

# Projekt inżynierski

## Aplikacja internetowa do analizy danych z eksperymentów real-time PCR

Autor: Jakub Porc

Kierujący pracą: Dr inż. Sebastian Student

# Wprowadzenie

- Eksperyment real-time PCR jest popularną techniką laboratoryjną ze względu na jego szerokie zastosowanie w wielu dziedzinach biologii molekularnej.
- Do analizy wyników tego doświadczenia konieczne jest zastosowanie metod komputerowych.
- Przetwarzanie danych z tego eksperymentu może być źródłem problemów dla osób nie mających doświadczenia w tej dziedzinie, dlatego istnieje konieczność stworzenia narzędzi ułatwiających i przyspieszających analizę.

# Cel projektu inżynierskiego

- Celem tego projektu jest stworzenie interaktywnej aplikacji internetowej, której zadaniem jest analiza danych z eksperymentu real-time PCR.
- Aplikacja powinna być prosta w obsłudze i pozwalać na szybką i wygodną prezentację wyników.
- Powinna wykorzystywać metodę Pfaffl i pozwalać na wykorzystanie wielu genów referencyjnych.

# Cel projektu inżynierskiego cd.

- Program powinien być łatwo dostępny.
- Interfejs użytkownika powinien być przejrzysty i czytelny.
- Zaplanowano również generowanie wykresów pudełkowych dla próbek.

# PCR i real-time PCR

- PCR to jedna z najbardziej popularnych technik w biologii molekularnej, dzięki której można otrzymać znaczną ilość określonego fragmentu DNA.
- Real-time PCR pozwala dodatkowo na pomiar ilości amplikonu w czasie dzięki pomiarowi fluorescencji pochodzącej od odpowiedniego barwnika lub sondy.

# Metody pomiaru poziomu ekspresji genów

## Metoda absolutna

- Pozwala na określenie dokładnej ilości materiału w określonych jednostkach.
- Wykorzystuje krzywą kalibracyjną sporządzoną dla znanych ilości materiału w próbkach.

## Metoda względna

- Pozwala na określenie stosunku ilości próbki względem próby referencyjnej.
- Wykorzystuje różne modele matematyczne do wyznaczenia różnicy.

# Metoda Pfaffl

- W przeciwieństwie do innych metod wykorzystuje się wartość wydajności reakcji.
- W pierwszym etapie dokonuje się obliczenia różnicy między kalibratorem i próbą.

$$\Delta Ct = Ct(cal) - Ct(test)$$

- Następnie określa się różnicę w ekspresji korzystając ze wzoru:

$$Różnica = \frac{E(gen)^{\Delta Ct(gen)}}{E(ref)^{\Delta Ct(ref)}}$$

# Opis i struktura aplikacji

- Aplikacja składa się z 3 skryptów i kilku obrazów.
- W ui zawarty jest wygląd interfejsu użytkownika.
- W serwerze wykonywane są generacje wyników.
- W skrypcie *funkcja1.R* zaimplementowano metodę Pfaffl.



# Wygląd aplikacji

## Simple qPCR

Step 1: Please select the file with efficiencies

Choose the file

Browse...

No file selected

Step 2: Please select any number of reference genes from the list

Step 3: Please choose the Real-time PCR data

Choose the files:

Browse...

No file selected

Step 4: Press this button to perform the analysis

## ANALYSE

Press this button in order to download the results

[Download the results](#)

### Tool description

## Result

## Boxplots - sample

This is a tool which allows to perform a simple and quick Real-time PCR data analysis. It uses the Pfaffl method for quantification and the efficiency values from the experiment.

The following files are necessary

- the file with gene list and efficiencies

	A	B
1	gene_name	efficiency
2	ACTB	1.94644
3	ATPGV1E1	1.88183
4	HADHA	1.95313
5	BIRC3	2.05523
6	COL2	1.92733
7	GPRC5B	1.85649
8	IL1A	1.96685
9	MEAP4	1.84396
10	POSTN	1.99025
11	TNFAIP3	2.05630

- the Real-time PCR data in SDS 2.3 format

SOS 2.5 AQ Results	1.0
Filename	p12_07_10_1014_XE_cDNA_NG
PlateID	
Assay Type	Absolute Quantification
Run DateTime	10/7/14 4:29:22 PM

### Sample Information

Index	Sample Name	Detector Name	Reporter	Task	QTY	Quantity	Baseline	Start	Baseline	Stop	Time	Unit
Worksheet	File	HW	SW	SW	MS	MS	MS	MS	MS	MS	MS	MS
121-121p1		Stanley	RM	Unknown	15.00070	6.11220070						
79	121-121p2	Stanley	RM	Unknown	15.00170	6.11220070						
80	121-121p3	Stanley	RM	Unknown	15.00170	6.11220070						
127	124-124p2	Stanley	RM	Unknown	15.17091	6.11220070						
151	123-123p1	Stanley	RM	Unknown	15.00763	6.11220070						
175	127-123p2	Stanley	RM	Unknown	15.28000	6.11220070						
180	14p1	Stanley	RM	Unknown	15.10454	6.11220070						

The file names of the data must have the following format: gene\_other

Where

gene - gene name

other - any additional information

The output table will be visible in the Results tab, while the boxplots will be available in the boxplots tab

If you wish to save the results, press the 'Download the results' button.

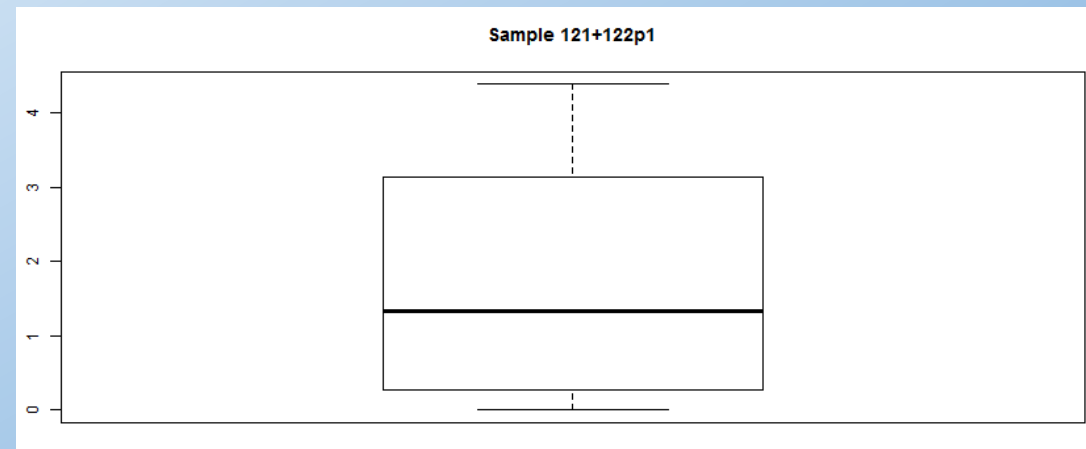
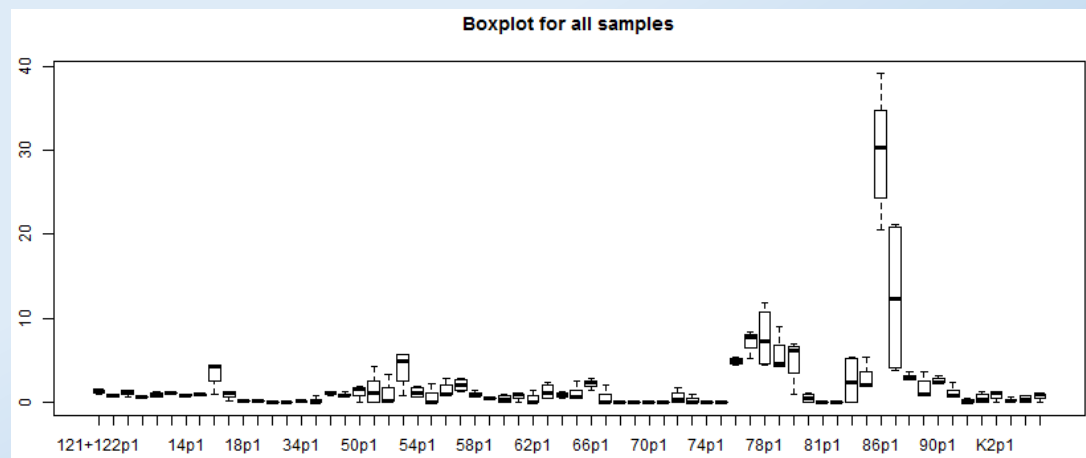
Note: The program assumes that each plate has the same samples.

# Wyniki

Sample Name	ACTB	BIRC3	CCL2	GPRC5B	IL1A
121+122p1	1.13847554938083	4.09670805212449	1.18658422822053	2.54323732167115	0.499772081750035
121+122p2	1.17492074198917	3.98931574054315	1.21664734804807	2.6109333684399	0.590748490415339
124+125p1	1.15624821596073	4.08076147659764	1.05310806074133	2.68936421291295	0.526463895631909
124+125p2	1.16899323750131	4.15685565304566	1.02679096044858	2.64778986812624	0.522335499928422
127+128p1	1.13380471169414	3.77652259983747	1.13019464399269	2.24209951431513	0.58064349514203
127+128p2	1.12167578948294	3.7032727385796	1.07165961457179	2.2455573286825	0.562865457314627
14p1	0.886077555670707	1.14388113290061	0.746996738378139	1.51665997559285	2.31788570510448
14p2	0.862868715957436	1.07418552375432	0.788032480846823	1.50331124226183	2.5052788997112
16p1	0.951143023320845	1.2242287450095	0.770420125119262	1.34913718715238	2.31901693902567
16p2	0.926639576219997	1.21601389414855	0.800804933908945	1.39555196823034	2.48688980783748
18p1	0.937292173201797	1.29847163892502	0.749144710430388	1.38246445296051	2.39455016927
18p2	0.915937582199359	1.22992234260313	0.762602679217617	1.45301827375883	2.83393756695742
32p1	0.942202500292248	0.13454058491008	2.08120740608717	0.896084904398088	2.69103735493894
32p2	0.961066118128227	0.142567262331688	2.00691759096615	0.876222850790427	2.3728610920583
34p1	1.1594165783702	0.16376502100224	1.98502646781579	0.92967343939727	1.8253123198375

	A	B	C	D	E	F	G	H	I	J
1	Sample.N	ATP6V1E1	BIRC3	CCL2	GPRC5B	IL1A	MFAP4	POSTN	TNFAIP3	
2	121+122p1	1.3961328	4.3849976	1.2700853	2.7222075	0.5349415	0.0030382	0.0009840	3.53297607380287	
3	121+122p2	1.4383226	4.2675733	1.3015093	2.7930478	0.6319536	0.0034087	0.0006793	3.67620023272936	
4	124+125p1	1.4336964	4.3840609	1.1313795	2.8892491	0.5655929	0.0074589	0.0011641	3.2412022103936	
5	124+125p2	1.4224908	4.4378992	1.0962119	2.8268060	0.5576504	0.0070125	0.0013299	2.90374070739887	
6	127+128p1	1.4953068	4.1414851	1.2394164	2.4587756	0.6367567	0.0035687	0.0020593	2.76505415638895	
7	127+128p2	1.4584193	4.0419480	1.1696660	2.4509202	0.6143411	0.0045538	0.0022192	2.48977604879883	
8	14p1	0.9547064	1.1726820	0.7658047	1.5548467	2.3762459	No data	No data	2.43392080312937	
9	14p2	0.9564937	1.1117106	0.8155612	1.5558272	2.5927971	No data	4.8531563	2.37161204819159	
10	16p1	0.9008285	1.2022497	0.7565884	1.3249156	2.2773827	No data	6.6368120	2.27310188463342	
11	16p2	0.9653676	1.2327239	0.8118093	1.4147292	2.5210638	No data	7.1865804	2.25609437406945	
12	18p1	0.9724208	1.3144948	0.7583891	1.3995241	2.4240990	No data	No data	2.05936987744992	
13	18p2	0.9662550	1.2520441	0.7763191	1.4791527	2.8849098	No data	5.5482597	2.33543698612568	
14	32p1	0.7651876	0.1255243	1.9417348	0.8360335	2.5106969	0.0004711	No data	0.832178908060378	
15	32p2	0.7838721	0.1332039	1.8751104	0.8186756	2.2170200	No data	No data	0.776757584506958	
16	34p1	0.6639746	0.1359956	1.6484292	0.7720304	1.5157974	0.0004166	No data	0.709231311104286	
17	34p2	0.6756958	0.1235301	1.6420250	0.7237864	1.7481487	0.0003999	No data	0.674678311639217	
18	36p1	0.6828213	0.1297885	1.7174490	0.7383558	1.9335169	No data	No data	0.750699830218832	
19	36p2	0.7507111	0.1492206	1.6989795	0.8050282	2.0812329	0.0006530	No data	0.881335980976727	
20	50p1	1.3093554	0.1730477	0.0002449	0.4618129	0.0157290	0.0014531	0.0025581	0.0446902853582785	
21	50p2	1.3127560	0.1674377	0.0001327	0.5106494	0.0183910	0.0007995	0.0011069	0.0451215737184783	

# Wyniki cd.



# Podsumowanie

- Wykonana aplikacja spełnia założenia i cele projektu.
- Stworzono przejrzysty i czytelny interfejs użytkownika.
- Analiza danych jest prosta i szybka.
- Wykorzystana metoda Pfaffl pozwala na wykorzystanie wielu genów referencyjnych.
- Aplikacja została umieszczona na serwerze.
- Program generuje wykresy pudełkowe ułatwiające analizę.