



18.9.2025

Návrh a vývoj **zákaznických** nástrojů pro **kvalitativní a** **kvantitativní analýzu** **laboratorních dat**

Ing. Helena Vítková, Ph.D.

vitkovah@vut.cz

Skupina bioinformatiky a systémové biologie BioSys_BUT

• vedoucí: **doc. Mgr. Ing. Karel Sedlář, Ph.D.**

• 5 odborných asistentů

Mgr. Bc. Darina Čejková, Ph.D.

Ing. Helena Vítková, Ph.D.

Sudeep Roy, Ph.D.

Ing. Markéta Jakubíčková, Ph.D.

Ing. Jana Musilová, Ph.D.



• 7 PhD studentů

Ing. Kristýna Heřmánková

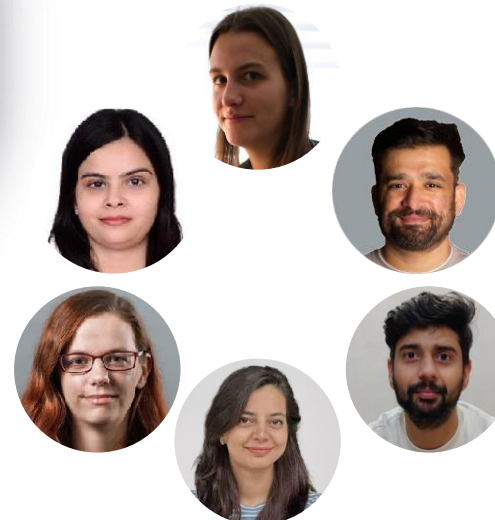
Ing. Kateřina Šabatová

Minoo Partovi Nasr

Mohammad Umair

M.Tech. Vaishali Pankaj

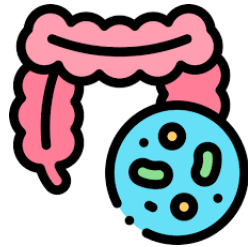
Mgr. Indeerjet Bhogal

UNIVERSITY of NEBRASKA
LINCOLN

**funkční
genomika**



**analýza
mikrobiomu**



***in silico*
návrh léčiv**



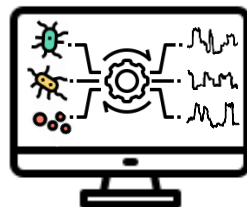
**systémová
biologie**



**zpracování
a analýza
sekvenačních dat**



**zpracování
genomických
signálů**



**genotypizace
bakteriálních
kmenů**

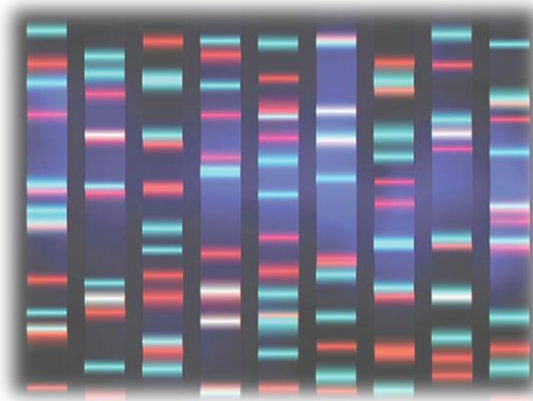


**tvorba nástrojů
pro rutinní
klinickou
mikrobiologii**



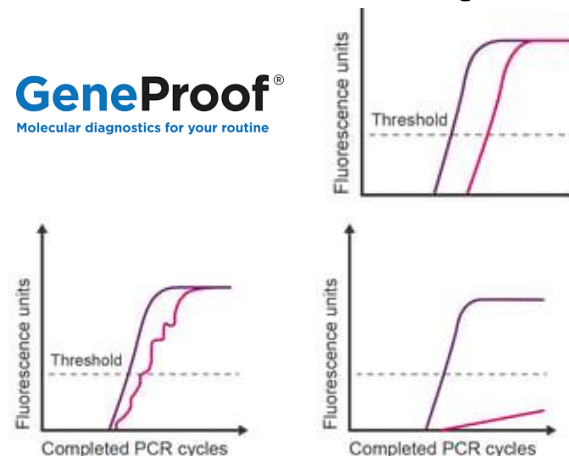
ZPRACOVÁNÍ ELEKTROFORETOGRAMU

- Úprava kvality obrazu
- Digitalizace výsledku
- Klasifikace vzorků
- Automatická čipová elektroforéza
- DNA fingerprinting



ANALÝZA AMPLIFIKAČNÍCH KŘIVEK qPCR

- Normalizace křivek
- Automatická klasifikace
- Autokorekce na základě průběhu měření
- Smluvní výzkum



ZPRACOVÁNÍ SEKVENAČNÍCH DAT

- Miniaturní mobilní sekvenátory
- Sekvence kdykoliv a kdekoliv
- Real-time analýza
- Rychlá diagnóza



Další nástroje

- Laboratorní databáze
- Převodní klíče
- Optimalizace volby markerů
- Statistické vyhodnocení

V ČEM TO UMÍME?

#BioSys_BUT

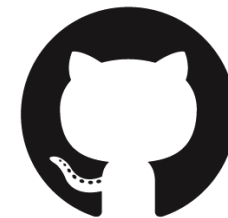
VÝVOJ:



UŽIVATELSKÁ ROZHRANÍ:



SDÍLENÍ:



GitHub

<https://github.com/BioSys-BUT/>

- ZADÁNÍ:

- Vyhodnocení vzájemné podobnosti vzorků formou dendrogramu
- Vzájemné srovnání více obrazů gelů

- PROBLÉMY:

- Znečištění vzorků
- Neuniformní proces přípravy gelu
- Nehomogenní průběh elektroforézy
- Kvalita snímku gelu



ZPRACOVÁNÍ ELEKTROFORETOGRAMU

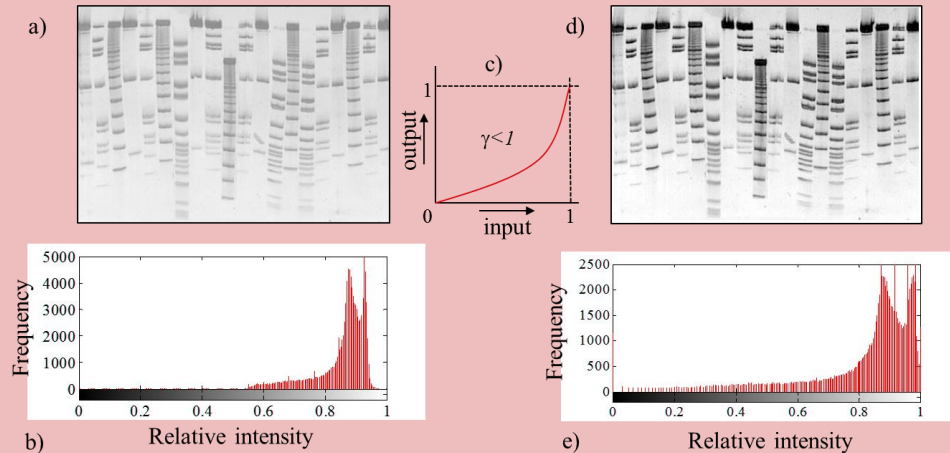
#BioSys_BUT

SW1: KLASIFIKACE VZORKŮ Z PLANÁRNÍ GELOVÉ ELEKTROFORÉZY

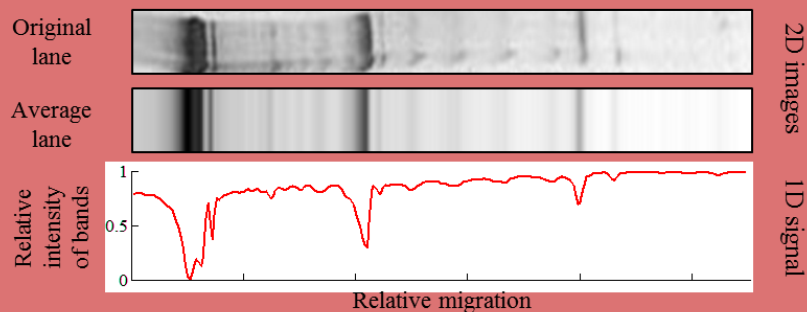
• I. Zpracování obrazu gelu

1. Zvýšení kvality obrazu:

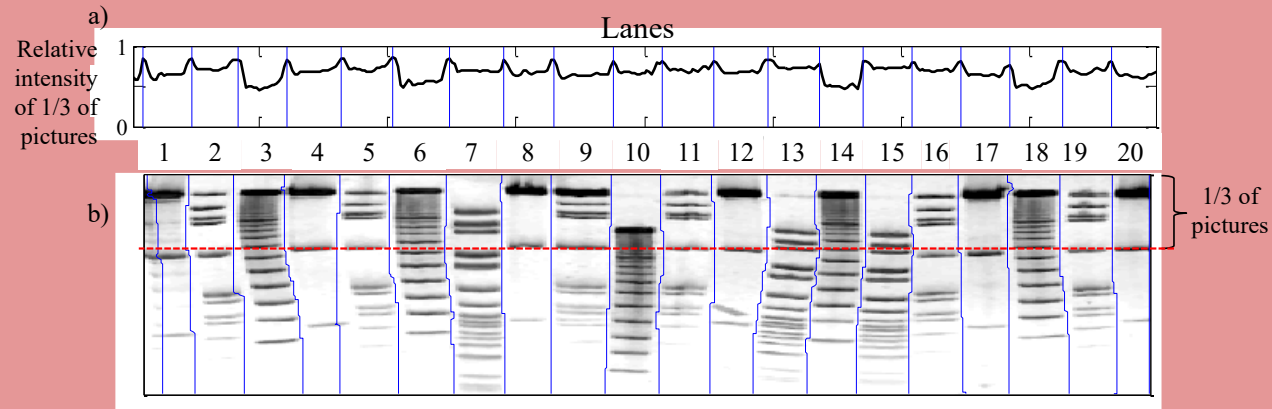
- transformace obrazu s gamma korekcí



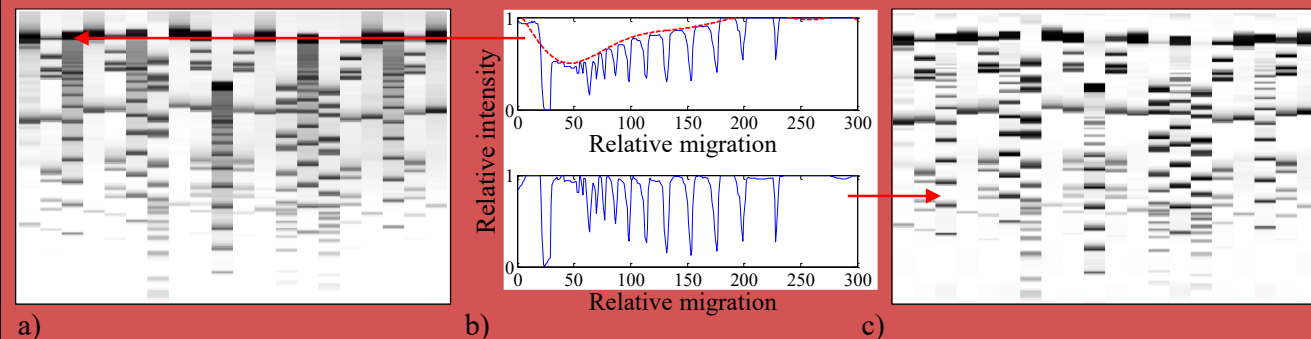
3. Transformace obrazu na signál



2. Automatická segmentace gelových linií



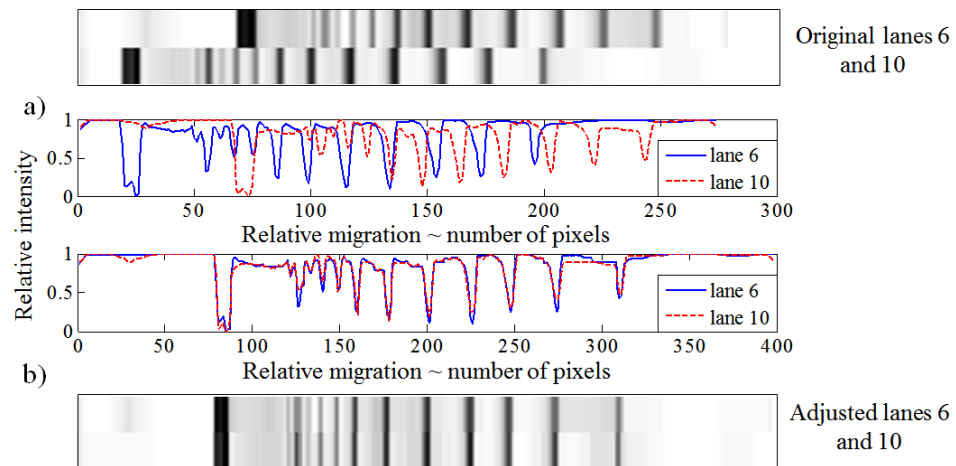
3. Odstranění nerovnoměrného pozadí (nečistoty)



• II. Klasifikace vzorků

1. Přizpůsobení vzájemné pozice bandů

- Aplikace dynamického borcení časové osy – dynamic time warping (DTW)



2. Stanovení vzájemné podobnosti všech párů vzorků

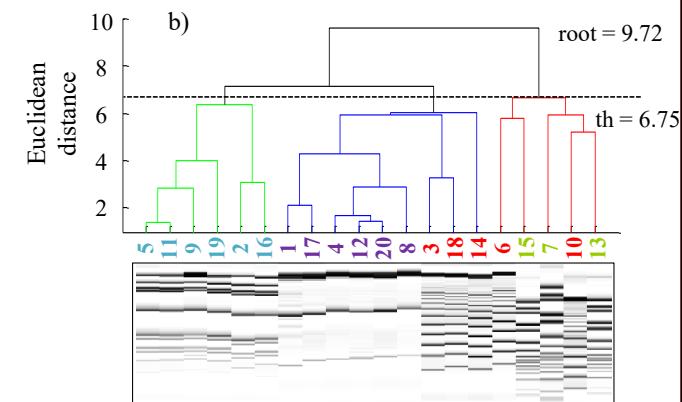
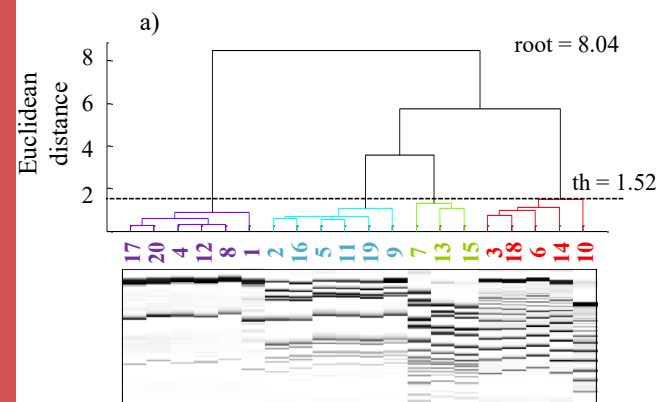
- Euklidova vzdálenost

$$d_{xy} = \sqrt{\sum_{i=1}^k (X_i - Y_i)^2}$$

3. Klasifikace formou dendrogramu

S DTW

bez DTW



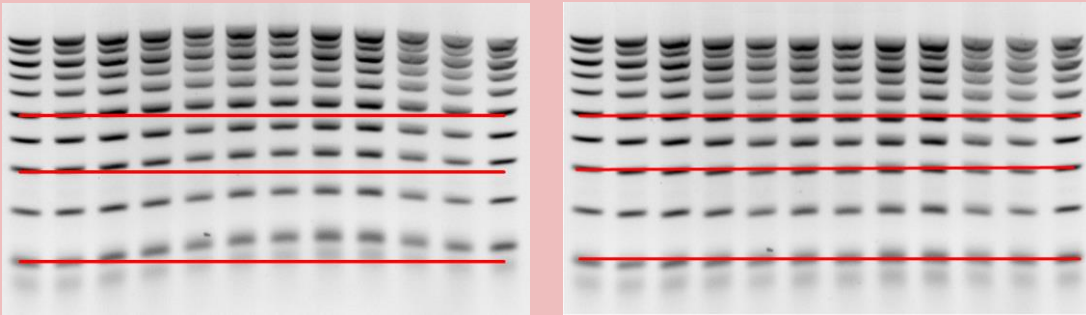
ZPRACOVÁNÍ ELEKTROFORETOGRAMU

#BioSys_BUT

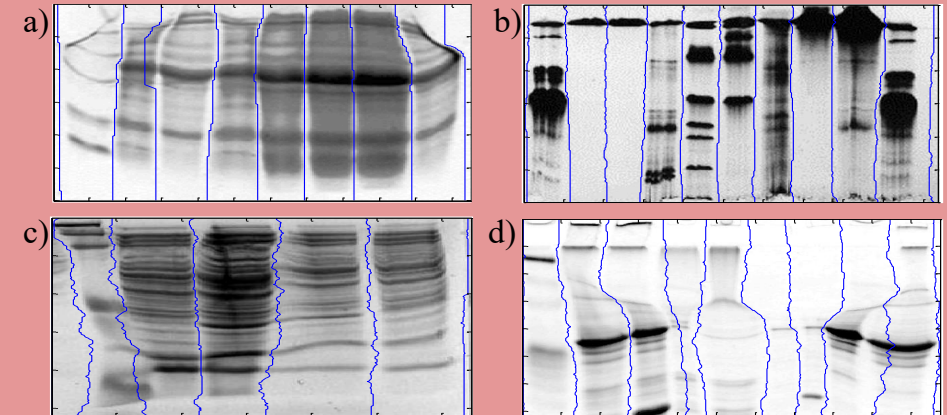
SW1: KLASIFIKACE VZORKŮ Z PLANÁRNÍ GELOVÉ ELEKTROFORÉZY

• Troubleshooting

1. „Smile“ efekt



3. Testování robustnosti segmentace

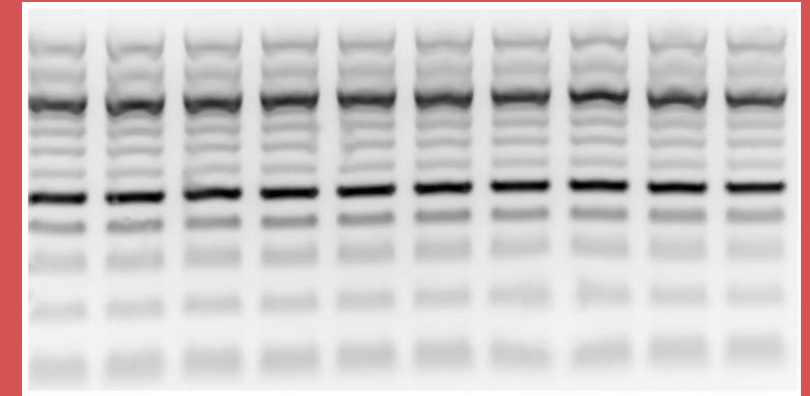
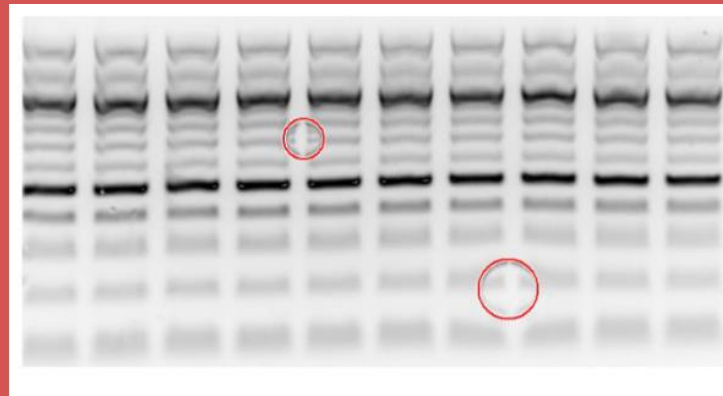
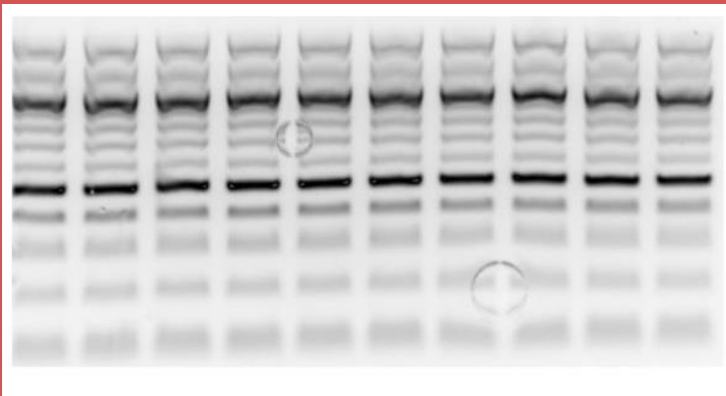


2. „Bubble“ efekt

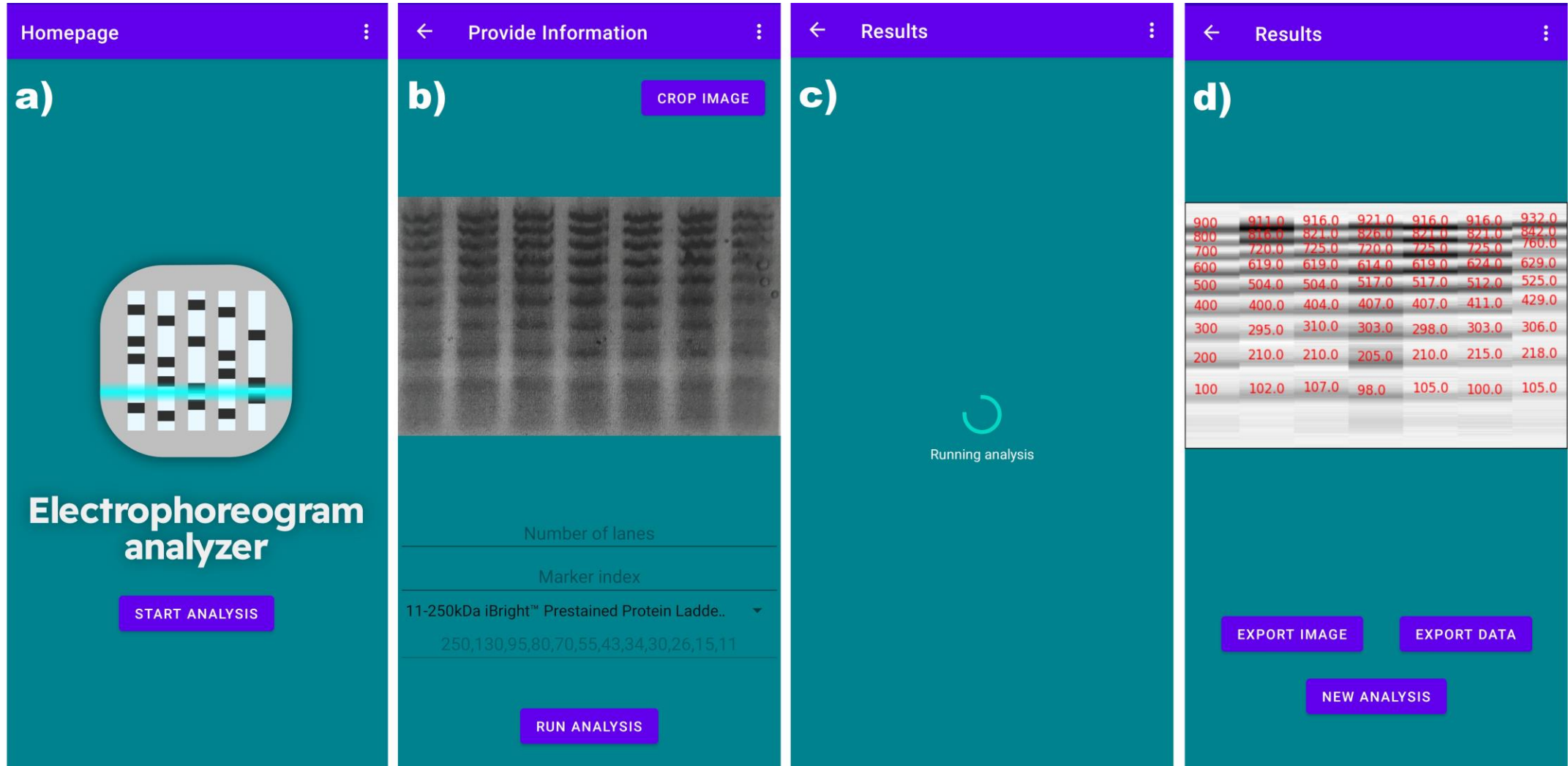
Originál

Detekce bublin

Výsledek korekce



Mobile APP: Anna Mojžišová - MOBILNÁ APLIKÁCIA NA SPRACOVANIE OBRAZU ELEKTROFORETICKÉHO GÉLU

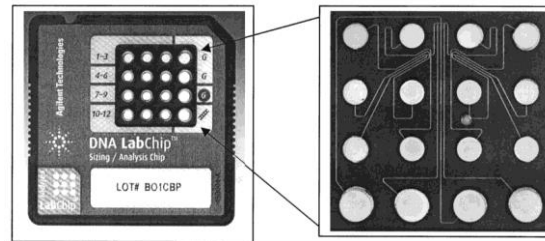


- ZADÁNÍ:

- Vyhodnocení vzájemné podobnosti vzorků formou dendrogramu
- Vzájemné srovnání více běhů

- PROBLÉMY:

- ~~Znečištění vzorků~~
- ~~Neuniformní proces přípravy gelu~~
- ~~Nehomogenní průběh elektroforézy~~
- ~~Kvalita snímku gelu~~
- Podobná délka fragmentů
→ potřeba vysokého rozlišení



ZPRACOVÁNÍ ELEKTROFORETOGRAMU

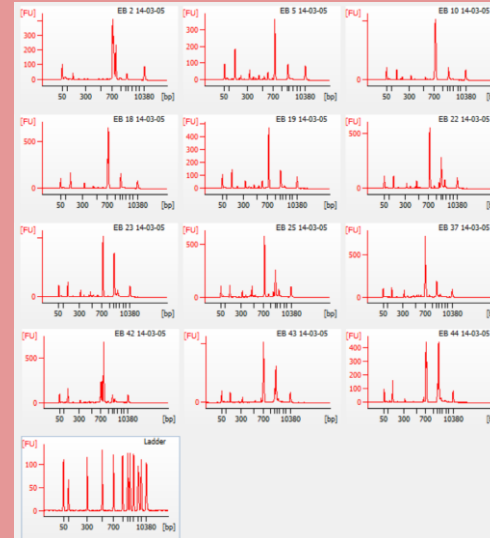
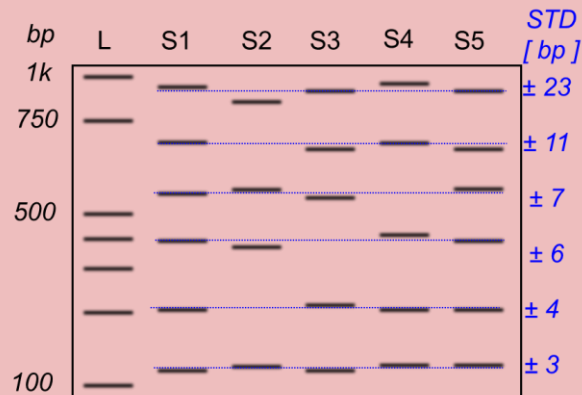
#BioSys_BUT

SW2: REP-PCR GENOMIC FINGERPRINTING s automatickou čipovou elektroforézou

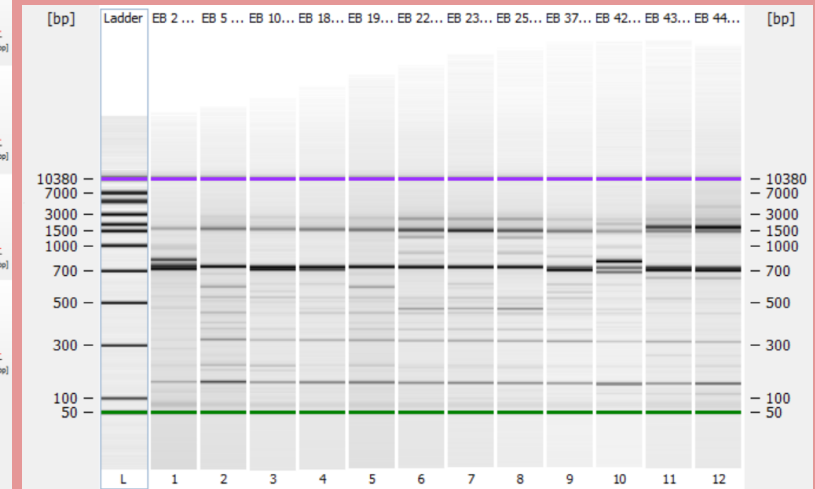
• I. ZAROVNÁNÍ BANDŮ

Bands position distorsion

S1	S2	S3	S4	S5
967	913	945	972	948
718	593	692	706	700
585	387	579	402	594
398	126	208	199	397
201		121	128	200
122				125



Electropherograms vs. Gel Image



Princip algoritmu:

S1	S2	S3	S4	S5
967	913	945	972	948
718	593	692	706	700
585	387	579	402	594
398	126	208	199	397
201		121	128	200
122				125

S1	S2	S3	S4	S5	mean
967	913	945	972	948	949
718		692	706	700	704
585	593	579		594	588
398	387		402	397	396
201		208	199	200	202
122	126	121	128	125	124

I.

S1	S2	S3	S4	S5
949	949	949	949	949
704		704	704	704
588	588	588		588
396	396		396	396
202		202	202	202
124	124	124	124	124

II.

S1	S2	S3	S4	S5
1	1	1	1	1
1	0	1	1	1
1	1	1	0	1
1	1	0	1	1
1	0	1	1	1
1	1	1	1	1

III.

ZPRACOVÁNÍ ELEKTROFORETOGRAMU

#BioSys_BUT

SW2: REP-PCR GENOMIC FINGERPRINTING s automatickou čipovou elektroforézou

• II. KLASIFIKACE VZORKŮ

Identity matrix	S1	S2	S3	S4	S5
	1	1	1	1	1
	1	0	1	1	1
	1	1	1	0	1
	1	1	0	1	1
	1	0	1	1	1
	1	1	1	1	1

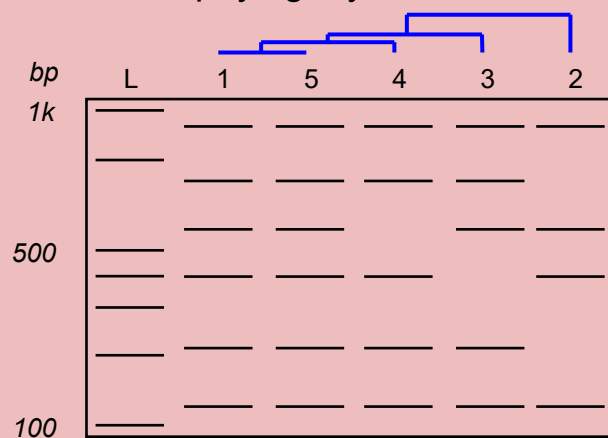
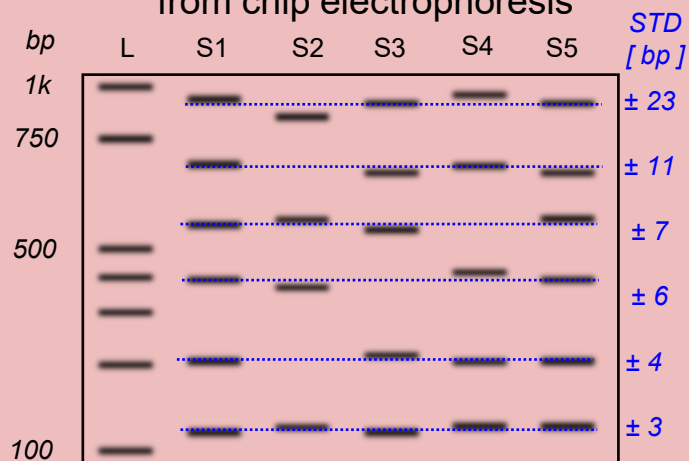
Pairwise distance matrix	S1	S2	S3	S4	S5
	*				
	2	*			
	1	3	*		
	1	3	2	*	
	0	2	1	1	*

→
Jaccard distance

ORIGINAL
Gel image

RESULT

Gel image after bands alignment
and phylogeny evaluation

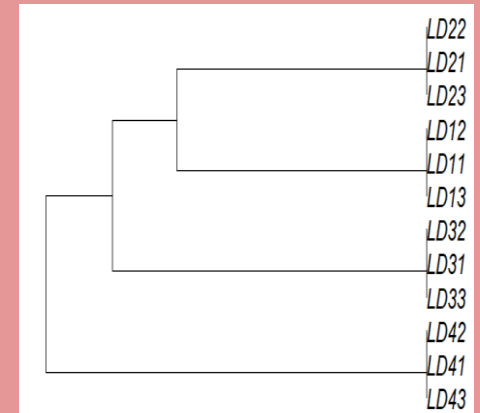


Zarovnání bandů – KALIBRACE



4 typy hmotnostních markerů:

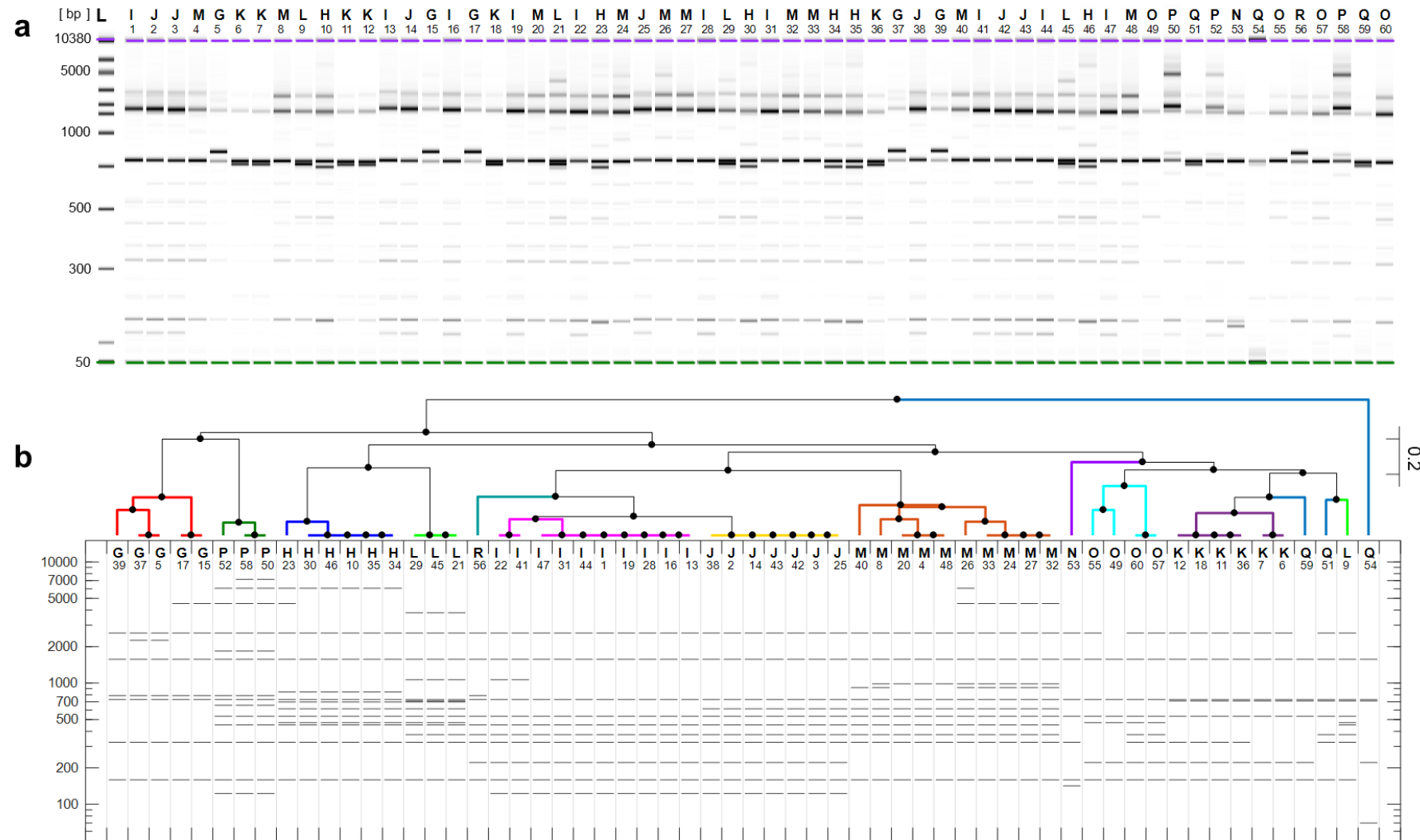
- GeneRuler_50bp_DNA Ladder
- GeneRuler_100bp_Plus
- GeneRuler_1kb_DNA Ladder
- OGeneRuler_1kb_DNA Ladder



• Výsledek klasifikace

Klinická data:

- 60 vzorků
- Pět běhů
- 12 bakteriálních kmenů
(velká písmena značí kmen)



ANALÝZA AMPLIFIKAČNÍCH KŘIVEK

SW3: SOFTWAREOVÝ NÁSTROJ PRO ANALÝZU AMPLIFIKAČNÍCH KŘIVEK QPCR

#BioSys_BUT

GeneProof[®]
Molecular diagnostics for your routine

- ZADÁNÍ:

- Klasifikace amplifikačních křivek probíhá ve skupinách deklarovaných výrobcem jako pozitivní (P) a negativní (N).
- Modul v jazyce Python integrovatelný do stávajícího obslužného SW zařízení.
- Smluvní výzkum

- ETAPY ŘEŠENÍ:

- Návrh klasifikačního algoritmu
- Optimalizace a verifikace algoritmu
- Tvorba finálního modulu



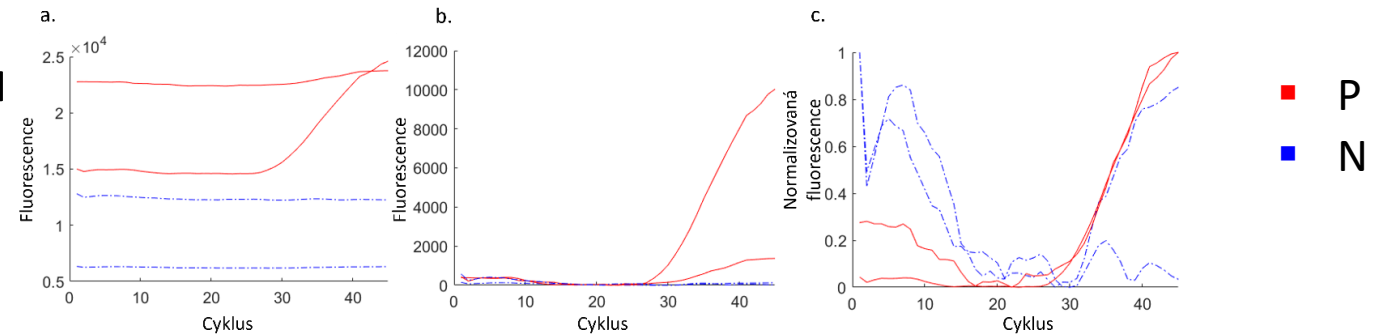
ANALÝZA AMPLIFIKAČNÍCH KŘIVEK

SW3: SOFTWAREVÝ NÁSTROJ PRO ANALÝZU AMPLIFIKAČNÍCH KŘIVEK QPCR

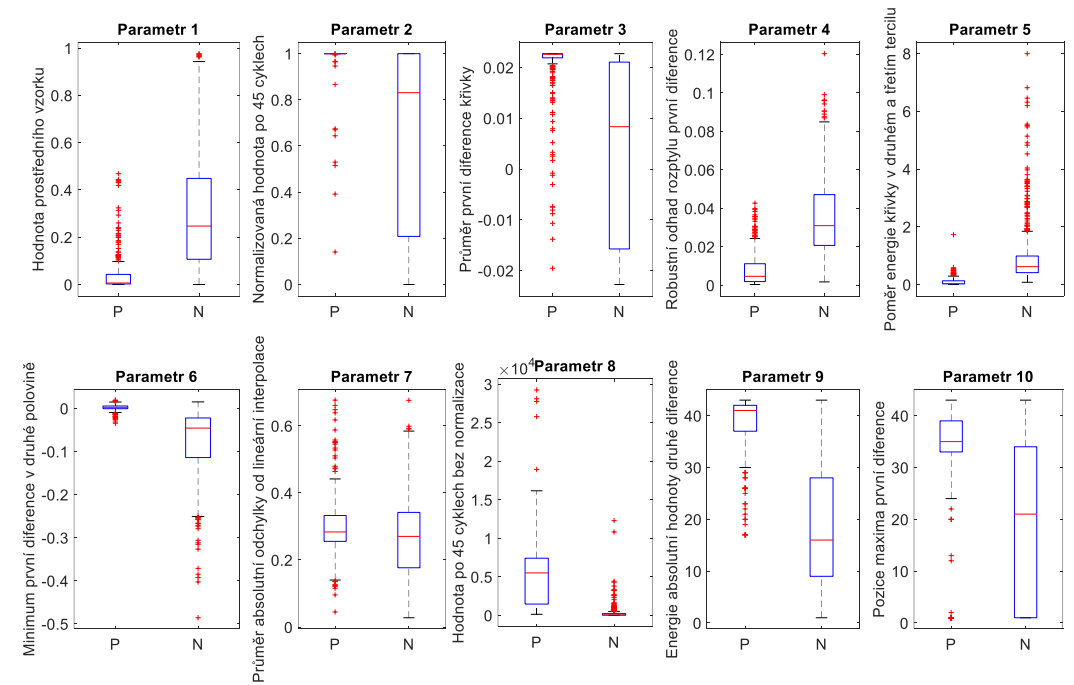
#BioSys_BUT

• I. ETAPA

- Návrh klasifikačního algoritmu
 - Volba klasifikačních parametrů
 - Nastavení prahových hodnot



Parametr	Popis parametru
P1	Hodnota prostředního normalizovaného vzorku
P2	Normalizovaná hodnota po 45 cyklech
P3	Průměr první difference křivky
P4	Robustní odhad rozptylu první difference křivky
P5	Poměr energie křivky v druhém a třetím tercilu (popisuje nárůst mezi prostředním plató a sigmoidálním nárůstem)
P6	Minimum první difference v druhé polovině křivky (minimální nárůst mezi 25-45 cyklem)
P7	Průměr absolutní odchylky od lineární interpolace křivky (křivka by neměla téměř lineárně růst)
P8	Hodnota po 45 cyklech bez normalizace (je použito i v 1 fázi klasifikace)
P9	Pozice maxima první difference (měla by být v druhé třetině křivky kde je nárůst nejstrmější)
P10	Energie absolutní hodnoty druhé difference (popisuje „kostrbatost“ křivky)

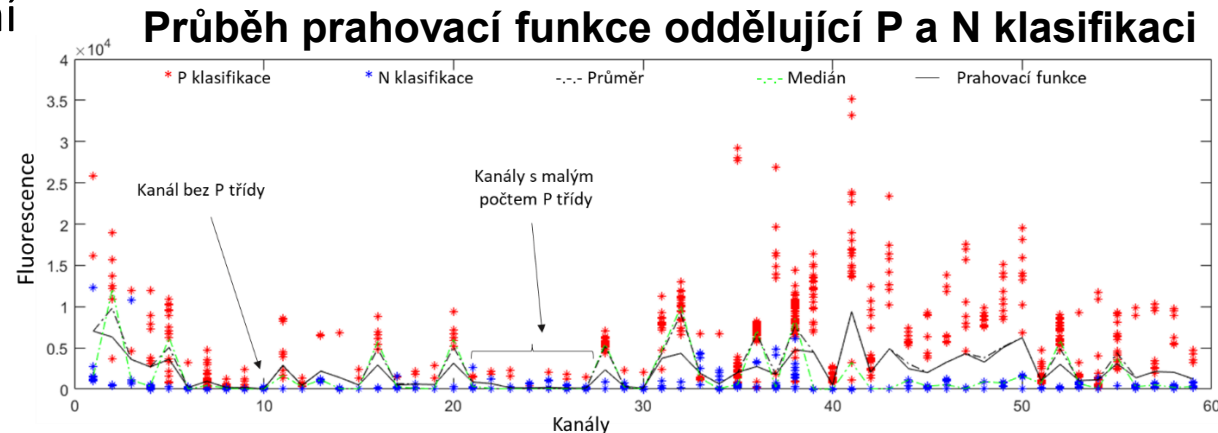
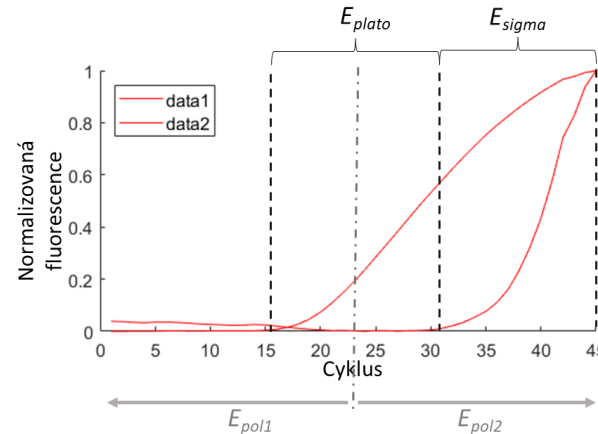


ANALÝZA AMPLIFIKAČNÍCH KŘIVEK

#BioSys_BUT

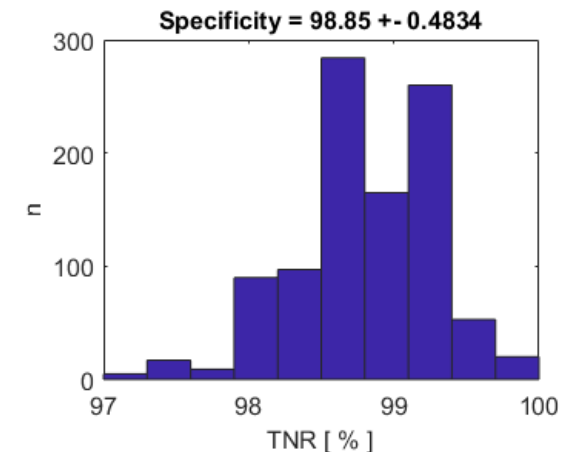
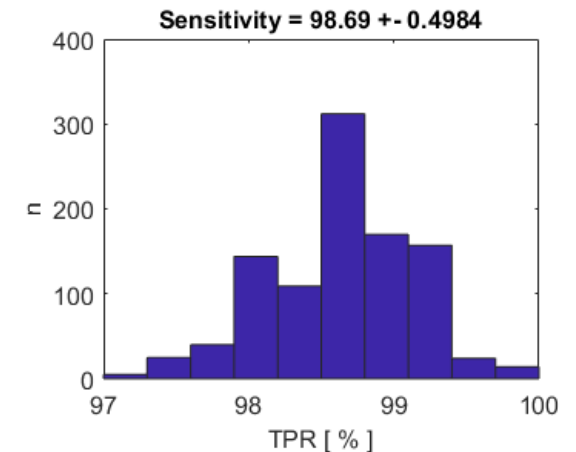
SW3: SOFTWAREVÝ NÁSTROJ PRO ANALÝZU AMPLIFIKAČNÍCH KŘIVEK QPCR

- II. ETAPA – III. ETAPA
- Optimalizace a verifikace algoritmu
 - Sestavení klasifikačního kritéria
 - Segmentace křivek
 - Formulace prahovací funkce
 - Statistické testování



	TPR [%]	TNR [%]	PVp [%]	PVn [%]	nV [%]	bACC [%]	F1 score [%]
První sada 294 křivek	97,12	98,92	98,06	98,40	1,36	98,02	97,58
Druhá sada 644 křivek	99,24	97,38	96,30	99,47	0	98,31	97,74

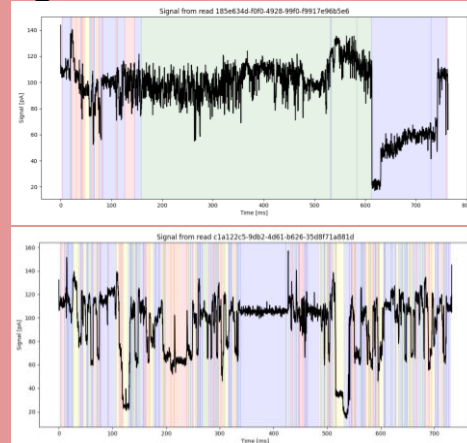
Histogramy kritérií kvality
klasifikace pro 1000
bootstrapových
pseudoreplikací:



- # NANOSLAST



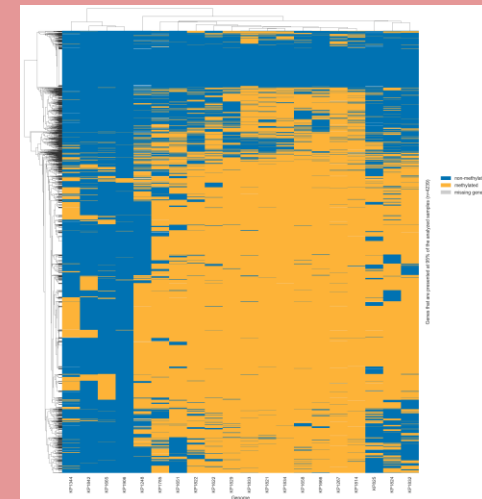
NANOSLAST
Nanopore Signal Local
Alignment and
Segmentation Tool



-

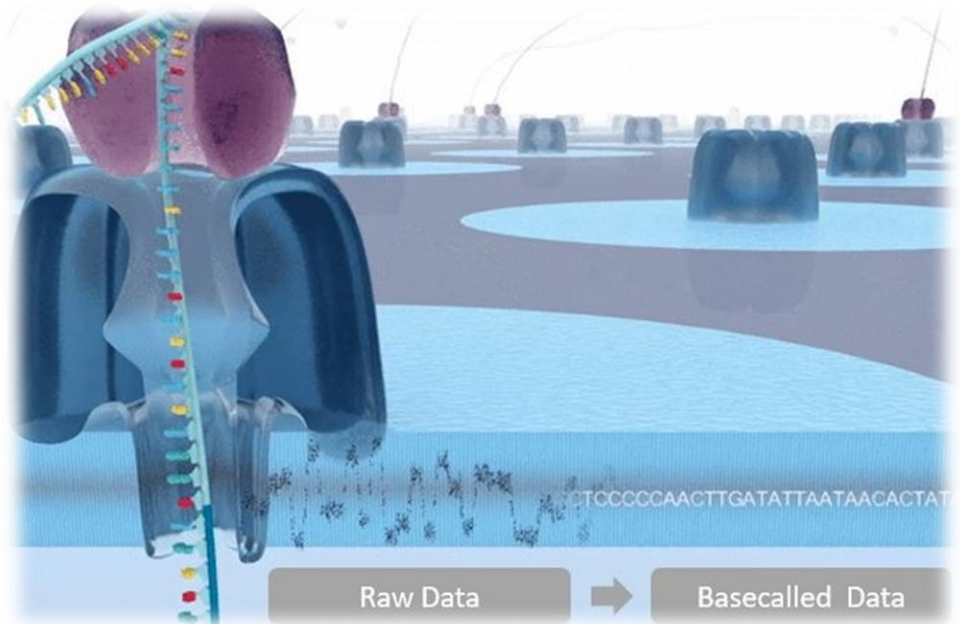
Largest Absolute Change in Total Methylation

CpG Site	CT-CCM (Total Methylation Count)	CT-DSM (Total Methylation Count)
cg48_0710 (Hb)	45	55
cg48_1448 (Cdk)	16	26
cg48_0880 (SP)	23	14
cg48_0830 (Hb)	13	5
cg48_0710 (Cdk)	12	5
cg48_0625 (Cdk)	19	12
cg48_0920 (SP)	15	9
cg48_1480 (Hb)	9	3
cg48_0920 (Cdk)	12	6
cg48_1410 (Hb)	10	4
cg48_1130 (Hb)	3	9
cg48_1440 (Hb)	18	12
cg48_0475 (Cdk)	9	15
cg48_1448 (Cdk)	10	16
cg48_0945 (Hb)	18	17

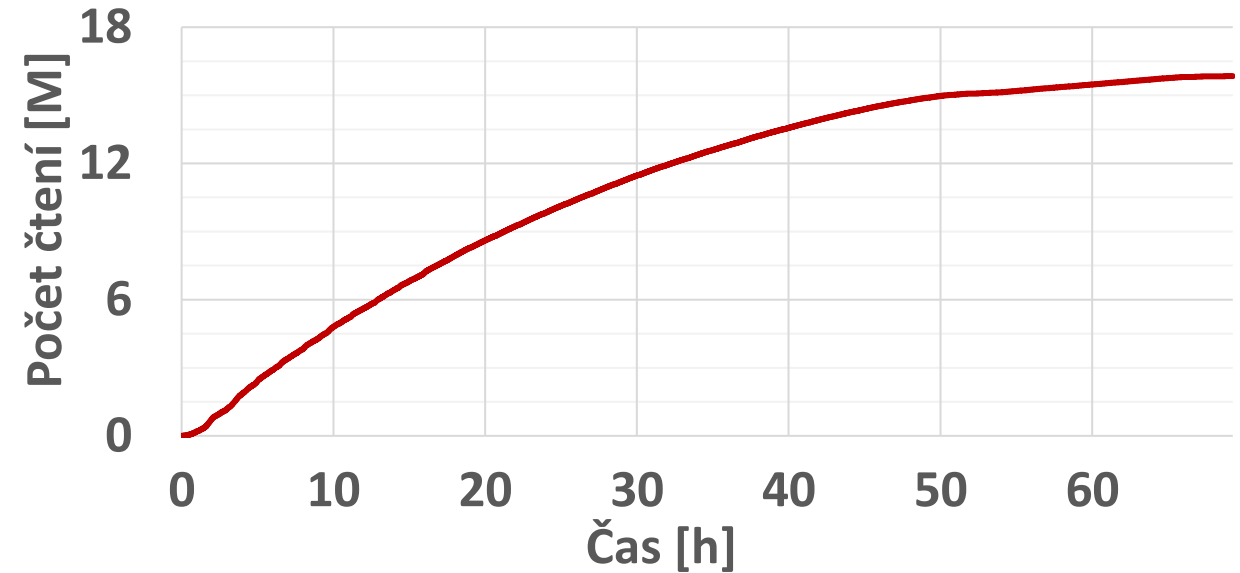


- Real-time přístup k datům
- Podpora režimu „Read-Until“
- 2 tisíce čtení již po jedné minutě

Diagnóza před koncem sekvenace!



Průběh sekvenace v čase



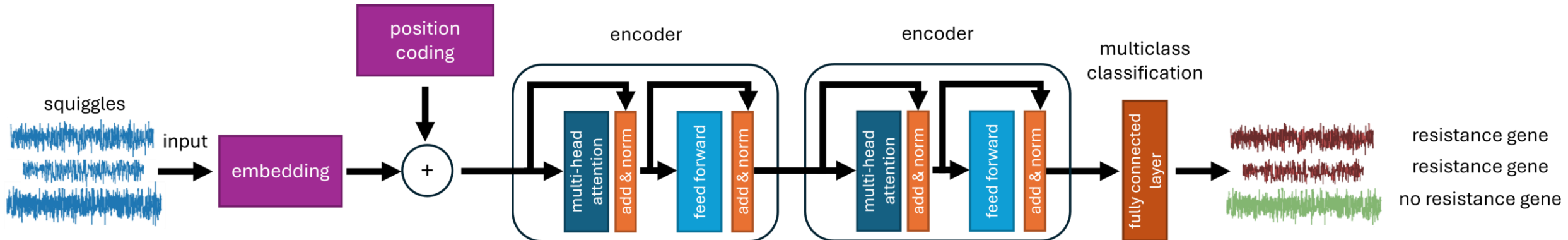
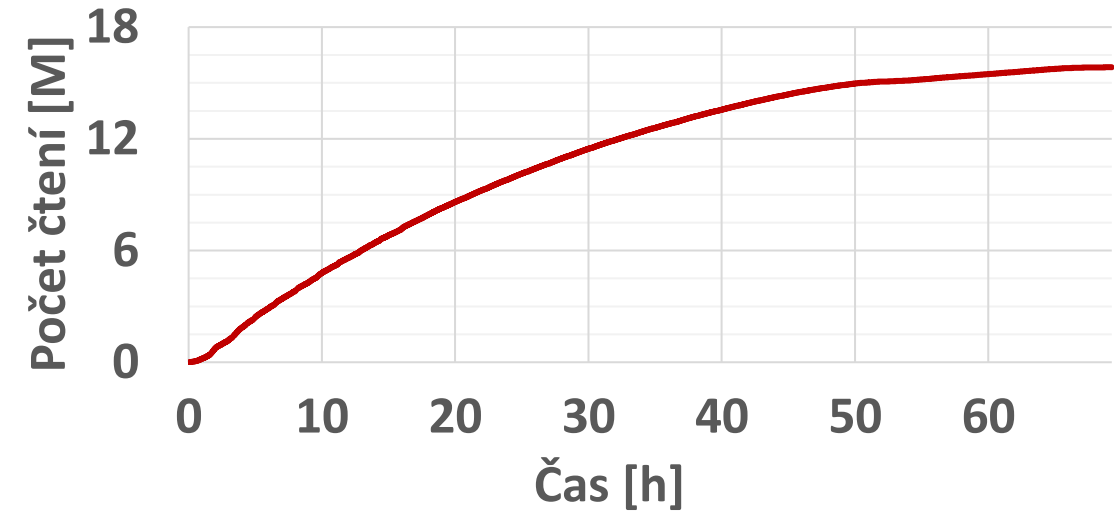
ZPRACOVÁNÍ SEKVENAČNÍCH DAT

NÁSTROJ PRO REAL-TIME DETEKCI GENŮ REZISTENCE A VIRULENCE

#BioSys_BUT

- real-time přístup k sekvenačním datům
- 2 tisíce čtení již po jedné minutě
- předběžné výsledky (v NT):
 - 1. minuta – první informace o rezistenci
 - 2. hodina – polovina detekovaných genů rezistencí
 - 16. hodina – detekovány všechny rezistence
- detekce přímo v surových signálech pomocí strojového učení:

Průběh sekvenace v čase





VYSOKÉ UČENÍ
TECHNICKÉ
V BRNĚ

FAKULTA ELEKTROTECHNIKY ústav
A KOMUNIKAČNÍCH biomedicínského
TECHNOLOGIÍ inženýrství

#BioSys_BUT

Skupina bioinformatiky a systémové biologie | BioSys_BUT



- doc. Ing. Mgr. Karel Sedlář, Ph.D.
- Mgr. Bc. Darina Čejková, Ph.D.
- Ing. Helena Vítková, Ph.D.
- Ing. Markéta Jakubíčková, Ph.D.
- Ing. Jana Musilová, Ph.D.
- Sudeep Roy, Ph.D.
- Ing. Kristýna Heřmánková
- Ing. Kateřina Šabatová
- Minoo Partovi Nasr
- Mohammad Umair
- M.Tech. Vaishali Pankaj
- Mgr. Indeerjet Bhogal

DĚKUJI ZA POZORNOST