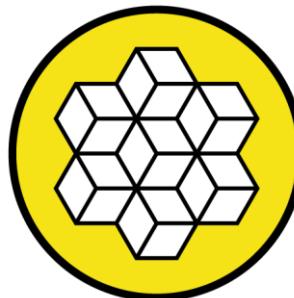


Sekvenace COVID-19 v BIOCEVu

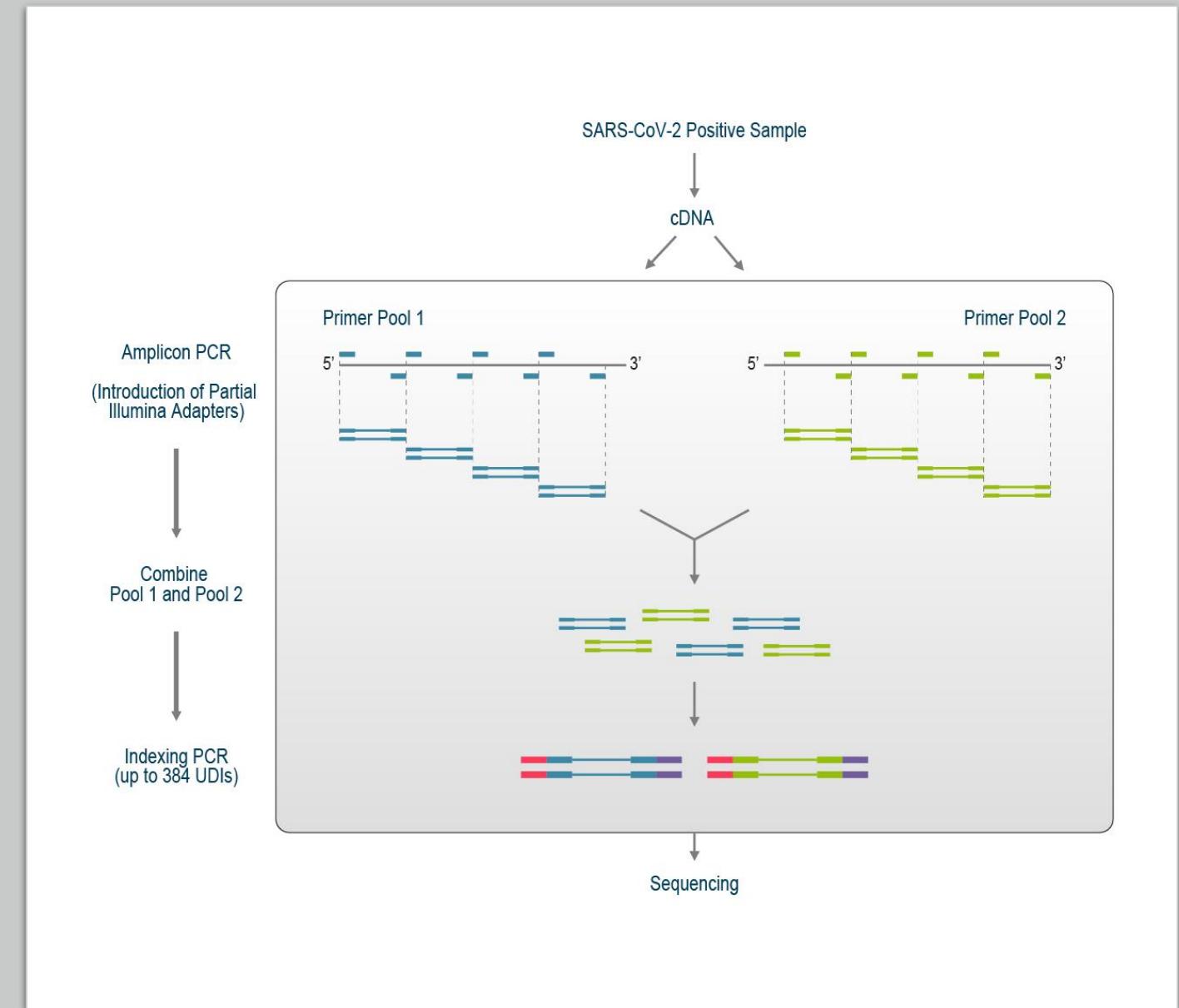
Ingrid Poláková, Štěpánka Hrdá, Blanka Hamplová, Zoltán Füssy,
Sebastian Treitli, Ruth Tachezy, Vladimír Hampl



BIOCEV

Princip celogenomového sekvenování COVID-19

- příklad Artic-v3



Izolace RNA

- Izolace RNA z primárních stěrů
 - izolace v SZÚ – různé kity (MagMAX Viral/Pathogen Nucleic Acid Isolation Kit, MagMAX CORE Nucleic Acid Purification Kit)
 - 96 % sekvenací OK (selhávaly vzorky s Ct nad 35)
 - izolace Biocev
 - kit Zybio Nucleic Acid Extraction Kit - sekvenace OK
 - kit Diana Nucleic Acid Extraction Kit - sekvenace KO
- Izolace RNA ze slin
 - izolace Zybio
 - Sekvenace OK pouze u jednoho ze tří vzorků (střední Ct)
 - bez izolace, pouze ošetřeno proteinázou K
 - Sekvenace OK pouze u jednoho ze tří vzorků (nízké Ct)

Zybio EXM 3000 - Nucleic Acid Isolation system

- extrakce 32 vzorků za 9 min



Features

Throughput	1-32
Process volume	30 µL-1000 µL
Recovery rate	≥98%
Stability	CV≤3%
Pollution control	UV sterilization
Filtration	HEPA Filter
Dimensions(L*W*H)	375mm*415mm*440mm
Weight	27kg

Zybio - Nucleic Acid Extraction Kit (Magnetic Bead Method); B200



2. Operation of Semi-Automatic Nucleic Acid Extraction System
2.1 For B-100/B-200:



1
Proteinase k 96-well plates

Add 15µL [Proteinase K] (B-100: 10µL) to the position A1~H1 and A7~H7 in order



2
Sample 96-well plates

Add 200 µL sample (B-100: 100µL) in order

Table 4 Running Program Setting

No.	Position	Name	Waiting Time (min)	Mixing Time (min)	Absorption Magnetic Beads Time(sec)	Mixture Velocity	Volume	Temperature State	Temperature (°C)
1	2	Move	0	0	30	Slow	150	Closed	0
2	1	Lysis	0	4	60	Slow	500	Heating for Lysis	55
3	3	Wash	0	1	60	Slow	600	Closed	0
4	6	Elution	0	2	30	Slow	50	Heating for Elution	80
5	1	Move	0	0	0	Slow	300	Closed	0

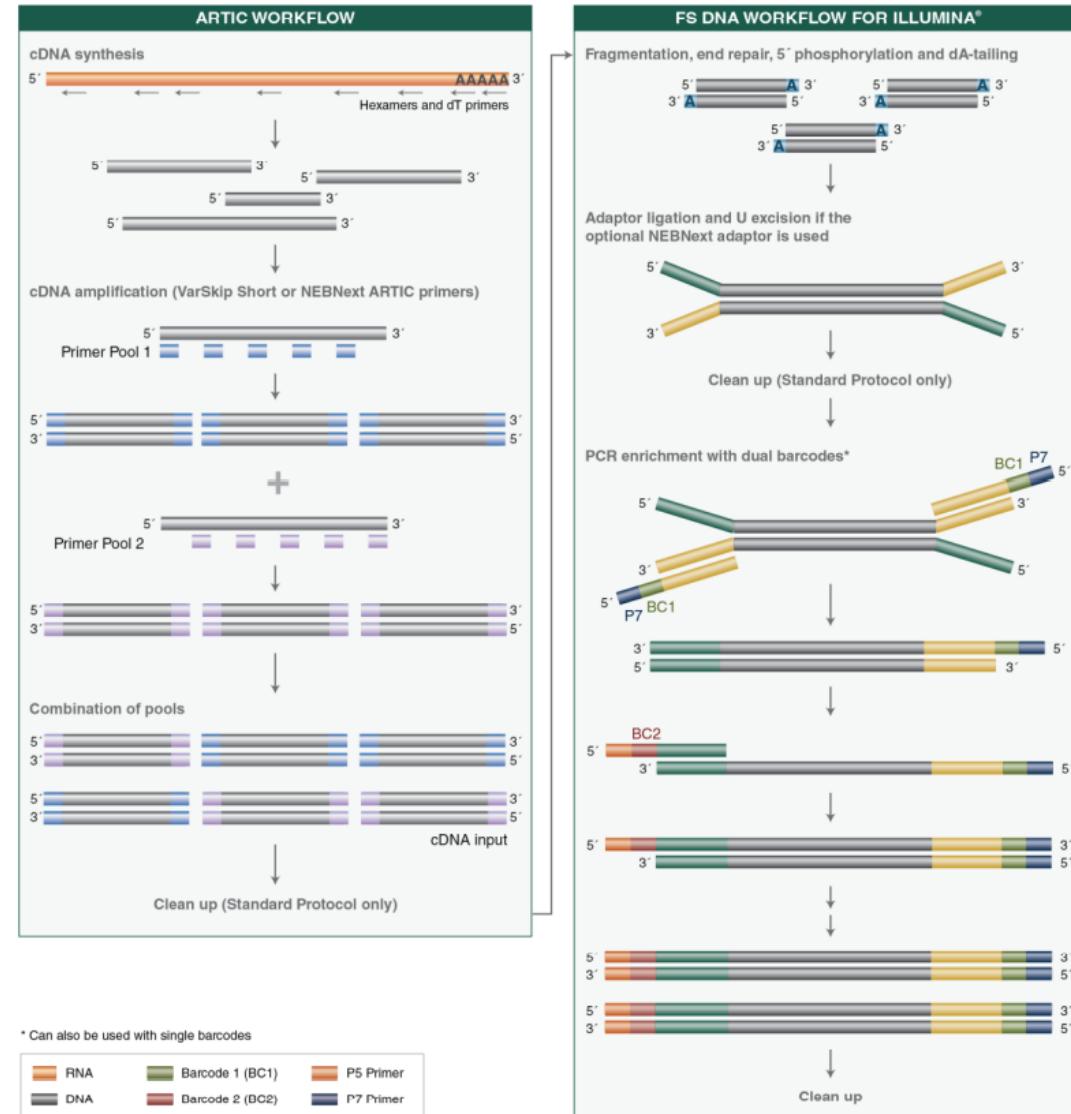
- izolovaná RNA – cca. 60-80 µl

Vzorky

- ideálně příjem inaktivovaných vzorků
- uložení vzorků až do doby úspěšně provedené sekvenace
- dlouhodobé skladování izolované RNA

Příprava knihoven v Biocevu

NEBNext® ARTIC SARS-CoV-2 FS Library Prep Kit (Illumina®)



NEBNext ARTIC Standard Protocol with cDNA Amplicon and Ligation Bead Cleanups (Three clean-up steps) - one plate/96 samples

cDNA Synthesis

RNA Sample	8 µl
LunaScript RT SuperMix reagent	2 µl
<i>Incubate reactions in a thermal cycler</i>	23 min
<i>Hands on time</i>	45 min

cDNA Amplification

Prepare two pools A and B	
cDNA	4.5 µl
Q5 Hot Start High-Fidelity 2X Master Mix	6.25 µl
ARTIC SARS-CoV-2 Primer Mix (A or B)	1.75 µl
<i>Incubate reactions in a thermal cycler</i>	3 hod 5 min
<i>Hands on time</i>	75 min

Combine the Pool A and Pool B PCR reactions for each sample.

Cleanup of cDNA Amplicons (560 bp)

Cleaning with magnetic beads with 0,8x ratio	
<i>Hands on time</i>	120 min

Fragmentation/End Prep (120 bp)

Pooled cDNA amplicons	13 µl
NEBNext Ultra II FS Reaction Buffer	3.5 µl
NEBNext Ultra II FS Enzyme Mix	1 µ
<i>Incubate reactions in a thermal cycler</i>	60 min
<i>Hands on time</i>	30 min

Adaptor Ligation

FS Reaction Mixture	17.5 µl
NEBNext Adaptor for Illumina	1.25 µl
NEBNext Ultra II Ligation Master Mix	15 µl
<i>Incubate reactions in a thermal cycler</i>	15 min

Ligation mixture

USER® Enzyme	1.5 µl
<i>Incubate reactions in a thermal cycler</i>	15 min
<i>Hands on time</i>	30 min

Cleanup of Adaptor-ligated cDNA

Cleaning with magnetic beads with 0,8x ratio	
<i>Hands on time</i>	120 min

PCR Enrichment of Adaptor-ligated DNA

Adaptor Ligated DNA Fragments	7.5 µl
NEBNext Library PCR Master Mix	12.5 µl
Index Primer Mix	5 µl
<i>Incubate reactions in a thermal cycler</i>	15 min
<i>Hands on time</i>	40 min

Cleanup of PCR Reaction

Cleaning with magnetic beads with 0,9x ratio	
<i>Hands on time</i>	120 min

Pooling samples

<i>Hands on time</i>	20 min
----------------------	--------

NEBNext ARTIC Standard Protocol without cDNA Amplicon and Ligation Bead Cleanups (One clean-up step) - one plate/96 samples

cDNA Synthesis

RNA Sample	8 µl
LunaScript RT SuperMix reagent	2 µl
<i>Incubate reactions in a thermal cycler</i>	23 min
<i>Hands on time</i>	45 min

cDNA Amplification

Prepare two pools A and B	
cDNA	4.5 µl
Q5 Hot Start High-Fidelity 2X Master Mix	6.25 µl
ARTIC SARS-CoV-2 Primer Mix (A or B)	1.75 µl
<i>Incubate reactions in a thermal cycler</i>	3 hod 5 min
<i>Hands on time</i>	75 min

Combine the Pool A and Pool B PCR reactions for each sample.

Cleanup of cDNA Amplicons

Diluting of cDNA Amplicons (560 bp)	
Pooled cDNA amplicons	1.3 µl
0.1X TE	11.7 µl
<i>Hands on time</i>	120 10 min

Fragmentation/End Prep (120 bp)

Pooled diluted cDNA amplicons	13 µl
NEBNext Ultra II FS Reaction Buffer	3.5 µl
NEBNext Ultra II FS Enzyme Mix	1 µl
<i>Incubate reactions in a thermal cycler</i>	60 min
<i>Hands on time</i>	30 min

Adaptor Ligation

FS Reaction Mixture	17.5 µl
NEBNext Adaptor for Illumina	1.25 µl
NEBNext Ultra II Ligation Master Mix	15 µl
<i>Incubate reactions in a thermal cycler</i>	15 min

Ligation mixture

USER® Enzyme	1.5 µl
<i>Incubate reactions in a thermal cycler</i>	15 min
<i>Hands on time</i>	30 min

Cleanup of Adapter-ligated cDNA

<i>Hands on time</i>	120 min
----------------------	---------

PCR Enrichment of Adapter-ligated DNA

Adaptor Ligated DNA Fragments	7.5 µl
NEBNext Library PCR Master Mix	12.5 µl
Index Primer Mix	5 µl
<i>Incubate reactions in a thermal cycler</i>	15 min
<i>Hands on time</i>	40 min

Cleanup of PCR Reaction

<i>Hands on time</i>	120-min
----------------------	---------

Pooling samples + Cleanup of pool

Cleaning with magnetic beads with 0,7x ratio	
<i>Hands on time</i>	35 min

Sekvenování knihoven v Biocevu

Miseq Illumina

MiSeq Reagent Kit v3 (150-cycle) MS-102-3001

25 MIL. PAIR-END READS

Nanášeno: 13 pM pool

Dosaženo: více než 30 mil. čtení v jednom runu

Pool 96 – 106 vzorků

Průměrně 300 – 400 tisíc čtení na vzorek

Většinou dobrá sekvenace i při 100 tisíc čtení na vzorek

Jarní kolo sekvenací:

Zpracovali jsme 673 vzorků a z toho jsme zdáně osekvenovali 647, což je více než 96 %.

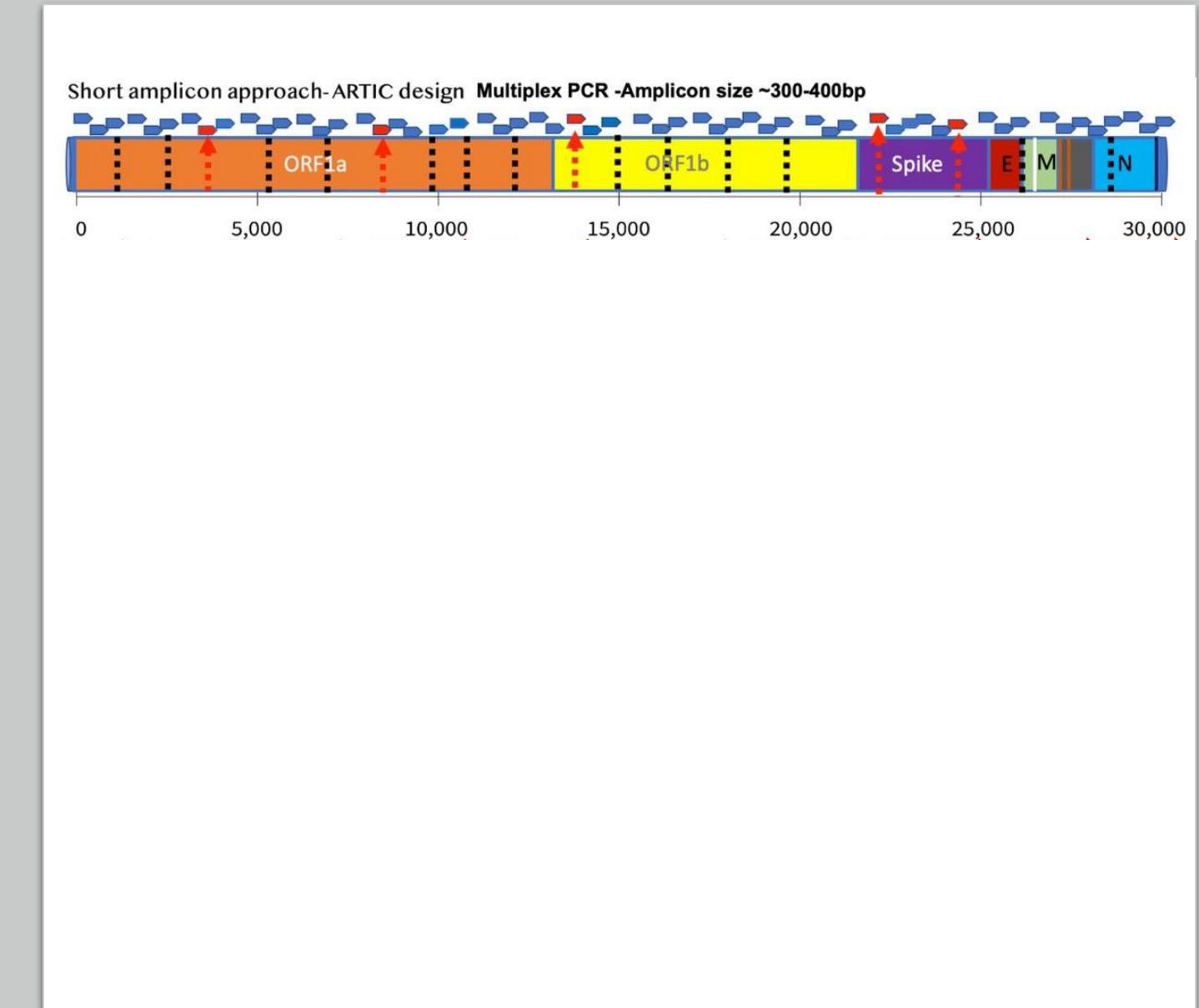
Princip celogenomového sekvenování COVID-19

- zpracování sekvenačních dat
- pipeline NF-Core viralrecon
- výstupy:
 - sekvence
 - variány (mutace)
 - kvalita readů, coverage (~99%), hloubka sekvenace (100x-3000x)

1. Download samples via SRA, ENA or GEO ids ([ENA](#) [FTP](#), [parallel-fastq-dump](#); *if required*)
2. Merge re-sequenced FastQ files ([cat](#); *if required*)
3. Read QC ([FastQC](#))
4. Adapter trimming ([fastp](#))
5. Variant calling
 - i. Read alignment ([Bowtie 2](#))
 - ii. Sort and index alignments ([SAMtools](#))
 - iii. Primer sequence removal ([iVar](#); *amplicon data only*)
 - iv. Duplicate read marking ([picard](#); *removal optional*)
 - v. Alignment-level QC ([picard](#), [SAMtools](#))
 - vi. Choice of multiple variant calling and consensus sequence generation routes ([VarScan 2](#), [BCFTools](#),
// [BCFTools](#), [BEDTools](#))
 - Variant annotation ([SnpEff](#), [SnpSift](#))
 - Consensus assessment report ([QUAST](#))
6. *De novo* assembly
 - i. Primer trimming ([Cutadapt](#); *amplicon data only*)
 - ii. Removal of host reads ([Kraken 2](#))
 - iii. Choice of multiple assembly tools ([SPAdes](#) // [metaSPAdes](#) // [Unicycler](#) // [minia](#))
 - Blast to reference genome ([blastn](#))
 - Contiguate assembly ([ABACAS](#))
 - Assembly report ([PlasmidID](#))
 - Assembly assessment report ([QUAST](#))
 - Call variants relative to reference ([Minimap2](#), [seqwish](#), [vg](#), [Bandage](#))
 - Variant annotation ([SnpEff](#), [SnpSift](#))
7. Present QC and visualisation for raw read, alignment, assembly and variant calling results ([MultiQC](#))

Princip celogenomového sekvenování COVID-19

- příklad Artic-v3 – nevýhody krátkých amplikonů



Přednastaveno pro snadné použití

```
cov_pipeline.sh -i /cesta/k/plateX/fastq/ -p plateX
```

- pracuje se všemi dostupnými sadami primerů, modifikovatelný podle aktuálních potřeb

Přednastaveno pro snadné použití

`cov_pipeline.sh -i /cesta/k/plateX/fastq/ -p plateX`

..	
.nextflow	Directory
failed	Directory
fastq	Directory
needs_investigation	Directory
results	Directory
upload	Directory
.nextflow.log	1.4 MB log-file
manual_checking.tsv	43 B tsv-file
merged.fasta	2.7 MB fasta-file
metadata_table.xlsx	51.3 KB Microsoft Excel...
nextclade.csv	77.4 KB csv-file
plate8_sample_sheet.csv	13.8 KB csv-file
results_pangolin.csv	7.5 KB csv-file
sample_sheet.xlsx	22.5 KB Microsoft Excel...
submission_file_GISAID.fasta	2.7 MB fasta-file
summary.tsv	6.2 KB tsv-file
Vzorky.xlsx	183.0 KB Microsoft Excel...

Minimální manuální úpravy

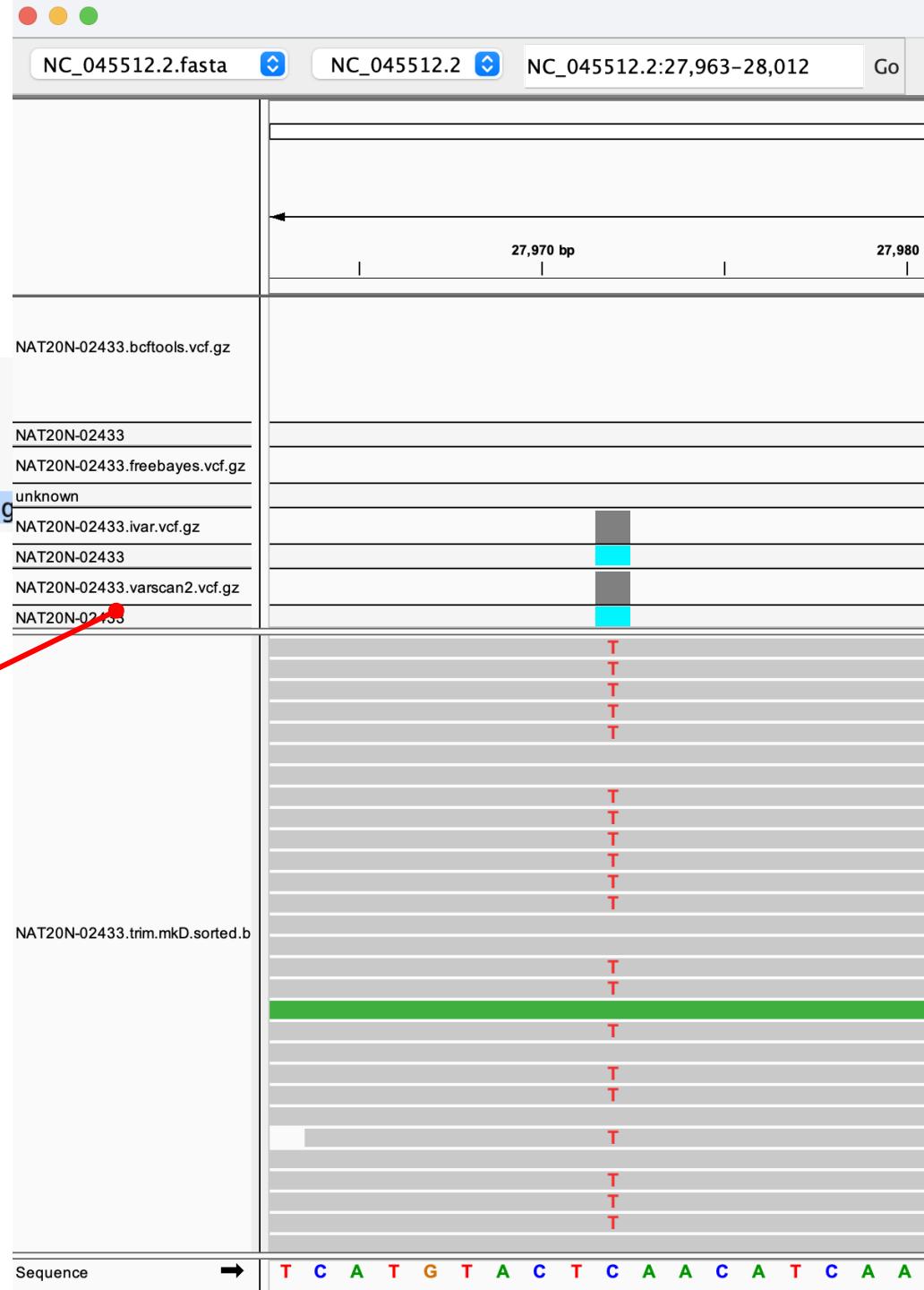
summary.tsv

NAT20N-02429	All four variant callers agree	0.992643	1317
NAT20N-02430	All four variant callers agree	0.992375	1285
NAT20N-02431	All four variant callers agree	0.997693	1349
NAT20N-02432	Freebayes, Ivar and Varscan agree	0.998161	1436
NAT20N-02433	Only Ivar and Varscan agree. BCFTools and Freebayes don't agree		

manual_checking.tsv

NAT20N-02384Δ	freebayes-
NAT20N-02419Δ	freebayes-
NAT20N-02425Δ	freebayes-
NAT20N-02427Δ	freebayes-
NAT20N-02433Δ	varsan2-

python manual_processing.py manual_checking.tsv
=> přesune finální soubory do ./upload/fasta



Na závěr

- izolace RNA zaběhnutá ze stěrů
- příprava knihoven z širokého rozmezí množství viru, časově efektivní
- viralrecon je flexibilní pipeline pro zpracování sekvenačních dat, kterou jsme vyladili s ohledem na automatizaci a jednoduchost
 - dostupná pro použití na serveru eltu1
 - kontakt fussyz@natur.cuni.cz or treitlis@natur.cuni.cz

Project „Enhancing Whole Genome Sequencing (WGS) and/or Reverse Transcription Polymerase Chain Reaction (RT-PCR) national infrastructures and capacities to respond to the Covid-19 pandemic in the European Union and European Economic Area“ had received funding from the European Centre for Disease Prevention and Control under the Grant Agreement number ECDC/HERA/2021/004 ECD.12218.



More information about the project: <http://www.szu.cz/ecdc-1>

NOTE:

Information included reflects only the authors' view and the Centre (ECDC) is not responsible for any use that may be made of the information it contains.