Aug 22, 2025

Dynamont

DOI
https://dx.doi.org/10.17504/protocols.io.x54v9528pl3e/v1

Jannes Spangenberg[1,2], Christian Höner zu Siederdissen[1,2], Manja Marz[1,2,3,4]

[1]Friedrich Schiller University Jena; [2]RNA Bioinformatics & High-Throughput Analysis;
[3]European Virus Bioinformatics Center 2; [4]FLI Leibniz Institute for Age Research

Dynamont

Jannes Spangenberg

DOI: https://dx.doi.org/10.17504/protocols.io.x54v9528pl3e/v1

Protocol Citation: Jannes Spangenberg, Christian Höner zu Siederdissen, Manja Marz 2025. Dynamont. **protocols.io** https://dx.doi.org/10.17504/protocols.io.x54v9528pl3e/v1

Protocol status: Working
We use this protocol and it's working

Created: August 22, 2025

Last Modified: August 22, 2025

Protocol Integer ID: 225246

Keywords: dynamont this protocol, dynamont, dynamont manuscript, zenodo, dataset

Funders Acknowledgements:
TMWWDG
Grant ID: FKZ5575/10-9
DFG EXC 2051
Grant ID: Project-ID 390713860
SFB1076
Grant ID: 3.A06
BMBF
Grant ID: 01GR2305B.TP7
European Research Council
Grant ID: CoG-101088027

Abstract

This protocol shows the commands used in the Dynamont manuscript to compare the datasets published in zenodo: https://zenodo.org/records/15853348

Dynamont is available on
GitHub: https://github.com/rnajena/dynamont
Conda: https://anaconda.org/jannessp/dynamont
PyPi: **https://pypi.org/project/dynamont/**

Benchmark Preparation Commands

1 # Extract 10.000 Random Reads

```
pod5 view <dataset.pod5> --ids --no-header -o all_ids.txt
sort --random-sort all_ids.txt | head --lines 10000 > 10k_ids.txt
pod5 filter <dataset.pod5> -o <dataset_r10k.pod5> --ids
10k_ids.txt
```

2 # Convert To Other Data Formats

```
blue-crab p2s <dataset_r10k.pod5> -o <dataset_r10k.blow5>
pod5 convert to_fast5 <dataset_r10k.pod5> -o fast5/
multi_to_single_fast5 -i fast5/ -s single_fast5/
```

3 # Basecalling
explicitely using rna002_70bps_hac\@v3 for RNA002 data

```
dorado basecaller sup -x cuda:0 <dataset_r10k.pod5> >
<dataset_r10k.bam>
samtools bam2fq <dataset_r10k.bam> > <dataset_r10k.fastq>
dorado summary <dataset_r10k.bam> > sequencing_summary.txt
```

4 # converting sequencing summary to tombo format (single fast5)

```
awk -F'\t' 'NR == 1 {print; next} {$1 = $2 ".fast5"; print}'
OFS='\t' sequencing_summary.txt > tombo_sequencing_summary.txt
```

5 # Mapping
preset = splice if RNA and h_sapiens, s_cerevisiae, e_coli, sarscov2
preset = lr:hq for DNA R10.4.1
preset = map-ont else
minimap2 -x -a | samtools view -hbF4 | samtools sort >
samtools index

Dynamont Segmentation Commands

6 # model can be added explicitely, otherwise default pore model is chosen

```
python segment.py --raw <path/to/pod5/dataset_r10k/> --basecalls
<dataset_r10k.bam> --mode basic -o <dynamont.csv> --pore <pore>
```

Dorado Segmentation Commands

7 # Basecalling with emit moves
explicitely using rna002_70bps_hac\@v3 for RNA002 data

```
dorado basecaller sup -x cuda:0 --emit-moves <dataset_r10k.pod5> >
<dataset_r10k_moves.bam>
```

8 # Extracting moves as segmentation borders

```
python extractDoradoMoves.py <dataset_r10k_moves.bam> -o
<dataset_r10k_moves.tsv>
```

f5c Segmentation Commands

9 # Indexing files

```
f5c index --slow5 <dataset_r10k.blow5> <dataset_r10k.fastq>
```

10 # Eventalign
added --rna in case of RNA

```
f5c Eventalign -b <dataset_r10k_mapping.bam> -g <ref.fa> -r
<dataset_r10k.fastq> --slow5 <dataset_r10k.blow5> --signal-index -
-collapse-events --pore <pore> --min-mapq 0 --summary
<dataset_r10k_event.sum> > <dataset_r10k_event.tsv>
```

11 # Resquiggle
added --rna in case of RNA

```
f5c Resquiggle --pore <pore> <dataset_r10k.fastq>
<dataset_r10k.blow5> > <dataset_r10k_resqu.tsv>
```

Tombo Segmentation Commands

12 # Indexing files

```
tombo preprocess annotate_raw_with_fastqs --fast5-basedir
single_fast5/ --fastq-filenames <dataset_r10k.fastq> --sequencing-
summary-filenames sequencing_summary.txt
```

13 # only executed on RNA002

```
tombo Resquiggle --q-score 0 --rna single_fast5/ <ref.fa>
```

Uncalled4 Segmentation Commands

14 # preset = splice if RNA and h_sapiens, s_cerevisiae, e_coli, sarscov2
preset = lr:hq for DNA R10.4.1
preset = map-ont else

```
dorado basecaller sup -x cuda:0 --reference <ref.fa> --mm2-opts "-
x <preset> --secondary=no" --emit-moves <dataset_r10k.pod5> >
<mapped_basecalls.bam>
samtools view -hbF 2304 <mapped_basecalls.bam> >
<primary_mapped_basecalls.bam>
uncalled4 align --ref <ref.fa> --reads <dataset_r10k.pod5> --bam-
in <primary_mapped_basecalls.bam> --tsv-out
<uncalled4_segmentation.tsv> --tsv-cols aln.read_id,dtw --min-aln-
length 1
```