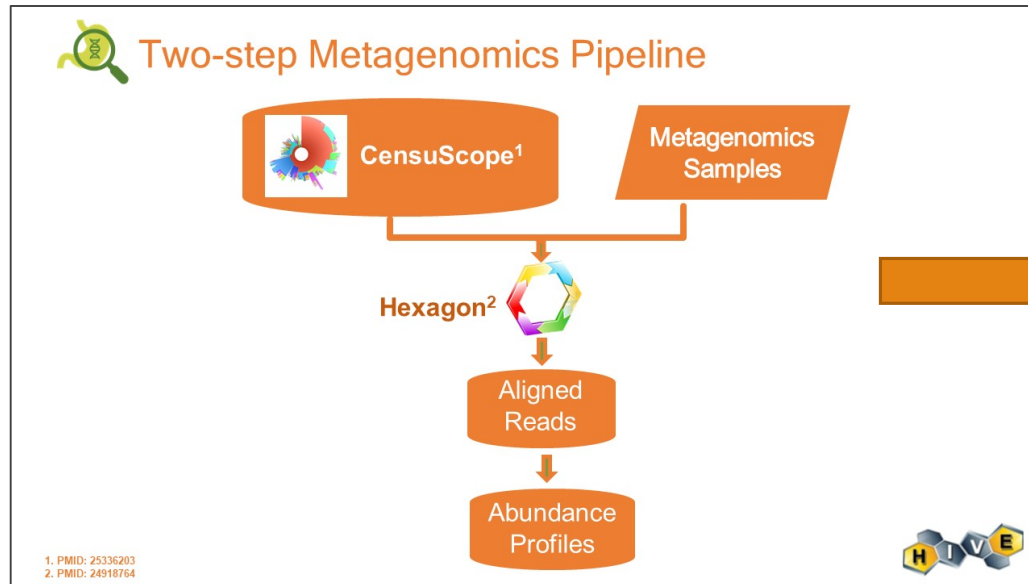# Introduction to Microbiome Bioinformatics: Building a Healthy Gut Database and Microbial Populations in Epilepsy

JONATHON KEENEY, PH.D

ASSISTANT RESEARCH PROFESSOR, GEORGE WASHINGTON UNIVERSITY

Establishing a Clinically Applicable Metagenomics Analysis Pipeline

Machine Learning and Predicative Modeling

- **Random Forests Algorithm**
- **Predictor Importance Plots**
- **Partial Dependency Plots**
- **Principle Component Analysis**

Expansion into New Areas of Microbiome Research: Epilepsy

The gut microbiota mediates the anti-seizure effects of the ketogenic diet

Christine A. Olson, Helen E. Vuong, Jessica M. Yano, Qingxing Y. Liang, David J. Nusbaum, and Elaine Y. Hsiao[*]

▸ Author information ▸ Copyright and License information Disclaimer

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed ▾   Advanced

Format: Abstract ▾   Send to ▾

The ketogenic diet influences taxonomic and functional composition of the gut microbiota in children with severe epilepsy.

Lindefeldt M[1], Eng A[2], Darban H[3], Bjerkner A[4], Zetterström CK[5], Allander T[4], Andersson B[3], Borenstein E[2,6,7,8,9], Dahlin M[1], Prast-Nielsen S[10].

⊖ Author information

1   1Neuropediatric Department, Astrid Lindgren Children's Hospital, Karolinska Hospital, Stockholm, Sweden.

Expansion into New Areas of Microbiome Research: Epilepsy

The gut microbiota mediates the anti-seizure effects of the ketogenic diet

Christine A. Olson, Helen E. Vuong, Jessica M. Yano, Qingxing Y. Liang, David J. Nusbaum, and Elaine Y. Hsiao*

▸ Author information ▸ Copyright and License information Disclaimer

The publisher's final edited version of this article is available at Cell
This article has been corrected. See Cell. 2018 July 12; 174(2): 497.
See commentary "A Gut Feeling About Seizures" in Epilepsy Curr, volume 18 on page 389.

See other articles in PMC that cite the published article.

PubMed.gov
US National Library of Medicine
National Institutes of Health

Format: Abstract ▾

NPJ Biofilms Microbiomes. 2019 Jan 23;5:5. doi: 10.1038/s41522-018-0073-2. eCollection 2019.

The ketogenic diet influences taxonomic and functional composition of the gut microbiota in children with severe epilepsy.

Lindefeldt M[1], Eng A[2], Darban H[3], Bjerkner A[4], Zetterström CK[5], Allander T[4], Andersson B[3], Borenstein E[2,6,7,8,9], Dahlin M[1], Prast-Nielsen S[10].

⊖ Author information
1 1Neuropediatric Department, Astrid Lindgren Children's Hospital, Karolinska Hospital, Stockholm, Sweden.
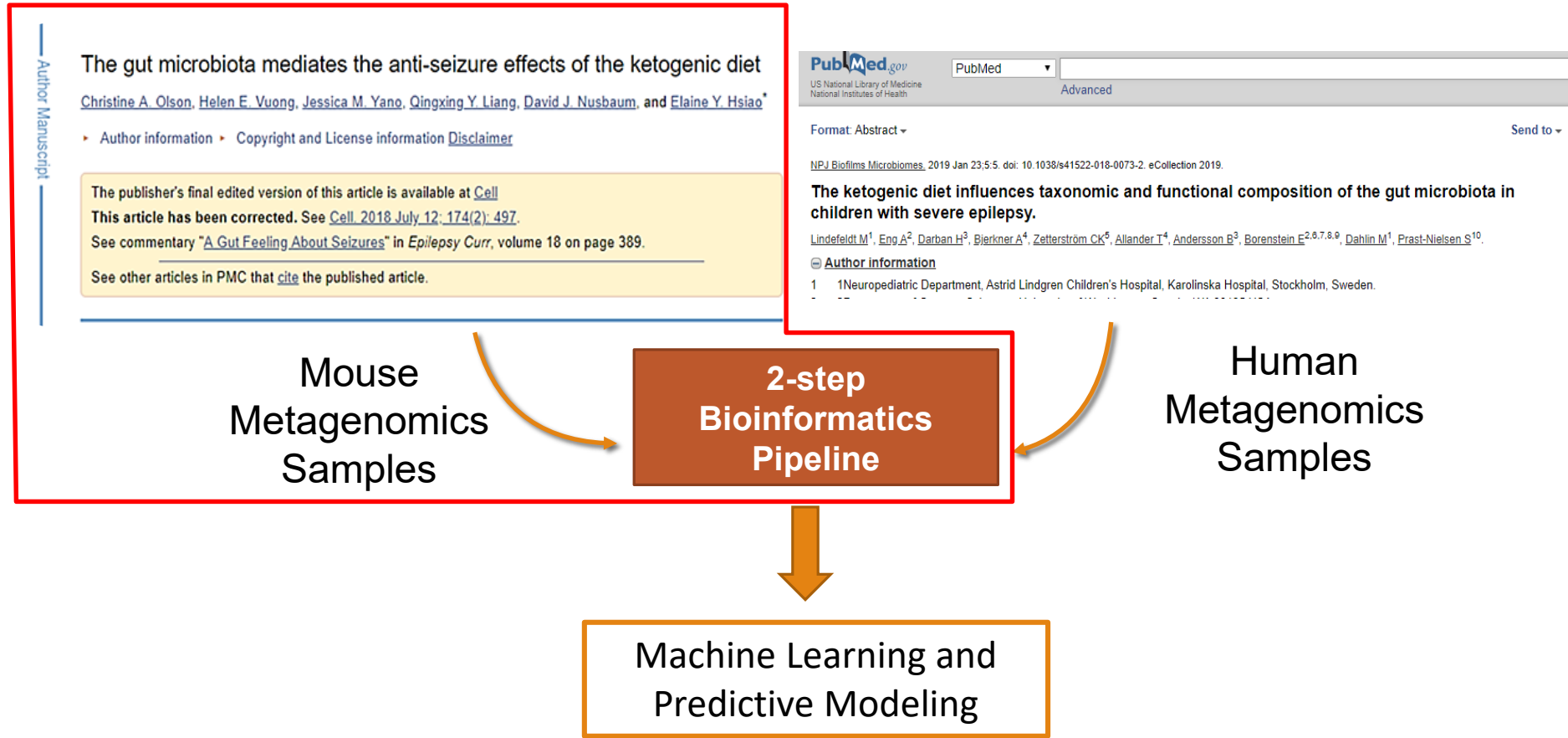
Mouse Metagenomics Samples

2-step Bioinformatics Pipeline

Human Metagenomics Samples

Machine Learning and Predictive Modeling

Expansion into New Areas of Microbiome Research: Epilepsy

- Mouse induced-seizure models of intractable epilepsy
- Experimental:
  - 6:1 protein:fat ratio KD diet
- Control:
  - Vitamin/mineral matched control diet
- Time points at 0, 4, 8, and 14 days
- The gut microbiome is necessary and sufficient for seizure protection

The gut microbiota mediates the anti-seizure effects of the ketogenic diet

Christine A. Olson, Helen E. Vuong, Jessica M. Yano, Qingxing Y. Liang, David J. Nusbaum, and Elaine Y. Hsiao[*]

▸ Author information ▸ Copyright and License information Disclaimer

The publisher's final edited version of this article is available at Cell

This article has been corrected. See Cell. 2018 July 12; 174(2): 497.

See commentary "A Gut Feeling About Seizures" in *Epilepsy Curr*, volume 18 on page 389.

See other articles in PMC that cite the published article.

Associated Data

▸ Supplementary Materials

SUMMARY

Go to: ☑

The ketogenic diet (KD) is used to treat refractory epilepsy, but the mechanisms underlying its neuroprotective effects remain unclear. Here we show that the gut microbiota is altered by the KD and required for protection against acute electrically-induced seizures and spontaneous tonic-clonic seizures in two mouse models. Mice treated with antibiotics or reared germ-free are resistant to KD-mediated seizure protection. Enrichment of, and gnotobiotic co-colonization with, KD-associated *Akkermansia* and *Parabacteroides* restores seizure protection. Moreover, transplantation of the KD gut microbiota, and treatment with *Akkermansia* and *Parabacteroides*, each confer seizure protection to mice fed a control diet.

Increased Species Diversity
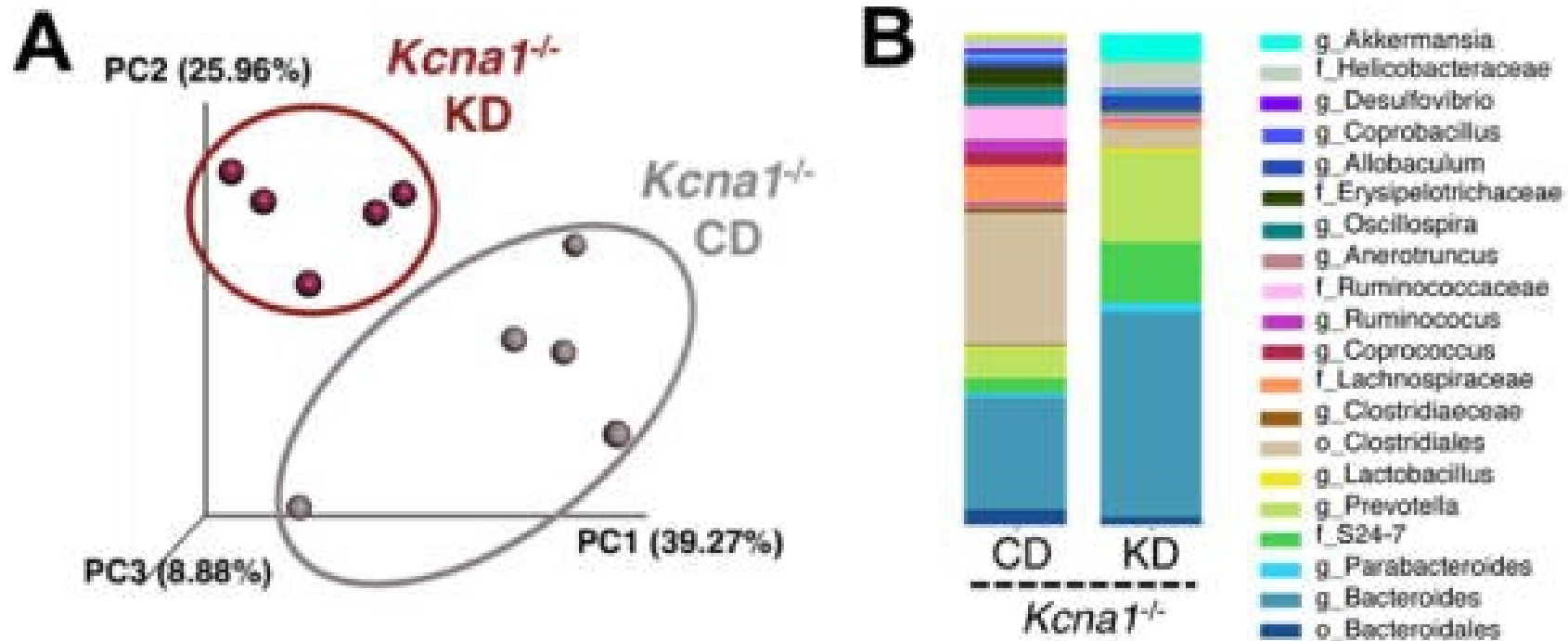
# High-throughput Integrated Virtual Environment (HIVE)

- Platform for bioinformatic analysis of very big data
  - GUI wrapper for command line scripts
- Highly parallelizable
  - Storage and compute are distributed to many powerful servers
- Developed by FDA
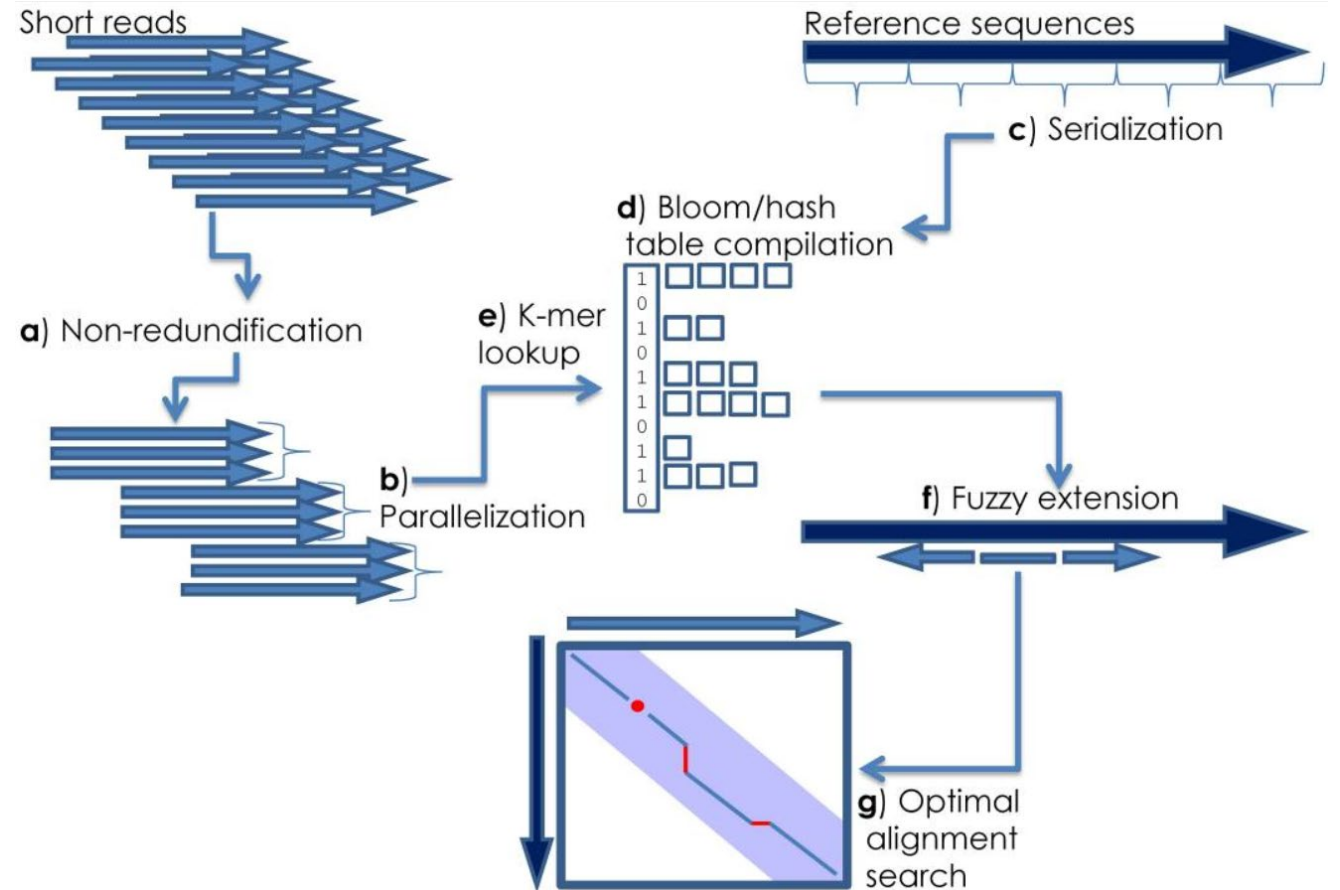 - GWU maintains only publicly available version
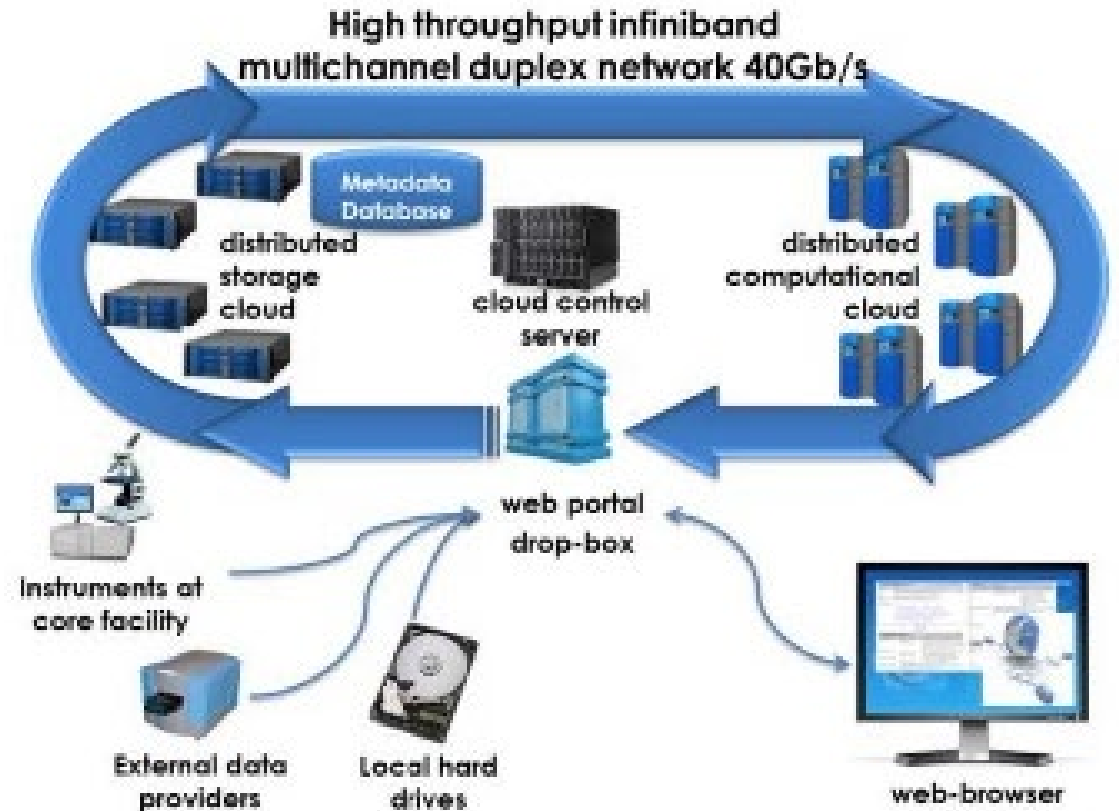
Step 1: Censuscope

- Sample a data set to estimate its composition
- Subsample Aggregating ("SubAgg")
  - Sample (m) the data set (n) such that m << n
  - Repeat i times
- "Wisdom of Crowds" effect
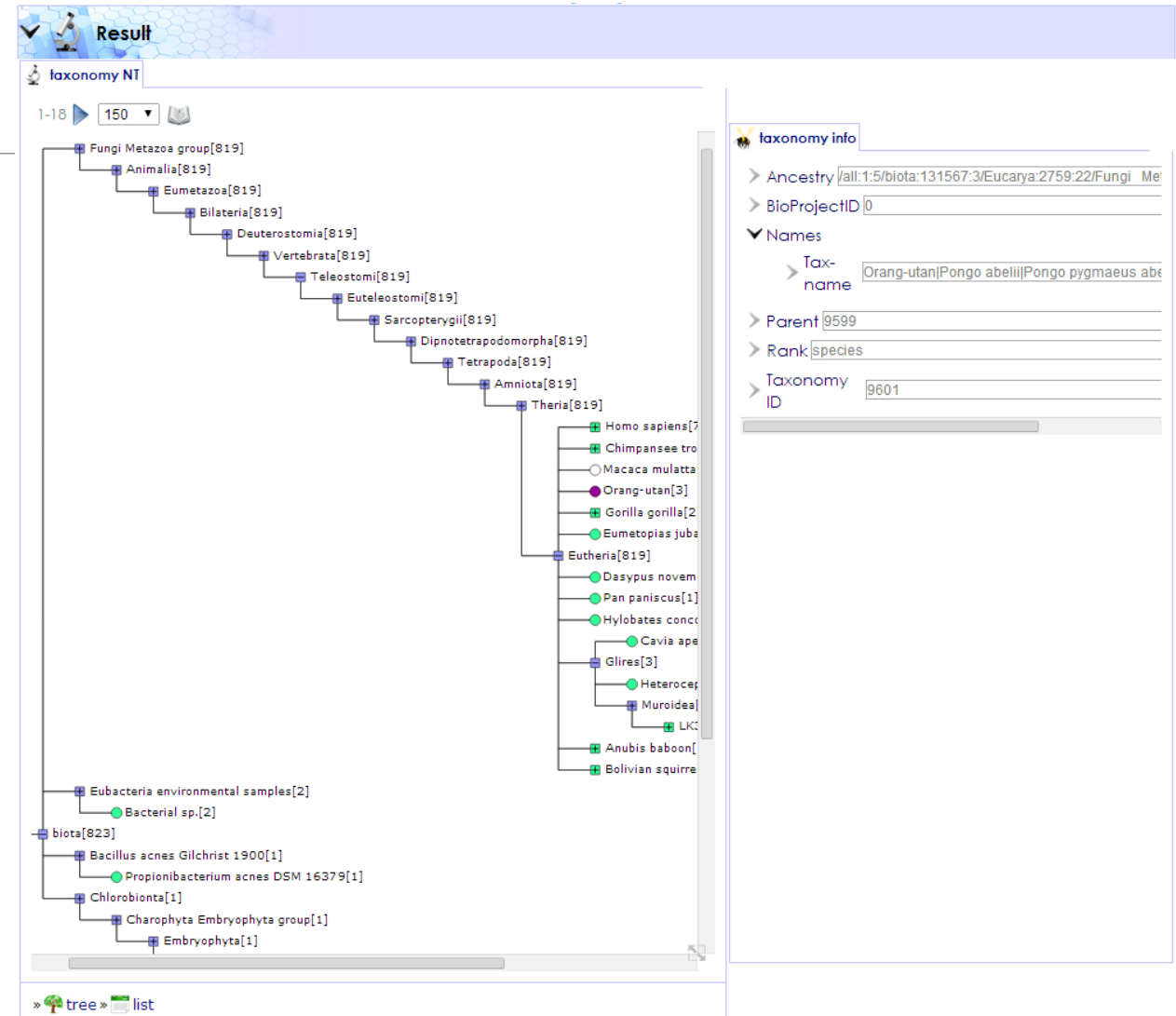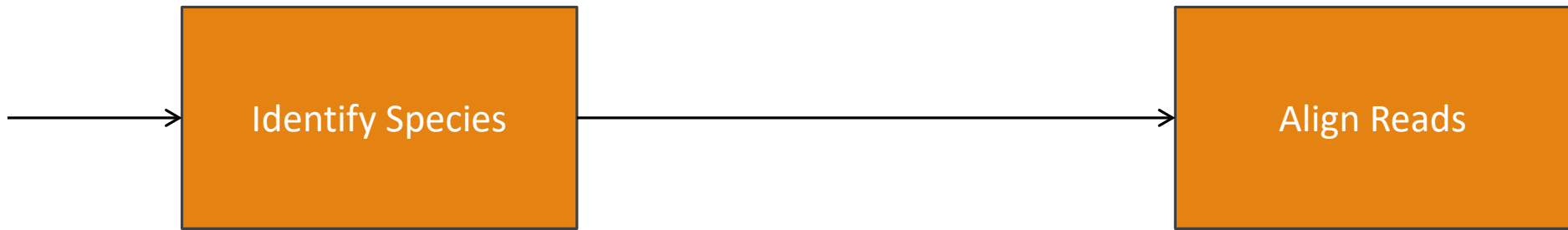
Step 2: Hexagon

- Based on Smith-Waterman alignment
  - Local alignment
- Modifications:
  - Non-redundification
    - Lower memory footprint
  - Highly parallelizable
    - Adapted for HIVE architecture
  - Builds a lookup table
    - K-mers are represented by binary digits
- Results in precomputed

Short reads

Reference sequences

c) Serialization

d) Bloom/hash table compilation

a) Non-redundification

e) K-mer lookup

b) Parallelization

f) Fuzzy extension

g) Optimal alignment search

2 Step Pipeline

Raw sequence data from SRA

Identify Species

Align Reads

Step 1: use **Censuscope** to identify species that are present in each sample. Retrieve the genomes of those species and compile them into one file that will be the reference database.

Step 2: Align the reads to the newly created reference database using **Hexagon**. The number of reads that map to each genome will estimate the abundance of each species.

https://hive.aws.biochemistry.gwu.edu/dna.cgi?cmd=home

Please go to HIVE...

id	Reference	Hits	% abundance	Hits Unique	Length	RPKM	Density
16025	NR_144616.1 Muribaculum intestinale strain YL27 16S ribosomal RNA, partialsequence.	26701	13.82%	12718	1460	94654.3	4559
14701	NR_043318.1 Alistipes onderdonkii strain WAL 8169 16S ribosomal RNA, partialsequence.	22589	11.69%	8363	1440	81189.6	3920.4
6589	NR_114079.1 Escherichia fergusonii strain NBRC 102419 16S ribosomal RNA,partial sequence.	14095	7.30%	8409	1467	49728	2401.4
19648	NR_116340.1 Brenneria alni strain pvfi20 16S ribosomal RNA, partial sequence.	11712	6.06%	772	1224	49524	2391.9
3889	NR_144745.1 Culturomica massiliensis strain Marseille-P2698 16S ribosomal RNA,partial sequence.	4297	2.22%	1604	1492	14906	716.9
0	Unaligned	2993	1.55%	2782	-	-	
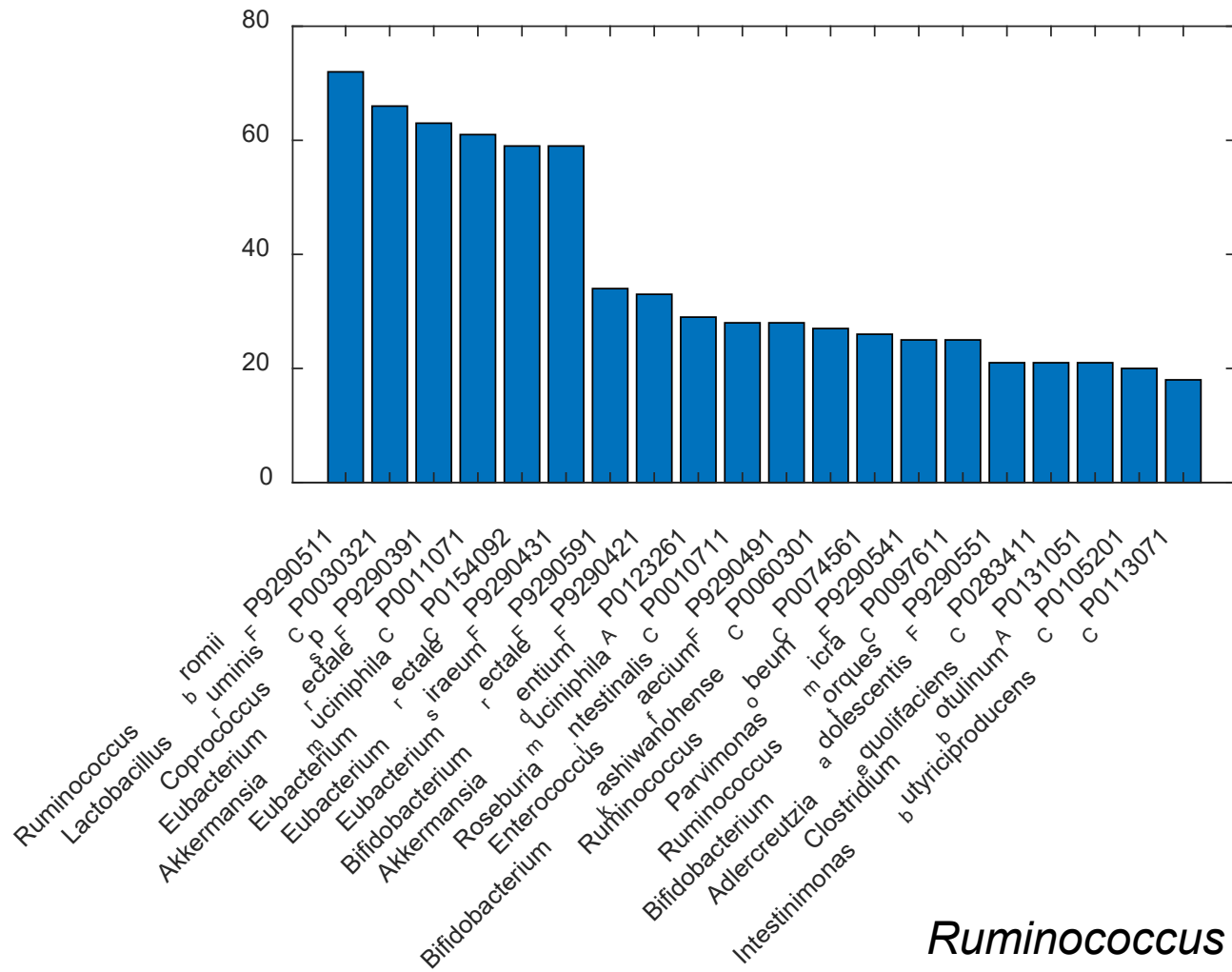
Predictor Importance Plot: contributors to seizure frequency response

Ruminococcus bromii and *Lactobacillus ruminis* are two most important predictors of seizure frequency response

Acknowledgements