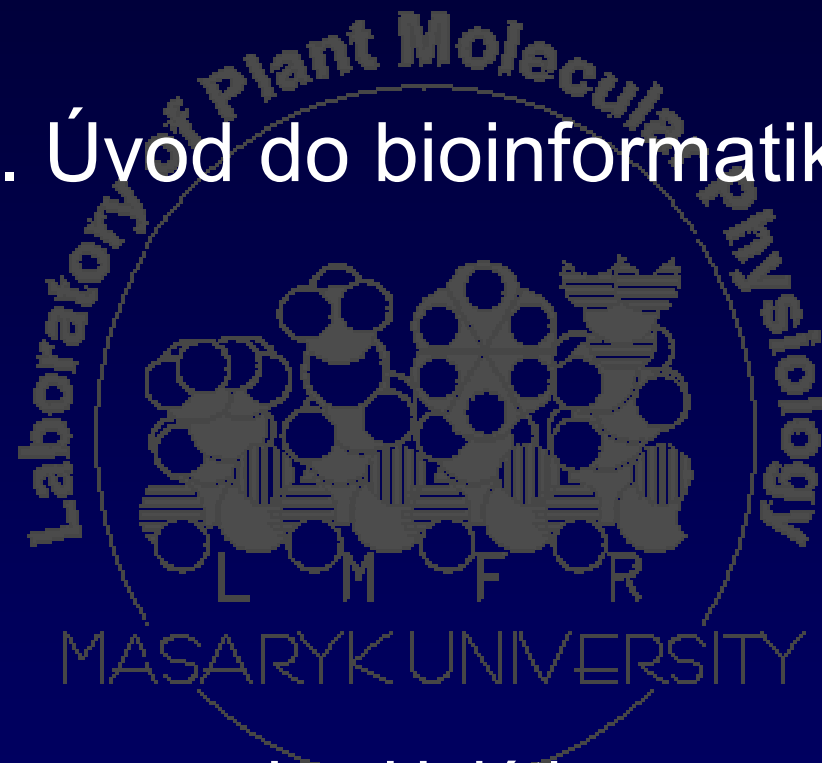


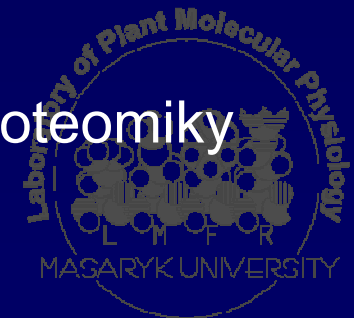
Základy genomiky

I. Úvod do bioinformatiky



Jan Hejátko

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



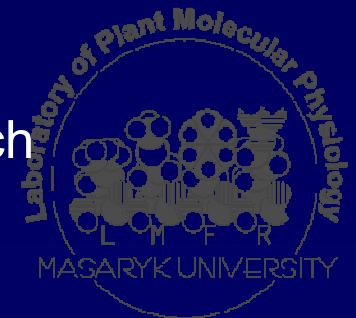
Základy genomiky I.

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



Základy genomiky I.

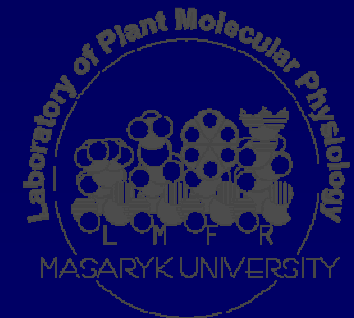
- Role BIOINFORMATIKY v současném pojetí FUNKČNÍ GENOMIKY
- Databáze
 - Spektrum „on-line“ zdrojů
 - PRIMÁRNÍ, SEKUNDÁRNÍ a STRUKTURÁLNÍ databáze
 - GENOMOVÉ zdroje
- Analytické nástroje
 - Vyhledávání homologií
 - Vyhledávání sekvenčních motivů, otevřených čtecích rámců, restrikčních míst....
 - Další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

:

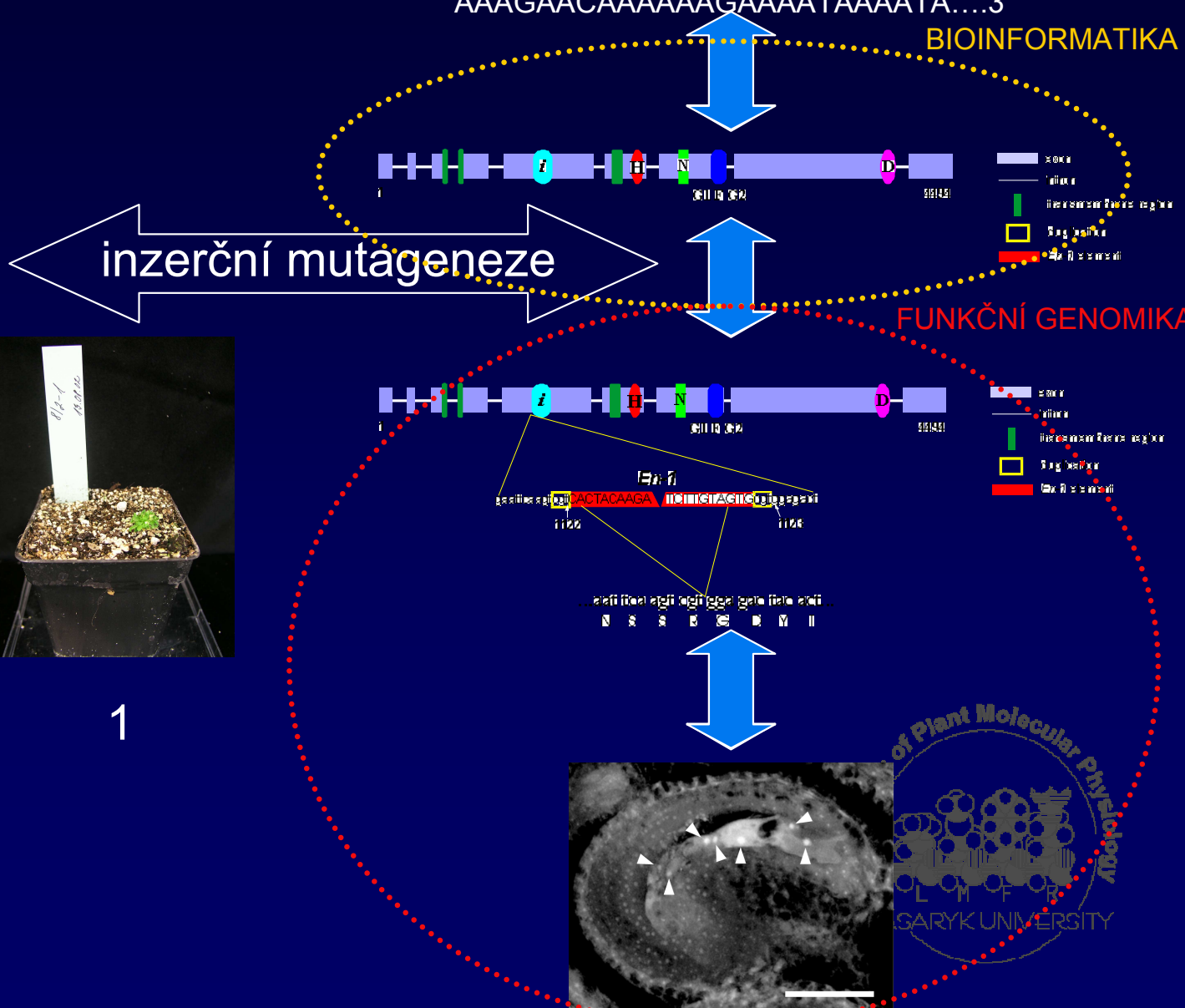
1



?

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

EMBNet Specialist Nodes

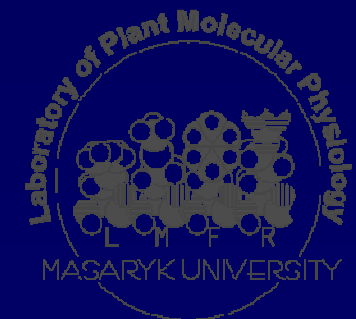
| | | |
|---------------------|-------------|-------------------------------------------------------------------------------------------|
| MIPS | Germany | http://www.mips.biochem.mpg.de/ |
| ICGEB | Italy | http://www.icgeb.trieste.it/ |
| Pharmacia Upjohn | Sweden | http://www.pnu.com/ |
| F.Hoffmann-La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

EMBNet Associate Nodes

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

USA Information Providers

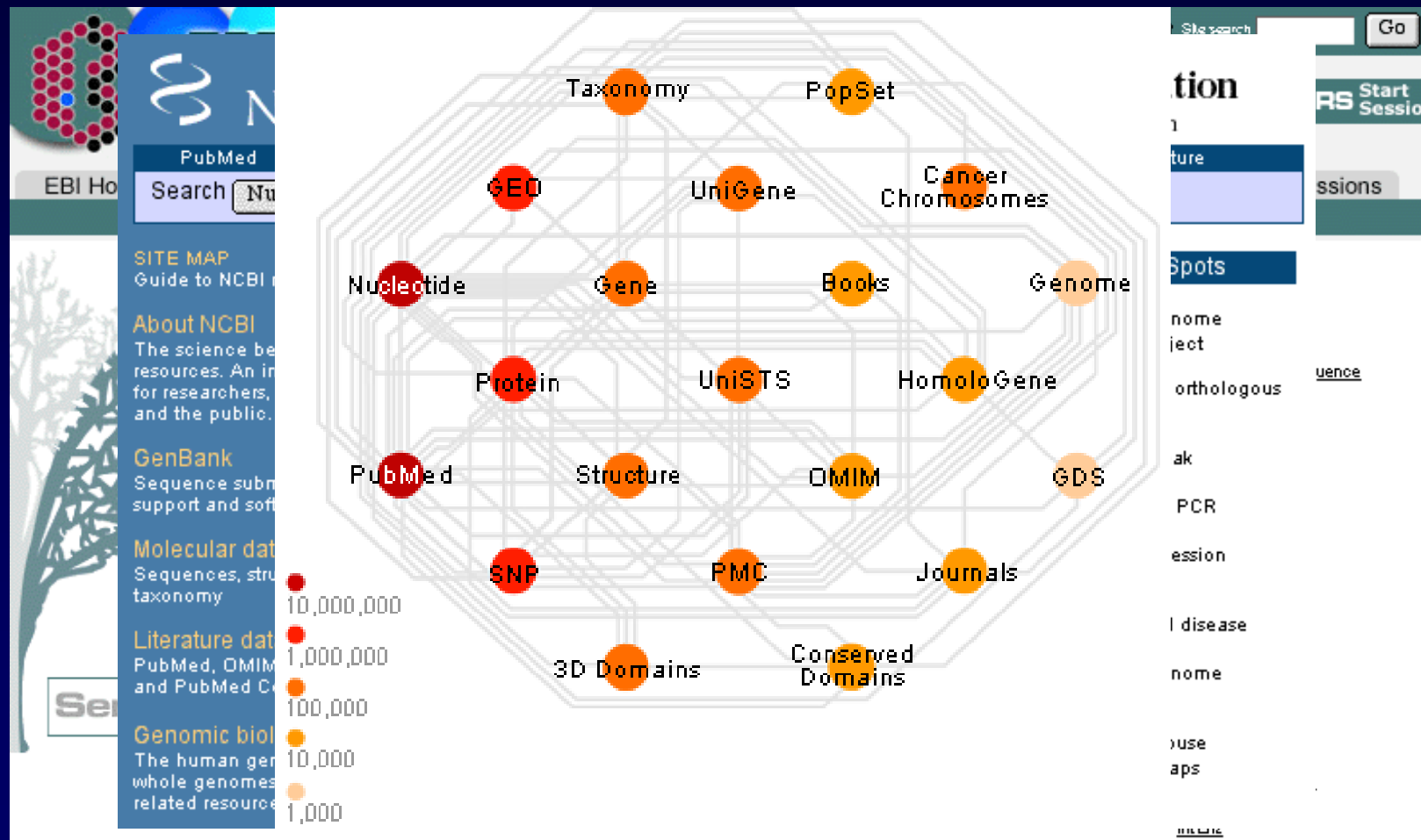
| | | |
|------|-----|-------------------------------------------------------------------------|
| NCBI | USA | http://www.ncbi.nlm.nih.gov/ |
| NLM | USA | http://www.nlm.nih.gov/ |
| NIH | USA | http://www.nih.gov/ |



Databáze

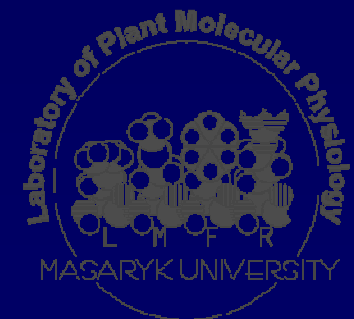
Spektrum on-line zdrojů

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



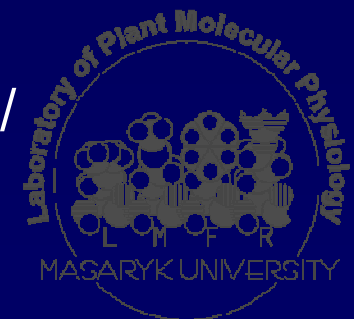
Základy genomiky I.

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



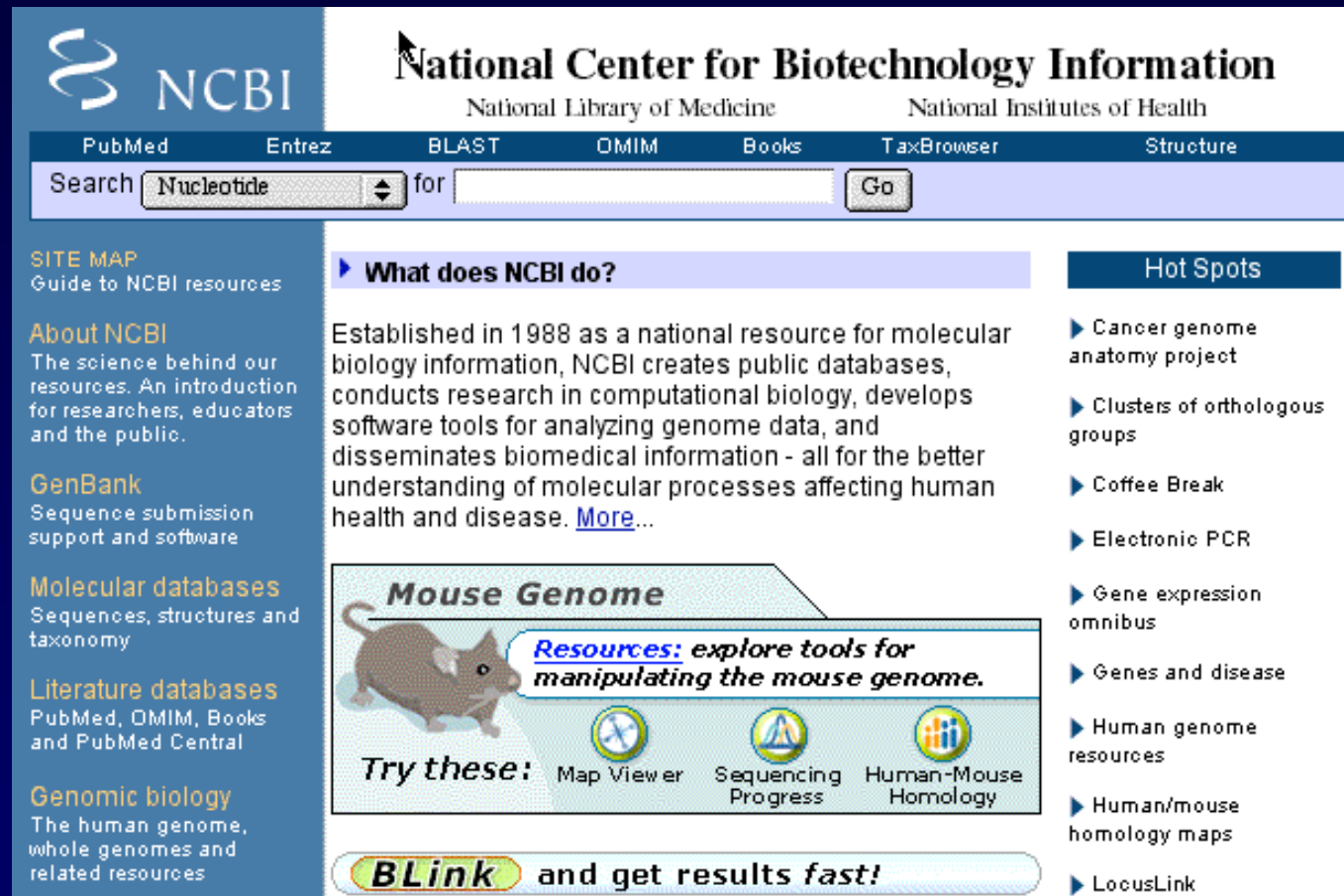
Primární databáze

- 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/>



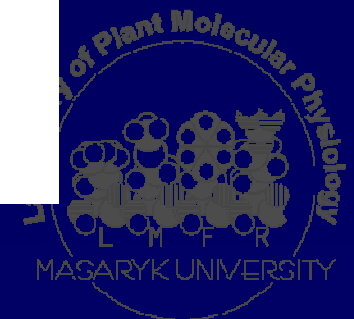
Primární databáze

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



The screenshot shows the NCBI homepage with the following elements:

- Header:** NCBI logo, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health.
- Navigation:** PubMed, Entrez, BLAST, OMIM, Books, TaxBrowser, Structure.
- Search:** Search Nucleotide for [] Go
- Left Sidebar:**
 - 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
- Main Content:**
 - 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...](#)
 - Mouse Genome:** *Resources: explore tools for manipulating the mouse genome.* Try these: Map Viewer, Sequencing Progress, Human-Mouse Homology.
 - BLink:** and get results fast!
- 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



Primární databáze

```
/translation="MNGRYSFTRQDFKTGAKPWSILALIWAAMIFAFMAVASWQDNAT
TQAILSQLRSINADSASLQRDVLRAHTGTVANVRPIISRLGALRKNLEDLKQLFRQSH
IVSEBNAQLLRQLEVSLNSADAAVAAPGAQNVRIQDLSLAFTRALSPLGKASTDQT
LEKPTLALASMLQPLRQPSPAISFBI SLELERLQKQRGLDEAPVRILAREGPIILELL
PQVKDLVNNIQTSDTAEIAENLQRECLEVYSLKNVVERSARIFLGSASVGLCLYITL
VYRLRKKTDWLARRLDYBELIKEIGVCFEGBAATSSAQAALRIIQRPDADTICALAL
VDHRRWAVETFGAKHPKPVWDDSVLRRI VSRKADERATVFR IISKKIVHLPLEIP
GLEILLAHKSTDKLI AVCCLGYQSYRPRPCQGEIQLELATACLCHYIDVRRKQTECD
VLARLLEHAQRLEAVGTLAGGIAHEFNILGSELGHAELAQNSVRSVTRRYIDYII
SSGDRAMLIIDQILTLERKQERMIKPPSEBELVTEIAPLLRMLPPNIELSFRPDMQ
SVIEGSEPLELQQVLINICKNASQAMTANGQIDIIISQAFLPVKKILAHGVMPPGDVYL
LSISDNGGGIPBAVLPHIFEPFPTTRARNGGTGLGLASVHGHI SAFPAGYIDVSTVGH
GTRPD IYLPPESSKEPVNPDSPFGRNKAPRGNGBI VALVEPDDLREAYEDKIAALGYE
PVGPRTPNRI RDWISKNEADLVMDQASLPEQSPNSVDLVLTASIIIGGNDLKMT
LSREDVT RDLYLPKPISSRTMAHAILT KIKT"
```

ORIGIN

```
1 atgaacggaa gatattcacc gacgcggcag gattttaaga caggcgcgaa gccttggtct
61 atattggccc ttatcgttgc tgeaatgatt ttocogttca tggcgggtgc gtcttggcag
121 gacaatgega ctaccccagg aatcctcage caactacgat cgat taacgc cgacagcgcc
181 tcaactgcage gogatgtact ccgcgctcac acgggcaccg tggcgaaacta ccgccccatt
241 atctccaggc tgggagctct gcggaagaat ctggaagatt tgaagcaatt atttagacaa
301 tctcatattg taagttagag caatgctgct caactgctac gccagctaga agtgtctcta
361 aattcggtct acgcgggcgt cgcgcctctt ggtgcgcaaa atgtacgctt gcaagattcg
421 ctggccagtt tcaactcgtc tttgagcagt ctccaggaa aagcctcaac cgatcagact
481 ttagaaaaac caacagaatt ggetagcagt atgctccaat ttcttcggca accaaagcccg
541 gctatttcat togagatcag ccttgaacta gagagggctcc aaaaaacaac cggtcttgat
601 gaagetcccc tgcgcatact tgcacgtgaa ggtcccaata tcttateget ttgcccacag
661 gtgaaaagtc tgggtgaacat gattcagacg tctgacaccg cagaaattgc ggagatgctg
721 cagcgcagat gtttggagggt ctatagcttg aaaaatgtag agggagcggag cgcacgtatc
781 tttcttgggt ccgcttcagt gggctcttgc ctctacatca tcaacttagt ctataggcta
841 cgcacaaaaa ccgatttggtt agcgcggcgt ttagattacg aagagctaat caaagagatc
901 ggaagtatgt ttgaaggtag ggcggccacc acgtcgtccg cgcacagctgc acttcgtatt
961 attcagcctc tctttgatgc cgatacgtgc cgttagctc tagtggacca tgaccgtaga
1021 tgggctgtcg aaacattcgg tgcgaaacac ccaaaacctg tgtgggacga cagcgtgcta
1081 cgcgaaatag tctctcgtac caaagcggac gaacgggcca cggatctccg catcatatcg
1141 tcgaaaaaaa togtacattt gectctcgaa atccaggtc tctcgatact actggctcac
1201 aaatccacag ataaactaat tgcggtttgt tcaactgggtt 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 tctcgatcag
1561 atcttgacgc tgagccgaaa acaggagcgc atgatcaagc catttagtgt ctcagagctt
1621 gtgaccgaaa togtccctct gctacgtatg gctcttcgcg caaacatcga gcttagtttc
1681 agatttgatc aaatgcagag cgtgatcgaa ggaagccgcg ttgaactca acaggtacta
1741 attaacatct gcaagaatgc tcccaagcc atgactgcaa atggtcaaat cgacatctac
1801 atcagccaaag cttttttacc agttaagaaa attctggcgc atggtgttat gccacctggc
1861 gactatgttc tctatctat tagcgacaat ggtggaggca tcccgaggc tgtgttacc
1921 cacatttttg aacctctctt tacgacacga gctcgcacac gtggaacggg tctcggcctt
1981 gcttctgtgc atggtcatat cagcgcgttt ggggttaca togacgttag tcaactgtt
2041 gggcatggga cgcgctttga catttatctc cctcgtctt ctaaggaacc cgtaaatcca
2101 gacagttttt toggccgcaa taaggcaccg cgtggaaaac gggagattgt gccacttgtt
2161 gagcccgatg acctcctcgc ggaaggctat gaagacaaga togcctctct aggatatgag
2221 ccggctcgggt tctgtacctt taatgaaatt cgcgattgga tttcaaaag caatgaaagc
2281 gatctggtea tggctgacaa agcgtctctt cctgaagatc aaagtcttaa tccctggat
2341 ttagtctca agaccgcctc catcatcatt ggcggaaatg atctcaaaat gacctttca
```

Proteinové sekundární databáze

- databáze funkčních nebo strukturálních *motivů* získaných srovnáváním 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 SIR 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 PMEMRKSVPFNYVQVREtAQSHQGTGLGLGIVQ
SLVRLMGGGIRITDKAMGekGTCPQFNVLIT
- >[PDOC50110 PS50110 RESPONSE_REGULATORY](#) Response regulatory domain [profile].
987 - 1085 RVLVVDNPFISRKVATGKLLKMGVSeVEQCDSGKRALRLVTRGLtqreeggsvdklpFDY
IFMDQMPMDGYRATREIRkvekSYGVRTPIIAVSGHD-----

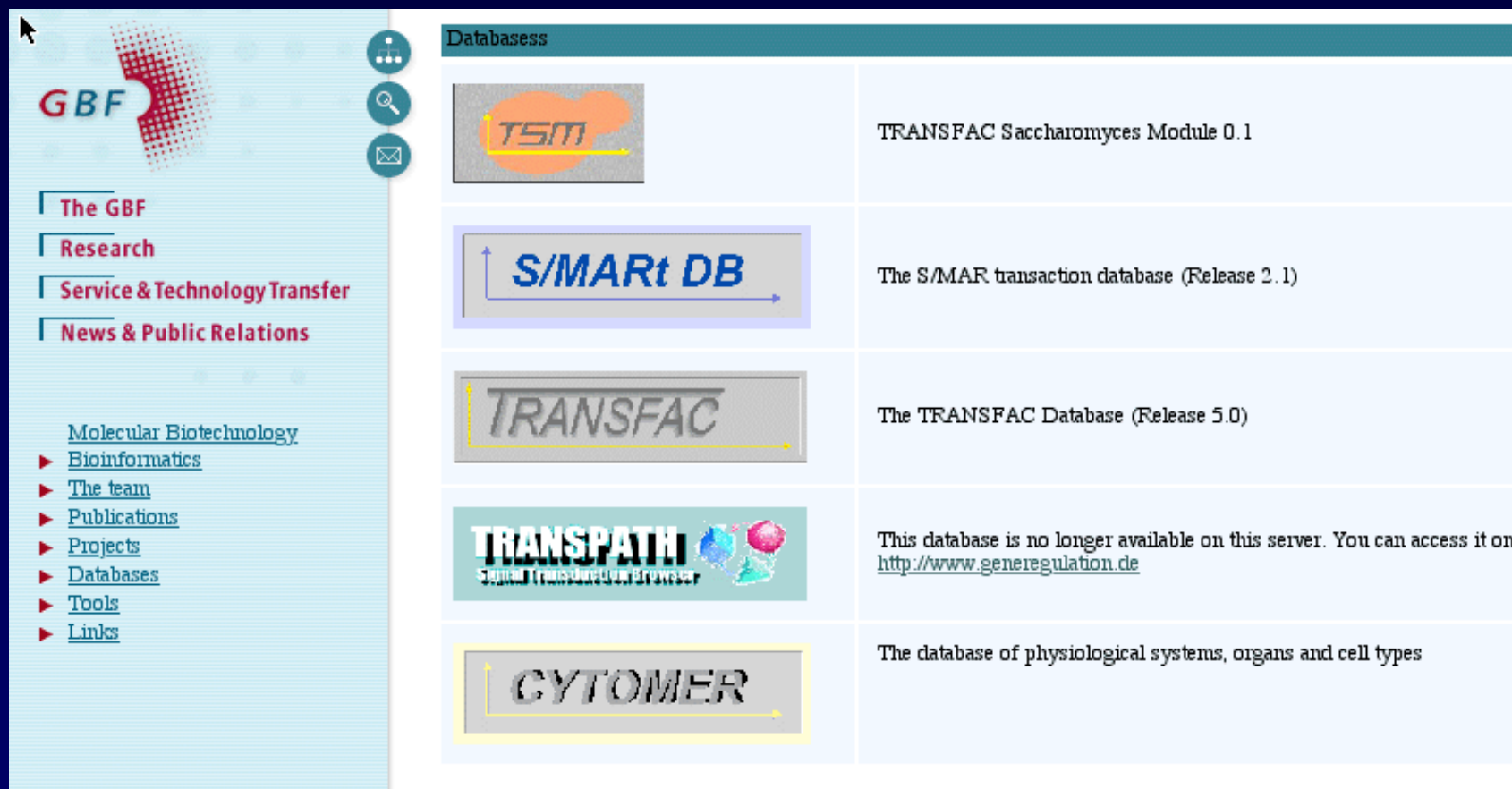
Below the results is a section titled "Graphical summary of hits (java applet)". It features a horizontal bar representing a protein sequence of 100 residues. Two red vertical cursors are positioned at the beginning and end of the sequence. A yellow bar highlights the "HIS_KIN" motif, and a red bar highlights the "RESPONSE_REGULATORY" motif. The interface includes "About" and "Prefs" buttons, and a "Zoom" control set to 1123.

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





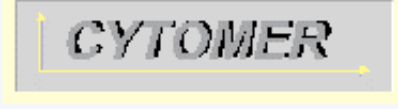
At the bottom of the screenshot, there is another navigation bar with the same links as the top, and a "START THE SCAN" button with a "RESET" button next to it.

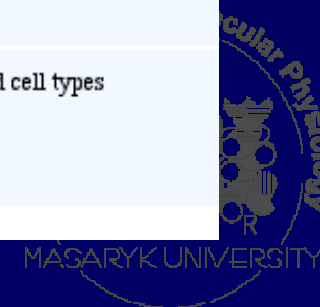
Sekundární databáze DNA

- 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 http://www.generegulation.de |
|  | 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](#)

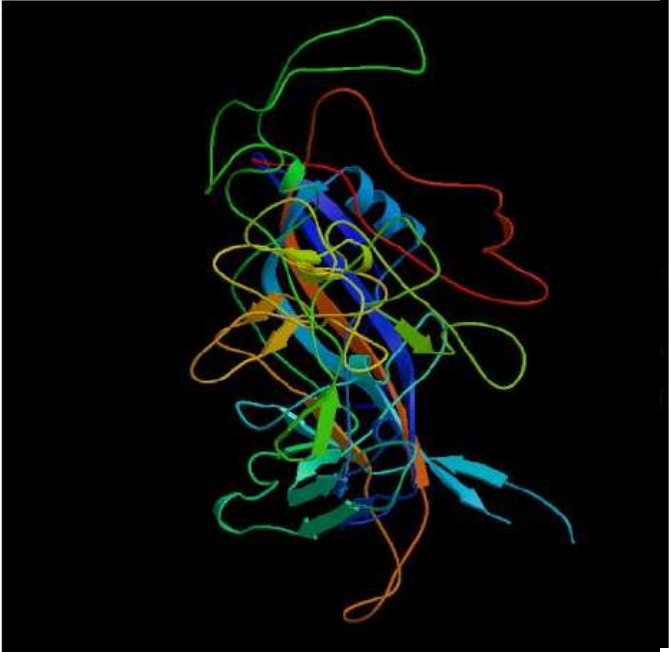
[Structural Neighbors](#)


[Geometry](#)

[Other Sources](#)

[Sequence Details](#)

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



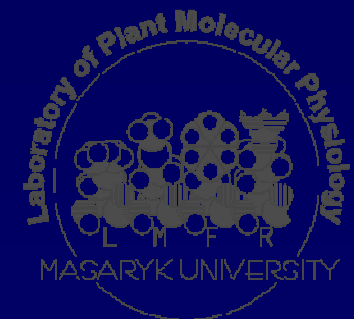


<http://www.rcsb.org/pdb/cgi/explore.cgi?job=graphics;pdbId=1P5Y;page=:pid=173561064349344&bio=1&opt=show&size=500> 12/29/2003



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

The image shows a screenshot of the National Center for Biotechnology Information (NCBI) website. At the top, the NCBI logo is on the left, and the text "National Center for Biotechnology Information" is centered, with "National Library of Medicine" and "National Institutes of Health" below it. A navigation bar contains links for PubMed, Entrez, BLAST, OMIM, Books, TaxBrowser, and Structure. Below this is a search bar with a dropdown menu set to "Nucleotide" and a "Go" button. On the left side, there is a "SITE MAP" section with links to "About NCBI", "GenBank", "Molecular databases", "Literature databases", and "Genomic biology". The "Genomic biology" link is circled in red. The main content area features a "What does NCBI do?" section with a paragraph about the center's mission. To the right is a "Hot Spots" section with a list of current projects. Below the main text is a "Mouse Genome" banner with a mouse illustration and icons for "Map Viewer", "Sequencing Progress", and "Human-Mouse Homology". At the bottom, there is a "BLink" button and the text "and get results fast!".

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

PubMed Entrez BLAST OMIM Books TaxBrowser Structure

Search **Nucleotide** for **Go**

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

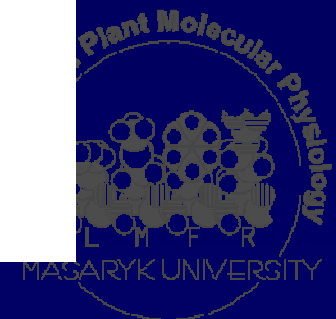
- ▶ Cancer genome anatomy project
- ▶ Clusters of orthologous groups
- ▶ Coffee Break
- ▶ Electronic PCR
- ▶ Gene expression omnibus
- ▶ Genes and disease
- ▶ Human genome resources
- ▶ Human/mouse homology maps
- ▶ LocusLink

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 Genomes

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

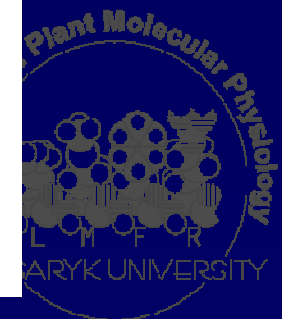
Arabidopsis thaliana genome view [BLAST search Arabidopsis genome](#)

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 | At | Gene | Accession | Description |
|-------|----|-----------|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1H | | At1g05180 | YUP8H12.21 | TIGR MIPS TAIR NM_100396 NP_172010 auxin-resistance protein AXR1 |
| 2H | | At1g08210 | T23G18.7 | TIGR MIPS TAIR NM_100695 NP_563808 expressed protein |
| 3H | | At1g11250 | T28P6.10 | TIGR MIPS TAIR NM_100997 NP_172591 syntaxin-related protein At-SYR1, putative |
| 4H | | At1g14670 | T5E21.15 | TIGR MIPS TAIR NM_101334 NP_172919 endomembrane protein, putative |
| 5H | | At1g17790 | F2H15.2 | TIGR MIPS TAIR NM_101642 NP_564037 expressed protein |
| 6H | | At1g21050 | T22I11.13 | TIGR MIPS TAIR NM_101958 NP_564130 expressed protein |
| 7H | | At1g24210 | F3I6.14 | TIGR MIPS TAIR NM_102267 NP_564212 expressed protein |
| 8H | | At1g28370 | F3M18.20 | TIGR MIPS TAIR NM_102603 NP_174159 ethylene-responsive element binding factor, putative |
| 9H | | At1g31885 | F5M6.28 | TIGR MIPS TAIR NM_102926 NP_174472 major intrinsic protein, putative |
| 10H | | At1g35670 | F15O4.8 | TIGR MIPS TAIR NM_103271 NP_174807 calcium-dependent protein kinase |
| 11H | | At1g48160 | F21D18.11 | TIGR MIPS TAIR NM_103712 NP_175250 signal recognition particle 19 kDa protein subunit, putative |
| 12H | | At1g51980 | F5F19.4 | TIGR MIPS TAIR NM_104079 NP_175610 mitochondrial processing peptidase alpha subunit, putative |
| 13H | | At1g55150 | T7N22.9 | TIGR MIPS TAIR NM_104388 NP_175911 ethylene-responsive RNA helicase, putative |
| 14H | | At1g60140 | T13D8.4 | TIGR MIPS TAIR NM_104705 NP_176221 trehalose-6-phosphate synthase, putative |
| 15H | | At1g63750 | F24D7.6 | TIGR MIPS TAIR NM_105052 NP_176562 putative disease resistance protein |
| 16H | | At1g67090 | F5A8.1 | TIGR MIPS TAIR NM_105379 NP_176880 ribulose-bisphosphate carboxylase small unit, putative |
| 17H | | At1g69870 | T17F3.10 | TIGR MIPS TAIR NM_105655 NP_564979 putative peptide transporter |
| 18H | | At1g72970 | F3N23.17 | TIGR MIPS TAIR NM_105955 NP_565050 expressed protein |
| 19H | | At1g75780 | T4O12.1 | TIGR MIPS TAIR NM_106228 NP_177706 tubulin beta-1 chain |
| 20H | | At1g78700 | F9K20.26 | TIGR MIPS TAIR NM_106517 NP_565187 expressed protein |

Region Shown:

out zoom in

ideogram master

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

Primární data-genomové zdroje

The screenshot displays the NCBI Nucleotide search interface. At the top, there are navigation tabs for PubMed, Nucleotide, Protein, Genome, Structure, PMC, Taxonomy, OMIM, and Books. A search bar is set to 'Nucleotide' with a 'Go' button. Below the search bar, there are options for 'Limits', 'Preview/Index', 'History', 'Clipboard', and 'Details'. A message states 'History has expired due to inactivity.' Below this, there are buttons for 'Display', 'Graphics', 'Show: 1', 'Send to', 'File', and 'Get Subsequence'. The search results show '1: NC_003070. Arabidopsis thali...[gi:22330780]' with a 'Links' button.

The main section is titled 'Reverse Complement Strand' and 'View on plus strand'. It includes a search for gene field, a 'Find' button, and checkboxes for 'CDS with gene and mRNA' (checked) and 'Hide sequence'. A 'Refresh' button is also present. The sequence is displayed as a horizontal bar with a scale from 3002869 to 1. A red box highlights a 1.5M region. Below the bar, several gene models are shown, including At1g05200, At1g05190, At1g05180, and At1g05170. A legend indicates: pink line for CDS, blue line for RNA, grey line for gene, and red line for sequence fragment shown.

The 'Sequence:' section shows a detailed view of the DNA sequence and its translation. The sequence is:


```

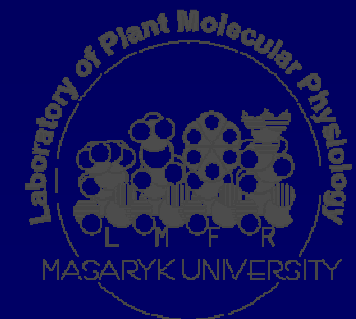
    1501830 CTTTTTGT TT ATCAGTTCAC CGGAGGCARA AATCGTCTCT CGCTTGAGCT CCGAAGATCC
    1501770 AACGAGTAAA AAGATCCAGG AGCATGTTG AACAGAGGCC AACAAATGGA CAACCTAAAA
    1501710 CCAAGTACGA TGCTCAGCTC AGGTATACAT ACTCTTTTTC CTAAACTCT ACTTCGACT
    1501650 ACTCTTGTCG GGAGGAAAT TGATGAGTA GTAATCGTTG GGTGACTGA TTAGATTTG
    1501590 GGGGAGSTA GGTCAAGCG CTTGGAGA AGCGAGTATC TGTTACTCA ATTGTGCCC
    
```

 The translation is:


```

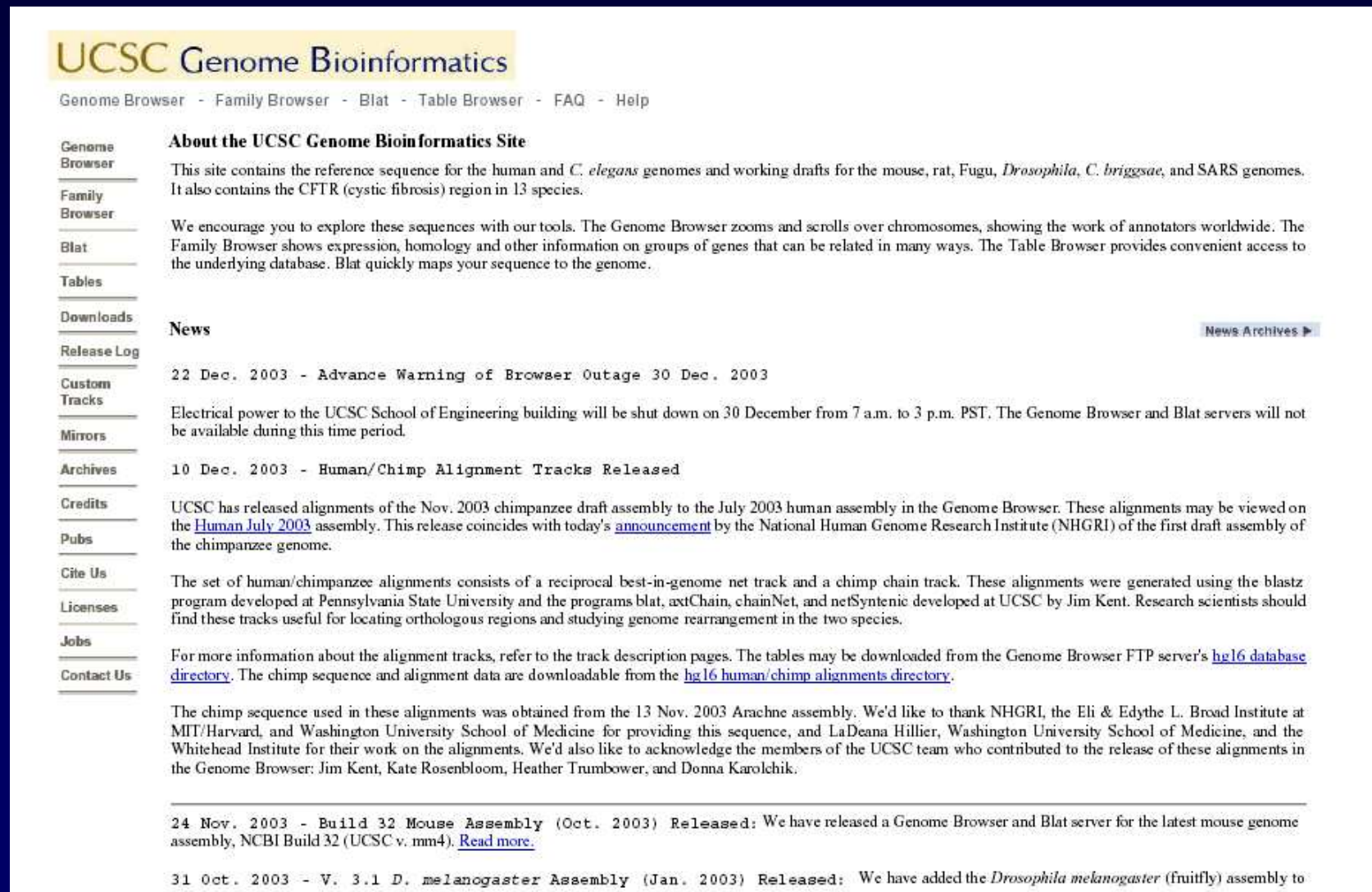
    Q A V K R S R R H V E E E P T M V E P K
    T K Y D R Q L R
    I W
    G E V G Q A A L E E A S I C L L N C G P
    
```

 Gene models for At1g05180 are shown with mRNA-protein id and CDS labels.



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
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

Mirrors

Archives

Credits

Pubs

Cite Us

Licenses

Jobs

Contact Us

About the UCSC Genome Bioinformatics Site

News

22 Dec. 2003 - Advance Warning of Browser Outage 30 Dec. 2003

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.

10 Dec. 2003 - Human/Chimp Alignment Tracks Released

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.

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.

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.

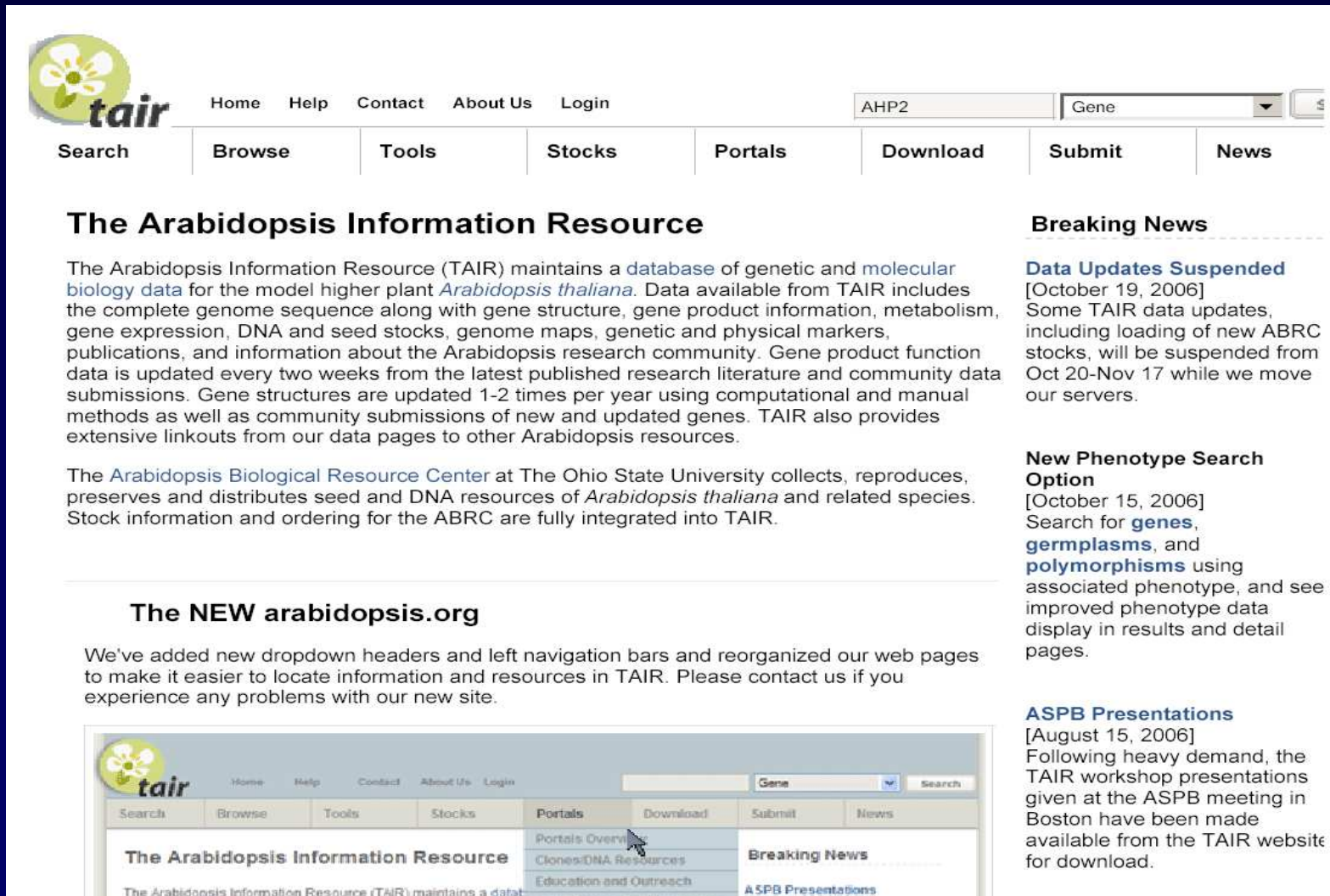
24 Nov. 2003 - Build 32 Mouse Assembly (Oct. 2003) Released: We have released a Genome Browser and Blat server for the latest mouse genome assembly, NCBI Build 32 (UCSC v. mm4). [Read more.](#)

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 buttons for Search, Browse, Tools, Stocks, Portals, Download, Submit, and News. The main content area is divided into two columns. The left column features the heading 'The Arabidopsis Information Resource' followed by a paragraph describing the database of genetic and molecular biology data for *Arabidopsis thaliana*. Below this is another paragraph about the Arabidopsis Biological Resource Center at The Ohio State University. The right column has a 'Breaking News' section with two items: 'Data Updates Suspended' (dated October 19, 2006) and 'New Phenotype Search Option' (dated October 15, 2006). At the bottom of the right column is a section for 'ASPB Presentations' (dated August 15, 2006). A smaller version of the website interface is shown at the bottom of the screenshot, with a mouse cursor pointing to the 'Portals' dropdown menu. 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.

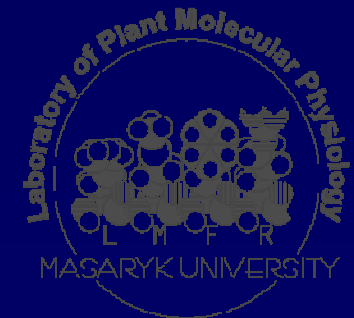
New Phenotype Search Option
[October 15, 2006]
Search for **genes**, **germplasms**, and **polymorphisms** using associated phenotype, and see improved phenotype data display in results and detail pages.

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

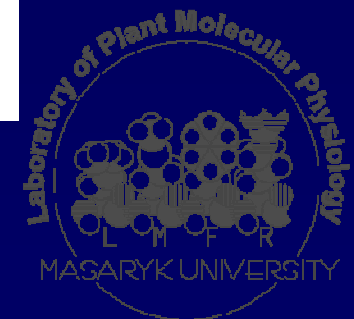
[Search](#)

```
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



Podstata algoritmu BLAST (Basic Local Alignment Search Tool)

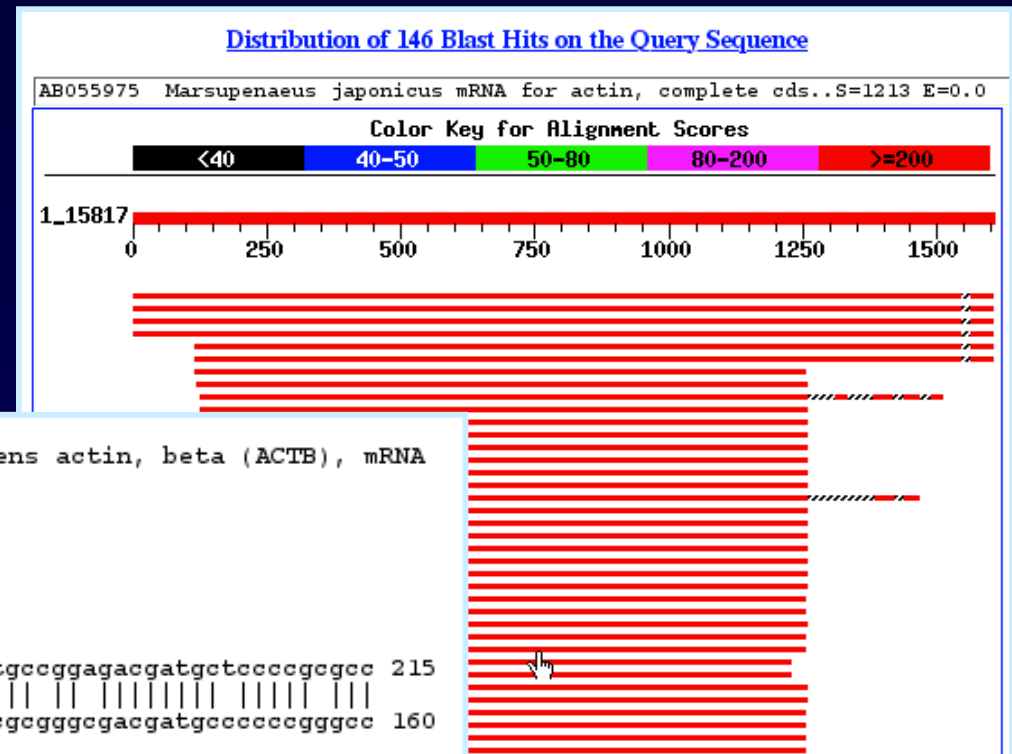
- Velikost vyhledávacího slova (word size): 7 - 15 bází
 - Primární podobnosti (seed matches)
 - Rozšiřování oblasti homologie doprava i doleva
- Hodnocení homologie pomocí matice PAM (Point Accepted Mutation) nebo BLOSUM (BLOcks Substitution Matrix)
- Zobrazení výsledků

Matrice PAM 250

| C | 12 | | | | | | | | | | | | | | | | | | | | |
|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|----|----|--|
| S | 0 | 2 | | | | | | | | | | | | | | | | | | | |
| T | -2 | 1 | 3 | | | | | | | | | | | | | | | | | | |
| P | -3 | 1 | 0 | 6 | | | | | | | | | | | | | | | | | |
| A | -2 | 1 | 1 | 1 | 2 | | | | | | | | | | | | | | | | |
| G | -3 | 1 | 0 | -1 | 1 | 5 | | | | | | | | | | | | | | | |
| N | -4 | 1 | 0 | -1 | 0 | 0 | 2 | | | | | | | | | | | | | | |
| D | -5 | 0 | 0 | -1 | 0 | 1 | 2 | 4 | | | | | | | | | | | | | |
| E | -5 | 0 | 0 | -1 | 0 | 0 | 1 | 3 | 4 | | | | | | | | | | | | |
| Q | -5 | -1 | -1 | 0 | 0 | -1 | 1 | 2 | 2 | 4 | | | | | | | | | | | |
| H | -3 | -1 | -1 | 0 | -1 | -2 | 2 | 1 | 1 | 3 | 6 | | | | | | | | | | |
| R | -4 | 0 | -1 | 0 | -2 | -3 | 0 | -1 | -1 | 1 | 2 | 6 | | | | | | | | | |
| K | -5 | 0 | 0 | -1 | -1 | -2 | 1 | 0 | 0 | 1 | 0 | 3 | 5 | | | | | | | | |
| M | -5 | -2 | -1 | -2 | -1 | -3 | -2 | -3 | -2 | -1 | -2 | 0 | 0 | 6 | | | | | | | |
| I | -2 | -1 | 0 | -2 | -1 | -3 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | 2 | 5 | | | | | | |
| L | -6 | -3 | -2 | -3 | -2 | -4 | -3 | -4 | -3 | -2 | -2 | -3 | -3 | 4 | 2 | 6 | | | | | |
| V | -2 | -1 | 0 | -1 | 0 | -1 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | 2 | 4 | 2 | 4 | | | | |
| F | -4 | -3 | -3 | -5 | -4 | -5 | -4 | -6 | -5 | -5 | -2 | -4 | -5 | 0 | 1 | 2 | -1 | 9 | | | |
| Y | 0 | -3 | -3 | -5 | -3 | -5 | -2 | -4 | -4 | -4 | 0 | -4 | -4 | -2 | -1 | -1 | -2 | 7 | 10 | | |
| W | -8 | -2 | -5 | -6 | -6 | -7 | -4 | -7 | -7 | -5 | -3 | 2 | -3 | -4 | -5 | -2 | -6 | 0 | 0 | 17 | |
| | C | S | T | P | A | G | N | D | E | Q | H | R | K | M | I | L | V | F | Y | W | |

MRKE
MRKE
MRKY
MRAK

Analytické nástroje



[gi|5016088|ref|NM_001101.2|](#) **LU** Homo sapiens actin, beta (ACTB), mRNA
Length = 1793

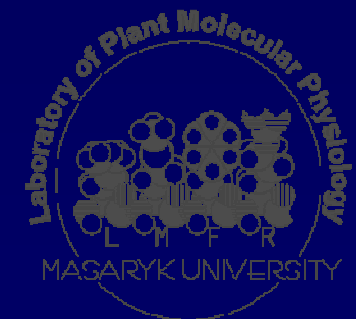
Score = 1110 bits (560), Expect = 0.0
Identities = 965/1100 (87%)
Strand = Plus / Plus

Query: 156 gtcgacaacggctctggcatgtgcaaggccggatttgccggagacgatgctccccgcgcc 215
|||
Sbjct: 101 gtcgacaacggctccggcatgtgcaaggccggcttcgogggacgatgccccccgggcc 160

Query: 216 gtcttcccacgatgtgggacgtccccgtcaccaggggtgtgatggcggcatgggcccag 275
|||
Sbjct: 161 gtcttcccctccacgtgtggggcgccccagggcaccagggcgtgatggcggcatgggcccag 220

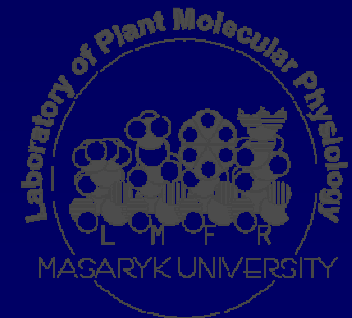
Query: 276 aaggactcgtacgtgggtgatgaggcgcagagcaagcgtggtatcctcaccctgaagtac 335
|||
Sbjct: 221 aaggattcctatgtgggagcagggcccagagcaagagaggcctcctcaccctgaagtac 280

Query: 336 cccattgagcacggtatcgtgaccaactgggacgatatggagaagatctggcaccacacc 395
|||
Sbjct: 281 cccatcgagcacggcatcgtcaccaactgggacgacatggagaaaatctggcaccacacc 340



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
GTTACTCAATTTGAATTAACAAAGGGGCAAATTTGACTATTTTGGCCCTTATATCTTTTGGTCACAAAAAC
ATAAAATATCCCATCCGAAATTCCAAATGGTCCATTATCGGCAAGTAGCTTTCTTTAATTATAGTTAGTT
GACAAAACACTATCAAGATATCATTATTATAATAAATAACTTCAAAGTCCATCATCTTAGCTGCCTCCTCA
GTAGAGCCGCCAGTAAAAATAAGACCGATCAAATAAAGCCGCCATTAAAAATAATGAATTTTAGGACTCTC
GATTGGCACGTAAGTGCCAAAACCTTTCCAATACTTTGTGCAACTTGGGGCTGCTAGGTTCTGAGCTTC
CAGATATGGGATATTTCTAAGTTTATCTCCTAAITTTACATCTCAACTAATATTAAGAAITTAACAGGTA
CAGCAAATCATAAAATTTTCTCTAAAGAAGACAAATGAATCCGGTTACTGATTCAATTGGCCTTTTTCAGAG
TCTGCATGCCATATTCACATAAGGGGTCGTTTGGTAC AAGAAAATAATAATAAATTTCTGGGATAGAATTT
GAGATTGCATTTATCTTGTGTTTAATTATAAGTATTAGCTAATTTT CAGAATAAAATTTTACACTAAAATAG
TAAAATCAACTATCACATGTAGAAGGTGGAATGGAATAGCTAATCCCATAGCCACTCACATAGAATATCC
TTATTTATCTCACTATTTTACCAAATGATCGGTTAGTCTTCATGAGAATCCAGTATCCTCAATAAATGCA
GTAAGAAGTTAGAAAATTTTCATTAATCAAITCATATAATTTAAAAATATTAGATATGGAGCACTTAAG
ATACAATAAAAAGATGTACCGTTAATAATAAAAAGATAAGATAGAGTTTTAAATAGGAAAAAAAAAACGGTT
CGAGACACTCTTATGGAAGGCGTTGTCTTCAAAGTAGATTCTCATTCAATTGCTCTGGTGC AATAGCAAAA
TGACATCTTACTCTTAAGATACAGCGAGCCACTCTACAATCTTCTATTGTATACTCAAATGAAAGTTTTA
GAGAACTTTCAAATCTCTCAACTACTTTTTAAGGGAATTCAAAATACGACCAATATTTATTACTTACTTAC
TTATAGTTAAATGATATGAATTTTTATTTTTAAATTTGAAATGAAAAATTTAAATTACTTGATTTAATATAA



Analytické nástroje

Regex pattern:

ctt. {1,32}ctt

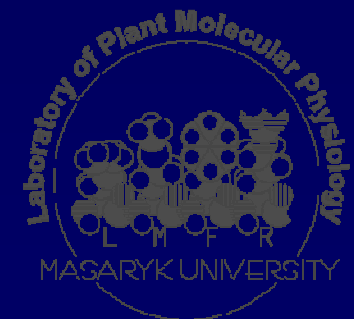
0 sequences were searched

1 match was found

Matches are indicated in blue

>170248

```
GAGCTCCCTTGGGGGGCAAGGGCAAACTTTTTTTGCTAAATGGAAAAATATTATACCAAGTGTITGTAATA
GTTACTCAATTTGAATTAACAAAGGGGCAAAATTTGACTATTTTTGCCCTTATATCTTTTTGGTCACAAAAAC
ATAAAATATCCCATCCGAAATTCCAAATGGTCCATTATCGGCAAGTAGCTTTCTTTTAAATTATAGTTAGTT
GACAAAACACTATCAAGATATCATTATTATAATAATACTTCAAAGTCCATCATCTTTAGCTGCCTCCTCA
GTAGAGCCGCCAGTAAAAATAAGACCGATCAAATAAAAAGCCGCCATTAAAAATAATGAATTTTAGGACTCTC
GATTTGGCAGTAAGTGCCAAAACCTTTCCAATACTTTTGCTGCAACTTTGGGGCTGCTAGGTTCTGAGCTTC
CAGATATGGGATATTTCTAAGTTTTATCTCCTAATTTACATCTCAACTAATATTAAAGAAATPAAACAGGTA
CAGCAAATCATAAAATTTTCTCTAAAGAAGACAATGAATCCGGTTACTGATTCATTTGGCCTTTTTAGAG
TCTGCATGCCATATTTACTAAGGGGTCGTTTTGGTACAAGAAATAATAATAATAATTTTGGGATAGAATTT
GAGATTGCATTTATCTTGTGTTTTAATTATAAGTATPAGCTAATTTTCAGAATAAAATTTTACTACTAAAATAG
TAAATCAACTATCACATGTAGAAGGTGGAATGGAATAGCTAATCCCATAGCCACTCACATAGAATATCC
TTATTTATCTCACTATTTTACCAAATGATCGGTTAGTCTTTCATGAGAATCCAGTATCCTCAATAAATGCA
GTAAGAAGTTFAGAAAAATTTTCATTAAATCAATTCATATAATTTAAAAATATTAGATATGGAGCACTTAAG
ATACAATAAAAAGATGTACCGTTAATAATAAAAAGATAAGATAGAGTTTTTAAATAGGAAAAAAAAAAAAACGGTT
CGAGACACTCTTATGGAAGGCGTTGCTTTCAAAGTAGATTCTCATTTCATTGCTCTGGTGCATAGCAAAA
TGACATCTTACTCTTAAGATACAGCGAGCCACTCTACAATCTTTCTATTGTATACTCAAATGAAAAGTTTTA
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

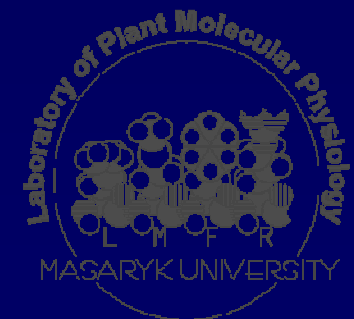
```
1 gagctcccttggggggcaagggcaaaactttttgctaaatggaaaaatattataccaagt 60
  V C N S Y S I * I N K G A N L T I L P L
61 gtttgtaatagttactcaatttgaattaacaaaggggcaaatttgactattttgcctta 120
```

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

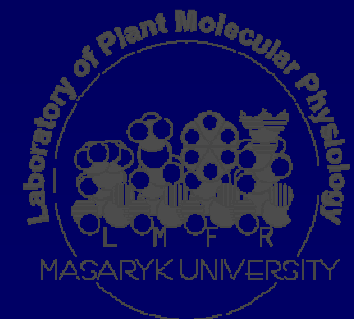
= Linear Map of Sequence:

```

          StyI
          BsaJI
        CviJI
        AluI
          SacI
        EcoICRI
          Bsp1286I
          BsiHKAI
        BanII  BslI
          \ \ \ \ \
1 gagctcccttggggggcaagggcaaaactttttgctaaatggaaaaatattataccaagt 60
  ctcgaggggaaccccccgttcccgttttgaaaaaacgatttacctttttataaatatggttca
    ^ * ^ * ^ * ^ * ^ * ^ * ^ *
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 gtttgtaatagttactcaatttgaattaacaaaggggcaaatttgactatgtttgcccetta 120
   caaacattatcaatgagttaaacttaattgtttccccgttttaaaactgataaaaacggggaat
     ^ * ^ * ^ * ^ * ^ * ^ * ^ * ^ * ^ *
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 CTTTGCTTGGTCTCTGCTTGACAACCTTCGAGTGGAGACTCGGCTACACTGCCCGTTTCGG 804655

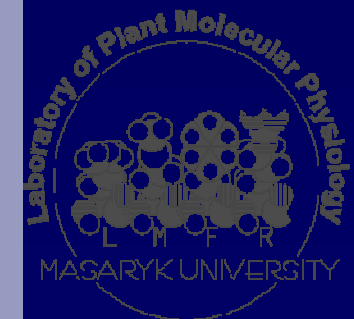
                2650      2660      2670      2680      2690      2700
24 .....AAATAGGT. 170381
1 ..... 11321163
2430 .....GAACAATT. 170248
1743 CAGTCAAATGATTGACAGAACTGCCAAAAACAAGCGAAAAATGGTAAAAAAAAAAAAAATTC 196886
2620 GATCGTCTATGTGGACTTCAATACTGTGAAGAGGTACGGCAAGGACTCAGGCTTCTGGTT 804655

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

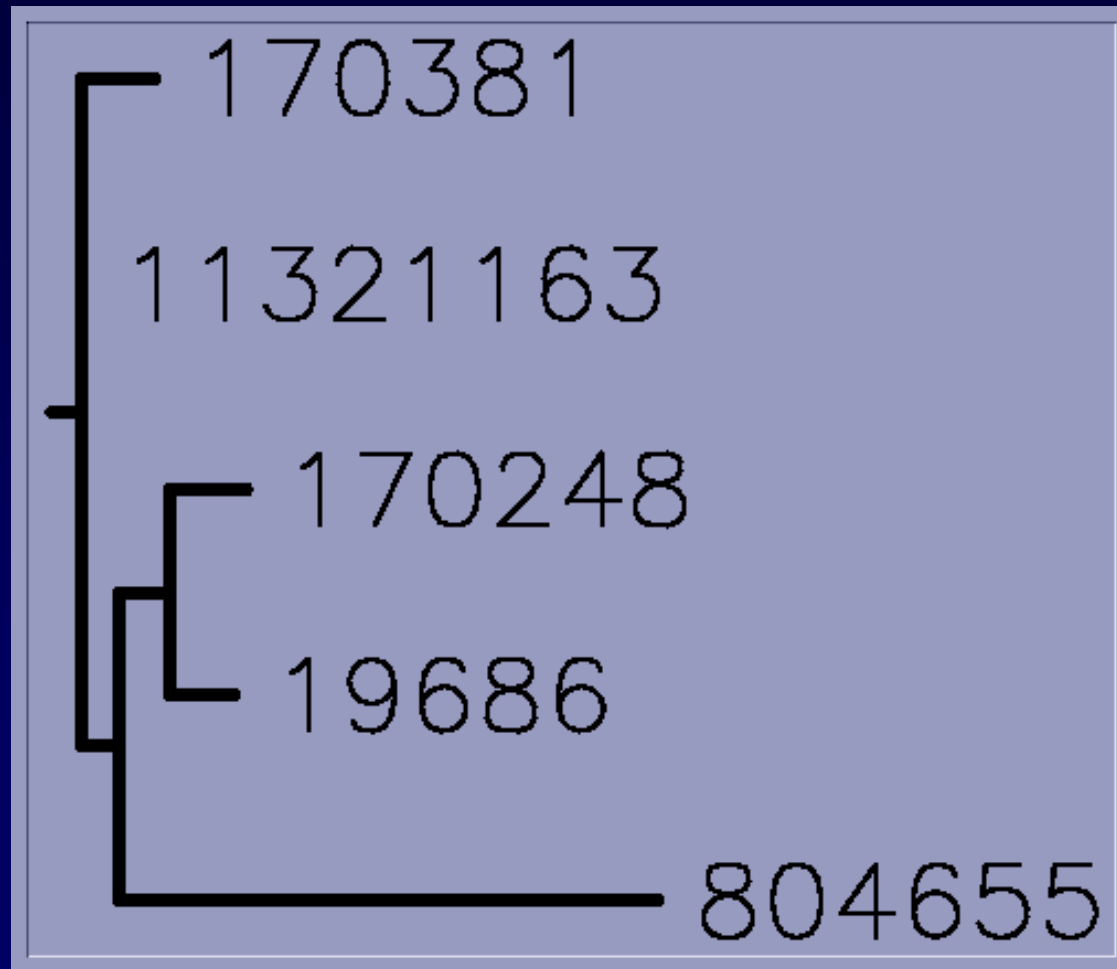
                2770      2780      2790      2800      2810      2820
79 AGAGGCTCAA...ATAGGTGT...TTGTTATGCAATCATGGCAACAAGCTTGGCATGAC 170381
1 T...ATCGGTGT...TTGGTATGCAATCATGGCAACAAGCTTGGCATGAC 11321163
2484 AGGGGCTCAATCAATAGGTGT...TTGCTATGCAATCATGGCAACAAGCTTGGCAATC 170248
1863 AGGGGCTCAATCAATAGGTGT...TTGCTATGCAATCATGGCAACAAGCTTGGCAATC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

                2830      2840      2850      2860      2870      2880
132 ATTCTGAAGTTATACAGC..TCTACAAGTCGAGAAACATTCGAAGACTCAGGCTTTATGA 170381
45 ATTCTGAAGTTATACAGC..TCTACAAGTCGAGAAACATTCGAAGACTCAGGCTTTATGA 11321163
2540 ATTCTGAAGTTATACAGC..TCTACAAGTCGAGAAACATTCGAAGACTCAGGCTTTATGA 170248
1919 ATTCTGAAGTTATACAGC..TCTACAAGTCGAGAAACATTCGAAGACTCAGGCTTTATGA 196886
2800 ACTTAGGGGCTGATTCGCTAAAGACGCACTGCGCAATATCCGCAGTCCGCAAAATGGCAG 804655

```




Analytické nástroje



Analytické nástroje

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

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

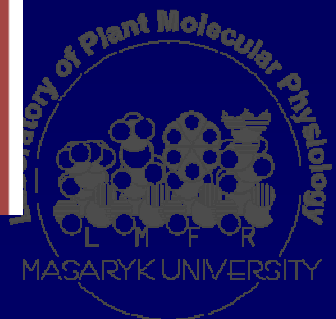

Primer 5

Primer 6

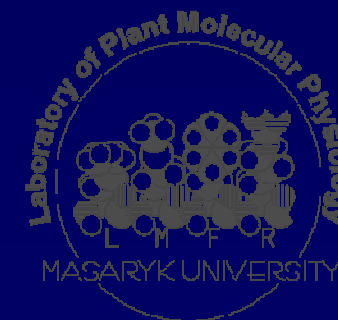
Primer 7

Primer 8

Annealing temperature

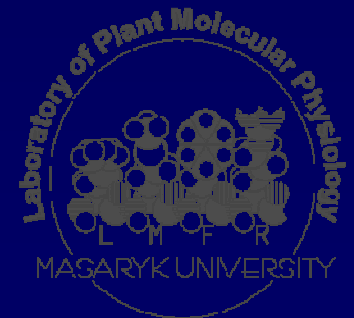


Analytické nástroje



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/>)

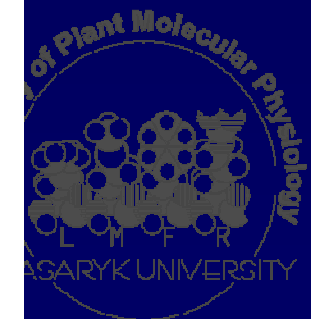
TIGR The Institute for Genomic Research

Software Home | Contact Us | TIGR Home

Software T

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

