

mgr Katarzyna Stefańska

Katedra i Zakład Informatyki i Statystyki

## Zalecenia dotyczące wyszukiwania otwartych ramek odczytu w sekwencji

### Wykorzystując ORFfinder

data przygotowania: 2024-04-22

## ŚCIEŻKI DOSTĘPU

### Podstawa teoretyczna

Otwarta ramka odczytu (ORF – open reading frame) to fragmenty sekwencji nukleotydowej, które mogą być przetłumaczone na białka. Do identyfikacji ORF można wykorzystać różne narzędzia, a ORFfinder jest jednym z popularnych narzędzi online do tego celu. Otwarta ramka odczytu zaczyna się kodonem START, kończy kodonem STOP i zawiera ciągłą sekwencję kodonów.

### Krok 1: Wczytaj sekwencję w programie ORFfinder

1. Otwórz dowolną przeglądarkę internetową na swoim urządzeniu.
2. Przejdź do strony ORFfinder: w pasku adresu przeglądarki wpisz adres strony <https://www.ncbi.nlm.nih.gov/orffinder/>.
3. Naciśnij klawisz Enter, aby przejść do strony ORFfinder.
4. Na stronie głównej ORFfinder znajduje się pole do wprowadzania sekwencji – możesz wczytać sekwencję poprzez wklejenie jej bezpośrednio do pola tekstowego, zaimportować plik zawierający sekwencję lub wpisać numer dostępu („accession number”) konkretnej sekwencji z GenBanku.

### Krok 2: Wybierz ustawienia wyszukiwania ORF i uruchom wyszukiwanie

1. Wybierz minimalną długość ramki odczytu (im mniejszy parametr zaznaczysz, tym więcej ramek zostanie wyszukanych).

- Wybierz rodzaj kodu genetycznego (dla sekwencji pochodzących od człowieka wybierz **Standard**).
- Wybierz kodon START.
- Uruchom wyszukiwanie przyciskiem **Submit**.

The screenshot shows the ORF finder web interface. It includes a text input field for the query sequence, a range selector (From/To), search parameters (Minimal ORF length, Genetic code, ORF start codon), and a list of genetic codes. A red box highlights the Submit button.

Annotations in the image:

- "wklejamy „accesion number” albo sekwencję nukleotydów" (we paste the "accession number" or the nucleotide sequence)
- "wybieramy, w jakim zakresie sekwencji mają być poszukiwane ORF" (we choose in what range of the sequence ORFs are to be searched)
- "zaznaczamy minimalną długość ORF" (we mark the minimal length of ORF)

### Krok 3: Analiza wyników

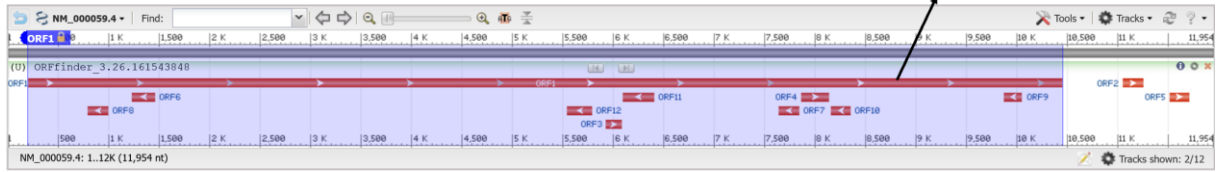
- Po zakończeniu procesu wyszukiwania zostaną wyświetlone znalezione ORF wraz z ich szczegółami, takimi jak pozycja startowa i końcowa, długość (w nukleotydach i aminokwasach).
- W celu pobrania i dalszej analizy konkretnej ramki odczytu, kliknij na wybraną ORF, następnie **Mark** i **Download marked set as Protein (FASTA)**.
- Jeśli chcesz zweryfikować, czy wybrana ramka odczytu koduje białko, kliknij przycisk **BLAST**. Po kliknięciu tego przycisku zostanie wykonane porównanie sekwencji aminokwasowej wybranej ramki odczytu do sekwencji dostępnych w bazie danych UniProtKB, co umożliwi weryfikację, czy dana ORF koduje białko.

## Open Reading Frame Viewer

Homo sapiens BRCA2 DNA repair associated (BRCA2), transcript variant 1, mRNA

ORFs found: 12 Genetic code: 1 Start codon: 'ATG' only

Graficzna reprezentacja znalezionych ORF (czerwone) na tle sekwencji zapytania (szara)



Sekwencja wybranej ORF

ORF1 (3418 aa) Display ORF as... **Mark**

```
>1<1|ORF1
MPTGSKERTFFFEIFKTRCNKADLGPISLHWFEELSSSEAPPYNSPEAES
EKNNNYEPNLKFKPQKPSVNLASTPIIFKEQGLTLPYQSPVKELDK
FKLDLGRWVPSRHKSLRTVTKHQDADDSVCLNSCLSESPVVLQCTH
VTPQRKSVCCSLFHTFVFKGRTPHHSESLGAEVOPRHSSSLAT
PPTLSSTLIVRNEEASETVFPHDTTANNKSYFSNHDESLKKNDRFIASV
TDSENTHREAAASHGFKTSQNSFKVNSCKDHIGKSHPNVLEDEVYETV
DTSSEDSFLCFKSRITNLQVITSKTRKTEHNAUDECESKINWKE
KYSFVSEVEPDNDPLDSNANQPFPESSGDKISKVVPVSLACENSQLTL
SGLNGAQHEKIPLHIESSCDQITSEKDLDTENKRRKDFLTSENSLPRIS
SLPSEKPLNEETVWNRKDEQHESEHTDCLNVAQZSOTSPVASSPQG
IKKSFIRIRESKETFNASFSGHHTDPNFKKTEASESGLIHTVCSQKE
```

Mark subset... Marked: 0 Download marked set as Protein FASTA

Label	Strand	Frame	Start	Stop	Length (nt   aa)
ORF1	+	2	200	10456	10257   3418
ORF11	-	3	6402	6094	309   102
ORF4	+	3	7866	8141	276   91
ORF12	-	3	5793	5542	252   83
ORF6	-	1	1454	1230	225   74
ORF2	+	2	11051	11260	210   69
ORF7	-	2	7843	7637	207   68
ORF8	-	2	991	791	201   66
ORF5	+	3	11514	11711	198   65
ORF10	-	3	8352	8161	192   63
ORF9	-	2	4865	4837	188   62

Lista wszystkich znalezionych ORF, z uwzględnieniem na jakiej nici się znajdują, miejscem STARTU i STOPU oraz długością podaną w nt i aa

ORF1 (3418 aa) Marked set (0)

SmartBLAST SmartBLAST best hit titles...

BLAST BLAST

BLAST Database: UniProtKB/Swiss-Prot (swissprot)