

# Úloha C1 – studijní materiál

## Plasmidy

Plasmid je označení pro samostatně se replikující kruhovou molekulu DNA velikosti zhruba 1 až 200 kb. Pro účely této úlohy si představíme plasmidy z pohledu rekombinantních technologií a rozdělíme si je na 2 typy: plasmidy klonovací a expresní

### 1. Klonovací plasmidy

Klonovací plasmidy jsou uzpůsobeny k tomu, aby byly z buněk snadno izolovatelné v dostatečném množství a kvalitě. K tomu je třeba použít kmeny buněk vhodné ke klonování. Kmeny určené pro klonování jsou pro tyto účely geneticky upraveny: jsou relativně nepatogenní, nedochází k degradaci plasmidové DNA působením nukleas, plasmidy jsou v buňkách přítomny v mnoha kopiích.

Klonovací plasmid se podle účelu skládá zejména z následujících částí:

- Replikační počátek (*ori*) – v tomto místě začíná replikace plasmidu
- Selekční marker – většinou se jedná o rezistenci k antibiotiku. Kultivací buněk v přítomnosti tohoto antibiotika zajistíme, že v kultuře budou růst pouze buňky obsahující cílový plasmid. Přítomnost antibiotika v mediu je pro buňku navíc znamením, že daný plasmid je pro její přežití důležitý. Buňka si ho tedy namnoží do mnoha kopií.
- Multiklonovací místo (MCS), někdy též polylinker – místo, do něž je vložen gen zájmu. MCS obsahuje rozpoznávací místa pro restriční endonukleasy, aby bylo možno s tímto genem manipulovat.

Dále může klonovací plasmid obsahovat tzv. reportérový gen, který nám pomáhá zjistit, zda došlo k začlenění DNA úseku do multiklonovacího místa.

### 2. Expresní plasmid

Expresní plasmid je uzpůsoben k produkci proteinů. Produkce proteinů probíhá v expresních kmenech buněk. Stejně jako u klonovacích kmenů je i zde komerčně dostupná široká škála kmenů lišících se například schopností tvořit disulfidické můstky, produkci proteinu je tak možné optimalizovat podle konkrétní potřeby.

Expresní plasmid obsahuje stejné elementy jako klonovací plasmid, ale má navíc elementy, které se vztahují k expresi genů. Patří sem promotor – část DNA, kde se zahajuje transkripce genu. Míra transkripce by navíc měla být externě regulovatelná, např. přidávkem induktorů exprese. Vyprodukovaný protein je dále třeba z buňky získat. Izolaci proteinu mohou usnadnit např. afinitní značky, které se k proteinům přidávají. Velice rozšířená je tzv. histidinová kotva, kdy se na jeden konec proteinu přidá alespoň 6 histidinů. Některé značky také mohou zvyšovat rozpustnost a stabilitu proteinu a značkou může být i další protein.

## Restriční endonukleasy

Restriční endonukleasy jsou hydrolytické enzymy rozpoznávající specifické sekvence (často palindromy) nukleových kyselin. V tomto místě dochází k hydrolýze fosfodiesterové vazby. Štěpení může být symetrické (tzv. tupé konce) nebo může na jednom z řetězců vznikat různě dlouhý převis (tzv. lepivé konce). Pro klonování je důležité, že štěpením dvou různých molekul DNA jedním enzymem vznikají konce, které jsou navzájem

komplementární a mohou být tedy snadno spojeny. Pojmenovávání těchto enzymů se řídí zvláštní nomenklaturou. Např. enzym **EcoRI**: **E** znamená, že byl enzym izolován z bakterie rodu *Escherichia*, **co** znamená druh *coli*, **R** je kmen RY13 a **I** značí, že jde o první nalezený restriční enzym v tomto rodu bakterie. V této úloze budeme technické informace o těchto enzimech hledat na stránkách firmy **New England Biolabs**, kde lze najít různé šikovné nástroje pro práci s restričními endonukleasami.

Restriční reakce probíhají většinou při 37 °C po dobu několika hodin a v prostředí, které je třeba nastavit pomocí různých pufrů. Volbou vhodného pufru lze zajistit optimální aktivitu enzymu a v mnoha případech lze výhodně štěpit dvěma různými enzymy v jedné reakci.

## Elektroforéza

Elektroforéza patří mezi základní biochemické separační metody. V této úloze se zaměříme na agarosovou gelovou elektroforézu, která se používá hlavně k separaci nukleových kyselin.

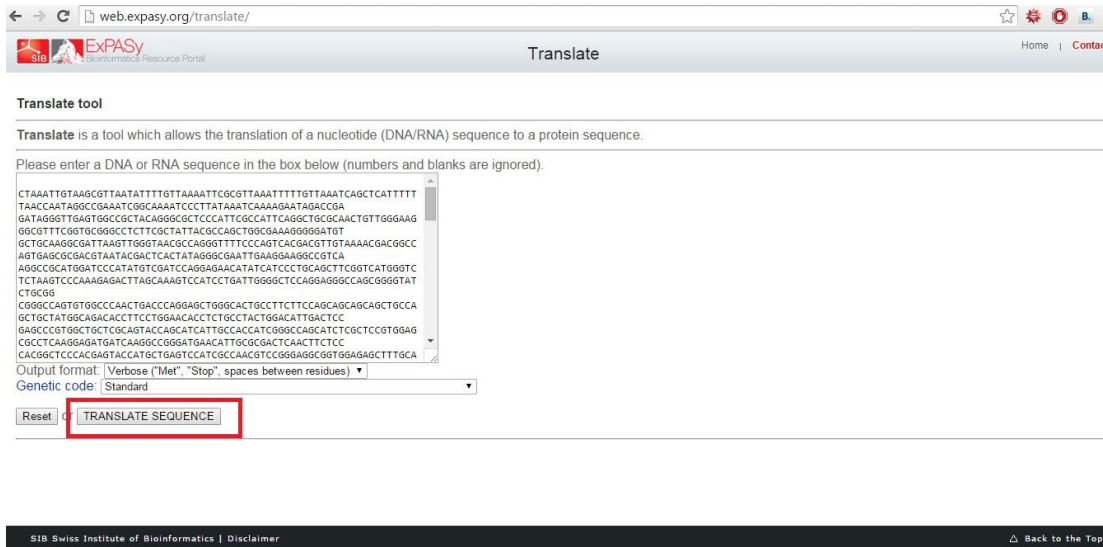
Molekula DNA i RNA je díky přítomnosti fosfátových skupin záporně nabitá a zavedeme-li ji tedy do stejnosměrného elektrického pole, začne se pohybovat k anodě. Tento děj ale musíme zasadit do nějakého stabilizujícího prostředí, které zabrání difuzi. V tomto případě použijeme polysacharid agarosu, která po rozvaření a zchlazení vytvoří gel. V gelu jsou přítomny póry (jejich velikost lze ovlivnit koncentrací agarosu), skrz které musí molekuly nukleových kyselin při migraci procházet. Přitom platí, že větší molekula bude muset překonávat větší odpor a během separace se tedy bude zbrzdovat. Po nějaké době se tedy fragmenty přítomné ve vzorku rozdělí podle velikosti – menší fragmenty doputují dál a naopak. Abychom mohli odhadnout velikost separovaných fragmentů, musíme na gel nanést také standard, tj. směs fragmentů o známé délce. Tyto standardy jsou opět komerčně dostupné a volí se podle očekávané velikosti separovaných molekul.

Separované molekuly je také třeba nějak vizualizovat. K tomu se používají např. fluorescenční značky, které se na nukleovou kyselinu navážou a společně potom pod UV světlem fluoreskují. Často je používán např. **ethidium-bromid**, který je ale velice silným mutagenem a je proto nahrazován jinými netoxickými barvivy (GelRed, SYBR Safe, a.j.) Kromě fluorescenčního značení lze použít také např. radioaktivní značení pomocí <sup>32</sup>P.

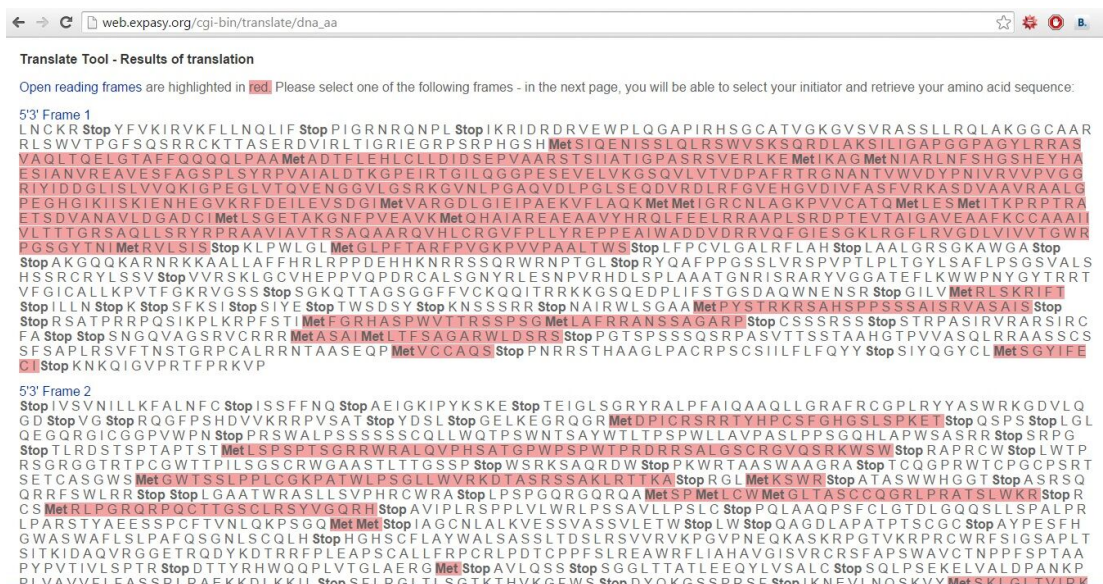
## Bioinformatické nástroje

Pokud známe sekvenci genu či proteinu, se kterým pracujeme, můžeme pomocí různých bioinformatických nástrojů zjistit mnoho užitečných informací, které při klonování použijeme. Celou řadu jich naleznete například na stránkách <http://expasy.org/>.

Pro přeložení genu do polypeptidového řetězce můžeme využít nástroj **Translate**. Do pole pro sekvenci byla vložena nukleotidová sekvence celého plasmidu pMK-RQ s naším genem. Program nyní bude v sekvenci hledat otevřené čtecí rámce (ORF-open reading frame), jejichž produktem mohou být různé proteiny.



Program po stisknutí tlačítka TRANSLATE SEQUENCE nabídne několik ORF, ze kterých si můžeme vybrat. Jelikož víme, že plasmid nese v MCS gen pro jeden protein, má smysl se zaměřit pouze na ORF, které obsahují nejdelší produkt. V tomto případě se jedná o první nalezený ORF.



Po výběru ORF a následném výběru začátku polypeptidového řetězce nástroj Translate poskytne sekvenci polypeptidu, se kterou lze dále pracovat.

web.expasy.org/cgi-bin/translate/dna\_sequences?/work/expasy/tmp/http/seqdna.15460,1,127

Translate

ID VIRT17988 Unreviewed; 574 AA.  
AC VIRT17988;  
DE Translation of nucleotide sequence generated on ExPASy  
DE on 03-oct-2015 by 82.144.148.106.  
CC -!- This virtual protein sequence will automatically be deleted  
CC from the server after a few days.  
DR SWISS-2DPAGE; VIRT17988; VIRTUAL.  
SQ SEQUENCE 574 AA; 38438896832932F5 CRC64.  
MSIQENISSL QLRSHWKSQ RDLAKSLIG APGGPAGYLR RASVAQLTQE LGTAFFQ000  
LPAAMADTFL EHLCLLDIDS EPVAARSTSI IATIGPASRS VERLKENIKA GNNIARLNFS  
HGSHEHYAES IAWREAVES FAGSPLSYRP VAIALDTKSP EIRTGILQGG PESEVELKVG  
SQQLVTPDPA FRTRGQANTV WNDYPIEVVYR VPVGGRIYED DGLSLVQPK IPEGLSLTGV  
ENGVLSGRK GVNLPGAQVD LPLGSEQDVR DLRFGEVGV DIVFASFVRK ASDVAVRAA  
LGPFGHGIKI ISKIENHEGV KRDFEILEVS DGINVARGDL GIEIPEAKVF LAQMPLIGRC  
NLAKPQVCA TQNLSEHITK PRPTRAETSD VANAVLDGAD CIIMLSGETAK GNPVPEAVKH  
QNAIAREEA AVYHRLQFEE LRRAPLSRD PTEVTATGAV EAFFKCCAAA IIVLTTRGS  
AQLLSRYRPR AAVIATVTRSA QAAQVHLER GVFFLLYREP PEAINADVDV RRVQFGIESS  
KLRGFLRVGD LVIVVGNRPP GSGYTNIRRV LSIS

//  
Sequence in FASTA format

[BLAST](#) BLAST submission on ExPASy/SIB

Sequence analysis tools: ProtParam, ProtScale, Compute pI/Mw, PeptideMass, PeptideCutter,

ScanProsite

Direct Submission to SWISS-MODEL

V poslední části úlohy se krátce seznámíme s nástrojem **BLAST** (Basic Local Alignment Search Tool). Tento nástroj po vložení sekvence nukleové kyseliny nebo proteinu prohledává sekvenční databáze a zjišťuje, které známé sekvence se našemu zadání nejvíce podobají. Tyto sekvence k sobě následně přiloží a vypočítá homologii a skóre, tedy to, jak moc se sekvence shodují.

Použijeme BLAST pro proteinovou sekvenci. Do rámečku byla vložena sekvence náhodného proteinu. Před začátkem vyhledávání je možno vyhledávání zúžit na jednotlivé databáze nebo organismy, můžeme také vybrat algoritmus, podle kterého bude vyhledávání probíhat. Pro naše účely můžeme nastavení ponechat na původních hodnotách.

blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE\_TYPE=BlastSearch&LINK\_LOC=blasthome

BLAST® Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

My NCBI [Sign In] [Register]

NCBI/BLAST/blastp suite Standard Protein BLAST

blastn blastp **blastx** tblastn tblastx

BLASTP programs search protein databases using a protein query. [more...](#) [Reset page](#) [Bookmark](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s)  Clear Query subrange

From

To

Or, upload file  Soubor nevybrán

Job Title

Enter a descriptive title for your BLAST search

Align two or more sequences

Choose Search Set

Database

Organism   Exclude

Optional

Exclude  Models (XM/XP)  Uncultured/environmental sample sequences

Optional

Entrez Query

BLAST začne po odeslání sekvence prohledávat databáze (může to chvíli trvat). Už během prohledávání může program v sekvenci detekovat určité motivy jako např. domény nebo superrodiny. V tomto případě program už po 5 sekundách poznal, že protein patří do rodiny hexokinas.



blast.ncbi.nlm.nih.gov/Blast.cgi

BLAST® Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

My NCBI [Sign In] [Register]

NCBI/BLAST/blastp suite/ Formatting Results - 0Z75E851014 [Formatting options]

Job Title: Protein Sequence (917 letters)

Putative conserved domains have been detected, click on the image below for detailed results.

Query seq. nucleotide binding site 125 250 375 500 625 750 875 917

Specific hits

Superfamilies

Multi-domains

Request ID: 0Z75E851014

Status: Searching

Submitted at: Sat Oct 3 05:15:26 2015

Current time: Sat Oct 3 05:15:31 2015

Time since submission: 00:00:05

This page will be automatically updated in 2 seconds

BLAST is a registered trademark of the National Library of Medicine.

Copyright | Disclaimer | Privacy | Accessibility | Contact | Send feedback

NCBI | NLM | NIH | DHHS

Čekání na blast.ncbi.nlm.nih.gov...

Po skončení prohledávání program ukáže barevný diagram shod a uvede také jednotlivé homologní sekvence, které našel. V tomto případě se jednalo o sekvenci lidské hexokinasy-2, která správně skončila na prvním místě se 100% sekvenční identitou. BLAST dále našel vysokou homologii např. s predikovanými hexokinasami z paviána, gorily nebo kosatky ☺.

blast.ncbi.nlm.nih.gov/Blast.cgi

Select: All None Selected:0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> hexokinase-2 [Homo sapiens]	1891	1891	100%	0.0	100%	NP_000180.2
<input type="checkbox"/> PREDICTED: hexokinase-2 isoform X1 [Pan troglodytes]	1889	1889	100%	0.0	99%	XP_001162535.1
<input type="checkbox"/> Human hexokinase II cDNA [Homo sapiens]	1889	1889	100%	0.0	99%	CAA86511.1
<input type="checkbox"/> unnamed protein product [Homo sapiens]	1887	1887	100%	0.0	99%	BAF83046.1
<input type="checkbox"/> HK2 [synthetic construct]	1881	1881	100%	0.0	99%	AIC62504.1
<input type="checkbox"/> hexokinase II [Homo sapiens]	1880	1880	100%	0.0	99%	CAA86476.2
<input type="checkbox"/> PREDICTED: hexokinase-2 [Gorilla gorilla gorilla]	1880	1880	100%	0.0	99%	XP_004029536.1
<input type="checkbox"/> PREDICTED: hexokinase-2 isoform X1 [Homo sapiens]	1878	1878	100%	0.0	98%	XP_005264337.1
<input type="checkbox"/> PREDICTED: hexokinase-2 [Nomascus leucogenys]	1877	1877	100%	0.0	99%	XP_003268759.1
<input type="checkbox"/> PREDICTED: hexokinase-2 [Papio anubis]	1874	1874	100%	0.0	99%	XP_003908903.1
<input type="checkbox"/> PREDICTED: hexokinase-2 [Chlorocebus sabaeus]	1871	1871	100%	0.0	99%	XP_007968290.1
<input type="checkbox"/> PREDICTED: hexokinase-2-like isoform 2 [Macaca mulatta]	1870	1870	100%	0.0	99%	XP_001111706.1