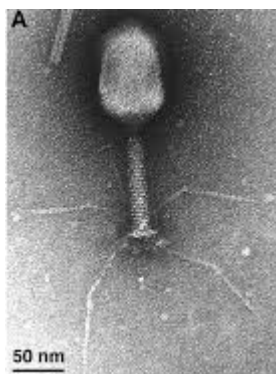
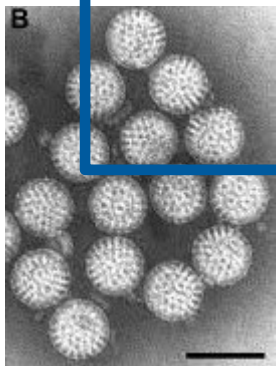


Sekvenování viromu pomocí NGS: PROČ a JAK

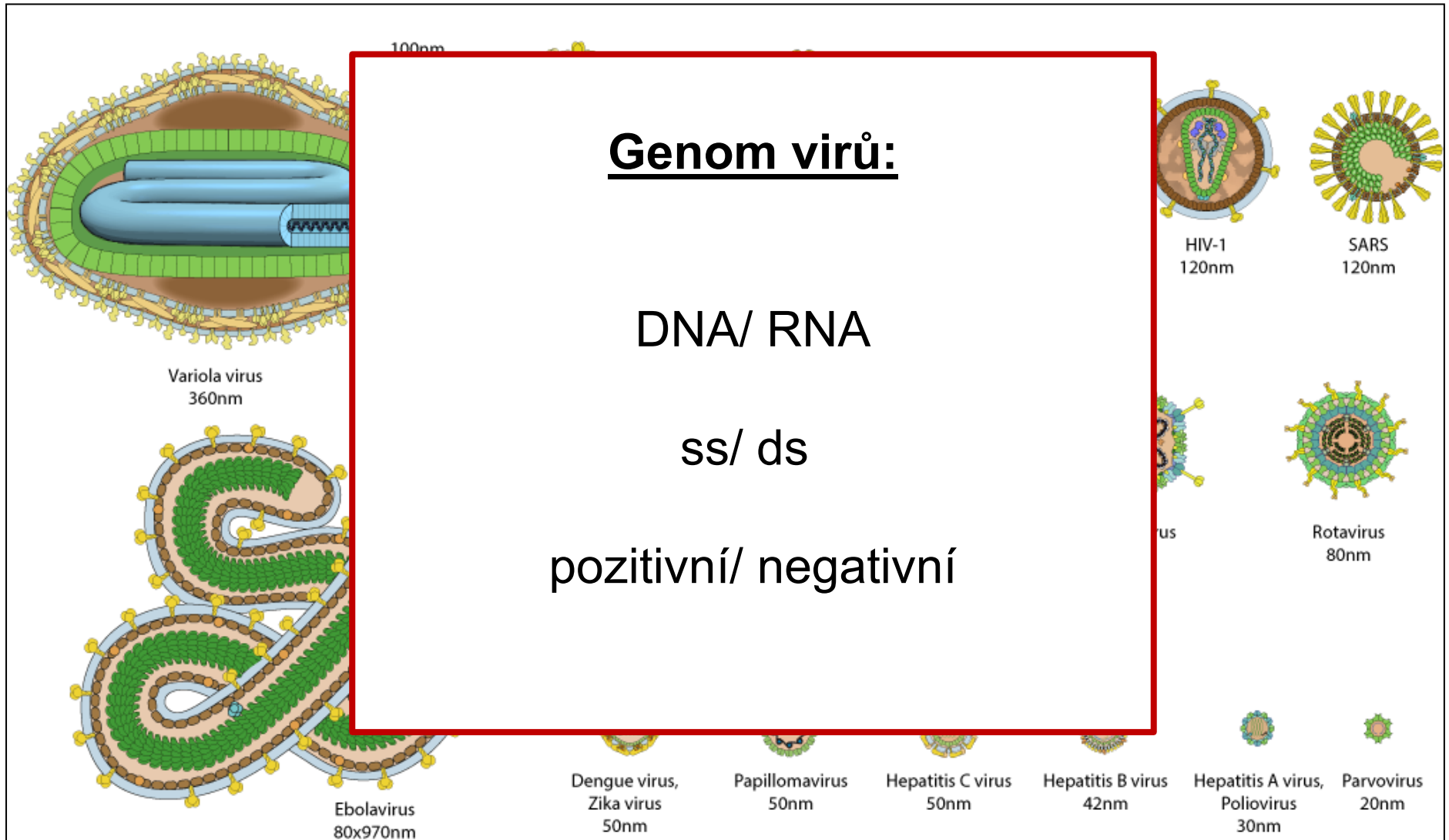


Mgr. Lenka Kramná
doc.MUDr. Ondřej Cinek PhD.

Laboratoř molekulární genetiky, Pediatrická klinika a Ústav lékařské mikrobiologie, FN Motol a 2. lékařská fakulta Univerzity Karlovy v Praze



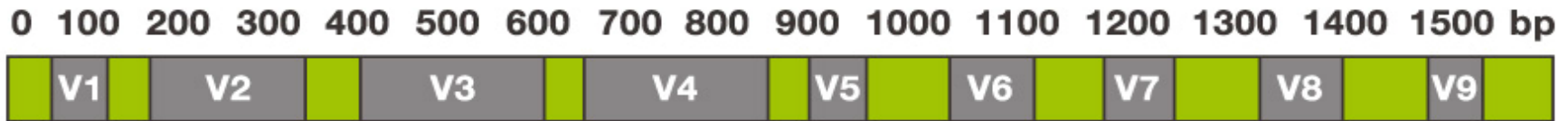
Sekvenování viromu je složité



Sekvenování viromu je složité

Bakterie:

Ize odlišit pomocí vybrané variabilní oblasti genu pro 16S



Viry:

neexistuje podobná oblast:

- u všech virů
- dostatečně variabilní

Sekvenace náhodných fragmentů (shotgun)

První užití metody

- Prof. Eric Delwart

JOURNAL OF VIROLOGY, May 2009, p. 4631–4641
0022-538X/09/\$08.00+0 doi:10.1128/JVI.02085-08
Copyright © 2009, American Society for Microbiology. All Rights Reserved.

Cardiov

Journal of General Virology (2010), 91, 74–86

Olga Blinkov
Sha

JOURNAL OF VIROLOGY, Jan. 2008, p. 311–320
0022-538X/08/\$08.00+0 doi:10.1128/JVI.01240-07
Copyright © 2008, American Society for Microbiology. All Rights Reserved.

J Infect Dis. 2009 January 15; 199(2): 196–200. doi:10.1086/598888

- nové viry u nemocných
- nové viry u zvířat
- ze vzorku stolice:
 - místo primární infekce
 - pokud infekce, vysoká kvantita virů

OPEN ACCESS Freely available online

 **PLOS** | PATHOGENS

Pearls

A Roadmap to the Human Virome

Eric Delwart^{1,2*}

Proč jsme to začali dělat:



Enterovirus: jedna z příčin Diabetu 1. typu?

1. Vzorek stolice:

- všechny enterovirové kmeny
- podívat se také do oblastí mimo VP1 (celý genom)
- na nalezené konkrétní genotypy navrhnout primery

2. Vzorek krve:

- nalézt tento konkrétní virus
- porovnat, zda odpovídá období nástupu diabetu u pacienta

Proč jsme to začali dělat:



Enterovirus: jedna z příčin Diabetu 1. typu?

Supernatant stolice:

- Pokud infekce, velmi vysoká virová nálož.

Zdravé děti:

Enterovirus: 97mil. kopií/ ul

Adenovirus: 2,5mil. kopií/ ul

BKV v moči: 10^{12} kopií/ ul

Lidských buněk: 0

Stolice:

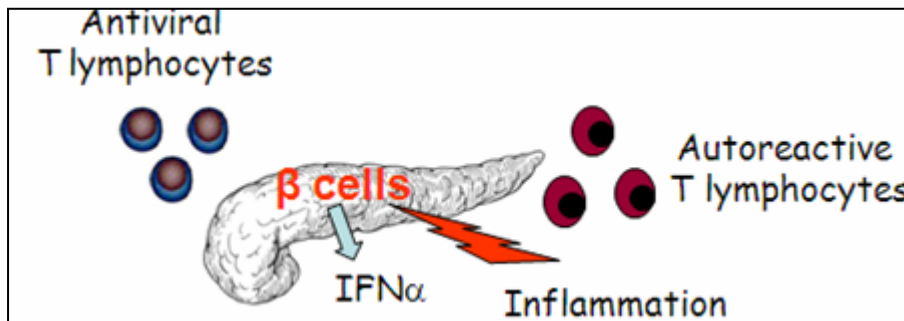
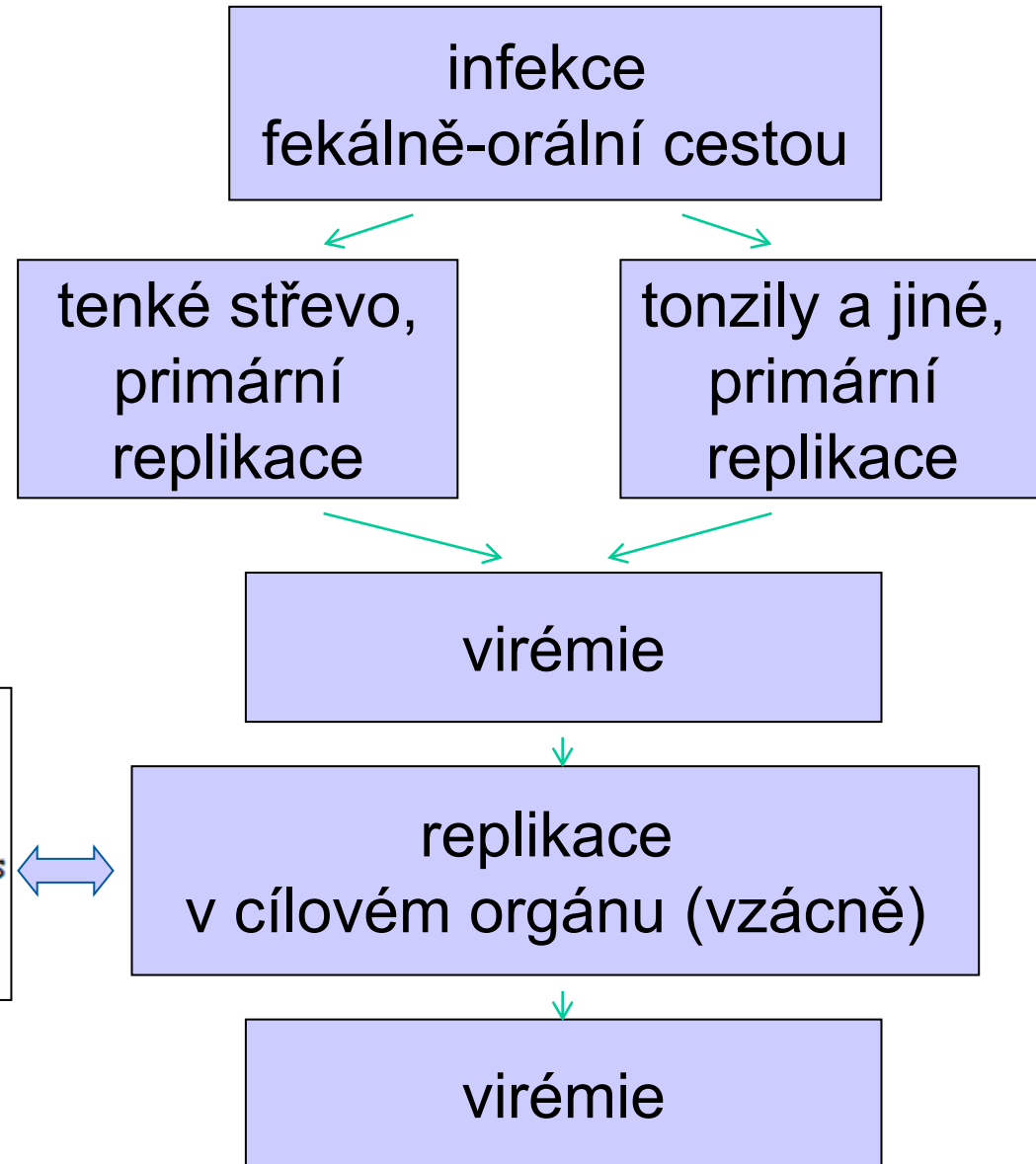
- 1/3 materiálu jsou lidské buňky

Krev:

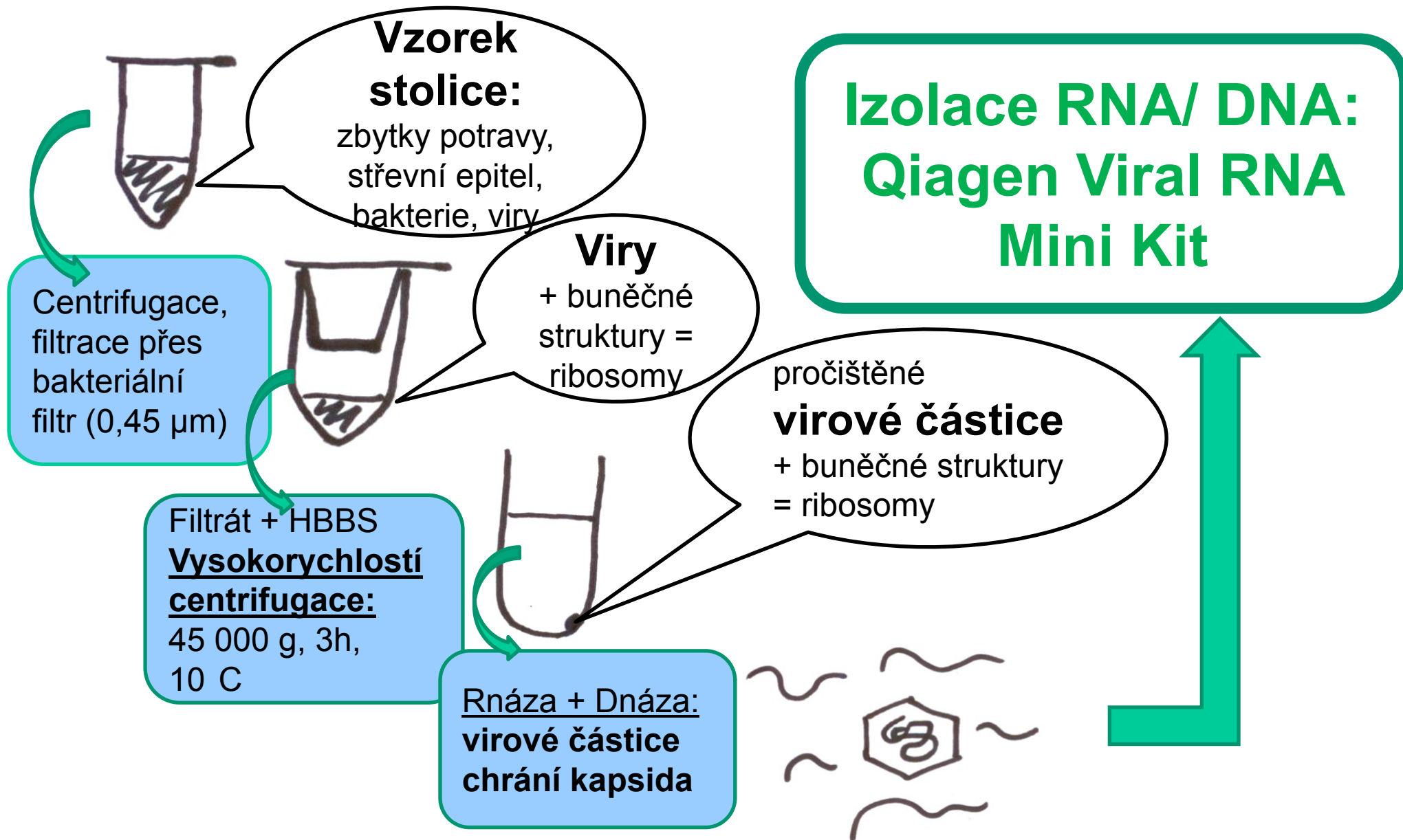
- Pokud infekce, nízká virová nálož (do 100 kopií/ ul = a to už je problém)
- Vysoká kontaminace lidskou DNA

Proč jsme to začali dělat:

Enterovirus: jedna z příčin Diabetu 1. typu?



Jak to probíhá v našich rukách



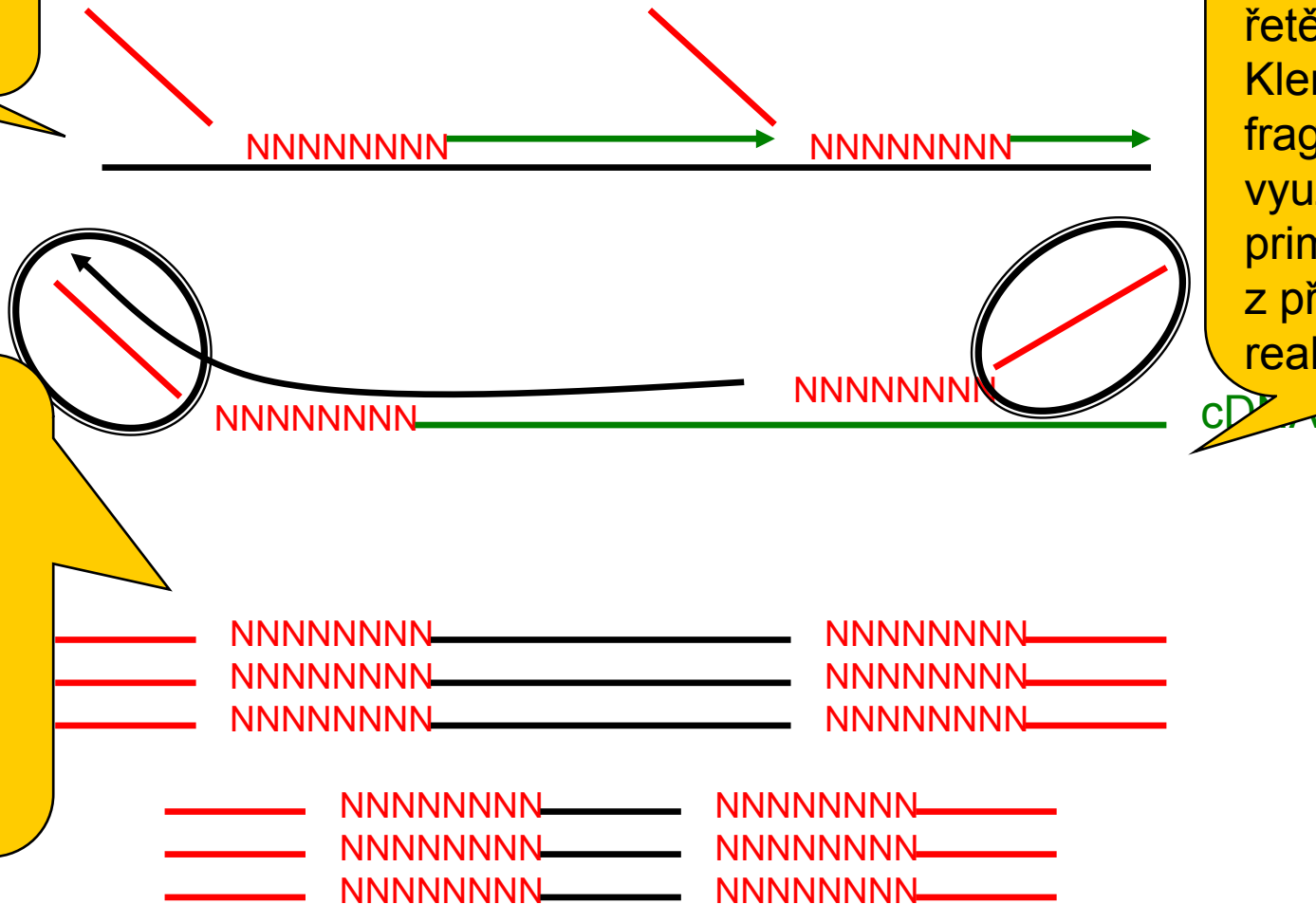
Jak to probíhá v našich rukách

Reverzní transkripce:
náhodné primery s tagem

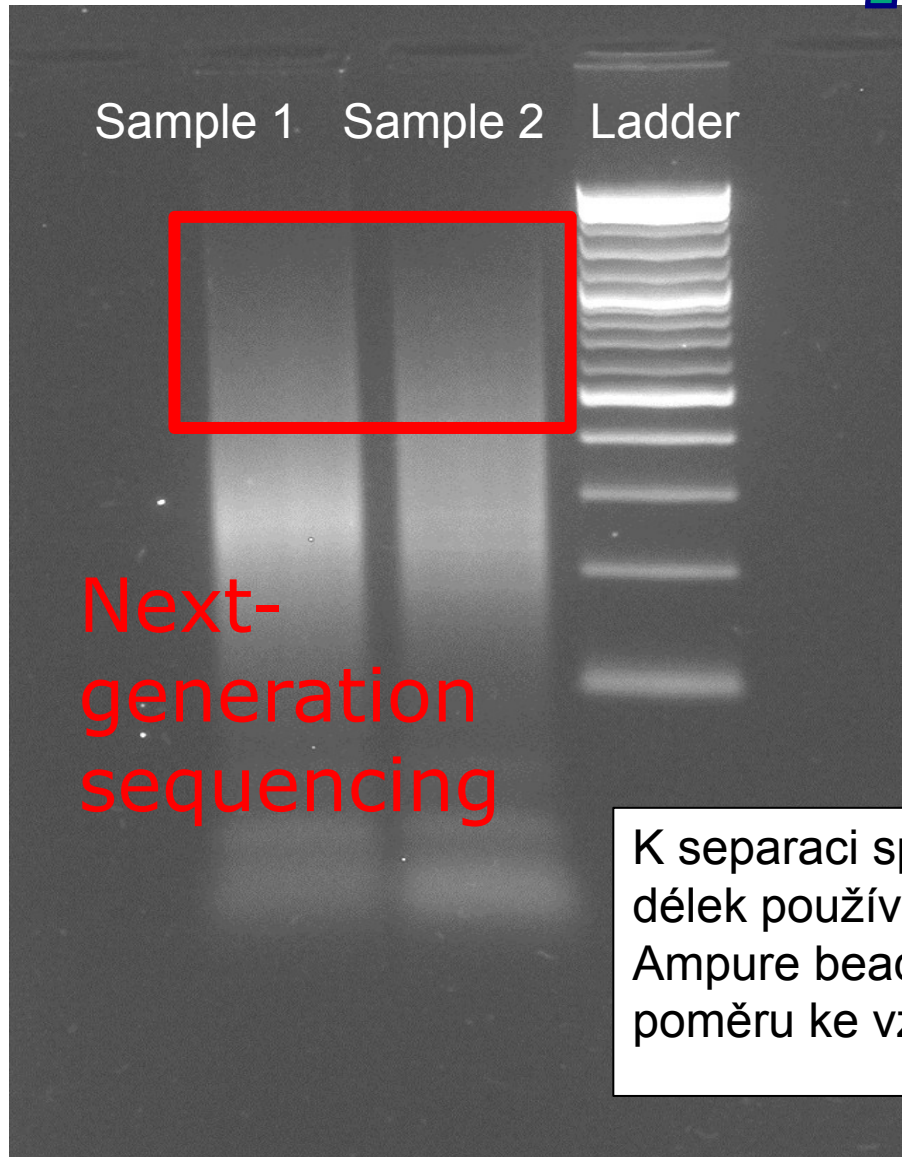
Amplifikace:
tagy jsou teď použity jako místo pro primer

(předtím přečištění EXO+SAP)

Syntéza druhého řetězce pomocí Klenowova fragmentu:
využije náhodné primery s tagem z předchozí reakce



Produkt k sekvenování



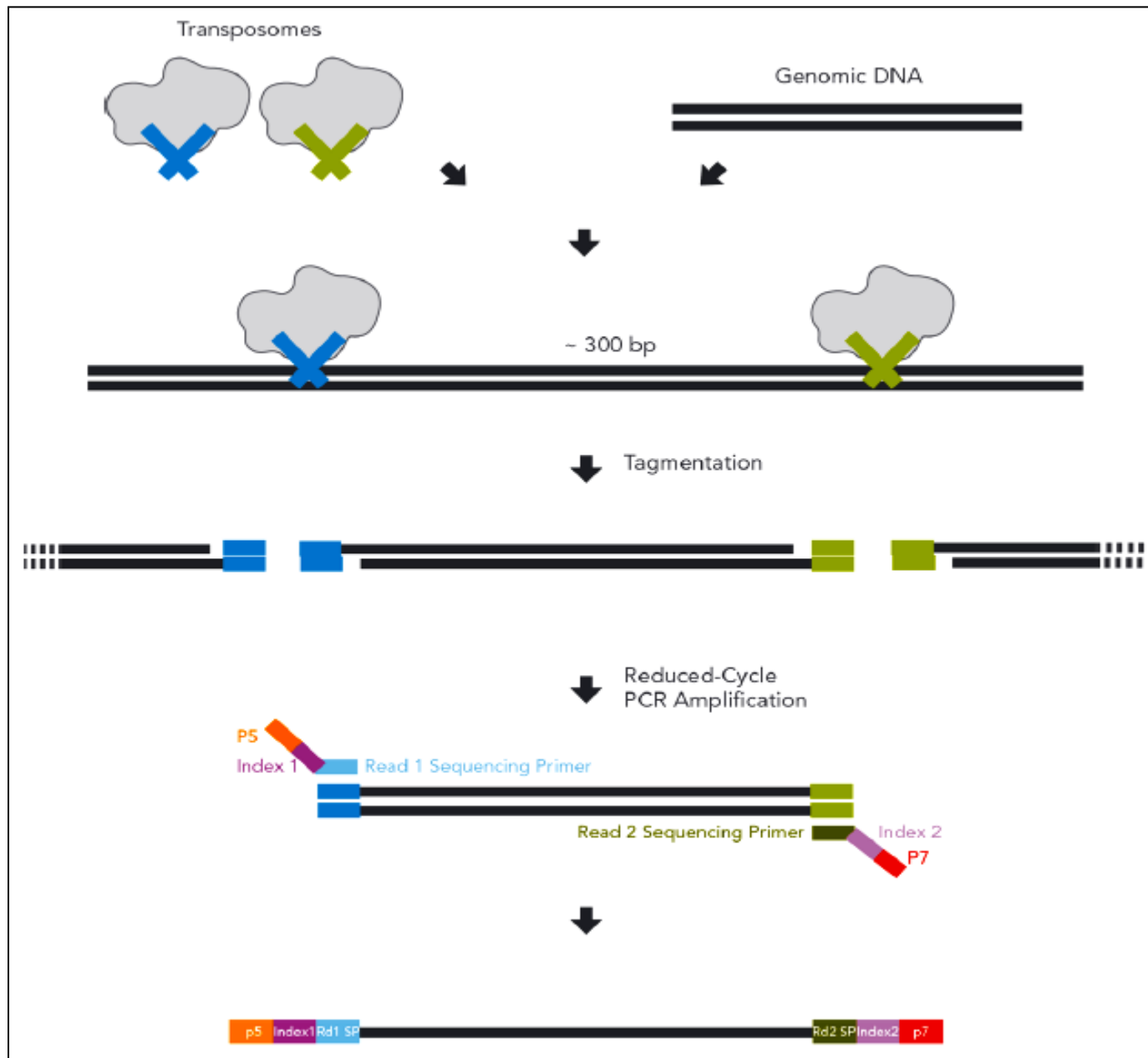
Next-generation sequencing

Nextera XT
MiSeq Illumina
(2X250)
24 vzorků / run
přibl. 15 mil.
sekvencí na run

K separaci správných délek používáme Ampure beads v poměru ke vzorku 2:1



Nextera XT DNA



- pouze 1ng vstupní DNA
- adaptory připojené pomocí enzymu, z každé strany uštipne 50nt, plus 12 cykl. PCR

Velvet: de-novo assembly



Konsenzus



Identifikace pomocí
BLASTu: omezení na celé
virové genomy

Analýza

Netranslatovaná oblast

Oblast zvětšení

Celý kontig

pokrytí



Analýza

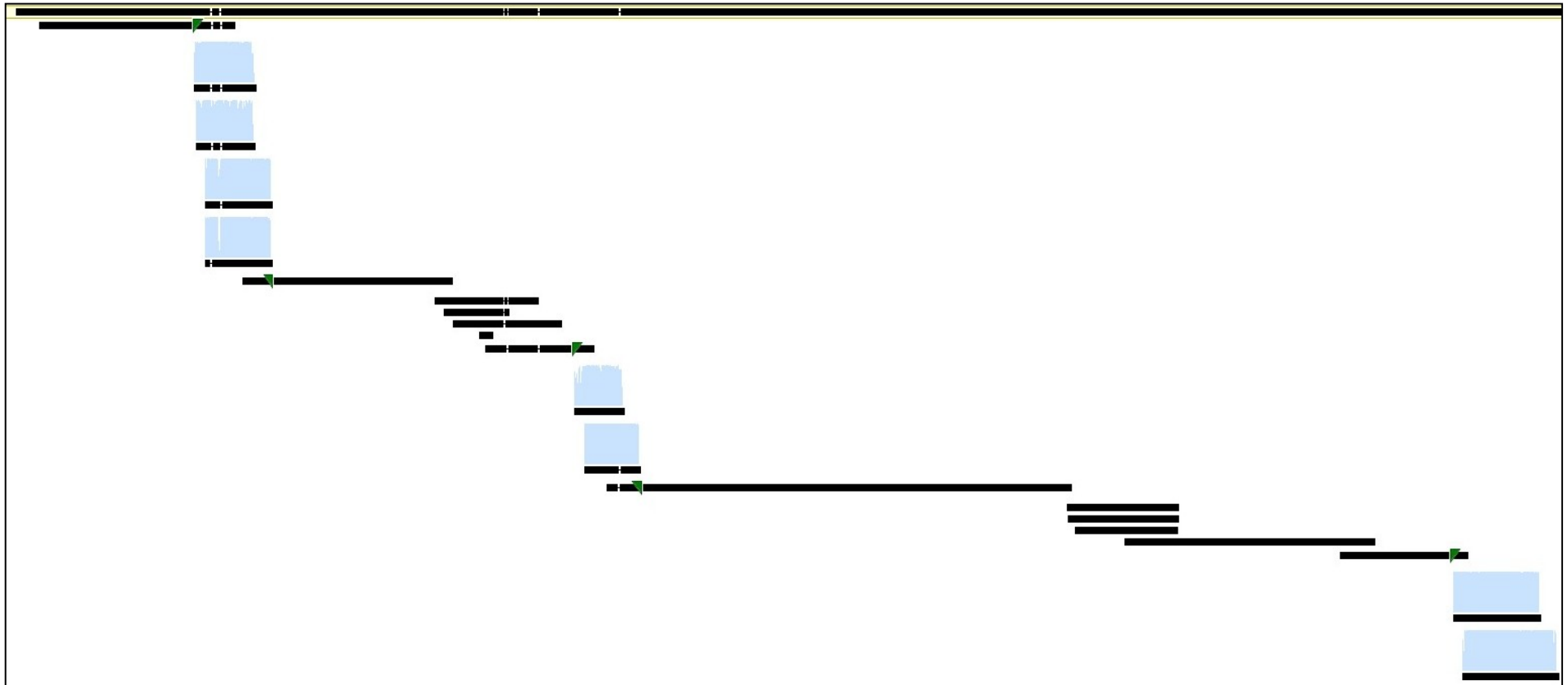


Kvazispecies:
evoluce



Získání celých virových genomů

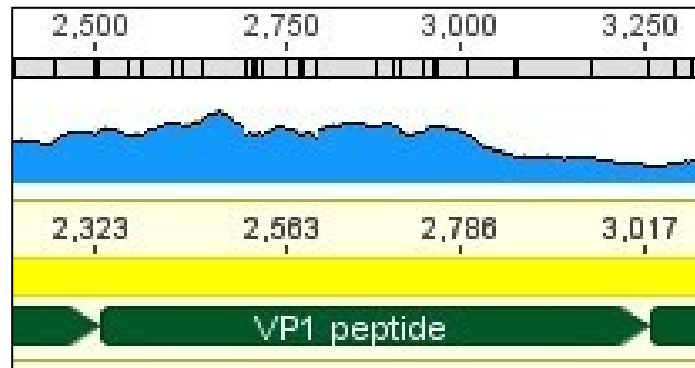
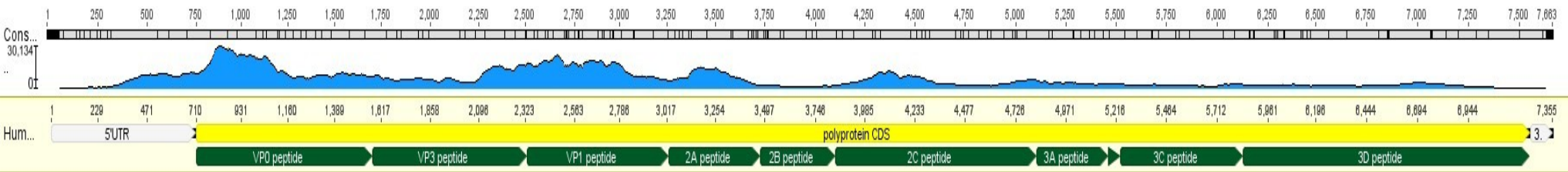
1. Assembly kontigů na referenční sekvenci
2. Návrh primerů
3. Dosekvenování mezer



Typizace picornavirů

Lidský parechovirus:

Human parechovirus 1, 40 000/ 100 000 reads

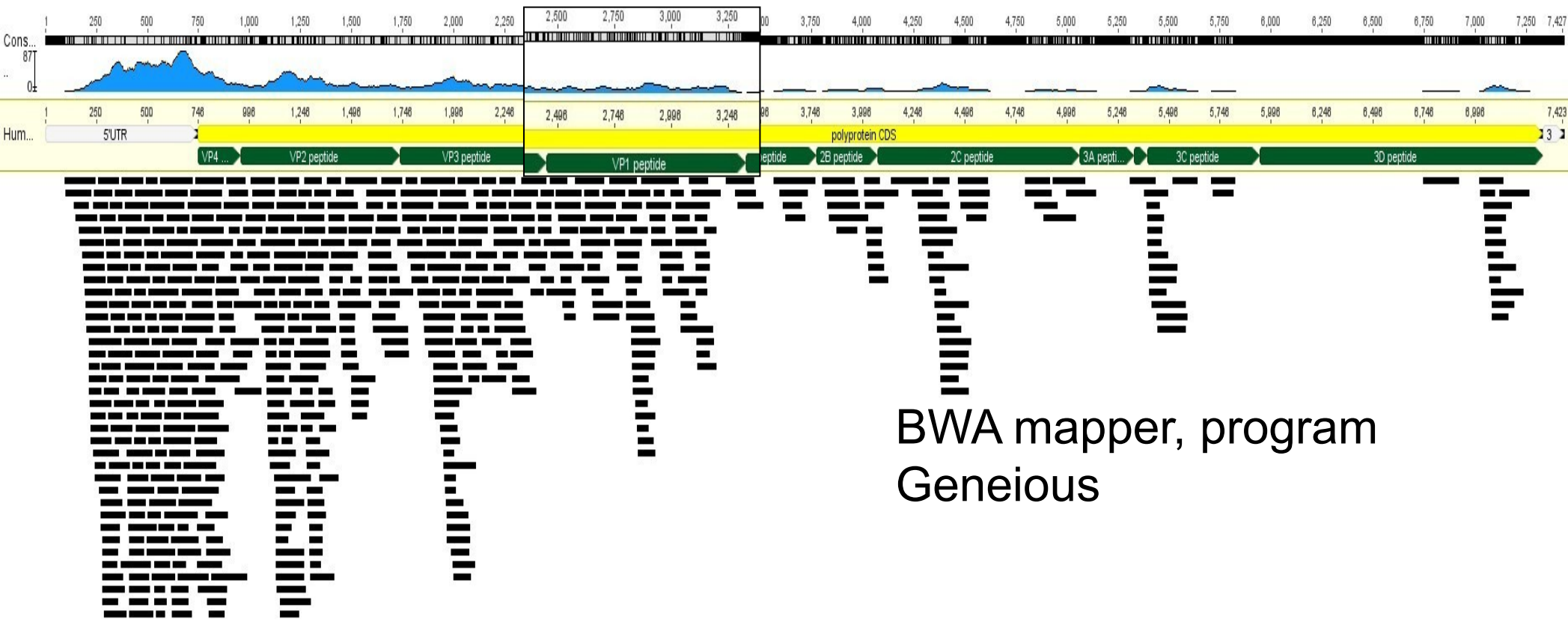


BWA mapper, program
Geneious

Typizace picornavirů

Lidský enterovirus:

Echovirus 25, 124/ 100 000 reads



Následná analýza

translace

oprava chyb assembly

The image displays a bioinformatics software interface for sequence analysis. The main window shows a DNA sequence alignment for 'geneious2' across five segments (1-110, 120-220, 230-330, 340-450, 460-565). For each segment, three frames (Frame 1, Frame 2, Frame 3) are shown. The sequence is color-coded by nucleotide (A, C, G, T). Below the sequence, three orange arrows indicate the reading frames. A right-hand panel titled 'Display' contains controls for 'Nucleotides' (checked), 'Complement' (unchecked), 'Translation' (checked), and 'Translation Options' (Frame: Forward frames, Genetic Code: Standard, Colors: ARND-). The interface also includes a vertical toolbar on the far right with icons for home, zoom, and other functions.

Následná analýza

translace

oprava chyb assembly

identifikace změn v AMK

```
uman coxsackievirus MSDPPPQVSVPFMSPASAYQWFYDGYPTFDDRPPQTSNRPYGGCPNNMLGTFAVRIVSKTPAERDLRVRVYMKLKHVRAWVPRPIRSQPYVLKNYPNYDGTQIVPSAKDRE
onsensus .....I.....N...C...D

uman coxsackievirus DIKNTGKFGQQSGAIYVGNRYRVVNRHLATHNDWTNLVWEDSTRDLLVSSTTAQGC DTIARCNCQTGVYYCNSRRKHYPVSFSPKPSLIFVEASEYYPARYQSHLMLAEGHS
onsensus .....F.....A..I...A.....K.....V.....

uman coxsackievirus EPGDCGGILRCQHGVVGIIVSTGGNGLVGFADV RDLLWLDEEAMEQGVSDYIKGLGDAFGTGFTDAVSREVEALKNHLLIGSEGAVDKILKNLVKLISALVIVIRSDYDMVT
onsensus .....E.....

uman coxsackievirus LTATLALIGCHGSPWAWIKSKTASILGIPMAQKQSASWLKKFNDMANAAGLEWISNKISKPIDWLKEKIIPAAKEKVEFLNNLKLPLLENQISNLEQSAASQDDLEVM
onsensus .....L.....S.....R.....E...A.

uman coxsackievirus FGNVSYLAHFCRKFQPLYATEAKRVYALEKRINNYMQFKSKHRIEPCLIIRGSPGTGKSLATGIARAIADKYHSSVYSLPPDPDHFHDGYKQQVVTVMDDLCQNPDKD
onsensus .....T...M.....
```

U čeho jsme nakonec skončili: Pilotní studie

930

Diabetes Care Volume 38, May 2015



CrossMark

Gut Virome Sequencing in Children With Early Islet Autoimmunity

Diabetes Care 2015;38:930–933 | DOI: 10.2337/dc14-2490

*Lenka Kramná,¹ Kateřina Kolářová,¹
Sami Oikarinen,² Juha-Pekka Pursiheimo,³
Jorma Ilonen,⁴ Olli Simell,⁵
Mikael Knip,^{6,7,8,9} Riitta Veijola,¹⁰
Heikki Hyöty,^{2,11} and Ondrej Cinek¹*

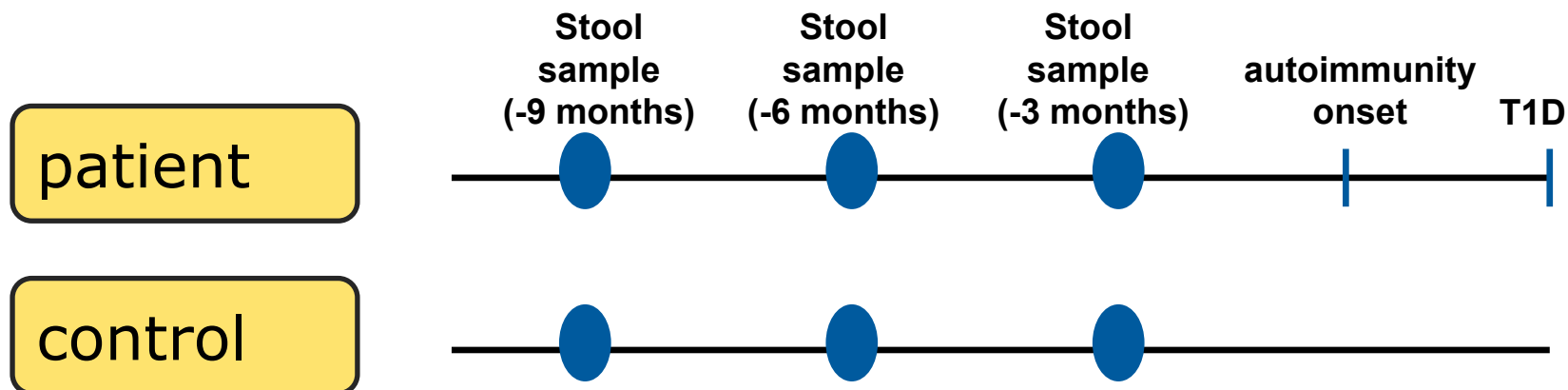
Pilotní studie

Finská DIPP (Diabetes Prediction and Prevention)
kohorta

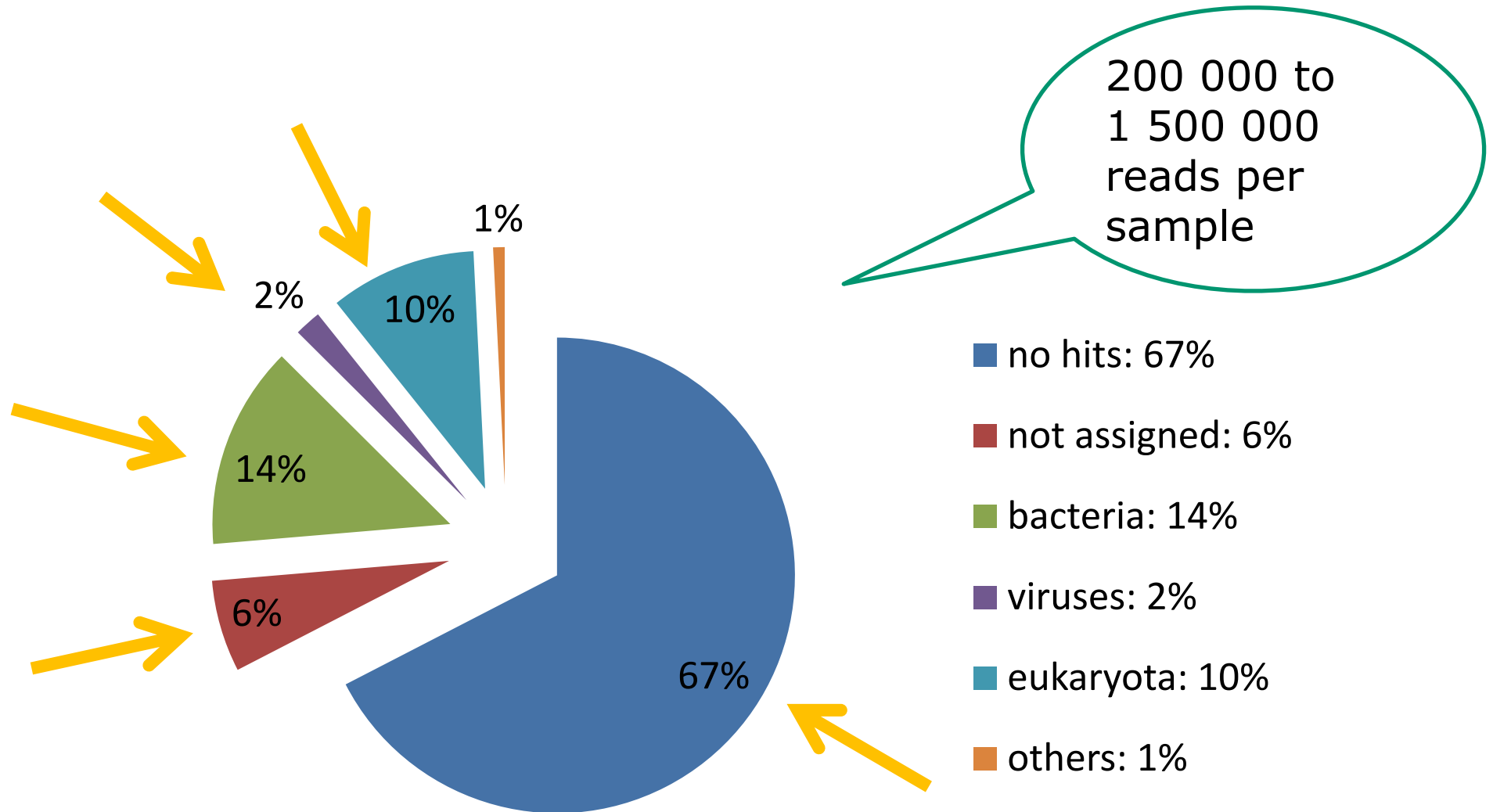
19 pacientů: následně T1D

19 kontrol: stejné genetické riziko, pohlaví, datum a
místo narození

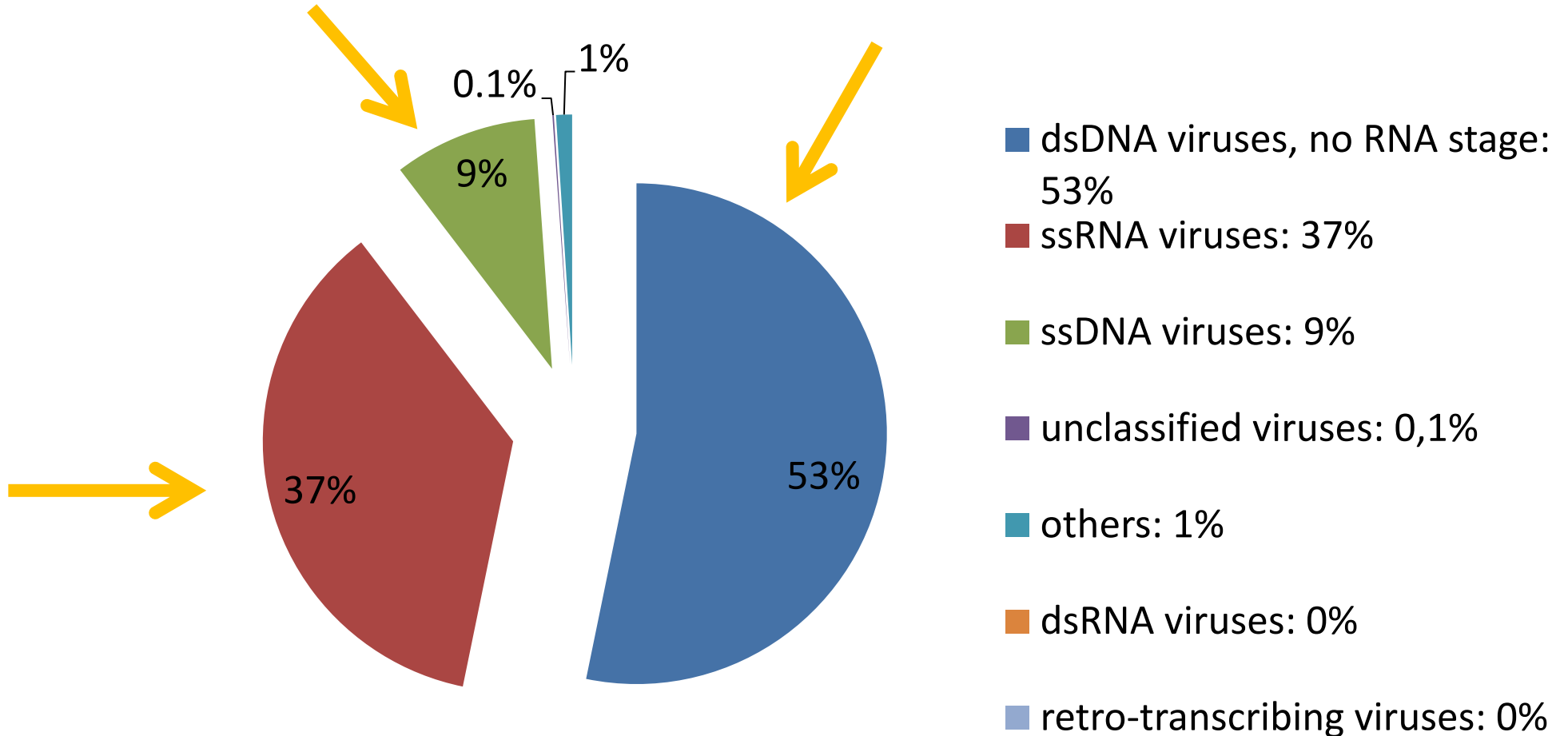
Celkem 96 vzorků stolic



Obsah sekvencí: MEGAN5



Obsah virových sekvencí



Výsledky

NGS: lidské viry

Virus	Případy (n=48)	Kontroly (n=48)	Asociace: OR [95% CI]
Enterovirus			
Parechovirus			
Bocavirus			
Sapovirus			
Anellovirus			

Lidské viry

- **Parvovirus**

Matching pair ID	-3months	
	case	control
1.	-	-
2.	-	-
3.	-	-
4.	-	-
5.	-	-
6.	10	-
7.	703	-
8.	1	19
9.	7	-
10.	-	-
11.	1	-
12.	9924	3
13.	-	1
14.	-	-
15.	-	-
16.	-	-
17.	1	-
18.	-	-
19.	-	-

Lidské viry

- **Parechovirus**

Matching pair ID	-6 months		-3 months	
	case	control	case	control
1.	5	-	-	-
2.	-	-	-	-
3.	-	-	-	-
4.	-	-	-	-
5.			-	-
6.	-	-	10	-
7.			703	-
8.	3041	3	1	19
9.	-	49	7	-
10.			-	-
11.	6	-	1	-
12.	8	-	39924	3
13.	-	2291	-	1
14.	-	-	-	-
15.	-	-	-	-
16.	-	-	-	-
17.	-	11	1	-
18.	1	-	-	-
19.	-	1	-	-

Lidské viry

- Parechovirus

infekce ve střevě
probíhají bouřlivě,
ale krátce

		-9 months		-6 months		-3 months	
Matching pair ID		case	control	case	control	case	control
1.				5	-	-	-
2.		-	-	-	-	-	-
3.		4	-	-	-	-	-
4.		-	-	-	-	-	-
5.						-	-
6.		1	1	-	-	10	-
7.						703	-
8.		10	7	3041	3	1	19
9.		-	1	-	49	7	-
10.						-	-
11.		-	1670	6	-	1	-
12.		-	2	8	-	39924	3
13.		3	-	-	2291	-	1
14.		-	2	-	-	-	-
15.		-	-	-	-	-	-
16.		-	-	-	-	-	-
17.		-	-	-	11	1	-
18.				1	-	-	-
19.				-	1	-	-

Výsledky: potvrzení výsledků NGS pomocí PCR

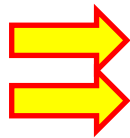
Virus	Případy (n=48)	Kontroly (n=48)	Asociace: OR [CI 95%]
Enterovirus			
Parechovirus			
Bocavirus			
Sapovirus			

Náš NGS protokol
detekuje viry v kvantitě
100 kopií / μ l
vyizolované NK a vyšší

Bakteriofágy

E. coli fág

fágy ve střevě jsou téměř výhradně temperátní a přetrvávají dlouho



Matching pair ID	-9 months		-6 months		-3 months	
	case	control	case	control	case	control
1.			-	-	3	-
2.	1	-	-	-	-	-
3.	-	-	2	-	-	-
4.	-	-	-	-	-	-
5.					-	-
6.	-	-	456	-	1	-
7.					-	1
8.	-	5	-	-	-	-
9.	-	-	-	-	-	-
10.					4	-
11.	-	249	-	1	-	86
12.	-	-	2	-	-	-
13.	2613	-	4152	-	358	-
14.	5365	12	1268	-	23935	2
15.	1	3	1	-	5	-
16.	-	-	-	-	73	-
17.	4	-	-	-	77	-
18.			-	-	-	36
19.			-	-	-	-

Co jsme si z toho odnesli

Bakterie a fágy ve střevě:

Střevo = největší imunitní orgán...

V otevřených prostředích:

zejména lýtické fágy:

„kill the winner“

„red queen coevolution“

Střevo: polouzavřený systém,
nejvíce lysogenní fágy (temperátní)

Infekce jedním fágem působí
protektivně proti infekci jiným fágem.

Bakteriální imunitní systém:

(CRISPR-CAS = části
fágových genomů do
bakteriální DNA)

Nedokážeme rozlišit.

Phages and the Evolution of Bacterial Pathogens:
from Genomic Rearrangements to
Lysogenic Conversion

Harald Brüssow,^{1*} Carlos Canchaya,¹ and
Wolf-Dietrich Hardt^{2*}

*Nestlé, Research Center, Vers-chez-les-Blanc, Lausanne,¹
and Institute for Microbiology, D-BIOL,
ETH Zürich, Zürich,² Switzerland*

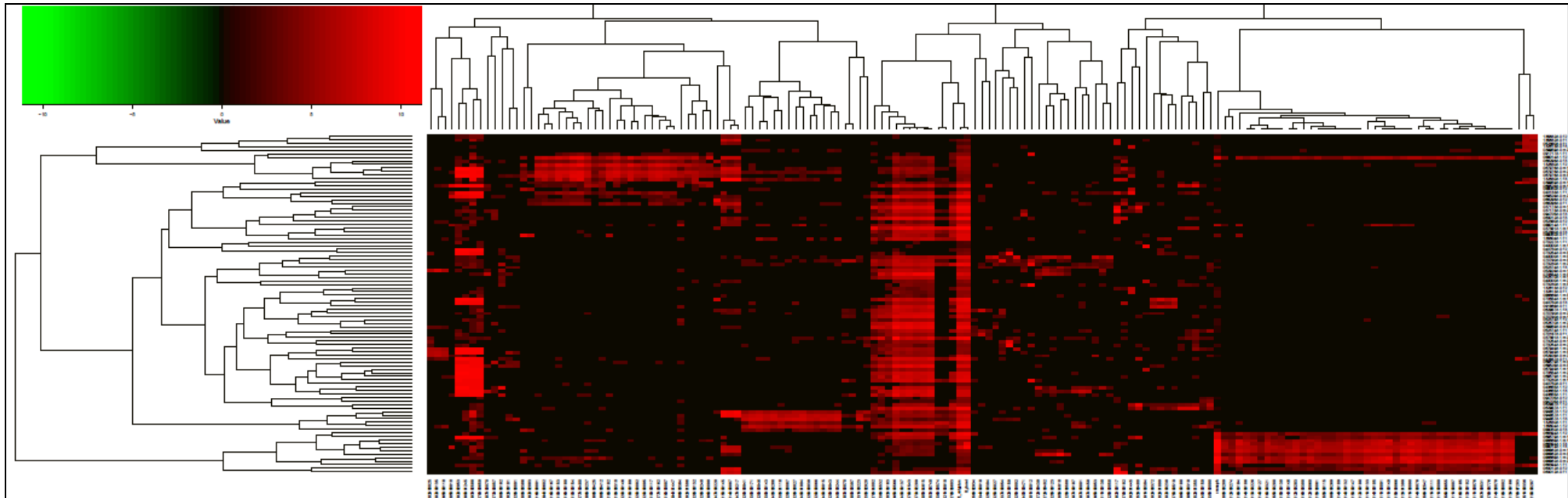
We currently have only about 200 complete phage genomes.

This number may seem large, but actually it is extremely small considering that phages outnumber bacteria in many environments by a factor of 10 and represent, with approximately 10³¹ tailed phage particles, numerically the largest share of biological material on Earth (240). Up to 10⁷ particles/ml were found in ocean water and sediment (30, 36, 176, 238, 240). Considering these vast numbers of phages and the number of bacteria coexisting in these niches, it has been estimated that 10²⁵ phage infections are initiated every second worldwide (177); this has probably occurred for the last 3 billion years.

Co jsme si z toho odnesli

Velký podíl neznámých sekvencí:

- ne všechny je důležité identifikovat
- vybrat jen ty, které jsou významně zvýšeny u případů oproti kontrolám



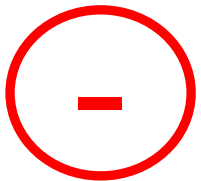
Problémy vs. výhody



Umožňuje identifikovat viry „nezkresleně“ (bez předchozích znalostí, co očekávat)

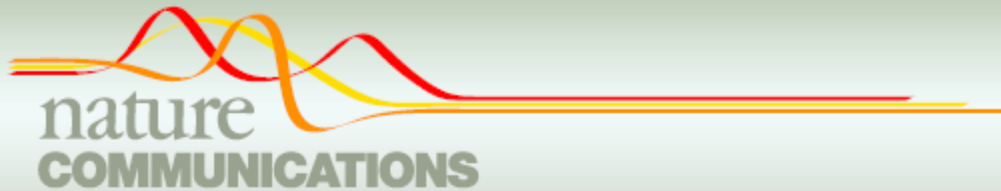
Bez předchozí izolace viru na buněčných kulturách

Kromě virových sekvencí je přítomna i bakteriální 16S



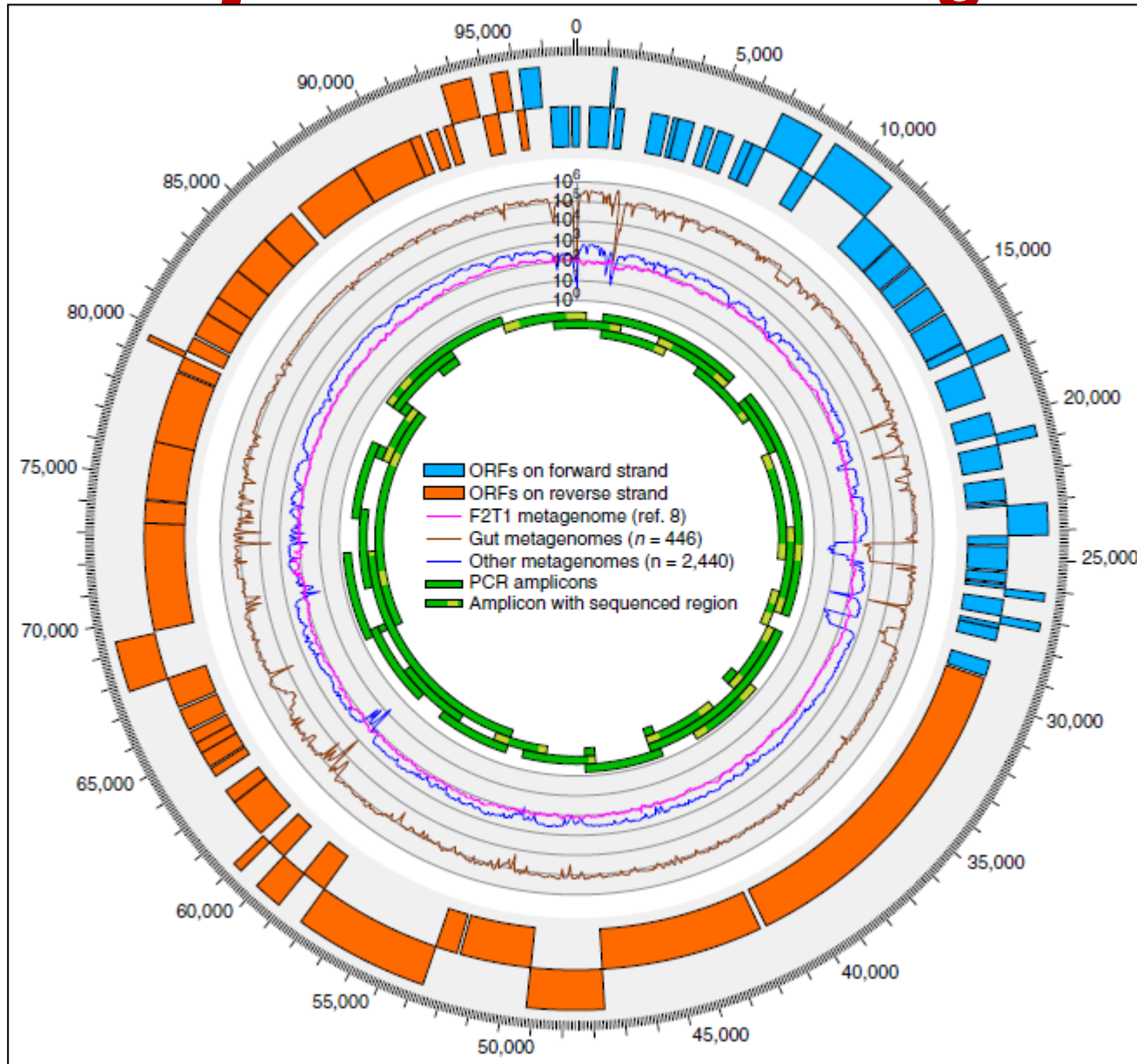
Méně citlivá než specifická PCR, zatím nevhodná pro klinické účely

Zajímavé: CrAssPhage



Metagenomics, or sequencing of the genetic material from a complete microbial community, is a promising tool to discover novel microbes and viruses. Viral metagenomes typically contain many unknown sequences. Here we describe the discovery of a previously unidentified bacteriophage present in the majority of published human faecal metagenomes, which we refer to as crAssphage. Its ~97 kbp genome is six times more abundant in publicly available metagenomes than all other known phages together; it comprises up to 90% and 22% of all reads in virus-like particle (VLP)-derived metagenomes and total community metagenomes, respectively; and it totals 1.68% of all human faecal metagenomic sequencing reads in the public databases. The majority of crAssphage-encoded proteins match no known sequences in the database, which is why it was not detected before. Using a new co-occurrence profiling approach, we predict a *Bacteroides* host for this phage, consistent with *Bacteroides*-related protein homologues and a unique carbohydrate-binding domain encoded in the phage genome.

Zajímavé: CrAssPhage



Praktické využití

- Neznámá epidemie průjmů u dětí v Motole
 - na oddělení XY častý výskyt průjmů neznámého původu
 - v létě, horko, otevřené okno, na parapety sedali holuby
 - po instalaci bodců holuby přestali létat a průjmy ustaly..
 - PCR negativní na všechny známé viry. V elektronovém mikroskopu částice podobné Calicivirům

Praktické využití

- Identifikace původu adenovirových infekcí v Motole
 - velmi podobné genomy = získání v nemocnici
 - rozdílné genomy = přinesli si z domu
 - Sangerové sekvenování (jen VP1), není schopno odhalit

Děkuji za pozornost 😊

