

# Metagenomika – příprava knihovny 16S rRNA

Petra Vídeňská, Ph.D.

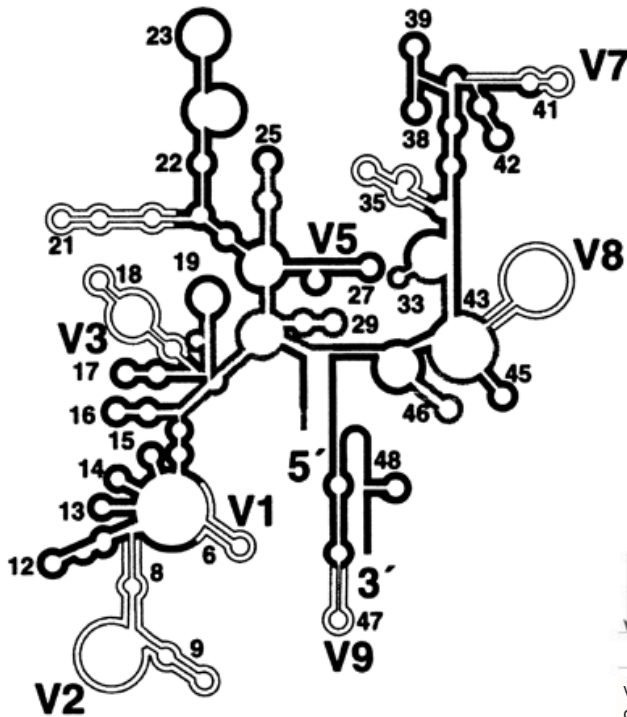
# Faktory ovlivňující výsledek sekvenace genu pro 16S rRNA

- Sběr vzorků
  - Čas vystavení vzduchu
  - Teplota
  - Postup
- Izolace DNA
  - Použitý kit a metoda disrupce bakteriálních buněk
  - Čas po který je vzorek rozmrazován před izolací
- Příprava knihovny
  - Výběr primerů
  - Polymeráza
  - Počet cyklů PCR (vznik chimér)
- Sekvenace
  - Výběr sekvenační platformy
  - Kvalita knihovny
- Analýza
  - Použitý quality trimming
  - Použité algoritmy pro výběr OTU
  - Algoritmus výběru reprezentativních sekvencí
  - Výběr databáze 16S rDNA

# Gen pro 16S rRNA

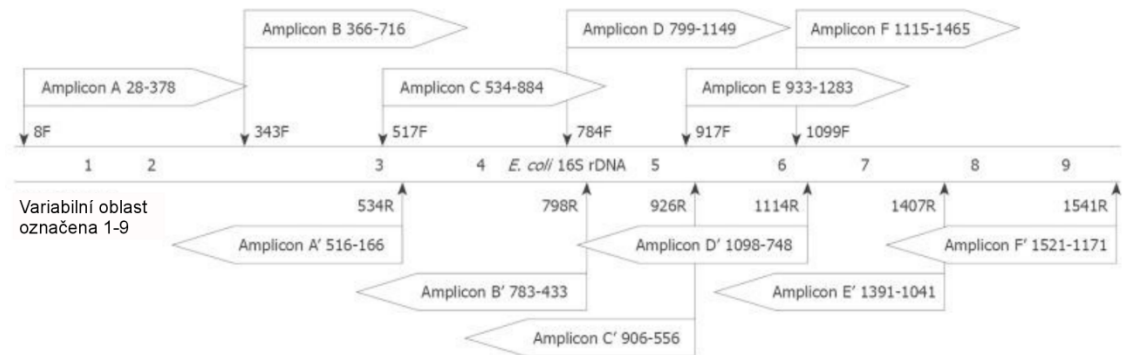
- Gen pro 16S rRNA využil roku 1977 Carl Woese ke zjištění fylogenetických vztahů mezi bakteriemi.
- Navrhl oddělit archea od prokaryot a navrhl rozdělení organismů do tří domén: *Archea*, *Bacteria* a *Eukaryota*
- Taxonomická klasifikace na základě 16S rRNA se využívá k vytvoření fylogenetických vztahů, odhadu diverzity, vyrovnanosti mikrobiálního společenství a struktury komunity
- 16S rRNA je nezbytnou součástí ribozomální podjednotky 30S, která je u prokaryot odpovědná za translaci RNA do proteinů.
- 3' konec 16S ribozomální RNA obsahuje anti-Shine-Dalgarnovu sekvenci, která váže do vazebného místa ribozomu 5' konec mRNA
- 16S rRNA napomáhá spojení obou ribozomálních podjednotek (30S a 50S) a utváří tak kompletní bakteriální ribozom.
- Obsahuje jednak konzervativní oblasti, na které lze navrhovat primery a jednak hypervariabilní oblasti umožňující rozlišení bakterií a určení diverzity bakterií ve vzorku

# 16S rRNA



## Sekundární struktura 16S rRNA

Nevybarvené oblasti odpovídají hypervariabilním oblastem V1 – V9.



**Vytvoření primerů *in silico*** Ukázka setu 6 primerů (A – F) pokrývajících vždy dvě variabilní oblasti a jejich teoretické forward i reverse čtení. Schéma ukazuje relativní pozici jednotlivých ampliconových čtení i jejich směr, stejně jako které báze budou začleněny do sekvence. Jako ukázkový templát posloužil gen pro 16S rRNA *E. coli* s přibližným vyznačením variabilních oblastí.

# Fylogeneze

<https://www.patricbrc.org/portal/portal/patric/Phylogeny?cType=taxon&cId=2>



- ORGANISMS
- DATA
- SERVICES
- TOOLS
- ABOUT

Login Not Registered? Sign Up  
Learn About Registering

Search

cellular organisms • Bacteria

Data Overview Tutorial »  
Download genome data »

- Overview
- Taxonomy
- Phylogeny
- Genome List
- Feature Table
- Specialty Genes
- Pathways
- Transcriptomics
- Interactions
- Diseases
- Literature

Chose an order to view its multi-gene phylogenetic tree in either phylogram or cladogram format. Use controls above the tree to show or hide support values, and to color the tree according to genus or species. Click on a species name to visit that species overview page. To learn more about interacting with trees and the method used to create them, please see [Phylogeny User Guide](#).

Show phylogenetic tree for the order **Legionellales**

Draw tree as

- phylogram
- cladogram

Color tree by

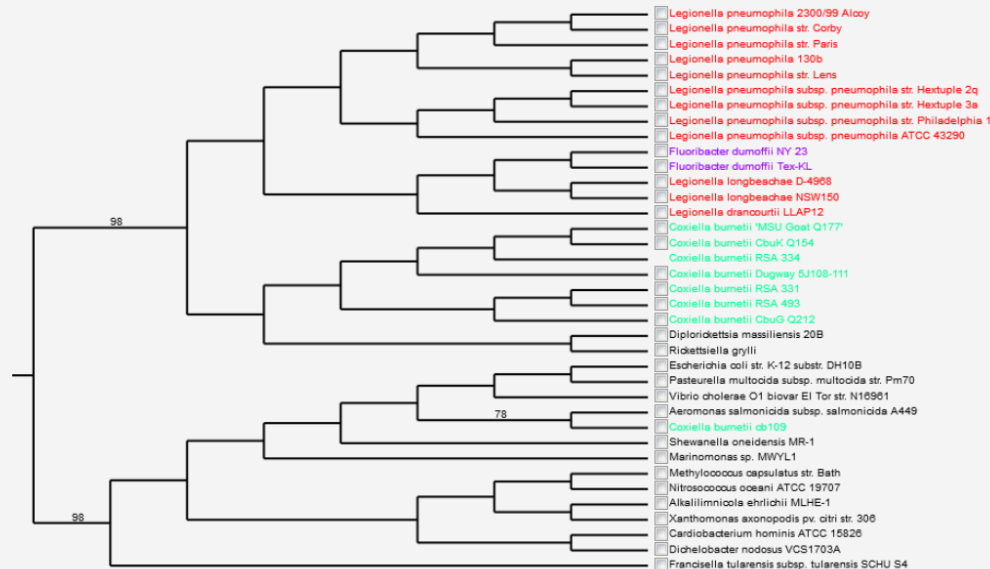
- species
- genus
- off

Turn support values

- on
- off

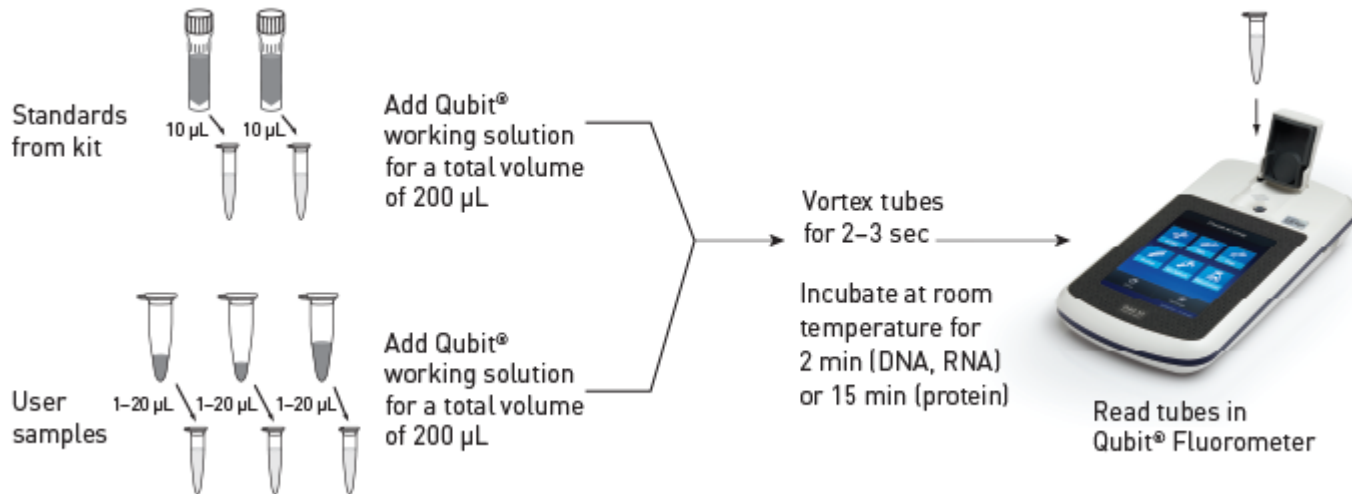
[Download tree as Newick file](#)

[Add Genome\(s\) to Workspace](#)



# Příprava knihovny- kontrola vyizolované DNA

- Klasický agarózový gel (TapeStation)
- Měření koncentrace
  - Spektrofotometricky (Nanodrop)
  - Fluorometricky (Qubit, PicoGreen)



# Výběr specifických primerů



Online Submissions: <http://www.wjgnet.com/1007-9327/office>  
wjg@wjgnet.com  
doi:10.3748/wjg.v16.i33.4135

World J Gastroenterol 2010 September 7; 16(33): 4135-4144  
ISSN 1007-9327 (print)  
© 2010 Baishideng. All rights reserved.

ORIGINAL ARTICLE

## Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiome

Carlos W Nossa, William E Oberdorf, Liying Yang, Jørn A Aas, Bruce J Paster, Todd Z DeSantis, Eoin L Brodie, Daniel Malamud, Michael A Poles, Zhiheng Pei

Carlos W Nossa, William E Oberdorf, Michael A Poles, Department of Medicine, New York University School of Medicine, New York, NY 10016, United States  
Liying Yang, Department of Pathology, New York University School of Medicine, New York, NY 10016, United States  
Jørn A Aas, Bruce J Paster, Department of Molecular Genetics, The Forsyth Institute, Boston, MA 02115, United States  
Jørn A Aas, Faculty of Dentistry, University of Oslo, PO Box 1052 Blindern, 0316 Oslo, Norway  
Bruce J Paster, Harvard School of Dental Medicine, Boston, MA 02115, United States  
Todd Z DeSantis, Eoin L Brodie, Lawrence Berkeley National Laboratory, Center for Environmental Biotechnology, Berkeley, CA 94720, United States  
Daniel Malamud, New York University College of Dentistry, New York, NY 10016, United States  
Zhiheng Pei, Department of Pathology and Medicine, New York University School of Medicine, New York, NY 10016, United States; Department of Veterans Affairs New York Harbor Health System, New York, NY 10010, United States  
Author contributions: Pei Z and Nossa CW designed this study; Nossa CW, Oberdorf WE, Yang L and Aas JA performed computational analyses; Nossa CW, Pei Z, Paster BJ, DeSantis TZ, Brodie EL, Malamud D and Poles MA participated in manuscript preparation.

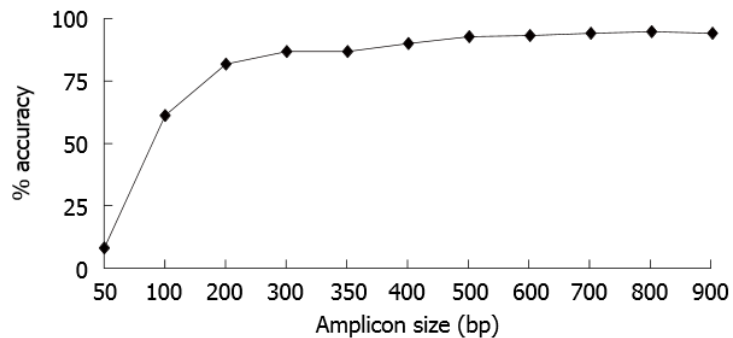
Supported by (in part) Grants UH2CA140233 from the Human Microbiome Project of the NIH Roadmap Initiative and National Cancer Institute; R01AI063477 from the National Institute of

primers for use in high throughput sequencing to classify bacteria isolated from the human foregut microbiome.

**METHODS:** A foregut microbiome dataset was constructed using 16S rRNA gene sequences obtained from oral, esophageal, and gastric microbiomes produced by Sanger sequencing in previous studies represented by 219 bacterial species. Candidate primers evaluated were from the European rRNA database. To assess the effect of sequence length on accuracy of classification, 16S rRNA genes of various lengths were created by trimming the full length sequences. Sequences spanning various hypervariable regions were selected to simulate the amplicons that would be obtained using possible primer pairs. The sequences were compared with full length 16S rRNA genes for accuracy in taxonomic classification using online software at the Ribosomal Database Project (RDP). The universality of the primer set was evaluated using the RDP 16S rRNA database which is comprised of 433306 16S rRNA genes, represented by 36 phyla.

**RESULTS:** Truncation to 100 nucleotides (nt) downstream from the position corresponding to base 28 in the *Escherichia coli* 16S rRNA gene caused misclassification of 87 (20.7%) of the 219 sequences compared

# Výběr specifických primerů



**Figure 1 Classification accuracy is dependent on amplicon size.** Full length sequences were trimmed to 900, 800, 700, 600, 500, 400, 350, 300, 200, 100, and 50 bases with each amplicon starting at *Escherichia coli* base 28. Each sequence trim was uploaded onto Ribosomal Database Project II classifier and the results at each taxonomical level were compared to results obtained using the full length sequence. Percent classification accuracies at the genus level for each amplicon size trim are shown.

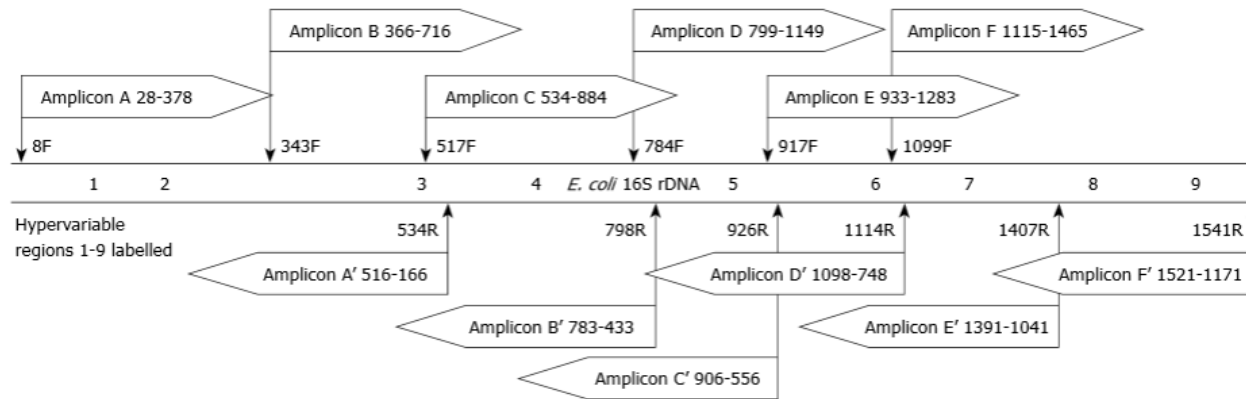
**Table 2 Amplicons designed for analysis**

Amplicon	Primers		Total length	HVR(s) included
	F	R		
A	8F	534R	527	1,2,3
B	343F	798R	456	3,4
C	517F	926R	410	3,4,5
D	784F	1114R	331	5,6
E	917F	1407R	491	6,7
F	1099F	1541R	443	7,8,9

HVR: Hypervariable region; F: Forward primers; R: Reverse primers.



# Výběr specifických primerů



**Figure 2 Design of amplicons for *in silico* evaluation.** Amplicons designed using the 6 universal primer sets as described in Table 3 were evaluated for theoretical forward and reverse reads. The schematic shows the relative position of each amplicon read and read direction, as well as which bases would be included in the sequence. The positional template is the *Escherichia coli* (*E. coli*) 16S rDNA gene with the approximate locations of the hypervariable regions labeled.

**Table 3 Accuracy of taxonomic classification of 219 foregut species using 350-bp sequences**

Amplicon	% accuracy compared to 900-bp amplicon/full length				
	Phylum	Class	Order	Family	Genus
<b>Forward reads</b>					
A	97.7/97.7	96.3/95.9	95.9/95.4	93.2/94.5	87.7/86.8
B	99.1/99.1	98.2/96.8	97.7/96.3	97.3/95.9	91.8/89.0
C	99.5/99.5	98.2/96.8	97.7/96.8	97.3/95.9	90.4/88.1
D	98.6/99.1	98.6/98.2	97.3/97.3	95.9/96.3	88.1/86.8
E	98.2/98.6	97.7/98.2	95.4/95.9	92.2/93.6	85.4/84.9
F	97.7/98.2	96.3/97.3	95.0/95.4	91.8/93.6	83.6/84.5
<b>Reverse reads</b>					
A'	98.6/98.6	97.3/97.7	97.3/97.7	95.0/96.3	90.0/90.9
B'	99.5/99.5	98.2/96.8	97.7/96.8	97.7/96.3	93.6/91.8
C'	99.5/99.5	98.2/96.8	97.7/96.8	97.3/95.9	91.3/90.0
D'	99.5/100	98.6/99.5	96.8/98.2	94.5/96.3	87.2/89.5
E'	98.2/98.6	96.8/98.2	94.5/96.3	91.3/94.1	83.6/87.2
F'	98.2/98.6	96.8/98.2	95.0/96.4	92.2/95.0	82.6/85.8

# Výběr specifických primerů

**Table 6** Taxonomic coverage of domain bacteria by primers 347F and 803R

Primer	Optimization	Total species	Coverage at mismatches <i>n</i> (%)			Total sequence	Coverage at mismatches <i>n</i> (%)		
			0	1	2		0	1	2
374F	Before	5165	3392 (65.7)	4835 (93.6)	5042 (97.6)	433 306	275 801 (63.7)	406 626 (93.8)	418 613 (96.6)
	After	5165	4703 (91.1)	4996 (96.7)	5114 (99.0)		433 306	391 695 (90.4)	418 832 (96.7)
803R	Before	5165	4584 (88.8)	5091 (98.6)	5159 (99.9)	433 306	352 827 (81.4)	417 612 (96.4)	430 967 (99.5)
	After	5165	4741 (91.8)	5131 (99.3)	5159 (99.9)		433 306	367 771 (84.9)	427 791 (98.7)

# Kontrola vybraných primerů

- <https://rdp.cme.msu.edu/probematch/search.jsp>
- F 5'-GGAGGCAGCAGTRRGAAT (347F)
- R 5'-CTACCRGGGTATCTAATCC (803R)

## Probe Match - Start

[\[ video tutorial | help \]](#)

---

### Search

Running Jobs: 0  
Pending Jobs: 0

Please enter your probe:

5'  3'

Probe targets   strand (- for forward primer, + for reverse primer).

Optional second probe

5'  3'

Probe targets   strand.

Restrict search to domain:   Bacteria

Restrict search to region relative to positions in E. coli (Bacteria or Archaea) or S. cerevisiae (Fungi):  to

---

### Quick Help (more help is available)


RDP's ProbeMatch finds sequences matching your probe in the RDP's database. A list of probes is available at sites such as Probase. Use programs such as Primrose or ARB to assist you with phylogenetic probe design.

**Probe:** Please enter your probe(s) in the 5' to 3' direction. Only IUPAC codons can be used, including ambiguity codons. A probe must be less than 64 bases in length.

**Restrict to Domain:** Search only Bacteria, Archaea or Fungi sequences, if desired.


**Restrict to Region:** Only search sequences with sequence data in the specified E. coli region (GenBank J01695) for Bacteria or Archaea, or S. cerevisiae (GenBank NC\_001144.5:455571..451786) for Fungi. There are many partial 16S sequences in the RDP database that may be missing data in the region your probe targets. This function can be useful if you're interested in what sequences your probe does not match. The coverage of the number of RDP sequences spanning each base position for each domain can be found on [resources](#).

Questions/comments: [rdpstaff@msu.edu](mailto:rdpstaff@msu.edu)



[top](#)

---

 **CENTER FOR MIKROBIAL ECOLOGY**  
@ Michigan State University

**MICHIGAN STATE UNIVERSITY** Advancing Knowledge.  
Transforming Lives.  
© 1992-2014 Michigan State University Board of Trustees, East Lansing MI 48824  
MSU is an affirmative-action, equal-opportunity employer

# Kontrola vybraných primerů

## Probe Match Result

Display depth:

Differences Allowed:

Probe: 5'ATTCCYYACTGCTGCCTCC 3'  
Target: 5'GGAGGCAGCAGTRRGGAAAT 3'

Probe: 5'CTACCRGGGTATCTAAICC 3'  
Target: 5'GGATTAGATACCCYGGTAG 3'

[\[start over\]](#) [\[help\]](#) [\[options\]](#)

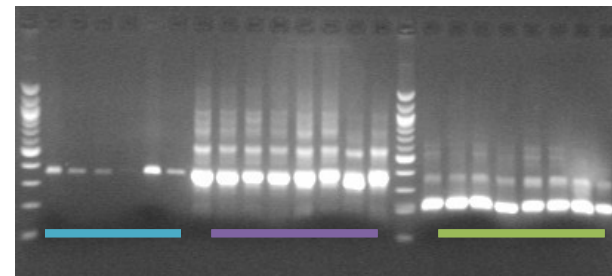
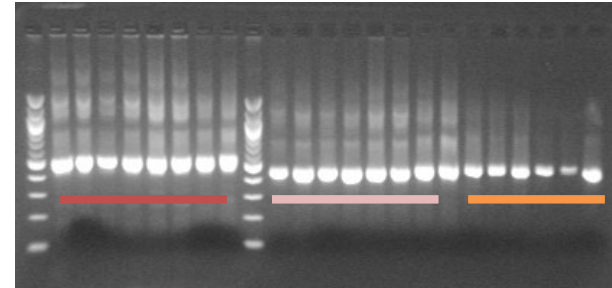
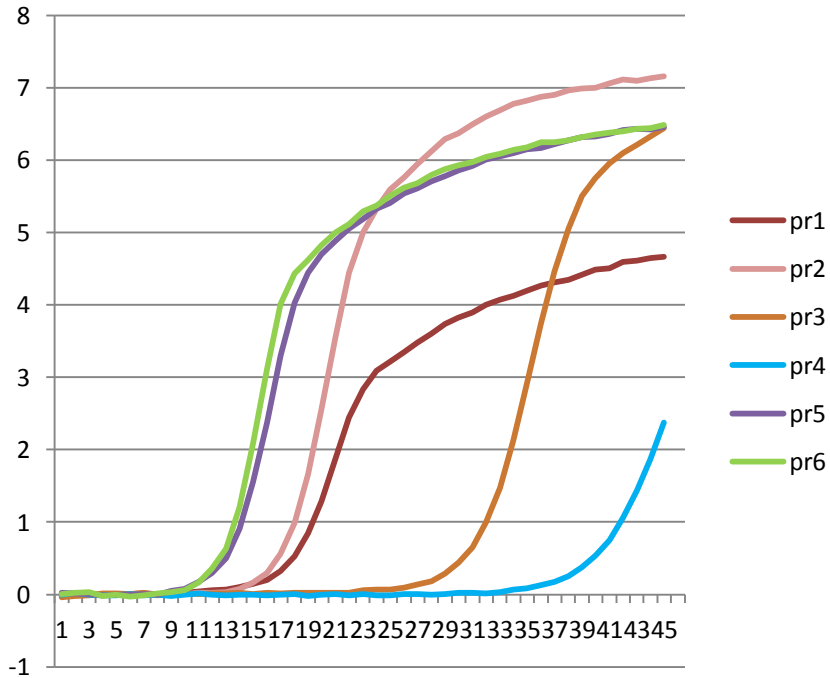
### Hierarchy View:

rootrank Root (1646694/3333501) (hits/total searched) [\[list results for this node\]](#)  
domain **Bacteria** (1646694/3070243)  
  phylum "Actinobacteria" (193042/349461)  
  phylum "Aquificae" (0/2085)  
  phylum "Bacteroidetes" (237709/388464)  
  phylum "Caldiserica" (127/361)  
  phylum "Chlamydiae" (11/1538)  
  phylum "Chlorobi" (736/1556)  
  phylum "Chloroflexi" (3122/42562)  
  phylum "Chrysiogenetes" (11/13)  
  phylum "Deferribacteres" (994/1462)  
  phylum "Deinococcus-Thermus" (0/5035)  
  phylum "Dictyoglomi" (0/71)  
  phylum "Elusimicrobia" (588/782)  
  phylum "Fibrobacteres" (1155/1863)  
  phylum "Fusobacteria" (12789/17964)  
  phylum "Gemmatimonadetes" (3057/5338)  
  phylum "Lentisphaerae" (9/2628)  
  phylum Nitrospirae (3167/8171)  
  phylum "Planctomycetes" (58/31651)  
  phylum "Proteobacteria" (600360/1031919)  
  phylum "Spirochaetes" (6/30375)  
  phylum "Synergistetes" (1612/2996)  
  phylum "Tenericutes" (3909/10618)  
  phylum "Thermodesulfobacteria" (0/249)  
  phylum "Thermotogae" (1/1199)  
  phylum BRC1 (0/628)  
  phylum Parcubacteria (0/942)  
  phylum Microgenomates (0/322)  
  phylum SR1 (857/1084)  
  phylum Candidatus Saccharibacteria (1392/7052)  
  phylum Latescibacteria (1/1162)  
  phylum "Armatimonadetes" (14/3243)  
  phylum "Verrucomicrobia" (44/27439)  
  phylum "Acidobacteria" (29054/74631)  
  phylum Firmicutes (524770/811625)  
  phylum Cyanobacteria/Chloroplast (670/70404)  
  phylum Marinimicrobia (963/1637)  
  phylum Aminicenantes (1722/2296)  
  phylum Omnitrophica (0/38)  
  phylum Acetothermia (0/89)  
  phylum Poribacteria (0/354)  
  phylum Atribacteria (0/168)  
  phylum Cloacimonetes (3/11224)  
  phylum Candidatus Calescamentes (0/5)  
  phylum candidate division WPS-1 (0/1224)  
  phylum candidate division WPS-2 (0/376)  
  phylum Hydrogenedentes (11/985)  
  phylum candidate division ZB3 (84/95)  
  phylum Ignavibacteriae (1107/2130)  
  phylum Nitrospinae (370/821)  
  Archaea Outgroup (0/1)  
  ▶ unclassified\_Bacteria (23179/111907)  
domain Archaea (0/154357)  
domain Fungi (0/108901)

# Výběr primerů z publikací

Název	enght	Tm	Sekvence	Poznámka	ozn
16S_V1/V3F_ERB	492	58	GAG TTT GAT CNT GGC TCA G	Erb-Downward at al., 2011 16S na pyrosequencing	pr1
16S_V1/V3R_ERB		58	GWN TTA CNG CGG CKG CTG	Erb-Downward at al., 2011 16S na pyrosequencing	
16S_V3/V4F_NOSS	456	60	GGA GGC AGC AGT RRG GAA T	Nossa et al., 2010 16S na pyrosequencing	pr2
16S_V3/V4R_NOSS		56	CTA CCR GGG TAT CTA ATC C	Nossa et al., 2010 16S na pyrosequencing	
16S_V3/V4F_CLAE	464	66	ACT CCT ACG GRA GGC AGC AG	Claesson et al., 2010 16S na pyrosequencing	pr3
16S_V3/V4R_CLAE		52	TAC NVG GGT ATC TAA TCC	Claesson et al., 2010 16S na pyrosequencing	
16S_V4/V5F_CLAE	363	53,3	AYT GGG YDT AAA GNG	Claesson et al., 2010 16S na pyrosequencing	pr4
16S_V4/V5R_CLAE		56	CCG TCA ATT YYT TTR AGT TT	Claesson et al., 2010 16S na pyrosequencing	
16S_V2/V3F_MCKE	345	50	CTG CTG CCT YCC GTA	McKenna at al., 2008 16S na pyrosequencing	pr5
16S_V2/V3R_MCKE		58	AGA GTT TGA TCC TGG CTC AG	McKenna at al., 2008 16S na pyrosequencing	
16S_univ-2F	-	-	GAG GAA GGI GIG GAI GAC GT	Tseng et al., 2003 pro real-time PCR	pr6
16S_univ-2R	-	-	AGI CCC GIG AAC GTA TTC AC	Tseng et al., 2003 pro real-time PCR	

# Real-time PCR s vybranými primery

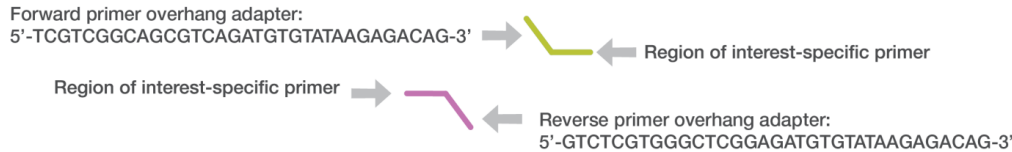


# Strategie přípravy amplikonů

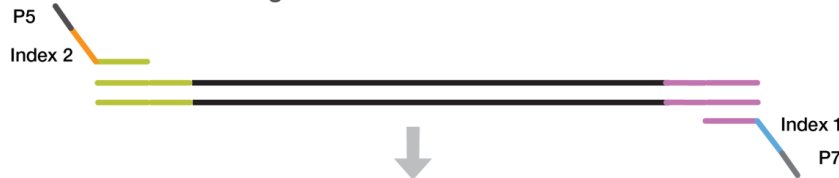
## 1. Dvoustupňová PCR

- primery lze nasyntetizovat s extenzí, na kterou se pak pomocí PCR přidají komerční indexy (např. Nextera XT)

PCR amplify template out of genomic DNA using region of interest-specific primers with overhang adapters



Attach indices and Illumina sequencing adapters using the Nextera<sup>®</sup> XT Index Kit



Normalize and pool libraries



Sequence

- Výhody
  - Přesný protokol, není potřeba žádná optimalizace
  - Vysoká efektivita připojování adaptorů
- Nevýhody
  - Vysoká cena
  - 96 indexů (lze i více, ale prodražuje se to)
  - Hodně vzorků se pooluje najednou, velká nevyrovnanost (s 200 vzorky nikdo nebude dělat qPCR, pooluje se pouze na základě výsledků fluorometrického stanovení koncentrace)
  - Dvoustupňová PCR

# Strategie přípravy ampliconů

**2. Dvoustupňová PCR s vnitřními tagy** – stejný postup jako předtím, ale před specifickým primerem je vložena značka

**16S Amplicon PCR Forward Primer = 5'**

**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG- N az NNN- značka (min. 5 nukleotidu)-  
spacer (2 nukleotidy)-CCTACGGGNGGCWGCAG**

**16S Amplicon PCR Reverse Primer = 5'**

**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG- N az NNN- značka (min. 5  
nukleotidu)-spacer (2 nukleotidy)-GACTACHVGGGTATCTAATCC**

**Extenze, oblast potřebnou pro pozdější přidání indexů**

**N – nukleotidy které přidáváme aby byla vyšší diverzita během sekvenování (např. pět primerů bude mít 1 N, pět primerů 2N, pět primerů 3N...)**

**značka – sekvence odlišující mezi sebou jednotlivé vzorky**

**spacer – odděluje značku od oblasti komplementární k oblasti zájmu**

**vlastní primer, komplementární k oblasti zájmu**



# Strategie přípravy amplikonů

## 2. Dvoustupňová PCR s vnitřními tagy

- Výhody
  - Po prvním poolování před indexací se pracuje již s méně vzorky – snadnost
  - Cena – ušetří se za indexy
  - Lépe vyrovnaná knihovna
- Nevýhody
  - dvoustupňová PCR

# Strategie přípravy ampliconů

## 3. Sekvence přímo produktů amplifikace

- Amplicony 100 – 500 bp – primery s extenzí přímo na sekvenování, musí obsahovat extenzi a index (jen jeden z nich)
- Caporaso a kol. (<http://www.earthmicrobiome.org/emp-standard-protocols/16s/>), pozor jiný sekvenační a indexační primer
  - Forward: AATGATACGGCGACCACCGAGATCTACAC TATGGTAATT GTGTGCCAGCMGCCGCGGTAA
  - Reverse: CAAGCAGAAGACGGCATACGAGAT Index AGTCAGTCAG CCGGACTACHVGGGTWTCTAAT
- Pro sekvenační kit s TruSeq indexy, viz. Nelson a kol., 2014 (Analysis, Optimization and Verification of Illumina - Generated 16S rRNA Gene Amplicon Surveys)
  - Forward:CAAGCAGAAGACGGCATACGAGAT- TruSeq indexy-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
  - Reverse:AATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGACGCTCTTCCGATCT-NNNN
- Podobně i Kozich a kol. (<http://aem.asm.org/content/suppl/2013/08/06/AEM.01043-13.DCSupplemental/zam999104626so1.pdf>)

# Strategie přípravy amplikonů

## 3. Sekvenace přímo produktů amplifikace

- Výhody
  - Rychlá příprava
  - Pouze jedna PCR reakce
  - Cena
- Nevýhody
  - Hodně vzorků se pooluje najednou, velká nevyrovnanost (s 200 vzorky nikdo nebude dělat qPCR, pooluje se pouze na základě výsledků fluorometrického stanovení koncentrace)

# Strategie přípravy amplikonů

## 4. Za pomocí TruSeq PCR free

([http://support.illumina.com/sequencing/sequencing\\_kits/truseq\\_dna\\_pcr\\_free\\_lt\\_sample\\_prep\\_kit.ilmn](http://support.illumina.com/sequencing/sequencing_kits/truseq_dna_pcr_free_lt_sample_prep_kit.ilmn))

- Protokol stejný jako pro chromozomální DNA, jen se nepoužije Covaris
- Výhody
  - oproti postupu 1. nemusí mít primer extenzi, lze ligovat na různě PCR amplikony
  - Není druhá PCR – méně chyb zanesených polymerázou
- Nevýhody
  - Delší a složitější příprava
  - Knihovna se hůře kontroluje (nízká koncentrace, 1. řetězcová DNA)

# Strategie přípravy amplikonů

## 5. Vytvoření vlastní dvoustupňové PCR –

- první sada primerů obsahuje vlastní extenzi a oblast zájmu, druhý pár primerů obsahuje kompatibilní oblast k první extenzi, index a extenzi nutnou k sekvenaci
- Výhody
  - Variabilita použití
  - Cena
- Nevýhody
  - Optimalizace

# Postup při dvoustupňové PCR s vnitřními tagy

- PCR reakce – s 2x KAPA HiFi HotStart ReadyMix, 3x 10  $\mu$ l, následné smíchání
- Kontrola produktu na gelu/Agilentu
- Přečištění AMPure Beads
- Změření koncentrace vzorků pomocí **fluorometru** (Qubit, PicoGreen)
- Příp. kontrola některých vzorků na Agilentu (kontrola odmytí primerů, že je pouze jeden produkt)

# Postup při dvoustupňové PCR s vnitřními tagy

- Poolování vzorků před indexací

Vzorek	Primery s vnitřní značkou – 1.PCR	Indexy - 2.PCR
1-20	1-20	1
21-40	1-20	2
41-60	1-20	3
...	...	...

- Přidání indexů - 2. PCR – nutná kontrola indexů (vytvoření sample sheetu) !!!

# Postup při dvoustupňové PCR s vnitřními tagy

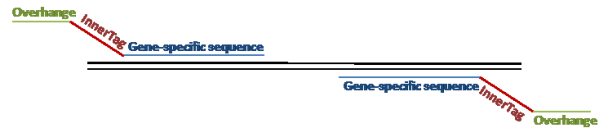
- Přečištění AMPure Beads
- Změření koncentrace vzorků pomocí fluorometru a/nebo qPCR
- Poolování vzorků před sekvenací
- Kontrola knihovny (Agilent, qPCR)
- Sekvenace na Illumina
  - Lépe mírně podklastrovat (hodně nízká diverzita knihovny)
  - Přidat PhiX (5-50%) pro zvýšení diverzity



# Postup při dvoustupňové PCR s vnitřními tagy

## Library preparation

Amplicon PCR (triplicates, after PCR pooled)



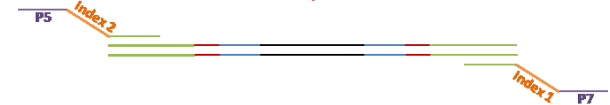
Index Group1	Index Group2	Index Group3	Index Group9
InnerTag1	InnerTag1	InnerTag1	InnerTag1
InnerTag2	InnerTag2	InnerTag2	InnerTag2
InnerTag3	InnerTag3	InnerTag3	InnerTag3
InnerTag16	InnerTag16	InnerTag16	InnerTag16

Gel electrophoresis, PCR Clean-up

Sample normalization (Fluorometrically), pooling

Index Group pool1	Index Group pool2	Index Group pool3	Index Group pool9

Index PCR



Gel electrophoresis, PCR Clean-up

Sample normalization (qPCR), pooling

Pooled 16S rRNA library

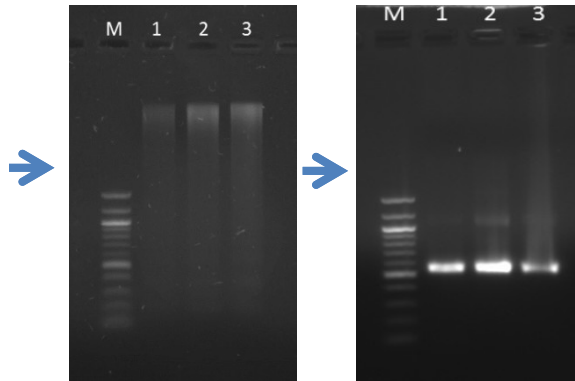
Check library (qPCR, Bioanalyzer)

Sequencing on MiSeq V3 chemistry, 2x300 bp (8pM library, 20 % PhiX DNA)

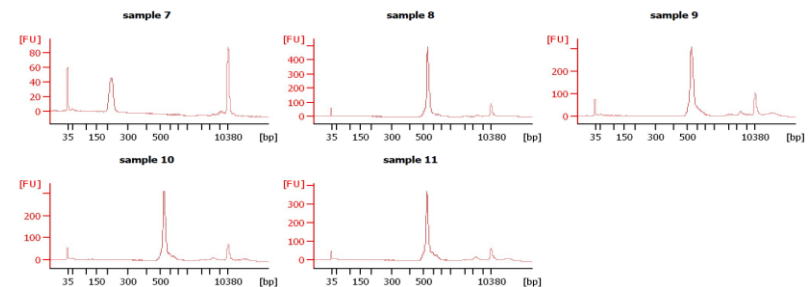
# Tipy a triky

- Vést si záznamy
  - Koncentrace vyizolované DNA
  - Gel s vyizolovanou DNA – ideálně nanášet stejně vyředěnou DNA
  - Gel s výsledky PCR, koncentrace po přečištění

Sample	Sample 1	Sample 2	Sample 3
Material	Buccal swab	Buccal swab	Buccal swab
Swab storage	No buffer, -20°C	Buffer BF1, 20°C	Buffer BF1, RT
Isolated DNA conc.	8.6 ng/μl	16.2 ng/μl	19.4 ng/μl

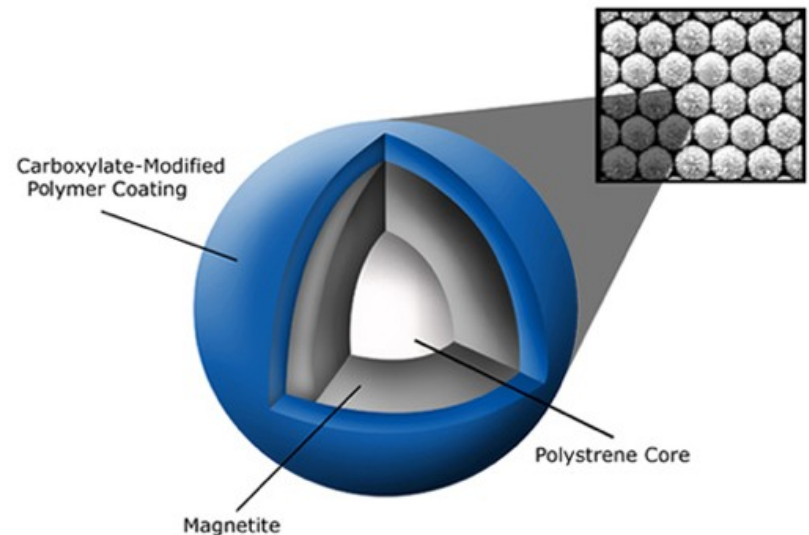


Sample	Sample 1	Sample 2	Sample 3
Concentration (ng/μl)	9.8	14.5	11.8
260/280	1.94	2.19	1.50
260/230	2.72	2.93	2.21



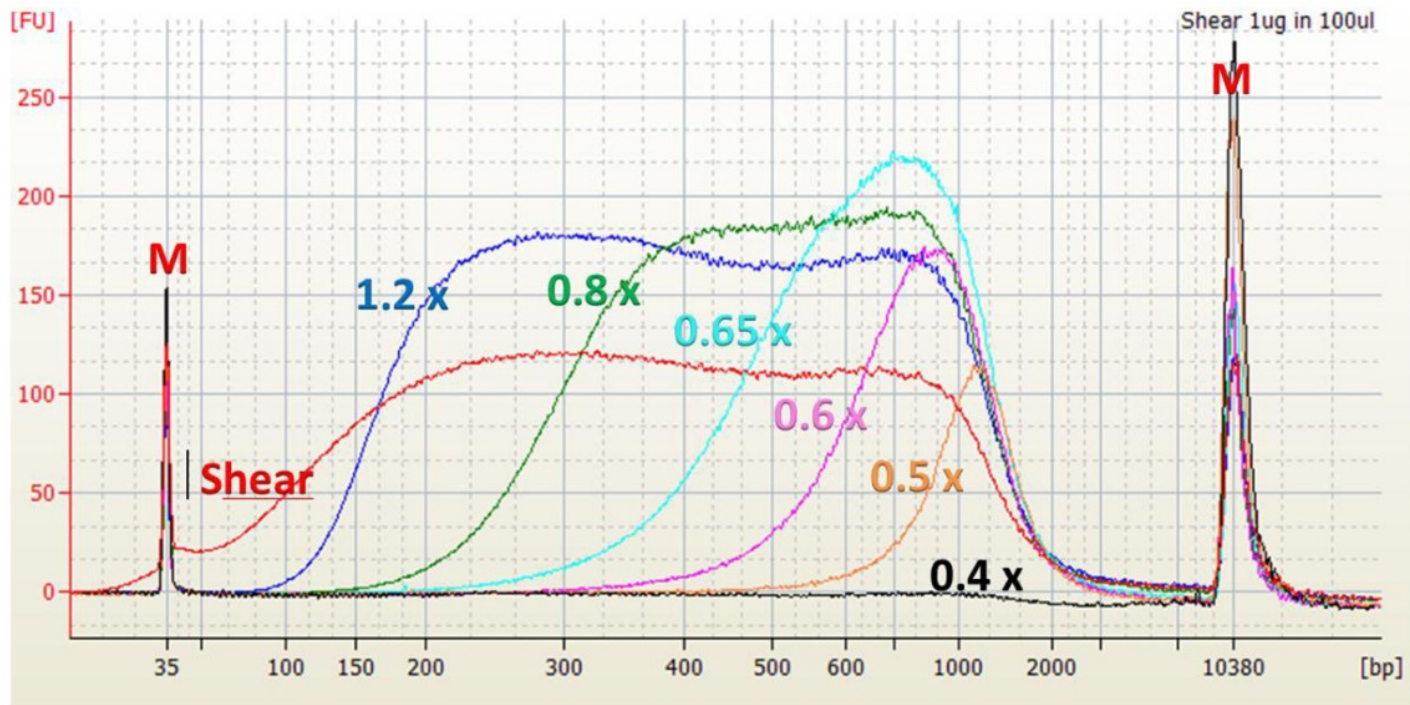
# Tipy a triky

- AMPpure beads x SPRI Select beads?
- AMPpure beads = SPRI Select beads
- <http://core-genomics.blogspot.cz/2012/04/how-do-spri-beads-work.html>



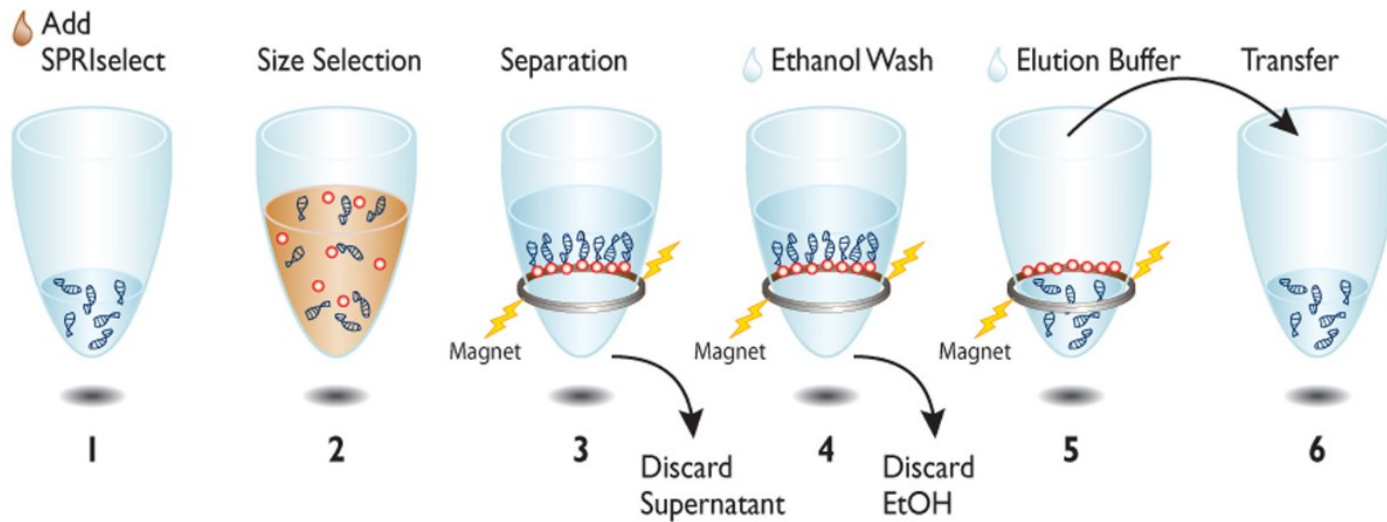
# Size selection

**Figure 1** Agilent High Sensitivity DNA chip Electropherogram.



# Size selection

## Left Side Size Selection Process Overview



# Size selection

	Definition	Diagram
<b>bp Region's % of Shear</b>	The maximum potential recovery of the Stock sample for the targeted size selection region.	
<b>Recovered % of bp Region</b>	Percent recovery of the targeted region	
<b>Recovered % of Shear</b>	Percent recovery of the Shear sample.	

# Kontrola délky fragmentů

## FlashGel (Lonza) x E-gel (Thermo Fisher Scientific)

[https://www.google.cz/search?q=e+gel,+flash+gel&ie=utf-8&oe=utf-8&gws\\_rd=cr&ei=Am3uVty6IYnZU47RnvgG#q=comparison+e+gel%2C+flash+gel](https://www.google.cz/search?q=e+gel,+flash+gel&ie=utf-8&oe=utf-8&gws_rd=cr&ei=Am3uVty6IYnZU47RnvgG#q=comparison+e+gel%2C+flash+gel)



FlashGel™ System

<http://www.lonza.com/products-services/bio-research/electrophoresis-of-nucleic-acids-and-proteins/nucleic-acid-electrophoresis/fast-electrophoresis-flashgel-system-for-dna-rna-and-recovery.aspx>



Load sample

Select desired run protocol and press "Go" button

Retrieve DNA ready for cloning

[https://tools.thermofisher.com/content/sfs/manuals/egelguide\\_man.pdf](https://tools.thermofisher.com/content/sfs/manuals/egelguide_man.pdf)

# FlashGel (Lonza) x E-gel (Thermo Fisher Scientific)

- 5 minute separation and recovery
- See bands in as little as 2 minutes Recover samples directly, without UV light, band excision or purification Real-time separation and documentation
- Watch band migration as it happens Photograph gels at the bench, without DNA damaging UV light Outstanding sensitivity and resolution
- The FlashGel™ System for fast electrophoresis consists of enclosed, disposable, precast agarose gel cassettes and a combination electrophoresis and transilluminator unit.
- 5 – 20 times more sensitive than ethidium bromide; detect <0.1 ng DNA or <10 ng total RNA



- FlashGel™ System showed the highest level of sensitivity at the lowest concentrations
- It also typically produced bands of the greatest intensity
- The only parameter at which the FlashGel™ System did not excel was clarity. However, clarity improved at lower concentrations. All of these parameters together can be seen in Figures 1 and 2. When taking all results into account, it appears that if lower concentrations of DNA samples are all that is available, the FlashGel™ System is the most appropriate choice. The E-Gel® Pre-cast Agarose Electrophoresis System yielded different results than the FlashGel™ system. Where the FlashGel™ system was weak, the E-Gel® Pre-cast Agarose Electrophoresis System excelled. The E-Gel® System was consistently clearer than the FlashGel™ system, except at lower concentrations. It showed less sensitivity and intensity than the FlashGel™ system and traditional agarose gel electrophoresis, as evident in Figures 1 and 2. The traditional self-cast agarose gels produced the most balanced results, remaining consistently strong across all measures of effectiveness. When separating the amplicons, the E-Gel® System received the highest ranking on the basis of its superior clarity.

# Kapilární elektroforéza

- Agilent – Bioanalyzer, TapeStation
- AATI - Fragment analyzer



# Bioanalyzer x TapeStation

## Comparison of systems for Sample QC in NGS

### 2100 Bioanalyzer

- High sensitivity
- Industry standard in xNA QC  
→ proven performance
- Low entry price
- Best-in-class software
- Versatile kit portfolio
- Matches well with automated DNA purification systems from Sage or Caliper

### 2200 TapeStation

- 96-well plate compatible for higher throughput
- Shorter hands-on-time and reduced user interaction due to higher degree of automation
- Very fast time-to-result due to parallel sample processing
- Scalable throughput at constant cost per sample

# Bioanalyzer x TapeStation

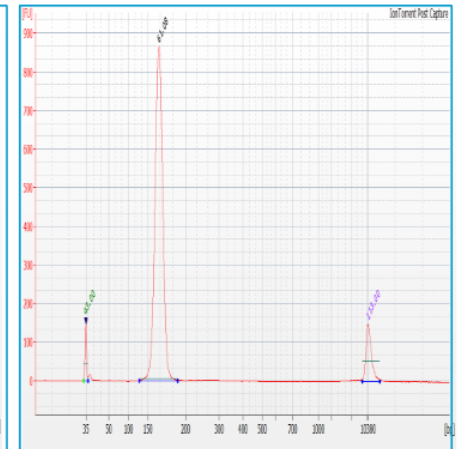
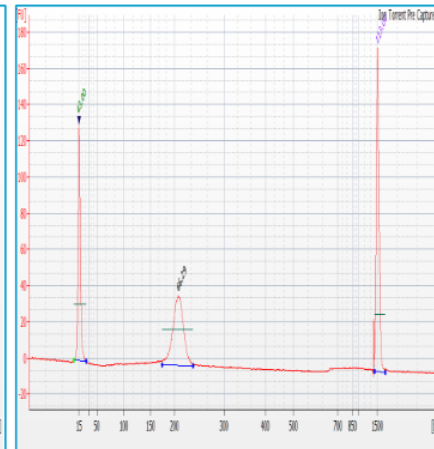
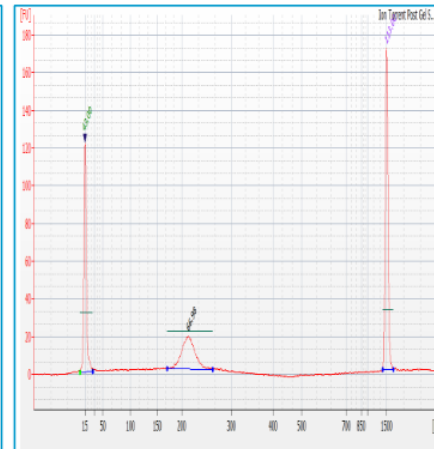
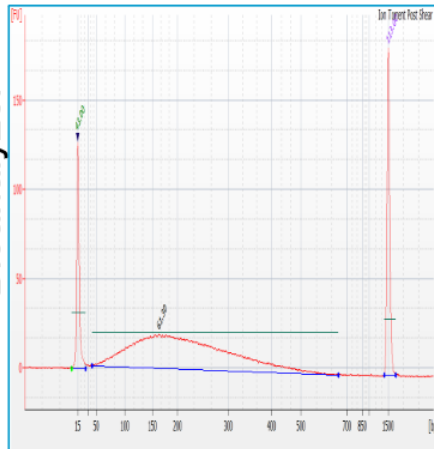
Post Shear

Post Gel Size Selection

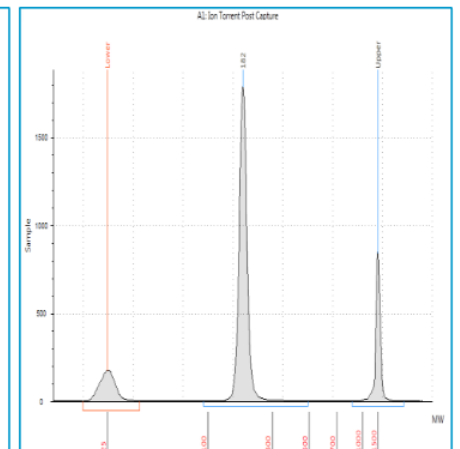
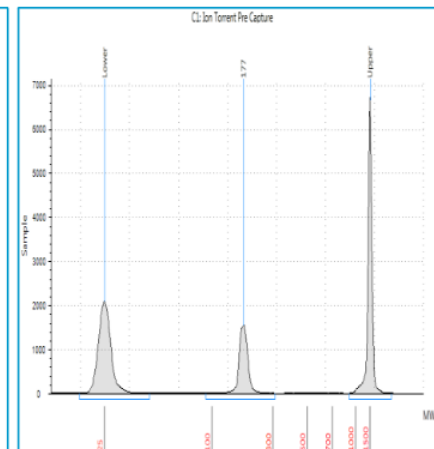
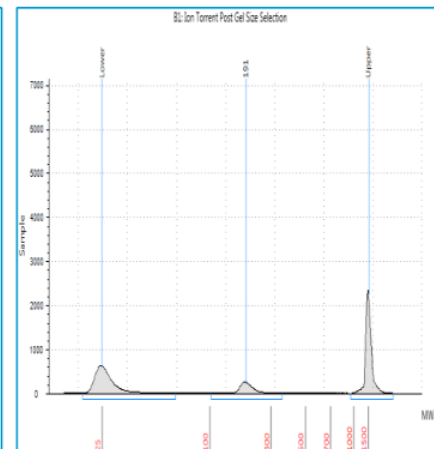
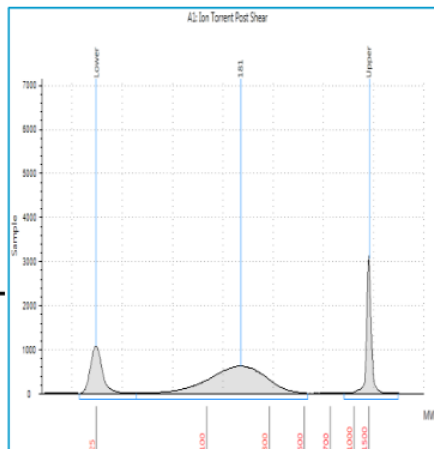
Pre Capture

Post Capture

Bioanalyzer



TapeStation

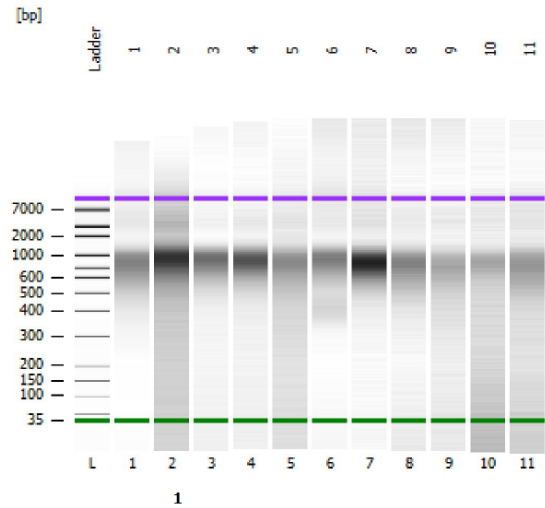


# Bioanalyzer

Assay Class: High Sensitivity DNA Assay  
Data Path: C:\...gh Sensitivity DNA Assay\_DE13806168\_2015-06-29\_14-44-06.xad

Created: 6/29/2015 2:44:05 PM  
Modified: 6/29/2015 3:48:23 PM

## Electrophoresis File Run Summary



### Instrument Information:

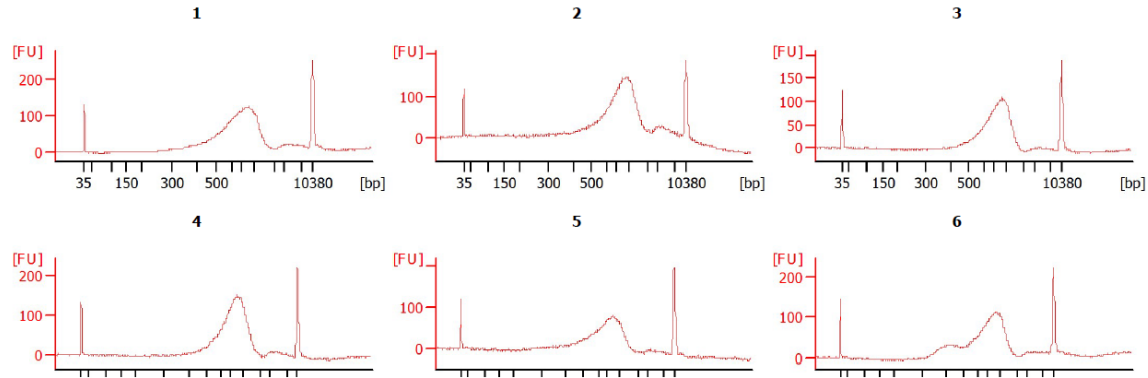
Instrument Name: DE13806168      Firmware: C.01.069  
Serial#: DE13806168      Type: G2939A

### Assay Information:

Assay Origin Path: C:\Program Files (x86)\Agilent\2100 bioanalyzer\2100 expert\assays\dsDNA\High Sensitivity DNA.xsy  
Assay Class: High Sensitivity DNA Assay  
Version: 1.03  
Assay Comments: Copyright © 2003-2010 Agilent Technologies

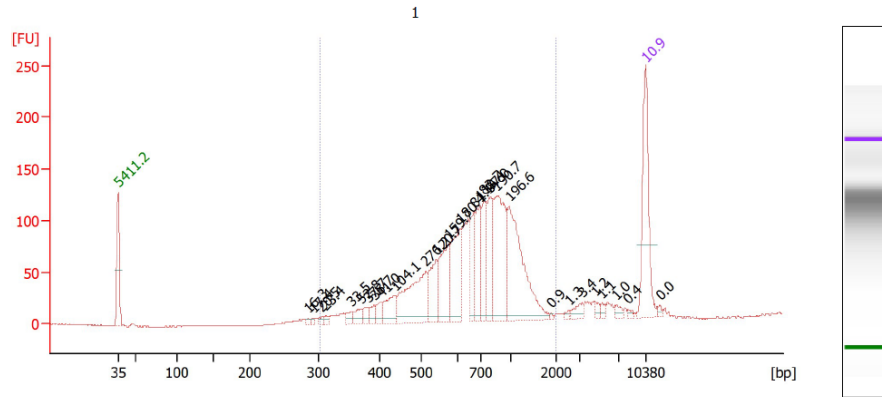
### Chip Information:

Chip Lot #:  
Reagent Kit Lot #:  
Chip Comments: S56 libraries 111TSP1/3



# Bioanalyzer

## Electropherogram Summary Continued ...



### Overall Results for sample 1 : 1

Number of peaks found: 28      Corr. Area 1: 1,681.0  
 Noise: 0.4

### Peak table for sample 1 : 1

Peak	Size [bp]	Conc. [pg/μl]	Molarity [pmol/l]	Observations
1	35	125.00	5,411.3	Lower Marker
2	284	3.06	16.3	
3	291	3.35	17.4	
4	304	4.12	20.5	
5	311	4.80	23.4	
6	353	7.81	33.5	
7	371	12.92	52.8	
8	382	9.51	37.7	
9	392	9.49	36.7	
10	404	10.91	41.0	
11	438	30.12	104.1	
12	511	93.07	276.2	
13	540	43.03	120.7	
14	572	60.31	159.7	
15	609	72.52	180.4	
16	670	35.91	81.1	
17	695	54.79	119.4	
18	720	20.84	83.7	

# Bioanalyzer

## Electropherogram Summary Continued ...

... Peak table for sample 1 : 1

Peak	Size [bp]	Conc. [pg/ $\mu$ l]	Molarity [pmol/l]	Observations
27	6,749	4.49	1.0	
28	8,093	2.11	0.4	
29	10,380	75.00	10.9	Upper Marker
30	12,227	0.00	0.0	

Region table for sample 1 : 1

From [bp]	To [bp]	Corr. Area	% of Total	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/ $\mu$ l]	Molarity [pmol/l]	Color
302	1,998	1,681.0	95	746	35.1	775.73	1,822.3	