



AARHUS UNIVERSITY

Explore the pattern of presence, abundance and diversity of the *Prevotella* genus in the gut microbiome of a diverse cohort of mammals.

By Ditte Torlyn

201408508

June 2021

Bioinformatics

Århus University

Supervised by Louise Bruun Thingholm, Mikkel Heide Schierup and Malte Rühlemann

Abstract

The gut microbiota and microbiome have recently become of great interest in relation to the human health, but many areas of the field are still uninvestigated.

This study investigates the patterns of presence, absence, abundance and diversity of the *Prevotella* genus in the gut microbiota of humans from Germany and Africa and wild chimpanzees, bonobo and gorillas from Africa and captive chimpanzees and orangutangs from zoological gardens in Germany. This was done with shotgun metagenomic sequencing data on 437 stool samples and extensive data processing including assembly of 18,355 metagenomic assembled genomes (MAGs or bins) with a pipeline made by Kiel University, followed by annotation with the tool GTDB-tk and abundance estimate with the tool MetaWRAP.

The genomes that annotated to the genus *Prevotella* were selected, quality controlled with CheckM and dereplicated with dRep. The MASH and ANI distances between pairs of genomes were computed and analyzed in RStudio in relation to the abundance's estimates and alpha and beta diversity measures.

The microbiotas of the German cohort were shown to be uniform and completely missing the *Prevotella* genus in some samples ($n=61$; $\sim 26\%$). The diversity of the microbiome was reduced in the German group compared to the other groups, but the difference in the microbiome between the samples was large, and the abundances were irregular across the group. The microbiota of Africans showed great diversity within each sample, but the group diversity was more compact than the apes' microbiota and their microbiome showed similar evolutionary distances as the Germans' microbiomes in terms of ANI distance. The microbiotas of captive chimps were a mixture of humans and wild chimps' microbiota. The microbiome of wild chimp and captive chimps showed to have greater evolutionary distance compared to the microbiome of human hosts.

These decreases of diversity and evolutionary distances in Germans could be explained by the western lifestyle and use of medications. The intermediate of the microbiota of captive chimpanzees between the microbiota of human and wild chimpanzees could be explained by the *Prevotella* species being able to transfer horizontally and a pure coevolution/cospeciation of great apes and the genus *Prevotella* might be untraceable, though it is notable that the different *Prevotella* clades are more abundant in one or a few host groups instead of being evenly abundant in all host groups.

Acknowledgment

Throughout the process of working with this project I have received a great deal of help and support. For that I will like to thank Ph.D. Louise Bruun Thingholm, who helped formulating the research question, aided great support when I hit dead ends and provided a great learning environment for personal development within the field of microbiome study with her inspiring and professional investment and engagement in the field. Malte Rühlemann has also been of substantial support regarding questions within the research field but foremost by providing the dataset that are the foundation for the thesis. Professor Mikkel Heide Schierup has been listening and asking questions that made me view the project in a broader perspective and was a great help when I was unsure in my achieving.

Regarding the study environment, a great thank to Marie Vibeke Vestergaard and Anne Mathiesen are also in its place. Besides the interesting study discussions, they have made the process enjoyable and made the long hours tolerable and brakes refreshing.

Content

Abstract	2
Acknowledgment.....	2
1. Introduction.....	4
2. Aim.....	5
3. Material	5
4. Method	6
Sample types and sequencing method	6
Assembly and binning.....	9
Annotation.....	10
Quality filtering.....	11
Dereplication	12
Distances and clustering.....	14
Mapping and Abundances.....	16
Alpha and beta diversity.....	17
5. Results and discussion	19
5.1 MASH cluster	19
5.2 ANI distances	20
5.3 MetaWRAP abundances.....	26
5.4 ANI in combination with abundances	30
5.5 Alpha diversity	31
5.6 Beta diversity.....	33
5.7 Ordination plots.....	34
6. Future studies.....	35
7. Appendix.....	36
8. Reference.....	48