

Materiály k tématu Revoluce, ne evoluce – jak se neutopit v záplavě dat aneb Jak voní déšť?

V článku Revoluce, ne evoluce – jak se neutopit v záplavě biologických dat byly popsány databáze, které ukládají velké množství zajímavých dat a jsou volně dostupné pro kohokoli s přístupem k internetu a základní znalostí angličtiny.

Zkuste se tedy ponořit do těchto databází a zjistit něco víc o zajímavých otázkách, které si můžeme položit. K cvičení budete potřebovat zařízení připojené k internetu, ideálně počítač. Většina zdrojů je k dispozici v anglickém jazyce. Pracujte ve dvou- až tříčlenných skupinách a diskutujte postup v rámci skupiny. Můžete pracovat i samostatně, ale bude to méně zábavné

Jak voní déšť?

Voní vám déšť? A přemýšleli jste někdy o tom proč déšť voní? Čistá voda je fascinující tekutina se spoustou úžasných vlastností, ale nevoní nijak.

Pokud déšť dopadá na suchou půdu, uvolňuje se vůně, které se říká petrichor.

Úkol 1: Vyhledejte etymologický význam tohoto původně řeckého slova **petrichor** např. za pomoci internetového vyhledavače či oblíbeného AI nástroje.

Na vůni petrichoru se podílí několik faktorů – pokud déšť doprovází bouřka, může vznikat ozon. K vůni se přidávají i mastné kyseliny (např. kyselina stearová nebo palmitová), které produkují rostliny a vylučují je do půdy, a také se při dešti do vzduchu formou aerosolu uvolňují molekuly geosminu, produkováné bakteriemi ze skupiny Actinomycetes.

Úkol 2. Najděte anglickou verzi Wikipedie pro heslo geosmin. Jak vypadá molekula geosminu?

Geosmin je molekula, na kterou je lidský čich vysazený – dovede ji totiž detekovat i v úplně minimálním množství – udává se, že lidský nos je schopen detekovat geosmin, i když se vyskytuje v koncentraci 5 částic na bilion (10^{12}) – to je přibližně něco jako čajová lžička v 200 olympijských plaveckých bazénech. Proč máme takovou citlivost zrovna ke geosminu? Jedna z hypotéz říká, že nám tato schopnost pomáhala hledat vzácné zdroje vody. Jak tedy dovede člověk zachytit vůni geosminu?

Teprve v loňském roce vědci zjistili, že je receptorem pro geosmin protein s kryptickým jménem Olfactory receptor 11A1 (Ball a kol. 2024). Pokusíme se tedy o něm dozvědět více.

Úkol 3:

Na internetu jděte na adresu www.uniprot.org a do vyhledávacího okénka vložte: **Olfactory receptor 11A1**. Hned první odkaz by měl být protein z člověka. Otevřete odkaz „kliknutím“ na unikátní identifikátor Q9GZK7.



The screenshot shows the UniProt entry for Olfactory receptor 11A1 (Q9GZK7). The page is divided into several sections. On the left, there is a navigation menu with tabs: Function, Names & Taxonomy, Subcellular Location, Disease & Variants, PTM/Processing, Expression, Interaction, Structure, Family & Domains, Sequence, and Similar Proteins. The main content area is titled 'Q9GZK7 · O11A1_HUMAN'. It includes a 'Function' section with the text 'Odorant receptor' and a 'Gene Ontology' section. The 'Amino acids' count is 315. The 'Protein existence' is 'Evidence at transcript level'. The 'Annotation score' is 35. The 'Tools' section includes links for Download, Add, Community curation, Add a publication, and Entry feedback.

Hned na začátku najdete i informaci o velikosti proteinu, pokud se podíváte na kolonku zvanou Amino acids

3a) Jak dlouhý je protein?

Protein má aminokyselin.

Pokud půjdete do části Subcellular Location, můžete se podívat, kde v buňce se protein vyskytuje – zda je rozptýlen rovnoměrně po buňce, nebo se velmi specificky nalézá jen v některé části buněk.

3b) Kde se protein nalézá v buňce?

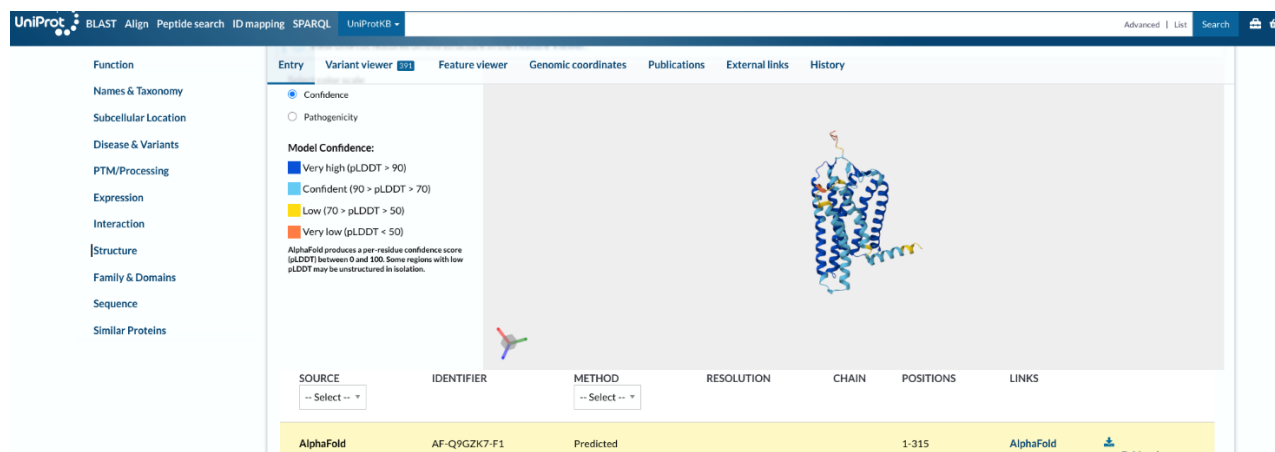
.....

Protein Olfactory receptor 11A1 spadá do jedné z největších lidských rodin proteinů, která má více než 500 členů a mimo jiné je zodpovědná za vnímání vůní. Jde o rodinu receptorů spřažených s tzv. G proteiny – tyto receptory se vyznačují tím, že membránou procházejí celkem 7×.

Přestože je Olfactory receptor 11A1 lidský protein a spadá do známé rodiny, víme o něm stále velmi málo a není tedy překvapením, že ani jeho 3D struktura nebyla dosud experimentálně určena – díky pokroku v predikci 3D struktury proteinů však máme k dispozici řadu nástrojů, které dovedou 3D strukturu spolehlivě predikovat.

My bychom se rádi dále dozvěděli, jak tento protein váže geosmin, a proto využijeme znovu informace v databázi Uniprot a v sekci Structure si můžeme prohlédnout predikovanou 3D

strukturu proteinu a také nalézt unikátní identifikátor predikované struktury programem AlphaFold.



3c) V databázi Uniprot v sekci Structure si prohlédněte strukturu – jde o membránový protein – helikální úseky procházejí mebránou. Spočítejte, kolik helikálních úseků protein má (ten malý helikální úsek kolmý na ostatní membránou neprochází).

Schopnost cítit geosmin však nemají jenom lidé, ale i další obratlovci, a dokonce i bezobratlí jako např. chvostokoci. Schopnost cítit vodu je velmi důležitá pro pouštní zvířata, jako je např. velbloud. Velbloud má homolog lidského proteinu Olfactory receptor 11A1.

V databázi Uniprot najděte sekvenci tohoto proteinu z velblouda dvouhrbého – protein má identifikátor **A0A8B8U6E4**.

3d) Má velbloud delší, nebo kratší Olfactory receptor 11A1 než člověk?

3e) Rádi bychom udělali predikci 3D struktury velbloudího homologu proteinu, abychom se mohli přesvědčit, zda jde o podobné proteiny. A proto budeme potřebovat sekvenci aminokyselin tohoto proteinu. Najděte ji v sekci Sequence na Uniprot stránce proteinu a poté v této sekci „klikněte“ na Download.

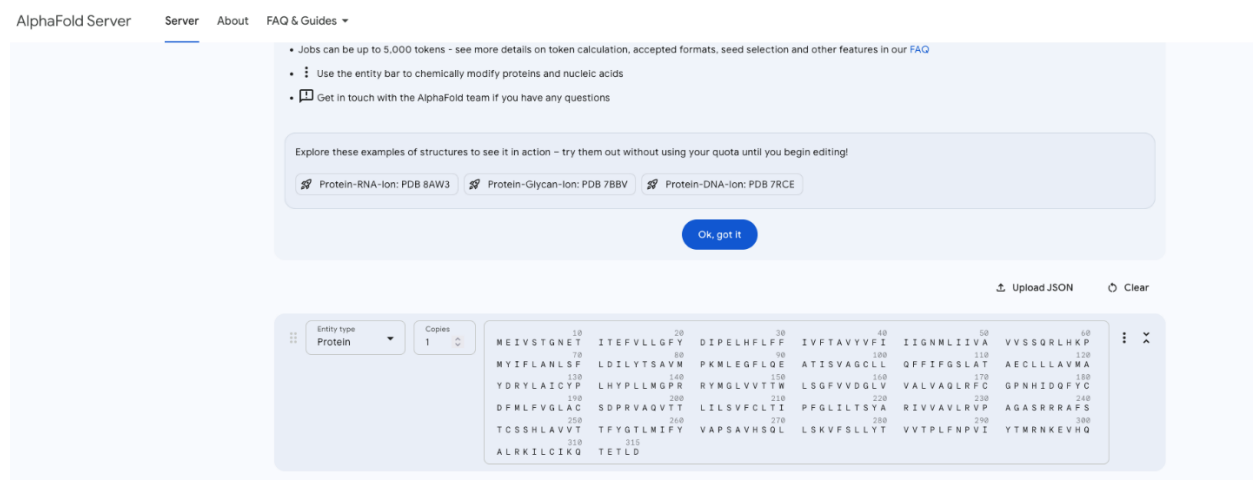
Sekvenci však můžete nalézt také zde:

```
>tr|A0A8B8U6E4|A0A8B8U6E4_CAMFR Olfactory receptor 11A1 OS=Camelus ferus
OX=419612 GN=LOC102514433 PE=4 SV=1
MSGCSRNGHDLRAGPPPQEVMDTLPEEFTRYGSPVHAQPLSEGSGELQHLGFLPVTLFL
AIYVIVRGNALIVLAVASSQTLHTPMYFFLGPFSLLEIGYTSNIAPQLLRSFLEGRETI
SLVSCFLQFYLFASLAAAECLLLSAMS YDCYLAICHPLHPALMSTLLCGCLAAGTWFCG
FSFSAFILALAAHLPLFPGSREIDHYFCDFAPVVGLFCGDVRVMWGAGVSISGFLTLPF
LLIVTSYAFILRTVLPVPSGRGKQKAFSTCFSHLSAVGVFYGTLLIVVYVAPTDMAPFLR
KAFSVLYTVFTPMVNP I IYSLKNQEVKGALLRLWGQLIPGHTPQRGIRQLLTSS
```

Sekvence proteinu je vše, co potřebujeme k predikci 3D struktury za pomoci AlphaFold serveru. Server najdete zde: <https://alphafoldserver.com/welcome>
AlphaFold server je volně dostupný, ale potřebujete mít nějaký účet u Google (nebo si jej vytvořit).

Do políčka Input vložíte sekvenci proteinu bez úvodních identifikátorů, tedy toto:

```
MSGCSRNGHDLRAGPPPQEVMDTLPEEFTRYGSPVHAQPGLSEGSGELQHLGFLPVTFLFL
AIYVIVIRGNALIVLAVASSQTLHTPMYFFLGPFSLLEIGYTSNIAPQLLRSFLEGRETI
SLVSCFLQFYLFASLAAAECLLSAMSYDCYLAICHPLHYPALMSTLLCGCLAAGTWFCG
FSFSAFILALAAHLPLFPGRSREIDHYFCDFAPVVGLFCGDVVRVMWAGVSI SGFLTLPF
LLIVTSYAFILRTVLPVPSGRGKQKAFSTCFSHLSAVGVFYGT LIVVYVAPT DHMAPFLR
KAFSVLYTVFTPMVNPIIYSLKNQEVKGALLRLWGQLIPGHTPQRGIRQLLTSS
```



A poté pokračujete na Continue and preview job. Poté bude vyzváni k pojmenování úkolu, tak jej nějak pojmenujte – velbloud, déšť nebo jakkoli jinak a poté označte Confirm and Submit job. A počkejte, než se sekvence namodeluje – obvykle to trvá jen několik málo minut. Po skončení si můžete strukturu prohlédnout vybráním jména úkolu myší.

3f) Pokud srovnáte obě struktury, jsou v nich místa či oblasti, kde jsou si struktury obou proteinů podobné?

3g) Je nějaká oblast, kde se struktury obou proteinů liší?

Díky dostupnosti sekvencí a možnosti vytvářet predikce 3D struktur se tak můžete přesvědčit, zda např. proteiny, které jsou příbuzné, i podobně vypadají, nebo zda se tvarově liší.

Pokud byste chtěli blíže zkoumat, jak podobné struktury jsou, pak je vhodné použít algoritmy strukturního srovnávání a programy sloužící k vizualizaci proteinových struktur – např. program Pymol nebo Chimera – v odkazech je i návod na strukturní alignment v programu Pymol.

Obdobným způsobem jako v tomto pracovním listě se můžete např. pokusit dozvědět více o mechanismech účinku kouzelných léků na hubnutí (např. Ozempic) nebo se podívat, jak moc

se liší proteiny fungující správně od těch nefunkčních, které vedou např. ke geneticky podmíněným chorobám.

Extra: Úkol 4

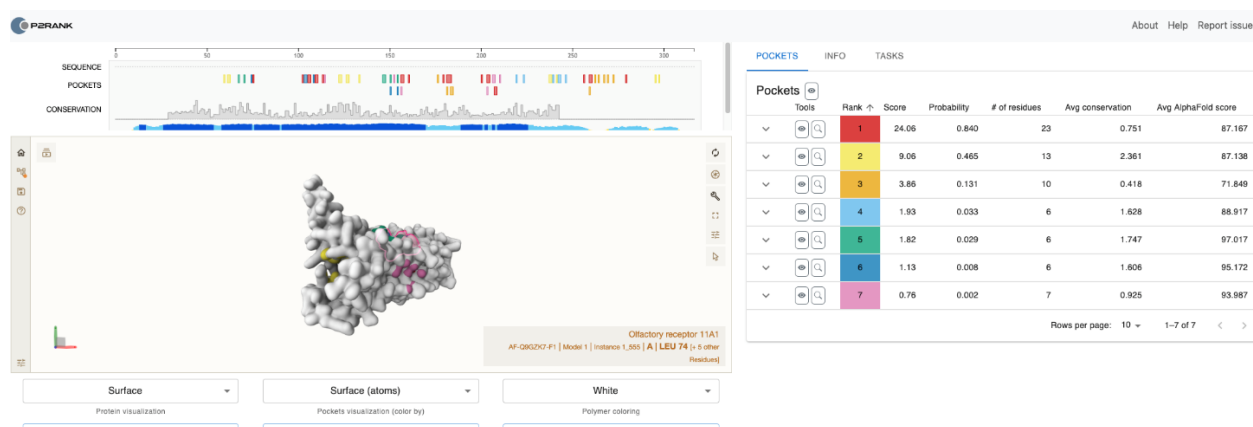
Pokud známe strukturu proteinu, můžeme se pokusit predikovat místo vazby geosminu a srovnat toto místo u proteinu člověka a velblouda, a to např. za pomoci nástroje P2RANK, který vyvinula skupina Davida Hokszy z Matematicko fyzikální fakulty Univerzity Karlovy. Nástroj najdete na <http://prankweb.cz>. P2RANK využívá strojové učení, aby na 3D struktuře proteinu predikoval vazebná místa malých molekul (Polák a kol. 2025).

Nejprve se vraťte na anglickou Wikipedii pro heslo geosmin a na ní najděte a zaznamenejte si zápis struktury zvaný SMILES.

Na stránkách <http://prankweb.cz> nejprve budeme hledat vazebné místo v proteinu člověka – struktura proteinu nebyla dosud experimentálně určena, ale můžeme použít také AlphaFold model.

V Select input model vyberte AlphaFold structure a do políčka vložte unikátní identifikátor proteinu v Uniprot databázi: Q9GZK7.

Ponechte základní nastavení a spusťte program (upozornění: na výsledek můžete čekat několik minut). Po dokončení se vám vlevo objeví struktura a vpravo tabulka s predikovanými vazebnými kapsami – nejpravděpodobněji je správná hned ta první v pořadí.



Pokud byste chtěli vidět, jak se do kapsy bude geosmin vázat, pak rozbalte nabídku u první kapsy a vyberte Create task – v rozbalovací nabídce Task type pak vyberte Docking. (Docking je metoda, která předpovídá, jak se malé molekuly vážou do proteinu). A zadejte SMILES zápis geosminu, který jste si dříve uložili – tedy toto: O[C@]12[C@H](CCC[C@]2(CCCC1)C)C. Poté vyberte Create task a za chvíli byste měli mít predikci hotovou. V tasks označte ikonu oka, která se objevila za slovem successful, o otevře se vám okno, kde bude zobrazena navázaná molekula geosminu do proteinu a v jejím okolí budou zvýrazněné aminokyseliny,

které ji vážou. Pokud na aminokyselinu najedete myší, uvidíte, jak se jmenuje a jaké má pořadí v sekvenci proteinu.

Následně můžete celý proces zopakovat i s proteinem z velblouda, jehož model jste vytvořili. Otevřete si nové okno v prohlížeči a v něm opět jděte na <http://prankweb.cz>, ale tentokrát vyberte Custom structure a poté AlphaFold model with conservation a nahrajte tam vámi uloženou strukturu z Alphafold serveru (možná bude zkomprimovaná, a tak složku rozbalte a vyberte soubor, který bude obsahovat vámi vybrané jméno a na konci bude mít 0.cif) a poté už pokračujte podle návodu výše.

4a) Je nejlepší vazebná kapsa shodná u člověka a velblouda?

4b) Podílejí se na vazbě geosminu stejné aminokyseliny?

Odkazy:

Petrichor:

<https://www.acs.org/content/dam/acsorg/education/students/highschool/chemistryclubs/infographics/petrichor-the-smell-of-rain.pdf>

<https://www.smithsonianmag.com/science-nature/what-makes-rain-smell-so-good-13806085/?no-ist>

<https://21století.cz/2024/09/12/geosmin-vune-krajiny-po-desti-jako-signal/>

<https://www.pocitamesvodou.cz/mate-radi-vuni-deste-a-vite-co-vam-to-vlastne-voni/>

<https://cs.wikipedia.org/wiki/Geosmin>

<https://en.wikipedia.org/wiki/Geosmin>

[doi:10.1093/chemse/17.1.23](https://doi.org/10.1093/chemse/17.1.23).

Geosmin receptor:

<https://pubs.acs.org/doi/pdf/10.1021/acs.jafc.4c01515>

Pymol a Chimera (vizualizace struktur):

<https://pymol.org/edu/> – odkaz na stažení školní verze Pymol zdarma

<https://www.youtube.com/watch?v=T0v1Dx6Yb-I> – video k srovnávání struktur

<https://www.cgl.ucsf.edu/chimera/download.html> – odkaz na stažení Chimery

P2RANK:

<https://doi.org/10.1093/nar/gkaf421>