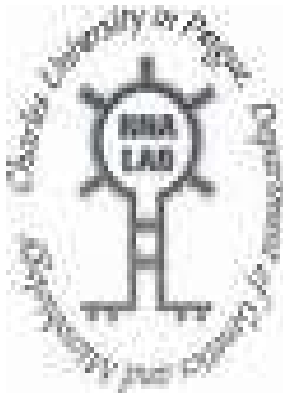


# **IRESite: Databáze experimentálně studovaných IRES struktur mRNA**

**Mokrejš M., Kříž V., Vopálenský V.,  
Mašek T., Pospíšek M.**

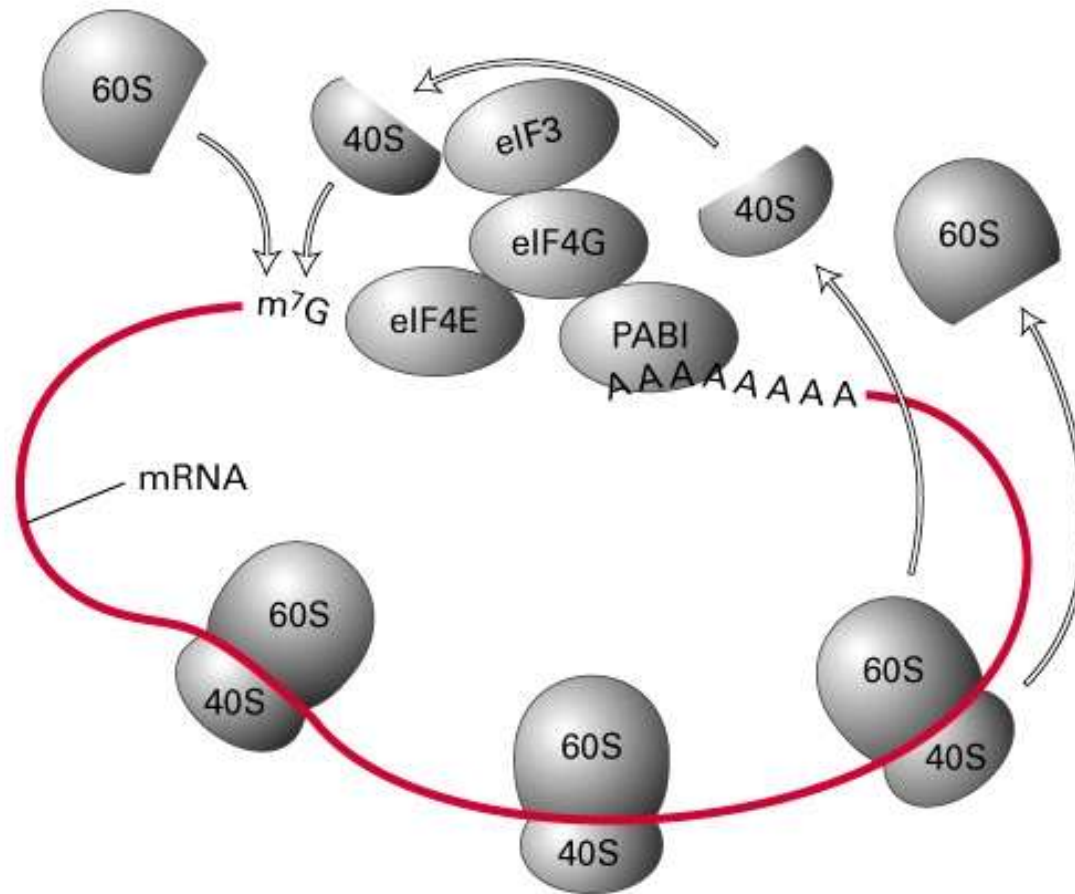


# **Internal ribosome entry site (IRES)**

- **Pouze u eukaryot**
- **Místo, kde ribozóm zahájí translaci mRNA**
- **Nezávisle na eukaryotické čepičce a eIF4F**
- **Zvláštní sekundární struktura mRNA, většinou v 5' UTR či lehce přesahující do kódující oblasti**
- **54 virových IRES: virus dětské obrny, virus hepatitidy A a C, HIV, BVDV, CSFV, FMDV**
- **70 buněčných mRNA také obsahuje IRES**

# Iniciace translace u eukaryot

Lodish (2000) Molecular biology of the cell



# **Buněčné mRNA obsahující IRES**

- **Růstové faktory (FGF2, PDGF, VEGF)**
- **Transkripční faktory (antennapedia, ultrabithorax, MYT2, NF-kB reprimují faktor NRF, ...)**
- **Onkogeny (c-myc, Pim-1, ...)**
- **transportéry/receptory (Cat1, ...)**
- **Translační faktory (eIF4G, eIF4GI, DAP5)**
- **Aktivátory apoptózy (Apaf1, ARC, RC3, ...)**
- **Další (BiP, La autoantigen, ornithine decarboxylase, ...)**

# IRESDb

- **IRESDb: The French Institute of Health and Medical Research (Bonnal et al., 2003)**
- <http://ifr31w3.toulouse.inserm.fr/IRESDatabase>
- **nedostatek informací o provedených experimentech**
- **nemožnost získat informaci o interagujících proteinech, sekundárních strukturách**
- **Celkový dojem: lepší přehled literatury**

# **Cíle projektu [www.iresite.org](http://www.iresite.org)**

- **Roztřídit publikované výsledky**
- **Překontrolovat výsledky**
- **Umožnit snadno porovnávat výsledky experimentů**
- **Umožnit uchovávat nová experimentální data**
- **Vhodně prezentovat sekvence, struktury a vškeré podstatné parametry jednotlivých experimentů**
- **Prohledávat podle různých kritérií**
- **=====> RELAČNÍ DATABÁZE**

# **SCHÉMA DATABÁZE**

- **221 sloupců ve 25 tabulkách**
- **92 biologických parametrů**
- **10 parametrů pro citace**
- **129 parametrů tvoří interní vazby**

# TYPOVÉ ROZLOŽENÍ

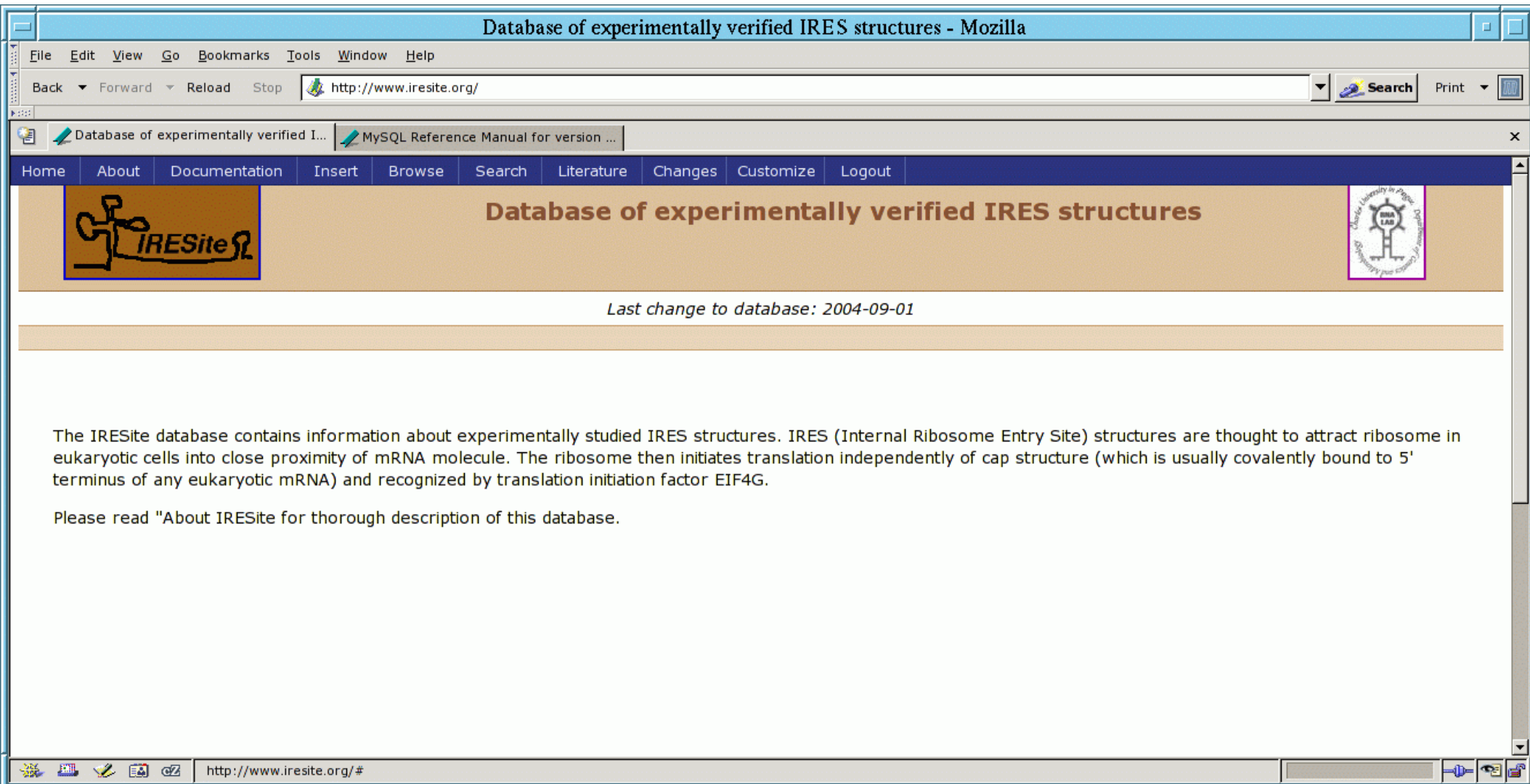
- **138 položek \*INT**
- **19 položek ENUM**
- **54 položek \*CHAR**
- **9 položek BLOB**
- **1 položka DATE**



# Části www rozhraní

- **Veřejná:**
  - prohlížení
  - prohledávání pomocí filtrů
- **Interní:**
  - **úprava a potvrzení položek ==> zveřejnění**
  - **stažení položek z veřejné části**
  - **smazání položek**

# Hlavní stránka: [www.iresite.org](http://www.iresite.org)



The screenshot shows a Mozilla browser window titled "Database of experimentally verified IRES structures - Mozilla". The address bar contains "http://www.iresite.org/". The browser's menu bar includes "File", "Edit", "View", "Go", "Bookmarks", "Tools", "Window", and "Help". The toolbar shows "Back", "Forward", "Reload", "Stop", "Search", and "Print". The page content features a navigation menu with "Home", "About", "Documentation", "Insert", "Browse", "Search", "Literature", "Changes", "Customize", and "Logout". The main heading is "Database of experimentally verified IRES structures", accompanied by the IRESite logo on the left and a circular logo of the "RNA LAB" at the "Charles University in Prague" on the right. Below the heading, it states "Last change to database: 2004-09-01". The main text describes the IRESite database as containing information about experimentally studied IRES structures, which are thought to attract ribosomes into close proximity of the mRNA molecule. It notes that the ribosome then initiates translation independently of the cap structure (usually covalently bound to the 5' terminus of any eukaryotic mRNA) and is recognized by the translation initiation factor EIF4G. A link is provided to read "About IRESite" for a thorough description of the database. The status bar at the bottom shows the URL "http://www.iresite.org/#".



Database of experimentally verified IRES structures - Mozilla

File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop <http://www.iresite.org/> Search Print

Database of experimentally verified I... MySQL Reference Manual for version ...

Home About Documentation Insert Browse Search Literature Changes Customize Logout

 **Database of experimentally verified IRES structures** 

*Last change to database: 2004-09-01*

The IRESite database contains information about experimentally studied IRES structures. IRES (Internal Ribosome Entry Site) structures are thought to attract ribosome in eukaryotic cells into close proximity of mRNA molecule. The ribosome then initiates translation independently of cap structure (which is usually covalently bound to 5' terminus of any eukaryotic mRNA) and recognized by translation initiation factor EIF4G.

Please read "About IRESite for thorough description of this database.

<http://www.iresite.org/#>

# Co extrahujeme z publikací?

- **Informace o plazmidu, promotoru, sekvenci, struktuře**
- **Pozici a typu reportérových genů**
- **Úrovní exprese během experimentu**
- **Organismus:**
  - **Kóduje mRNA**
  - **Nebo jeho IRES bylo někam naklonováno**
  - **Nebo v něm došlo k translaci studované mRNA**
  - **Nebo v něm interagují proteiny se studovanou mRNA**

# Ukázkový přehled položek

## Pending entries to be revised and committed

<input type="checkbox"/>	IRESite IRES Exp Id	IRES name	Gene name	IRES from gene	Sequence of mRNA is	IRESite IRES Seq Id	IRESite IRES 2Dstruct Id	Translated in	Translation method	Type of mRNA	mRNA length	IRES length	Origin of this natural mRNA	Origin of IRES in this engineered mRNA	E-mail
<input type="checkbox"/>	2	HCV-1b		221513	engineered	4		<i>Homo sapiens BHK-21</i>	<i>in vivo</i>	linear bicistronic	2967	339		viral	Vaclav Vopalensky
<input type="checkbox"/>	4	HCV-1b		221513	engineered	4		<i>Homo sapiens HuH7</i>	<i>in vivo</i>	linear bicistronic	2967	339		viral	Vaclav Vopalensky
<input type="checkbox"/>	1	mmHIF1a	Hif1a		natural	2		<i>Mus musculus</i>		linear monocistronic	3973	257	genomic		Martin Pospisek

# Struktura jednotlivých položek

- **2 typy položek:**
  - **buněčné/virové mRNA obsahující IRES (natural)**
  - **uměle vytvořené mRNA obsahující IRES (engineered):**
    - **obsažená IRES je původem buněčná/virová**
    - **obsažená IRES je uměle vytvořena (typicky náhodná sekvence)**



## Database of experimentally verified IRES structures



*Last change to database: 2004-09-01*

Secondary structures can be recorded using "bracket" style. Consider the following example:

```

stem axis
  v
  |
  |
  TTT
  G | A
  A | C
  A | T
  A | T
  T | A
  C | G
  A |
  G | C
  C | G
  G | C
5'-AAAAA | CCCCC-3'
  |

```

The above structure represented in bracket notation:

```

.....(((.((((.....)))))).....
-----^ bulge (protruding A base at position 9)

```

The following simple rule can be used even for highly branched structures: paired bases being on the left-hand side from stem axis should be represented by a left bracket, unpaired bases using a dot and paired bases being on the right-hand side of a stem should be represented by right bracket.



## Database of experimentally verified IRES structures



Last change to database: 2004-09-01

### Nucleic acid

IRESite\_Id:6

#### Date

2004-07-28

Quality of mRNA sequence reported here.

[our\\_best\\_guess](#)

Occurrence of biopolymer studied in this experiment

[engineered](#)

Origin of IRES in this engineered mRNA

[viral](#)

Sequence of experimentally studied mRNA (as +DNA), i.e. fully spliced mRNA without promoter or vector sequences.

ATGGAGAAAAAATCACTGGATATACCAACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTTCAGTCAGTTGCTCAATGTACCTATAACAGACCGTTTCAGCTGGATATTACGGCCTTTTAAAGACCGTAAAGAAAAAATAAGCACAGGTTT  
 ^1 \_\_\_\_\_ ^10 \_\_\_\_\_ ^20 \_\_\_\_\_ ^30 \_\_\_\_\_ ^40 \_\_\_\_\_ ^50 \_\_\_\_\_ ^60 \_\_\_\_\_ ^70 \_\_\_\_\_ ^80 \_\_\_\_\_ ^90 \_\_\_\_\_ ^100 \_\_\_\_\_ ^110 \_\_\_\_\_ ^120 \_\_\_\_\_ ^130 \_\_\_\_\_ ^140 \_\_\_\_\_ ^150 \_\_\_\_\_ ^160 \_\_\_\_\_

IRESite\_Seq\_Id:31

Description (one-line) of this mRNA (may differ from description in [GenBank!](#)) (one-line)

[BVDV genome IRES](#)

Function of this biopolymer in experiment was

[mRNA](#)

Biopolymer type and strand

[+DNA](#)

OPTIONAL: RFAM entry for this mRNA

OPTIONAL: Remarks (255 chars maximum)

#### Reference

\* Chon S. K., Perez D. R., Donis R. O. (1998) 370-382. Virology. 251(2): 370-382

\* Colett M. S., Larson R., Gold C., Strick D., Anderson D. K., Purchio A. F. (1988) 191-199. Virology. 165(1): 191-199



## Database of experimentally verified IRES structures



*Last change to database: 2004-09-01*

### Search in IRES

**Gene name**

**Organism**

**Reporter gene**

**Promoter name**

**description**

**Authors (for example Pestova T.V., Hellen C.U.)**

**Journal (Mol. Cel. Biol., PNAS U.S.A., J. Cell. Biol. ... or type in Unpublished)**

**Title**

Search



# Database of experimentally verified IRES structures



Last change to database: 2004-09-01

## Description of mRNA

### Functional status of IRES in this mRNA

- true\_IRES
- possibly\_IRES
- possibly\_not\_IRES
- defective\_IRES
- not\_IRES
- unclear
- not\_enough\_info

AAACAACCTCGGCCTGGCGGGTCGCA GCGCCGCGGGGACCGCGAGGCGACCGGGAGCCGCTGGGACCCAGGCTGCGCGCCCTCGGCCTGCCCGGGAGCCCGGCCAGTCGGAGTCCCGGGCCGAGCGGGGCTTGCCCTCCCTCTCAGCGTTCCTCTGTGCGCAGTGC

21 \_\_\_\_\_ ^31 \_\_\_\_\_ ^41 \_\_\_\_\_ ^51 \_\_\_\_\_ ^61 \_\_\_\_\_ ^71 \_\_\_\_\_ ^81 \_\_\_\_\_ ^91 \_\_\_\_\_ ^101 \_\_\_\_\_ ^111 \_\_\_\_\_ ^121 \_\_\_\_\_ ^131 \_\_\_\_\_ ^141 \_\_\_\_\_ ^151 \_\_\_\_\_ ^161 \_\_\_\_\_ ^171 \_\_\_\_\_ ^181 \_\_\_\_\_

(rev.) within ORF1. Supposedly enter 'Bag1' to be consistent with previously entered data. If you type in a different value, also previously entered gene name will change

base range (first 3 bases should be ATG or equivalent, but do NOT include stop codon bases)

### One line description of first gene (cistron 1)

Mus musculus Bcl2-associated athanogene 1 (Bag1), mRNA.

### Evidence of IRES function

### Evidence of IRES existence

### Your mRNA:

CCGGTAGCCACCGTGCCTGTGAACAACAACCTCGGCCTGGCGGGTCGCA GCGCCGCGGGGACCGCGAGGCGACCGGGAGCCGCTGGGACCCAGGCTGCGCGCCCTCGGCCTGCCCGGGAGCCCGGCCAGTCGGAGTCCCGGGCCGAGCGGGGCTTGCCCTCCCTCTCAGCGTTCCTCTGTGCGCAGTGC

^1 \_\_\_\_\_ ^11 \_\_\_\_\_ ^21 \_\_\_\_\_ ^31 \_\_\_\_\_ ^41 \_\_\_\_\_ ^51 \_\_\_\_\_ ^61 \_\_\_\_\_ ^71 \_\_\_\_\_ ^81 \_\_\_\_\_ ^91 \_\_\_\_\_ ^101 \_\_\_\_\_ ^111 \_\_\_\_\_ ^121 \_\_\_\_\_ ^131 \_\_\_\_\_ ^141 \_\_\_\_\_ ^151 \_\_\_\_\_ ^161 \_\_\_\_\_ ^171 \_\_\_\_\_ ^181 \_\_\_\_\_

### Absolute start position of IRES structure in this mRNA

### Absolute stop position of IRES structure in this mRNA

### OPTIONAL: Remarks (255 chars maximum)

# Vkládání položek

- **pomocí HTML formulářů**
- **načítání dat o mRNA z Genbank přes Internet**
- **angl. scroll-bars umožňují:**
  - **jednotnost obsažených dat**
  - **zrychlení vkládání podobných položek**
- **načítání citací z Pubmed přes Internet (PMID)**
- **vše funguje ze všech www prohlížečů**

# **Aktuální stav**

- **Lze vkládat oba typy položek (natural, engineered)**
- **Lze upravit a následně potvrdit položky**
- **Lze prohlížet zveřejněné položky**
- **Lze stáhnout z veřejné části položky**
- **Lze smazat jakékoli položky**

# **Ponaučení ze špatného designu databáze I.**

- **Poloprázdné tabulky**
  - **Řešení: větší množství malých tabulek**
- **Index sloupců**
  - **Ne vždy je přítomnost indexu urychlením, ve sloupci by mely být alespoň 3 typy hodnot aby index měl smysl**
- **BLOB**
  - **Umístit do tabulek s minimem sloupců**

# **Ponaučení ze špatného designu databáze II.**

- **Příklady: sekvence, struktura, rozložení cistronů na polycistronické mRNA, organismus, citace či kombinace citací**
  - **Řešení: Duplikovaná data umístit do zvláštní tabulky a odkazovat se vícenásobně na stejný řádek**

# Multiplicita v popisu vícecistronické mRNA

- `CREATE TABLE cistron (`
- `cistron_id int(11) unsigned NOT NULL auto_increment,`
- `base_range varchar(10) default NULL,`
- `description blob,`
- `gene_name varchar(10) default NULL,`
- `function blob,`
- `version int(1) unsigned default NULL,`
- `PRIMARY KEY (cistron_id),`
- `KEY gene_name (gene_name)`
- `) TYPE=InnoDB;`
  
- `INSERT INTO cistron VALUES (2, '15-950', 'renilla luciferase', 'Rluc', '', NULL);`
- `INSERT INTO cistron VALUES (3, '1302-2951', 'firefly luciferase', 'FFluc', '', NULL);`

- CREATE TABLE mapp\_cistron (
  - mapp\_cistron\_id int(11) unsigned NOT NULL auto\_increment,
  - cistron\_1\_id int(11) unsigned default NULL,
  - cistron\_2\_id int(11) unsigned default NULL,
  - cistron\_3\_id int(11) unsigned default NULL,
  - cistron\_4\_id int(11) unsigned default NULL,
  - cistron\_5\_id int(11) unsigned default NULL,
  - PRIMARY KEY (mapp\_cistron\_id),
  - KEY cistron\_1\_id (cistron\_1\_id),
  - KEY cistron\_2\_id (cistron\_2\_id),
  - KEY cistron\_3\_id (cistron\_3\_id),
  - KEY cistron\_4\_id (cistron\_4\_id),
  - KEY cistron\_5\_id (cistron\_5\_id)
- ) TYPE=InnoDB;
  
- INSERT INTO mapp\_cistron VALUES (2,2,3,NULL,NULL,NULL);

***Martin MOKREJŠ***  
***Viničná 5***  
***128 44, PRAHA 2***



***el. pošta: mokrejsm***  
***zavináč***  
***ribosome.natur.cuni.cz***  
***http ://www.iresite.org***

***Poděkování:***

***Vítězslav Kříž (TU Liberec)***  
***Vašek Vopálenský***  
***Tomáš Mašek***  
***Martin Pospíšek***  
***studenti laboratoře biochemie RNA***