

# 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

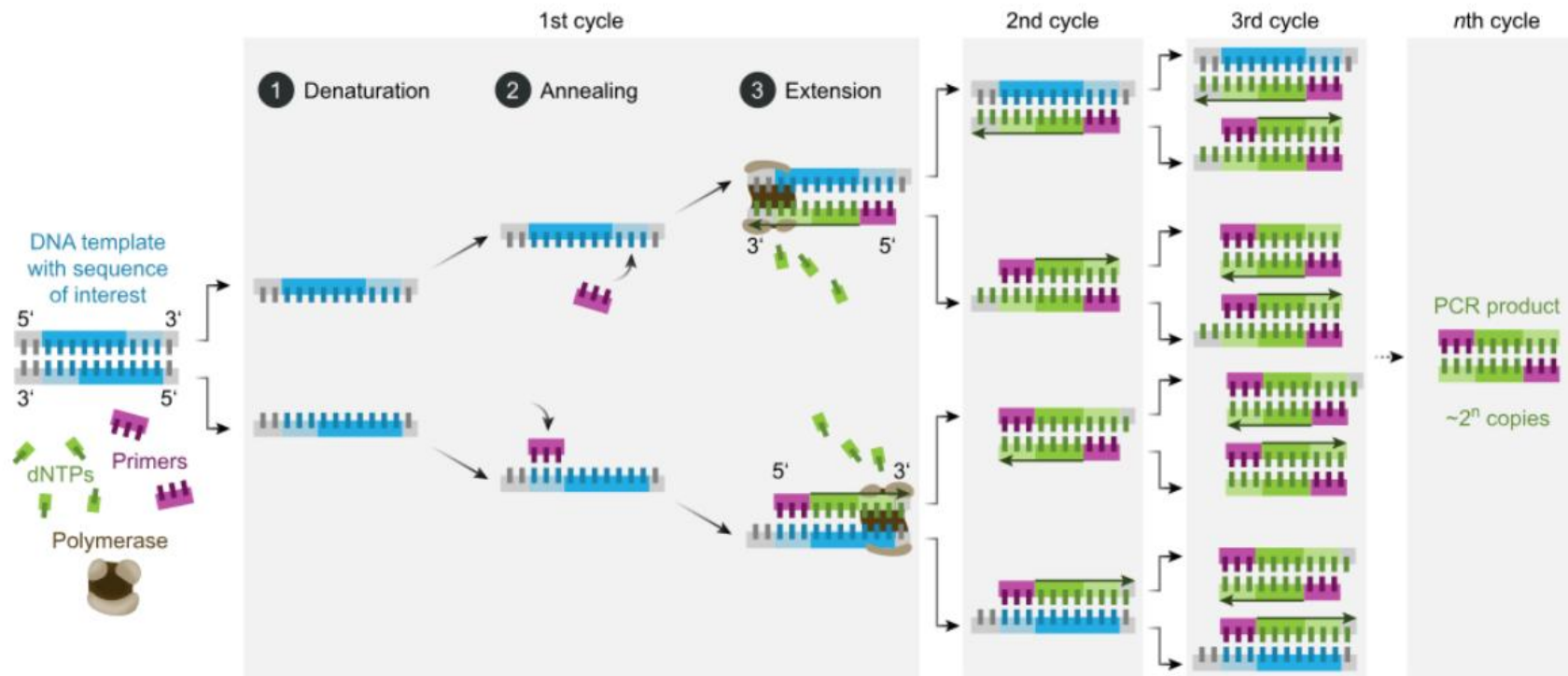
microRNA

Genomes....

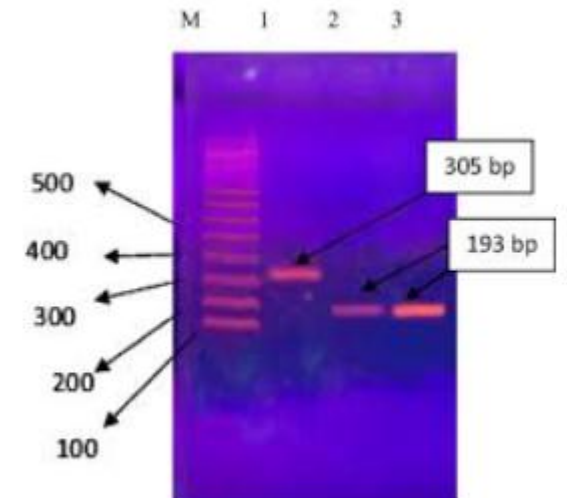
....

# 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...

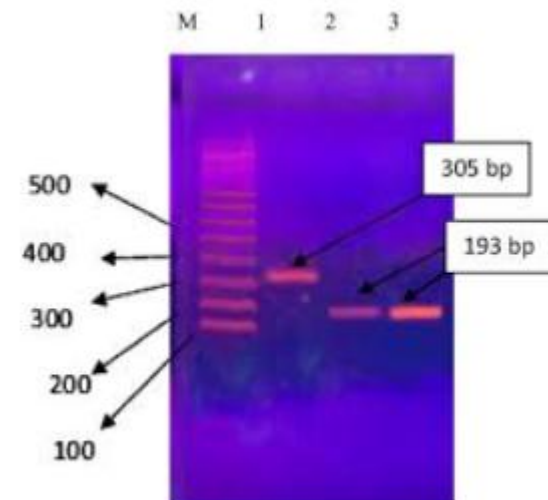
**Product:** various length, between two primers

**Primers:** manual design

-forward-check by OligoCalc

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

→ „End point“ detection



# 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!)

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Search Advanced Help

GenBank Send to: Change region shown Customize view

### Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), transcript variant 1, mRNA

NCBI Reference Sequence: NM\_000903.2

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

Go to:

LOCUS	NM_000903	2601 bp	mRNA	linear	FRI 29-MAR-2018
DEFINITION	Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), transcript variant 1, mRNA.				
ACCESSION	NM_000903				
VERSION	NM_000903.2				
KEYWORDS	RefSeq.				
SOURCE	Homo sapiens (human)				
ORGANISM	<a href="#">Homo sapiens</a> Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 2601)				
AUTHORS	Cheng X, Liu F, Liu H, Wang G and Hao H.				
TITLE	Enhanced glycometabolism as a mechanism of NQO1 potentiated growth of NSCLC revealed by metabolomic profiling				
JOURNAL	Biochem. Biophys. Res. Commun. 496 (1), 31-36 (2018)				
PUBMED	<a href="#">29291405</a>				
REMARK	GeneRIF: Taken together, we proposed that NQO1 could potentiate NSCLC cell proliferation by enhancing cellular glycometabolism, and				

Analyze this sequence  
Run BLAST  
**Pick Primers**  
Highlight Sequence Features  
Find in this Sequence  
Show in Genome Data Viewer

Articles about the NQO1 gene  
Redox modulation of NQO1. [PLoS One. 2018]  
Enhanced glycometabolism as a mechanism of NQO1 po [Biochem Biophys Res Commun. 2018]  
RNA-binding activity of TRIM25 is mediated by its PRY/SPRY domain and is requi [BMC Biol. 2017]  
See all...

# Specific gene detection

## Gene detection

- Specificity matters! (Primer BLAST!)

**Primer-BLAST** *A tool for finding specific primers*

NCBI/ **Primer-BLAST**: Finding primers specific to your PCR template (using Primer3 and BLAST).

[Reset page](#) [Save search parameters](#) [Retrieve recent results](#) [Publication](#) [Tips for finding specific primers](#)

**PCR Template**

Enter accession, gi, or FASTA sequence (A refseq record is preferred) [Clear](#)

NM\_000903.2

Range

Forward primer  From  To  [Clear](#)

Reverse primer

Or, upload FASTA file

**Primer Parameters**

Use my own forward primer (5'→3' on plus strand)  [Clear](#)

Use my own reverse primer (5'→3' on minus strand)  [Clear](#)

PCR product size

Min  Max

# of primers to return

Primer melting temperatures ( $T_m$ )

Min  Opt  Max  Max  $T_m$  difference  [Clear](#)

**Exon/intron selection**

A refseq mRNA sequence as PCR template input is required for options in the section [Clear](#)

Exon junction span  [Clear](#)

Exon junction match

Exon at 5' side  Exon at 3' side

# Specific gene detection

## Gene detection

- Specificity matters! (Primer BLAST!)

Primer-BLAST *A tool for finding specific primers*

NCBI/ Primer-BLAST: Finding primers specific to your PCR template (using Primer3 and BLAST).

[Reset page](#) [Save search parameters](#) [Retrieve recent results](#) [Publication](#) [Tips for finding specific primers](#)

**PCR Template**

Enter accession, gi, or FASTA sequence (A refseq record is preferred) [Clear](#)

NM\_000903.2

Range

Forward primer From 196 To [Clear](#)

Reverse primer 1016

Or, upload FASTA file [Procházet...](#)

**Primer Parameters**

Use my own forward primer (5'→3' on plus strand) [Clear](#)

Use my own reverse primer (5'→3' on minus strand) [Clear](#)

PCR product size Min 200 Max 500

# of primers to return 10

Primer melting temperatures (T<sub>m</sub>) Min 57.0 Opt 60.0 Max 63.0 Max T<sub>m</sub> difference 3 [Clear](#)

**Exon/intron selection**

A refseq mRNA sequence as PCR template input is required for options in the section [Clear](#)

Exon junction span No preference [Clear](#)

Exon junction match Exon at 5' side Exon at 3' side

when CDS displayed



# Specific gene detection

## Gene detection

- Specificity matters! (Primer BLAST!)

Note: Parameter values that differ from the default are highlighted in yellow

### Primer Pair Specificity Checking Parameters

**Specificity check** ☒ Enable search for primer pairs specific to the intended PCR template

**Search mode** Automatic

**Database** Refseq mRNA

**Exclusion** ☐ Exclude predicted Refseq transcripts (accession with XM, XR prefix) ☐ Exclude uncultured/environmental sample sequences

**Organism** 9606  
Enter an organism name (or organism group name such as enterobacteriaceae, rodents), taxonomy id or select from the suggestion list as you type.  
[Add more organisms](#)

**Entrez query (optional)**

**Primer specificity stringency** Primer must have at least 2 total mismatches to unintended targets, including at least 2 mismatches within the last 5 bps at the 3' end. Ignore targets that have 6 or more mismatches to the primer.

**Max target size** 4000

**Allow splice variants** ☒ Allow primer to amplify mRNA splice variants (requires refseq mRNA sequence as PCR template input)

**Get Primers** ☐ Show results in a new window ☒ Use new graphic view

[Advanced parameters](#)

Note: Parameter values that differ from the default are highlighted in yellow

Only human sequences

# Specific gene detection

## Gene detection

- Specificity matters! (Primer BLAST!)

Primer-BLAST

Primer-Blast results

NCBI/ Primer-BLAST : results: Job id=DwXQRHukdgxRNmAzbVNEARdIVTM6W04uOw [more...](#)

**Input PCR template** [NM\\_000903.2](#) Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1), transcript variant 1, mRNA  
Range 196 - 1016  
**Specificity of primers** Primer pairs are specific to input template as no other targets were found in selected database: Refseq mRNA (Organism limited to Homo sapiens)  
[Other reports](#) [Search Summary](#)

**Graphical view of primer pairs**

Genes

NP\_000894.1

exon

Primer pairs for job DwXQRHukdgxRNmAzbVNEARdIVTM6W04uOw

Primer 1

Primer 2

Primer 3

Primer 4

Primer 5

Primer 6

Primer 7

Primer 8

Primer 9

Primer 10

NM\_000903.2: 72..1.1K (1.1Kbp)

Tools

Tracks

Tracks shown: 3/12

# 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

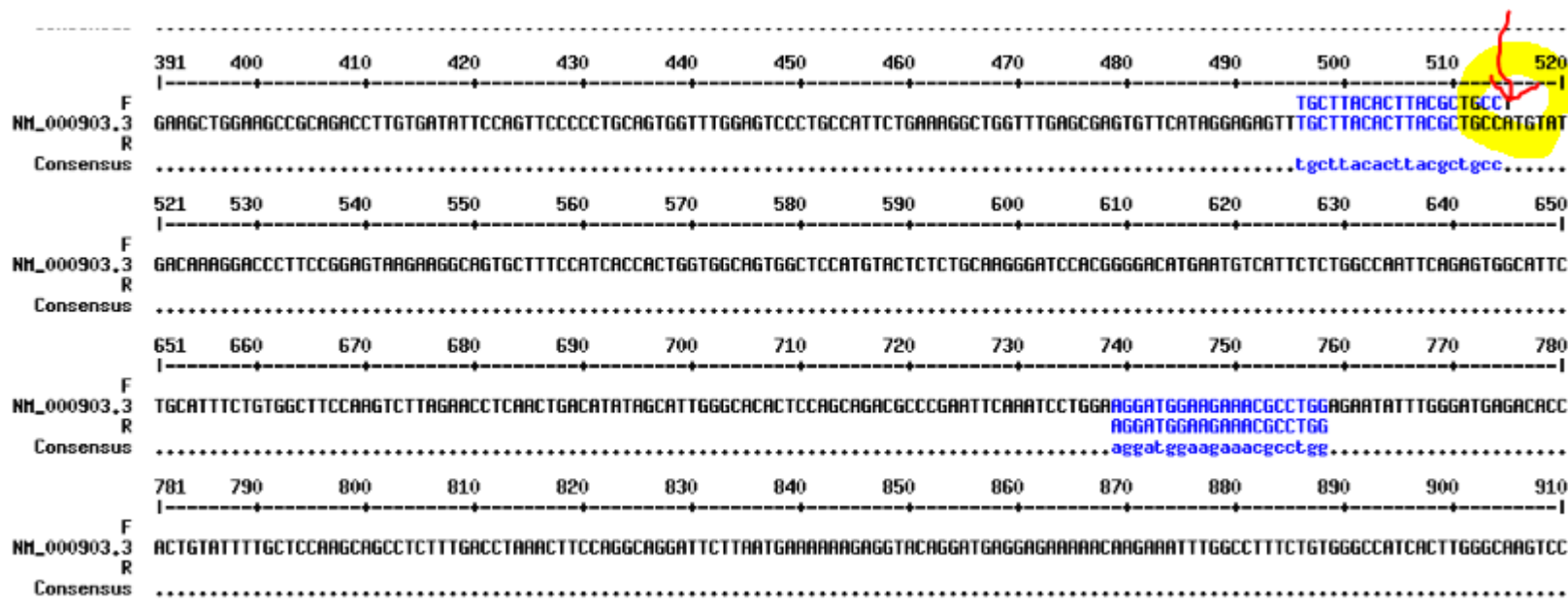
Verify if the primers can be used for detection of NQO1 ([NM\\_000903.3](#)) (use multalin)

# 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

Primer-BLAST

A tool for finding specific primers

Finding primers specific to your PCR template (using Primer3 and BLAST).

Primers for target on one template

Primers common for a group of sequences

Retrieve recent results

Publication

Tips for finding specific primers

PCR Template

Enter accession, gi, or FASTA sequence (A refseq record is preferred) ? 

Clear

NM\_000903.3

Or, upload FASTA file 

Zvolit soubor

 Nevybrán žádný soubor

Range ? 

Clear

Forward primer 

From

To

Reverse primer 

From

To

Primer Parameters

Use my own forward primer (5'->3' on plus strand) 

TGCTTACACTTACGCTGCCT

Clear

Use my own reverse primer (5'->3' on minus strand) 

CCAGGCGTTTCTCCATCCT

Clear

PCR product size 

Min

70

Max

1000

# of primers to return 

10

Primer melting temperatures (T<sub>m</sub>) 

Min

57.0

Opt

60.0

Max

63.0

Max T<sub>m</sub> difference

3

 ?

Exon/intron selection

A refseq mRNA sequence as PCR template input is required for options in the section ?

Exon junction span 

No preference

 ?

Exon junction match 

Min 5' match

7

Min 3' match

4

Max 3' match

8

Minimal and maximal number of bases that must anneal to exons at the 5' or 3' side of the junction ?

☐ Primer pair must be separated by at least one intron on the corresponding genomic DNA ?

Intron inclusion 

Min

Max

Intron length range 

Min

Max

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

# Primers verification – control

---

**Primer-BLAST** » JOB ID:lpXIMMEBzKnrk1yWUfZ4pCvtaZYG\_nKLBw

Primer-BLAST Results ?

❗ Specified left primer cannot be found in template...make sure this primer is on the plus strand of your template.

FOLLOW NCBI



Connect with NLM



National Library of Medicine  
8600 Rockville Pike  
Bethesda, MD 20894

Web Policies  
FOIA  
HHS Vulnerability Disclosure

Help  
Accessibility  
Careers

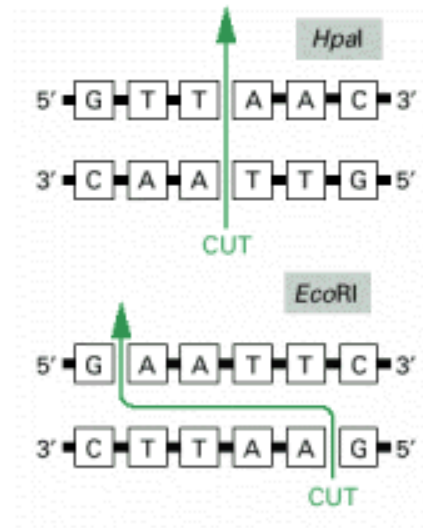
NLM | NIH | HHS | USA.gov



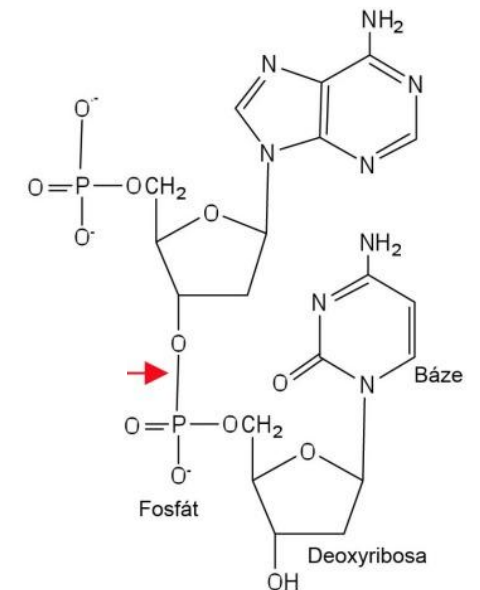
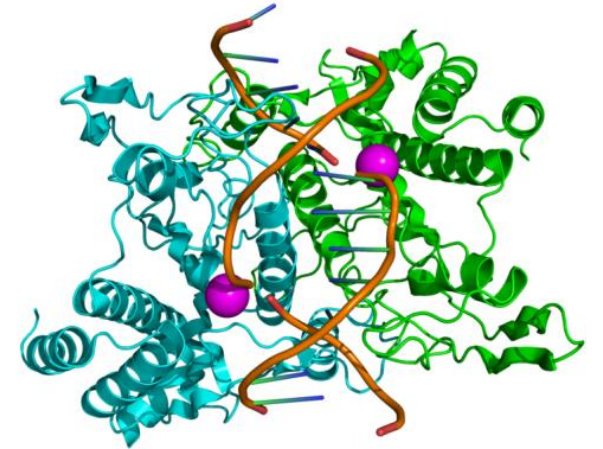
# Restriction analysis

## Restriction endonucleases (molecular scissors)

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



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



# 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 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
ctaaattgtaagcgttaatatTTTTgttaaaattcgcggttaaattttgttaaatacagctca
TTTTtaaccaataggccgaaatcggaataatcccttataaatcaaaagaatagaccgaga
tagggttgagtgtgttccagtttggaacaagagtgccactattaaagaacgtggactcaa
cgtcaaaggcgcaaaaccgtctatcaggcgatggccactacgtgaaccatcacccata
tcaagtttttggggtcgaggtgccgtaaacgtaaacggaaccctaaggaggagccccc
```

Please check the [browser compatibility page](#) before using this program.

Submit

Clear

Reset

- Treat sequences as  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"

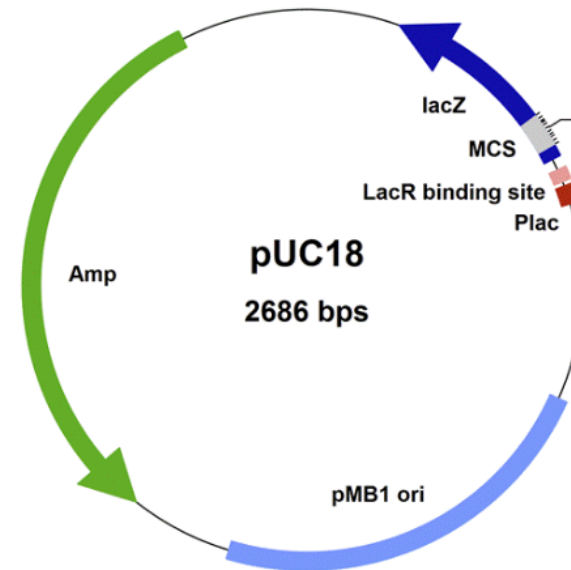
Site:	Positions:
AatI agg cct	none
AatII gacgt c	none
Acc16I tgc gca	480, 2269
AccII cg cg	36, 412, 432, 456, 622, 624, 664, 795, 1001, 1003, 1201, 1782, 2112, 2605, 2937
AccIII t ccgga	none
AcI aa cgtt	2273, 2646
AcVI cac gtg	none
AfaI gt ac	758, 2527
AfeI agc gct	none
AfII c ttaag	none
AgeI a ccggt	none
AhlI a ctagt	684
Alw44I 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
Ascl qalcacacc	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  
Sall  
XbaI  
BamHI  
AvaI  
SmaI  
XmaI  
Acc65I  
KpnI  
BamII  
Eco53k  
SacI  
ApoI  
EcoRI

# Cloning-Fragment analysis

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

## Program SMS: „Restriction digest“

**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

**Sequence Manipulation Suite:**  
**Restriction Digest**

Restriction Digest cleaves a DNA sequence in a virtual restriction d

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

```
CGCCTGGAGAATATTTGGGATGAGACCACTGTATTTTGTCCAAGCAGCCTCTTT
GACCTAAACTTCC
AGGCAGGATTCTTAATGAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTG
GCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCAACTGACAACCAGATCAAAGCTAGAAAATGA
```

Submit

Clear

Reset

Treat sequences as

linear

molecules.

Digest with

AluI agl

and

linear

and

nothing

and

nothing

\*This page requires JavaS

\*You can mirror this page or use it off-line.

Mon Nov 6 02:56:29 2017

Valid XHTML 1.0; Valid CSS

nothing  
AatI agg|cct  
AatII gacgt|c  
Acc16I tgc|gca  
AccII cg|cg  
AccIII t|ccgga  
AclI aa|cgtt  
AclI cac|gtg  
AfaI gt|ac  
AfeI agc|gct  
AflII c|ttaag  
AgaI a|ccggt  
AhlI a|ctagt  
Alw441 g|tgcac  
AluI ag|ct  
Aor51HI agc|gct  
ApaI gggcc|c  
ApaLI g|tgcac  
AscI gg|cgcgcc  
AseI at|taat

or three restriction enzymes. The resulting fragments are sorted by size, and they are given a  
that produced them. You can digest linear or circular molecules, and even a mixture of  
st to determine the fragment sizes you will see when you perform a digest in the lab.  
t limit is 100,000,000 characters.

[new window](#) | [home](#) | [citation](#)

# Cloning-Fragment analysis

## NQ01:CDS-restriction summary

>NM\_000903.3:122-946 Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQ01), transcript variant 1, mRNA  
ATGGTCGGCAGAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTATGCCATGAAGG  
AGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCGGACCTCTATGCCATGAACCT  
CAATCCCATCATTTCCAGAAAGGACATCACAGGTAACCTGAAGGACCCCTGCGAACTTTCAGTATCCTGCC  
GAGTCTGTTCTGGCTTATAAAGAGGCCATCTGAGCCAGATATTGTGGCTGAACAAAAGAGCTGGAAG  
CCGCAGACCTTGTGATATCCAGTTCCTCCCTGCAGTGGTTTGGAGTCCCTGCCATTCGTAAGAGGCTGGT  
TGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCCTCCGGAGT  
AAGAAGGCAGTGCTTTCCATCACCACCTGGTGGCAGTGGCTCCATGTACTCTCTGCAAGGGATCCACGGGG  
ACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCATTTCTGTGGCTTCCAAGTCTTAGAACC  
TCAACTGACATATAGCATTGGGCACACTCCAGCAGACGCCGAATTCAAATCCTGGAAGGATGGAAGAAA  
CGCTGGAGAAATATTGGGATGAGACACCACTGTATTTTGTCTCCAAGCAGCCTCTTTGACCTAAACTCC  
AGGCAGGATTCTTAATGAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGG  
CCATCACTTGGGCAAGTCCATCCCACTGACACCAAGATCAAAGCTAGAAAAATGA

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

## NQ01:CDS-restriction 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

```
CGCCTGGAGAAATATTGGGATGAGACACCACTGTATTTTGTCTCAAGCAGCCTCTTT
GACCTAAACTTCC
AGGCAGGATTCCTTAATGAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTG
GCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCACTGCAACCCAGATCAAAGCTAGAAAAATGA
```

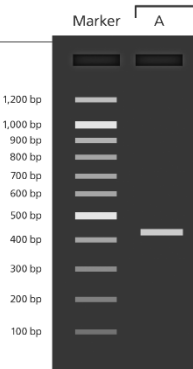
Submit Clear Reset

- Treat sequences as  molecules.
- Digest with  and  and .

>414 bp linear fragment from linear parent NM\_000903.3:122-946  
ATGGTCGGCAGAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTAT  
GCCATGAAGGAGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCG  
GACCTCTATGCCATGAACCTCAATCCCATCATTTCCAGAAAGGACATCACAGGTAACCTG  
AAGGACCCCTGCGAACTTTCAGTATCCTGCCAGTCTGTTCTGGCTTATAAAGAGGCCAT  
CTGAGCCAGATATTGTGGCTGAACAAAAGAGCTGGAAGCCCGACAGCCTTGTGATATTC  
CAGTTCCTCCCTGCAGTGGTTTGGAGTCCCTGCCATTCGTAAGGCTGGTTTGAGCGAGTG  
TTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCCTTC

>411 bp linear fragment from linear parent NM\_000903.3:122-946  
CGGAGTAAGAAGGCAGTGCTTTCCATCACCACCTGGTGGCAGTGGCTCCATGTACTCTCTG  
CAAGGGATCCACGGGGACATGAATGTCTCTCTGGCCAATTGAGAGTGGCATTCTGCAT  
TTCTGTGGCTTCCAAGTCTTAGAACCTCAACTGACATATAGCATTGGGCACACTCCAGCA  
GACGCCCGAATTCAAATCCTGGAAGGATGGAAGAAACGCCTGGAGAAATATTGGGATGAG  
ACACCACTGTATTTTGTCTCCAAGCAGCCTCTTTGACCTAAACTTCCAGGCAGGATTCCTTA  
ATGAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGGCCAT  
CACTTGGGCAAGTCCATCCCACTGACAAACCAATCAAAGCTAGAAAAATGA

A: 414+ 411nt  
NQ01:CDS (825nt)



## EcoRI

>602 bp linear fragment from linear parent NM\_000903.3:122-946  
ATGGTCGGCAGAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTAT  
GCCATGAAGGAGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCG  
GACCTCTATGCCATGAACCTCAATCCCATCATTTCCAGAAAGGACATCACAGGTAACCTG  
AAGGACCCCTGCGAACTTTCAGTATCCTGCCAGTCTGTTCTGGCTTATAAAGAGGCCAT  
CTGAGCCAGATATTGTGGCTGAACAAAAGAGCTGGAAGCCCGACAGCCTTGTGATATTC  
CAGTTCCTCCCTGCAGTGGTTTGGAGTCCCTGCCATTCGTAAGGCTGGTTTGAGCGAGTG  
TTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCCTCCGGAGT  
AAGAAGGCAGTGCTTTCCATCACCACCTGGTGGCAGTGGCTCCATGTACTCTCTGCAAGGG  
ATCCACGGGGACATGAATGTCTCTGGCCAATTGAGAGTGGCATTCTGCATTTCTGT  
GGCTTCCAAGTCTTAGAACCTCAACTGACATATAGCATTGGGCACACTCCAGCAGACGCC  
CG

>223 bp linear fragment from linear parent NM\_000903.3:122-946  
AATTCAAATCCTGGAAGGATGGAAGAAACGCCTGGAGAAATATTGGGATGAGACACCACT  
GTATTTTGTCTCAAGCAGCCTCTTTGACCTAAACTCCAGGCAGGATTCTTAATGAAAAA  
AGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGGCCCTCACTTGGG  
CAAGTCCATCCCACTGACAAACCAATCAAAGCTAGAAAAATGA

B: 602+ 223nt  
NQ01:CDS (825nt)

# Cloning-Fragment analysis

- Treat sequences as  molecules.
- Digest with  and  and .

## NQ01:CDS-restiction digest

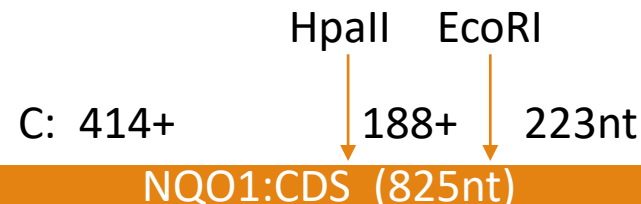
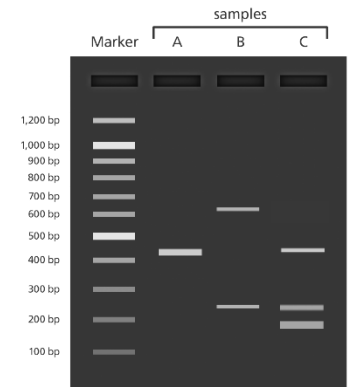
```
>NM_000903.3:122-946 Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1),
transcript variant 1, mRNA
ATGGTCGGCAGAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTATGCCATGAAGG
AGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCGGACCTCTATGCCATGAACCT
CAATCCCATCATTTCCAGAAAGGACATCACAGSTAACTGAAGGACCCCTGCGAACTTTCAGTATCCTGCC
GAGTCTGTTCTGGCTTATAAAGAAGGCCATCTGAGCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAG
CCGACACCTTGTGATATTCAGTTCCCTTGCAGTGGTTTGGAGTCCCTGCCATTCTGAAAGGCTGGTT
TGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCCTCCCGAGT
AAGAAGGCAGTGTCTTCCATCACCCTGGTGGCAGTGGCTCCATGTACTCTCTGCAAGGGATCCACGGGG
ACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCATTCTCTGGGCTTCCAAGTCTTAGAACC
TCAACTGACATATAGCATTGGGCACACTCCAGCAGACGCCCGAATTCAAATCTGGAAGGATGGAAGAAA
CGCCTGGAGATATTGGGATGAGACACCACTGATTTTGTCTCCAAGCAGCCTCTTTGACCTAACTTCC
AGGCAGGATTCTTAATGAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCACTGACACCAAGATCAAAGCTAGAAAAATGA
```

HpaII c cgg	415	→
EcoRI g aatc	603	→
AluI ag ct	274, 815	

```
>414 bp linear fragment from linear parent NM_000903.3:122-946
ATGGTCGGCAGAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTAT
GCCATGAAGGAGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCG
GACCTCTATGCCATGAACCTTCAATCCCATCATTTCCAGAAAGGACATCACAGGTAACTG
AAGGACCCCTGCGAACTTTCAGTATCCTGCCGAGTCTGTTCTGGCTTATAAAGAAGGCCAT
CTGAGCCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAGCCGACACCTTGTGATATTC
CAGTTCCCCCTGCAGTGGTTTGGAGTCCCTGCCATTCTGAAAGGCTGGTTTGGAGCGAGTG
TTCATAGGAGAGTGTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCCTTC
```

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

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





# Cloning-Fragment analysis

## 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

```
CGCCTGGAGAATATTGGGATGAGACCACTGTATTTGTCCCAAGCAGCCTCTT
GACCTAAACTTCC
AGGCAGGATTCTTAATGAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTG
GCCCTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCACTGACAACAGATCAAAGCTAGAAAATGA
```

Submit Clear Reset

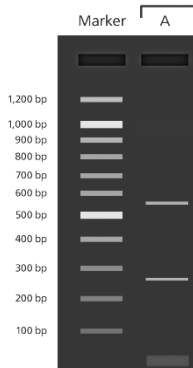
- Treat sequences as  molecules.
- Digest with  and  and .

```
>541 bp linear fragment from linear parent NM_000903.3:122-946
CTGGAAGCCGAGACCTTTGTGATATTCAGTTCCTCCCTGCAGTGGTTTGGAGTCCCTGCC
ATTTCTGAAAGGCTGGTTTGAGCGAGTGTTCATAGGAGAGTTTGTCTACACTTACGCTGCC
ATGTATGACAAAGGACCTTCCGGAGTAAGAAGGCAGTGTCTTCCATCACCAGTGGTGGC
AGTGGCTCCATGTACTCTCTGCAAGGGATCCACGGGGACATGAATGTCATTCTCTGGCCA
ATTCAGAGTGGCATTCTGCATTTCTGTGGCTTCCAAGTCTTAGAACCTCAACTGACATAT
AGCATTGGGCACACTCCAGCAGACGCCCGAATTCAAATCCTGGAAGGATGGAAGAAACGC
CTGGAGAATATTTGGGATGAGACACCACTGTATTTTGTCTCCAAGCAGCCTCTTTGACCTA
AACTTCCAGGCAGGATTCTTAATGAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAA
TTTGGCCTTTCTGTGGGCCATCACTTGGGCAAGTCCATCCCACTGACAACAGATCAAA
G
```

```
>273 bp linear fragment from linear parent NM_000903.3:122-946
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTAT
GCCATGAAGGAGGCTGTCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCG
GACCTCTATGCCATGAACCTCAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTG
AAGGACCCCTGCGAATCTTCAGTATCCTGCCGAGTCTGTTCTGGCTTATAAAGAAGGCCAT
CTGAGCCCAGATATTGTGGCTGAACAAAAGAAG
```

>11 bp linear fragment from linear parent NM\_000903.3:122-946  
CTGGAAGCTTGA

A: 273+ 541+ 11nt  
NQO1:CDS (825nt)



## NQO1:CDS-restriction digest

```
>NM_000903.3:122-946 Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1),
transcript variant 1, mRNA
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTATGCCATGAAG
AGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCGGACCTCTATGCCATGAACCT
CAATCCCATCATTTCCAGAAAGGACATCACAGGTAACTGAAGGACCCCTGCGAACTTTCAGTATCCTGCC
GAGTCTGTTCTGGCTTATAAAGAAGGCCATCTGAGCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAG
CCGACAGCCTTGTGATATTCAGTTCCTCCCTGCACTGGTGGAGTCCCTGCCATTCTGAAAGGCTGGTT
TGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCCTCCGAGT
AAGAAGGCAGTGTCTTCCATCACCAGTGGTGGCAGTGGCTCCATGTACTCTCTGCAAGGGATCCACGGGG
ACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCATTTCTGTGGCTTCCAGTCTTAGAACC
TCAACTGACATATAGCATTGGGCACACTCCAGCAGACGCCCGAATTCAAATCCTGGAAGGATGGAAGAAA
CGCCTGGAGATATTGGGATGAGACACCACTGTATTTTGTCTCCAAGCAGCCTCTTTCACCTAAACTCC
AGGCAGGATTCTTAATGAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCACTGACAACCAAGATCAAAGCTAGAAAATGA
```

HpaII c|cgg 415

EcoRI g|aatc 603

AluI ag|ct 274, 815

A: 273+ 541+ 11nt  
NQO1:CDS (825nt)

# Cloning-Fragment analysis

- Treat sequences as  molecules.
- Digest with  and  and .

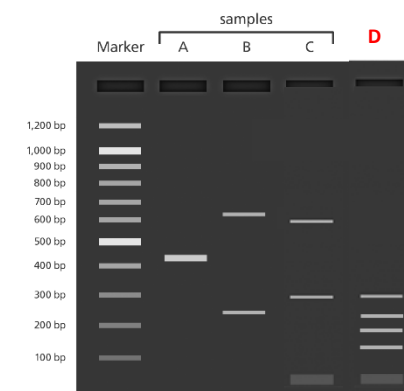
```
>273 bp linear fragment from linear parent NM_000903.3:122-946
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTAT
GCCATGAAGGAGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCG
GACCTCTATGCCATGAACCTCAATCCCATCATTTCCAGAAAGGACATCACAGGTAAACTG
AAGGACCCCTGCGAACTTTCAGTATCCTGCCGAGTCTGTTCTGGCTTATAAAGAAGGCCAT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

```
>212 bp linear fragment from linear parent NM_000903.3:122-946
AATTCAAATCCTGGAAGGATGGAAGAAACGCCTGGAGAATATTTGGGATGAGACCCACT
GTATTTTGTCTCCAAGCAGCCTCTTTGACCTAAACTCCAGGCAGGATTCTTAATGAAAAA
AGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGCCATCACTTGGG
CAAGTCCATCCCAACTGACAACCATCAAAAG
```

```
>188 bp linear fragment from linear parent NM_000903.3:122-946
CGGAGTAAGAAGGCAGTGCTTTCCATCACCCTGGTGGCAGTGGCTCCATGTACTCTCTG
CAAGGGATCCACGGGGACATGAATGTATTCTCTGGCCAATTCAGAGTGGCATTCTGCAT
TTCTGTGGCTTCCAAGTCTTAGAACCTCACTGACATATAGCATTGGGCACACTCCAGCA
GACGCCCG
```

```
>141 bp linear fragment from linear parent NM_000903.3:122-946
CTGGAAGCCGACAGCTTGTGATATTCCAGTTCCTCCCTGCAGTGGTTTGGAGTCCCTGCC
ATTCTGAAAGGCTGGTTTGGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCC
ATGTATGACAAAGGACCTTC
```

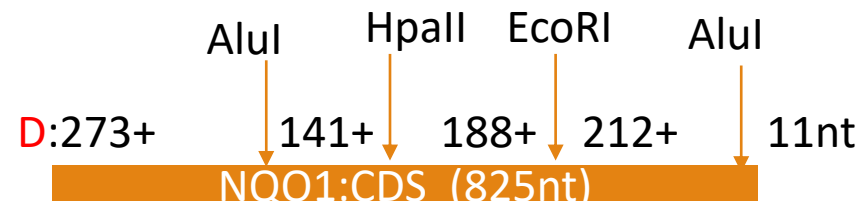
```
>11 bp linear fragment from linear parent NM_000903.3:122-946
CTAGAAATGA
```



## NQO1:CDS-restiction digest

```
>NM_000903.3:122-946 Homo sapiens NAD(P)H quinone dehydrogenase 1 (NQO1),
transcript variant 1, mRNA
ATGGTCGGCAGAAGAGCACTGATCGTACTGGCTCACTCAGAGAGGACGTCCTTCAACTATGCCATGAAG
AGGCTGCTGCAGCGGCTTTGAAGAAGAAAGGATGGGAGGTGGTGGAGTCGGACCTCTATGCCATGAACCT
CAATCCCATCATTTCCAGAAAGGACATCACAGSTAACTGAAGGACCCCTGCGAACTTTCAGTATCCTGCC
GAGTCTGTTCTGGCTTATAAAGAAGGCCATCTGAGCCAGATATTGTGGCTGAACAAAAGAAGCTGGAAG
CCGACAGCCTTGTGATATCCAGTTCCTCCCTGCAGTGGTTTGGAGTCCCTGCCATTCTGAAAGGCTGGTT
TGAGCGAGTGTTCATAGGAGAGTTTGCTTACACTTACGCTGCCATGTATGACAAAGGACCCCTCCGAGT
AAGAAGGCAGTGCTTTCCATCACCCTGGTGGCAGTGGCTCCATGTACTCTCTGCAAGGGATCCACGGGG
ACATGAATGTCATTCTCTGGCCAATTCAGAGTGGCATTCTGCATTCTCTGTTGGCTTCCAAGTCTTAGAACC
TCAACTGACATATAGCATTGGGCACACTCCAGCAGACGCCCGAATTCAAATCCTGGAAGGATGGAAGAAA
CGCCTGGAGATATTGGGATGAGACACCACTGTATTTTGGCTCCAAGCAGCCTCTTTCAGCTAACTCC
AGGCAGGATTCTTAATGAAAAAAGAGGTACAGGATGAGGAGAAAAACAAGAAATTTGGCCTTTCTGTGGG
CCATCACTTGGGCAAGTCCATCCCACTGACAAACAGATCRAAGCTAGAAAAATGA
```

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





## 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?