Metagenomika – virom a eukaryota

Petra Vídeňská, Ph.D.

Eukaryota

- Problém s velikými rozdíly (dělají se zejména plísně, kvasinky, vyšší houby, paraziti a protozoa)
- Lze využít různé markery – 18S rDNA, ITS 1/2, D2…
- Která oblast je nejlepší na co – literatura
- Nyní se nejvíce využívá celometagenomové sekvenování

Ukázka využití genu pro 18S rRNA

ARTICLE IN PRESS

PARINT-01455; No of Pages 4

Parasitology International xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Parasitology International

journal homepage: www.elsevier.com/locate/parint

A novel method to assess the biodiversity of parasites using 18S rDNA Illumina sequencing; parasitome analysis method

Akina Hino [1], Haruhiko Maruyama, Taisei Kikuchi *

Division of Parasitology, Faculty of Medicine, University of Miyazaki, Miyazaki 889-1692, Japan

ARTICLE INFO

Article history:
Received 27 December 2015
Received in revised form 6 January 2016
Accepted 14 January 2016
Available online xxxx

Keywords:
Next generation sequencer (NGS)
Parasite diversity
Metagenome
18S ribosomal RNA

ABSTRACT

Understanding parasite diversity has important implications in several research fields, including ecology, evolutionary biology, and epidemiology. Here, we introduce a novel method to assess the biodiversity of parasites—especially those in the host alimentary tract—using an 18S rDNA-based metagenomic approach. The method is easy and quick compared to conventional methods, and does not require dissections of host bodies or identification skills for various parasite species. The use of a "next generation sequencer" in this method allows us to perform the assessment in a high throughput manner, which will increase our knowledge of parasite diversity.

Ukázka využití genu pro 18S rRNA

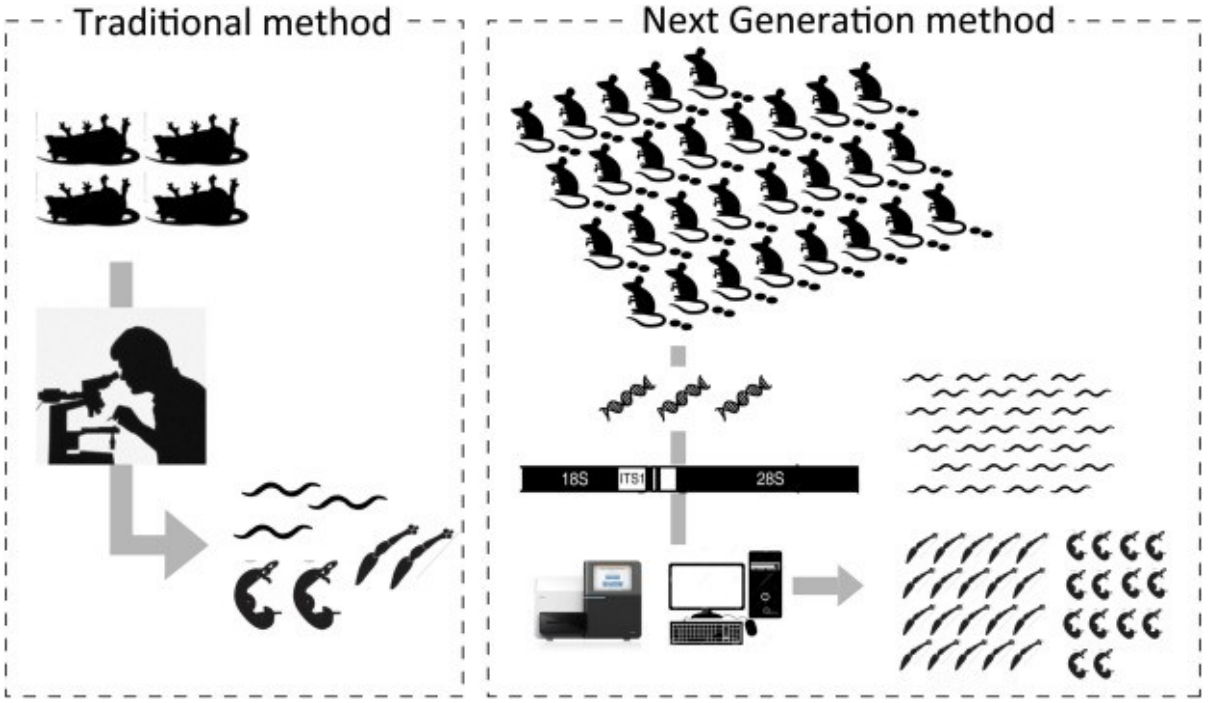


Fig. 1. An illustration to compare the traditional method and the 18S rRNA based metagenome approach using next generation sequencers. With 18S rRNA based metagenomics, parasites can be detected and identified in a high-throughput manner.

Ukázka využití genu pro 18S rRNA

Table 2. Primers and oligonucleotides used in this method.

Primer	Sequence
Illumina_Euk_1391f PCR Primer*	AATGATACGGCGACCACCGAGATCTACAC TATCGCCGTT CG GTACACACCGCCCGTC
Illumina_EukBr PCR primer sequence**	CAAGCAGAAGACGGCATACGAGAT XXXXXXXXXXXX AGTCAGTCAG CA TGATCCTTCTGCAGGTTCACCTAC
Mammal blocking primer	GCCCGTCGCTACTACCGATTGGIIIIITTAGTGAGGCCCT-[C3 Spacer]
Euk_illumina_read1_seq_primer	TATCGCCGTT CG GTACACACCGCCCGTC
Euk_illumina_read2_seq_primer	AGTCAGTCAG CA TGATCCTTCTGCAGGTTCACCTAC
Euk_illumina_index_seq_primer	GTAGGTGAACCTGCAGAAGGATCA TG CTGACTGACT

* Space-delimited sequences indicate, from left to right, 5′ Illumina adaptor, forward primer pad, forward primer linker and forward primer.

** Space-delimited sequences indicate, from left to right, reverse complement of 3′ Illumina adapter, Golay barcode, reverse primer pad, reverse primer linker and reverse primer. Golay barcodes designated by Xs, allowing multiple samples to be distinguishable, are available at http://www.earthmicrobiome.org/ [12].

Ukázka využití genu pro 18S rRNA

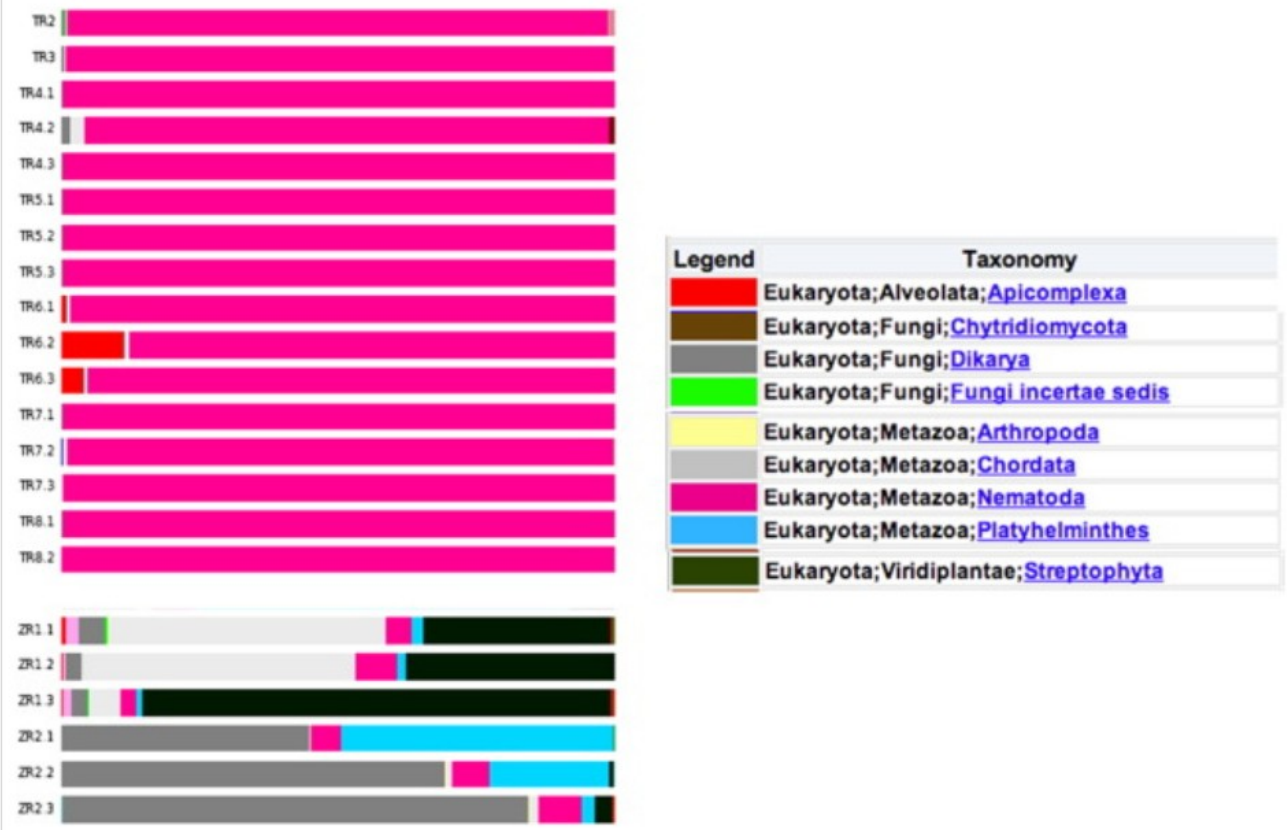
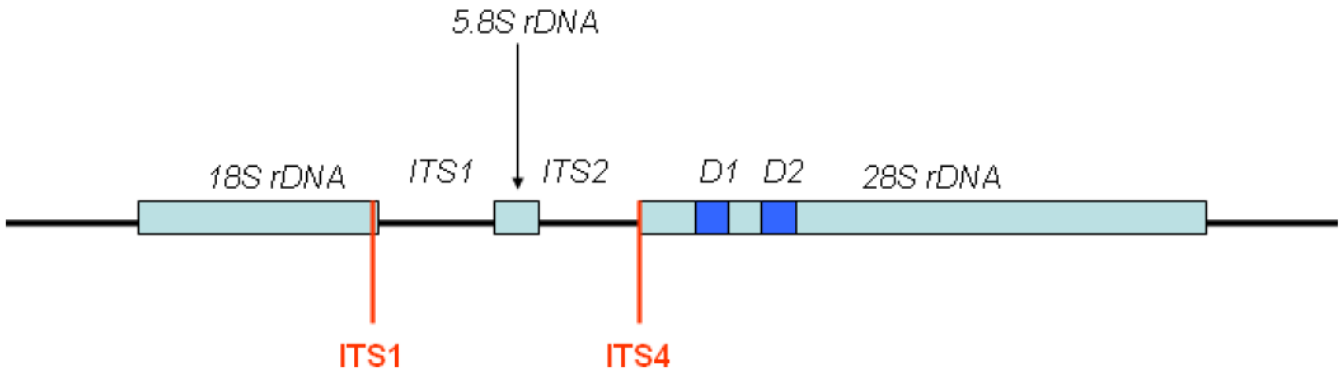


Fig. 3.

An example of QIIME phylum-level classification of the 18S rDNA Illumina sequencing data. QIIME package provides several tools to visualise the classification results in various types of charts or plots. In this stacked bar charts, each horizontal bar represents relative abundance of each eukaryotic taxon in a sample, indicating that parasite compositions are significantly different between TR-labelled samples and ZR-labelled samples.

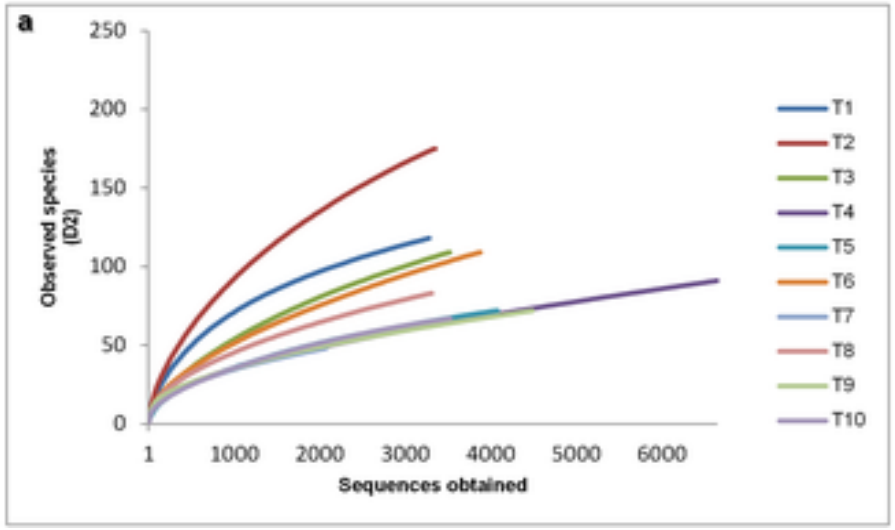
Mikrobiální diverzita na *Vitis vinifera*

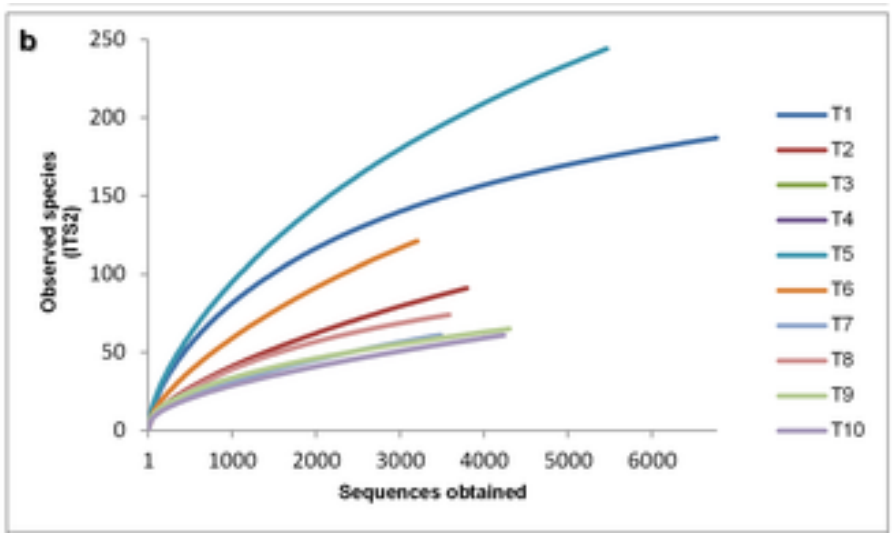
- sběr zdravých i napadených listů *V. vinifera* cv Tempranillo
- odběr 10x od května do června → T1-T10
- uskladnění listů při -80 C, izolace DNA
- příprava knihovny oblasti V6 16S rDNA pro prokaryotickou populaci
- příprava knihovny ITS2 a D2 pro eukaryotickou populaci

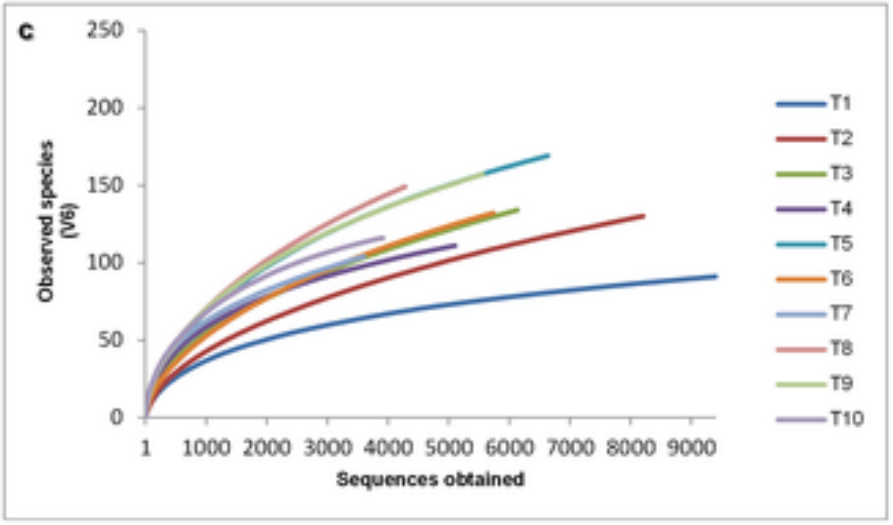


Gryndler, 2013

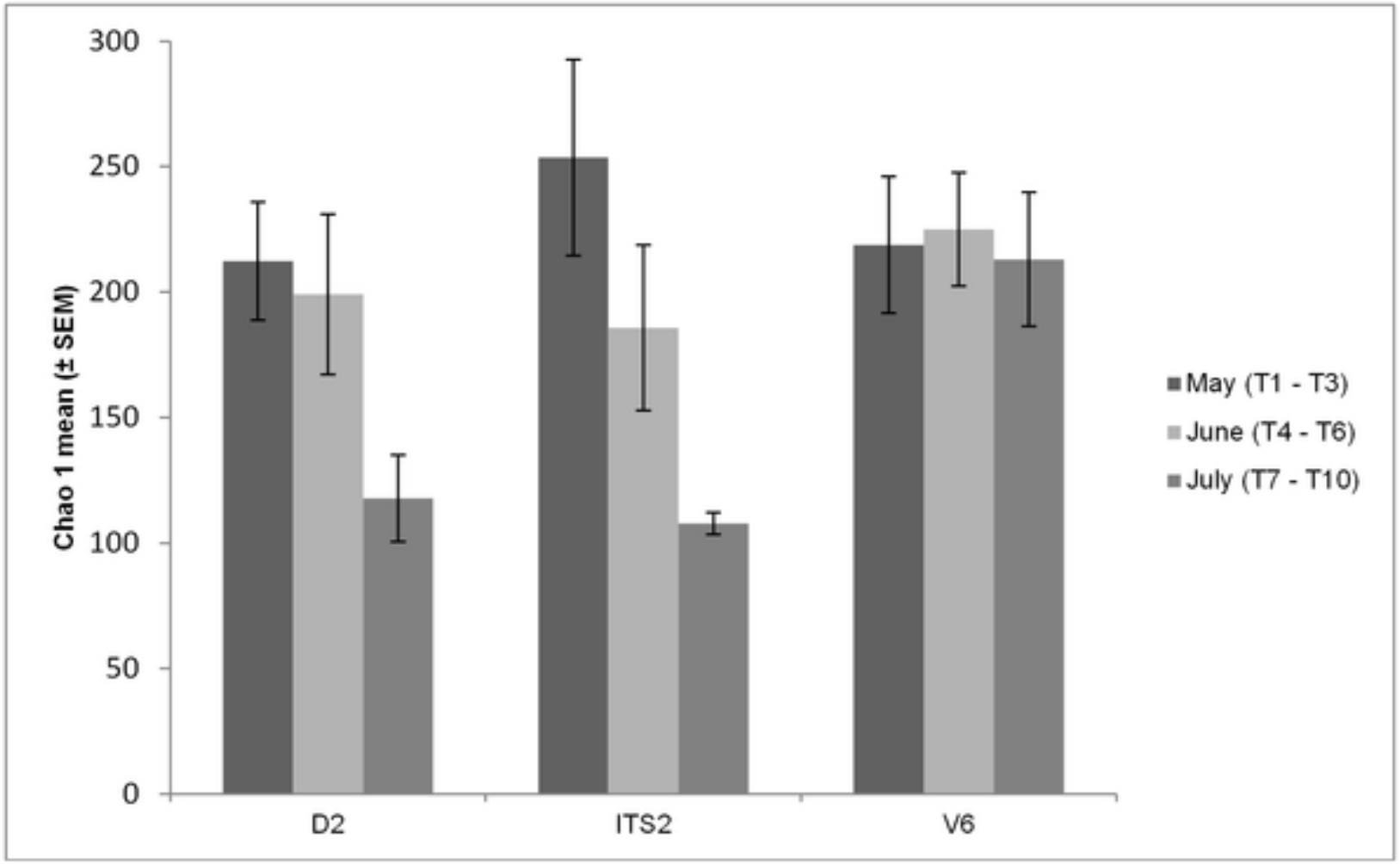
Rarefakční křivky



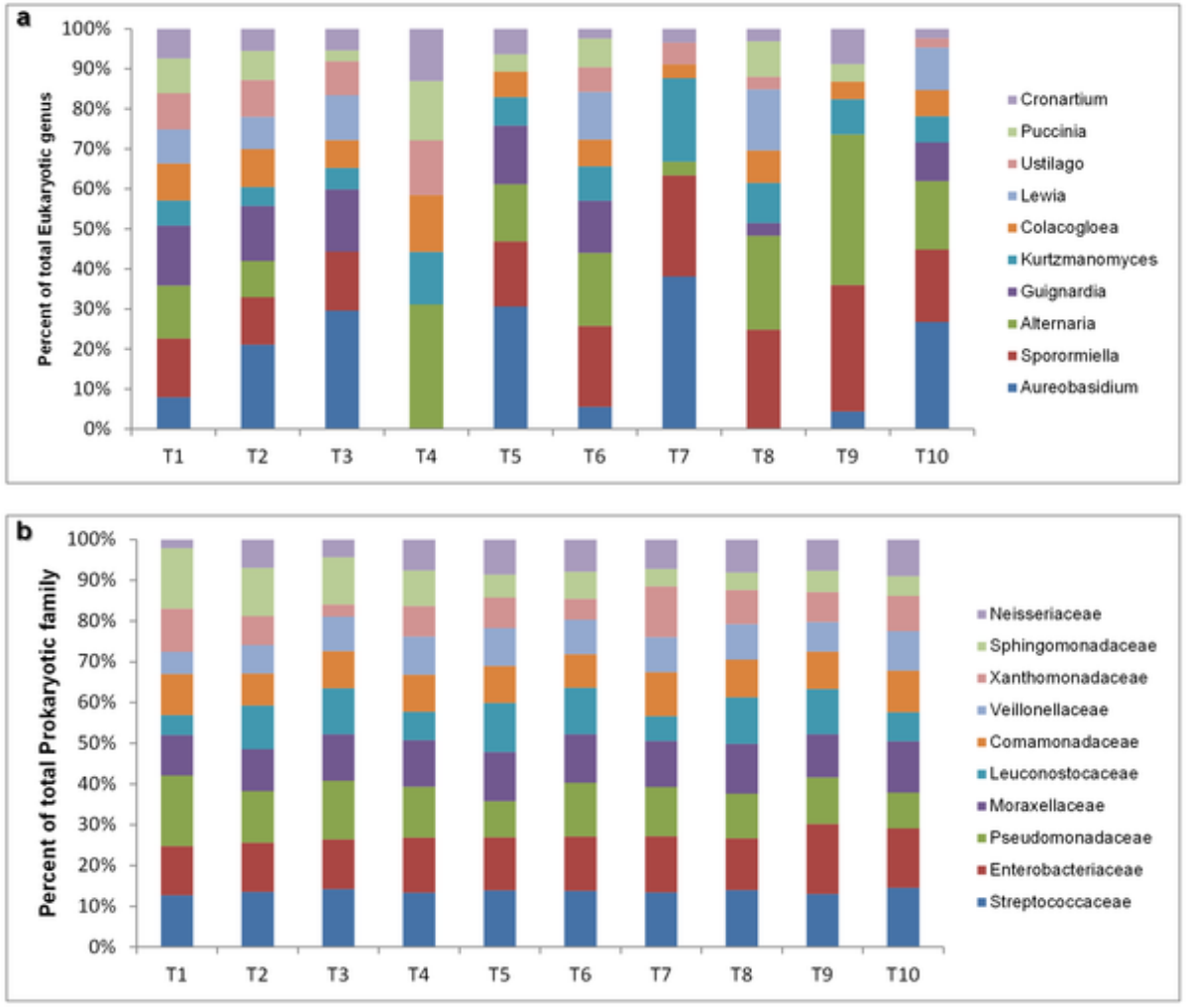




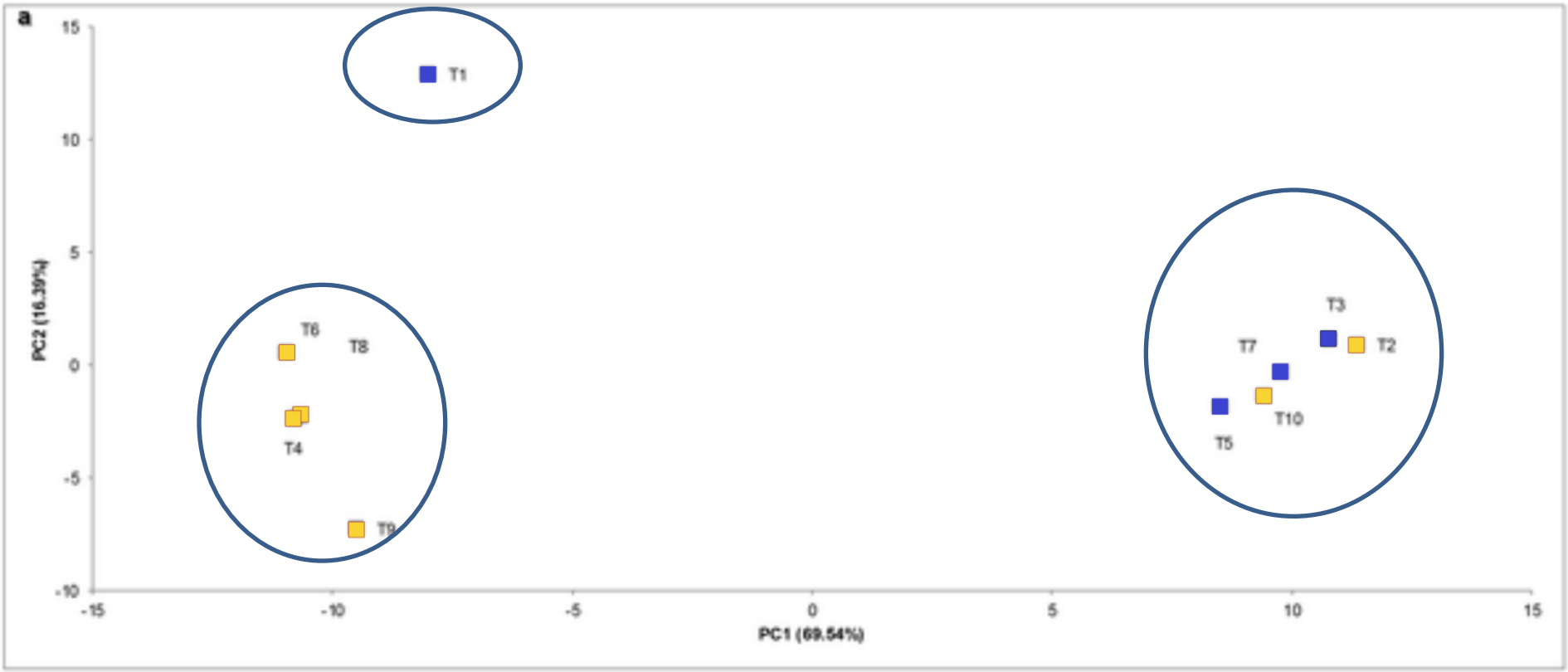
Dynamika biodiverzity



Distribuce eukaryontních i prokaryontních společentví v průběhu času



PCA Biplot mikrobiální komunity



Virom

- Neexistuje univerzální konzervovanou oblast
- Není sjednocená metodika
- Problém RNA x DNA viry
- Ve vzorku je většinou virální DNA/RNA minimum

Srovnání metod

SCIENTIFIC REPORTS

OPEN

Received: 19 August 2015
Accepted: 15 October 2015
Published: 12 November 2015

Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis

Nádia Conceição-Neto[1,2], Mark Zeller[1], Hanne Lefrère[1], Pieter De Bruyn[1], Leen Beller[1], Ward Deboutte[1], Claude Kwe Yinda[1,2], Rob Lavigne[3], Piet Maes[2], Marc Van Ranst[2], Elisabeth Heylen[1,*] & Jelle Matthijnssens[1,2,*]

A major limitation for better understanding the role of the human gut virome in health and disease is the lack of validated methods that allow high throughput virome analysis. To overcome this, we evaluated the quantitative effect of homogenisation, centrifugation, filtration, chloroform treatment and random amplification on a mock-virome (containing nine highly diverse viruses) and a bacterial mock-community (containing four faecal bacterial species) using quantitative PCR and next-generation sequencing. This resulted in an optimised protocol that was able to recover all viruses present in the mock-virome and strongly alters the ratio of viral versus bacterial and 16S rRNA genetic material in favour of viruses (from 43.2% to 96.7% viral reads and from 47.6% to 0.19% bacterial reads). Furthermore, our study indicated that most of the currently used virome protocols, using small filter pores and/or stringent centrifugation conditions may have largely overlooked large viruses present in viromes. We propose NetoVIR (Novel enrichment technique of VIRomes), which allows for a fast, reproducible and high throughput sample preparation for viral metagenomics studies, introducing minimal bias. This procedure is optimised mainly for faecal samples, but with appropriate concentration steps can also be used for other sample types with lower initial viral loads.

Postup

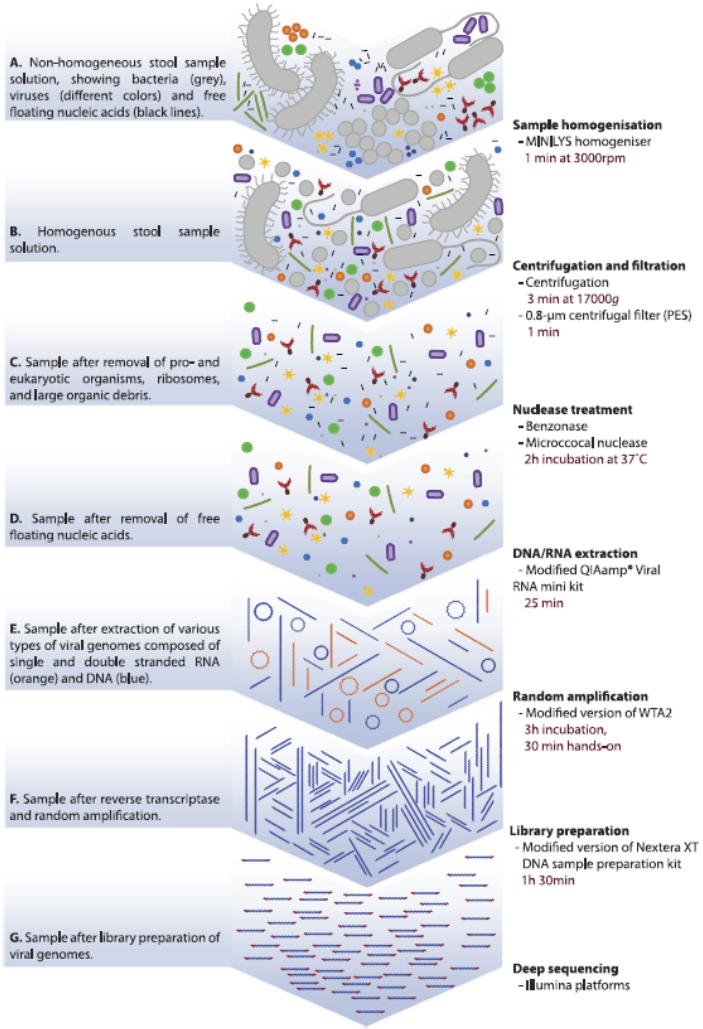


Figure 1. Schematic concise description of the proposed NetoVIR protocol. Estimations of incubation time and total time for each step are shown. On average, the protocol takes 8 h to complete. A detailed protocol is described in Protocol S1 (Supplementary information).

Vliv homogenizace

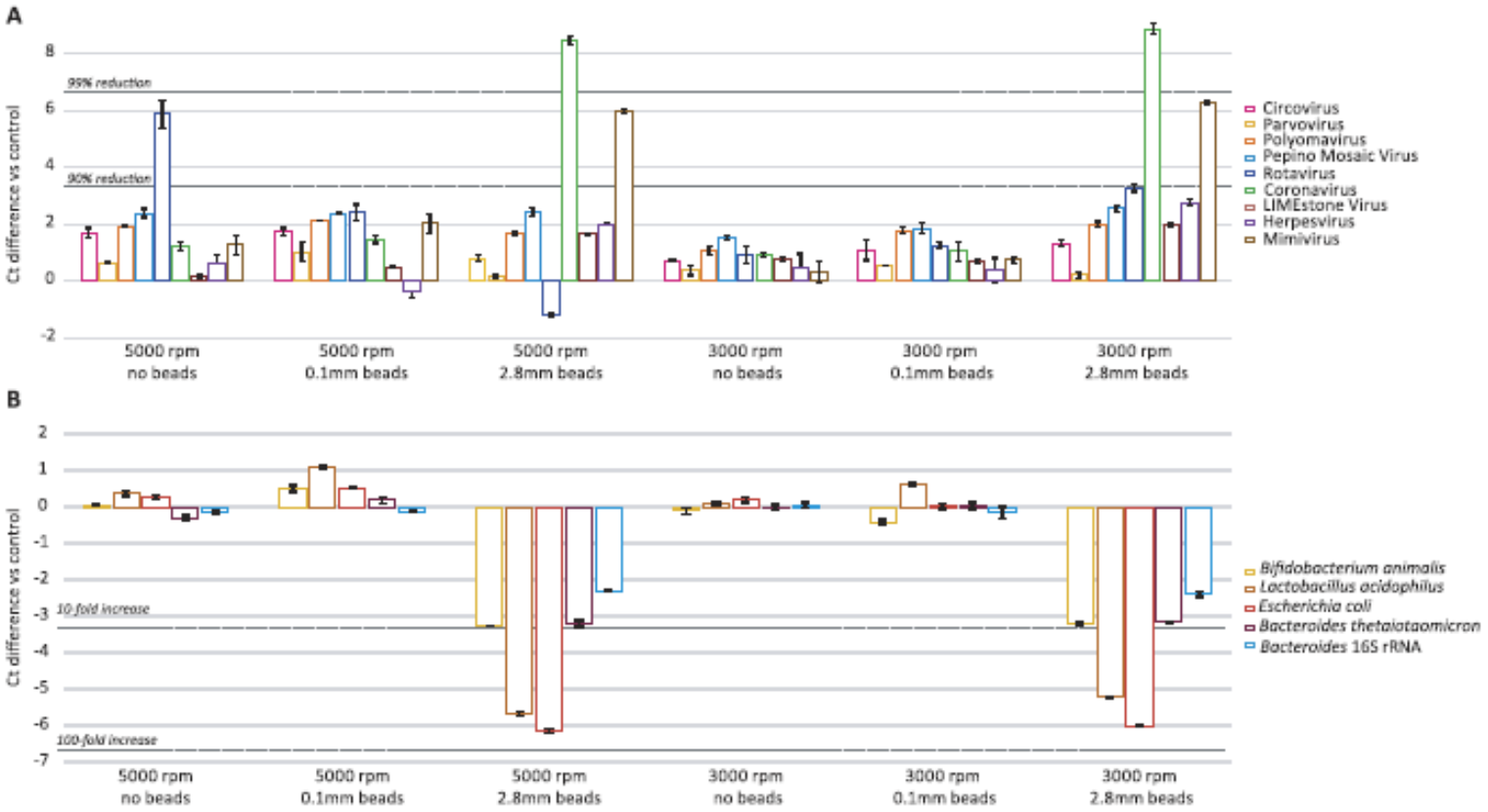


Figure 2. Ct differences vs control for different homogenisation experiments performed on the mock-virome (A) and on the bacterial mock-community and *Bacteroides* 16S rRNA (B). Standard deviations are based on three qPCR replicates.

Vliv centrifugace

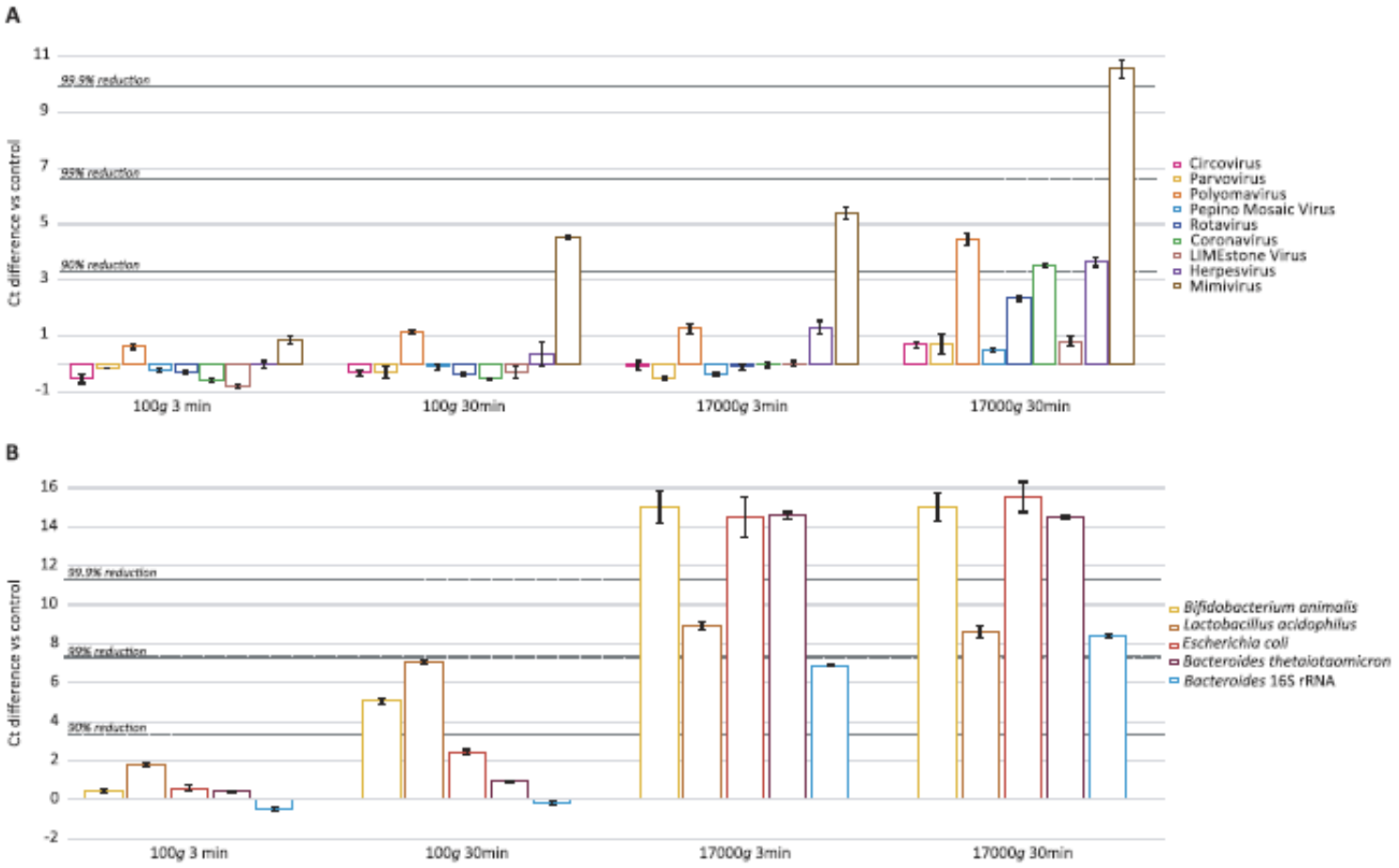


Figure 3. Ct differences vs control for centrifugation conditions tested on the mock-virome (A), on the bacterial mock-community and *Bacteroides* 16S rRNA (B). Standard deviations of the qPCR replicates are displayed.

Vliv filtrace

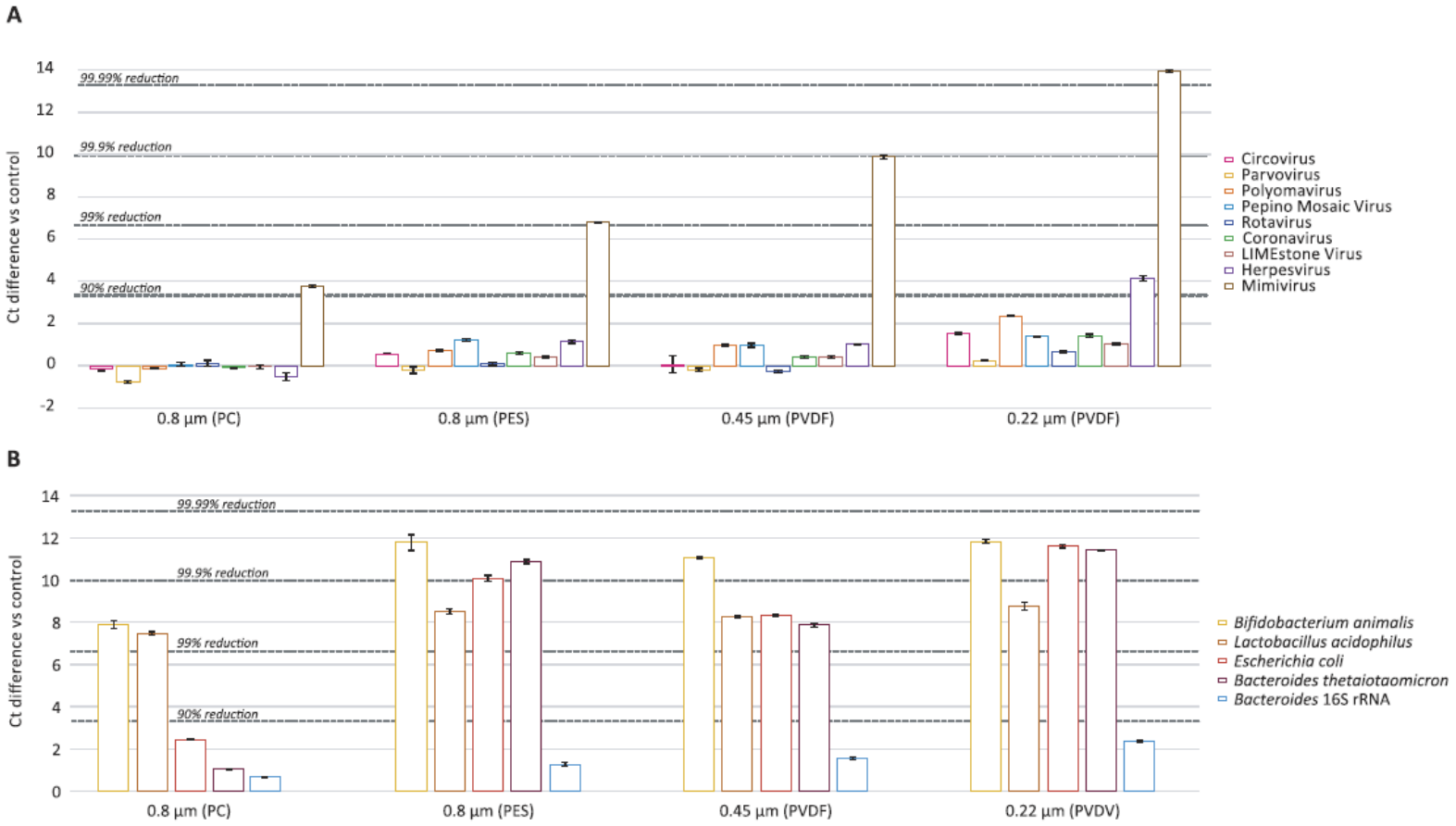


Figure 4. Ct differences vs control for filtration experiments performed on the mock-virome (A), bacterial mock-community and *Bacteroides* 16S rRNA (B). Standard deviations of the qPCR replicates are displayed.

Vliv ošetření chloroformem

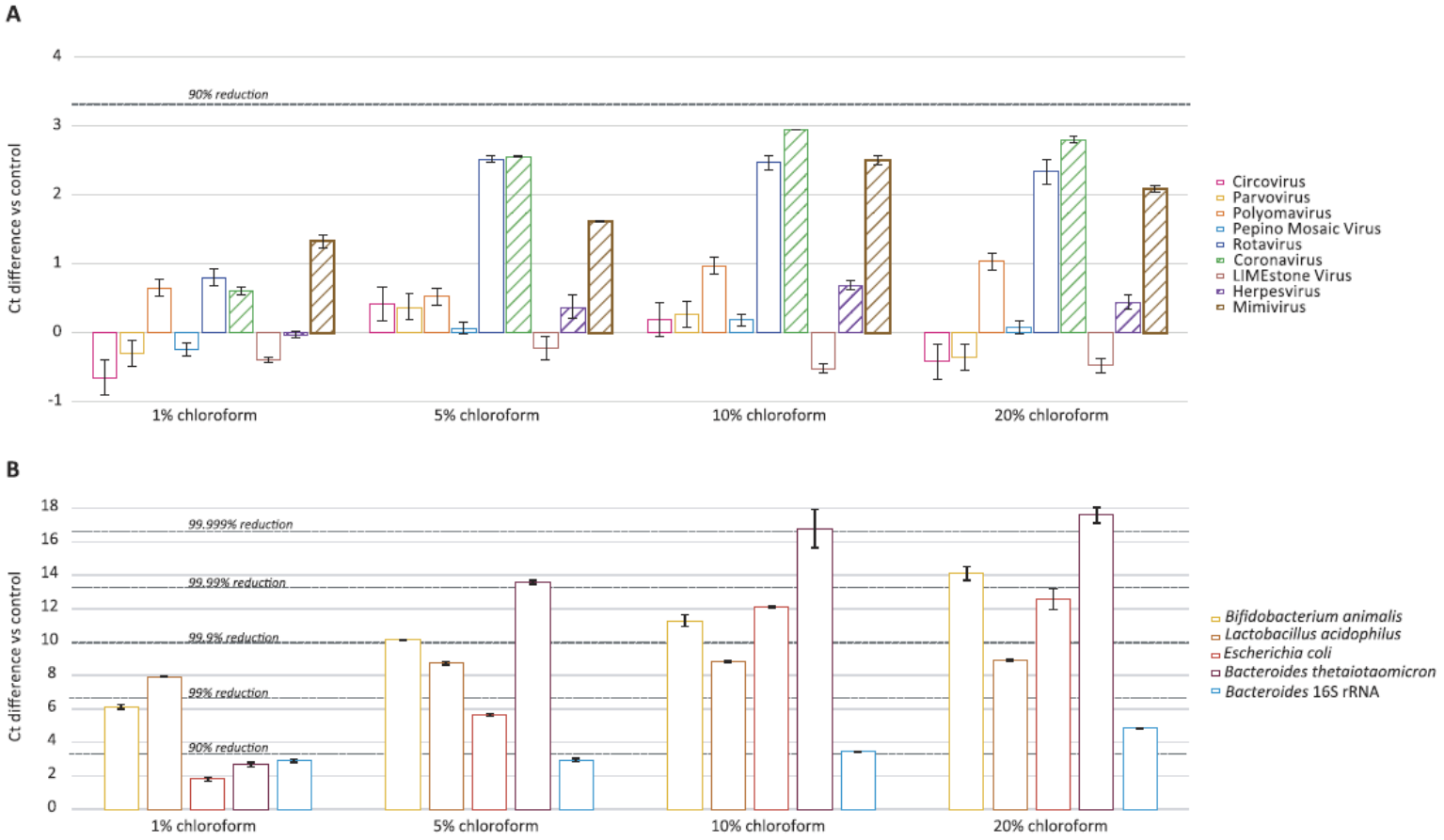


Figure 5. Ct differences vs control for chloroform treatment experiments performed on the mock-virome (A), bacterial mock-community and *Bacteroides* 16S rRNA (B). Enveloped viruses are depicted with a pattern. Standard deviations of the qPCR replicates are displayed.

Vliv amplifikace s náhodnými hexamery

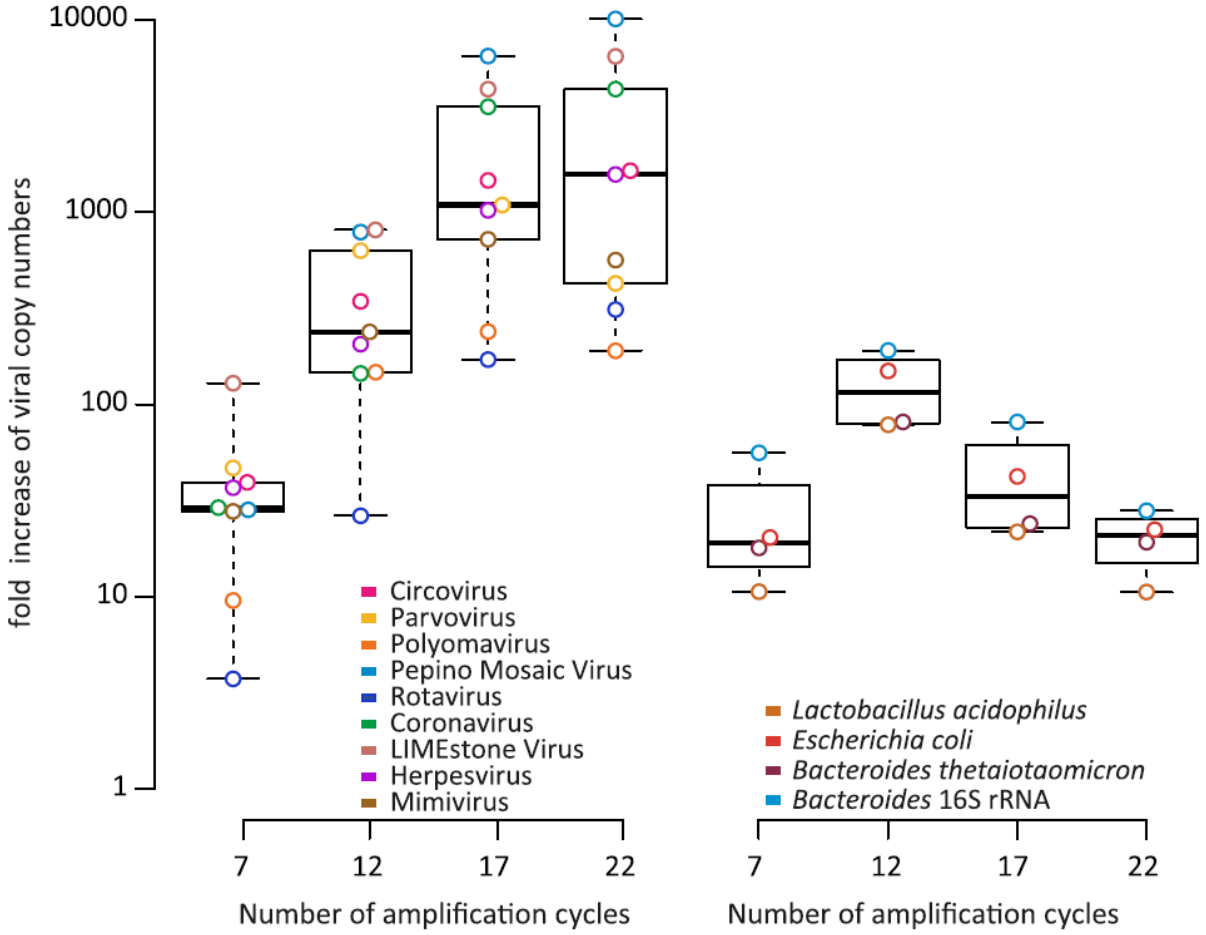


Figure 6. Fold increase vs control for random amplification experiments performed on the mock-virome, mock bacterial community and *Bacteroides* 16S rRNA. *Bifidobacterium animalis* is not shown since no amplification was observed.

Srovnání metod

Journal of Virological Methods 195 (2014) 194–204

Contents lists available at ScienceDirect

Journal of Virological Methods

journal homepage: www.elsevier.com/locate/jviromet

Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery

Richard J. Hall*, Jing Wang, Angela K. Todd, Ange B. Bissielo, Seiha Yen, Hugo Strydom, Nicole E. Moore, Xiaoyun Ren, Q. Sue Huang, Philip E. Carter, Matthew Peacey

Institute of Environmental Science and Research, at the National Centre for Biosecurity & Infectious Disease, 66 Ward Street, Wallaceville, Upper Hutt 5018, New Zealand

Article history:
Received 23 May 2013
Received in revised form 26 August 2013
Accepted 29 August 2013
Available online 13 September 2013

Keywords:
Metagenomic
Virus
Purification
Enrichment

A B S T R A C T

The discovery of new or divergent viruses using metagenomics and high-throughput sequencing has become more commonplace. The preparation of a sample is known to have an effect on the representation of virus sequences within the metagenomic dataset yet comparatively little attention has been given to this. Physical enrichment techniques are often applied to samples to increase the number of viral sequences and therefore enhance the probability of detection. With the exception of virus ecology studies, there is a paucity of information available to researchers on the type of sample preparation required for a viral metagenomic study that seeks to identify an aetiological virus in an animal or human diagnostic sample. A review of published virus discovery studies revealed the most commonly used enrichment methods, that were usually quick and simple to implement, namely low-speed centrifugation, filtration, nuclease-treatment (or combinations of these) which have been routinely used but often without justification. These were applied to a simple and well-characterised artificial sample composed of bacterial and human cells, as well as DNA (adenovirus) and RNA viruses (influenza A and human enterovirus), being either non-enveloped capsid or enveloped viruses. The effect of the enrichment method was assessed by both quantitative real-time PCR and metagenomic analysis that incorporated an amplification step. Reductions in the absolute quantities of bacteria and human cells were observed for each method as determined by qPCR, but the relative abundance of viral sequences in the metagenomic dataset remained largely unchanged. A 3-step method of centrifugation, filtration and nuclease-treatment showed the greatest increase in the proportion of viral sequences. This study provides a starting point for the selection of a purification method in future virus discovery studies, and highlights the need for more data to validate the effect of enrichment methods on different sample types, amplification, bioinformatics approaches and sequencing platforms. This study also highlights the potential risks that may attend selection of a virus enrichment method without any consideration for the sample type being investigated.

Srovnání metod

Virus enrichment process prior to sequencing in metagenomic studies on human and animal samples.

Paper title	Author	Year	Journal	Aim of study	Sample	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Amplification	Sequencing
A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species	Allander et al.	2001	PNAS	Development of a method for discovery of unknown viruses and elimination of contaminat-ing host DNA. Allowed the discovery of novel bovine parvoviruses.	Serum	0.22 μm centrifugal filtration at 2,000 × g	DNAse I 2 U/μL^{-1} at 37°C for 2 hours	–	–	–	–	–	Sequence-Independent Amplification	Cloning and sanger-method
Metagenomic analyses of an uncultured viral community from human faeces	Breitbart et al.	2003	Journal of Bacteriology	Metagenomic analyses of an uncultured viral community from human faeces	Faecal suspension	≈ 100 μm Nitex filter	100 kDa tangential flow filter	Density gradient ultracentrifugation, CsCl 1.38 to 1.5 g mL^{-1} fraction collected	–	–	–	–	Sequence-independent Amplification	Cloning and sanger-method
Identification of a new human coronavirus	van der Hoek et al.	2004	Nature Medicine	Identification of a new human coronavirus	Suspension of LLC-MK2 cells	10 minutes at 13,500 rpm in microcentrifuge	DNAse I for 45 minutes at 37°C (Ambion)	–	–	–	–	–	VIDISCA	cDNA-AFLP, cloning and sanger-method
Viral genome sequencing by random priming methods	Djikeng et al.	2008	BMC Genomics	Development of a method for rapid sequencing of whole genomes from new viruses	Bacterial growth media, plasma, cell culture supernatant, faecal suspension	Low speed centrifugation?	0.22 μm filtration	149,000 × g (unspecified time)	2 U mL^{-1} DNAse I and/or 10 μg mL^{-1} RNAse A at 37°C for 1 hour	–	–	–	Sequence-Independent Amplification	Cloning and sanger-method
A highly divergent Picornavirus in a marine mammal	Kapoor et al.	2008	Journal of Virology	Unidentified virus cultured from a seal	Supernatant from infected Vero cell culture	5000 × g for 10 minutes	0.45 μm filtration	39,800 × g for 3 hours at 10°C	0.2 U μL^{-1} Turbo DNAse (Ambion) at 37°C for 90 minutes	–	–	–	Sequence-Independent Amplification	Cloning and sanger-method
Rapid identification of known and new RNA viruses from animal tissues	Victoria et al.	2008	PLOS Pathogens	Unidentified viruses cultured in suckling mouse brains	Brain tissue homogenate from mice	2500 rpm at 4°C for 20 minutes	0.45 μm filtration	22,000 × g 2 hours at 4°C	14 U Turbo DNAse (Ambion), 20 U benzonase (Novagen) and 20 U Roche One Phromega) at 37°C for 90 minutes in 1X Drase buffer (Ambion)	–	–	–	Sequence-Independent Amplification	Cloning and sanger-method
Discovery of a novel single-stranded DNA virus from a Sea Turtle Fibropapilloma by using viral metagenomics	Ng et al.	2009	Journal of Virology	To investigate and purifiy the viruses associated with sea turtle fibropa-pillomatosis (FP)	External fibropapil-loma homogenate	10,000 × g at 4°C for 10 minutes	0.45 μm filtration (Millipore)	Density CsCl gradient ultracentrifugation 60,000 × g at 4°C for 3 hours, 1.2 to 1.5 g mL^{-1} fraction collected	Incubated viral fraction for 10 minutes with 0.2 volumes chloroform	Supernatant removed from chloroform and incubated with 2.5U Dnase I (Sigma-Aldrich)	–	–	Sequence-Independent Amplification	Cloning and sanger-method

Low-speed centrifugation.
Filtration (excludes tangential flow).
Ultracentrifugation.
Nuclease treatment.
Unclassified method.

Srovnání metod

Quantity and proportion of sequence reads with a positive BLASTN hit against the model organism groups used in the virus discovery metagenomic dataset, comparing the effect of different virus enrichment methods.

Treatment	Total number of sequence reads[a]	Metazoa		*Enterobacteriaceae*		Adenovirus		Influenza		Enterovirus	
		BLASTN hits	% total	BLASTN hits	% total	BLASTN hits	% total	BLASTN hits	% total	BLASTN hits	% total
No treatment	1,980,878	39,481	1.99%	1,523,514	76.9%	40	0.002%	20	0.001%	3,221	0.16%
Centrifuge	2,010,717	48,719	2.42%	1,486,125	73.9%	0	0.000%	77	0.004%	14,805	0.74%
Filtration	1,941,626	65,746	3.39%	1,334,433	68.7%	6	0.000%	110	0.006%	22,731	1.17%
Nuclease	1,821,828	5,148	0.28%	1,421,268	78.0%	17	0.001%	14	0.001%	2,532	0.14%
2-step treatment[b]	1,730,569	53,421	3.09%	1,199,232	69.3%	14	0.001%	57	0.003%	18,712	1.08%
3-step treatment[b]	1,417,803	26,856	1.89%	857,873	60.5%	16	0.001%	161	0.011%	67,227	4.74%

[a] Combined total number of sequence reads for two independent physical replicates which were also run on different Illumina MiSeq flowcells. This figure represents the collapsed sequencing data, therefore redundant reads are not represented more than once.

[b] Serial applications of treatment methods. The 2-step method consisted of centrifugation then filtration. The 3-step method consisted of centrifugation, filtration then nuclease-treatment.