

EvoPAD Summer School „Bioinformatics“



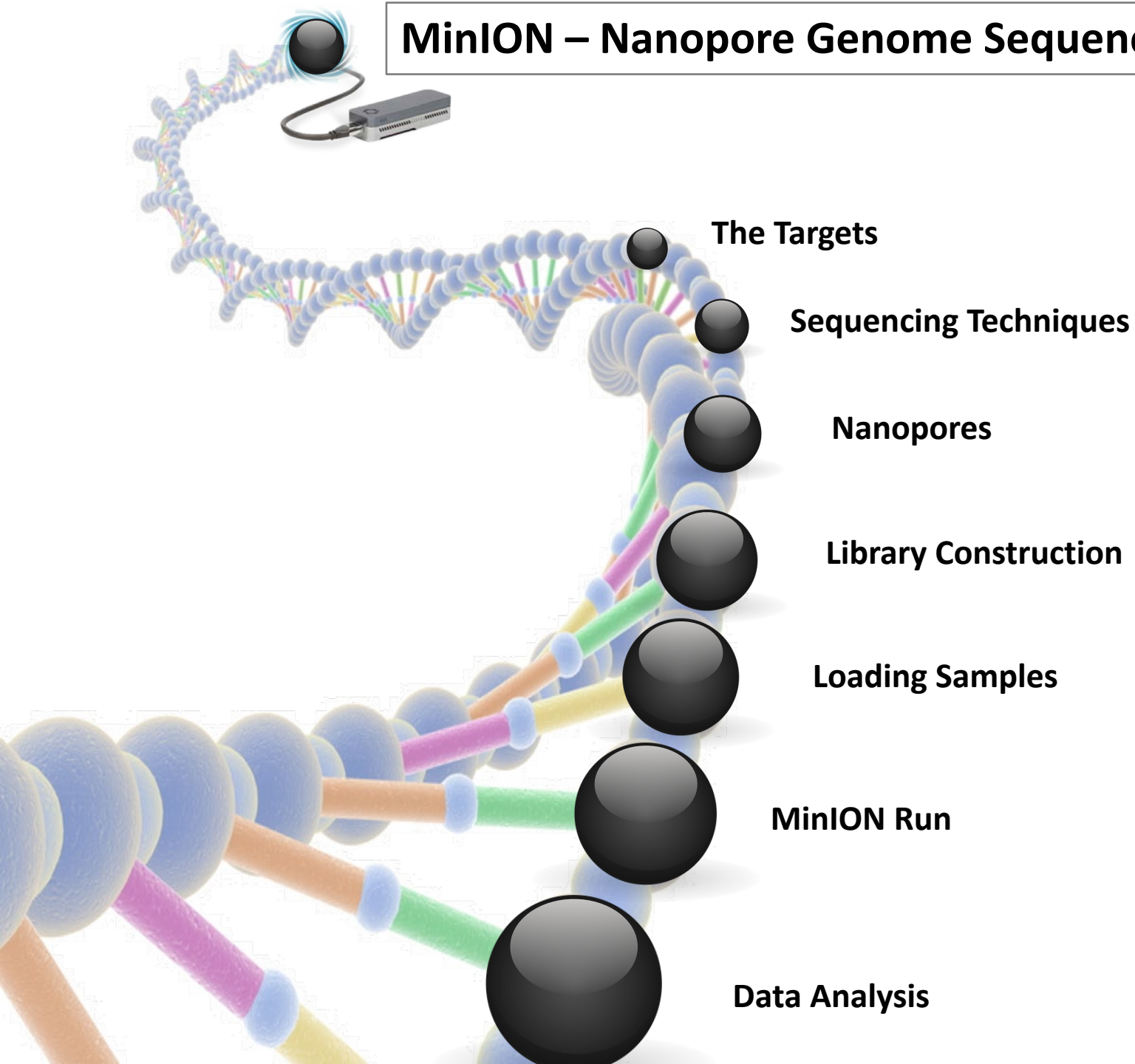
MinION Genome Sequencing



Jürgen Schmitz
Martin Kiefmann

10-13. September 2019 K6 Seminarhotel, Kirschallee 6, 38820 Halberstadt

MinION – Nanopore Genome Sequencing



The Targets

Sequencing Techniques

Nanopores

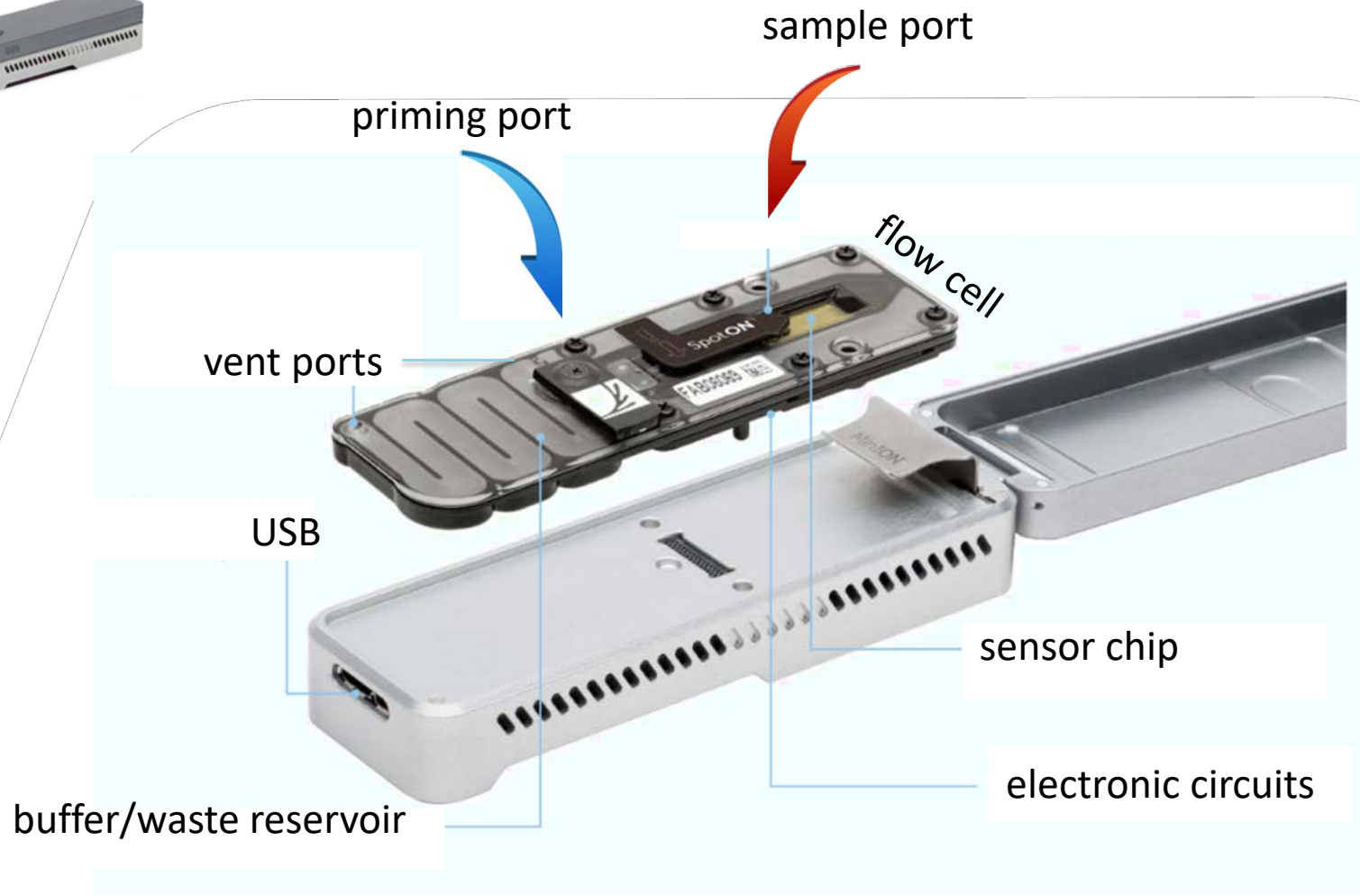
Library Construction

Loading Samples

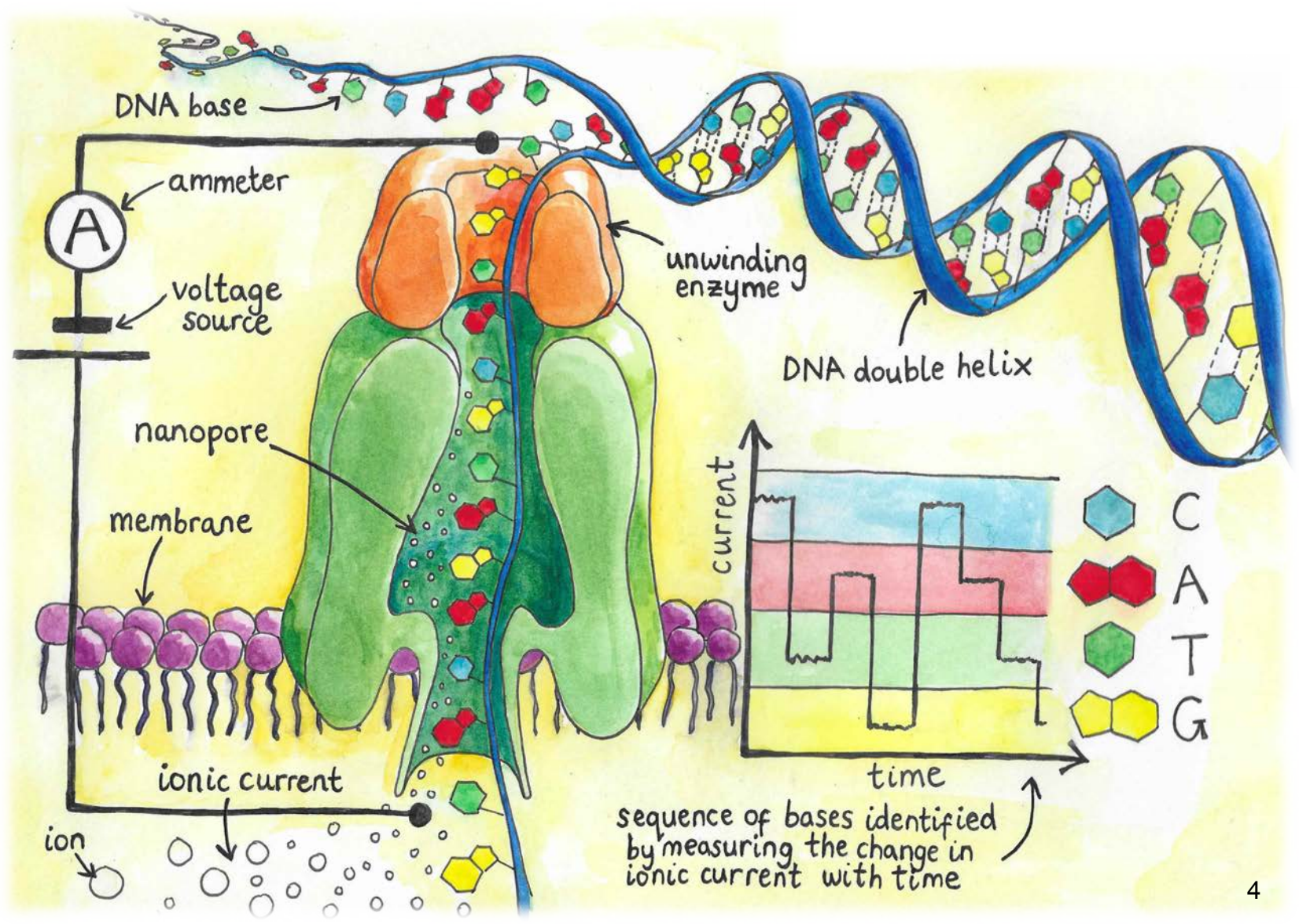
MinION Run

Data Analysis

PLUGANDPLAY



Principles of Nanopore Sequencing





MinION DNA extraction





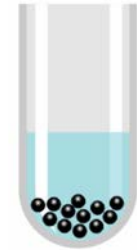
- Magnetic beads for efficient DNA binding
- Minimal sharing
- High range, high quality DNA for Nanopore (200-500kb)

Smart DNA prep (m) for MinION



Proteolytic Lysis of Tissue (1-100 mg)

1. Cut the tissue into small pieces and place in a 1.5 ml tube
2. Add 400 μ l **Lysis Solution CBV** and 40 μ l **Proteinase K**
3. Vortex shortly and incubate at 55°C for **1-3 hs** in thermal shaker (1,200 rpm)
4. After lysis centrifuge 5 min (max speed), transfer supernatant to a new tube
5. Add 1 μ l **RNase A** (10 mg/ml), digestion **10 min** at RT, transfer to **SE tube**



Binding DNA to SE Smart Extraction Macro Beads

6. Add 40 μ l **Binding Optimizer** and 350 μ l **2-Propanol** in the **SE tube**
7. Place SE tube into a thermal shaker (3 min; 1,400 rpm)
8. Place SE tube into magnetic rack for separating the SE Macro Beads
9. Discard supernatant

Washing and removing of alcohol

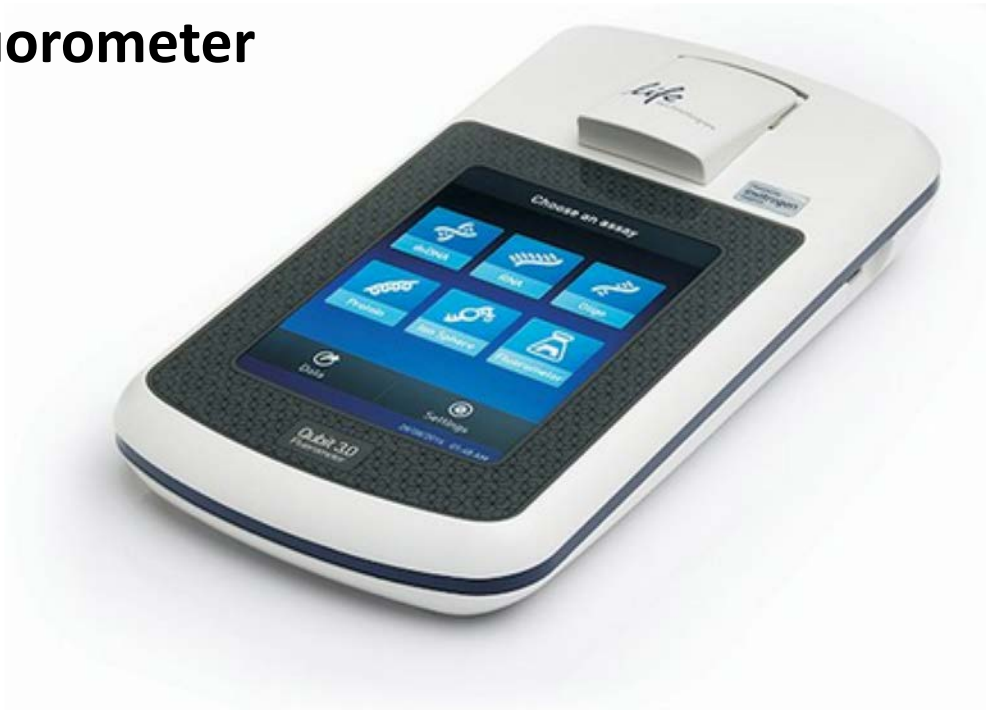
10. Add 800 μ l **Wasching Solution LS** (invert rack 5X)
 11. Discard supernatant – leaving SE tubes in magnetic rack
 - 2X | 12. Add 800 μ l of 80% ethanol and wash beads (invert rack 5X)
 13. Discard supernatant – leaving SE tubes in magnetic rack
- Remove the ethanol **15 min** 65°C in thermal cycler, lid open 400 rpm

Elution of DNA

14. Add 200 μ l – 1,000 μ l Elution Buffer **15 min** at 65°C, thermal cycler 1,000 rpm
15. Place SE tube into the magnetic rack and transfer DNA supernatant into a new tube



● Qubit™ 4 Fluorometer



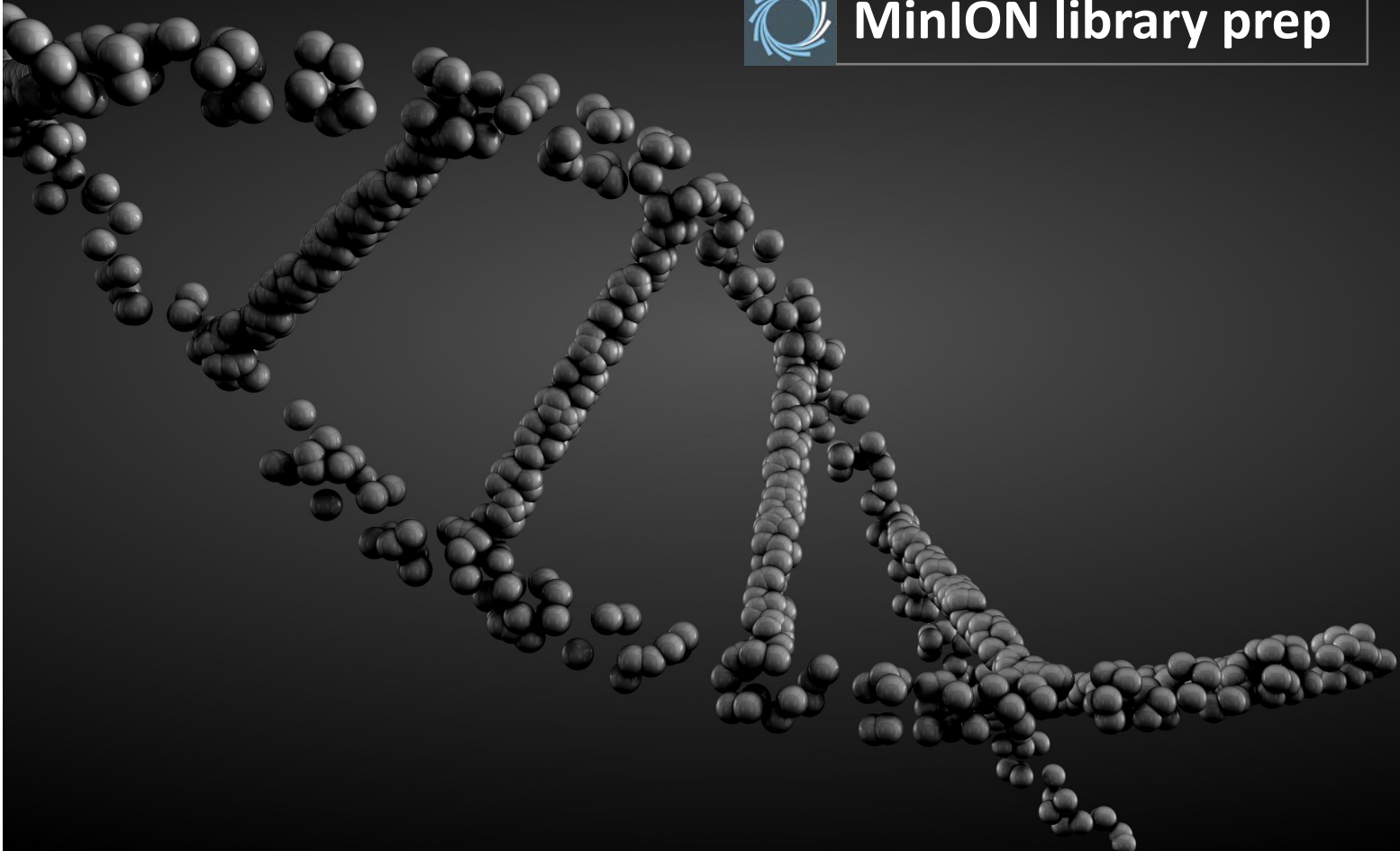
Quantify 1 μ l eluted sample

(15.) Average fragment size > 30 kb (electrophoresis)

16. Input mass (Qubit 1 μ g)

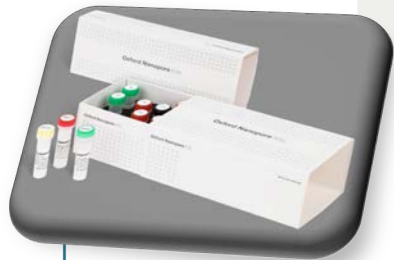


MinION library prep





SQK-LSK109 Ligation Sequencing Kit



60 min



High molecular weight gDNA



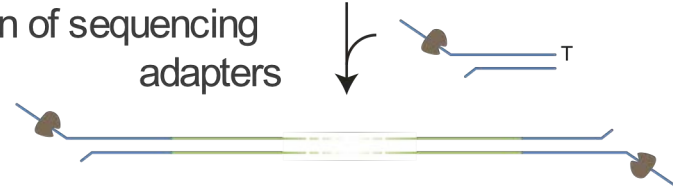
Optional fragmentation
(for <100 ng input)



repair
and end-prep



Ligation of sequencing
adapters



Loading

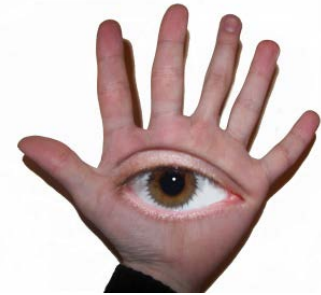


1D (directional) PCR-free gDNA



● Prepare the DNA

1. 1 μg genomic DNA in a DNA LoBind tube
2. Adjust volume to 49 μl with Nuclease-free water
3. Mixing by inversion (avoid shaking)
4. Spin down briefly in a microfuge



● End-prep and nick repair

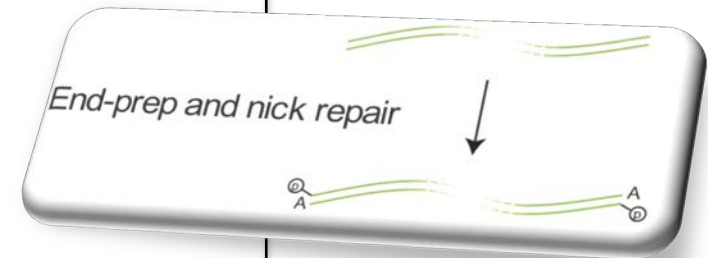
5. Thaw DNA CS (DCS) at RT, spin down, mix by pipetting, place on ice
6. Prepare the NEBNext FFPE DNA Repair Mix and NEB Next End repair/dA-tailing Module reagents and place on ice (see 5.)

Mix in a 0.2 ml thin-walled PCR tube:

- | | |
|-------------------|-----------------------------------|
| 1 μl | DNA CS |
| 47 μl | DNA |
| 3.5 μl | NEBNext FFPE DNA Repair Buffer |
| 2 μl | NEBNext FFPE DNA Repair Mix |
| 3.5 μl | Ultra II End-prep reaction buffer |
| 3 μl | Ultra II End-prep enzyme mix |

Mix gently by flicking the tube, and spin down

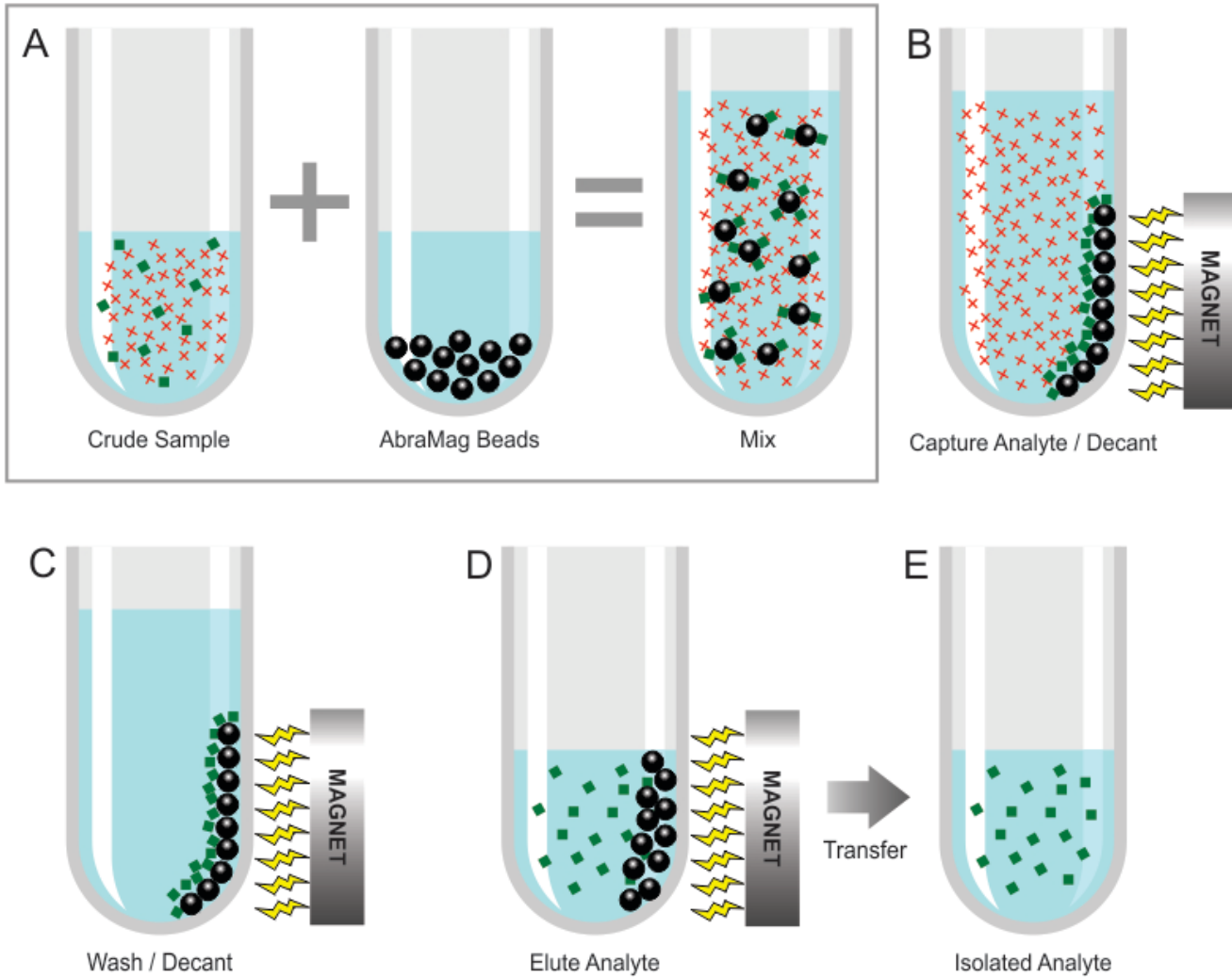
Using a thermal cycler, incubate at 20°C for 5 Min and 65°C 5 min

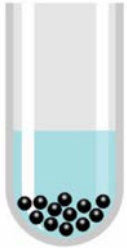


MinION library prep



Magnetic beads and magnetic rack





Clean and extract the repaired DNA

1. Resuspend the **AMPure XP** beads by vortexing
2. Transfer the DNA sample to a clean 1.5 ml DNA LoBind tube
3. Add 60 μ l of resuspended **AMPure XP** beads to the end-prep reaction
4. Mix by flicking the tube
5. Incubate on a Hula mixer (rotating) for **5-15 min** at RT
6. Prepare 500 μ l of fresh **70% ethanol** in Nuclease-free water
7. Spin down the sample and pellet on a magnet
8. Keep the tube on the magnet and pipette off the supernatant

2X

9. Keep on magnet, wash beads with 200 μ l of freshly prepared **70% ethanol** without disturbing the pellet
10. Remove the 70% ethanol using a pipette and discard

11. Spin down and place the tube back on the magnet
12. Pipette off any residual ethanol. Allow to dry for ca. 30 seconds
13. Remove tube from magnetic rack and resuspend pellet in 61 μ l water
14. Incubate for **5 min** at RT
15. Pellet the beads on a magnet until the eluate is clear and colourless
16. Remove and retain 61 μ l of eluate into a clean 1.5 ml LoBind tube

- Safe stop point 4°C -



● Quantification



Quantify 1 μ l eluted sample using a Qubit fluorometer

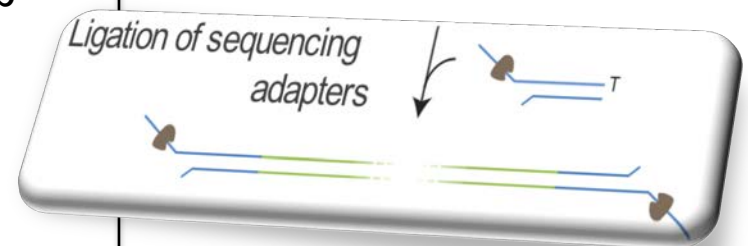


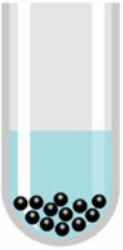
Ligation of sequencing adapters

1. Spin down **Adapter Mix (AMX)** and **T4 Ligase (E6056)**. Place on ice
2. Thaw **Ligation Buffer (LNB)** at RT, spin down and mix by pitpetting. Place on ice
3. Thaw the **Elution Buffer (EB)** at RT, mix by vortexing, spin down. Place on ice
4. To enrich DNA fragments of 3 kb or longer, thaw one tube of **L Buffer (LFB)**
To retain DNA fragments shorter than 3 kb, thaw one tube of **S Buffer (SFB)**
5. Mix by vortexing, spin down and place on ice

6. In a 1.5 LoBind tube mix:
 - 60 μ l DNA sample from previous step
 - 25 μ l **Ligase Buffer (LNB)**
 - 10 μ l NEBNext Quick **T4 DNA Ligase**
 - 5 μ l **Adapter Mix (AMX)**

Mix gently by flicking the tube, and spin down
Incubate **10 min** at RT





Clean and extract the adapter-ligated DNA

7. Prepare the **AMPure XP beads** for use. Resuspend by vortexing
8. Add 40 μ l of resuspended **AMPure XP** beads to the reaction and flick tube
9. Incubate on Hula mixer for **5-45 min** at RT
10. Spin down the sample and pellet on a magnet. Keep the tube on the magnet
11. Remove the supernatant using a pipett and discard

2X

12. Wash the beads by adding either
 - 250 μ l **Long Fragment Buffer (LFB)**
 - or 250 μ l **Short Fragment Buffer (SFB)**
 - Flick the beads to resuspend, then return the tube to the magnetic rack
 - allow the beads to pellet
 - Remove the supernatant using a pipett and discard

13. Spin down and place the tube back on the magnet
14. Pipette off any residual supernatant
15. Allow to try for ca. 30 seconds
16. Remove the tube from the magnetic rack and resuspend pellet in 15 μ l **EB buffer**
17. Incubate for **10-30 min** at 37°C
18. Pellet the beads on a magnet until the eluate is clear and colourless



● Quantify

19. Remove and retain 15 μ l of eluate (Lib) into a clean 1.5 μ l LoBind tube
20. Dispose of the pelleted beads
21. Quantify 1 μ l eluted sample using a Qubit fluorometer



22. The prepared library is used for loading into the flow cell
23. Store the library on ice until ready to load



MinION library prep



SQK-LSK109

Ligation Sequencing Kit



High molecular weight gDNA



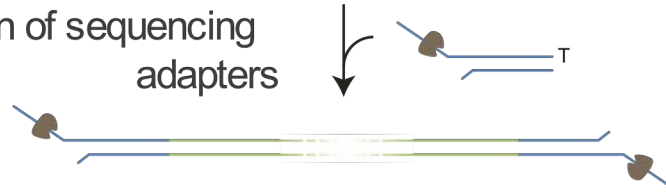
Optional fragmentation
(for <100 ng input)



repair
and end-prep



Ligation of sequencing
adapters



Loading

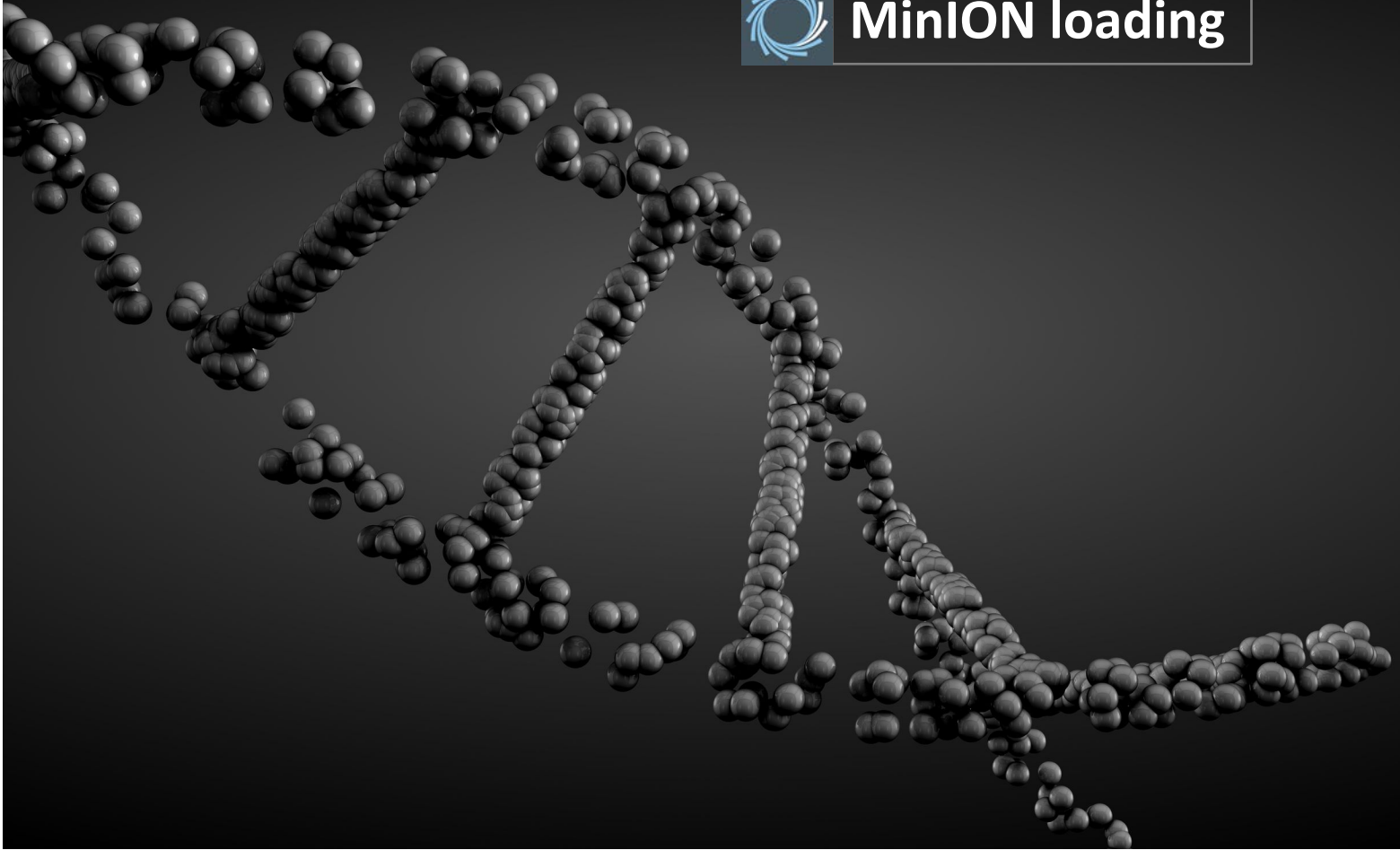


60 min





MinION loading



MinION loading



● Loading strategy

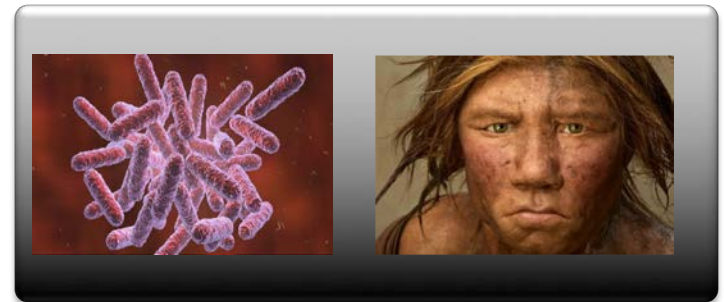
first 6 h

next 6 h



first 6 h

next 6 h



Flow cell 1



Flow cell 2



● MinION – Set up the MinION flow cell and host computer

1. Open the MinKNOW GUI from the desktop icon
2. Establish a local internet connection (firmware confirmation)
3. Turn off sleep mode of computer
4. Choose the flow cell type from the selector box (FLO-MIN106)
5. Click Start test
6. Check the number of active pores available in System History panel
- should be >800 from 4 X 512





● Priming and loading the SpotON flow cell

1. Thaw the **Sequencing Buffer (SQB)**, **Loading Beads (LB)**, **Flush Tether (FLT)** and **Flush Buffer (FLB)** at RT, place on ice
2. Mix the **Sequencing Buffer (SQB)** and **Flush Buffer (FLB)** tubes by vortexing
3. Spin down the **Flush Tether (FLT)** tube, mix by pipetting, return to ice
4. Open the lid of the nanopore sequencing device and slide the flow cell's **priming port cover** clockwise so that the priming port is visible



priming port

5. Care must be taken when drawing back buffer from the flow cell. The array of pores must be covered by buffer at all time. **Removing more than 20-30 μ l** damaging the pores in the array
6. After opening the priming port, check for small bubble under the cover
Draw back a small volume to **remove any bubble** (a few μ ls)
7. Set a P1000 pipette to 200 μ l
8. Insert the tip into the priming port
9. Turn the wheel until the dial shows **220-230 μ l**, or until you see a small volume of buffer entering the pipette tip



10. Prepare the flow cell priming mix: add 30 μl of thawed and mixed flush Tether (FLT) directly to the tube of thawed and mixed Flush Buffer (FLB), and mix by pipetting
11. Load 800 μl of the priming mix into the flow cell via the **priming port**



12. Avoid air bubbles, wait for 5 min
13. Thoroughly mix the content of the Loading Beads (LB) by pipetting
The Loading Beads (LB) tube contains a suspension of beads. These beads settle very quickly. It is vital that they are **mixed!!!** immediately before use
14. In a new tube, prepare the library for loading as follows:

37.5 μl	Sequencing Buffer (SQB)
25.5 μl	Loading Beads (LB), mixed immediately before use
<u>12</u> μl	DNA library
75 μl	

MinION loading



15. Gently lift the **SpotON sample port** cover to make the SpotON sample port accessible



16. Load slowly!!! 200 μ l of the priming mix into the flow cell via the **priming port!!!!** (NOT the SpotON sample port). Avoiding airbubbles



17. Mix the prepared library gently by pipetting up and down just prior loading

18. Add 75 μ l of sample to the flow cell via the **SpotON sample port** in a dropwise fashion. Ensure each drop flows into the port before adding the next



19. Gently close the spotON sample port cover, making sure the bung enters the SpotON port, close the priming port and replace the MinION lid.



MinION run





● Data strategy



MS_GP



coati

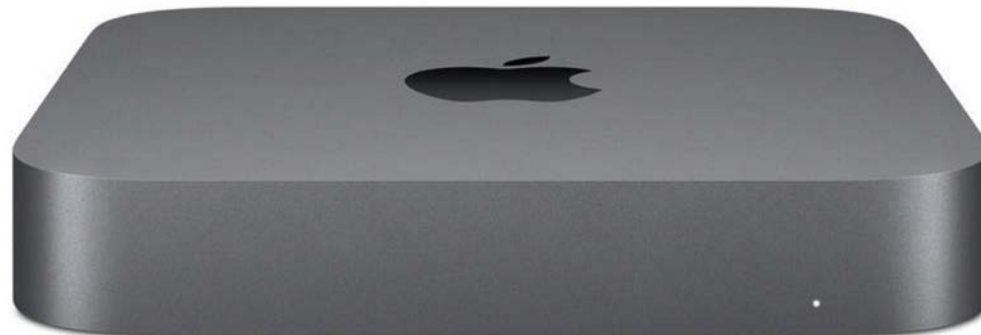


Ec3



human


Names:



Data deposited: /Library/MinKNOW/data

MinION run



1. Double-click the **MinKNOW** icon  at the desktop to open the MinKNOW GUI
2. Choose the flow cell type from the selector box. Then check „Available“ box
FLO-MIN106
3. Click the „**New Experiment**“ button at the bottom left of the GUI

On the new experiment popup screen, select the running parameters for your Experiment from the individual tabs

Output settings – **FASTQ**: The number of basecalls that MinKNOW will write in a single file. By default this is set to 4000

Output settings – **FAST5**: The number of files that MinKNOW will write to a single folder. Default 4000

Basecalling tab to ON – if using basecalling via MinKNOW (abalone)

Basecalling tab to Off – if using basecalling via guppy later

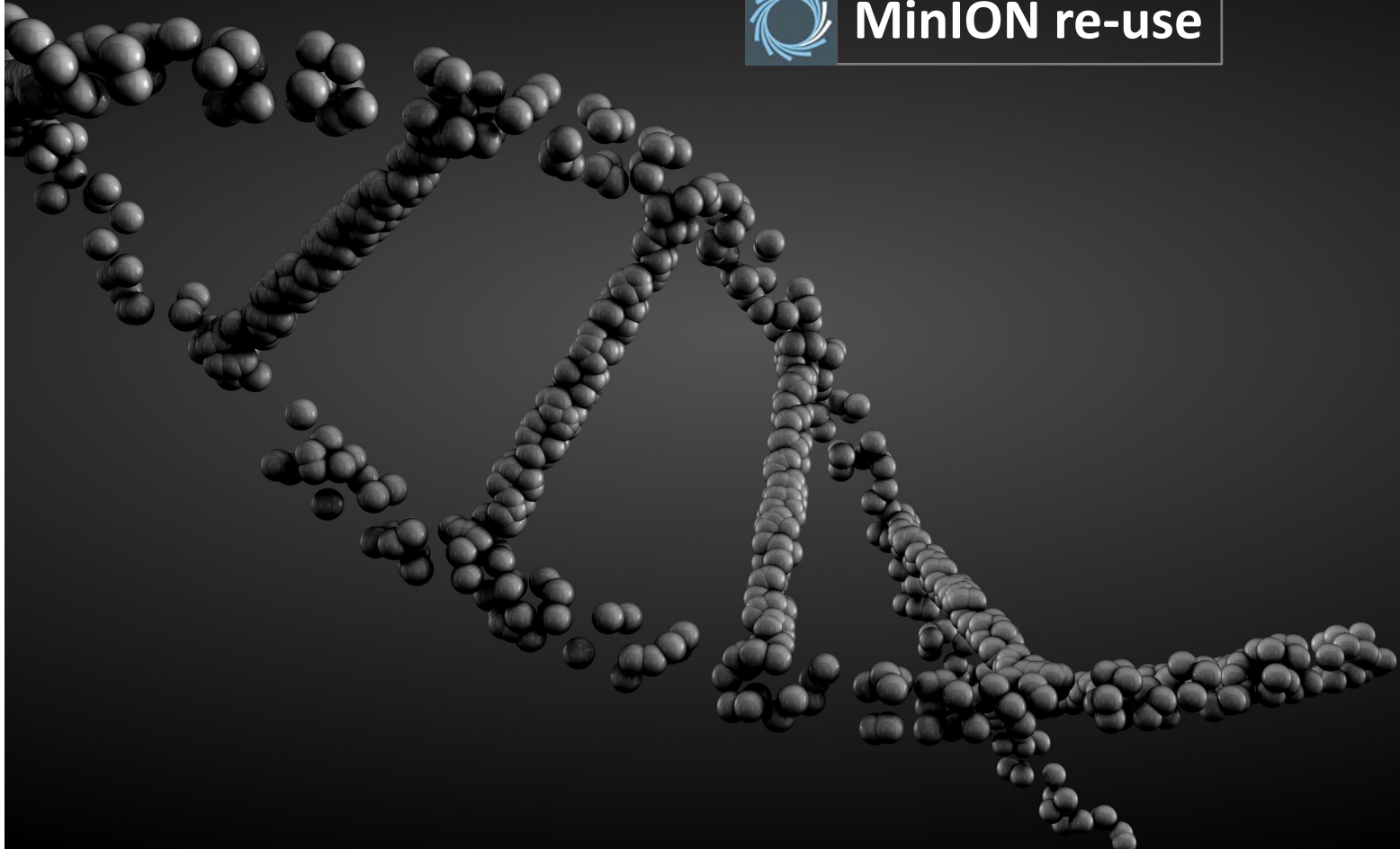
Click start run

Allow the script to run to completion

The MinKNOW Experiment page will indicate the progress of the script
This can be accessed through the „Experiment“ tab (top right of the screen)
Monitor messages in the message panel in the MinKNOW GUI



MinION re-use



MinION re-use flow cell



● Wash Kit (EXP-WSH002) – to re-using flow cells

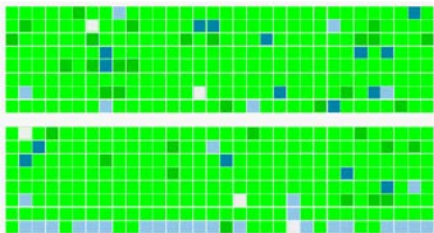
Solution A	1000 μ l	2 tubes	12X	sequesters the library strands from membrane
Solution B	500 μ l	4		removes the sequestered library refreches FC
Storage Buffer	1600 μ l	4		

1. Open the **priming port** cover to check that buffer is continuous
2. Ensure that the SpotON sample loading port is closed
3. Gilson P1000 **adding 150 μ l Solution A** through priming port
6. Wait **10 min**
7. Add **150 μ l Solution B**
8. Start a **new MinKNOW experiment**

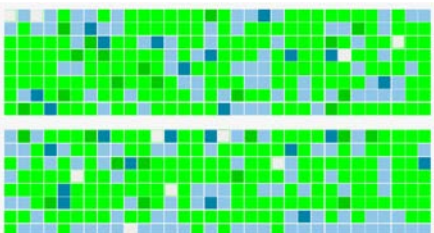
Later use

7. Slowly add 500 μ l of Storage Buffer through priming port
8. Close priming port. Remove buffer from the waste section of the flow cell through either of waste ports
9. Store flow cell at 4-8°C

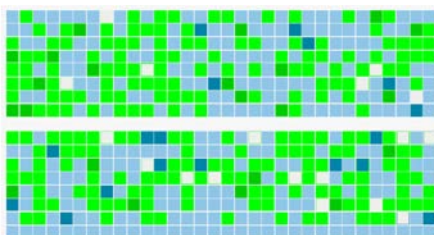
■ pore active/512



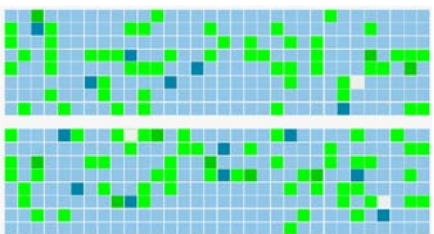
5 h



18h



28h



45h

■ pore not longer available/512

Sequencing Running...

READS: 172.1M ESTIMATED BASES: 1.9 Gb QUEUED: 53% CALLED: 17% BASES: 616.66 Mb

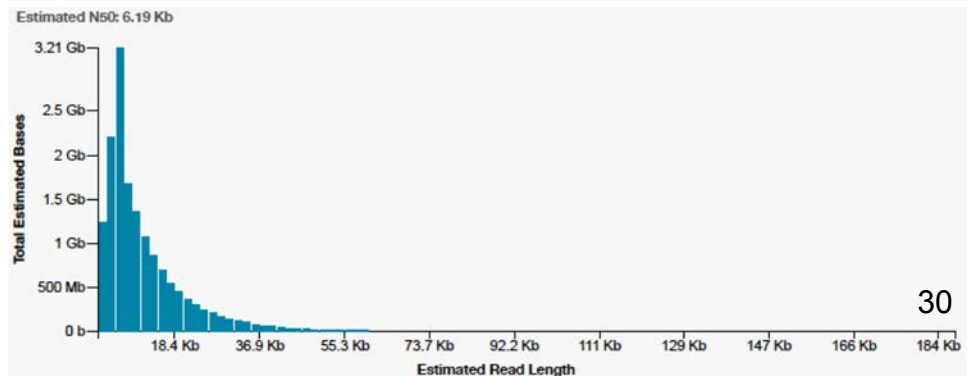
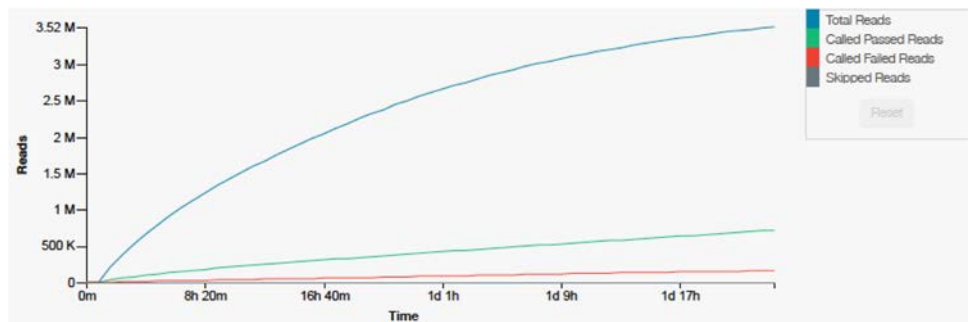
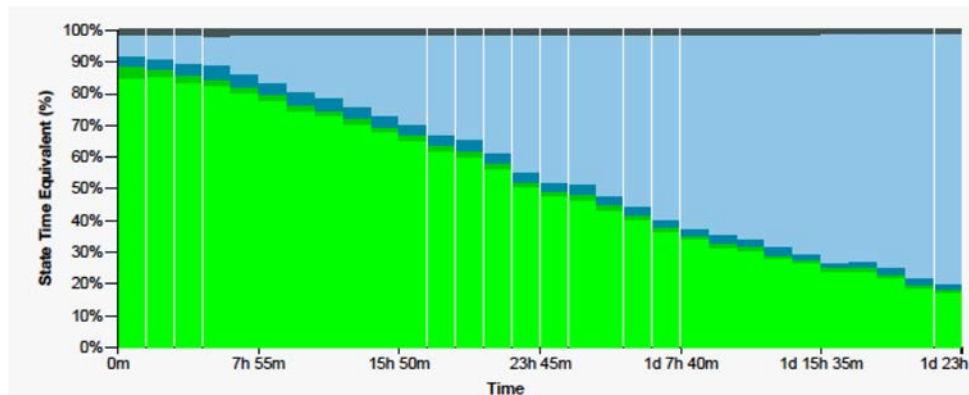
Position: M429066 Experiment group: Nanosieve Sample ID: Nanosieve Flow cell ID: F43C4093 Flow cell product code: FLD-MN106 Kit ID: SQK-LBK109 Current output directory: /Library/Miniconda3/base/... Basecall model: Fast basecall...

7 Hours 26 Minutes / 48 Hours

15% [Progress bar]

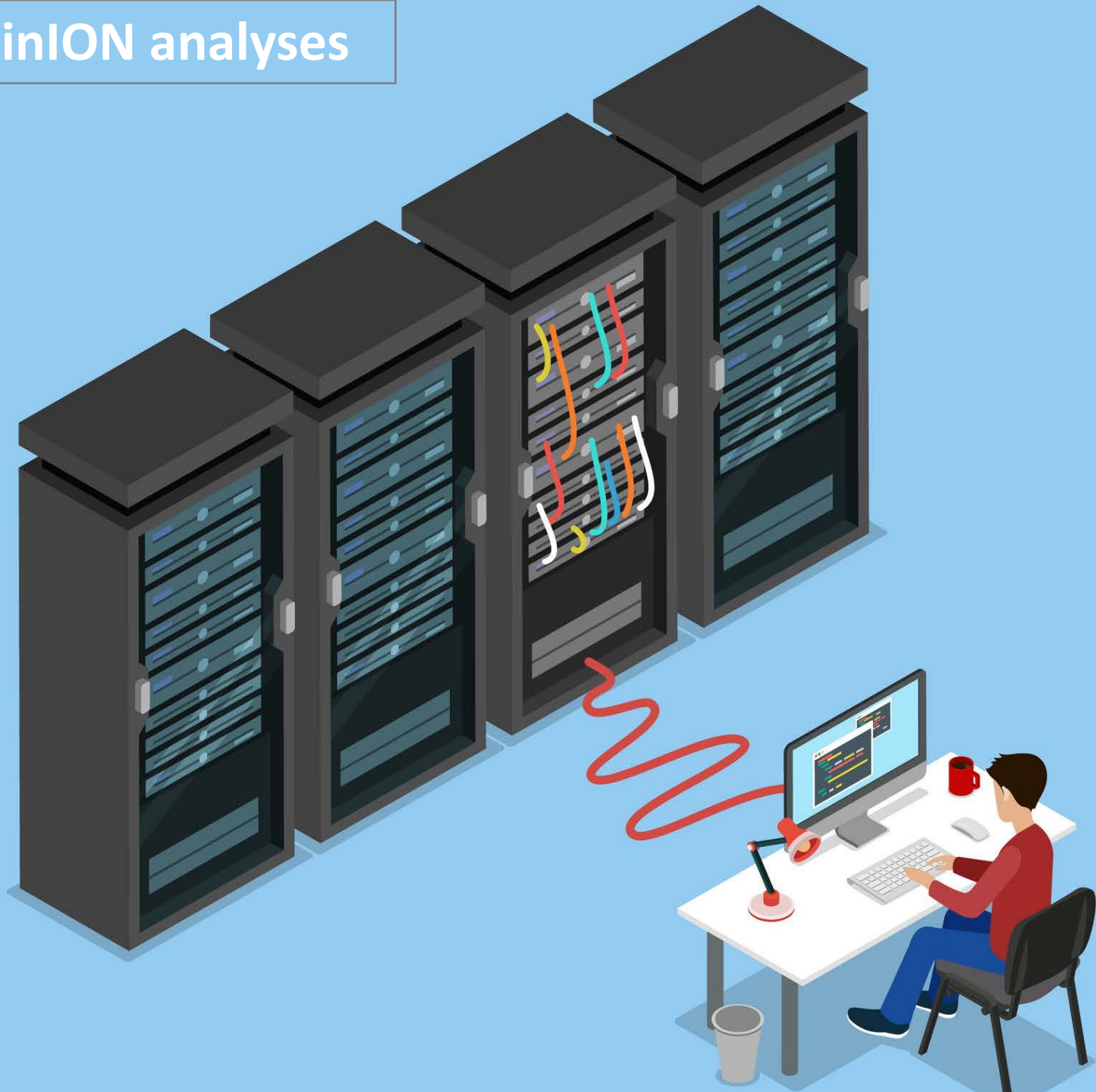
Stop run [Red button] Pause run [Yellow button] **Export PDF Report [Blue button]**

TOTAL RUN TIME: 7h 27m 56s VOLTAGE: -175 mV

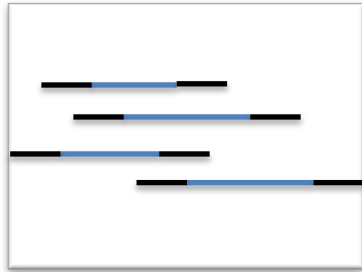




MinION analyses



MinION analyses

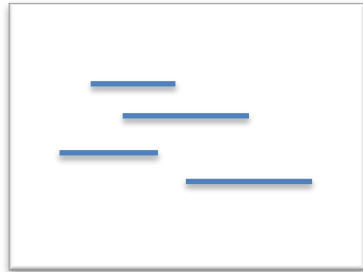


MinION
Sequencing

reads
fast5



48 h



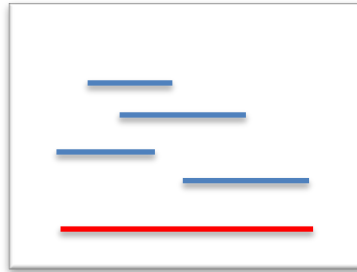
MinKNOW/Guppy
Quality control



reads
fastq



7 d
GPUs ⚡



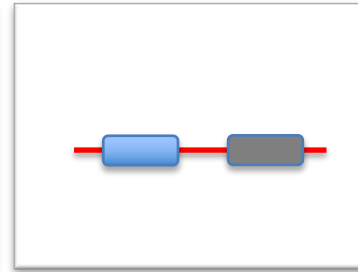
Canu
Assembly



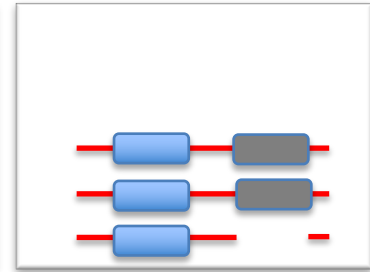
contigs
fasta



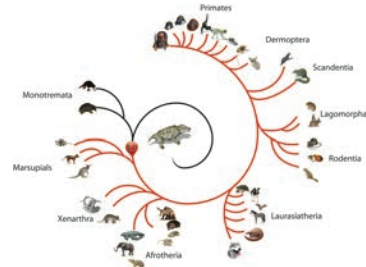
10 d



BLAST-RM
Annotation



UCSC
Comparison





● **Server access**

```
ssh -p 40044 extern@retrogenomics2.uni-muenster.de  
password: eXtern01  
location: /home/extern
```



● **Transfer > server (terminal client)**

```
scp -o "Port 40044" -r /Users/juergenschmitz/Desktop/Transfer/*.* extern@retrogenomics2.uni-muenster.de:/home/extern/data/coati/fast5
```

source client target server

● **Transfer < server (terminal client)**

```
scp -o "Port 40044" extern@retrogenomics2.uni-muenster.de:/home/extern/data/coati/fast5/\*.\* /Users/juergenschmitz/Desktop/Transfer
```

source server target client

ssh = secure shell network protocol

scp = secure copy protocol

-o = option

-r = copy all files in a folder recursively

VPN:

zmbepprak

ZMBE.2019n

MinION analyses



UNIX commands:

mc

visual shell



terminal

```
juergenschmitz — mc [extern@retrogenomics2-work]:~/data/coati — ssh -p 40044 extern@retrogenomics2.uni-muenster.de — 202x31
```

Left	File	Command	Options	Right																																														
< ~/data/coati				< ~/data/coati/output_canu																																														
<table border="1"><thead><tr><th>Name</th><th>Size</th><th>Modify</th><th>time</th></tr></thead><tbody><tr><td>..</td><td></td><td></td><td></td></tr><tr><td>UP—DIR</td><td></td><td>Sep 6</td><td>19:40</td></tr><tr><td>/RM</td><td>2</td><td>Sep 3</td><td>20:19</td></tr><tr><td>/fast5</td><td>6</td><td>Sep 4</td><td>10:00</td></tr><tr><td>/fastq</td><td>6</td><td>Sep 5</td><td>15:43</td></tr><tr><td>/output_canu</td><td>3</td><td>Sep 8</td><td>13:58</td></tr><tr><td>/output_guppy</td><td>46</td><td>Sep 8</td><td>14:17</td></tr></tbody></table>	Name	Size	Modify	time	..				UP—DIR		Sep 6	19:40	/RM	2	Sep 3	20:19	/fast5	6	Sep 4	10:00	/fastq	6	Sep 5	15:43	/output_canu	3	Sep 8	13:58	/output_guppy	46	Sep 8	14:17		<table border="1"><thead><tr><th>Name</th><th>Size</th><th>Modify</th><th>time</th></tr></thead><tbody><tr><td>..</td><td></td><td></td><td></td></tr><tr><td>UP—DIR</td><td></td><td>Sep 4</td><td>11:30</td></tr><tr><td>mar.contigs.fasta</td><td>636240K</td><td>Sep 8</td><td>13:58</td></tr></tbody></table>	Name	Size	Modify	time	..				UP—DIR		Sep 4	11:30	mar.contigs.fasta	636240K	Sep 8	13:58
Name	Size	Modify	time																																															
..																																																		
UP—DIR		Sep 6	19:40																																															
/RM	2	Sep 3	20:19																																															
/fast5	6	Sep 4	10:00																																															
/fastq	6	Sep 5	15:43																																															
/output_canu	3	Sep 8	13:58																																															
/output_guppy	46	Sep 8	14:17																																															
Name	Size	Modify	time																																															
..																																																		
UP—DIR		Sep 4	11:30																																															
mar.contigs.fasta	636240K	Sep 8	13:58																																															
/output_canu			6543G/6696G (97%)	mar.contigs.fasta			6543G/6696G (97%)																																											

Hint: Want to do complex searches? Use the External Panelize command.
extern@retrogenomics2-work:~/data/coati>

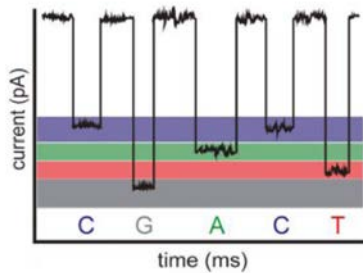
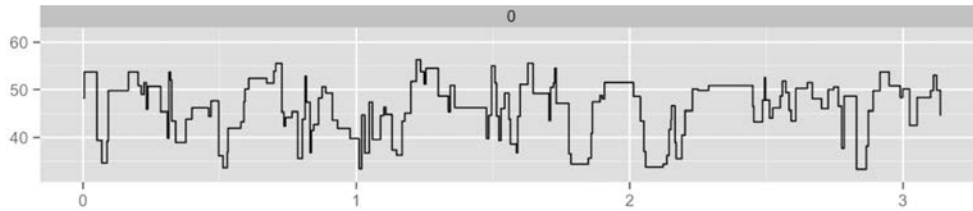
```
1 help 2 menu 3 view 4 edit 5 copy 6 renNov 7 mkdir 8 delete 9 pullDn 10 quit
```

MinION analyses



FAST5 files

Data out of Oxford Nanopore Technology (ONT)
Raw electrical signal in HDF5 (Hierarchical Data Format)
Main data: „squiggles“ = pico-amp measurements at micropores





FASTQ files

```
@0243e9e9-5442-4df3-aa14-fdfd6d2ba214 flow_cell_id=FAK33926 protocol_group_id=Maria1 sample_id=maria1
ATCAGTATTGCTTCGTTTCGTTACGTATTGCTGGCAGCAGGTGCGGGCGTACTTCAGCCTGGCGTTCTGTGACTTCTGCCTTAAGTTT
+
'-8#.,(+%%+1631:66#0(,54<96<564593-14)+&()&*&%$)*+-7>50%3++02-23>999:9788>9<.:05/,&++<<89<6-.*(2,+3;1
```

@ ... = sequence identifier and description

ACTG ... = sequence

+ = separator

!) 3 =... = quality value

ASCII	p
!	1
)	0.1
3	0.01
=	0.001
H	0.0001

Probability of incorrect base



● FASTA files

```
>coati_FAK34099_9e4d2b4f8a8663b7a95159919ce03578e435d1d4_0  
ATCAGTATTGCTTCGTTTCGGTTACGTATTGCTGGCAGCAGGTGCGGGCGTACTTCAGCCTGGCGTTCTGTGACTTCTGCCTTAAGTTT
```

> ... = header

ACTG ... = sequence

● **Transfer FASTQ to FASTA:** `sed -n '1~4s/^@/>/p;2~4p' INFILE.fastq > OUTFILE.fasta`

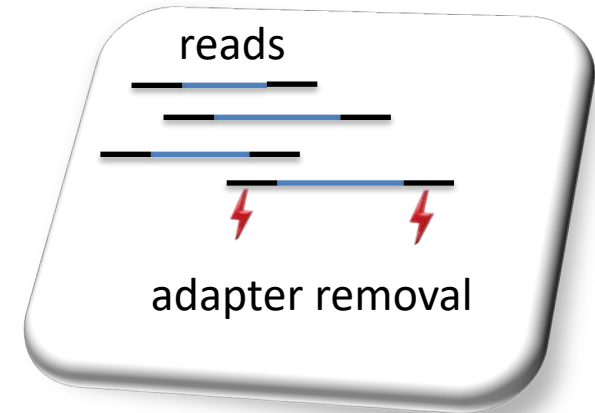


● Guppy basecalling

Adapter Strand Trimming: Removal of sequencing adapter

Input: .fast5 files containing raw data

Output: .fast5 and FASTQ files basecalled



Command line (server):

source fast5

output

```
guppy_basecaller --input_path /home/extern/data/coati/fast5/ --save_path /home/extern/data/coati/output_guppy --  
config/appl/src/guppy/ont-guppy-cpu/data --flowcell FLO-MIN106 --kit SQK-LSK109 --num_callers 10 --recursive
```

guppy config file

used flowcell

used kit

threads

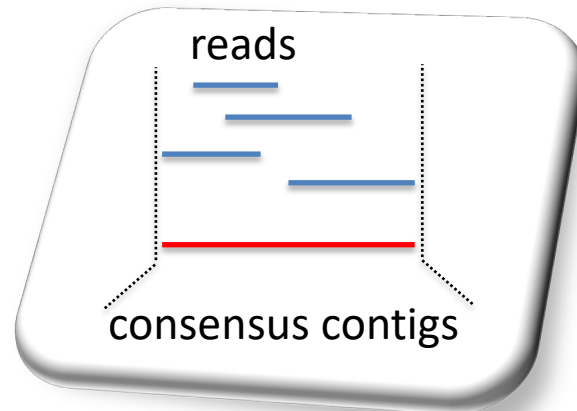
including subfolders

Canu assembly

Reads-to-contigs (derive consensus sequences)

Input: .fastq or fasta

Output: .fasta



Command line:

> 6X coverage 1 node genome size memory/threads restrictions MinION settings save folder

canu stopOnLowCoverage=6 useGrid=false genomeSize=2.4g maxMemory 256 maxThreads 24 correctedErrorRate=0.16 -p coatiCANU -d /home/extern/data/coati/output_canu -nanopore-raw /home/extern/data/coati/fastq/for_coati/ *.fastq

save folder

data folder

10 days - 56 threads - 521 Gb

```
-- Finished on Sat Sep 7 19:43:48 2019 (10 seconds) with 6100.424 GB
-----
-- Finished stage 'generateOutputs', reset canuIteration.
-- Assembly 'mar' finished in '/data/projects/ASSEMBLY/nasenbaer'.
-- Summary saved in 'mar.report'.
-- Sequences saved:
-- Contigs      -> 'mar.contigs.fasta'
-- Unassembled -> 'mar.unassembled.fasta'
-- Unitigs     -> 'mar.unitigs.fasta'
-- Read layouts saved:
-- Contigs     -> 'mar.contigs.layout'.
-- Unitigs    -> 'mar.unitigs.layout'.
-- Graphs saved:
-- Contigs    -> 'mar.contigs.gfa'.
-- Unitigs   -> 'mar.unitigs.gfa'.
-- Bye.
```

0.6 Gb contigs – 0.6 Gb unique reads

```
drwxrwxr-x 2 juergen user 688 Sep 7 19:43 canu-logs
drwxrwxr-x 2 juergen user 2 Aug 30 13:35 canu-scripts
drwxrwxr-x 6 juergen user 10 Sep 6 19:02 correction
drwxrwxr-x 2 juergen user 2 Aug 30 13:35 haplotype
-rw-rw-r-- 1 juergen user 651509421 Sep 7 19:43 mar.contigs.fasta
-rw-rw-r-- 1 juergen user 8498310 Sep 7 19:43 mar.contigs.gfa
-rw-rw-r-- 1 juergen user 9524804661 Sep 7 19:41 mar.contigs.layout
-rw-rw-r-- 1 juergen user 59123961 Sep 7 19:41 mar.contigs.layout.readToTig
-rw-rw-r-- 1 juergen user 51970305 Sep 7 19:41 mar.contigs.layout.tigInfo
-rw-rw-r-- 1 juergen user 3435934573 Sep 6 19:05 mar.correctedReads.fasta.gz
-rw-rw-r-- 1 juergen user 27307 Sep 7 19:43 mar.report
drwxrwxr-x 3 juergen user 29 Sep 7 06:11 mar.seqStore
-rw-rw-r-- 1 juergen user 226520 Aug 30 13:39 mar.seqStore.err
-rw-rw-r-- 1 juergen user 224202 Aug 30 13:35 mar.seqStore.ssi
-rw-rw-r-- 1 juergen user 2536456933 Sep 7 06:14 mar.trimmedReads.fasta.gz
-rw-rw-r-- 1 juergen user 4090474584 Sep 7 19:43 mar.unassembled.fasta
-rw-rw-r-- 1 juergen user 2230197 Sep 7 19:43 mar.unitigs.bed
-rw-rw-r-- 1 juergen user 663418894 Sep 7 19:43 mar.unitigs.fasta
-rw-rw-r-- 1 juergen user 16870520 Sep 7 19:43 mar.unitigs.gfa
-rw-rw-r-- 1 juergen user 1365885494 Sep 7 19:42 mar.unitigs.layout
-rw-rw-r-- 1 juergen user 20378745 Sep 7 19:42 mar.unitigs.layout.readToTig
-rw-rw-r-- 1 juergen user 2865999 Sep 7 19:42 mar.unitigs.layout.tigInfo
drwxrwxr-x 5 juergen user 7 Sep 7 06:12 trimming
drwxrwxr-x 10 juergen user 16 Sep 7 19:38 unitigging
```



NCBI Resources ▾ How To ▾ maglet My NCBI Sign Out

NCBI National Center for Biotechnology Information

All Databases Search

- NCBI Home
- Resource List (A-Z)
- All Resources
- Chemicals & Bioassays
- Data & Software
- DNA & RNA
- Domains & Structures
- Genes & Expression
- Genetics & Medicine
- Genomes & Maps
- Homology
- Literature
- Proteins
- Sequence Analysis
- Taxonomy
- Training & Tutorials
- Variation

Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News](#) | [Blog](#)

Submit

Deposit data or manuscripts into NCBI databases



Download

Transfer NCBI data to your computer



Learn

Find help documents, attend a class or watch a tutorial



Develop

Use NCBI APIs and code libraries to build applications



Analyze

Identify an NCBI tool for your data analysis task



Research

Explore NCBI research and collaborative projects



Popular Resources

- PubMed
- Bookshelf
- PubMed Central
- PubMed Health
- BLAST
- Nucleotide
- Genome
- SNP
- Gene
- Protein
- PubChem

NCBI Announcements

New video on the NCBI YouTube channel: [Viral resources at NCBI](#)

11 Feb 2016

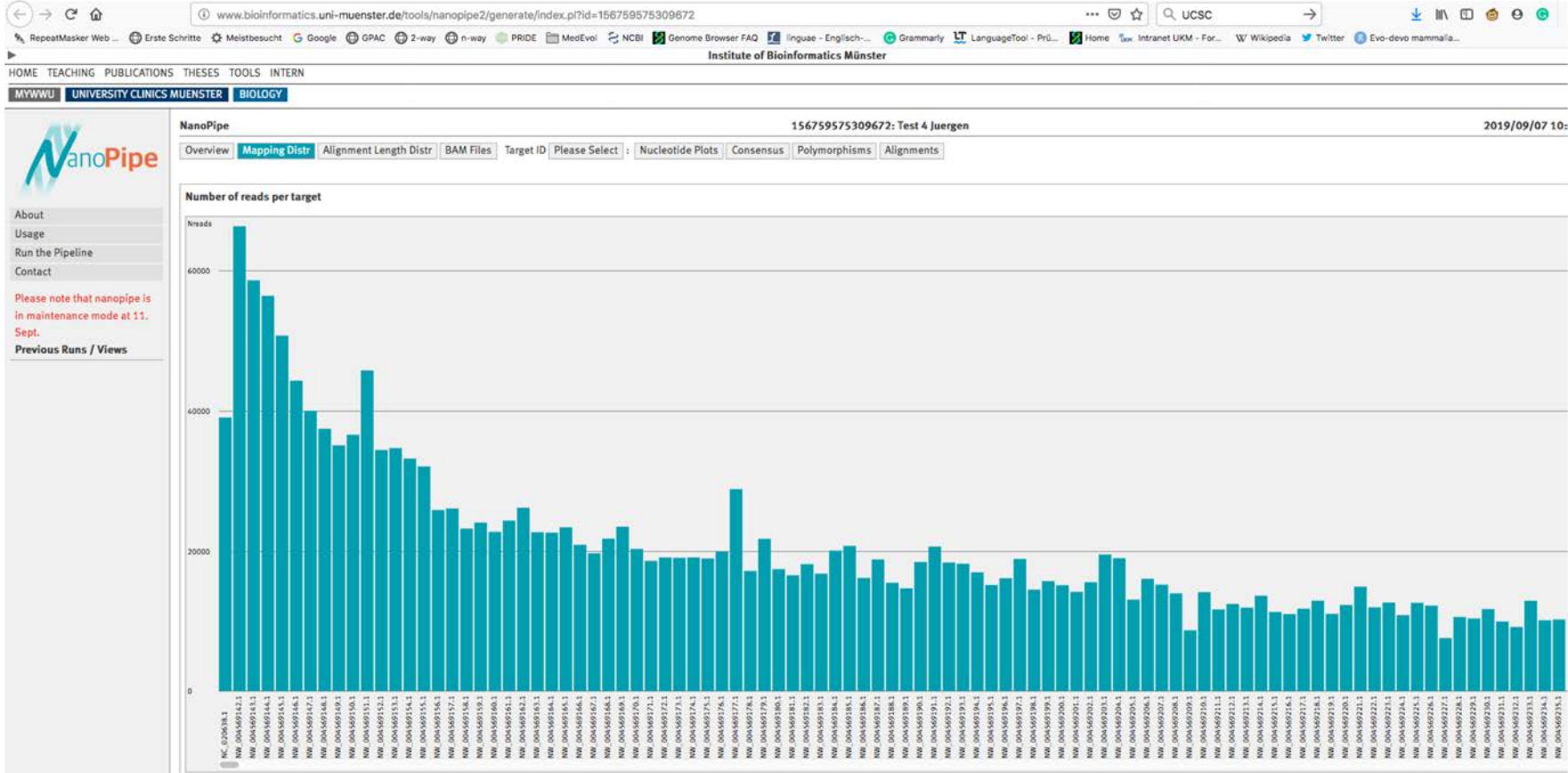
In the newest video on the NCBI YouTube channel [Viral resources at](#)

NCBI to assist Louisiana State University in South and Southeast regional genomics hackathon

MinION analyses



NanoPipe



MinION analyses



alignment
PhyDE 0.994

The character sets

Name	Set Description	Visible	Color
cema	8815-9648	<input checked="" type="checkbox"/>	Red
clpP	9649-10527	<input checked="" type="checkbox"/>	Blue
inFA	10528-10896	<input checked="" type="checkbox"/>	Green
matK	10897-12783	<input checked="" type="checkbox"/>	Yellow
ndhA	12784-13914	<input checked="" type="checkbox"/>	Purple
ndhB	13915-15474	<input checked="" type="checkbox"/>	Pink

alignment view showing sequence alignment for 49 taxa across positions 8930 to 9010. The alignment is color-coded by character set. A yellow box highlights a 'relevant region to discuss' between positions 8930 and 8940. A grey box highlights a 'relevant region to discuss' between positions 8940 and 8950. The taxa list on the left includes: 1 Acoamer, 2 Aocaul, 3 Ambrella, 4 Anethum, 5 Arabidopsis, 6 Atropa, 7 Brassica, 8 Buxus, 9 Calycanthus, 10 Chloranthus, 11 Citrus, 12 Coffea, 13 Cucumis, 14 Cycas, 15 Daucus, 16 Dioscorea, 17 Drimys, 18 Ehretia, 19 Etaeis, 20 Eucalyptus, 21 Ginkgo, 22 Glycine, 23 Gossypium, 24 Helianthus, 25 Hordeum, 26 Illicium, 27 Ipomoea, 28 Jasminum, 29 Lactuca, 30 Litiodendron, 31 Lotus, 32 Medicago, 33 Musa, 34 Nycsyl, 35 Nictab, 36 Nictom, 37 Nuphar, 38 Nymphaea, 39 Oenothera, 40 Oryza, 41 Panax, 42 Passiflora, 43 Pelargonium, 44 Phalaenopsis, 45 Pinus, 46 Piper, 47 Populus, 48 Ranunculus, 49 Saccharum.

taxa: 68 / chars: 85590



UCSC comparative genomics

UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly

chr22:36,226,210-36,239,954 13,745 bp.

UCSC Genome Browser interface showing tracks for RefSeq, RepeatMasker, Phylo-P, and multi-species alignments. The interface includes a search bar, navigation controls, and track options.

move start < 2.0 >

Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position. Press "*" for keyboard shortcuts.

track search default tracks default order hide all add custom tracks track hubs configure multi-region reverse resize refresh

collapse all Use drop-down controls below and press refresh to alter tracks displayed. Tracks with lots of items will automatically be displayed in more compact modes. expand all

Mapping and Sequencing refresh

Base Position dense	P12 Fix Patches pack	P12 Alt Haplotypes pack	P12 Assembly hide	Centromeres hide	P12 Chromosome Band hide
Clone Ends hide	FISH Clones hide	P12 Gap hide	P12 GC Percent hide	GRC Contigs hide	GRC Incident hide
Hg19 Diff hide	P12 INSDC hide	LRG Regions hide	Mappability... hide	P12 RefSeq Acc hide	Restr Enzymes hide
Scaffolds hide	Short Match hide	STS Markers hide			

Genes and Gene Predictions refresh

P12 GENCODE v29 full	NCBI RefSeq hide	P12 Other RefSeq pack	P12 Updated All GENCODE... hide	P12 AUGUSTUS hide	CCDS hide
CRISPR 10K... hide	Gened Genes hide	P12 Genscan Genes hide	KMC Genes Mapped hide	LRG Transcripts hide	MANE select v0.6 hide



● Generate 2-way genome alignments

retrogenomics.uni-muenster.de/tools/twoway/generate/index.pl?

RepeatMasker Web ... Erste Schritte Meistbesucht Google GPAC 2-way n-way PRIDE MedEvol NCBI Genome Browser FAQ lingueae - Englisch-... Grammarly

Retrogenomics WWU Muenster

About Tools Intern

2-Way Genome-Alignments

2-way 2019/09/08 14:54

Previous Requests

ID	The id of a previous request
----	------------------------------

Target Data

Upload file Distribution Fasta file with repeats masked in lowercase letters (max 4GB)

or select clade and species Select a clade... Select a species...

Query Data

Upload file Distribution Fasta file with repeats masked in lowercase letters (max 4GB)

or select clade and species Select a clade... Select a species...

Additional Parameters

Title	<input type="text"/>	Job title
Min Length	100	Min sequence length (minimum 50)

[+] Input Parameters (Click to see/change)

Please note, depending on your data the run can take hours or even days!

© Retrogenomics WWU Muenster



coati



MS_GP



human



Ec3

● RepeatMasking

```
./repeatmasker -nolow -xsmall -pa 10 -species mammalia /home/extern/coati/RM
```

```
./repeatmasker -nolow -xsmall -pa 10 -species rodentia /home/extern/guinea_pig/RM
```

```
./repeatmasker -nolow -xsmall -pa 10 -species human /home/extern/human/RM
```

```
./RepeatFinder.pl -f <input file in multiple fasta format> -i <identification name>
```

FUN
AHEAD

