Introduction to applied bioinformatics

PETRA MATOUŠKOVÁ

2024/2025

8/9

„Nucleotide bioinformatics III“

Retrieving nucleotide sequences from databases (Genbank/NCBI)

Feature analysis: statistics, reverse complement, restriction analysis

Translation, identifying open reading frame

PCR primer design - detection, rt-PCR

Secondary structure prediction

Sequence comparison, unknown sequence identification

Single Nucleotide Polymorphisms

DNA sequencing

Gene expression

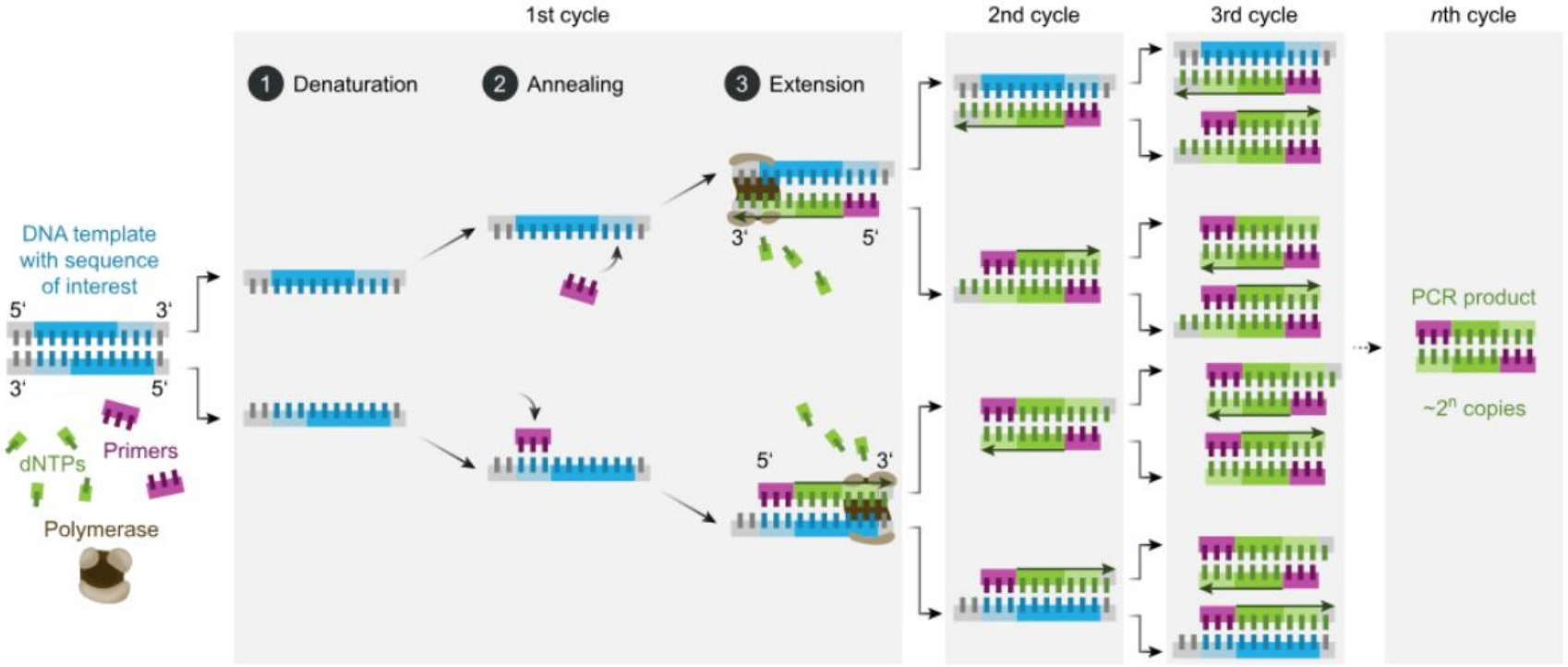
microRNA

Genomes….

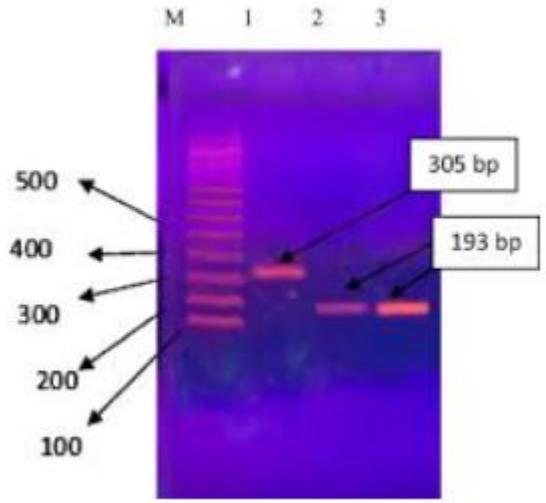
….

Polymerase chain reaction

Polymerase chain reaction (PCR) → **amplification of required gene**, between two primers



→ „End point“ detection



[Polymerase chain reaction (PCR) (article) | Khan Academy]

Polymerase chain reaction

Amplification of required gene → **for cloning, functional studies…**

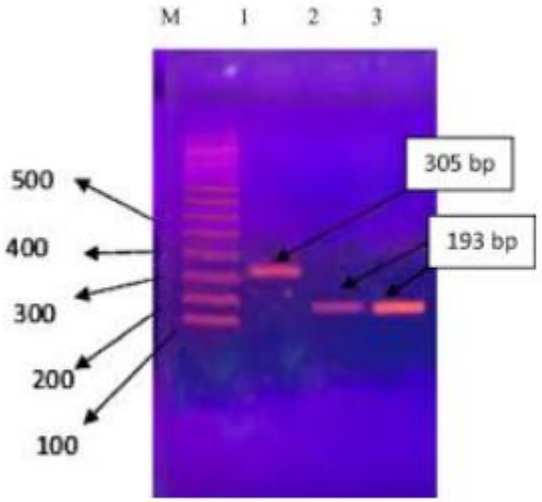
Product: various length, between two primers

Primers: **manual design**

→ „End point“ detection

-forward-check by OligoCalc

-reverse -from reverse complement sequence and check by OligoCalc)



Polymerase chain reaction (PCR) (article) | Khan Academy

Polymerase chain reaction

Specific gene detection

Primers: anywhere along the sequence !

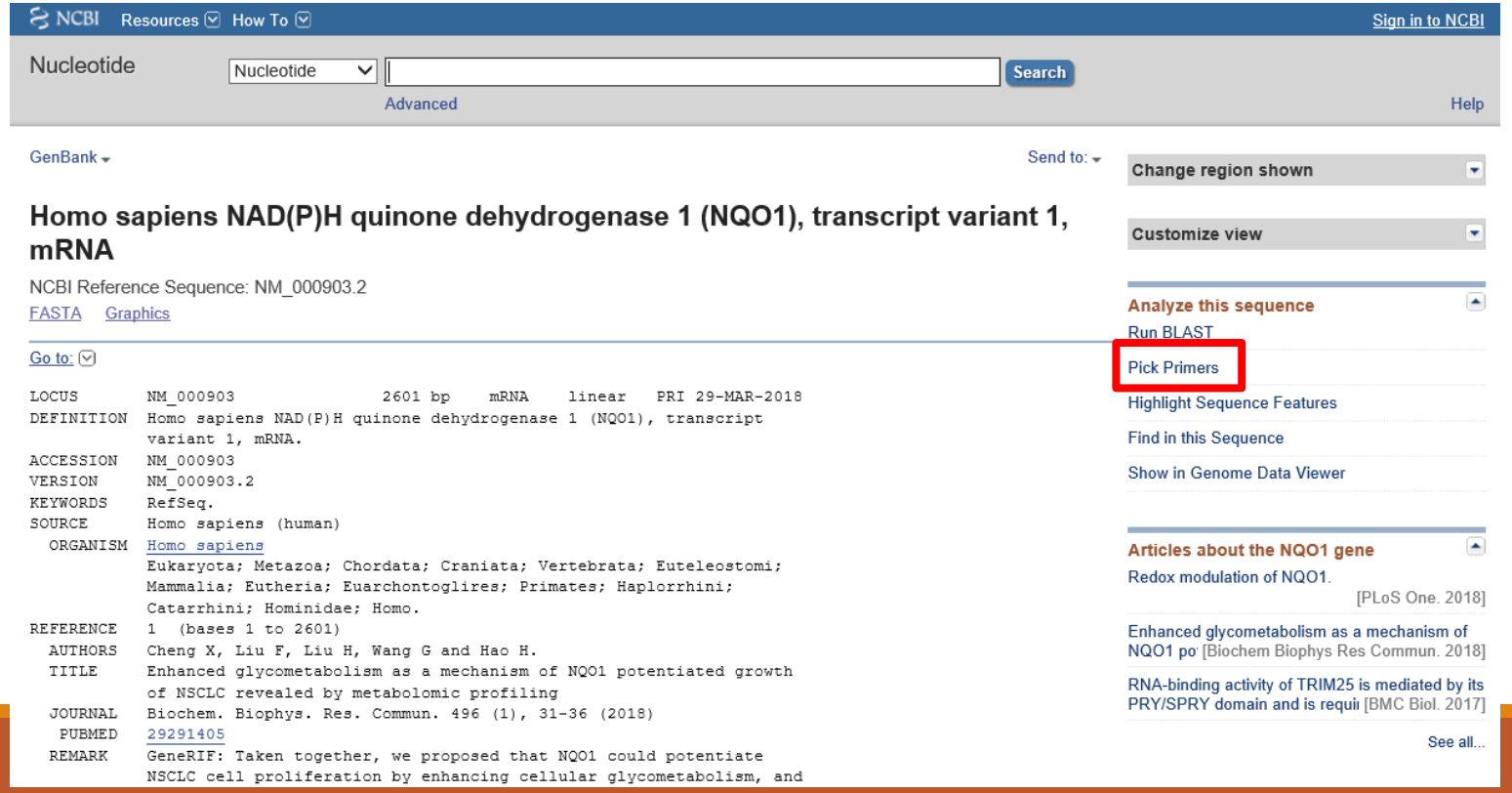
Product: **200 - 500nt** (good separation for detection of length)

But primers: must be **specific** for tested organism → **BLAST**

Specific gene detection

Gene detection

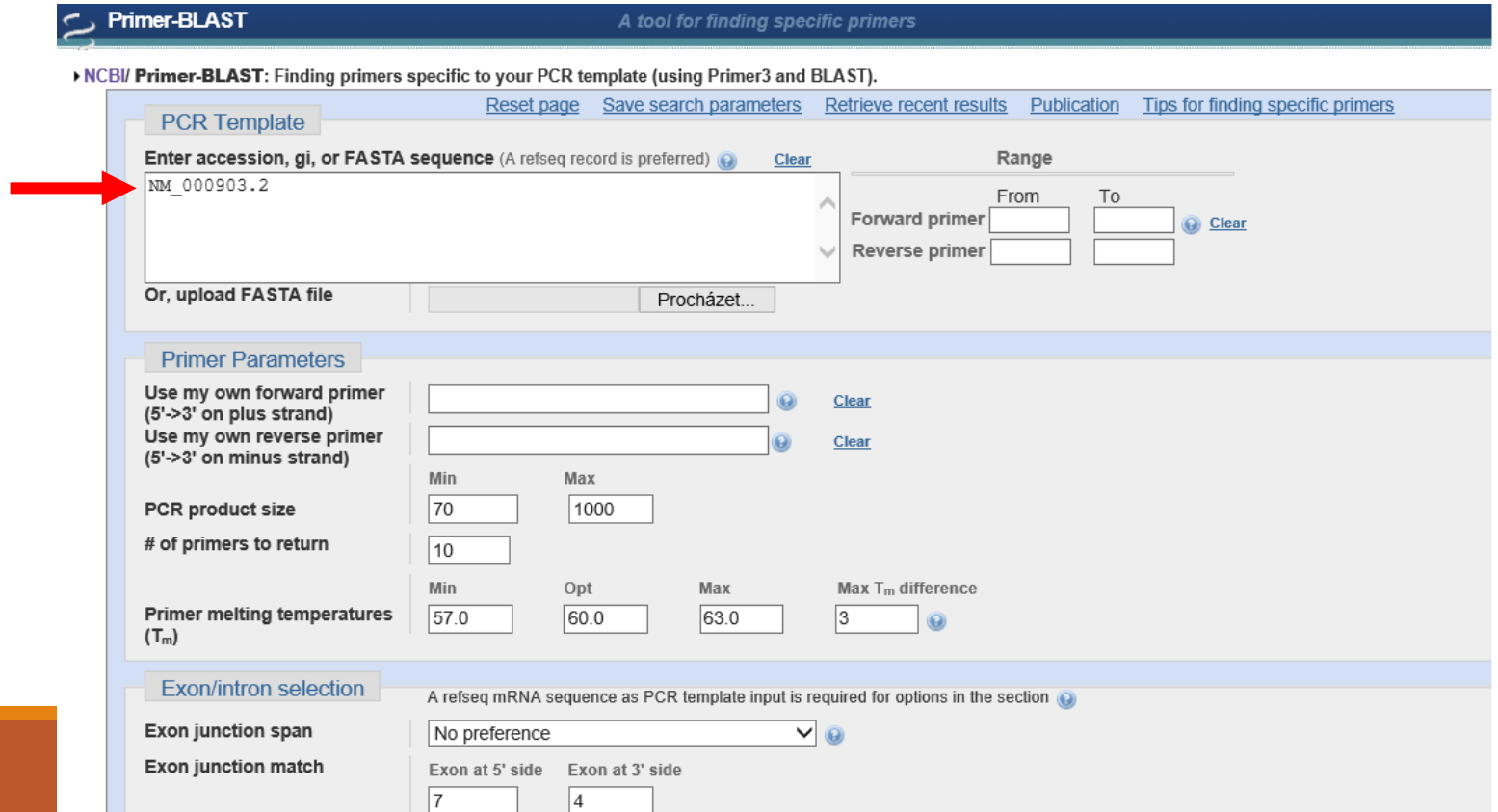
- Specificity matters! (Primer BLAST!)



Specific gene detection

Gene detection

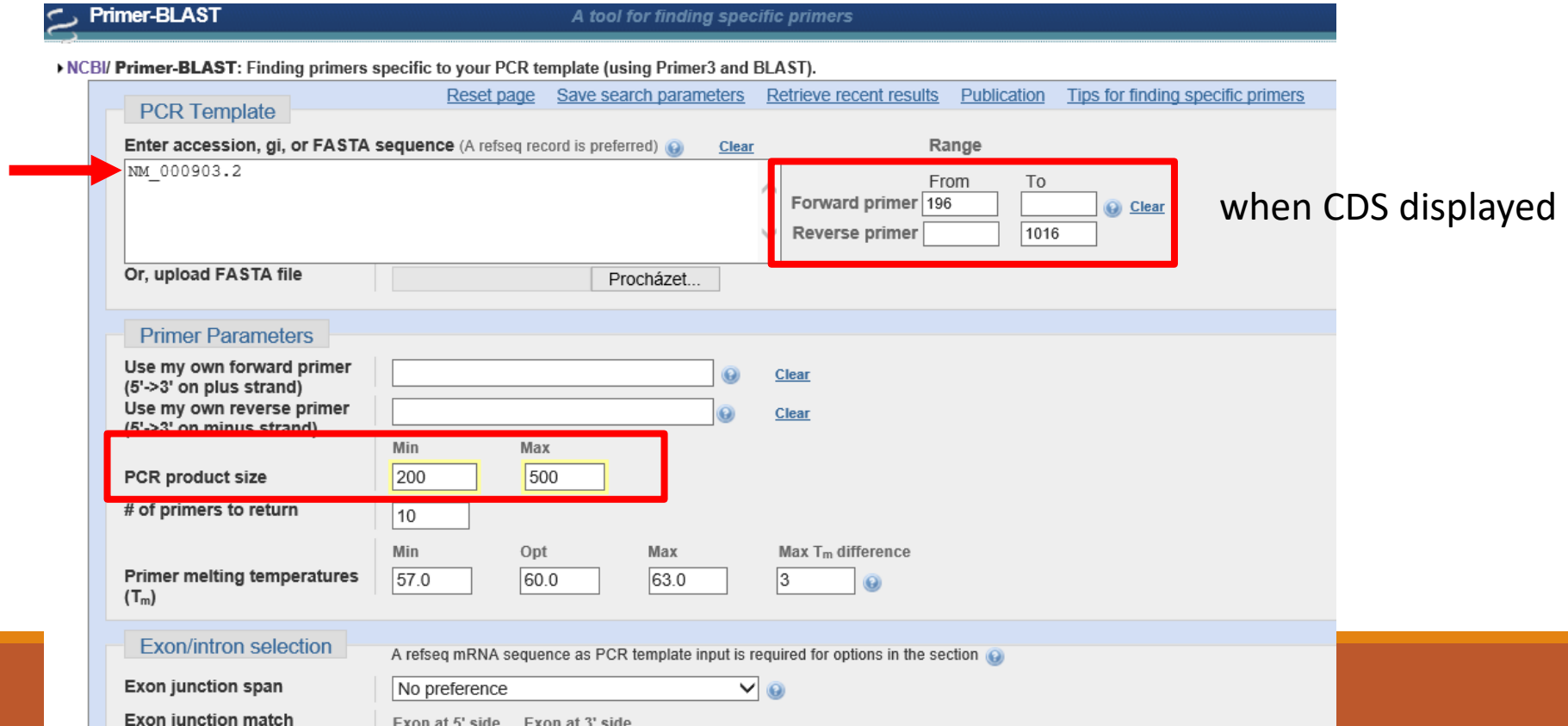
- Specificity matters! (Primer BLAST!)



Specific gene detection

Gene detection

- Specificity matters! (Primer BLAST!)

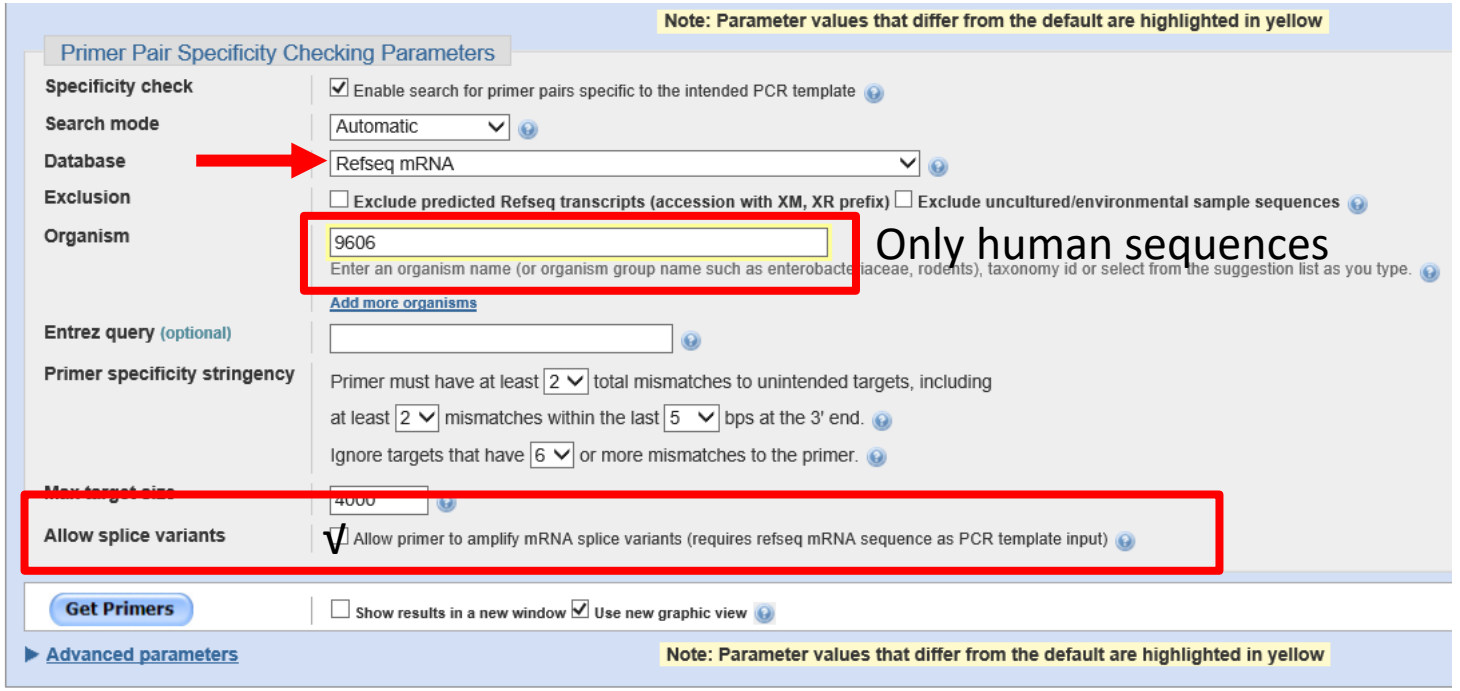


when CDS displayed

Specific gene detection

Gene detection

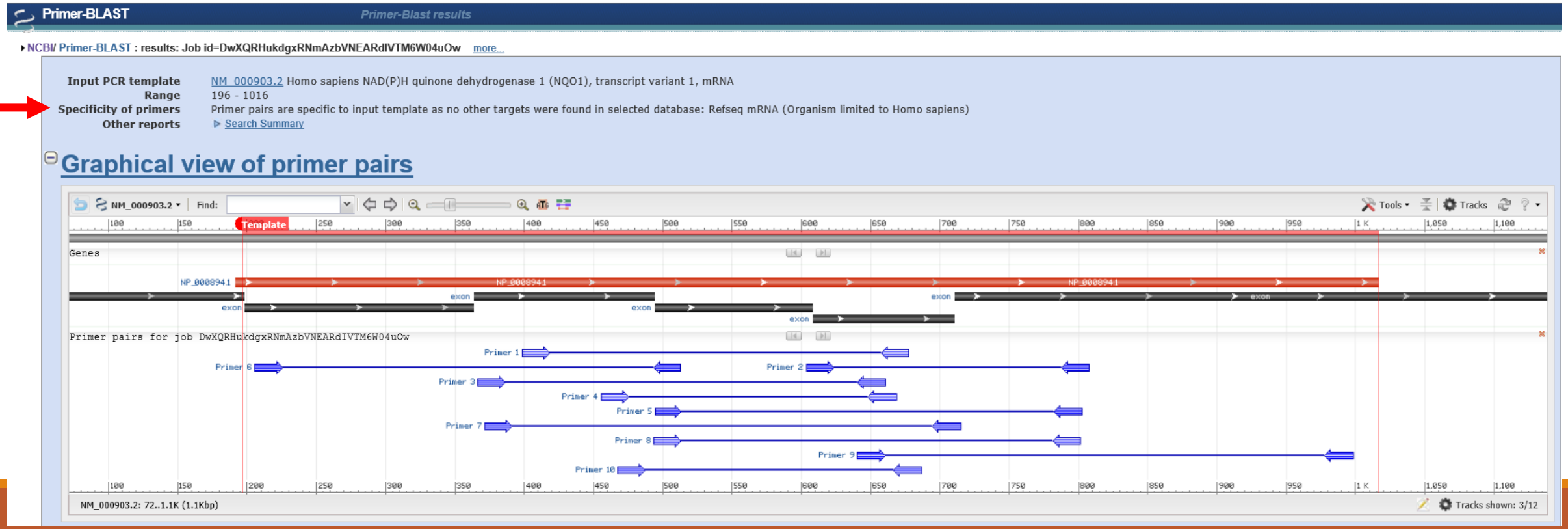
- Specificity matters! (Primer BLAST!)



Specific gene detection

Gene detection

- Specificity matters! (Primer BLAST!)



Specific gene detection

Gene detection

- Specificity matters! (Primer BLAST!)

Detailed primer reports

Primer pair 1

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	GCCGAGTCTGTTCTGGCTTA	Plus	20	399	418	59.75	55.00	3.00	2.00
Reverse primer	GTGGATCCCTTGCAGAGAGT	Minus	20	677	658	59.09	55.00	6.00	2.00
Product length	279								

Products on intended target

>NM_000903.2 Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), transcript variant 1, mRNA

```
product length = 279
Forward primer  1    GCCGAGTCTGTTCTGGCTTA  20
Template        399  ....................  418

Reverse primer  1    GTGGATCCCTTGCAGAGAGT  20
Template        677  ....................  658
```

Products on allowed transcript variants

>NM_001025434.1 Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), transcript variant 3, mRNA

```
product length = 165
Forward primer  1    GCCGAGTCTGTTCTGGCTTA  20
Template        399  ....................  418

Reverse primer  1    GTGGATCCCTTGCAGAGAGT  20
Template        563  ....................  544
```

Practical part….

Design primers **specific** to detect your sequence using „pick primers“
(and BLAST).

(Hw) Align the primers designed by the program (Pick primers) with your CDS and mRNA

Primers verification

F:TGCTTACACTTACGCTGCCT

R:CCAGGCGTTTCTTCCATCCT

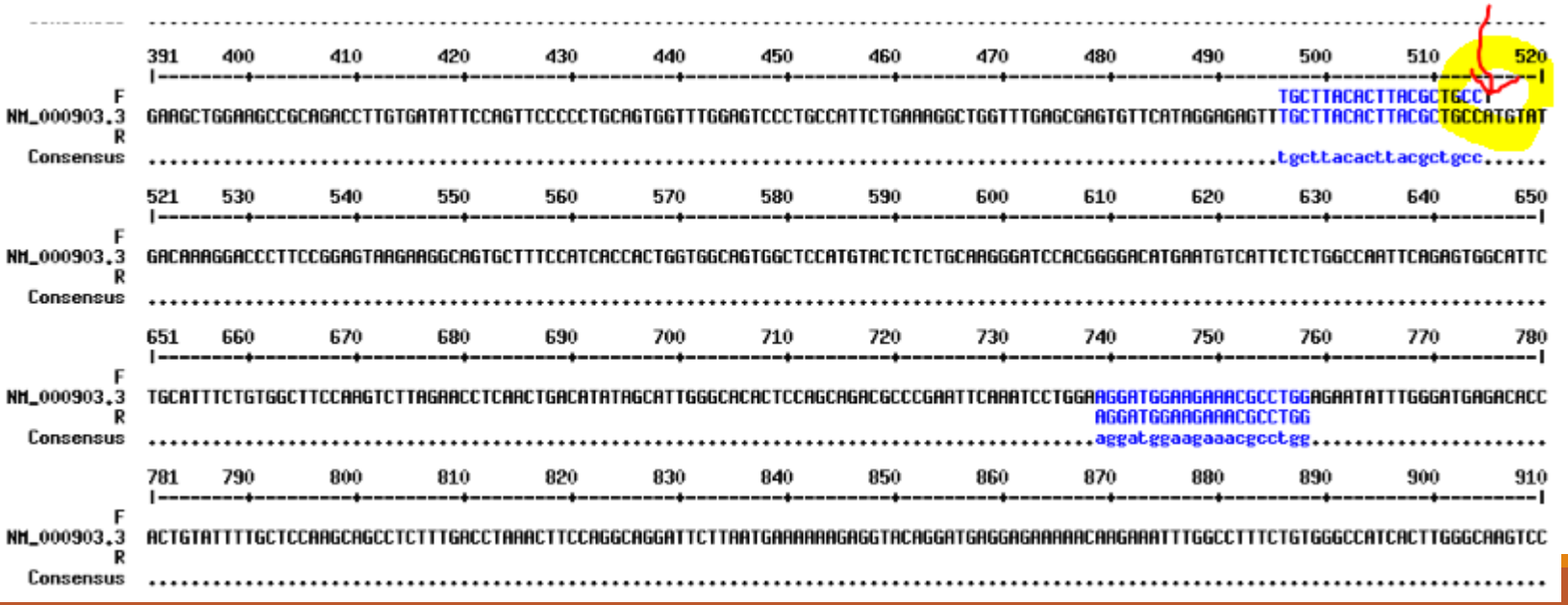
Verify if the primers can be used for detection of NQO1 (NM_000903.3) (use multalin)

Primers verification

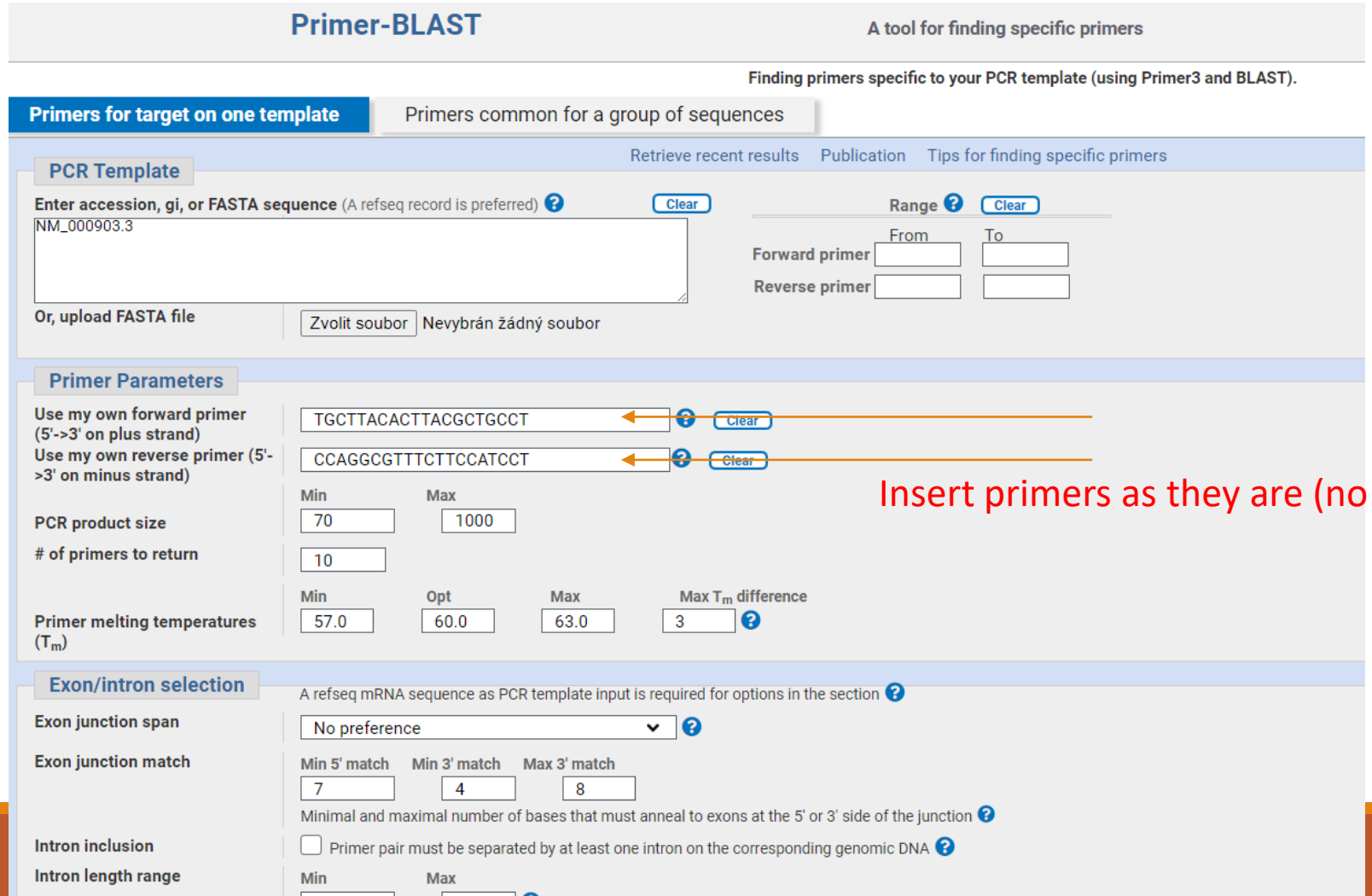
F:TGCTTACACTTACGCTGCCT

R:CCAGGCGTTTCTTCCATCCT

Verify if the primers can be used for detection of NQO1 (NM_000903.3) (use multalin)

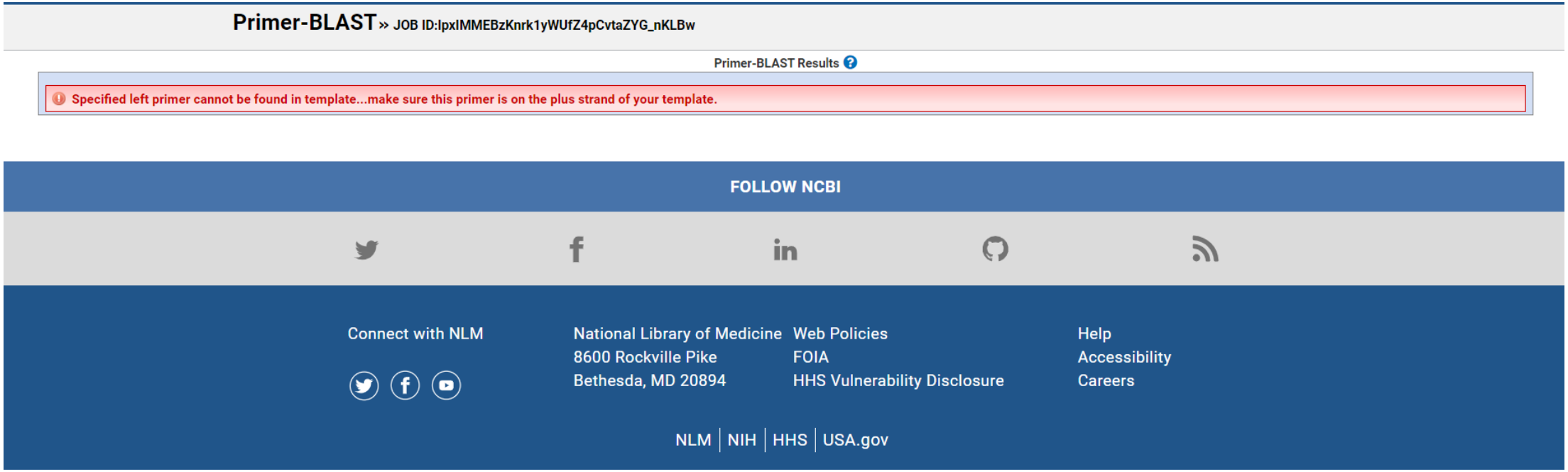


Primers verification – control

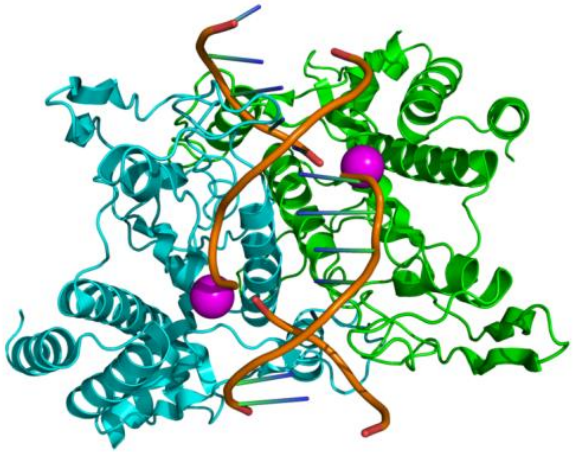


Insert primers as they are (no R-reverse complement)

Primers verification – control

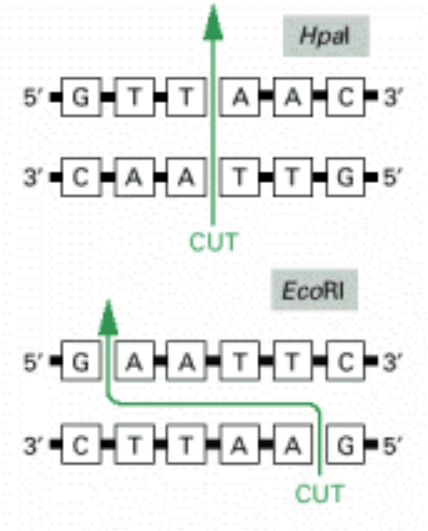


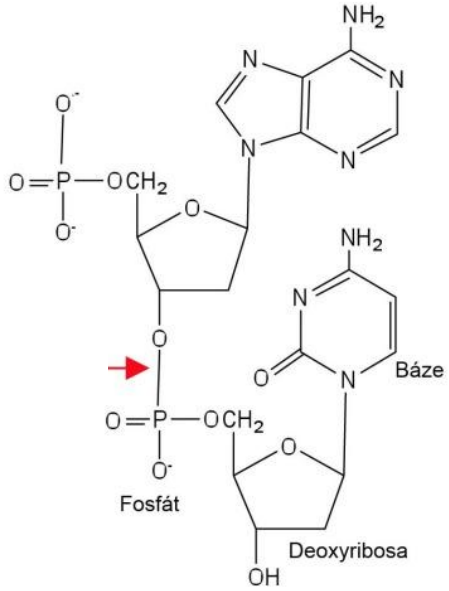
Restriction analysis



Restriction endonucleases (molecular scissors)

- cleaves both strands in specific symetric site – palindrom
- creating blunt or cohesive ends

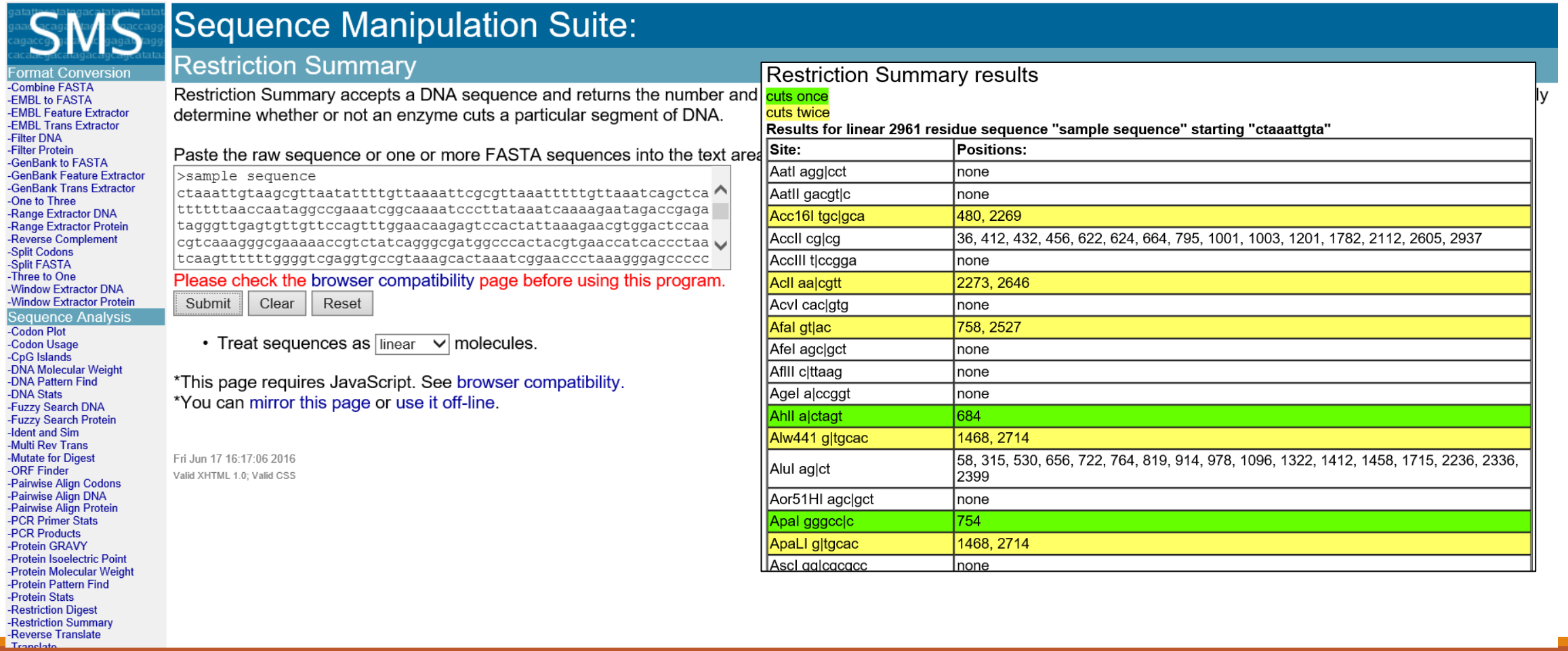




- usage for DNA fragmentation, restriction analysis of mutations, cloning, etc.

https://www.youtube.com/watch?v=piGOuud3f40

Restriction analysis



SMS

Format Conversion
-Combine FASTA
-EMBL to FASTA
-EMBL Feature Extractor
-EMBL Trans Extractor
-Filter DNA
-Filter Protein
-GenBank to FASTA
-GenBank Feature Extractor
-GenBank Trans Extractor
-One to Three
-Range Extractor DNA
-Range Extractor Protein
-Reverse Complement
-Split Codons
-Split FASTA
-Three to One
-Window Extractor DNA
-Window Extractor Protein

Sequence Analysis
-Codon Plot
-Codon Usage
-CpG Islands
-DNA Molecular Weight
-DNA Pattern Find
-DNA Stats
-Fuzzy Search DNA
-Fuzzy Search Protein
-Ident and Sim
-Multi Rev Trans
-Mutate for Digest
-ORF Finder
-Pairwise Align Codons
-Pairwise Align DNA
-Pairwise Align Protein
-PCR Primer Stats
-PCR Products
-Protein GRAVY
-Protein Isoelectric Point
-Protein Molecular Weight
-Protein Pattern Find
-Protein Stats
-Restriction Digest
-Restriction Summary
-Reverse Translate
-Translate

Sequence Manipulation Suite:

Restriction Summary

Restriction Summary accepts a DNA sequence and returns the number and ly
determine whether or not an enzyme cuts a particular segment of DNA.

Paste the raw sequence or one or more FASTA sequences into the text area

```
>sample sequence
ctaaattgtaagcgttaatattttgttaaaattcgcgttaaatttttgttaaatcagctca
ttttttaaccaataggccgaaatcggcaaaatcccttataaatcaaaagaatagaccgaga
tagggttgagtgttgttccagtttggaacaagagtccactattaaagaacgtggactccaa
cgtcaaagggcgaaaaaccgtctatcagggcgatggcccactacgtgaaccatcaccctaa
tcaagtttttttggggtcgaggtgccgtaaagcactaaatcggaaccctaaagggagccccc
```

Please check the browser compatibility page before using this program.

Submit Clear Reset

- Treat sequences as linear molecules.

*This page requires JavaScript. See browser compatibility.
*You can mirror this page or use it off-line.

Fri Jun 17 16:17:06 2016
Valid XHTML 1.0; Valid CSS

Restriction Summary results

cuts once
cuts twice

Results for linear 2961 residue sequence "sample sequence" starting "ctaaattgta"

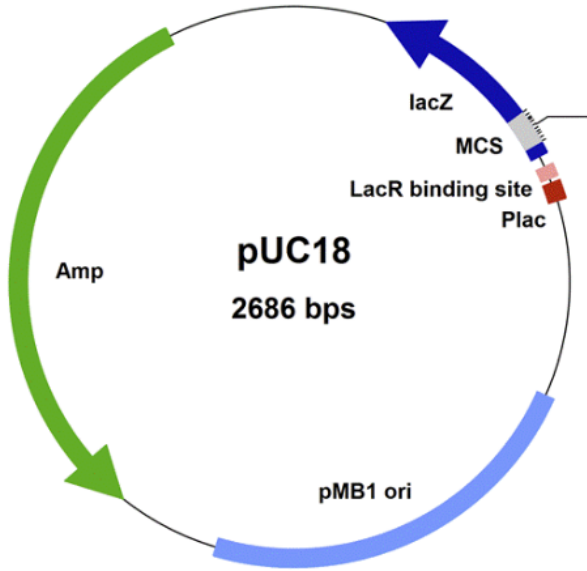
Site:	Positions:
AatI agg|cct	none
AatII gacgt|c	none
Acc16I tgc|gca	480, 2269
AccII cg|cg	36, 412, 432, 456, 622, 624, 664, 795, 1001, 1003, 1201, 1782, 2112, 2605, 2937
AccIII t|ccgga	none
AclI aa|cgtt	2273, 2646
AcvI cac|gtg	none
AfaI gt|ac	758, 2527
AfeI agc|gct	none
AflII c|ttaag	none
AgeI a|ccggt	none
AhlI a|ctagt	684
Alw441 g|tgcac	1468, 2714
AluI ag|ct	58, 315, 530, 656, 722, 764, 819, 914, 978, 1096, 1322, 1412, 1458, 1715, 2236, 2336, 2399
Aor51HI agc|gct	none
ApaI gggcc|c	754
ApaLI g|tgcac	1468, 2714
AscI gg|cgcgcc	none

Practical part

Analyze your sequence for presence of RE target sites.

- check REs from multiple cloning site of vector:

HW: EcoRI, XbaI, BamHI

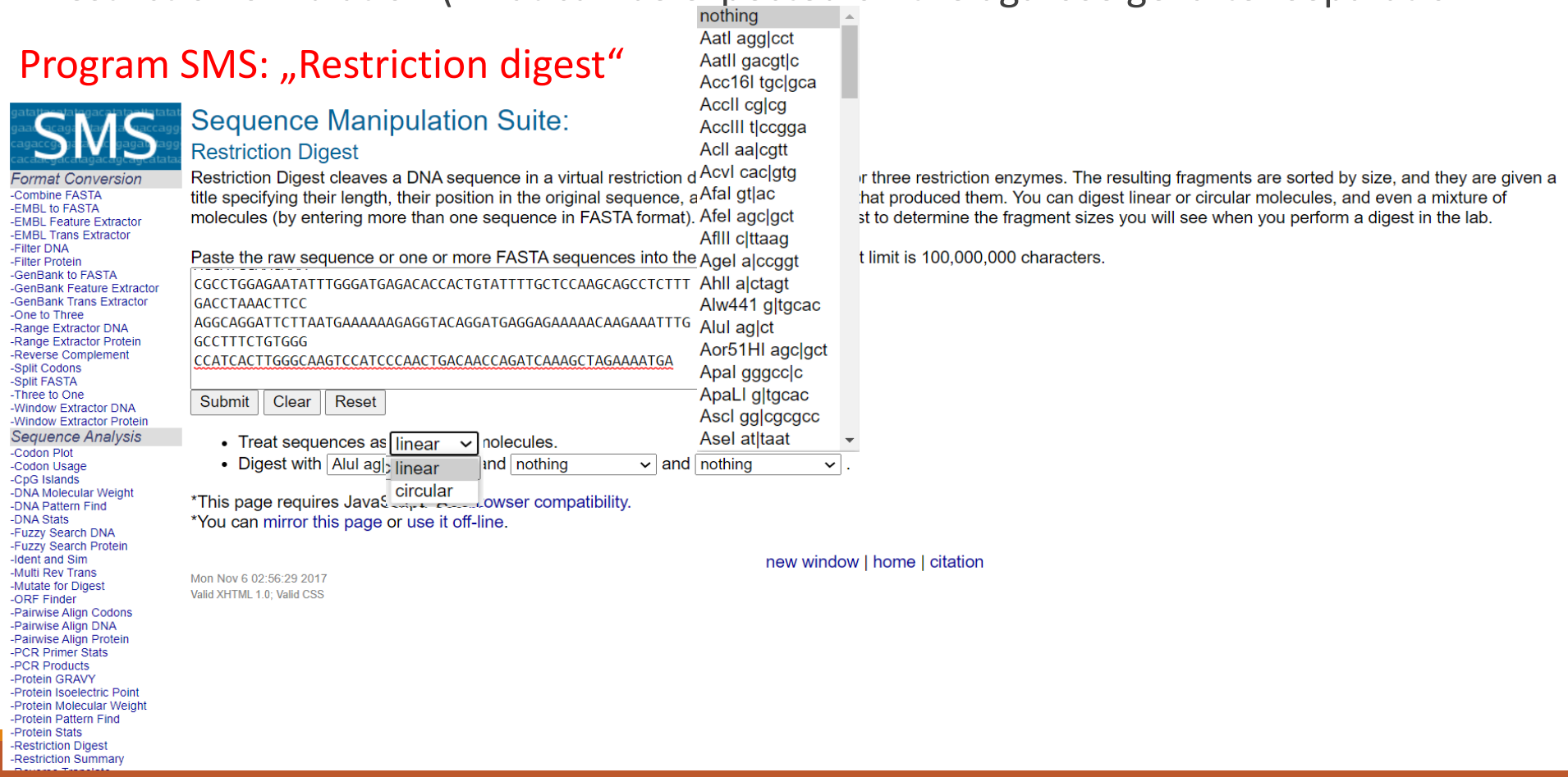


HindIII
SphI
SbfI
PstI
BspMI
AccI
HincII
SalI
XbaI
BamHI
AvaI
SmaI
XmaI
Acc65I
KpnI
BanII
Eco53k
SacI
ApoI
EcoRI

Cloning-Fragment analysis

= restriction simulation (what can be expected on the agarose gel after separation

Program SMS: „Restriction digest“



Cloning-Fragment analysis

NQO1:CDS-restiction summary

>NM_000903.3:122-946 Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), transcript variant 1, mRNA

```
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTATGCCATGAAGG
AGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCGGACCTCTATGCCATGAACTT
CAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTGAAGGACCCTGCGAACTTTCAGTATCCTGCC
GAGTCTGTTCTGGCTTATAAAGAAGGCCATCTGAGCCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAG
CCGCAGACCTTGTGATATTCCAGTTCCCCCTGCAGTGGTTTGGAGTCCCTGCCATTCTGAAAGGCTGGTT
TGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCTTCCGGAGT
AAGAAGGCAGTGCTTTCCATCACCACTGGTGGCAGTGGCTCCATGTACTCTCTGCAAGGGATCCACGGGG
ACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCATTTCTGTGGCTTCCAAGTCTTAGAACC
TCAACTGACATATAGCATTGGGCACACTCCAGCAGACGCCCGAATTCAAATCCTGGAAGGATGGAAGAAA
CGCCTGGAGAATATTTGGGATGAGACACCACTGTATTTTGCTCCAAGCAGCCTCTTTGACCTAAACTTCC
AGGCAGGATTCTTAATGAAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCAACTGACAACCAGATCAAAGCTAGAAAATGA
```

HpaII c|cgg	415
EcoRI g|aattc	603
AluI ag|ct	274, 815

NQO1:CDS-restiction digest

Restriction Digest

Restriction Digest cleaves a DNA sequence in a virtual restriction digest, with one, two, or title specifying their length, their position in the original sequence, and the enzyme sites th molecules (by entering more than one sequence in FASTA format). Use Restriction Digest

Paste the raw sequence or one or more FASTA sequences into the text area below. Input

```
CGCCTGGAGAATATTTGGGATGAGACACCACTGTATTTTGCTCCAAGCAGCCTCTTT
GACCTAAACTTCC
AGGCAGGATTCTTAATGAAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTG
GCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCAACTGACAACCAGATCAAAGCTAGAAAATGA
```

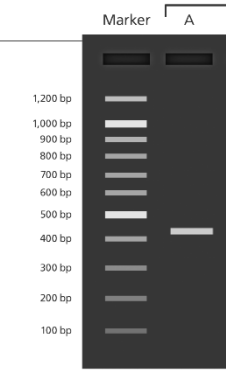
Submit | Clear | Reset

- Treat sequences as linear molecules.
- Digest with HpaII c|cgg and nothing and nothing .

```
>414 bp linear fragment from linear parent NM_000903.3:122-946
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTAT
GCCATGAAGGAGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCG
GACCTCTATGCCATGAACTTCAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTG
AAGGACCCTGCGAACTTTCAGTATCCTGCCGAGTCTGTTCTGGCTTATAAAGAAGGCCAT
CTGAGCCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAGCCGCAGACCTTGTGATATTC
CAGTTCCCCCTGCAGTGGTTTGGAGTCCCTGCCATTCTGAAAGGCTGGTTTGAGCGAGTG
TTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCTTC

>411 bp linear fragment from linear parent NM_000903.3:122-946
CGGAGTAAGAAGGCAGTGCTTTCCATCACCACTGGTGGCAGTGGCTCCATGTACTCTCTG
CAAGGGATCCACGGGGACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCAT
TTCTGTGGCTTCCAAGTCTTAGAACCTCAACTGACATATAGCATTGGGCACACTCCAGCA
GACGCCCGAATTCAAATCCTGGAAGGATGGAAGAAACGCCTGGAGAATATTTGGGATGAG
ACACCACTGTATTTTGCTCCAAGCAGCCTCTTTGACCTAAACTTCCAGGCAGGATTCTTA
ATGAAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGGCCAT
CACTTGGGCAAGTCCATCCCAACTGACAACCAGATCAAAGCTAGAAAATGA
```

A: 414+ 411nt

NQO1:CDS (825nt)



EcoRI

```
>602 bp linear fragment from linear parent NM_000903.3:122-946
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTAT
GCCATGAAGGAGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCG
GACCTCTATGCCATGAACTTCAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTG
AAGGACCCTGCGAACTTTCAGTATCCTGCCGAGTCTGTTCTGGCTTATAAAGAAGGCCAT
CTGAGCCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAGCCGCAGACCTTGTGATATTC
CAGTTCCCCCTGCAGTGGTTTGGAGTCCCTGCCATTCTGAAAGGCTGGTTTGAGCGAGTG
TTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCTTCCGGAGT
AAGAAGGCAGTGCTTTCCATCACCACTGGTGGCAGTGGCTCCATGTACTCTCTGCAAGGG
ATCCACGGGGACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCATTTCTGT
GGCTTCCAAGTCTTAGAACCTCAACTGACATATAGCATTGGGCACACTCCAGCAGACGCC
CG

>223 bp linear fragment from linear parent NM_000903.3:122-946
AATTCAAATCCTGGAAGGATGGAAGAAACGCCTGGAGAATATTTGGGATGAGACACCACT
GTATTTTGCTCCAAGCAGCCTCTTTGACCTAAACTTCCAGGCAGGATTCTTAATGAAAAA
AGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGGCCATCACTTGGG
CAAGTCCATCCCAACTGACAACCAGATCAAAGCTAGAAAATGA
```

B: 602+ 223nt

NQO1:CDS (825nt)

Cloning-Fragment analysis

- Treat sequences as linear molecules.
- Digest with EcoRI g|aattc and HpaII c|cgg and nothing .

NQO1:CDS-restiction digest

>NM_000903.3:122-946 Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), transcript variant 1, mRNA

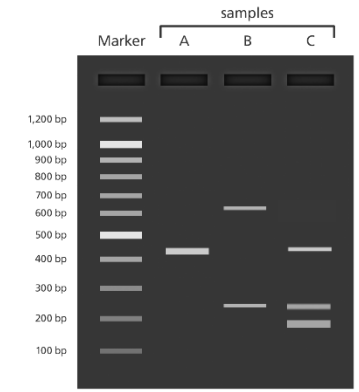
```
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTATGCCATGAAGG
AGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCGGACCTCTATGCCATGAACTT
CAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTGAAGGACCCTGCGAACTTTCAGTATCCTGCC
GAGTCTGTTCTGGCTTATAAAGAAGGCCATCTGAGCCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAG
CCGCAGACCTTGTGATATTCCAGTTCCCCCTGCAGTGGTTTGGAGTCCCTGCCATTCTGAAAGGCTGGTT
TGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCTTCCGGAGT
AAGAAGGCAGTGCTTTCCATCACCACTGGTGGCAGTGGCTCCATGTACTCTCTGCAAGGGATCCACGGGG
ACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCATTTCTGTGGCTTCCAAGTCTTAGAACC
TCAACTGACATATAGCATTGGGCACACTCCAGCAGACGCCCGAATTCAAATCCTGGAAGGATGGAAGAAA
CGCCTGGAGAATATTTGGGATGAGACACCACTGTATTTTGCTCCAAGCAGCCTCTTTGACCTAAACTTCC
AGGCAGGATTCTTAATGAAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCAACTGACAACCAGATCAAAGCTAGAAAATGA
```

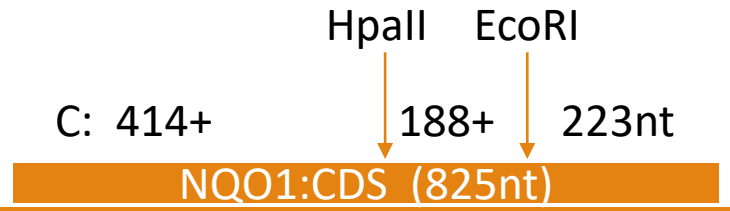
HpaII c|cgg 415

EcoRI g|aattc 603

AluI ag|ct 274, 815

```
>414 bp linear fragment from linear parent NM_000903.3:122-946
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTAT
GCCATGAAGGAGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCG
GACCTCTATGCCATGAACTTCAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTG
AAGGACCCTGCGAACTTTCAGTATCCTGCCGAGTCTGTTCTGGCTTATAAAGAAGGCCAT
CTGAGCCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAGCCGCAGACCTTGTGATATTC
CAGTTCCCCCTGCAGTGGTTTGGAGTCCCTGCCATTCTGAAAGGCTGGTTTGAGCGAGTG
TTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCTTC

>223 bp linear fragment from linear parent NM_000903.3:122-946
AATTCAAATCCTGGAAGGATGGAAGAAACGCCTGGAGAATATTTGGGATGAGACACCACT
GTATTTTGCTCCAAGCAGCCTCTTTGACCTAAACTTCCAGGCAGGATTCTTAATGAAAAA
AGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGGCCATCACTTGGG
CAAGTCCATCCCAACTGACAACCAGATCAAAGCTAGAAAATGA

>188 bp linear fragment from linear parent NM_000903.3:122-946
CGGAGTAAGAAGGCAGTGCTTTCCATCACCACTGGTGGCAGTGGCTCCATGTACTCTCTG
CAAGGGATCCACGGGGACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCAT
TTCTGTGGCTTCCAAGTCTTAGAACCTCAACTGACATATAGCATTGGGCACACTCCAGCA
GACGCCCG
```





Cloning-Fragment analysis

NQO1:CDS-restiction digest

```
>NM_000903.3:122-946 Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1),
transcript variant 1, mRNA
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTATGCCATGAAGG
AGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCGGACCTCTATGCCATGAACTT
CAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTGAAGGACCCTGCGAACTTTCAGTATCCTGCC
GAGTCTGTTCTGGCTTATAAAGAAGGCCATCTGAGCCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAG
CCGCAGACCTTGTGATATTCCAGTTCCCCCTGCAGTGGTTTGGAGTCCCTGCCATTCTGAAAGGCTGGTT
TGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCTTCCGGAGT
AAGAAGGCAGTGCTTTCCATCACCACTGGTGGCAGTGGCTCCATGTACTCTCTGCAAGGGATCCACGGGG
ACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCATTTCTGTGGCTTCCAAGTCTTAGAACC
TCAACTGACATATAGCATTGGGCACACTCCAGCAGACGCCCGAATTCAAATCCTGGAAGGATGGAAGAAA
CGCCTGGAGAATATTTGGGATGAGACACCACTGTATTTTGCTCCAAGCAGCCTCTTTGACCTAAACTTCC
AGGCAGGATTCTTAATGAAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCAACTGACAACCAGATCAAAGCTAGAAAATGA
```

HpaII c|cgg	415

EcoRI g|aattc	603

AluI ag|ct	274, 815

Restriction Digest

Restriction Digest cleaves a DNA sequence in a virtual restriction digest, with one, two, or title specifying their length, their position in the original sequence, and the enzyme sites th molecules (by entering more than one sequence in FASTA format). Use Restriction Digest

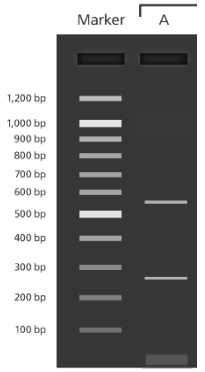
Paste the raw sequence or one or more FASTA sequences into the text area below. Input

```
CGCCTGGAGAATATTTGGGATGAGACACCACTGTATTTTGCTCCAAGCAGCCTCTTT
GACCTAAACTTCC
AGGCAGGATTCTTAATGAAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTG
GCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCAACTGACAACCAGATCAAAGCTAGAAAATGA
```

Submit Clear Reset

- Treat sequences as linear molecules.
- Digest with AluI ag|ct and nothing and nothing .

```
>541 bp linear fragment from linear parent NM_000903.3:122-946
CTGGAAGCCGCAGACCTTGTGATATTCCAGTTCCCCCTGCAGTGGTTTGGAGTCCCTGCC
ATTCTGAAAGGCTGGTTTGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCC
ATGTATGACAAAGGACCCTTCCGGAGTAAGAAGGCAGTGCTTTCCATCACCACTGGTGGC
AGTGGCTCCATGTACTCTCTGCAAGGGATCCACGGGGACATGAATGTCATTCTCTGGCCA
ATTCAGAGTGGCATTCTGCATTTCTGTGGCTTCCAAGTCTTAGAACCTCAACTGACATAT
AGCATTGGGCACACTCCAGCAGACGCCCGAATTCAAATCCTGGAAGGATGGAAGAAACGC
CTGGAGAATATTTGGGATGAGACACCACTGTATTTTGCTCCAAGCAGCCTCTTTGACCTA
AACTTCCAGGCAGGATTCTTAATGAAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAA
TTTGGCCTTTCTGTGGGCCATCACTTGGGCAAGTCCATCCCAACTGACAACCAGATCAAA
G

>273 bp linear fragment from linear parent NM_000903.3:122-946
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTAT
GCCATGAAGGAGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCG
GACCTCTATGCCATGAACTTCAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTG
AAGGACCCTGCGAACTTTCAGTATCCTGCCGAGTCTGTTCTGGCTTATAAAGAAGGCCAT
CTGAGCCCAGATATTGTGGCTGAACAAAAGAAG

>11 bp linear fragment from linear parent NM_000903.3:122-946 [
CTAGAAAATGA
```



A: 273+ 541+ 11nt

NQO1:CDS (825nt)

Cloning-Fragment analysis

- Treat sequences as linear molecules.
- Digest with EcoRI g|aattc and HpaII c|cgg and AluI ag|ct .

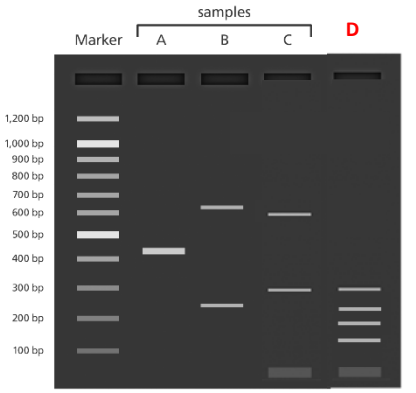
NQO1:CDS-restiction digest

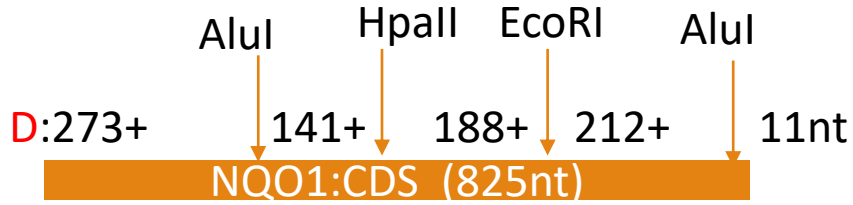
>NM_000903.3:122-946 Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), transcript variant 1, mRNA

```
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTATGCCATGAAGG
AGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCGGACCTCTATGCCATGAACTT
CAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTGAAGGACCCTGCGAACTTTCAGTATCCTGCC
GAGTCTGTTCTGGCTTATAAAGAAGGCCATCTGAGCCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAG
CCGCAGACCTTGTGATATTCCAGTTCCCCCTGCAGTGGTTTGGAGTCCCTGCCATTCTGAAAGGCTGGTT
TGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCTTCCGGAGT
AAGAAGGCAGTGCTTTCCATCACCACTGGTGGCAGTGGCTCCATGTACTCTCTGCAAGGGATCCACGGGG
ACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCATTTCTGTGGCTTCCAAGTCTTAGAACC
TCAACTGACATATAGCATTGGGCACACTCCAGCAGACGCCCGAATTCAAATCCTGGAAGGATGGAAGAAA
CGCCTGGAGAATATTTGGGATGAGACACCACTGTATTTTGCTCCAAGCAGCCTCTTTGACCTAAACTTCC
AGGCAGGATTCTTAATGAAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCAACTGACAACCAGATCAAAGCTAGAAAATGA
```

HpaII c|cgg	415
EcoRI g|aattc	603
AluI ag|ct	274, 815

```
>273 bp linear fragment from linear parent NM_000903.3:122-946
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTAT
GCCATGAAGGAGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCG
GACCTCTATGCCATGAACTTCAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTG
AAGGACCCTGCGAACTTTCAGTATCCTGCCGAGTCTGTTCTGGCTTATAAAGAAGGCCAT
CTGAGCCCAGATATTGTGGCTGAACAAAAGAAG

>212 bp linear fragment from linear parent NM_000903.3:122-946
AATTCAAATCCTGGAAGGATGGAAGAAACGCCTGGAGAATATTTGGGATGAGACACCACT
GTATTTTGCTCCAAGCAGCCTCTTTGACCTAAACTTCCAGGCAGGATTCTTAATGAAAAA
AGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGGCCATCACTTGGG
CAAGTCCATCCCAACTGACAACCAGATCAAAG

>188 bp linear fragment from linear parent NM_000903.3:122-946
CGGAGTAAGAAGGCAGTGCTTTCCATCACCACTGGTGGCAGTGGCTCCATGTACTCTCTG
CAAGGGATCCACGGGGACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCAT
TTCTGTGGCTTCCAAGTCTTAGAACCTCAACTGACATATAGCATTGGGCACACTCCAGCA
GACGCCCG

>141 bp linear fragment from linear parent NM_000903.3:122-946
CTGGAAGCCGCAGACCTTGTGATATTCCAGTTCCCCCTGCAGTGGTTTGGAGTCCCTGCC
ATTCTGAAAGGCTGGTTTGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCC
ATGTATGACAAAGGACCCTTC

>11 bp linear fragment from linear parent NM_000903.3:122-946 F
CTAGAAAATGA
```





Practical part

- Using **Restriction summary** find RE which cuts you sequence 1x and another that cuts 2x

- simulate restriction in **Restriction Digest** what will you see on the gel after such digestion?

Homework 8

Work with „your“ nucleotide sequence.

1) Find primers for the detection of your gene that would be specific (other transcript variants allowed).
2) Check the primers by multalin (CDS,mRNA and F, R)
3) Find out if and how many times cut your CDS following REs: EcoRI, BamHI, XbaI
4) Find out an REs that cut just once and on that cuts twice. Simulate the restriction digest of your CDS by both enzymes, what will be the products?