

Presentation Slides

Guidelines and Examples

January 3, 2017

Saarland University
Chair for Clinical Bioinformatics



Introduction

Purpose of these slides

- ✓ Slide layout, color scheme and font
- ✓ Figures, tables, formulas
- ✓ Other tips and references

✗ Exact description of how slides should look like

👉 There is **no** single perfect solution for any presentation

Layout, Color Scheme and Font

Some guidelines for the layout

Basic elements

- Section/slide title – for “global” orientation
- Slide counter – helps in navigation during Q&A

Optional elements

- Presentation title and date
- Institution name/logo
- ...

All these elements should not take too much **space**

- I.e. main part ($\geq 80\%$) is reserved for the slide content
- N.B. the slide real-estate is limited; use it efficiently

Some guidelines for the color scheme

- ☞ Appropriate **contrast** between text and background
 - Dark background/light font or Light background/dark font
 - May depend on the presentation room or beamer – **test early**, if possible
 - If not sure use light background and dark font
- ☞ No "aggressive" color combinations, e.g. red and green
- ☞ **Few** main colors, no "rainbow" color scheme
 - Optional: same color scheme for layout and graphics
- ☞ **Plain background** is in most cases the best choice

Some guidelines for the font

- ☞ Neutral and easy to read, **not** *"something like that"*
- ☞ Sufficiently large, e.g. 18pt for main text
- ☞ Avoid using only capital letters
- ☞ Use a different style for titles/keywords, e.g. in **bold**

Figures, Tables, Formulas

- ☞ Captions/numbering (optional)
- ☞ Appropriate figure **resolution**
 - All relevant details should be recognizable
 - Figures should **not** be pixelated
- ☞ **Consistent** size/placement (alignment)
 - If multiple figures on same slide
- ☞ **Omit** irrelevant details
 - E.g. screenshots – select only relevant area
- ☞ **Highlight** important details
- ☞ Add image **source**
 - Below the figure or on a separate reference slide

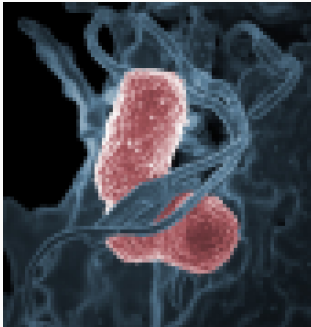
Figures: How it COULD be done



Figure 1: *Figure caption*

Figure 2: *Another figure caption*

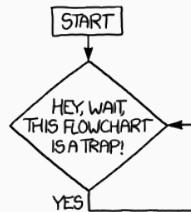
Figures: How it should NOT be done



Figures: When to use?



How something looks like



Workflow



Illustration of a method, e.g. NGS

Results/data as graphs

- ☞ Captions/numbering (optional)
- ☞ **Structure** the table
 - E.g. header, row blocks, column blocks
- ☞ **Highlight** important data by font family or color
- ☞ **Omit irrelevant** information
 - **Condense** the information; do **not** simply copy & paste
 - More space for important data
- ☞ **Consistent** alignment within the cells

Tables: How it COULD be done



X	Y	Z
block 1	...	...
block 1	...	...
block 1	...	...
block 2	...	...
block 2	Highlight	...
block 2	...	...

Table 1: Example table

Tables: How it should NOT be done



X	Y	Z	An unimportant column	A	B	C
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...

- 👉 Use specific **font** for variables
 - Do **not** use the same font as for regular text
- 👉 It is advisable to use formula editor/environment
- 👉 **Explain** the variables in the formula

Formulas: How it COULD be done



This is an example equation for variable X :

$$X = a \times b + Y - \alpha$$

where Y is the input variable and a , b and α are constants

Formulas: How it should NOT be done



This is an example equation for variable X:

$$X = a * b + Y - \text{alpha}$$

Structure

Title slide

- ☞ Presentation **title**, e.g. title of the publication
 - If presenting a paper add author names
 - If author list is too long, abbreviate it, e.g. last names of first three authors *et al.*
- ☞ Your **name**
- ☞ Presentation **purpose**, e.g. seminar
- ☞ Presentation **date**
- ☞ Optional: institution/group, logo

Outline/Overview

- ☞ Main **points** of the presentation
- ☞ **Avoid** general and uninformative keywords, i.e. no "Introduction", "Methods" etc.
- ☞ Instead, **concisely** motivate the main problem, how it is addressed, and what are the main improvements/conclusions

Methods

- 👉 Essential theoretical background
- 👉 E.g. workflow, structure, formulas etc.
- 👉 If necessary, provide reasonable simplifications for the purpose of the presentation but know the intricate details for the Q&A

Results

- 👉 Used data: How was it generated? Processed?
- 👉 Analysis steps
- 👉 Relevant results (figures, tables)
- 👉 **Emphasize** and highlight key results

Summary

- ☞ Summary of the approach
- ☞ Summary of the results
- ☞ Conclusions
- ☞ Advantages/drawbacks
- ☞ Future work

Other

☞ References:

- All used papers, books, articles, image sources etc.
- `google.com` or `wikipedia.org` are **not sufficient**, yet may be used as **starting points** of your literature research

☞ Optional: "Thank you" slide

☞ Appendix: Additional information and/or results, information requiring dedicated/explicit explanation etc.

Other suggestions

- If the paper includes **supplemental material** – get it, it may contain additional explanations and important information
- Look through the **references** used in the paper – some things not explained in the paper may be found in there
- Preparing a presentation includes almost always additional **literature research** (apart from the mentioned references)

Appearance

- ☞ Export slides to **PDF** if sharing them with someone
 - Same appearance on different operating systems
 - For non-standard fonts, make sure to **embed** them in the PDF
- ☞ Avoid **crowded** slides:
 - **Reduce content**: As little as possible and as much as necessary
 - **Split** into multiple slides, if necessary
- ☞ Choose **appropriate** way of representing information
 - E.g., visualize **complex** things by a **figure** instead of text/formula (if appropriate)
- ☞ Use **simple examples** for your explanations
 - You may have non-experts in the audience
- ☞ **Highlight** important details in text, figures, tables, formulas

- 👉 **No sentences** or **long phrases** (exceptions: quotes)
- 👉 **Paraphrase**, i.e. no copy-paste from paper
- 👉 Check appropriate use of whitespace – it distracts otherwise
 - **One** whitespace **after** a punctuation mark, e.g., comma or period
 - **No** whitespace **before** a punctuation mark

- **Practice** your talk in front of someone, it will help you to
 - Estimate the length (a bit shorter preferred over much longer)
 - Get additional feedback, possibly from non-experts
 - **Talk to the audience**, do not look at the slides all the time
 - If necessary, use flashcards or other notes that you can put in front of you
 - N.B. do **not** simply read from your notes either
 - Try to **avoid** making distracting gestures, sounds (e.g. "Um")
 - **Explain** all needed details when showing
 - Formulas → What is computed? Variables?
 - Figures → What is plotted/drawn? Legend? Axes?
 - Tables → Column/row content?
- before** interpreting
- **Be prepared** to explain all terms/methods/statements on your slides

References

Useful Tools, Links, and Literature

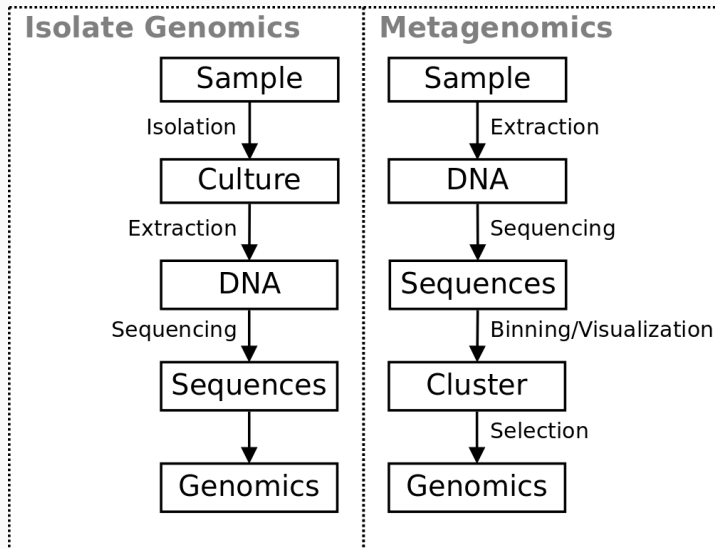
- Ten Secrets to Giving a Good Scientific Talk
- Elsevier: How to give a dynamic scientific presentation
- Butterick's Practical Typography - Presentations
- L^AT_EX
 - A high-quality typesetting system for writing scientific articles, theses, presentations
 - Freely available, there are many tutorials
 - Many features, e.g. mathematical environment, bibliography, automatic numbering for figures, tables and equations
- Gimp
 - Freely available
 - Can be used to create high-resolution screenshots from PDFs

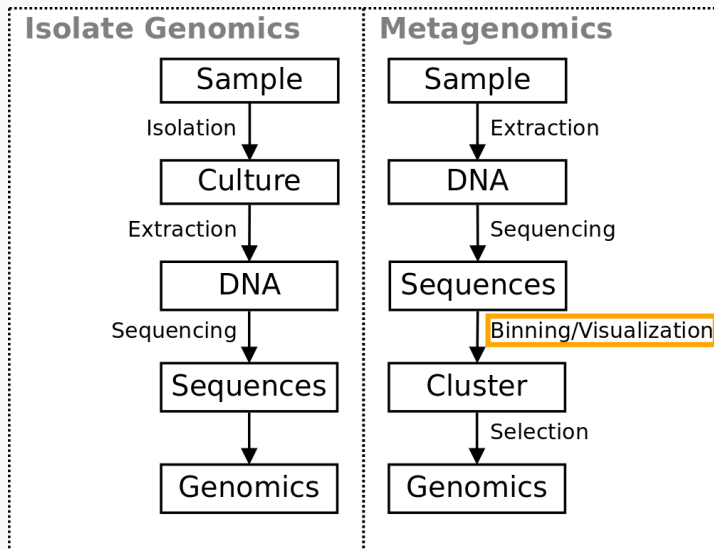
Presentation Example

Masterseminar

Visualization of metagenomic data utilizing coverage information

Dominik Wermke





GCTGC GGC GGT GCG TGG CAT CCG GGG CAC GGT GGA GAT GCG GAG



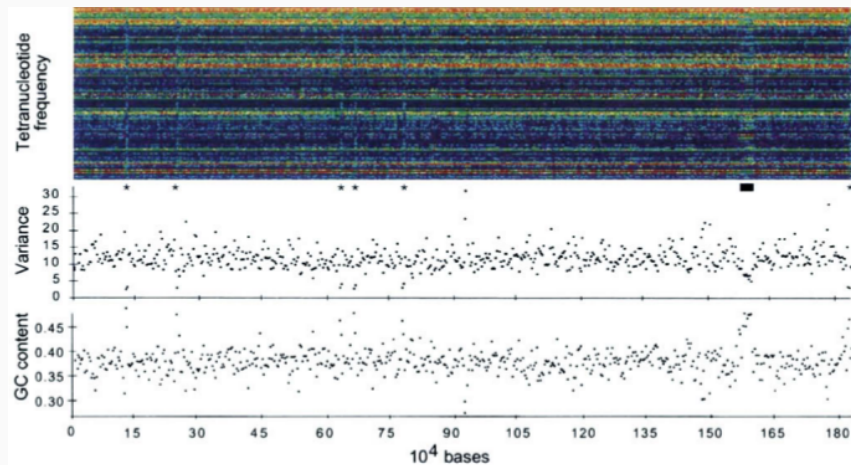
GGT GCG GCG GGT GCG
 GTG CGG CGG GTG CGT
 TGC GGC GGG TGC GTG



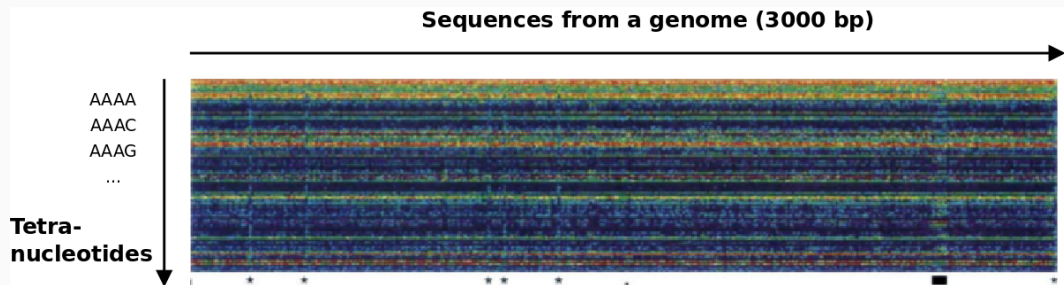
CGG	CGT	GGC	GCG	GGG	GGT	GTG	TGC
2	1	1	3	1	2	3	2

k	2	3	4	5
dimension	16	64	256	1024

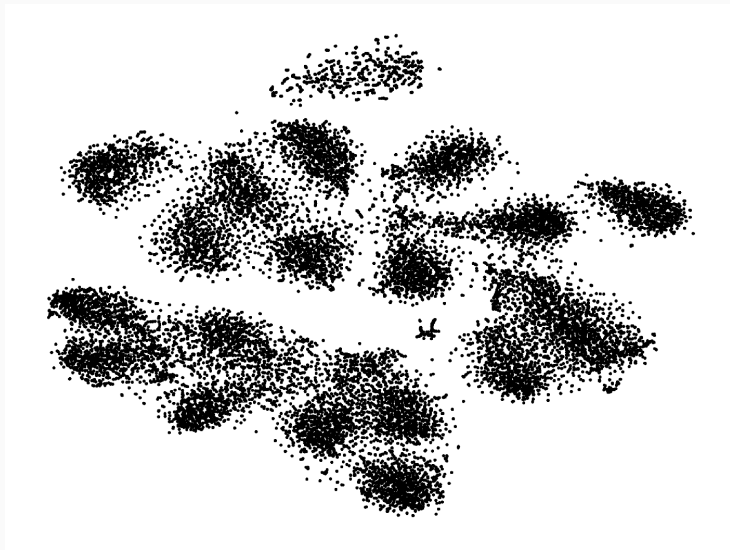
K-mers

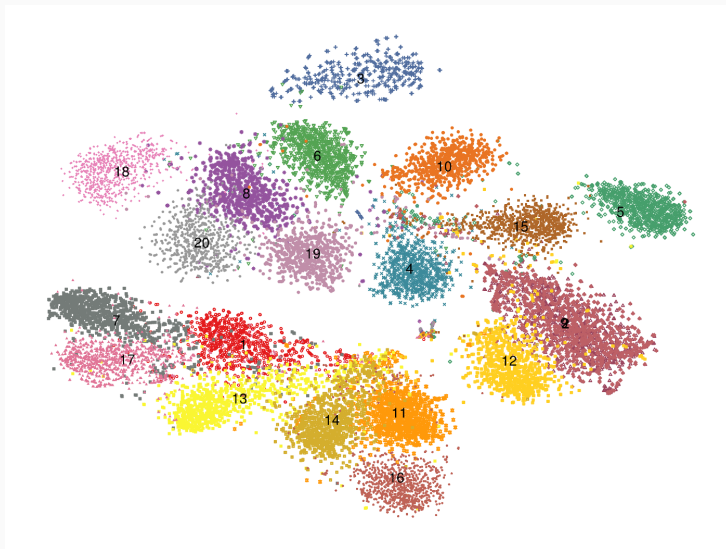


[1]



[1]





Classes:

○	(1) <i>Leifsonia xyli</i>
△	(2) <i>Escherichia coli</i> UT189
+	(3) <i>Candidatus Carsonella ruddii</i>
×	(4) <i>Haemophilus influenzae</i>
◇	(5) <i>Bacillus amyloliquefaciens</i>
▽	(6) <i>Brachyspira hyodysenteriae</i>
⊠	(7) <i>Geodermatophilus obscurus</i>
*	(8) <i>Rickettsia prowazekii</i>
⬠	(9) <i>Escherichia coli</i> str. 'clone D i14'
⊕	(10) Uncultured Termite group 1 bacterium
⊗	(11) <i>Maricaulis maris</i>
⊞	(12) <i>Marinobacter psychrophilus</i>
⊠	(13) <i>Methylobacterium nodulans</i>
⊞	(14) <i>Aminobacter aminovorans</i>
■	(15) <i>Bacillus clausii</i>
●	(16) <i>Bordetella hinzii</i>
▲	(17) <i>Brachybacterium faecium</i>
◆	(18) <i>Campylobacter subantarcticus</i>
●	(19) <i>Candidatus Amoebophilus asiaticus</i>
●	(20) <i>Baumannia cicadellinicola</i>

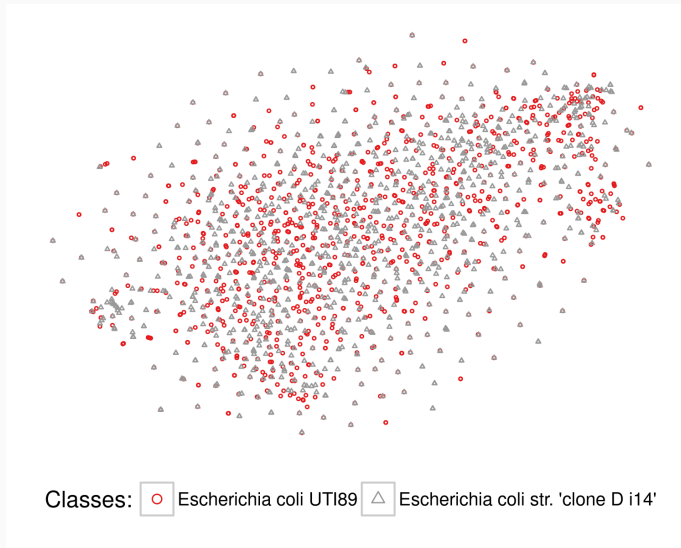
Visualization



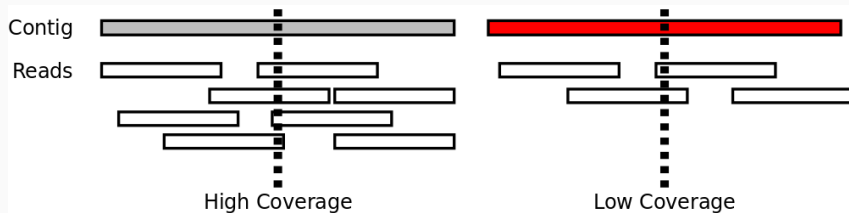
Classes:

- | | |
|---|--|
| ○ | (1) <i>Leifsonia xyli</i> |
| △ | (2) <i>Escherichia coli</i> UT189 |
| + | (3) <i>Candidatus Carsonella ruddii</i> |
| × | (4) <i>Haemophilus influenzae</i> |
| ◇ | (5) <i>Bacillus amyloliquefaciens</i> |
| ▽ | (6) <i>Brachyspira hyodysenteriae</i> |
| ⊠ | (7) <i>Geodermatophilus obscurus</i> |
| * | (8) <i>Rickettsia prowazekii</i> |
| ⬠ | (9) <i>Escherichia coli</i> str. 'clone D i14' |
| ⊕ | (10) Uncultured Termite group 1 bacterium |
| ⊗ | (11) <i>Maricaulis maris</i> |
| ⊞ | (12) <i>Marinobacter psychrophilus</i> |
| ⊠ | (13) <i>Methylobacterium nodulans</i> |
| ⊞ | (14) <i>Aminobacter aminovorans</i> |
| ■ | (15) <i>Bacillus clausii</i> |
| ● | (16) <i>Bordetella hinzii</i> |
| ▲ | (17) <i>Brachybacterium faecium</i> |
| ◆ | (18) <i>Campylobacter subantarcticus</i> |
| ● | (19) <i>Candidatus Amoebophilus asiaticus</i> |
| ● | (20) <i>Baumannia cicadellinicola</i> |

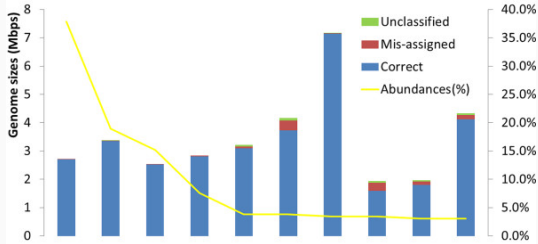
Visualization



Coverage

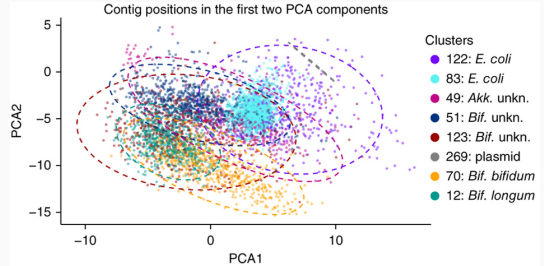


MaxBin:



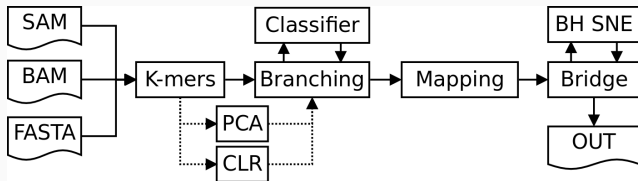
[2]

CONCOCT:

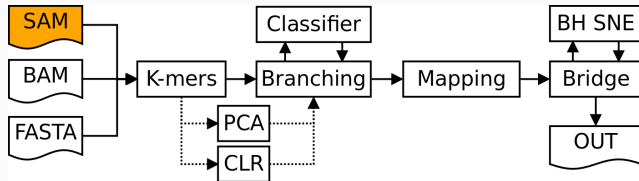


[3]

Implementation

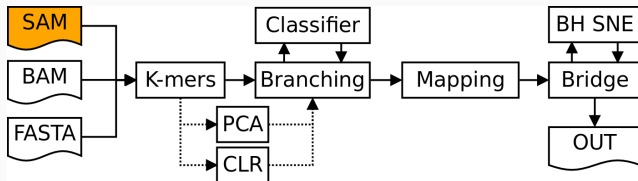


Implementation



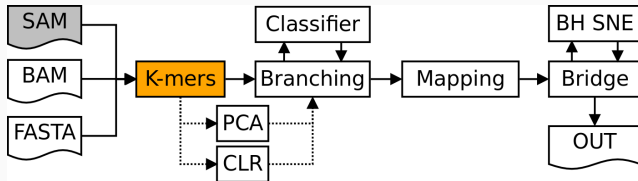
1	0	1	483	42	100M	*	0	0	GCAGTCGCGTATTCATCGCG	
2	16	1	372	42	100M	*	0	0	TGCTCGAACGTCTGCAAAAG	
3	0	1	646	42	100M	*	0	0	CAGCGCCGCTGATGGCGCGG	
4	16	2	61	42	100M	*	0	0	GAGTTTTAACGACGAACGTG	
5	0	2	390	42	100M	*	0	0	AGCGGATGAGCTTTGCGCTG	

Implementation



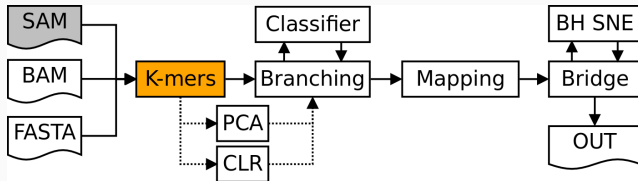
1	0	1	483	42	100M	*	0	0	GCAGTCGCGTATTCATCGCG	
2	16	1	372	42	100M	*	0	0	TGCTCGAACGTCTGCAAAAG	
3	0	1	646	42	100M	*	0	0	CAGCGCCGCTGATGGCGCGG	
4	16	2	61	42	100M	*	0	0	GAGTTTTAACGACGAACGTG	
5	0	2	390	42	100M	*	0	0	AGCGGATGAGCTTTGCGCTG	

Implementation

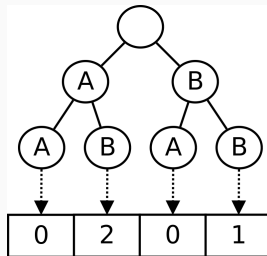


1	0	1	483	42	100M	*	0	0	GCAGTCGCGTATTCATCGCG	
2	16	1	372	42	100M	*	0	0	TGCTCGAACGTCTGCAAAAG	
3	0	1	646	42	100M	*	0	0	CAGCGCCGCTGATGGCGCGG	
4	16	2	61	42	100M	*	0	0	GAGTTTTAACGACGAACGTG	
5	0	2	390	42	100M	*	0	0	AGCGGATGAGCTTTGCGCTG	

Implementation

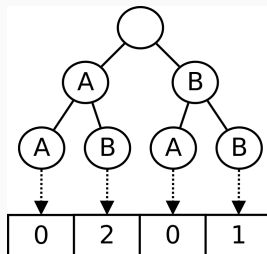


```
GCAGTCGCGTATTCATCGCG
TGCTCGAACGTCTGCAAAAG
CAGCGCCGCTGATGGCGCGG
GAGTTTTTAACGACGAACGTG
AGCGGATGAGCTTTGCGCTG
```



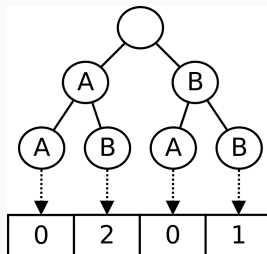
GCAGTCGCGTATTCATCGCG
TGCTCGAACGTCTGCAAAAG
CAGCGCCGCTGATGGCGCGG
GAGTTTTAACGACGAACGTG
AGCGGATGAGCTTTGCGCTG

Implementation



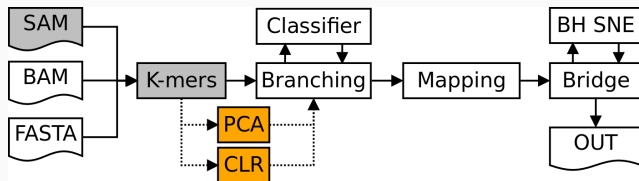
AAA	AAC	AAG	AAT	ACA	ACC	ACG	ACT	...
0	0	0	0	0	0	0	0	...
2	1	1	0	0	0	1	0	...
0	0	0	0	0	0	0	0	...
0	2	0	0	0	0	3	0	...
0	0	0	0	0	0	0	0	...

Implementation



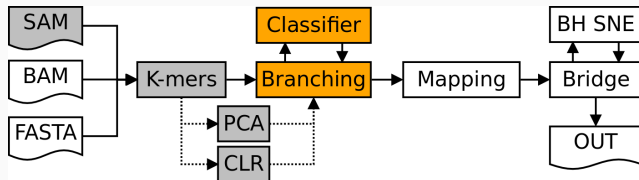
AAA	AAC	AAG	AAT	ACA	ACC	ACG	ACT	...
0	0	0	0	0	0	0	0	...
0.111	0.056	0.056	0	0	0	0.056	0	...
0	0	0	0	0	0	0	0	...
0	0.111	0	0	0	0	0.167	0	...
0	0	0	0	0	0	0	0	...

Implementation



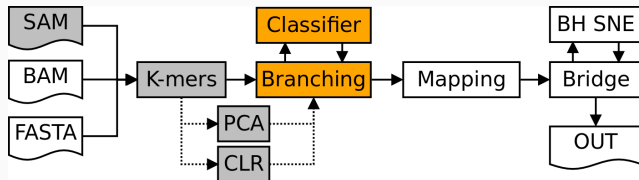
AAA	AAC	AAG	AAT	ACA	ACC	ACG	ACT	...
0	0	0	0	0	0	0	0	...
0.111	0.056	0.056	0	0	0	0.056	0	...
0	0	0	0	0	0	0	0	...
0	0.111	0	0	0	0	0.167	0	...
0	0	0	0	0	0	0	0	...

Implementation



AAA	AAC	AAG	AAT	ACA	ACC	ACG	ACT	...
0	0	0	0	0	0	0	0	...
0.111	0.056	0.056	0	0	0	0.056	0	...
0	0	0	0	0	0	0	0	...
0	0.111	0	0	0	0	0.167	0	...
0	0	0	0	0	0	0	0	...

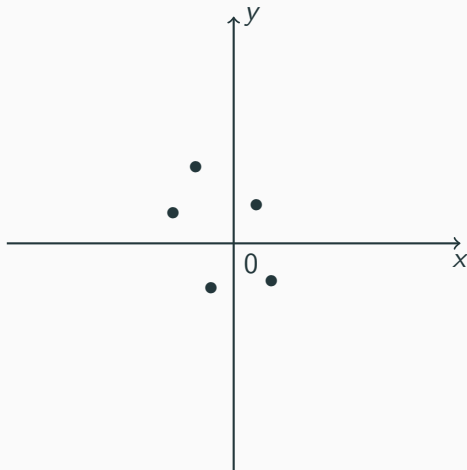
Implementation



ID	AAA	AAC	AAG	AAT	ACA	ACC	ACG	ACT	...
1	0	0	0	0	0	0	0	0	...
2	0.111	0.056	0.056	0	0	0	0.056	0	...
3	0	0	0	0	0	0	0	0	...
4	0	0.111	0	0	0	0	0.167	0	...
5	0	0	0	0	0	0	0	0	...

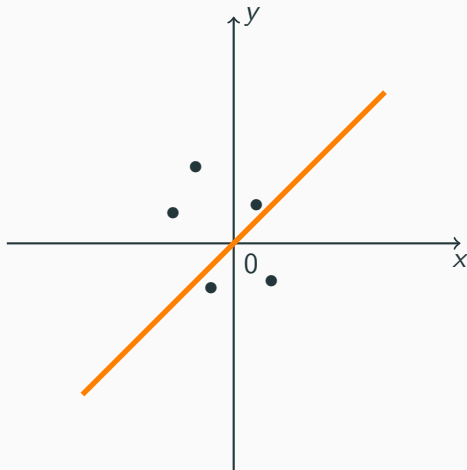
Implementation

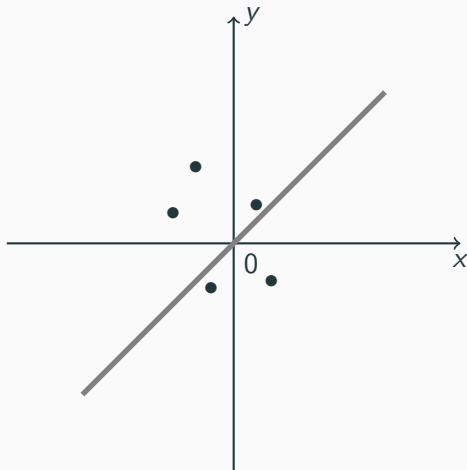
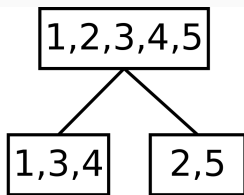
1,2,3,4,5



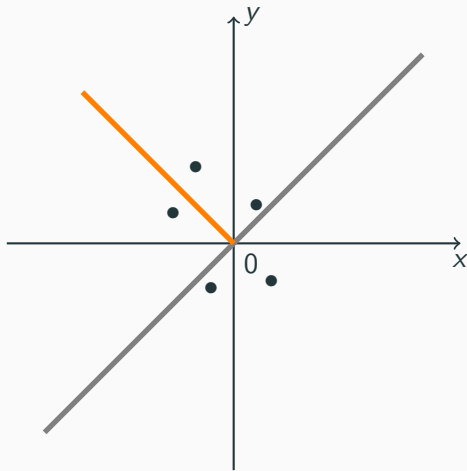
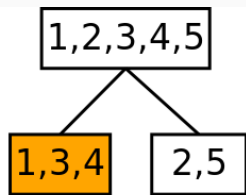
Implementation

1,2,3,4,5

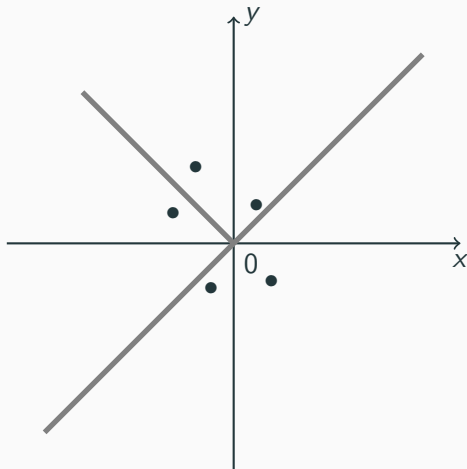
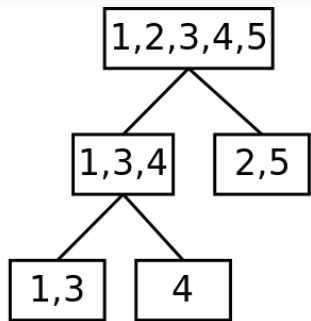




Implementation



Implementation



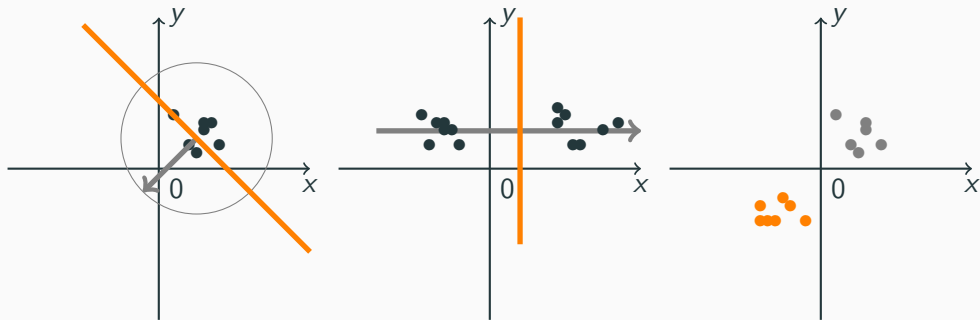
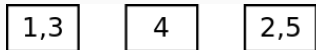
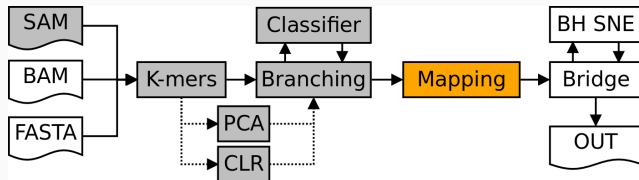


Figure 3: Mean classifier (left), Variance classifier (middle), 2-means (right)

Implementation



Implementation

1	0	1	483	42	100M	*	0	0	GCAGTCGCGTATTCATCGCG	
2	16	1	372	42	100M	*	0	0	TGCTCGAACGTCTGCAAAAG	
3	0	1	646	42	100M	*	0	0	CAGCGCCGCTGATGGCGCGG	
4	16	2	61	42	100M	*	0	0	GAGTTTTAACGACGAACGTG	
5	0	2	390	42	100M	*	0	0	AGCGGATGAGCTTTGCGCTG	

Contig ID	Read IDs
-----------	----------

1	1, 2, 3
---	---------

2	4, 5
---	------

1,3

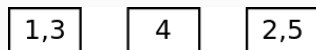
4

2,5

Implementation

Contig ID	Subspace 1	Subspace 2	Subspace 3
1	2 reads	0 reads	1 reads
2	0 reads	1 reads	1 reads

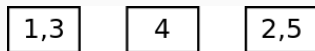
Contig ID	Read IDs
1	1, 2, 3
2	4, 5



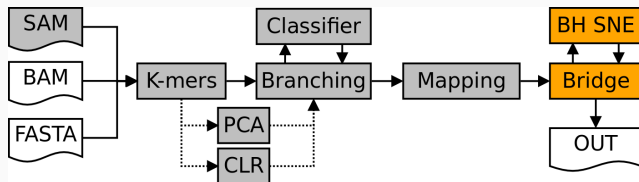
Implementation

	x1	x2	x3
Contig 1	2	0	1
Contig 2	0	1	1

Contig ID	Read IDs
1	1, 2, 3
2	4, 5

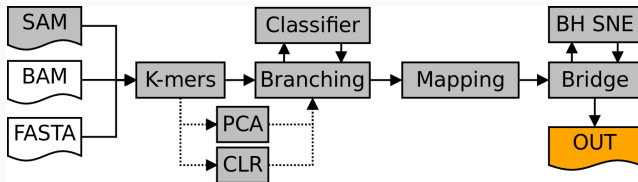


Implementation



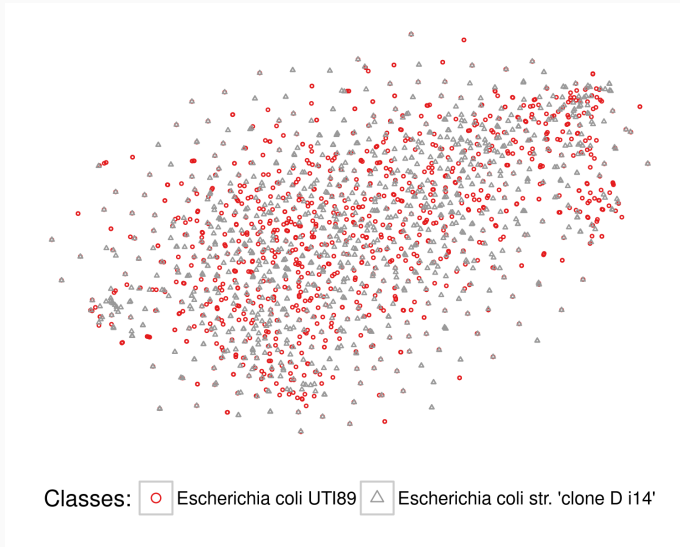
	x1	x2	x3
Contig 1	2	0	1
Contig 2	0	1	1

Implementation

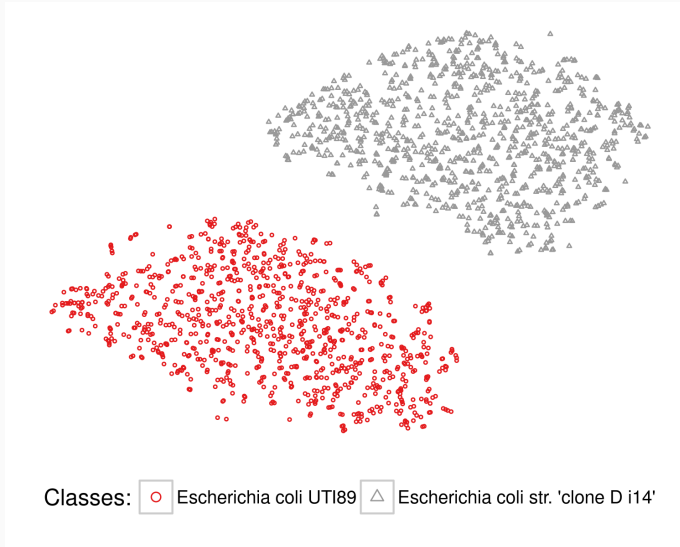


	x	y
Contig 1	0.345	1.234
Contig 2	-0.456	2.341

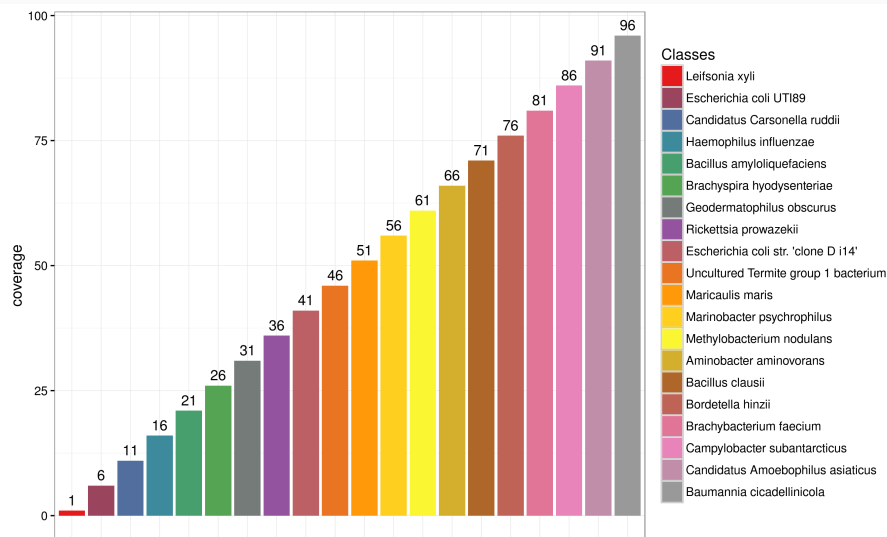
Two classes



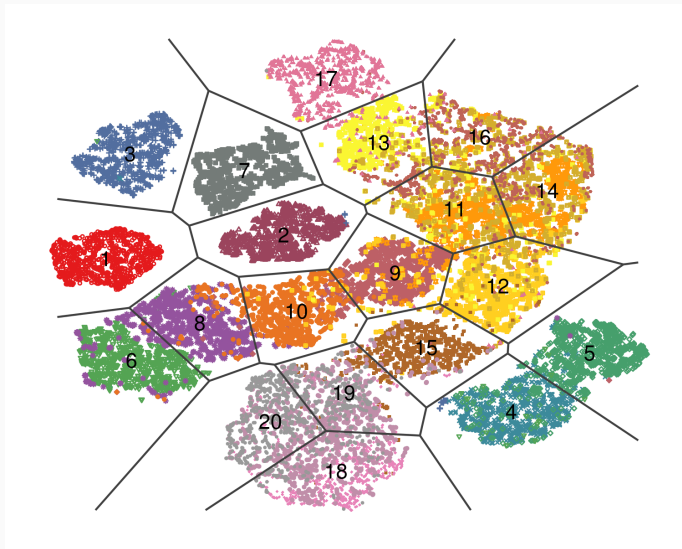
Two classes



Data Set 1 (Linear)



Data Set 1 (Linear)



Classes:

- | | |
|---|--|
| ○ | (1) <i>Leifsonia xyli</i> |
| △ | (2) <i>Escherichia coli</i> UT189 |
| + | (3) <i>Candidatus Carsonella ruddii</i> |
| × | (4) <i>Haemophilus influenzae</i> |
| ◇ | (5) <i>Bacillus amyloliquefaciens</i> |
| ▽ | (6) <i>Brachyspira hyodysenteriae</i> |
| ⊠ | (7) <i>Geodermatophilus obscurus</i> |
| * | (8) <i>Rickettsia prowazekii</i> |
| ⬠ | (9) <i>Escherichia coli</i> str. 'clone D i14' |
| ⊕ | (10) Uncultured Termite group 1 bacterium |
| ⊗ | (11) <i>Maricaulis maris</i> |
| ⊞ | (12) <i>Marinobacter psychrophilus</i> |
| ⊠ | (13) <i>Methylobacterium nodulans</i> |
| ⊠ | (14) <i>Aminobacter aminovorans</i> |
| ■ | (15) <i>Bacillus clausii</i> |
| ● | (16) <i>Bordetella hinzii</i> |
| ▲ | (17) <i>Brachybacterium faecium</i> |
| ◆ | (18) <i>Campylobacter subantarcticus</i> |
| ● | (19) <i>Candidatus Amoebophilus asiaticus</i> |
| ● | (20) <i>Baumannia cicadellinicola</i> |

Data Set 1 (Linear)

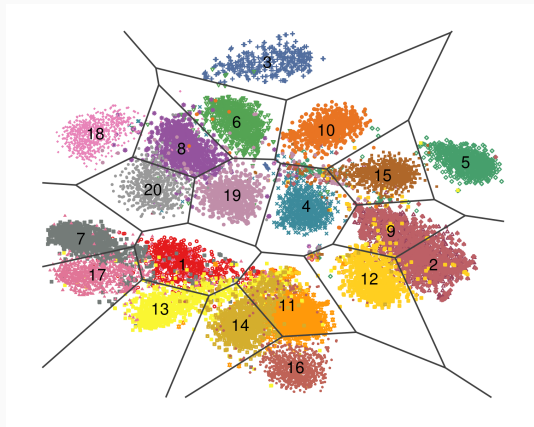


Figure 4: VizBin

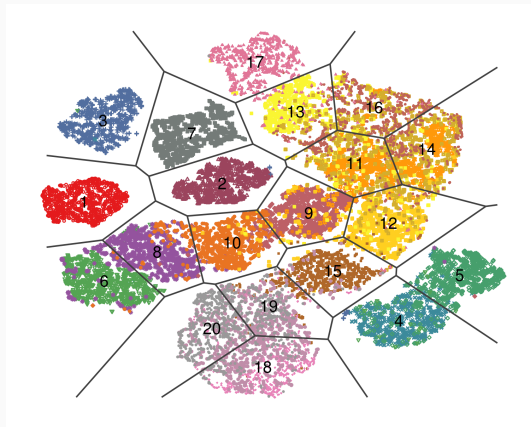


Figure 5: Implementation

Data Set 1 (Linear)

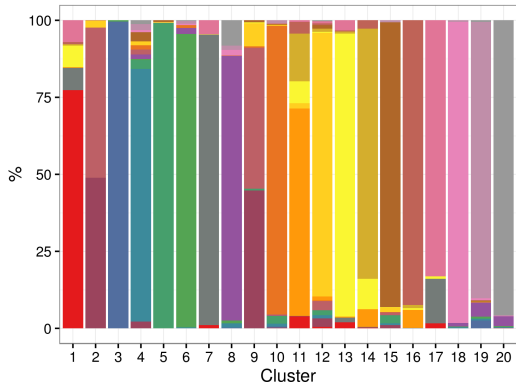


Figure 6: VizBin (Mean: 84.97%)

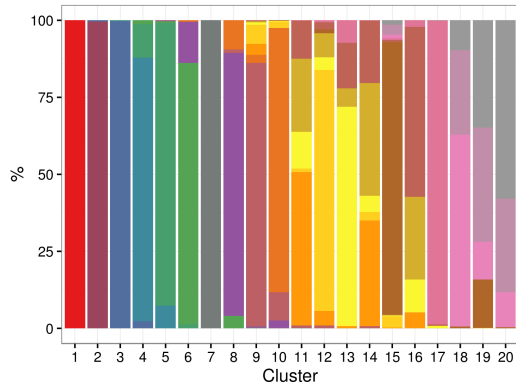
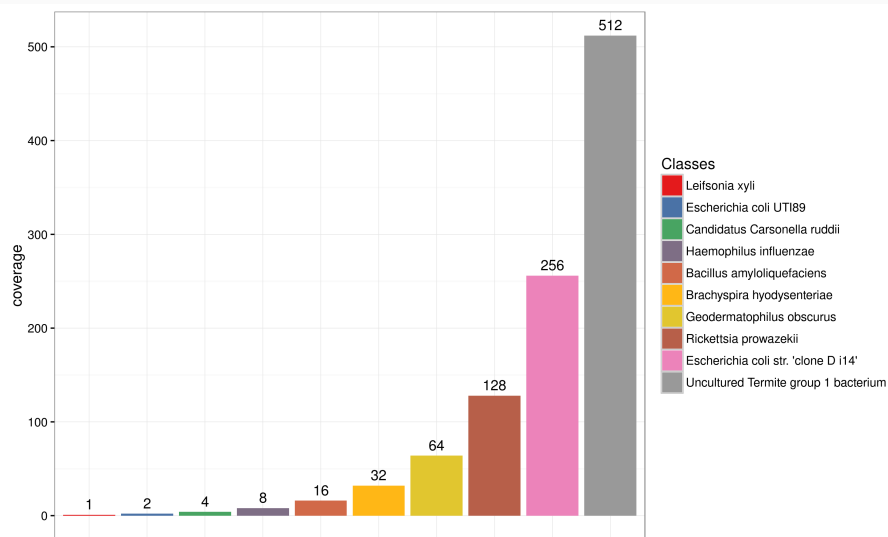
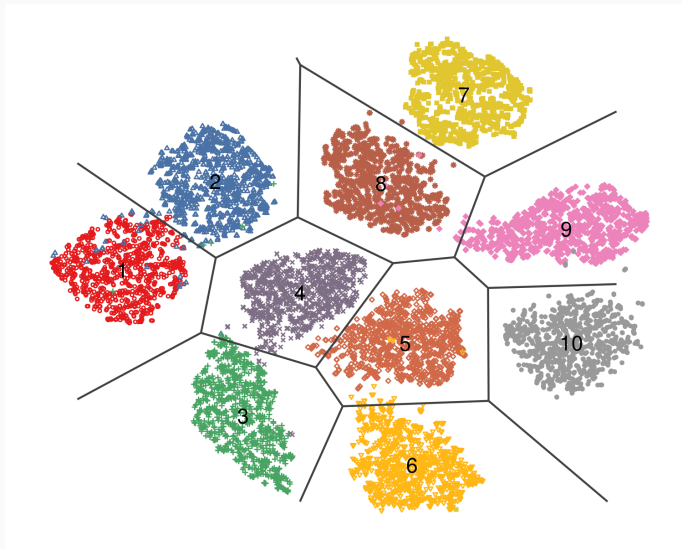


Figure 7: Implementation (Mean: 77.67%)

Data Set 2 (Exponential)



Data Set 2 (Exponential)



Classes:

- | | |
|---|--|
| ○ | (1) <i>Leifsonia xyli</i> |
| △ | (2) <i>Escherichia coli</i> UTI89 |
| + | (3) <i>Candidatus Carsonella ruddii</i> |
| × | (4) <i>Haemophilus influenzae</i> |
| ◇ | (5) <i>Bacillus amyloliquefaciens</i> |
| ▽ | (6) <i>Brachyspira hyodysenteriae</i> |
| ■ | (7) <i>Geodermatophilus obscurus</i> |
| * | (8) <i>Rickettsia prowazekii</i> |
| ◆ | (9) <i>Escherichia coli</i> str. 'clone D i14' |
| ⊕ | (10) Uncultured Termite group 1 bacterium |

Data Set 2 (Exponential)

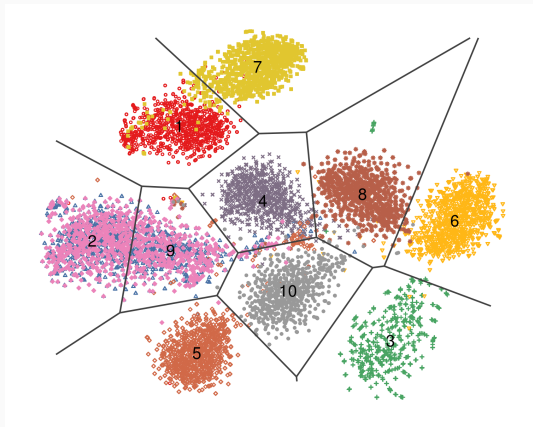


Figure 8: VizBin

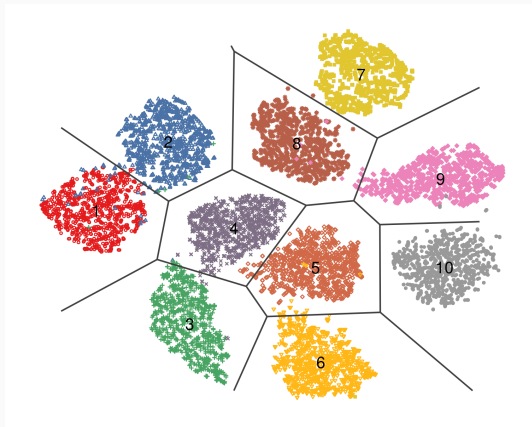


Figure 9: Implementation

Data Set 2 (Exponential)

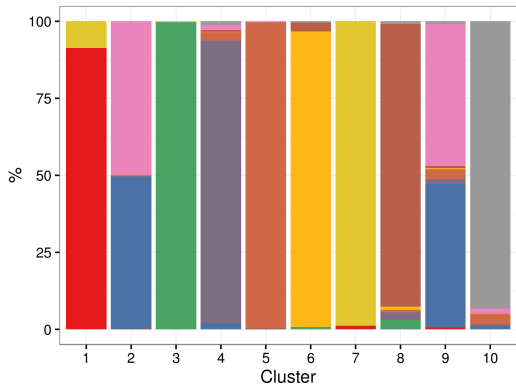


Figure 10: VizBin (Mean: 85.78%)

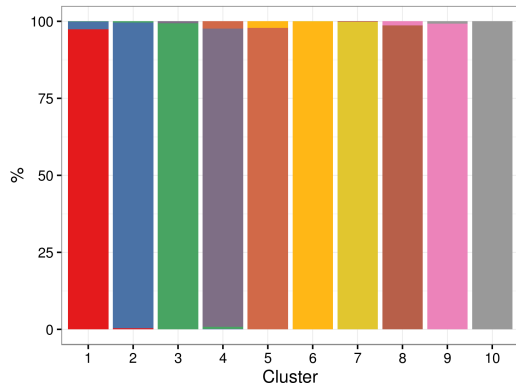
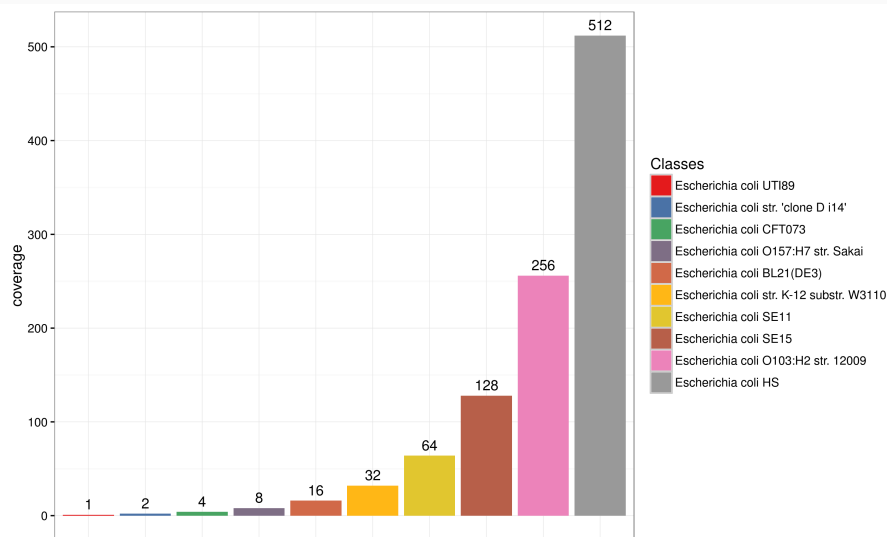
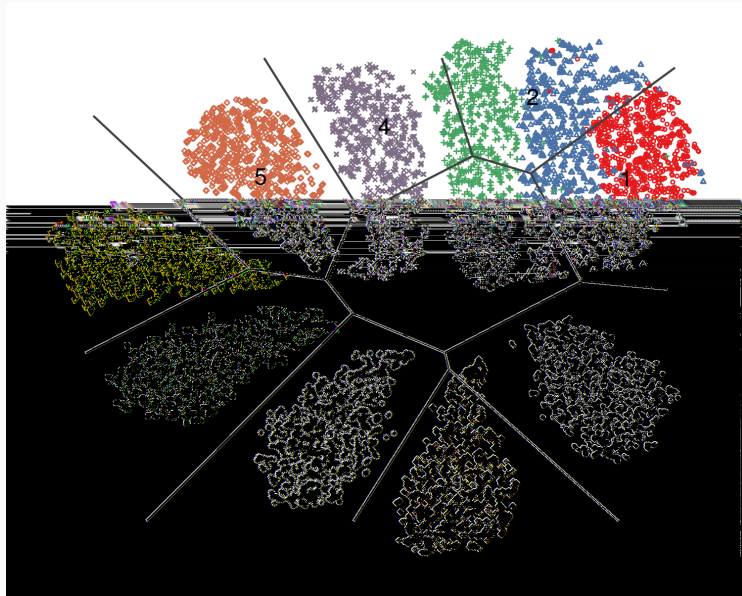


Figure 11: Implementation (Mean: 98.83%)

Data Set 3 (Similar Organisms)



Data Set 3 (Similar Organisms)



Classes:

- | | |
|---|--|
| ○ | (1) Escherichia coli UTI89 |
| △ | (2) Escherichia coli str. 'clone D i14' |
| + | (3) Escherichia coli CFT073 |
| × | (4) Escherichia coli O157:H7 str. Sakai |
| ◇ | (5) Escherichia coli BL21(DE3) |
| ▽ | (6) Escherichia coli str. K-12 substr. W3110 |
| ⊠ | (7) Escherichia coli SE11 |
| ✱ | (8) Escherichia coli SE15 |
| ⬠ | (9) Escherichia coli O103:H2 str. 12009 |
| ⊕ | (10) Escherichia coli HS |

Data Set 3 (Similar Organisms)

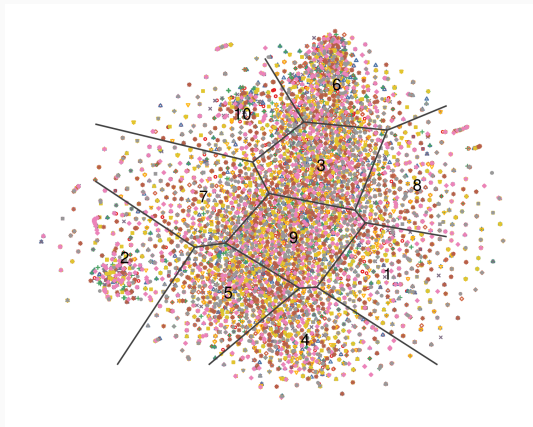


Figure 12: VizBin

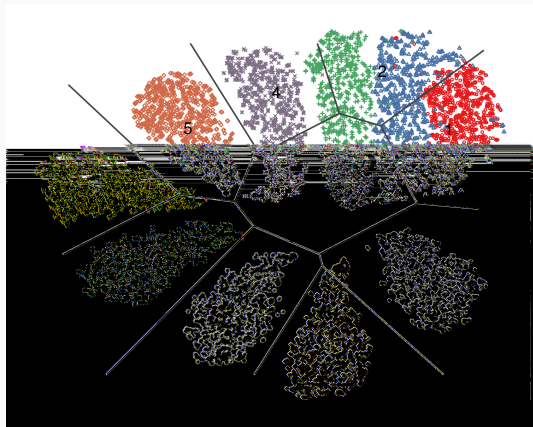


Figure 13: Implementation

Data Set 3 (Similar Organisms)

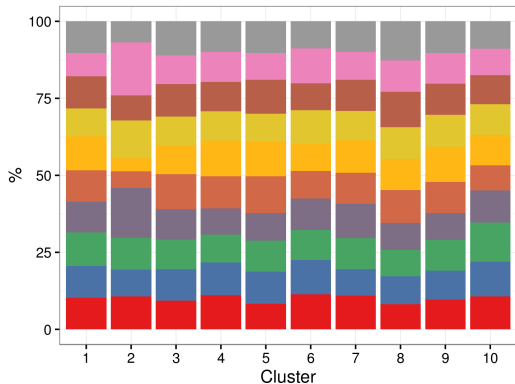


Figure 14: VizBin (Mean: 9.83%)

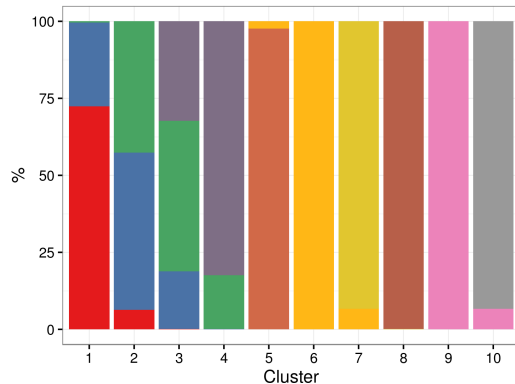
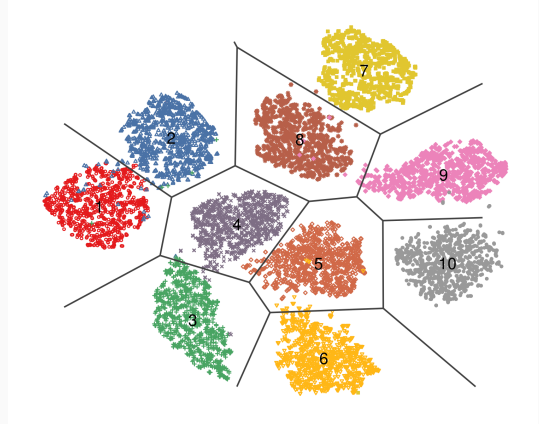
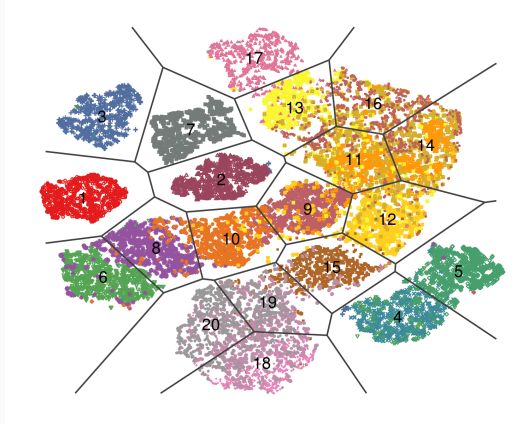
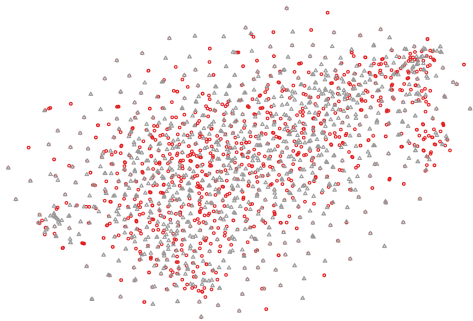


Figure 15: Implementation (Mean: 83.89%)

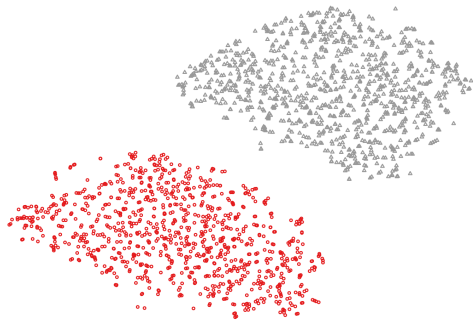
Conclusion



Conclusion



Classes: ○ Escherichia coli UTI89 △ Escherichia coli str. 'clone D i14'



Classes: ○ Escherichia coli UTI89 △ Escherichia coli str. 'clone D i14'



Noble PA, Citek RW and Ogunseitan OA.

Tetranucleotide frequencies in microbial genomes.

Electrophoresis

19(4), p. 528–35



Yu-Wei Wu, Yung-Hsu Tang, Susannah G Tringe, Blake A Simmons and Steven W Singer

MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm.

Microbiome

2:26



J.Alneberg et al.

Binning metagenomic contigs by coverage and composition.

Nature Methods

11, p. 1144–1146

Appendix

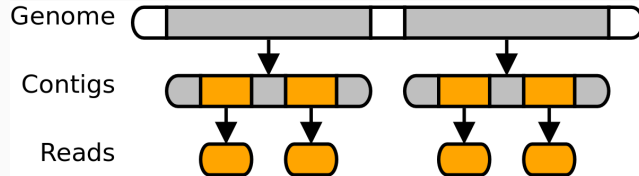
Name	Assembly Accession
<i>Leifsonia xyli</i> CTCB07	GCF_000007665.1_ASM766v1
<i>Candidatus Carsonella ruddii</i> HT	GCF_000287295.1_ASM28723v1
<i>Haemophilus influenzae</i> 86-028NP	GCF_000016485.1_ASM1648v1
<i>Bacillus amyloliquefaciens</i> TA208	GCF_000195515.1_ASM19551v1
<i>Brachyspira hyodysenteriae</i> WA1	GCF_000022105.1_ASM2210v1
<i>Geodermatophilus obscurus</i> DSM 43160	GCF_000025345.1_ASM2534v1
<i>Rickettsia prowazekii</i> Dachau	GCF_000277225.1_ASM27722v1
Uncultured Termite group 1 bacterium Rs-D17	GCF_000146025.2_ASM14602v1
<i>Maricaulis maris</i> MCS10	GCF_000014745.1_ASM1474v1
<i>Marinobacter psychrophilus</i> 20041	GCF_001043175.1_ASM104317v1
<i>Methylobacterium nodulans</i> ORS 2060	GCF_000022085.1_ASM2208v1
<i>Aminobacter colombiense</i> DSM 12261	GCF_000025885.1_ASM2588v1
<i>Bacillus clausii</i> KSM-16	GCF_000009825.1_ASM982v1
<i>Bordetella hinzii</i> F582	GCF_001078275.1_ASM107827v1
<i>Brachybacterium faecium</i> DSM 4810	GCF_000023405.1_ASM2340v1
<i>Campylobacter subantarcticus</i> LMG 24374	GCF_000816265.1_ASM81626v1
<i>Candidatus Amoebophilus asiaticus</i> 5a2	GCF_000020565.1_ASM2056v1
<i>Candidatus Baumannia cicadellincola</i> Hc	GCF_000013185.1_ASM1318v1

continued on next slide

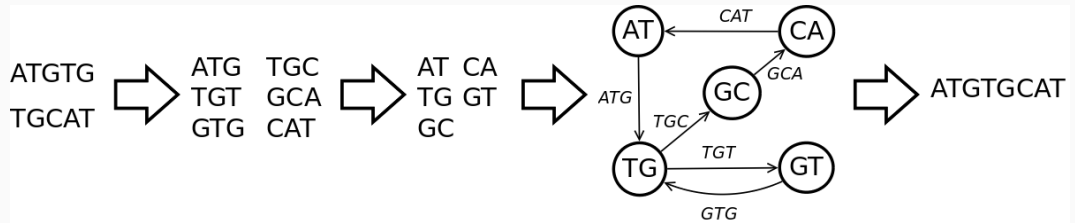
Appendix

Name	Assembly Accession
<i>Escherichia coli</i> UTI89	GCF_000013265.1_ASM1326v1
<i>Escherichia coli</i> clone D i14	GCF_000233895.1_ASM23389v1
<i>Escherichia coli</i> CFT073	GCF_000007445.1_ASM744v1
<i>Escherichia coli</i> O157:H7 str. Sakai	GCF_000008865.1_ASM886v1
<i>Escherichia coli</i> BL21(DE3)	GCF_000009565.1_ASM956v1
<i>Escherichia coli</i> str. K-12 substr. W3110	GCF_000010245.2_ASM1024v1
<i>Escherichia coli</i> SE11	GCF_000010385.1_ASM1038v1
<i>Escherichia coli</i> SE15	GCF_000010485.1_ASM1048v1
<i>Escherichia coli</i> O103:H2 str. 12009	GCF_000010745.1_ASM1074v1
<i>Escherichia coli</i> HS	GCF_000017765.1_ASM1776v1

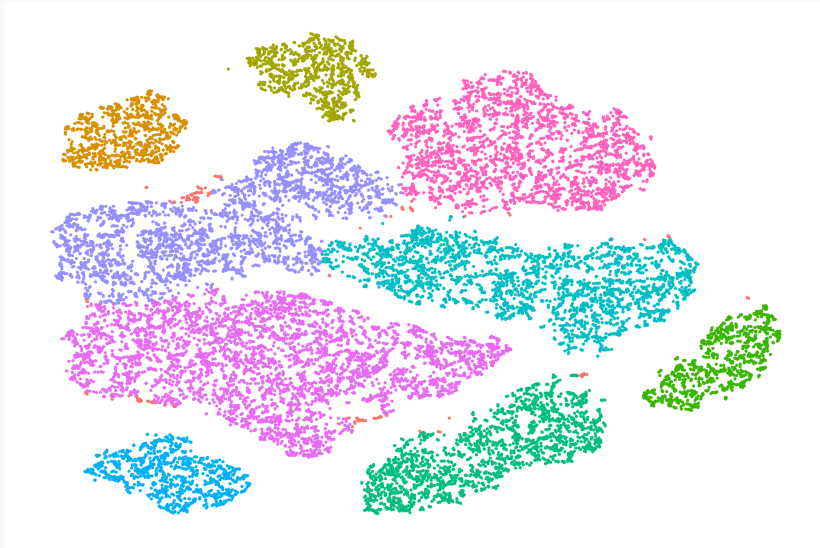
Appendix



Appendix



Appendix



Appendix

High Dimensional:

$$p_{i|i} = 0 \quad p_{j|i} = \frac{\exp -d(x_i, x_j)^2 / 2\sigma_i^2}{\sum_{k \neq i} \exp -d(x_i, x_k)^2 / 2\sigma_i^2} \quad (1)$$

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N} \quad (2)$$

Low Dimensional (Student-t):

$$q_{ij} = \frac{\left(1 + \|y_i - y_j\|^2\right)^{-1}}{\sum_{k \neq l} \left(1 + \|y_k - y_l\|^2\right)^{-1}} \quad (3)$$

Combined:

$$\text{KL}(P||Q) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}} \quad (4)$$