

HIV-virus i retslokalet

Af: *Kresten Cæsar Torp, Aalborghus Gymnasium*

Materialet inddrager dele af: Regassa, Laura; Naowarra Cheeptham og Michèle Shuster: Murder by HIV?, NSTA, Link: [Murder by HIV? Grades 9-12 Edition | NSTA](#)

Retssagen

I 1995 blev en lidt speciel sag indbragt for retten i Louisiana, USA. Som en del af bevisførelsen valgte man at anvende en fylogenetisk analyse af HIV-virus fra offeret.

Sommeren før, i juli 1994, havde offeret, en sygeplejerske, afbrudt et forhold til sin kæreste, en læge. 4. august mødte han imidlertid op på hendes adresse, og gav hende en indsprøjtning med vitamin B-12 mod hendes vilje. Han havde også tidligere givet hende indsprøjtninger med vitamin B-12, men denne gang var indsprøjtningen meget smertefuld.

I januar 1995 blev hun testet positiv for HIV, og hun anklagede sin tidligere kæreste for at have inficeret hende med HIV med sprøjten i august. Han blev anklaget for uagtsomt manddrab.

Offeret var bloddonor, og var i den forbindelse testet negativ for HIV adskillige gange, senest i april 1994. Hun var om sygeplejerske en enkelt gang kommet i berøring med slim fra en AIDS-patient, men var efterfølgende testet negativ for HIV. Hun havde haft sex med i alt 7 mænd, som alle blev testet negative for HIV.

Det viste sig, at en af ekskærestens patienter, der havde AIDS, havde fået taget en blodprøve lige inden 4. august 1994, men prøven var ikke blevet sendt videre til kontrol. De tilhørende papirer var ikke udfyldt korrekt, og politiet fandt dem i en kasse med journaler fra 1982 i et depot.

Hvordan matcher man HIV-sekvenser?

Hvordan bestemmer man, om offerets HIV-virus stammer fra patienten, der fik taget blodprøven? Der er flere overvejelser:

- HIV's genom består af RNA, som først skal oversættes til DNA, før det kan sekventeres.
- RNA muterer hurtigere end DNA. Virus fra en patient vil derfor afvige lidt fra hinanden.

Man anvender følgende procedure:

1. RNA fra HIV-virus oprenses fra personen.
2. RNA oversættes til DNA. Det sker ved en tilbagetranskription, vha. enzymet Revers Transkriptase
3. DNA sekventeres.

Herefter undersøges om virus-DNA fra offeret matcher virus-DNA fra patienten.

Denne sammenligning skal dog tage højde for at der er sket mutationer. Derfor sammenlignes den med virus-DNA fra andre HIV-smittede i området.

4. Sammenligningen foretages ved at lave alignment af DNA-sekvenserne og ved hjælp af deres genetiske afstand opstille et fylogenetisk træ.

Opgaver

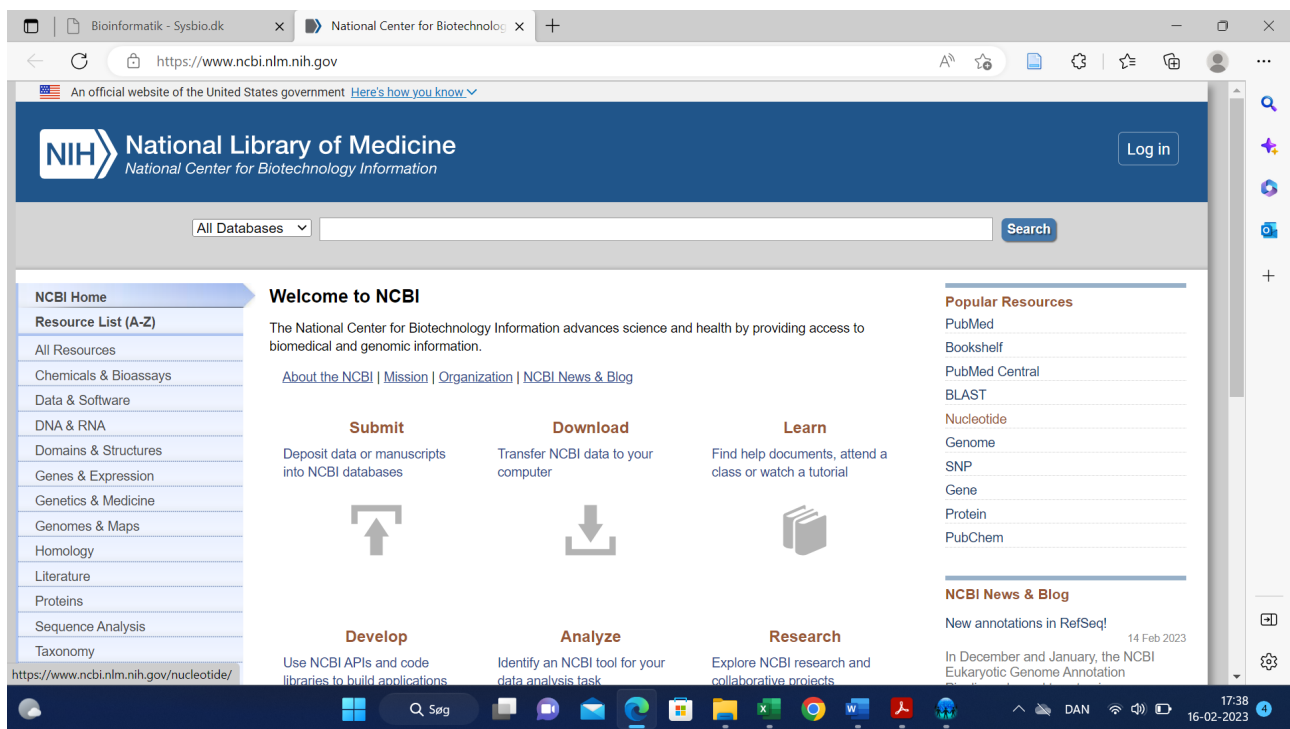
- Forklar, principperne i hvordan man kan foretage procedure 1-3. Anvend dine bøger. Det vil flere af processerne givetvis være beskrevet.
- Udarbejd det fylogenetiske træ for sagen. Du skal følge fremgangsmåden nedenfor.
- Beskriv det fylogenetiske træ du får ud.
- Forklar, om træet underbygger, at offerets virus stammer fra patientens blodprøve.
- Sammenlign resultaterne fra den retsgenetiske test og de øvrige oplysninger i sagen. Diskuter, om I tænker, at sagen endte med en domfældelse af kæresten.

Udarbejdelse af fylogenetisk træ

Første skridt er at finde de sekvenser som skal sammenlignes. Dem samler man i en fil, en form for bibliotek. De skal have et format, som kan genkendes af programmet, når de skal sammenlignes. Selve sammenligningen kaldes et alignment, og man kan ud fra det konstruere et fylogentisk træ. Her gør vi det i programmet UGENE.

Oprettelse af bibliotek

1. Gå ind på NCBI's database Genbank. Link: [National Center for Biotechnology Information \(nih.gov\)](https://www.ncbi.nlm.nih.gov)
2. Vælg "Nucleotide" i menuen til højre.

The image is a screenshot of a web browser displaying the National Center for Biotechnology Information (NCBI) website. The browser's address bar shows the URL https://www.ncbi.nlm.nih.gov. The website header includes the NIH logo and the text "National Library of Medicine National Center for Biotechnology Information". Below the header is a search bar with a dropdown menu set to "All Databases" and a "Search" button. On the left side, there is a vertical navigation menu with various categories such as "Resource List (A-Z)", "All Resources", "Chemicals & Bioassays", "Data & Software", "DNA & RNA", "Domains & Structures", "Genes & Expression", "Genetics & Medicine", "Genomes & Maps", "Homology", "Literature", "Proteins", "Sequence Analysis", and "Taxonomy". The main content area features a "Welcome to NCBI" message and several service tiles: "Submit" (Deposit data or manuscripts into NCBI databases), "Download" (Transfer NCBI data to your computer), "Learn" (Find help documents, attend a class or watch a tutorial), "Develop" (Use NCBI APIs and code libraries to build applications), "Analyze" (Identify an NCBI tool for your data analysis task), and "Research" (Explore NCBI research and collaborative projects). On the right side, there are "Popular Resources" (PubMed, Bookshelf, PubMed Central, BLAST, Nucleotide, Genome, SNP, Gene, Protein, PubChem) and "NCBI News & Blog" with a recent article titled "New annotations in RefSeq!". The Windows taskbar is visible at the bottom of the screenshot.

3. Find den sekvens, vi vil sammenligne med de øvrige. Det gør man vha. en accession number, dvs. et ID, som sekvensen har i databasen. Kopier Accession Number: **AY156807**.
4. Indsæt nummeret i søgefeltet.

https://www.ncbi.nlm.nih.gov/nucleotide/

Nucleotide AY156807

Nucleotide

The Nucleotide database is a collection of sequences from several sources, including GenBank, RefSeq, TPA and PDB. Genome, gene and transcript sequence data provide the foundation for biomedical research and discovery.

Using Nucleotide

- Quick Start Guide
- FAQ
- Help
- GenBank FTP
- RefSeq FTP

Nucleotide Tools

- Submit to GenBank
- LinkOut
- E-Utilities
- BLAST
- Batch Entrez

Other Resources

- GenBank Home
- RefSeq Home
- Gene Home
- SRA Home
- INSDC

Tryk "Search".

https://www.ncbi.nlm.nih.gov/nuccore/AY156807

HIV-1 clone V2.MIC.RT from USA reverse transcriptase (pol) gene, partial cds

GenBank: AY156807.1

[FASTA](#) [Graphics](#) [PopSet](#)

Go to: ☺

LOCUS AY156807 805 bp DNA linear VRL 22-OCT-2002

DEFINITION HIV-1 clone V2.MIC.RT from USA reverse transcriptase (pol) gene, partial cds.

ACCESSION AY156807

VERSION AY156807.1

KEYWORDS -

SOURCE Human immunodeficiency virus 1 (HIV-1)

ORGANISM Human immunodeficiency virus 1

Viruses; Riboviria; Pararnavirae; Artverviricota; Revtraviricetes; Orthovirales; Retroviridae; Orthoretrovirinae; Lentivirus.

REFERENCE 1 (bases 1 to 805)

AUTHORS Metzker, M.L., Mindell, D.P., Liu, X.M., Ptak, R.G., Gibbs, R.A. and

[Analyze this sequence](#)

- Run BLAST
- Pick Primers
- Highlight Sequence Features
- Find in this Sequence

[Retrovirus Resource](#)

Retrieve, view, and download retroviral genomic and protein sequences.

[Related information](#)

Nu får du en række oplysninger frem om sekvensen. Du skal imidlertid bruge selve DNA-sekvensen og du skal have den i et filformat, som programmet kan anvende senere. Filformatet er FASTA, som du finder lige under overskriften.

5. Tryk på FASTA.

GenBank: AY156807.1

>AY156807.1 HIV-1 clone V2.MIC.RT from USA reverse transcriptase (pol) gene, partial cds

```

STAGSACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCACCTTTAAATTTTC
CGATAAGTCCTATTGAAACTGTACCAGTAAATTAAGCCAGGAATGGATGGCCCAAAGTTTAAACAATG
GCCACTGACGAGAGAAAAAATAAAGCATTAGTAGAAATTTGTACAGAAATAGAAAAAGGGAAGAAAAATT
TCAAAAAATTTGGGCTGAAAAATCCATACAATCTCCAGTATTTGCCATAAGAAAAAACAAGTACTAGAT
GGAGAAAAATTTAGTAGATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAAT
ACCACATCCTGCAGGGTTAAAAAGAAAAAATCAGTAACAGTGTGGATGGGTGATGCATATTTTCA
GTTCCCTTAGATAAAGAGTTTCAGGAAGTATACAGTATTTACCAATACCTAGTATAAACAATGAGACACCA
AGATTAGATACAGTACAATGTGCTTCCACAGGGATGGAAGGATCACCAGCAATATTTCCAAAGTAGCAT
GACAAAAATCTTAGAGCCTTTTAGAAAAACAATCCAGACATAGTTATCTCAATACATGGATGATCTG
TATGTAGGACTGACTTAGAAATAGGCAGCATAGAATAAAAAACAGAGAACTAAGACAACATCTGTGA
AGTGGGGATTTTTCACACCAGACAAAAACACGAAAGAACCTCCATTCGGTGGATGGGTATGAAC
CCATCTGATAAATGGACGTACAGCTATAGTGC

```

Første del indeholder en række oplysninger om sekvensen. Dem vender vi lige med...

Marker kun selve DNA-sekvensen (GTAG.....GTGC) og kopier den.

Gå tilbage til Genbanks hovedside.

Vælg BLAST (lige under "Nucleotide")

BLAST®

Home Recent Results Saved Strategies Help

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

NEWS

ElasticBLAST 1.0.0 is Now available!

ElasticBLAST version 1.0.0 has support for faster cheaper disks at AWS and better supports Kubernetes on GCP!

Mon, 09 Jan 2023 [More BLAST news...](#)

Web BLAST

Nucleotide BLAST
nucleotide to nucleotide

blastx
translated nucleotide to protein

tblastn
protein to translated nucleotide

Protein BLAST
protein to protein

BLAST er et program som anvendes til at søge sekvenser i databasen, som ligner den sekvens, man indtaster (Query).

6. Tryk på Nucleotide BLAST.

7. Indsæt sekvensen fra før.

BLASTN programs search nucleotide databases using a nucleotide query. more...

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#)

AGTGGGGATTTTTCACACCAGACGAAAAACACCAGAAAAGAACCTCCATTC
CGTTGGATGGGTATGAACT
CCATCCTGATAAATGGACAGTACAGCCATATAGTGC

Query subrange

From

To

Or, upload file Der er ingen fil valgt [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

Align two or more sequences [?](#)

Choose Search Set

Database Standard databases (nr etc.) rRNA/ITS databases Genomic + transcript databases Betacoronavirus

Nucleotide collection (nr/nt) [?](#)

Organism Optional exclude

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

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

Limit to Optional Sequences from type material

Entrez Query Optional [YouTube](#) [Create custom database](#)

8. Under "Choose Search Set" vælg: "Standard database" og "Nucleotide collection (nr/nt)".

9. Under "Program Selection": Vælg "Somewhat similar"

Job Title

Enter a descriptive title for your BLAST search [?](#)

Align two or more sequences [?](#)

Choose Search Set

Database Standard databases (nr etc.) rRNA/ITS databases Genomic + transcript databases Betacoronavirus

Nucleotide collection (nr/nt) [?](#)

Organism Optional exclude

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

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

Limit to Optional Sequences from type material

Entrez Query Optional [YouTube](#) [Create custom database](#)

Program Selection

Optimize for

Highly similar sequences (megablast)

More dissimilar sequences (discontiguous megablast)

Somewhat similar sequences (blastn)

Choose a BLAST algorithm [?](#)

Search database Nucleotide collection (nr/nt) using Blastn (Optimize for somewhat similar sequences)

Show results in a new window

Algorithm parameters

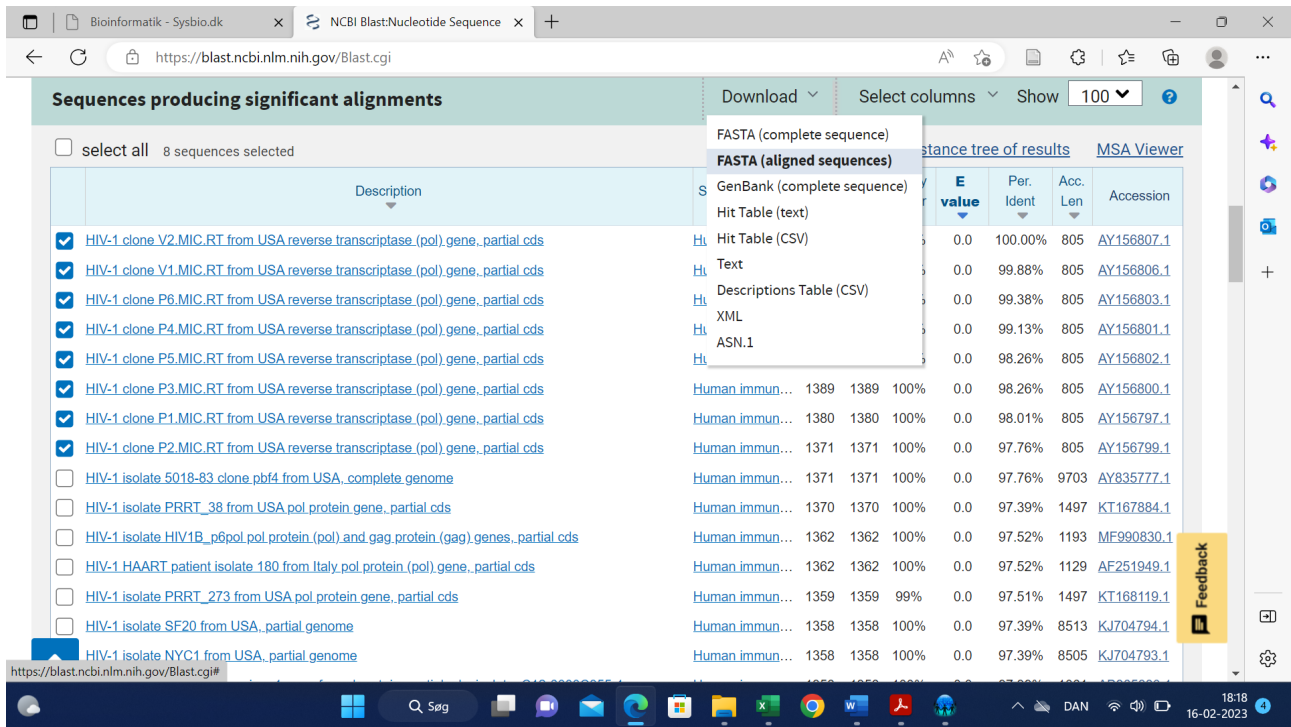
10. Tryk: "BLAST". Det tager lidt tid af lave matchet i databasen. Husk at den indeholder flere mia. sekvenser.

11. Scroll lidt ned, om du finder en liste over de hits i databasen, som minder mest om den indtastede. Marker de første 8. De stammer fra retssagen:

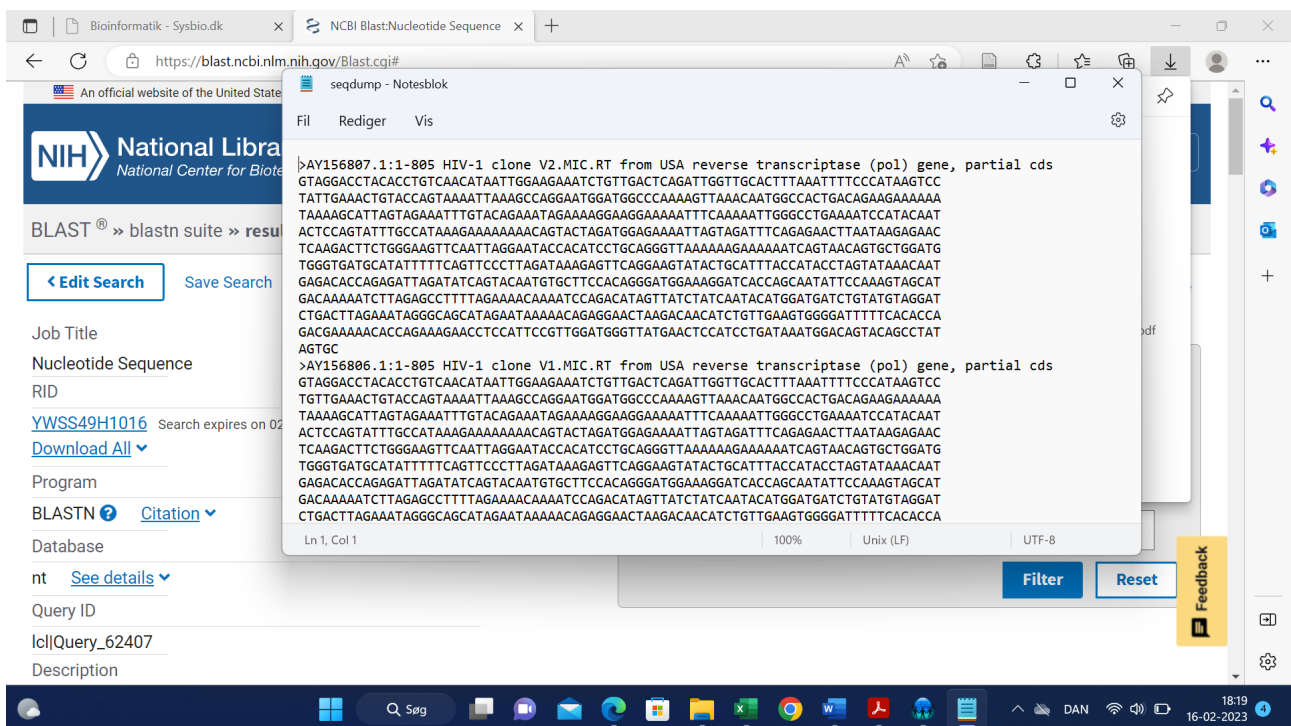
- Klon V1 og Klon V2 er virusvarianter fra offeret (victim)
- Klonerne P1-P6 stammer fra patienten.
- "MIC.RT." angiver metoden: Revers Transkriptase
- "(pol) gene" skyldes, at sekvensen kun er et udsnit af HIV-genomet, nemlig pol-genet.

12. Download filerne ved at trykke på "Download".

13. Vælg "FASTA (aligned sequence)".



14. Gem filerne som en tekstfil (.txt). Word ol. Du'r ikke.



15. For at undersøge, om virus fra offer og patient er tæt beslægtede, må man sammenligne med virus fra andre personer, der ikke har været involverede i sagen. Suppler filen med sekvenser fra tre andre personer fra USA og tre fra samme stat (Louisiana). Sekvensernes acquisition number er:

- AY835777 >USA_1
- AY835778 >USA_2
- AY835769 >USA_3
- AY156793 >LA_1
- AY156789 >LA_2
- AY156788 >LA_3

16. Nu skal filen rettes til, så sekvenserne får håndterbare navne (første linje inden sekvensen starter). De vigtigste regler er:

Sekvensen skal altid starte med ">"

Der må ikke være mellemrum. Brug "_"

Du kan fx omdøbe sekvenserne til:

Acquisition number	Nyt navn
AY156807	>victim_clone1
AY156806	>victim_clone2
AY156803	>patient_clone1
AY156801	>patient_clone2
AY156802	>patient_clone3
AY156800	>patient_clone4
AY156797	>patient_clone5
AY156799	>patient_clone6
AY835777	>USA_1
AY835778	>USA_2
AY835769	>USA_3
AY156793	>LA_1
AY156789	>LA_2
AY156788	>LA_3

Du har nu en FASTA-fil du kan anvende til at lave det fylogenetiske træ.

Alignment og konstruktion af fylogenetisk træ i UGENE

UGENE er et program, som kan sammenligne sekvenser ved alignment og konstruere fylogenetiske træer. Begge dele foregår automatiseret, hvilket er nødvendigt.

For at forstå hvad der sker, kan du lave opgaverne i filen: *Træer med anvendelse i biologi*.

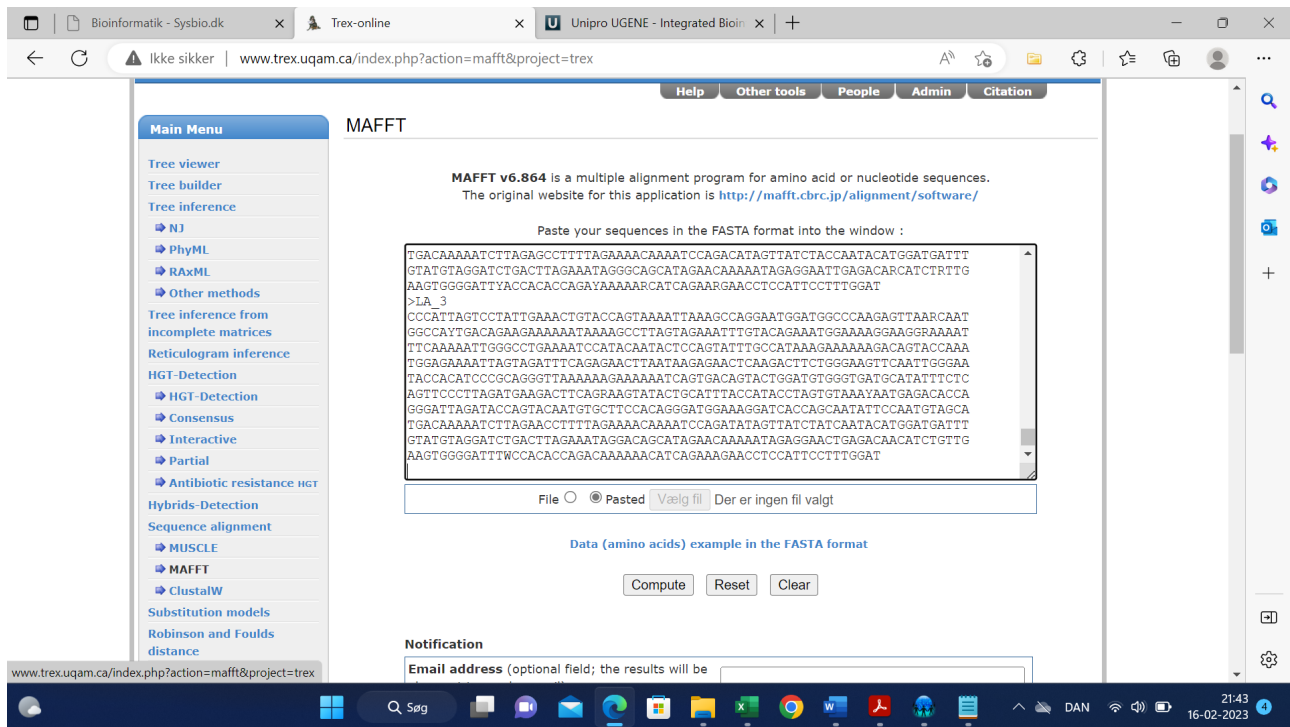
Download UGENE fra hjemmesiden: [Unipro UGENE - Integrated Bioinformatics Tools](#)

Alignment og konstruktion af fylogenetisk træ i T-REX

T-Rex er et onlineprogram, som kan sammenligne sekvenser ved alignment og konstruere fylogenetiske træer. Begge dele foregår automatiseret, hvilket er nødvendigt.

For at forstå hvad der sker, kan du lave opgaverne i filen: *Træer med anvendelse i biologi*.

1. Åben FASTA-filen "HIV_RT_sekvenser_til_fylogeni".
2. Marker og kopier filens indhold.
3. Gå ind på hjemmesiden: <http://www.trex.uqam.ca/>
4. Vælg "MAFFT" i menuen til venstre.
5. Indsæt filen i feltet.



The screenshot shows the MAFFT online interface. The browser address bar is www.trex.uqam.ca/index.php?action=mafft&project=trex. The page title is "MAFFT". The main content area contains the following text:

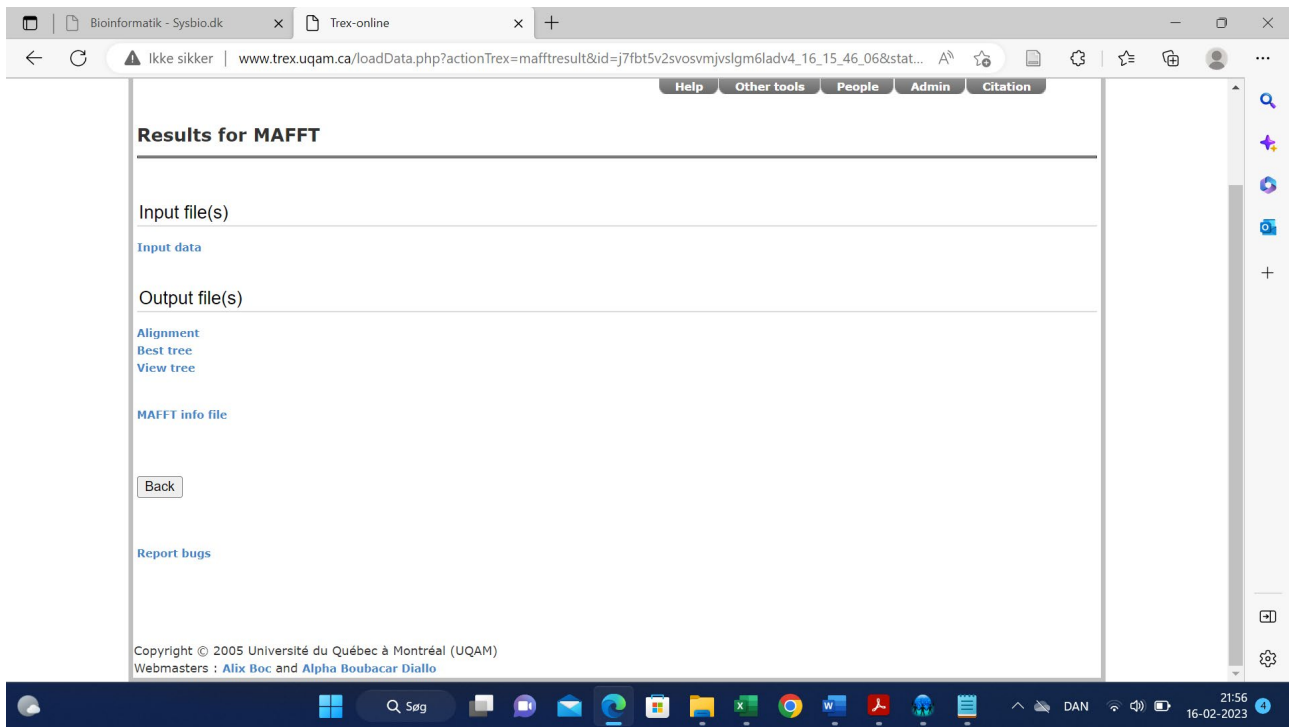
MAFFT v6.864 is a multiple alignment program for amino acid or nucleotide sequences.
The original website for this application is <http://mafft.cbrc.jp/alignment/software/>

Paste your sequences in the FASTA format into the window :

```
TGACAAAAATCTTAGAGCCTTTAGAAAACAAAATCCAGACATAGTTATCTACCANTACATGGATGATTT
GTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAACAAAATAGAGGAATTGAGACARCATCTRTTG
AAGTGGGGATTYACCACACCAGAYAAAAARCATCAGAARGAACCTCCATTCCCTTGGAT
>LA_3
CCCATTAGTCCTATTGAAACTGTACCAGTAAAATTAAGCCAGGAATGGATGGCCCAAGAGTTAARCAAT
GGCCAYTGACAGAGAAAAATAAAAGCCTTAGTAGAAATTTGTACAGAAATGGAAAAAGGAAAGRAAAAT
TTCAAAAATTTGGGCTGAAAATCCATACAACTCCAGTATTTGCCATAAAGAAAAAGACAGTACCAAAA
TSGAGAAAATTAGTAGATTTGAGAGAATTAATAAGAGAATCAAGACTTCTGGGAAGTTCAATTTGGSA
TACCACATCCCGCGGGTTAAAAGAAAAAAATCAGTGACAGTACTGGATGTGGGTGATGCAATTTTCTC
AGTCCCTTAGATGAAAGTTGAGAAATGATAGTACTGATTTACCATAGTGTAAATGATGAGACACCA
GGATTAGATACCAGTAAATGTCTTCCACAGGATGGAAGGATCCACAGCAATTCATCCATGTAGCA
TGACAAAAATCTTAGAACCTTTAGAAAACAAAATCCAGATATAGTTATCTCATTCATGGATGATTT
GTATGTAGGATCTGACTTAGAAATAGGACAGCATAGAACAAAATAGAGGAATGAGACARCATCTRTTG
AAGTGGGGATTTWCCACACCAGACAAAAARCATCAGAARGAACCTCCATTCCCTTGGAT
```

Below the text area, there is a "File" button, a "Pasted" indicator, and a "Vælg fil" button. Below that, there is a "Data (amino acids) example in the FASTA format" section with "Compute", "Reset", and "Clear" buttons. At the bottom, there is a "Notification" section with an "Email address" field and a note: "Email address (optional field; the results will be".

6. Tryk "Compute".



The screenshot shows a web browser window with two tabs: "Bioinformatik - Sysbio.dk" and "Trex-online". The address bar shows the URL: www.trex.uqam.ca/loadData.php?actionTrex=mafftresult&id=j7ft5v2svosvmjvslgm6ladv4_16_15_46_06&stat.... The page content is titled "Results for MAFFT" and includes the following sections:

- Input file(s)**
- Input data**
- Output file(s)**
- Alignment**
- [Best tree](#)
- [View tree](#)
- [MAFFT info file](#)
-
- [Report bugs](#)

At the bottom of the page, there is a copyright notice: "Copyright © 2005 Université du Québec à Montréal (UQAM)" and "Webmasters : [Alix Boc](#) and [Alpha Boubacar Diallo](#)". The browser's taskbar at the bottom shows the Windows logo, search bar, and various application icons. The system tray on the right indicates the time as 21:56 and the date as 16-02-2023.

7. Nu kan resultatet vises i forskellige formater under "Output". Vælg "View tree".

Du kan nu se resultatet.