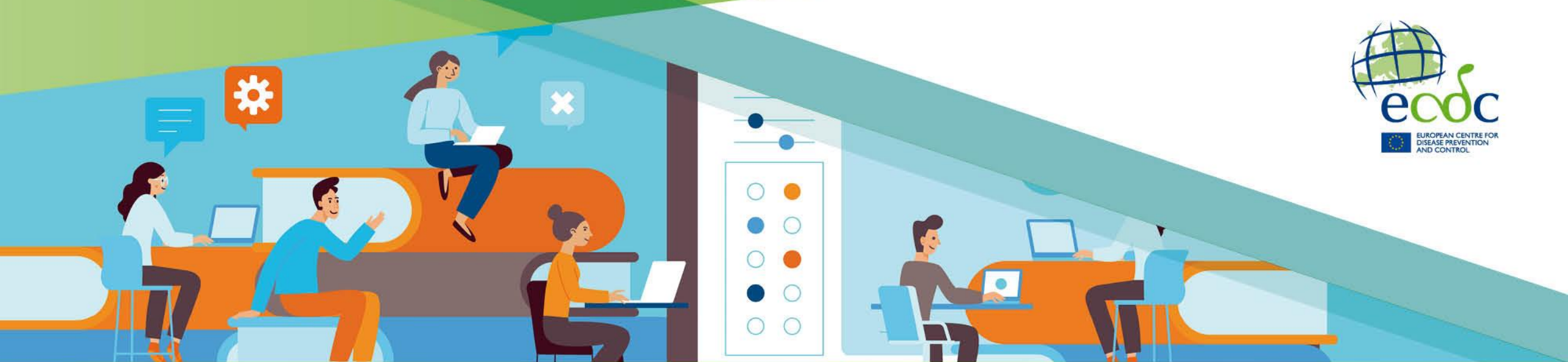




GenEpi-BioTrain Virtual Training 27

# Clinical application of Nanopore for diagnostic 16S rDNA sequencing analysis

2026-06-05

Anna Norén, Elin Loo, Lili Andersson-Li, Lina Guerra Blomqvist, Samuel Lampa, Sofia Stamouli

# Extraction for Oxford Nanopore

- Sample material:  
Cerebrospinal fluids and joint fluids
- Extraction kit:  
TANBead Gram Bacteria DNA
- Pretreatment according to the kit:  
Samples are incubated with  
Proteinase K and Lysozyme  
at **56°C** for **25 min**
- Maelstrom 4810 Automated Nucleic Acid Extractor,  
TANBead

Pipette image By J.N. Eskra - Own work, CC BY-SA 4.0,  
[commons.wikimedia.org/w/index.php?curid=77326666](https://commons.wikimedia.org/w/index.php?curid=77326666)



# Content overview

- Introduction
- Session 1: 16S sequencing background and analysis tools
  - Part 1: Background: 16S biology, Nanopore sequencing
  - Part 2: Open-source tools for ONT 16S sequencing data analysis
  - Intro to exercises
- Session 2: Interpretation of results
  - Part 1: Reporting, alignment quality metrics
  - Part 2: Clinical interpretation
  - Recap of exercises

# Session 2

## Interpretation of results

# Aim and learning objectives

## Aim

This session covers the reporting and interpretation of Nanopore full-length 16S rDNA sequencing results, focusing on how findings are presented and assessed for confidence and clinical relevance.

## Specific learning objectives

At the end of this session, you will be able to:

- Understand how the reporting tools in TranaVy helps guide interpretation
- Understand how to compare sample findings against negative controls and spike-ins to guide interpretation
- Understand how to evaluate confidence in species predictions using probability metrics
- Understand how to interpret raw alignment metrics and their limitations
- Recognize some situations where species-level assignment is uncertain

# Conflict of interest

Declaration of no conflict of interest

# Code of conduct

This training environment is a safe space where we:

- Use respectful language;
- Practice empathy among ourselves;
- Critique ideas, not individuals;
- Respect the right to be silent;
- Favour open dialogue where there are no “stupid” questions;
- Welcome expression of all views in an inclusivity spirit.

ECDC has **zero tolerance policy** for harassment, discrimination and infringement of the [Training Code of conduct](#)



**Why is the 16S rRNA gene good to use for bacterial identification?**





# Why is the 16S rRNA gene good to use for bacterial identification?



**In which situations can 16S sequencing be especially useful compared with traditional culture methods?**





**In which situations can 16S sequencing be especially useful compared with traditional culture methods?**



# What is EMU?





# What is EMU?



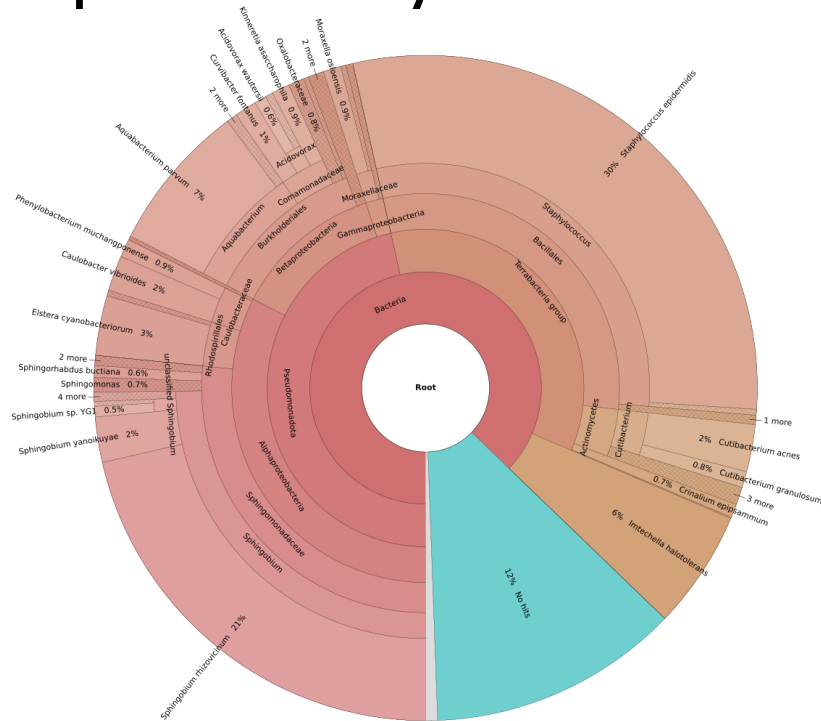
**What is the role of TRANA in Nanopore 16S analysis?**





# What is the role of TRANA in Nanopore 16S analysis?

- Reporting tool for generating static HTML reports from results produced by the TRANA pipeline
- Complementary to KRONA output



## 16S TRANA Report

Sample Name: SAMPLE\_01  
Date: 2026-05-05

Run: RUN\_01  
Pipeline version: TRANA v0.6.0

### Summary Statistics

|                                     |           |
|-------------------------------------|-----------|
| Number of reads before downsampling | 18898     |
| Mean Read Length                    | 1525.3 bp |
| Mean Read Quality                   | 16.5      |
| Median Read Length                  | 1545.0 bp |
| Median Read Quality                 | 17.2      |
| Read Length N50                     | 1548.0 bp |
| STDEV Read Length                   | 78.7 bp   |
| Total Bases                         | 28.8 Mbp  |
| Number of mapped reads              | 18898     |

### Phred Quality Score (Q)

Q10: 10% error rate  
Q20: 1% error rate  
Q30: 0.1% error rate  
Q40: 0.01% error rate

|                         |               |                    |
|-------------------------|---------------|--------------------|
| < Q15<br>Not acceptable | > Q15<br>Good | > Q20<br>Excellent |
|-------------------------|---------------|--------------------|

### Abundance

Species with low abundance (<0.5%) are shown in light grey.

#### SAMPLE\_01

| row | abundance | species                    | genus          | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|----------------------------|----------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae    | 1282    | 5585                  | 99.30%                  | 96.33%                  |
| 2   | 24.39%    | Sphingobium rhizovicinum   | Sphingobium    | Sphingomonadaceae    | 432308  | 4045                  | 98.61%                  | 98.07%                  |
| 3   | 8.27%     | Aquabacterium parvum       | Aquabacterium  | nan                  | 70584   | 1371                  | 98.45%                  | 100.00%                 |
| 4   | 6.58%     | Imtechella halotolerans    | Imtechella     | Flavobacteriaceae    | 1165090 | 1091                  | 98.77%                  | 100.00%                 |
| 5   | 3.30%     | Elstera cyanobacterium     | Elstera        | Rhodospirillaceae    | 2022747 | 548                   | 98.63%                  | 99.57%                  |
| 6   | 2.61%     | Cutibacterium acnes        | Cutibacterium  | Propionibacteriaceae | 1747    | 433                   | 98.78%                  | 95.77%                  |
| 7   | 2.51%     | Sphingobium yanoikuyae     | Sphingobium    | Sphingomonadaceae    | 13690   | 417                   | 98.54%                  | 95.91%                  |
| 8   | 1.96%     | Caulobacter vibrioides     | Caulobacter    | Caulobacteraceae     | 155892  | 325                   | 98.80%                  | 96.63%                  |
| 9   | 1.23%     | Curvibacter fontanus       | Curvibacter    | Comamonadaceae       | 247486  | 203                   | 96.85%                  | 98.72%                  |

Purple rows indicate spike species

Green rows indicate species not found in negative control or with a normalised abundance ratio  $\geq 25$ .

# TranaVy

- Main features
  - Identifiers
  - Summary statistics
  - Abundance tables

## 16S TRANA Report

Sample Name: SAMPLE\_01

Date: 2026-05-05

Run: RUN\_01

Pipeline version: TRANA v0.6.0

### Summary Statistics

|                                     |           |
|-------------------------------------|-----------|
| Number of reads before downsampling | 18898     |
| Mean Read Length                    | 1525.3 bp |
| Mean Read Quality                   | 16.5      |
| Median Read Length                  | 1545.0 bp |
| Median Read Quality                 | 17.2      |
| Read Length N50                     | 1548.0 bp |
| STDEV Read Length                   | 78.7 bp   |
| Total Bases                         | 28.8 Mbp  |
| Number of mapped reads              | 18898     |

### Phred Quality Score (Q)

Q10: 10% error rate  
Q20: 1% error rate  
Q30: 0.1% error rate  
Q40: 0.01% error rate

|                         |               |                    |
|-------------------------|---------------|--------------------|
| < Q15<br>Not acceptable | > Q15<br>Good | > Q20<br>Excellent |
|-------------------------|---------------|--------------------|

### Abundance

Species with low abundance (<0.5%) are shown in light grey.

#### SAMPLE\_01

| row | abundance | species                    | genus          | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|----------------------------|----------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae    | 1282    | 5585                  | 99.30%                  | 96.33%                  |
| 2   | 24.39%    | Sphingobium rhizovincium   | Sphingobium    | Sphingomonadaceae    | 432308  | 4045                  | 98.61%                  | 98.07%                  |
| 3   | 8.27%     | Aquabacterium parvum       | Aquabacterium  | nan                  | 70584   | 1371                  | 98.45%                  | 100.00%                 |
| 4   | 6.58%     | Imtechella halotolerans    | Imtechella     | Flavobacteriaceae    | 1165090 | 1091                  | 98.77%                  | 100.00%                 |
| 5   | 3.30%     | Elstera cyanobacteriorum   | Elstera        | Rhodospirillaceae    | 2022747 | 548                   | 98.63%                  | 99.57%                  |
| 6   | 2.61%     | Cutibacterium acnes        | Cutibacterium  | Propionibacteriaceae | 1747    | 433                   | 98.78%                  | 95.77%                  |
| 7   | 2.51%     | Sphingobium yanoikuyae     | Sphingobium    | Sphingomonadaceae    | 13690   | 417                   | 98.54%                  | 95.91%                  |
| 8   | 1.96%     | Caulobacter vibrioides     | Caulobacter    | Caulobacteraceae     | 155892  | 325                   | 98.80%                  | 96.63%                  |
| 9   | 1.23%     | Curvibacter fontanus       | Curvibacter    | Comamonadaceae       | 247486  | 203                   | 96.85%                  | 98.72%                  |

Purple rows indicate spike species

Green rows indicate species not found in negative control or with a normalised abundance ratio  $\geq 25$ .

#### Negative control

| row | abundance | species                   | genus         | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|---------------------------|---------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Sphingobium rhizovincium  | Sphingobium   | Sphingomonadaceae    | 432308  | 5631                  | 98.66%                  | 98.07%                  |
| 2   | 10.84%    | Aquabacterium parvum      | Aquabacterium | nan                  | 70584   | 1813                  | 98.45%                  | 100.00%                 |
| 3   | 10.80%    | Elstera cyanobacteriorum  | Elstera       | Rhodospirillaceae    | 2022747 | 1807                  | 98.71%                  | 99.57%                  |
| 4   | 7.53%     | Sphingobium yanoikuyae    | Sphingobium   | Sphingomonadaceae    | 13690   | 1260                  | 98.62%                  | 95.98%                  |
| 5   | 5.38%     | Imtechella halotolerans   | Imtechella    | Flavobacteriaceae    | 1165090 | 900                   | 98.85%                  | 100.00%                 |
| 6   | 2.94%     | Caulobacter vibrioides    | Caulobacter   | Caulobacteraceae     | 155892  | 492                   | 98.68%                  | 96.11%                  |
| 7   | 2.59%     | Cutibacterium acnes       | Cutibacterium | Propionibacteriaceae | 1747    | 432                   | 98.79%                  | 95.71%                  |
| 8   | 1.90%     | Curvibacter fontanus      | Curvibacter   | Comamonadaceae       | 247486  | 317                   | 96.77%                  | 98.72%                  |
| 9   | 1.56%     | Kinneretia asaccharophila | Kinneretia    | Comamonadaceae       | 582607  | 260                   | 98.28%                  | 100.00%                 |

Purple rows indicate spike species

# TranaVy

- Main features
  - Identifiers
  - Summary statistics
  - Abundance tables

## 16S TRANA Report

Sample Name: SAMPLE\_01

Date: 2026-05-05

Run: RUN\_01

Pipeline version: TRANA v0.6.0



### Summary Statistics

|                                     |           |
|-------------------------------------|-----------|
| Number of reads before downsampling | 18898     |
| Mean Read Length                    | 1525.3 bp |
| Mean Read Quality                   | 16.5      |
| Median Read Length                  | 1545.0 bp |
| Median Read Quality                 | 17.2      |
| Read Length N50                     | 1548.0 bp |
| STDEV Read Length                   | 78.7 bp   |
| Total Bases                         | 28.8 Mbp  |
| Number of mapped reads              | 18898     |

### Phred Quality Score (Q)

Q10: 10% error rate

Q20: 1% error rate

Q30: 0.1% error rate

Q40: 0.01% error rate

< Q15  
Not acceptable

> Q15  
Good

> Q20  
Excellent

|   |        |                            |                |                      |         |      |        |         |
|---|--------|----------------------------|----------------|----------------------|---------|------|--------|---------|
| 1 | 33.67% | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae    | 1282    | 5585 | 99.30% | 96.33%  |
| 2 | 24.39% | Sphingobium rhizovincium   | Sphingobium    | Sphingomonadaceae    | 432308  | 4045 | 98.61% | 98.07%  |
| 3 | 8.27%  | Aquabacterium parvum       | Aquabacterium  | nan                  | 70584   | 1371 | 98.45% | 100.00% |
| 4 | 6.58%  | Imtechella halotolerans    | Imtechella     | Flavobacteriaceae    | 1165090 | 1091 | 98.77% | 100.00% |
| 5 | 3.30%  | Elstera cyanobacterium     | Elstera        | Rhodospirillaceae    | 2022747 | 548  | 98.63% | 99.57%  |
| 6 | 2.61%  | Cutibacterium acnes        | Cutibacterium  | Propionibacteriaceae | 1747    | 433  | 98.78% | 95.77%  |
| 7 | 2.51%  | Sphingobium yanoikuyae     | Sphingobium    | Sphingomonadaceae    | 13690   | 417  | 98.54% | 95.91%  |
| 8 | 1.96%  | Caulobacter vibrioides     | Caulobacter    | Caulobacteraceae     | 155892  | 325  | 98.80% | 96.63%  |
| 9 | 1.23%  | Curvibacter fontanus       | Curvibacter    | Comamonadaceae       | 247486  | 203  | 96.85% | 98.72%  |

Purple rows indicate spike species

Green rows indicate species not found in negative control or with a normalised abundance ratio  $\geq 25$ .

### Negative control

| row | abundance | species                   | genus         | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|---------------------------|---------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Sphingobium rhizovincium  | Sphingobium   | Sphingomonadaceae    | 432308  | 5631                  | 98.66%                  | 98.07%                  |
| 2   | 10.84%    | Aquabacterium parvum      | Aquabacterium | nan                  | 70584   | 1813                  | 98.45%                  | 100.00%                 |
| 3   | 10.80%    | Elstera cyanobacterium    | Elstera       | Rhodospirillaceae    | 2022747 | 1807                  | 98.71%                  | 99.57%                  |
| 4   | 7.53%     | Sphingobium yanoikuyae    | Sphingobium   | Sphingomonadaceae    | 13690   | 1260                  | 98.62%                  | 95.98%                  |
| 5   | 5.38%     | Imtechella halotolerans   | Imtechella    | Flavobacteriaceae    | 1165090 | 900                   | 98.85%                  | 100.00%                 |
| 6   | 2.94%     | Caulobacter vibrioides    | Caulobacter   | Caulobacteraceae     | 155892  | 492                   | 98.68%                  | 96.11%                  |
| 7   | 2.59%     | Cutibacterium acnes       | Cutibacterium | Propionibacteriaceae | 1747    | 432                   | 98.79%                  | 95.71%                  |
| 8   | 1.90%     | Curvibacter fontanus      | Curvibacter   | Comamonadaceae       | 247486  | 317                   | 96.77%                  | 98.72%                  |
| 9   | 1.56%     | Kinnetetia asaccharophila | Kinnetetia    | Comamonadaceae       | 582607  | 260                   | 98.28%                  | 100.00%                 |

Purple rows indicate spike species

# TranaVy

- Main features
  - Identifiers
  - Summary statistics
  - Abundance tables
- Sample and negative control side-by-side comparison

## 16S TRANA Report



Sample Name: SAMPLE\_01

Date: 2026-05-05

Run: RUN\_01

Pipeline version: TRANA v0.6.0

### Summary Statistics

|                                     |           |
|-------------------------------------|-----------|
| Number of reads before downsampling | 18898     |
| Mean Read Length                    | 1525.3 bp |
| Mean Read Quality                   | 16.5      |
| Median Read Length                  | 1545.0 bp |
| Median Read Quality                 | 17.2      |
| Read Length N50                     | 1548.0 bp |
| STDEV Read Length                   | 78.7 bp   |
| Total Bases                         | 28.8 Mbp  |
| Number of mapped reads              | 18898     |

### Phred Quality Score (Q)

Q10: 10% error rate

Q20: 1% error rate

Q30: 0.1% error rate

Q40: 0.01% error rate

|                         |               |                    |
|-------------------------|---------------|--------------------|
| < Q15<br>Not acceptable | > Q15<br>Good | > Q20<br>Excellent |
|-------------------------|---------------|--------------------|

### Abundance

Species with low abundance (<0.5%) are shown in light grey.

#### SAMPLE\_01

| row | abundance | species                    | genus          | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|----------------------------|----------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae    | 1282    | 5585                  | 99.30%                  | 96.33%                  |
| 2   | 24.39%    | Sphingobium rhizovincium   | Sphingobium    | Sphingomonadaceae    | 432308  | 4045                  | 98.61%                  | 98.07%                  |
| 3   | 8.27%     | Aquabacterium parvum       | Aquabacterium  | nan                  | 70584   | 1371                  | 98.45%                  | 100.00%                 |
| 4   | 6.58%     | Imtechella halotolerans    | Imtechella     | Flavobacteriaceae    | 1165090 | 1091                  | 98.77%                  | 100.00%                 |
| 5   | 3.30%     | Elstera cyanobacterium     | Elstera        | Rhodospirillaceae    | 2022747 | 548                   | 98.63%                  | 99.57%                  |
| 6   | 2.61%     | Cutibacterium acnes        | Cutibacterium  | Propionibacteriaceae | 1747    | 433                   | 98.78%                  | 95.77%                  |
| 7   | 2.51%     | Sphingobium yanoikuyae     | Sphingobium    | Sphingomonadaceae    | 13690   | 417                   | 98.54%                  | 95.91%                  |
| 8   | 1.96%     | Caulobacter vibrioides     | Caulobacter    | Caulobacteraceae     | 155892  | 325                   | 98.80%                  | 96.63%                  |
| 9   | 1.23%     | Curvibacter fontanus       | Curvibacter    | Comamonadaceae       | 247486  | 203                   | 96.85%                  | 98.72%                  |

Purple rows indicate spike species

Green rows indicate species not found in negative control or with a normalised abundance ratio  $\geq 25$ .

#### Negative control

| row | abundance | species                   | genus         | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|---------------------------|---------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Sphingobium rhizovincium  | Sphingobium   | Sphingomonadaceae    | 432308  | 5631                  | 98.66%                  | 98.07%                  |
| 2   | 10.84%    | Aquabacterium parvum      | Aquabacterium | nan                  | 70584   | 1813                  | 98.45%                  | 100.00%                 |
| 3   | 10.80%    | Elstera cyanobacterium    | Elstera       | Rhodospirillaceae    | 2022747 | 1807                  | 98.71%                  | 99.57%                  |
| 4   | 7.53%     | Sphingobium yanoikuyae    | Sphingobium   | Sphingomonadaceae    | 13690   | 1260                  | 98.62%                  | 95.98%                  |
| 5   | 5.38%     | Imtechella halotolerans   | Imtechella    | Flavobacteriaceae    | 1165090 | 900                   | 98.85%                  | 100.00%                 |
| 6   | 2.94%     | Caulobacter vibrioides    | Caulobacter   | Caulobacteraceae     | 155892  | 492                   | 98.68%                  | 96.11%                  |
| 7   | 2.59%     | Cutibacterium acnes       | Cutibacterium | Propionibacteriaceae | 1747    | 432                   | 98.79%                  | 95.71%                  |
| 8   | 1.90%     | Curvibacter fontanus      | Curvibacter   | Comamonadaceae       | 247486  | 317                   | 96.77%                  | 98.72%                  |
| 9   | 1.56%     | Kinneretia asaccharophila | Kinneretia    | Comamonadaceae       | 582607  | 260                   | 98.28%                  | 100.00%                 |

Purple rows indicate spike species

# TranaVy

- Main features
  - Identifiers
  - Summary statistics
  - Abundance tables
- Sample and negative control side-by-side comparison

## 16S TRANA Report



Sample Name: SAMPLE\_01

Date: 2026-05-05

Run: RUN\_01

Pipeline version: TRANA v0.6.0

### Summary Statistics

|                                     |           |
|-------------------------------------|-----------|
| Number of reads before downsampling | 18898     |
| Mean Read Length                    | 1525.3 bp |
| Mean Read Quality                   | 16.5      |
| Median Read Length                  | 1545.0 bp |
| Median Read Quality                 | 17.2      |
| Read Length N50                     | 1548.0 bp |
| STDEV Read Length                   | 78.7 bp   |
| Total Bases                         | 28.8 Mbp  |
| Number of mapped reads              | 18898     |

### Phred Quality Score (Q)

Q10: 10% error rate  
Q20: 1% error rate  
Q30: 0.1% error rate  
Q40: 0.01% error rate

|                         |               |                    |
|-------------------------|---------------|--------------------|
| < Q15<br>Not acceptable | > Q15<br>Good | > Q20<br>Excellent |
|-------------------------|---------------|--------------------|

### Abundance

Species with low abundance (<0.5%) are shown in light grey.

#### SAMPLE\_01

| row | abundance | species                    | genus          | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|----------------------------|----------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae    | 1282    | 5585                  | 99.30%                  | 96.33%                  |
| 2   | 24.39%    | Sphingobium rhizovincium   | Sphingobium    | Sphingomonadaceae    | 432308  | 4045                  | 98.61%                  | 98.07%                  |
| 3   | 8.27%     | Aquabacterium parvum       | Aquabacterium  | nan                  | 70584   | 1371                  | 98.45%                  | 100.00%                 |
| 4   | 6.58%     | Imtechella halotolerans    | Imtechella     | Flavobacteriaceae    | 1165090 | 1091                  | 98.77%                  | 100.00%                 |
| 5   | 3.30%     | Elstera cyanobacterium     | Elstera        | Rhodospirillaceae    | 2022747 | 548                   | 98.63%                  | 99.57%                  |
| 6   | 2.61%     | Cutibacterium acnes        | Cutibacterium  | Propionibacteriaceae | 1747    | 433                   | 98.78%                  | 95.77%                  |
| 7   | 2.51%     | Sphingobium yanoikuyae     | Sphingobium    | Sphingomonadaceae    | 13690   | 417                   | 98.54%                  | 95.91%                  |
| 8   | 1.96%     | Caulobacter vibrioides     | Caulobacter    | Caulobacteraceae     | 155892  | 325                   | 98.80%                  | 96.63%                  |
| 9   | 1.23%     | Curvibacter fontanus       | Curvibacter    | Comamonadaceae       | 247486  | 203                   | 96.85%                  | 98.72%                  |

Purple rows indicate spike species

Green rows indicate species not found in negative control or with a normalised abundance ratio  $\geq 25$ .

#### Negative control

| row | abundance | species                   | genus         | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|---------------------------|---------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Sphingobium rhizovincium  | Sphingobium   | Sphingomonadaceae    | 432308  | 5631                  | 98.66%                  | 98.07%                  |
| 2   | 10.84%    | Aquabacterium parvum      | Aquabacterium | nan                  | 70584   | 1813                  | 98.45%                  | 100.00%                 |
| 3   | 10.80%    | Elstera cyanobacterium    | Elstera       | Rhodospirillaceae    | 2022747 | 1807                  | 98.71%                  | 99.57%                  |
| 4   | 7.53%     | Sphingobium yanoikuyae    | Sphingobium   | Sphingomonadaceae    | 13690   | 1260                  | 98.62%                  | 95.98%                  |
| 5   | 5.38%     | Imtechella halotolerans   | Imtechella    | Flavobacteriaceae    | 1165090 | 900                   | 98.85%                  | 100.00%                 |
| 6   | 2.94%     | Caulobacter vibrioides    | Caulobacter   | Caulobacteraceae     | 155892  | 492                   | 98.68%                  | 96.11%                  |
| 7   | 2.59%     | Cutibacterium acnes       | Cutibacterium | Propionibacteriaceae | 1747    | 432                   | 98.79%                  | 95.71%                  |
| 8   | 1.90%     | Curvibacter fontanus      | Curvibacter   | Comamonadaceae       | 247486  | 317                   | 96.77%                  | 98.72%                  |
| 9   | 1.56%     | Kinneretia asaccharophila | Kinneretia    | Comamonadaceae       | 582607  | 260                   | 98.28%                  | 100.00%                 |

Purple rows indicate spike species

Spike species

# TranaVy

- Main features
  - Identifiers
  - Summary statistics
  - Abundance tables
- Sample and negative control side-by-side comparison

## 16S TRANA Report



Sample Name: SAMPLE\_01

Date: 2026-05-05

Run: RUN\_01

Pipeline version: TRANA v0.6.0

### Summary Statistics

|                                     |           |
|-------------------------------------|-----------|
| Number of reads before downsampling | 18898     |
| Mean Read Length                    | 1525.3 bp |
| Mean Read Quality                   | 16.5      |
| Median Read Length                  | 1545.0 bp |
| Median Read Quality                 | 17.2      |
| Read Length N50                     | 1548.0 bp |
| STDEV Read Length                   | 78.7 bp   |
| Total Bases                         | 28.8 Mbp  |
| Number of mapped reads              | 18898     |

### Phred Quality Score (Q)

Q10: 10% error rate  
Q20: 1% error rate  
Q30: 0.1% error rate  
Q40: 0.01% error rate

|                         |               |                    |
|-------------------------|---------------|--------------------|
| < Q15<br>Not acceptable | > Q15<br>Good | > Q20<br>Excellent |
|-------------------------|---------------|--------------------|

### Abundance

Species with low abundance (<0.5%) are shown in light grey.

SAMPLE\_01

| row | abundance | species                    | genus          | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|----------------------------|----------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae    | 1282    | 5585                  | 99.30%                  | 96.33%                  |
| 2   | 24.39%    | Sphingobium rhizovincinum  | Sphingobium    | Sphingomonadaceae    | 432308  | 4045                  | 98.61%                  | 98.07%                  |
| 3   | 8.27%     | Aquabacterium parvum       | Aquabacterium  | nan                  | 70584   | 1371                  | 98.45%                  | 100.00%                 |
| 4   | 6.58%     | Imtechella halotolerans    | Imtechella     | Flavobacteriaceae    | 1165090 | 1091                  | 98.77%                  | 100.00%                 |
| 5   | 3.30%     | Elstera cyanobacteriorum   | Elstera        | Rhodospirillaceae    | 2022747 | 548                   | 98.63%                  | 99.57%                  |
| 6   | 2.61%     | Cutibacterium acnes        | Cutibacterium  | Propionibacteriaceae | 1747    | 433                   | 98.78%                  | 95.77%                  |
| 7   | 2.51%     | Sphingobium yanoikuyae     | Sphingobium    | Sphingomonadaceae    | 13690   | 417                   | 98.54%                  | 95.91%                  |
| 8   | 1.96%     | Caulobacter vibrioides     | Caulobacter    | Caulobacteraceae     | 155892  | 325                   | 98.80%                  | 96.63%                  |
| 9   | 1.23%     | Curvibacter fontanus       | Curvibacter    | Comamonadaceae       | 247486  | 203                   | 96.85%                  | 98.72%                  |

Purple rows indicate spike species

Green rows indicate species not found in negative control or with a normalised abundance ratio  $\geq 25$ .

### Negative control

| row | abundance | species                   | genus         | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|---------------------------|---------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Sphingobium rhizovincinum | Sphingobium   | Sphingomonadaceae    | 432308  | 5631                  | 98.66%                  | 98.07%                  |
| 2   | 10.84%    | Aquabacterium parvum      | Aquabacterium | nan                  | 70584   | 1813                  | 98.45%                  | 100.00%                 |
| 3   | 10.80%    | Elstera cyanobacteriorum  | Elstera       | Rhodospirillaceae    | 2022747 | 1807                  | 98.71%                  | 99.57%                  |
| 4   | 7.53%     | Sphingobium yanoikuyae    | Sphingobium   | Sphingomonadaceae    | 13690   | 1260                  | 98.62%                  | 95.98%                  |
| 5   | 5.38%     | Imtechella halotolerans   | Imtechella    | Flavobacteriaceae    | 1165090 | 900                   | 98.85%                  | 100.00%                 |
| 6   | 2.94%     | Caulobacter vibrioides    | Caulobacter   | Caulobacteraceae     | 155892  | 492                   | 98.68%                  | 96.11%                  |
| 7   | 2.59%     | Cutibacterium acnes       | Cutibacterium | Propionibacteriaceae | 1747    | 432                   | 98.79%                  | 95.71%                  |
| 8   | 1.90%     | Curvibacter fontanus      | Curvibacter   | Comamonadaceae       | 247486  | 317                   | 96.77%                  | 98.72%                  |
| 9   | 1.56%     | Kinneretia asaccharophila | Kinneretia    | Comamonadaceae       | 582607  | 260                   | 98.28%                  | 100.00%                 |

Purple rows indicate spike species

Significant species

- Normalized abundance ratio

$$\text{Normalized abundance ratio} = \frac{\text{Abundance}_{\text{species}_s}}{\text{Abundance}_{\text{spike}_s}} \bigg/ \frac{\text{Abundance}_{\text{species}_c}}{\text{Abundance}_{\text{spike}_c}}$$

Where *\_s* signify in sample and *\_c* signify in control



16S TRANA Report

Sample Name: SAMPLE\_01

Date: 2026-05-05

Run: RUN\_01

Pipeline version: TRANA v0.6.0

Summary Statistics

|                                     |           |
|-------------------------------------|-----------|
| Number of reads before downsampling | 18898     |
| Mean Read Length                    | 1525.3 bp |
| Mean Read Quality                   | 16.5      |
| Median Read Length                  | 1545.0 bp |
| Median Read Quality                 | 17.2      |
| Read Length N50                     | 1548.0 bp |
| STDEV Read Length                   | 78.7 bp   |
| Total Bases                         | 28.8 Mbp  |
| Number of mapped reads              | 18898     |

Phred Quality Score (Q)

Q10: 10% error rate

Q20: 1% error rate

Q30: 0.1% error rate

Q40: 0.01% error rate

< Q15

> Q15

> Q20

Not acceptable

Good

Excellent

Abundance

Species with low abundance (<0.5%) are shown in light grey.

SAMPLE\_01

| row | abundance | species                    | genus          | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|----------------------------|----------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae    | 1282    | 5585                  | 99.30%                  | 96.33%                  |
| 2   | 24.39%    | Sphingobium rhizovicinum   | Sphingobium    | Sphingomonadaceae    | 432308  | 4045                  | 98.61%                  | 98.07%                  |
| 3   | 8.27%     | Aquabacterium parvum       | Aquabacterium  | nan                  | 70584   | 1371                  | 98.45%                  | 100.00%                 |
| 4   | 6.58%     | Imtechella halotolerans    | Imtechella     | Flavobacteriaceae    | 1165090 | 1091                  | 98.77%                  | 100.00%                 |
| 5   | 3.30%     | Elstera cyanobacteriorum   | Elstera        | Rhodospirillaceae    | 2022747 | 548                   | 98.63%                  | 99.57%                  |
| 6   | 2.61%     | Cutibacterium acnes        | Cutibacterium  | Propionibacteriaceae | 1747    | 433                   | 98.78%                  | 95.77%                  |
| 7   | 2.51%     | Sphingobium yanoikuyae     | Sphingobium    | Sphingomonadaceae    | 13690   | 417                   | 98.54%                  | 95.91%                  |
| 8   | 1.96%     | Caulobacter vibrioides     | Caulobacter    | Caulobacteraceae     | 155892  | 325                   | 98.80%                  | 96.63%                  |
| 9   | 1.23%     | Curvibacter fontanus       | Curvibacter    | Comamonadaceae       | 247486  | 203                   | 96.85%                  | 98.72%                  |

Purple rows indicate spike species

Green rows indicate species not found in negative control or with a normalised abundance ratio ≥ 25.

Negative control

| row | abundance | species                   | genus         | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|---------------------------|---------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Sphingobium rhizovicinum  | Sphingobium   | Sphingomonadaceae    | 432308  | 5631                  | 98.66%                  | 98.07%                  |
| 2   | 10.84%    | Aquabacterium parvum      | Aquabacterium | nan                  | 70584   | 1813                  | 98.45%                  | 100.00%                 |
| 3   | 10.80%    | Elstera cyanobacteriorum  | Elstera       | Rhodospirillaceae    | 2022747 | 1807                  | 98.71%                  | 99.57%                  |
| 4   | 7.53%     | Sphingobium yanoikuyae    | Sphingobium   | Sphingomonadaceae    | 13690   | 1260                  | 98.62%                  | 95.98%                  |
| 5   | 5.38%     | Imtechella halotolerans   | Imtechella    | Flavobacteriaceae    | 1165090 | 900                   | 98.85%                  | 100.00%                 |
| 6   | 2.94%     | Caulobacter vibrioides    | Caulobacter   | Caulobacteraceae     | 155892  | 492                   | 98.68%                  | 96.11%                  |
| 7   | 2.59%     | Cutibacterium acnes       | Cutibacterium | Propionibacteriaceae | 1747    | 432                   | 98.79%                  | 95.71%                  |
| 8   | 1.90%     | Curvibacter fontanus      | Curvibacter   | Comamonadaceae       | 247486  | 317                   | 96.77%                  | 98.72%                  |
| 9   | 1.56%     | Kinneretia asaccharophila | Kinneretia    | Comamonadaceae       | 582607  | 260                   | 98.28%                  | 100.00%                 |

Purple rows indicate spike species

# TranaVy

- Normalized abundance ratio

$$\text{Normalized abundance ratio} = \frac{\frac{\text{Abundance}_{\text{species}_s}}{\text{Abundance}_{\text{spike}_s}}}{\frac{\text{Abundance}_{\text{species}_c}}{\text{Abundance}_{\text{spike}_c}}}$$

Where  $_s$  signify in sample and  $_c$  signify in control

$$\text{Normalized abundance ratio} = \frac{33.67}{6.58} \div \frac{?}{?} = ?$$

## 16S TRANA Report

Sample Name: SAMPLE\_01

Date: 2026-05-05

Run: RUN\_01

Pipeline version: TRANA v0.6.0

### Summary Statistics

|                                     |           |
|-------------------------------------|-----------|
| Number of reads before downsampling | 18898     |
| Mean Read Length                    | 1525.3 bp |
| Mean Read Quality                   | 16.5      |
| Median Read Length                  | 1545.0 bp |
| Median Read Quality                 | 17.2      |
| Read Length N50                     | 1548.0 bp |
| STDEV Read Length                   | 78.7 bp   |
| Total Bases                         | 28.8 Mbp  |
| Number of mapped reads              | 18898     |

### Phred Quality Score (Q)

Q10: 10% error rate  
Q20: 1% error rate  
Q30: 0.1% error rate  
Q40: 0.01% error rate

|                         |               |                    |
|-------------------------|---------------|--------------------|
| < Q15<br>Not acceptable | > Q15<br>Good | > Q20<br>Excellent |
|-------------------------|---------------|--------------------|

### Abundance

Species with low abundance (<0.5%) are shown in light grey.

#### SAMPLE\_01

| row | abundance | species                    | genus          | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|----------------------------|----------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae    | 1282    | 5585                  | 99.30%                  | 96.33%                  |
| 2   | 24.39%    | Sphingobium rhizovicinum   | Sphingobium    | Sphingomonadaceae    | 432308  | 4045                  | 98.61%                  | 98.07%                  |
| 3   | 8.27%     | Aquabacterium parvum       | Aquabacterium  | nan                  | 70584   | 1371                  | 98.45%                  | 100.00%                 |
| 4   | 6.58%     | Intechella halotolerans    | Intechella     | Flavobacteriaceae    | 1165090 | 1091                  | 98.77%                  | 100.00%                 |
| 5   | 3.30%     | Elstera cyanobacterium     | Elstera        | Rhodospirillaceae    | 2022747 | 548                   | 98.63%                  | 99.57%                  |
| 6   | 2.61%     | Cutibacterium acnes        | Cutibacterium  | Propionibacteriaceae | 1747    | 433                   | 98.78%                  | 95.77%                  |
| 7   | 2.51%     | Sphingobium yanoikuyae     | Sphingobium    | Sphingomonadaceae    | 13690   | 417                   | 98.54%                  | 95.91%                  |
| 8   | 1.96%     | Caulobacter vibrioides     | Caulobacter    | Caulobacteraceae     | 155892  | 325                   | 98.80%                  | 96.63%                  |
| 9   | 1.23%     | Curvibacter fontanus       | Curvibacter    | Comamonadaceae       | 247486  | 203                   | 96.85%                  | 98.72%                  |

Purple rows indicate spike species

Green rows indicate species not found in negative control or with a normalised abundance ratio  $\geq 25$ .

#### Negative control

| row | abundance | species                   | genus         | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|---------------------------|---------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Sphingobium rhizovicinum  | Sphingobium   | Sphingomonadaceae    | 432308  | 5631                  | 98.66%                  | 98.07%                  |
| 2   | 10.84%    | Aquabacterium parvum      | Aquabacterium | nan                  | 70584   | 1813                  | 98.45%                  | 100.00%                 |
| 3   | 10.80%    | Elstera cyanobacterium    | Elstera       | Rhodospirillaceae    | 2022747 | 1807                  | 98.71%                  | 99.57%                  |
| 4   | 7.53%     | Sphingobium yanoikuyae    | Sphingobium   | Sphingomonadaceae    | 13690   | 1260                  | 98.62%                  | 95.98%                  |
| 5   | 5.38%     | Intechella halotolerans   | Intechella    | Flavobacteriaceae    | 1165090 | 900                   | 98.85%                  | 100.00%                 |
| 6   | 2.94%     | Caulobacter vibrioides    | Caulobacter   | Caulobacteraceae     | 155892  | 492                   | 98.68%                  | 96.11%                  |
| 7   | 2.59%     | Cutibacterium acnes       | Cutibacterium | Propionibacteriaceae | 1747    | 432                   | 98.79%                  | 95.71%                  |
| 8   | 1.90%     | Curvibacter fontanus      | Curvibacter   | Comamonadaceae       | 247486  | 317                   | 96.77%                  | 98.72%                  |
| 9   | 1.56%     | Kinneretia asaccharophila | Kinneretia    | Comamonadaceae       | 582607  | 260                   | 98.28%                  | 100.00%                 |

Purple rows indicate spike species

# TranaVy

- Normalized abundance ratio
  - Cut-off value: 25

$$\text{Normalized abundance ratio} = \frac{\frac{\text{Abundance}_{\text{species}_s}}{\text{Abundance}_{\text{spike}_s}}}{\frac{\text{Abundance}_{\text{species}_c}}{\text{Abundance}_{\text{spike}_c}}}$$

Where *\_s* signify in sample and *\_c* signify in control

$$\text{Normalized abundance ratio} = \frac{33.67}{6.58} \div \frac{0.47}{5.38} = 58.57$$

|    |       |                            |                |                   |      |    |        |        |
|----|-------|----------------------------|----------------|-------------------|------|----|--------|--------|
| 29 | 0.47% | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae | 1282 | 79 | 99.07% | 96.34% |
|----|-------|----------------------------|----------------|-------------------|------|----|--------|--------|

## 16S TRANA Report



Sample Name: SAMPLE\_01

Date: 2026-05-05

Run: RUN\_01

Pipeline version: TRANA v0.6.0

### Summary Statistics

|                                     |           |
|-------------------------------------|-----------|
| Number of reads before downsampling | 18898     |
| Mean Read Length                    | 1525.3 bp |
| Mean Read Quality                   | 16.5      |
| Median Read Length                  | 1545.0 bp |
| Median Read Quality                 | 17.2      |
| Read Length N50                     | 1548.0 bp |
| STDEV Read Length                   | 78.7 bp   |
| Total Bases                         | 28.8 Mbp  |
| Number of mapped reads              | 18898     |

### Phred Quality Score (Q)

Q10: 10% error rate  
Q20: 1% error rate  
Q30: 0.1% error rate  
Q40: 0.01% error rate

|                         |               |                    |
|-------------------------|---------------|--------------------|
| < Q15<br>Not acceptable | > Q15<br>Good | > Q20<br>Excellent |
|-------------------------|---------------|--------------------|

### Abundance

Species with low abundance (<0.5%) are shown in light grey.

SAMPLE\_01

| row | abundance | species                    | genus          | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|----------------------------|----------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae    | 1282    | 5585                  | 99.30%                  | 96.33%                  |
| 2   | 24.39%    | Sphingobium rhizovicium    | Sphingobium    | Sphingomonadaceae    | 432308  | 4045                  | 98.61%                  | 98.07%                  |
| 3   | 8.27%     | Aquabacterium parvum       | Aquabacterium  | nan                  | 70584   | 1371                  | 98.45%                  | 100.00%                 |
| 4   | 6.58%     | Intechella halotolerans    | Intechella     | Flavobacteriaceae    | 1165090 | 1091                  | 98.77%                  | 100.00%                 |
| 5   | 3.30%     | Elstera cyanobacterium     | Elstera        | Rhodospirillaceae    | 2022747 | 548                   | 98.63%                  | 99.57%                  |
| 6   | 2.61%     | Cutibacterium acnes        | Cutibacterium  | Propionibacteriaceae | 1747    | 433                   | 98.78%                  | 95.77%                  |
| 7   | 2.51%     | Sphingobium yanoikuyae     | Sphingobium    | Sphingomonadaceae    | 13690   | 417                   | 98.54%                  | 95.91%                  |
| 8   | 1.96%     | Caulobacter vibrioides     | Caulobacter    | Caulobacteraceae     | 155892  | 325                   | 98.80%                  | 96.63%                  |
| 9   | 1.23%     | Curvibacter fontanus       | Curvibacter    | Comamonadaceae       | 247486  | 203                   | 96.85%                  | 98.72%                  |

Purple rows indicate spike species

Green rows indicate species not found in negative control or with a normalised abundance ratio ≥ 25.

### Negative control

| row | abundance | species                   | genus         | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|---------------------------|---------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Sphingobium rhizovicium   | Sphingobium   | Sphingomonadaceae    | 432308  | 5631                  | 98.66%                  | 98.07%                  |
| 2   | 10.84%    | Aquabacterium parvum      | Aquabacterium | nan                  | 70584   | 1813                  | 98.45%                  | 100.00%                 |
| 3   | 10.80%    | Elstera cyanobacterium    | Elstera       | Rhodospirillaceae    | 2022747 | 1807                  | 98.71%                  | 99.57%                  |
| 4   | 7.53%     | Sphingobium yanoikuyae    | Sphingobium   | Sphingomonadaceae    | 13690   | 1260                  | 98.62%                  | 95.98%                  |
| 5   | 5.38%     | Intechella halotolerans   | Intechella    | Flavobacteriaceae    | 1165090 | 900                   | 98.85%                  | 100.00%                 |
| 6   | 2.94%     | Caulobacter vibrioides    | Caulobacter   | Caulobacteraceae     | 155892  | 492                   | 98.68%                  | 96.11%                  |
| 7   | 2.59%     | Cutibacterium acnes       | Cutibacterium | Propionibacteriaceae | 1747    | 432                   | 98.79%                  | 95.71%                  |
| 8   | 1.90%     | Curvibacter fontanus      | Curvibacter   | Comamonadaceae       | 247486  | 317                   | 98.77%                  | 98.72%                  |
| 9   | 1.56%     | Kinnetetia asaccharophila | Kinnetetia    | Comamonadaceae       | 582607  | 260                   | 98.28%                  | 100.00%                 |

Purple rows indicate spike species

- Customization
  - Spike species
  - Cut-off value for normalized abundance ratio
  - Optional metrics:
    - probability-scores, -p
    - alignment-metrics, -m

## 16S TRANA Report

Sample Name: SAMPLE\_01

Date: 2026-05-05

Run: RUN\_01

Pipeline version: TRANA v0.6.0

### Summary Statistics

|                                     |           |
|-------------------------------------|-----------|
| Number of reads before downsampling | 18898     |
| Mean Read Length                    | 1525.3 bp |
| Mean Read Quality                   | 16.5      |
| Median Read Length                  | 1545.0 bp |
| Median Read Quality                 | 17.2      |
| Read Length N50                     | 1548.0 bp |
| STDEV Read Length                   | 78.7 bp   |
| Total Bases                         | 28.8 Mbp  |
| Number of mapped reads              | 18898     |

### Phred Quality Score (Q)

Q10: 10% error rate  
Q20: 1% error rate  
Q30: 0.1% error rate  
Q40: 0.01% error rate

|                         |               |                    |
|-------------------------|---------------|--------------------|
| < Q15<br>Not acceptable | > Q15<br>Good | > Q20<br>Excellent |
|-------------------------|---------------|--------------------|

### Abundance

Species with low abundance (<0.5%) are shown in light grey.

#### SAMPLE\_01

| row | abundance | species                    | genus          | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|----------------------------|----------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae    | 1282    | 5585                  | 99.30%                  | 96.33%                  |
| 2   | 24.39%    | Sphingobium rhizovincium   | Sphingobium    | Sphingomonadaceae    | 432308  | 4045                  | 98.61%                  | 98.07%                  |
| 3   | 8.27%     | Aquabacterium parvum       | Aquabacterium  | nan                  | 70584   | 1371                  | 98.45%                  | 100.00%                 |
| 4   | 6.58%     | Imtechella halotolerans    | Imtechella     | Flavobacteriaceae    | 1165090 | 1091                  | 98.77%                  | 100.00%                 |
| 5   | 3.30%     | Elstera cyanobacterium     | Elstera        | Rhodospirillaceae    | 2022747 | 548                   | 98.63%                  | 99.57%                  |
| 6   | 2.61%     | Cutibacterium acnes        | Cutibacterium  | Propionibacteriaceae | 1747    | 433                   | 98.78%                  | 95.77%                  |
| 7   | 2.51%     | Sphingobium yanoikuyae     | Sphingobium    | Sphingomonadaceae    | 13690   | 417                   | 98.54%                  | 95.91%                  |
| 8   | 1.96%     | Caulobacter vibrioides     | Caulobacter    | Caulobacteraceae     | 155892  | 325                   | 98.80%                  | 96.63%                  |
| 9   | 1.23%     | Curvibacter fontanus       | Curvibacter    | Comamonadaceae       | 247486  | 203                   | 96.85%                  | 98.72%                  |

Purple rows indicate spike species

Green rows indicate species not found in negative control or with a normalised abundance ratio  $\geq 25$ .

#### Negative control

| row | abundance | species                   | genus         | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|---------------------------|---------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Sphingobium rhizovincium  | Sphingobium   | Sphingomonadaceae    | 432308  | 5631                  | 98.66%                  | 98.07%                  |
| 2   | 10.84%    | Aquabacterium parvum      | Aquabacterium | nan                  | 70584   | 1813                  | 98.45%                  | 100.00%                 |
| 3   | 10.80%    | Elstera cyanobacterium    | Elstera       | Rhodospirillaceae    | 2022747 | 1807                  | 98.71%                  | 99.57%                  |
| 4   | 7.53%     | Sphingobium yanoikuyae    | Sphingobium   | Sphingomonadaceae    | 13690   | 1260                  | 98.62%                  | 95.98%                  |
| 5   | 5.38%     | Imtechella halotolerans   | Imtechella    | Flavobacteriaceae    | 1165090 | 900                   | 98.85%                  | 100.00%                 |
| 6   | 2.94%     | Caulobacter vibrioides    | Caulobacter   | Caulobacteraceae     | 155892  | 492                   | 98.68%                  | 96.11%                  |
| 7   | 2.59%     | Cutibacterium acnes       | Cutibacterium | Propionibacteriaceae | 1747    | 432                   | 98.79%                  | 95.71%                  |
| 8   | 1.90%     | Curvibacter fontanus      | Curvibacter   | Comamonadaceae       | 247486  | 317                   | 96.77%                  | 98.72%                  |
| 9   | 1.56%     | Kinneretia asaccharophila | Kinneretia    | Comamonadaceae       | 582607  | 260                   | 98.28%                  | 100.00%                 |

Purple rows indicate spike species

# TranaVy

- Customization
  - Spike species
  - Cut-off value for normalized abundance ratio
  - Optional metrics:
    - probability-scores, -p
    - alignment-metrics, -m

## 16S TRANA Report

Sample Name: SAMPLE\_01

Date: 2026-05-05

Run: RUN\_01

Pipeline version: TRANA v0.6.0

### Summary Statistics

|                                     |           |
|-------------------------------------|-----------|
| Number of reads before downsampling | 18898     |
| Mean Read Length                    | 1525.3 bp |
| Mean Read Quality                   | 16.5      |
| Median Read Length                  | 1545.0 bp |
| Median Read Quality                 | 17.2      |
| Read Length N50                     | 1548.0 bp |
| STDEV Read Length                   | 78.7 bp   |
| Total Bases                         | 28.8 Mbp  |
| Number of mapped reads              | 18898     |

### Phred Quality Score (Q)

Q10: 10% error rate  
Q20: 1% error rate  
Q30: 0.1% error rate  
Q40: 0.01% error rate

|                         |               |                    |
|-------------------------|---------------|--------------------|
| < Q15<br>Not acceptable | > Q15<br>Good | > Q20<br>Excellent |
|-------------------------|---------------|--------------------|

### Abundance

Species with low abundance (<0.5%) are shown in light grey.

#### SAMPLE\_01

| row | abundance | species                    | genus          | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|----------------------------|----------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Staphylococcus epidermidis | Staphylococcus | Staphylococcaceae    | 1282    | 5585                  | 99.30%                  | 96.33%                  |
| 2   | 24.39%    | Sphingobium rhizovincium   | Sphingobium    | Sphingomonadaceae    | 432308  | 4045                  | 98.61%                  | 98.07%                  |
| 3   | 8.27%     | Aquabacterium parvum       | Aquabacterium  | nan                  | 70584   | 1371                  | 98.45%                  | 100.00%                 |
| 4   | 6.58%     | Imtechella halotolerans    | Imtechella     | Flavobacteriaceae    | 1165090 | 1091                  | 98.77%                  | 100.00%                 |
| 5   | 3.30%     | Elstera cyanobacterium     | Elstera        | Rhodospirillaceae    | 2022747 | 548                   | 98.63%                  | 99.57%                  |
| 6   | 2.61%     | Cutibacterium acnes        | Cutibacterium  | Propionibacteriaceae | 1747    | 433                   | 98.78%                  | 95.77%                  |
| 7   | 2.51%     | Sphingobium yanoikuyae     | Sphingobium    | Sphingomonadaceae    | 13690   | 417                   | 98.54%                  | 95.91%                  |
| 8   | 1.96%     | Caulobacter vibrioides     | Caulobacter    | Caulobacteraceae     | 155892  | 325                   | 98.80%                  | 96.63%                  |
| 9   | 1.23%     | Curvibacter fontanus       | Curvibacter    | Comamonadaceae       | 247486  | 203                   | 96.85%                  | 98.72%                  |

Purple rows indicate spike species

Green rows indicate species not found in negative control or with a normalised abundance ratio  $\geq 25$ .

#### Negative control

| row | abundance | species                   | genus         | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|-----|-----------|---------------------------|---------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 1   | 33.67%    | Sphingobium rhizovincium  | Sphingobium   | Sphingomonadaceae    | 432308  | 5631                  | 98.66%                  | 98.07%                  |
| 2   | 10.84%    | Aquabacterium parvum      | Aquabacterium | nan                  | 70584   | 1813                  | 98.45%                  | 100.00%                 |
| 3   | 10.80%    | Elstera cyanobacterium    | Elstera       | Rhodospirillaceae    | 2022747 | 1807                  | 98.71%                  | 99.57%                  |
| 4   | 7.53%     | Sphingobium yanoikuyae    | Sphingobium   | Sphingomonadaceae    | 13690   | 1260                  | 98.62%                  | 95.98%                  |
| 5   | 5.38%     | Imtechella halotolerans   | Imtechella    | Flavobacteriaceae    | 1165090 | 900                   | 98.85%                  | 100.00%                 |
| 6   | 2.94%     | Caulobacter vibrioides    | Caulobacter   | Caulobacteraceae     | 155892  | 492                   | 98.68%                  | 96.11%                  |
| 7   | 2.59%     | Cutibacterium acnes       | Cutibacterium | Propionibacteriaceae | 1747    | 432                   | 98.79%                  | 95.71%                  |
| 8   | 1.90%     | Curvibacter fontanus      | Curvibacter   | Comamonadaceae       | 247486  | 317                   | 96.77%                  | 98.72%                  |
| 9   | 1.56%     | Kinneretia asaccharophila | Kinneretia    | Comamonadaceae       | 582607  | 260                   | 98.28%                  | 100.00%                 |

Purple rows indicate spike species



**Why do we use the spike abundance when comparing species abundance across sample and negative control?**





**Why do we use the spike abundance when comparing species abundance across sample and negative control?**

# Probability scores

- Data for this is already available in the Emu output (in `*_read-assignment-distributions.tsv`, see below)
- Reflects the probability, per read, after Emu's processing

|                                      | 432308 | 13690               | 2565554               | 2282397 | 219301 | 70584                 | 450365                 | 1165090 |
|--------------------------------------|--------|---------------------|-----------------------|---------|--------|-----------------------|------------------------|---------|
| 886864c8-580c-4361-95d0-603325931d2e | 1.0    |                     |                       |         |        |                       |                        |         |
| 94b2366b-72a3-4fd5-a264-a0798464e976 |        | 0.8481517474373064  | 0.1518482525626937    |         |        |                       |                        |         |
| 2c47bf95-a083-4e44-adda-3bb5c303b9e3 |        | 0.9982541030218708  | 0.00174589697812922   |         |        |                       |                        |         |
| f2b74119-1937-4332-aae1-e704fa8f6601 |        | 0.9982542075621543  | 0.0017457924378456334 |         |        |                       |                        |         |
| 20493355-0bd9-40c6-b212-1868100936e5 |        | 0.8481517474373036  | 0.15184825256269643   |         |        |                       |                        |         |
| 582af236-c98e-42bc-8bba-5db79dda09f4 |        | 0.9982540686283654  | 0.0017459313716346515 |         |        |                       |                        |         |
| 1c697f82-0cd1-417b-a4b0-c4b89ad4f75d |        | 0.998250206374969   | 0.0017497936250311394 |         |        |                       |                        |         |
| 34178578-0066-4bc9-b54f-7d7c1d8b1148 |        | 0.05144463171396726 | 0.9485553682860327    |         |        |                       |                        |         |
| 15651727-61d8-4027-ba7f-da92d1a085de |        | 0.9982541376405205  | 0.0017458623594794073 |         |        |                       |                        |         |
| 289faf6d-0450-48dd-a596-60f41309dbc2 |        |                     |                       | 1.0     |        |                       |                        |         |
| 083d64ff-c64f-45b0-a607-fb7774943445 |        |                     |                       |         | 1.0    |                       |                        |         |
| 4eb80702-e578-4413-a145-020dc1460ea4 |        |                     |                       |         |        | 1.0                   |                        |         |
| f786dd56-436c-4444-a240-e4d1a52970f9 |        |                     |                       |         |        |                       | 4.1811006260320675e-42 |         |
| 9c63b015-9dbe-47cd-9cc5-fa3b0e8e6341 |        |                     |                       |         |        | 1.0                   |                        |         |
| 7f2b1337-cafe-47b2-b942-ea4c2db6eebe |        |                     |                       |         |        | 1.0                   |                        |         |
| e72099b0-5c08-4117-837f-e15c9036ce4c |        |                     |                       |         |        | 7.151725076902957e-71 | 1.0                    |         |
| 117d0800-8ec9-4e1b-a3ae-d2e1c2e32590 |        |                     |                       |         |        | 0.9999999996500353    | 3.499647373365163e-10  |         |
| aec00191-dded-4661-836e-4cdbeb4e804b |        |                     |                       |         |        | 1.0                   |                        |         |
| 94117a53-d30f-43c4-934d-f593d7002a01 |        |                     |                       |         |        | 1.0                   |                        |         |
| 821fe7b6-6f3e-49f3-9c58-b5d677c683fe |        |                     |                       |         |        | 1.0                   |                        |         |
| 1ea754c2-cb6c-49e5-9291-c8f48e468a05 |        |                     |                       |         |        | 1.0                   |                        |         |
| 00afea5a-cfd7-4bdd-8aa5-48e521ea190b |        |                     |                       |         |        | 1.0                   | 3.8329989378240703e-42 |         |
| 43685156-b879-4cdf-83ac-a5ce0af23736 |        |                     |                       |         |        | 1.0                   |                        |         |
| b0d1ac29-65b2-4952-896f-d962ad4809f7 |        |                     |                       |         |        |                       |                        | 1.0     |
| 22c72173-2838-4223-8e0f-5614f8f25d12 |        |                     |                       |         |        |                       |                        | 1.0     |
| :                                    |        |                     |                       |         |        |                       |                        | 1.0     |

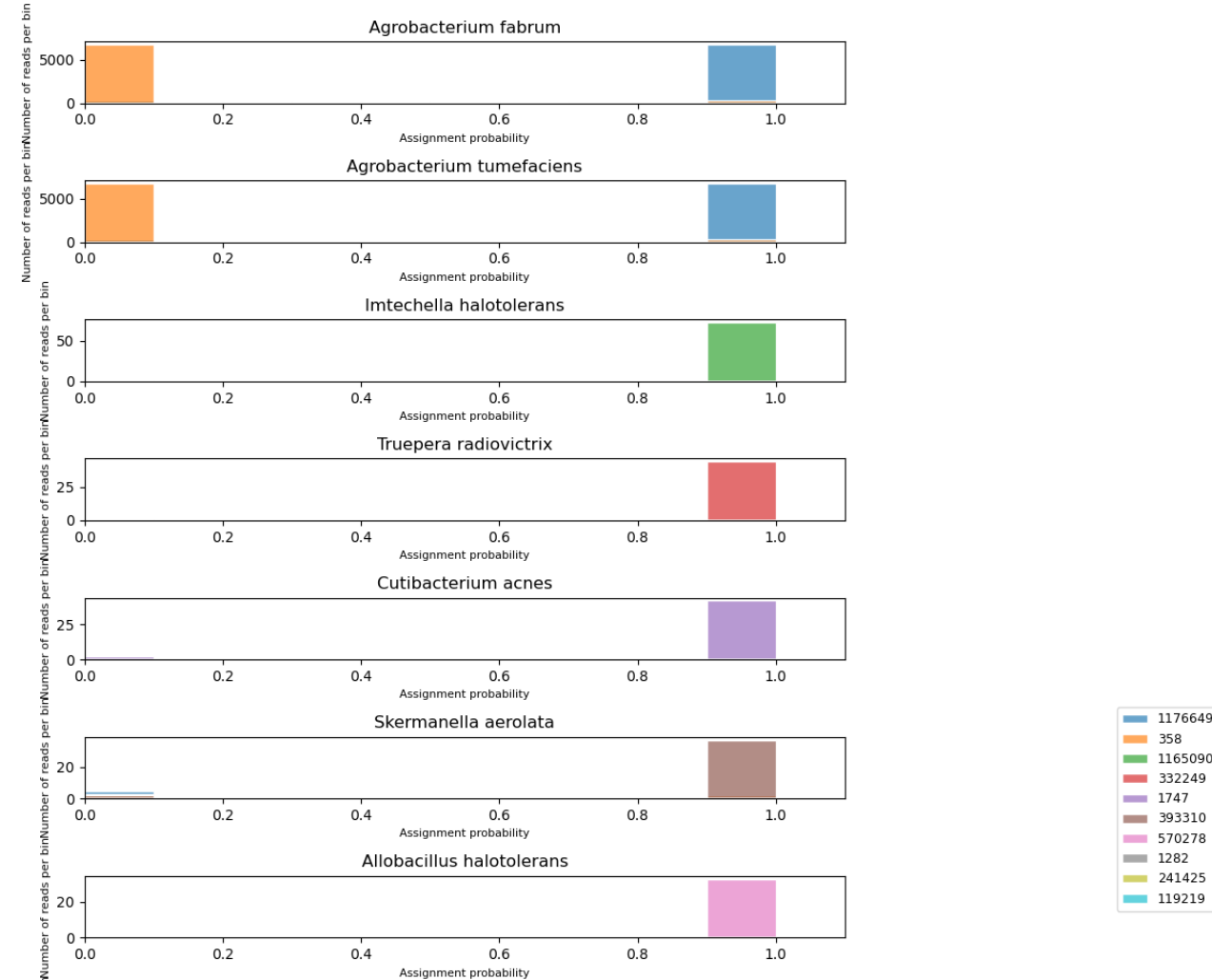
# Probability scores

## Calculating median probability:

- For each species  $s_1$ , in the order of decreasing total abundance:
  - Select reads with a non-null, non-zero species assignment
  - For each of these reads, take the median of probabilities for the initially selected species  $s_1$

## Plotting probability distribution:

- For each species, in the order of decreasing total abundance:
  - Extract reads with a non-null, non-zero species assignment
  - For each species for which these reads have values (can be more than one), plot a histogram



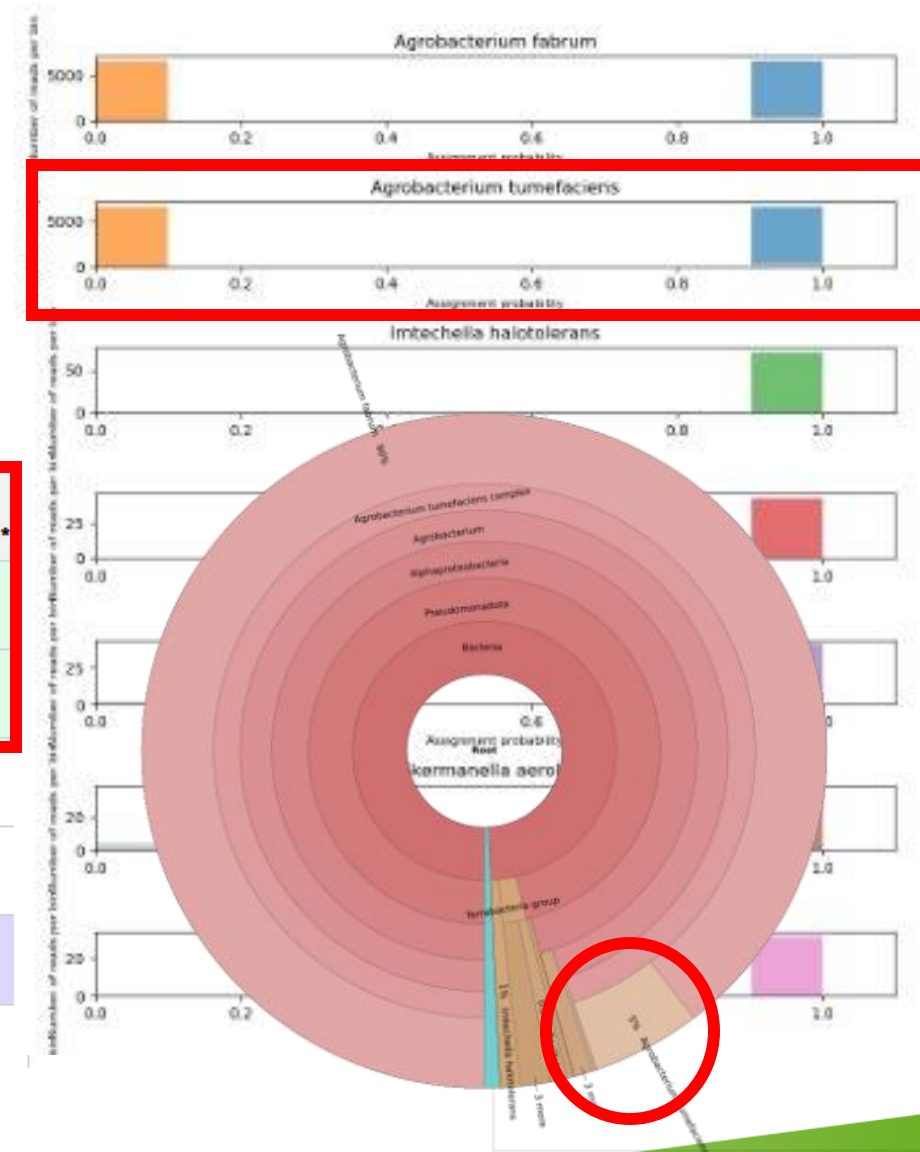
# Median/mean probability score

- Surfaces cases where the abundance is spread across multiple closely related species
- Highlights how the abundance in Emu is a combination of probability per read, and number of supporting reads

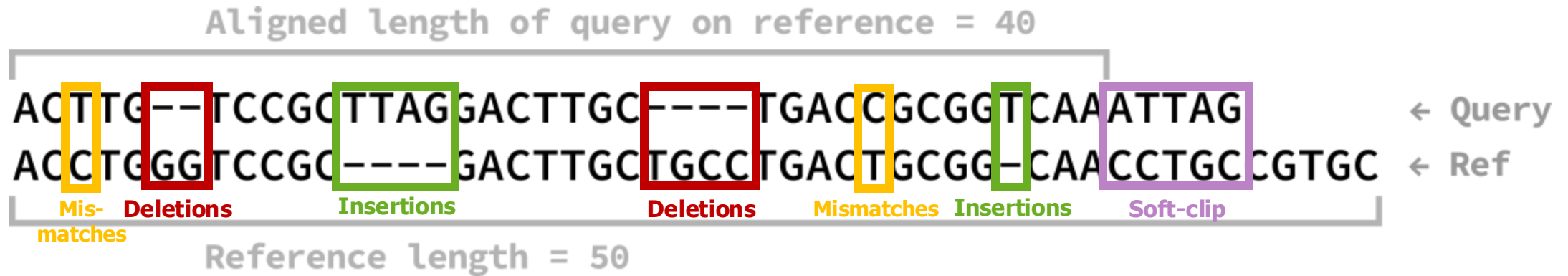
|   | abundance | species                   | genus         | family            | tax id  | estimated read counts | median probability* | mean probability* |
|---|-----------|---------------------------|---------------|-------------------|---------|-----------------------|---------------------|-------------------|
| 0 | 83.25%    | Agrobacterium fabrum      | Agrobacterium | Rhizobiaceae      | 1176649 | 16554.019105          | 0.999999            | 0.941425          |
| 1 | 4.64%     | Agrobacterium tumefaciens | Agrobacterium | Rhizobiaceae      | 358     | 922.091590            | 0.000001            | 0.052448          |
| 2 | 2.97%     | Sphingobium rhizovincinum | Sphingobium   | Sphingomonadaceae | 432308  | 590.933353            | 1.000000            | 0.856425          |
| 3 | 1.26%     | Aquabacterium parvum      | Aquabacterium | nan               | 70584   | 249.999999            | 1.000000            | 0.980392          |
| 4 | 1.07%     | Imtechella halotolerans   | Imtechella    | Flavobacteriaceae | 1165090 | 212.000000            | 1.000000            | 0.986047          |
| 5 | 0.83%     | Sphingobium               | Sphingobium   | Sphingomonadaceae | 13690   | 165.220973            | 1.000000            | 0.897940          |

Purple rows indicate spike species

Green rows indicate species not found in negative control



# Raw alignment quality metrics



$$\text{aligned identity} = \frac{b_M}{b_M + b_X + b_I + b_D}$$

$$\text{aligned span} = \frac{b_{\text{aligned reference sequence}}}{b_{\text{reference}}} = \frac{b_{\text{query}} + b_D - b_I - b_S}{b_{\text{reference}}}$$

$b_M$  = matched bases

$b_X$  = mismatched bases

$b_I$  = inserted bases

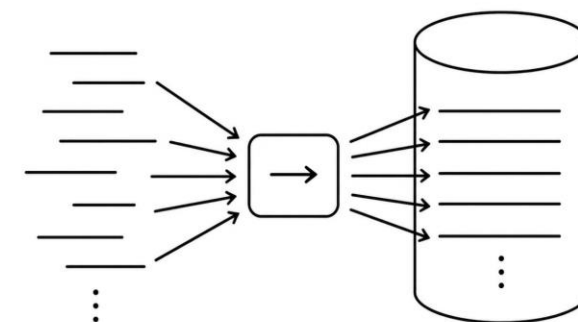
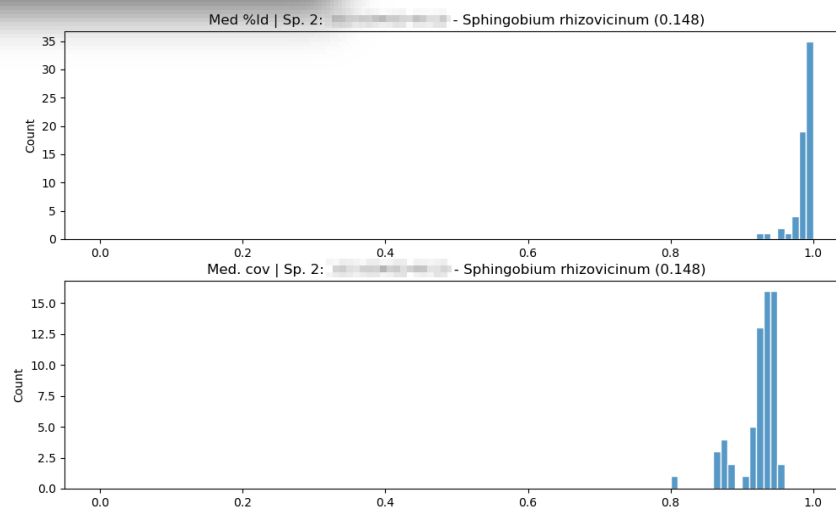
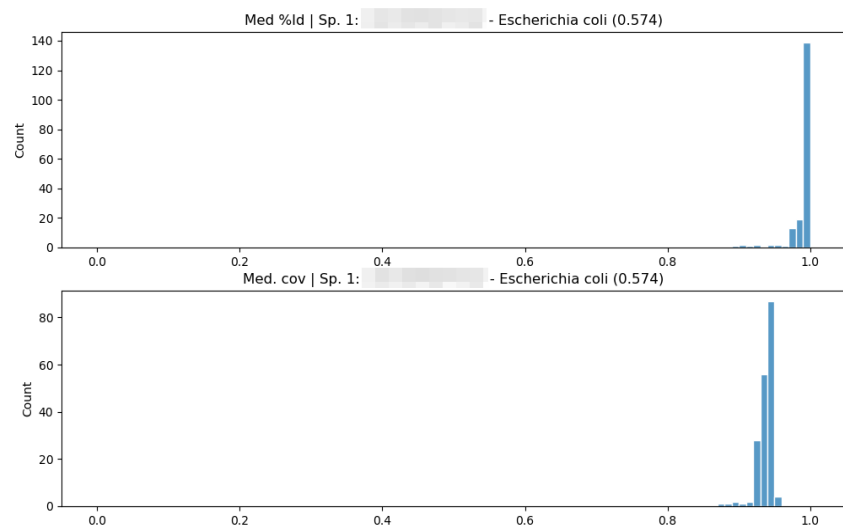
$b_D$  = deleted bases

$b_S$  = soft-clipped bases

$b_{\text{query}}$  = query bases

$b_{\text{reference}}$  = reference bases

# Raw alignment metrics as distributions



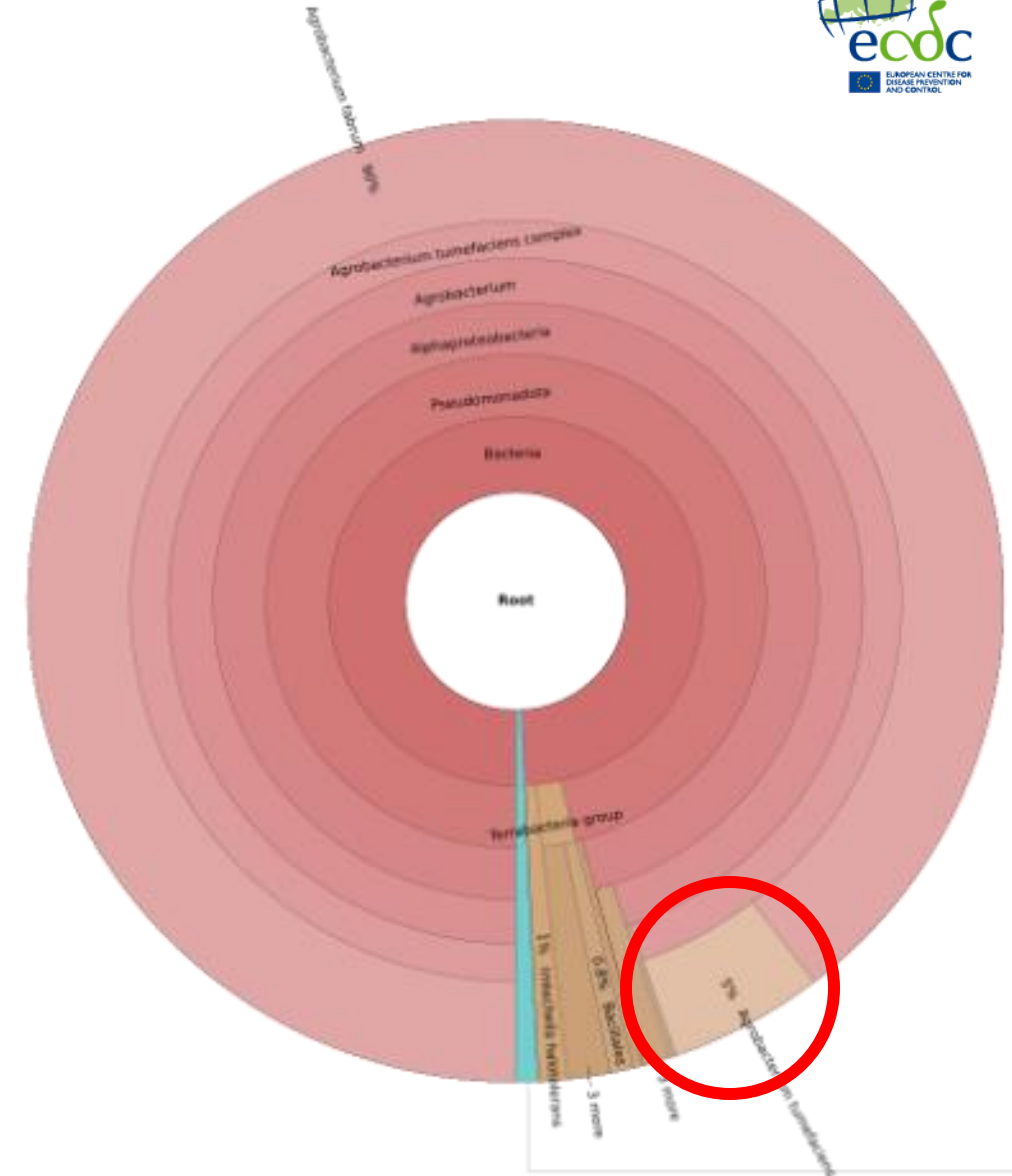
Reminder: Emu aligns *all* reads against *all* sequences in the database, resulting in a set of alignments rather than a single alignment.

# Raw Alignment Metrics

- Raw alignment metrics
  - Median percent identity
  - Median percent coverage (span)
- Calculated on raw alignments  
=> ONT errors not removed  
(lower identity expected)



|   | abundance | species                   | genus         | family               | tax id  | estimated read counts | median aligned identity | median aligned coverage |
|---|-----------|---------------------------|---------------|----------------------|---------|-----------------------|-------------------------|-------------------------|
| 0 | 90.18%    | Agrobacterium fabrum      | Agrobacterium | Rhizobiaceae         | 1176649 | 6827                  | 99.58%                  | 96.15%                  |
| 1 | 5.22%     | Agrobacterium tumefaciens | Agrobacterium | Rhizobiaceae         | 358     | 395                   | 99.38%                  | 87.17%                  |
| 2 | 0.96%     | Imtechella halotolerans   | Imtechella    | Flavobacteriaceae    | 1165090 | 73                    | 99.31%                  | 100.00%                 |
| 3 | 0.58%     | Truepera radiovictrix     | Truepera      | Trueperaceae         | 332249  | 44                    | 99.35%                  | 96.07%                  |
| 4 | 0.55%     | Cutibacterium acnes       | Cutibacterium | Propionibacteriaceae | 1747    | 42                    | 99.50%                  | 95.77%                  |



# Some things to keep in mind

- Indels currently counted in bases
- Counting in events might have benefits in some cases
- BLAST percent identity skips indels altogether
  - Counts only percent matching of "aligned" bases - bases with a corresponding base on the other sequence

```
> ref
Length=45

Score = 17.7 bits (9), Expect = 0.006
Identities = 9/9 (100%), Gaps = 0/9 (0%)
Strand=Plus/Plus

Query  15  GACTTGCTG  23
        |||||
Sbjct  13  GACTTGCTG  21

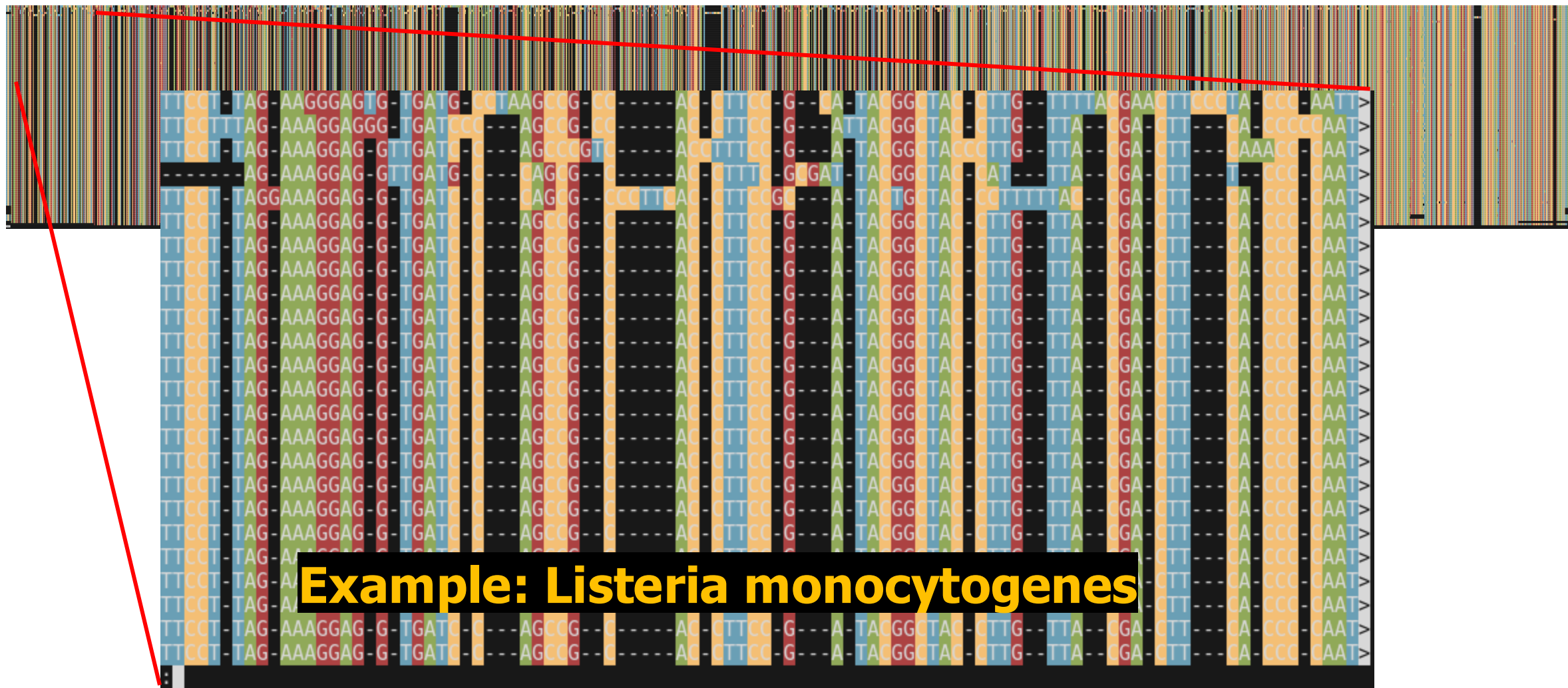
Score = 14.0 bits (7), Expect = 0.080
Identities = 12/14 (86%), Gaps = 1/14 (7%)
Strand=Plus/Plus

Query  21  CTGACCGCGGTCAA  34
        ||||| |||| |||
Sbjct  23  CTGACTGCGG-CAA  35

Score = 12.2 bits (6), Expect = 0.29
Identities = 6/6 (100%), Gaps = 0/6 (0%)
Strand=Plus/Plus

Query  5  GTCCGC  10
        |||||
Sbjct  7  GTCCGC  12
:_
```

# Within-species variation in the Emu database





**What is one typical pattern to be suspicious about in Emu results?**





**What is one typical pattern to be suspicious about in Emu results?**



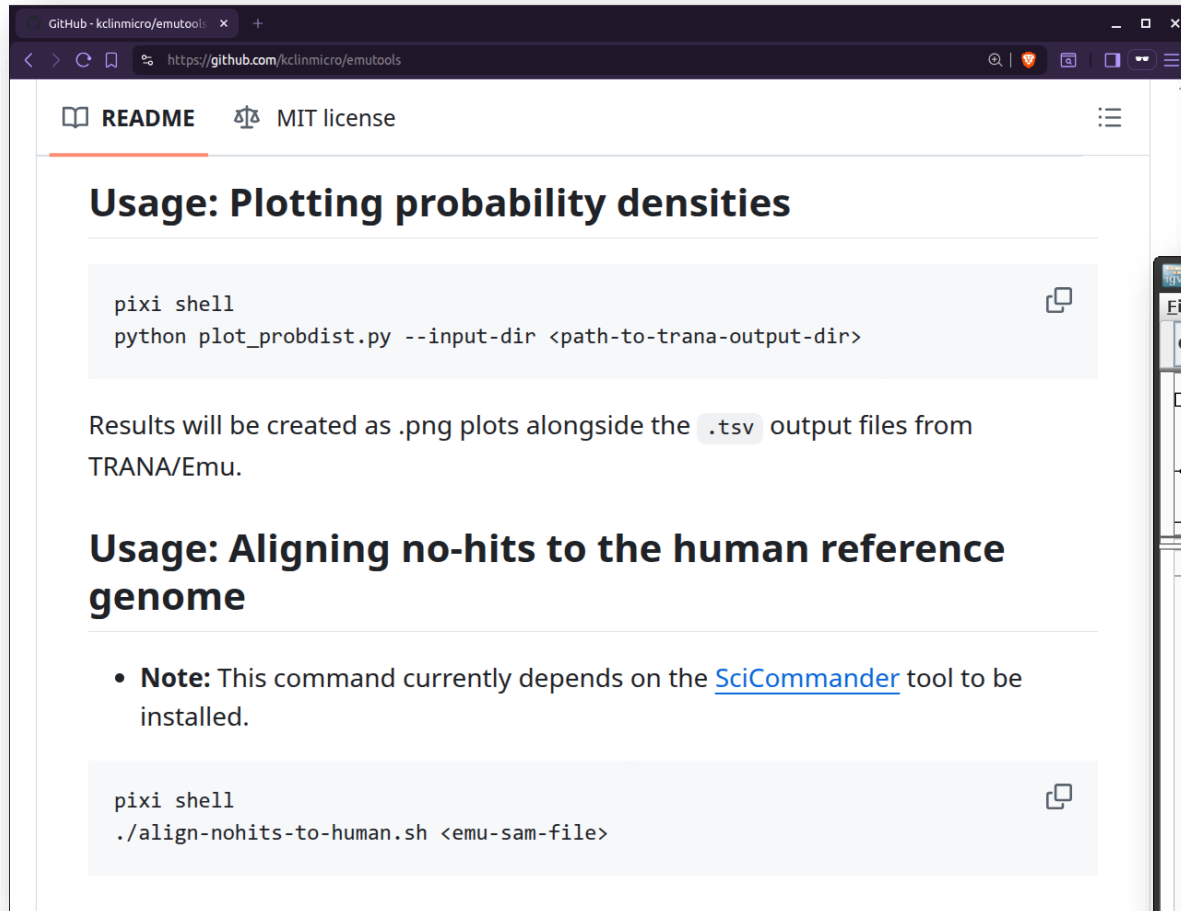
**Why do we need to take the  
median value of metrics from  
Emu?**





# Why do we need to take the median value of metrics from Emu?

# The emutools library



GitHub - kclinmicro/emutools

README MIT license

## Usage: Plotting probability densities

```
pixi shell
python plot_probdist.py --input-dir <path-to-trana-output-dir>
```

Results will be created as .png plots alongside the .tsv output files from TRANA/Emu.

## Usage: Aligning no-hits to the human reference genome

- **Note:** This command currently depends on the [SciCommander](#) tool to be installed.

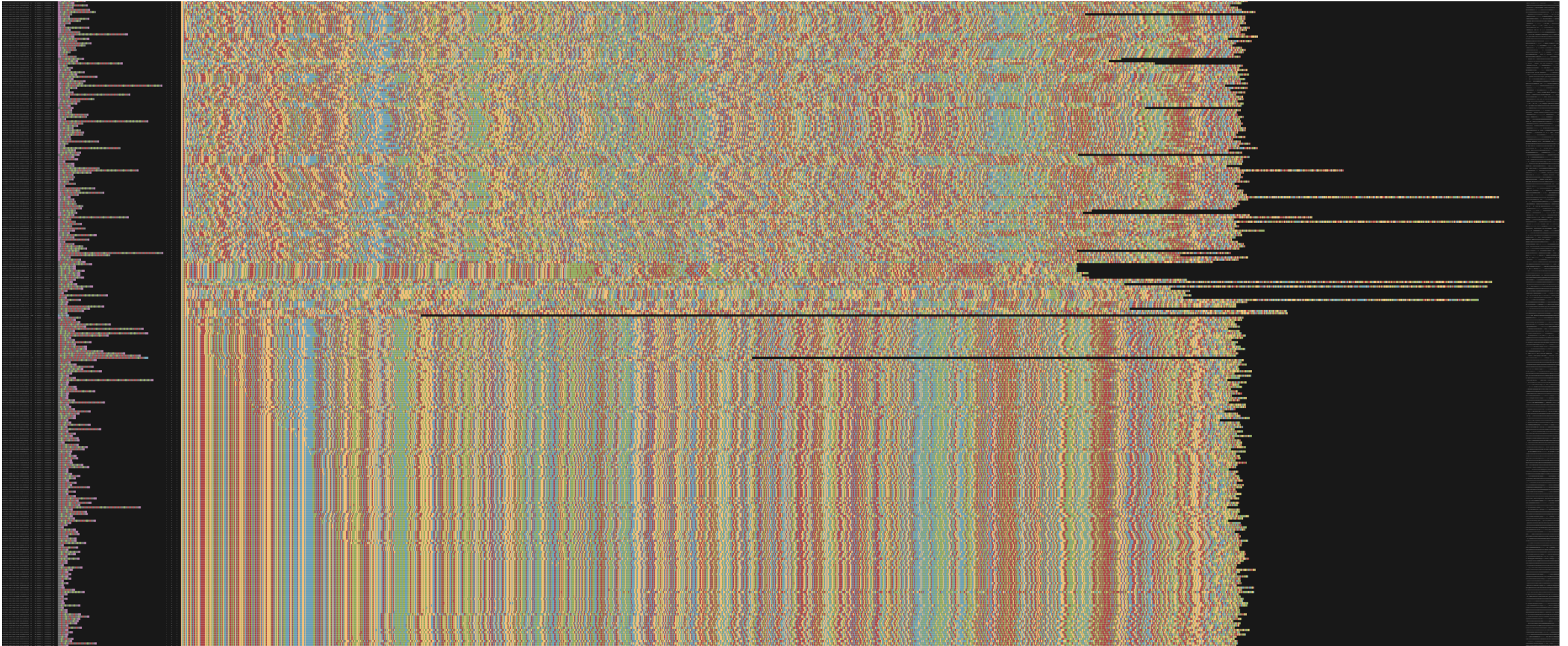
```
pixi shell
./align-nohits-to-human.sh <emu-sam-file>
```

- Collects various scripts for troubleshooting & visualization of Emu output
- [github.com/kclinmicro/emutools](https://github.com/kclinmicro/emutools)



Early-release software:  
Rough edges expected!

# Sometimes just some sorting and coloring reveals a lot



```
samtools view sample01<...>.aln_human.bam | sort -k 10,10 | head -n 100 | sek sam | column -t -s $'\t' | less -S
```

Tool used for coloring: [github.com/samuell/sek](https://github.com/samuell/sek)

# In summary

List of learning points in this session:

- Report interpretation requires more than abundance alone
- Negative controls are essential for identifying true findings
- Spike-normalized abundance ratios help prioritize relevant species
- Probability scores helps surface uncertain species level predictions
- Alignment-based metrics add further information to help spot uncertain species predictions
- Emu might distribute abundance on more than one closely related species in some situations

# Recommended material to explore further

- Aja-Macaya, Pablo, et al. 2025. **"Nanopore Full Length 16S rRNA Gene Sequencing Increases Species Resolution in Bacterial Biomarker Discovery."** *Scientific Reports* 15 (1): 26486. <https://doi.org/10.1038/s41598-025-10999-8>.
- Brunet, Sofia, et al. 2025. **"Nationwide Multicentre Study of Nanopore Long-Read Sequencing for 16S rRNA-Species Identification."** *European Journal of Clinical Microbiology & Infectious Diseases* 44 (8): 1907–16. <https://doi.org/10.1007/s10096-025-05158-w>.



**Break till 10:25**

***Don't leave this meeting  
Just turn off the camera***

# Reusing this content



This work is licensed under the Creative Commons (<http://creativecommons.org>) Attribution 4.0 Generic license



You are free:

- to share – to copy, distribute and transmit the work
- to remix – to adapt the work

Under the following conditions; you must attribute the work in the manner specified by the author or licensor but not:

- in any way that suggests that they endorse you or your use of the work
- in any way that suggests you are author of the work.

*The authors of this slide/page encourage educators, trainers, and professionals to include this slide within their documents and presentations for rightful attribution of their works and thus also allow it to be easily shared. (ECV1-29/9/11)*

[Terms of usage \(Learning Portal\)](#)

# Acknowledgements

**The creation of this training material was commissioned by ECDC to Karolinska University Hospital with the direct involvement of Anna Norén, Elin Loo, Lili Andersson-Li, Lina Guerra Blomqvist, Samuel Lampa, Sofia Stamouli**

**The revision and update of this training material was commissioned by ECDC to Karolinska University Hospital with the direct involvement of Anna Norén, Elin Loo, Lili Andersson-Li, Lina Guerra Blomqvist, Samuel Lampa, Sofia Stamouli**