# Cvičení - Bioinformatické databáze

**Srpkovitá anémie (Sickle Cell Disease)**

Srpková anémie je genetická nemoc, která je nejčastější u Italů, Řeků, Afrických a Latinských Američanů. Projevuje se změnou tvaru buněk uvnitř červených krvinek. Buňky jsou pak méně flexibilní a mohou uvíznout v kapilárách, což způsobí blokaci toku krve do orgánů. Lidé se srpkovou anémií mohou mít symptomy, jako jsou například: zažloutlé oči a kůže, bledá pleť, bolesti kostí a kloubů, anémie nebo opožděný růst. Dožívají se věku 40 – 60 let.

**Hledání v databázích nukleotidových sekvencí**

Vyhledejte mRNA lidského genu beta-globinu (Homo sapiens hemoglobin, beta, HBB), který je součástí většího proteinu hemoglobinu. Použijte databázi GenBank. Určete a stáhněte kódující oblasti tohoto genu a sekvenci proteinu, který je tímto genem kódován.

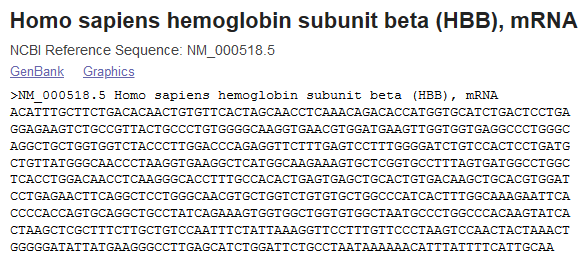
1. **Jaký přístupový kód má tento gen?**

ACCESSION: NM\_000518

URL: https://www.ncbi.nlm.nih.gov/nuccore/NM\_000518.5

1. Určete kódující sekvenci genu beta-globin a stáhněte ji ve FASTA formátu.

<https://www.ncbi.nlm.nih.gov/nuccore/NM_000518.5?report=fasta>



1. Jaký je přístupový kód proteinové sekvence kódované genem beta-globin?

NP\_000509 - https://www.ncbi.nlm.nih.gov/protein/4504349

1. Stáhněte sekvenci proteinu kódovaného genem beta-globin ve FASTA formátu.

URL: https://www.ncbi.nlm.nih.gov/protein/P68871.2?report=fasta



1. Stáhněte článek věnující se genetickému onemocnění beta-talasemie (beta-thalassemia), jenž je důsledkem mutace v lidském hemoglobinu.

**Hledání v databázích proteinových sekvencí**

Vyhledejte záznam o proteinu beta-globin (Homo sapiens hemoglobin, beta, HBB) v proteinové databázi UniProtKB/Swiss-Prot.

1. Jaký přístupový kód má tento protein?

Kód: P68871

<https://www.uniprot.org/uniprot/P68871>

1. Z kolika aminokyselin je tento protein tvořen?

147 aminokyselin

1. Jakou funkci má tento protein?

Zajišťuje přenos kyslíku z plic

1. Do jaké rodiny tento protein patří?

1.A.107.1.2 the pore-forming globin (globin) family

1. Jaký efekt má mutace E7V?

Produkce HbS (spojené s nemocí srpkové anémie)

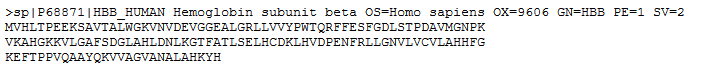
URL: <https://www.ncbi.nlm.nih.gov/clinvar/variation/15333/>

1. Zjistěte RS identifikátor pro mutaci E7V.

rs334

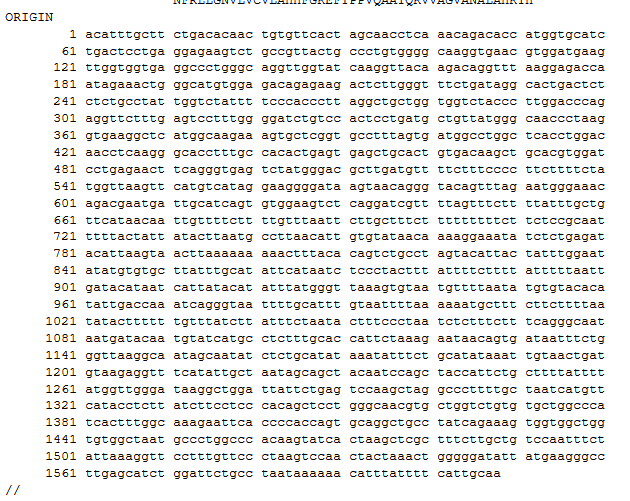
1. Stáhněte sekvenci tohoto proteinu ve FASTA formátu.

<https://www.uniprot.org/uniprot/P68871.fasta>



1. Stáhněte z GenBank DNA sekvenci, která kóduje tento protein.

<https://www.ncbi.nlm.nih.gov/nuccore/NC_000011.10?report=genbank&from=5225464&to=5227071&strand=true>



Informace o SNP mutacích

1. Ověřte patogenicitu v databázi Clinvar.

[https://www.ncbi.nlm.nih.gov/clinvar?term=((30372[AlleleID])OR(30214[AlleleID]))](https://www.ncbi.nlm.nih.gov/clinvar?term=((30372%5bAlleleID%5d)OR(30214%5bAlleleID%5d)))

1. Povšimněte si rozdílných pozic v různých verzích genomových map.

Pozice 5227002 v GRCh38.p12

Pozice 5248232 v GRCh37.p13

1. V jakém regionu lidského genomu se mutace nachází?

Kódovací region

**Porovnání textových vyhledávacích systémů**

Prohlédněte si seznam získaných výsledků. Nalezněte záznam spojený s mutací beta-globinu pro nemoc alpha-thalassemia v databázi OMIM.

1. Prohlédněte si seznam získaných výsledků. Nalezněte záznam spojený s mutací beta-globinu pro nemoc alpha-thalassemia v dtabázi OMIM.

<https://omim.org/entry/604131>

**Informace o genomových projektech**

V databázi Genomes OnLine Database (GOLD) zjistěte informace o dokončených a probíhajících genomových projektech.

1. Kolik bakteriálních a eukaryotických genomů bylo dosud osekvenováno a publikováno?

32 233 eukaryotních organismů

163 524 bakteriálních organismů

<https://gold.jgi.doe.gov/>

**Vizualizace struktur molekul**

V největší strukturní databázi PDB hledejte informace o proteinu beta-globin.

1. V databázi PDB najděte článek popisující lidský beta-globin (tip: PDB-101, human hemoglobin). Podívejte se na animaci ukazující rozdíl mezi okysličenou a neokysličenou verzí. V článku nalezněte odkazy na PDB záznamy s okysličenou a neokysličenou verzí a také na verzi s genovou mutací způsobující srpkovitou anémii.

<https://pdb101.rcsb.org/motm/41>

1. Srovnejte rozlišení a R-faktory výše nalezených struktur. Kterou ze struktur lze považovat za nejkvalitnější.

|  |  |  |
| --- | --- | --- |
|  | R-faktor | Rozlišení |
| Okysličená verze: | 22.3 % | 2.1 A |
| Neokysličená verze: | 16 % | 1.74 A |
| Verze s genovou mutací: | 16.5 % | 2.05 A |

Za nejkvalitnější lze považovat neokysličenou verzi, jelikož má nejlepší rozlišení a nejnižší hodnotu R-faktoru.

1. Stáhněte pdb soubor libovolné struktury.

<https://files.rcsb.org/view/1BUW.pdb>

1. Stáhněte sekvenci ve formátu FASTA. Použijte staženou sekvenci pro vyhledání struktury Beta-Globinu (tip: advanced search).

Advanced options -> Sequence -> BLAST

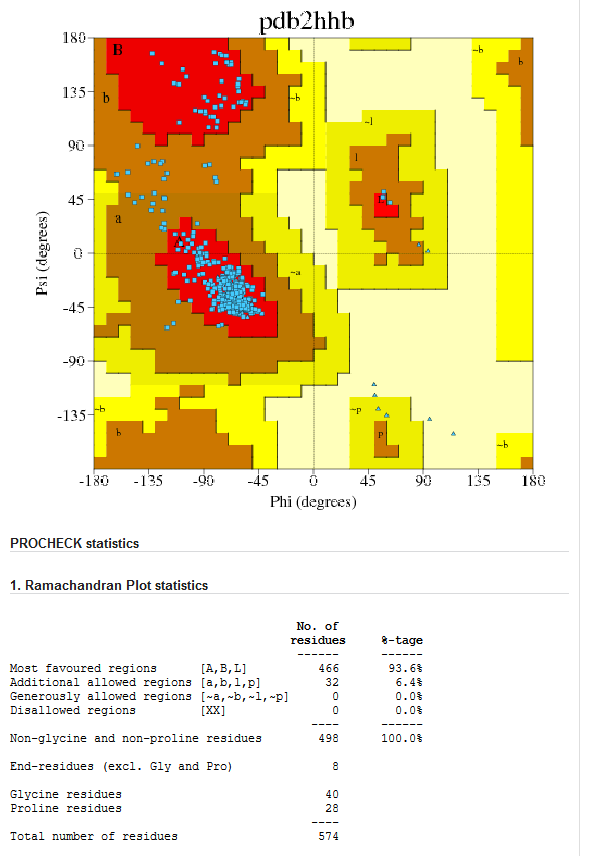
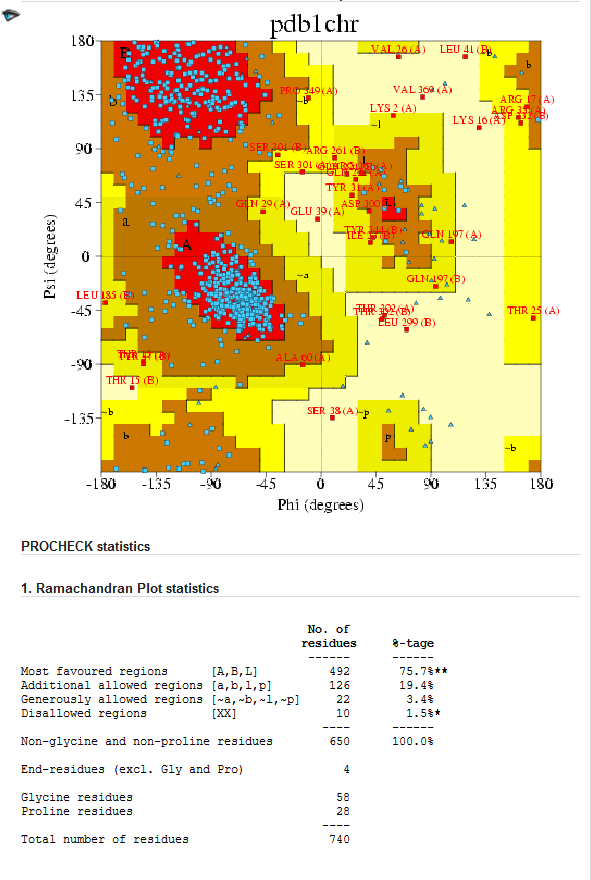
Databáze PDBSum

V databázi PDBSum vyhledejte strukturu 2HHB.

<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl>

1. Prohlédněte si Ramachandrův diagram. Jedná se o dobře definovanou strukturu? Srovnejte s jinými strukturami (např. 1CHR).

Ano, jedná se o dobře definovanou strukturu, jelikož má model v nejvýhodnějších regionech více než 90 %. Oproti tomu struktura 1CHR má pouze 75.7 % a tedy se jedná o nepravděpodobnou strukturu.

1. V záložce Proteins si prohlédněte informace o sekundární struktuře. Jaké z nich můžeme vyvodit závěry? Nevím.
2. Jaké další informace lze dohledat v PDBSum?

Titulek, zdroj, R-faktor, rozlišení, autory, odkazy, ligandy, rozštěpy atd.