

Reconciliation between Operational Taxonomic Units and Species Boundaries

Mohamed Mysara (SCK•CEN, VUB, KU Leuven)

Peter Vandamme (UGent)

Natalie Leys (SCK•CEN)

Jeroen Raes (KU Leuven)

Pieter Monsieurs (SCK•CEN)

Unit of Microbiology, Belgian Nuclear Research Centre SCK•CEN

Department of Biomedical Science, KU Leuven

Department of Bioscience Engineering, Vrije Universiteit Brussel VUB

Department of Biochemistry and microbiology, Ghent University.



Two different mode to use DynamiC

- Training Mode:

[specified via `_m train`]

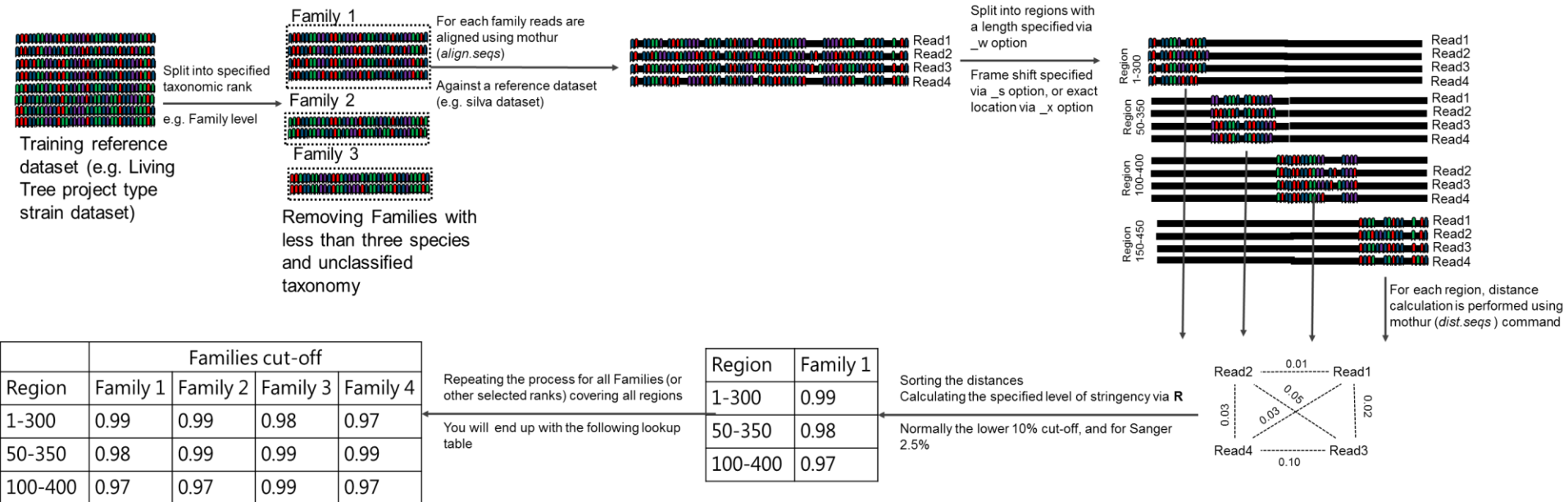
To build a lookup table with the modified cut-off per region within 16S rRNA for each taxonomic rank

- Test Mode:

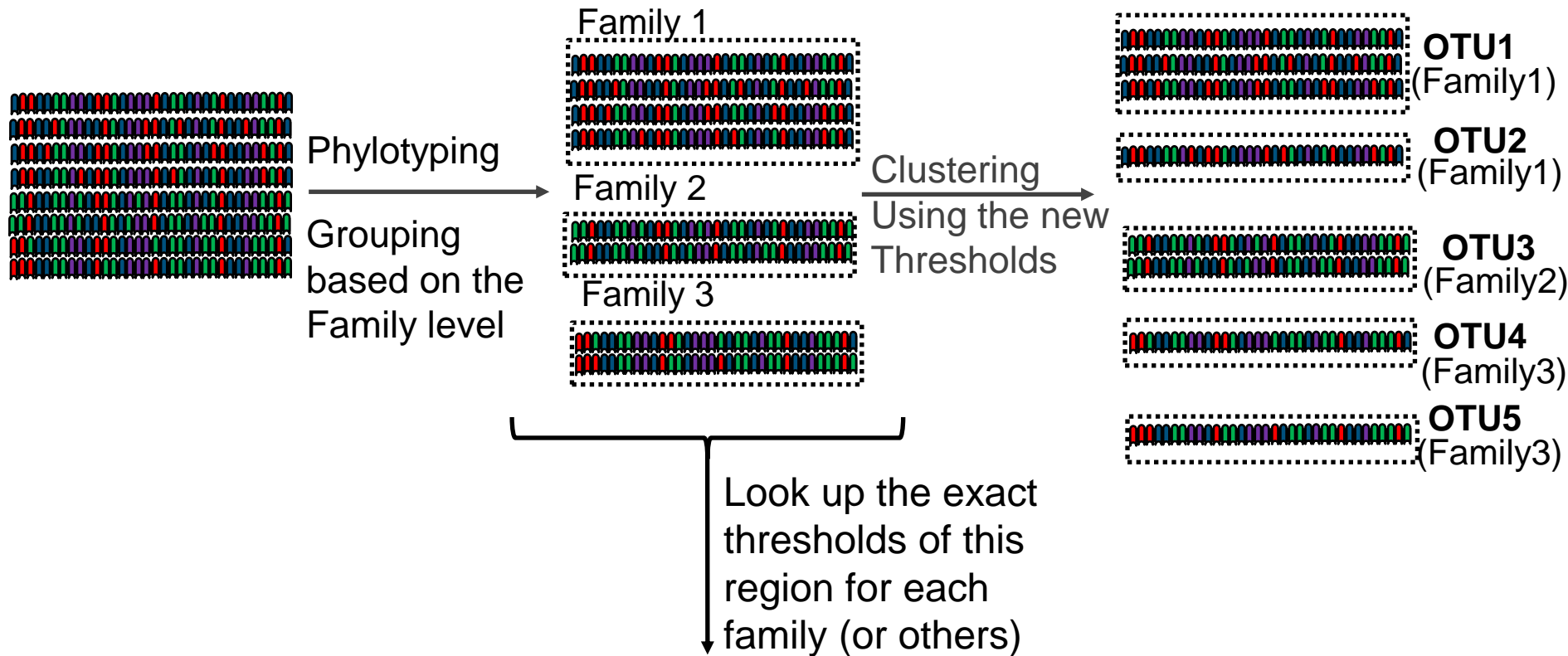
- [specified via `_m test`]

To utilize the calculated cut-offs within the lookup table to cluster your 16s rRNA gene sequencing reads.

Concept of DynamiC Training mode



Concept of DynamiC testing mode



The lookup table calculated in the training mode

	Families cut-off			
Region	Family 1	Family 2	Family 3	Family 4
1-300	0.99	0.99	0.98	0.97
50-350	0.98	0.99	0.99	0.99
100-400	0.97	0.97	0.99	0.97

Requirements

- To use DynamiC you need to have:

- R statistical language installed
- Perl installed

They can be installed from <https://www.perl.org/>, <https://www.r-project.org/> and <http://www.mothur.org/> respectively.

- Also you need to download:

- Reference dataset for calculating the Lookup table such as living tree project (type strain) dataset (with the taxonomic information)
- SILVA alignment dataset

They can be downloaded from

<http://www.arb-silva.de/projects/living-tree/>

http://www.mothur.org/wiki/Silva_reference_files

Running the Training mode

- `./DynamiC.run`

- `_m train`

- `_f /YOUR_PATH/LTP.fasta` [as fasta downloaded from LTP]

- `_t /YOUR_PATH/LTP.families` [as taxonomy extracted from LTP]

- `_r /YOUR_PATH/silva.seed_v119.align` [as downloaded from mothur]

- `_o /YOUR_OUTPUT_PATH`

- `_w 500` [Window size]

- `_s 100` [Frame shift]

- `_u 0.01` [Upper allowed cut-off]

- `_h 0.1` [level of stringency]

- `_p 10` [Number of processors]

- Other options exists, please see the readme for more information

- The main output of the training mode is the lookup table, **it can be found in the Final folder created within the output directory**

Running the Testing mode

- `./DynamiC.run`

- `_m test`

- `_w /YOUR_PATH/Final/LookUp_table` [created via training mode]

- `_t /YOUR_PATH/LTP.families` [as taxonomy extracted from LTP]

- `_r /YOUR_PATH/silva.seed_v119.align` [as downloaded from mothur]

- `_o /YOUR_OUTPUT_PATH`

- `_f seqs.unique.fasta` [fasta file of your dereplicated sequences]

- `_n seqs.names` [name file indicating the dereplication information, created via `unique.seqs` command]

- `_t seqs.unique.Tax.wang.taxonomy` [Taxonomy file calculated via `classify.seqs` command in mothur using the same data used for training the LTP]

- `_k 8 _o HMP_Results`

- Other options exists, please see the readme for more information

- The main output of the testmode is the list file, **it can be found in the Final folder created within the output directory**

Questions and Citation

- If you have any question please contact us on these emails:
 - Mohamed.mysara@gmail.com
 - pieter.monsieurs@sckcen.be
- Or via Github
 - <https://github.com/M-Mysara/DynamiC/wiki>

● Citing

If you are going to use DynamiC, please cite it with the included software (mothur):

- **Mysara M., P. Vandamme, N. Leys, J. Raes and P. Monsieurs, 2016, Reconciliation between Operational Taxonomic Units and Species Boundaries, in preparation.**
- **Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and environmental microbiology 75:7537–41.**

Copyright © 2014 - SCK•CEN

PLEASE NOTE!

This presentation contains data, information and formats for dedicated use ONLY and may not be copied, distributed or cited without the explicit permission of the SCK•CEN. If this has been obtained, please reference it as a "personal communication. By courtesy of SCK•CEN".

SCK•CEN

Studiecentrum voor Kernenergie
Centre d'Etude de l'Energie Nucléaire
Belgian Nuclear Research Centre

Stichting van Openbaar Nut
Fondation d'Utilité Publique
Foundation of Public Utility

Registered Office: Avenue Herrmann-Debrouxlaan 40 – BE-1160 BRUSSELS
Operational Office: Boeretang 200 – BE-2400 MOL

