ML4Microbiome workshop

Classifying microbiome and resistome using machine learning

Aldert Zomer
24 May 2022

Utrecht University

Veterinary Medicine

World Health Organization

OIE World Organisation for Animal Health

Who am I?

- Microbiologist and bioinformatician – My computer is my lab.
- PhD University of Groningen, Oscar Kuipers, 2002-2
  - Lactic acid bacteria transcriptomics / genomics

1 genome
2 years
100k € / genome

- University College Cork, Ireland, Douwe van Sinderen, 2006-2010
  - Gut bacteria transcriptomics / genomics
  - Metagenomics
- RadboudUMC, Peter Hermans/Marien de Jonge, 2010-2014
  - Pneumococcus population genomics and gene essentiality (Tn-Seq)
  - Nasopharyngeal microbiome
  - Bacterial and host transcriptomics during infection
- Utrecht University, Jaap Wagenaar, 2014 –
  - Bacterial population (meta) genomics: *Staphylococcus*, *E. co*
  - Effect of antimicrobials on resistome and microbiome

1000s of genomes
6 weeks
<40€ / genome

Bacterial population genomics

- *Campylobacter*
  - Campylobacter fetus Birgitta Duim and Linda van der Graaf
  - Discerning Environmental Pathways of Campylobacter Transmission (DEPiCT, DiSCOVER) with Lapo Mughini-Gras (RIVM)
  - Campylobacter plasmids Linda, Alison
- *Salmonella / E. coli*
  - Plasmid/strain transmission (STARCS, Full-Force)
  - Carbapenamase, colistin (JPI-TRIUMPH, BEWARE, COINCIDE)
  - ExPEC population genetics for vaccine development (Janssen)
  - APEC population genetics with Sjaak de Wit and Royal GD (COLIBRI)
  - Effect of residual antibiotics concentrations on selection of resistance (Resrisk)
- *Staphylococcus*
  - *S. schleiferi/coagulans* and *pseudintermedius*
  - *S. aureus* vaccine development (Janssen)
- Other
  - *B. pertussis* vaccination policy drives population genetics (Janssen)
  - First generation cephalosporins with David Speksnijder (MSD)


Linda van der Graaf
-van Bloois


Aram Swinkels


Jeroen Leus


Soe Yu Naing

janssen | PHARMACEUTICAL COMPANIES OF Johnson & Johnson

Metagenomics

- How does antimicrobial treatment and fecal transfer affect the resistome of horses – Equine sciences UU
- Rehabilitation of seals and their microbiome and resistome – UMCG
- Outcompeting MRSA from piglet noses (ExcludeMRSA/MRSAprevent) with Birgitta Duim
- Nanopore sequencing for rapid microbial diagnostics – VMDC, Els Broens
- Effect of diet change on the gut microbial community of a migratory shorebird – NIOZ KNAW Jan van Gils
- Quantifying the complex drivers of antibiotic resistance transfer by reusing (meta-) genomic databases Bas Dutilh -TBB

Roosmarijn Luiken

Mathijs Theelen

Ana Rubio Garcia

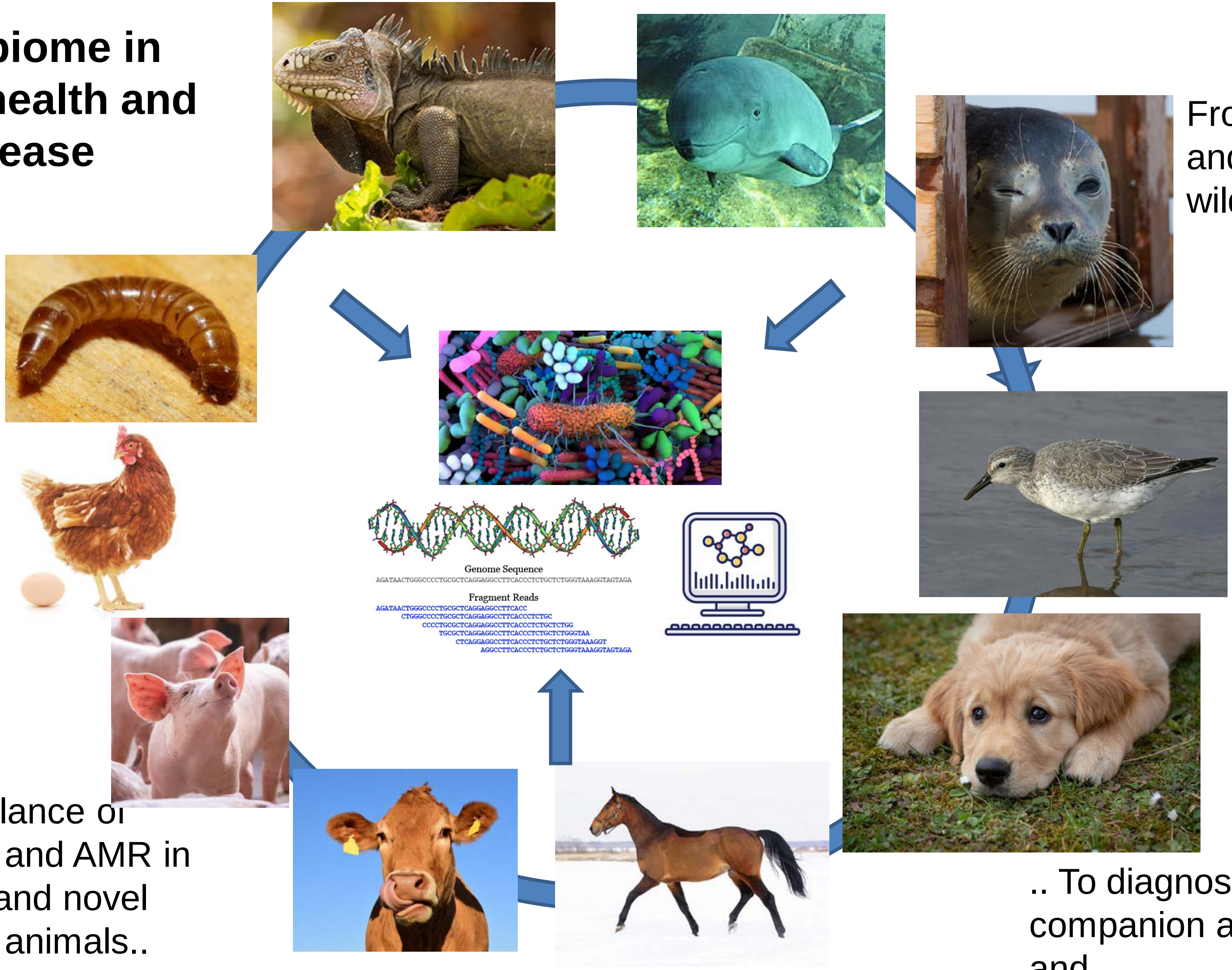
Abel Vlasblom

Marloes van Dijk

Marie de Wilde

Rick Beeloo

Microbiome in animal health and disease



From preservation and health status of wild animals

.. to surveillance of pathogens and AMR in traditional and novel production animals..

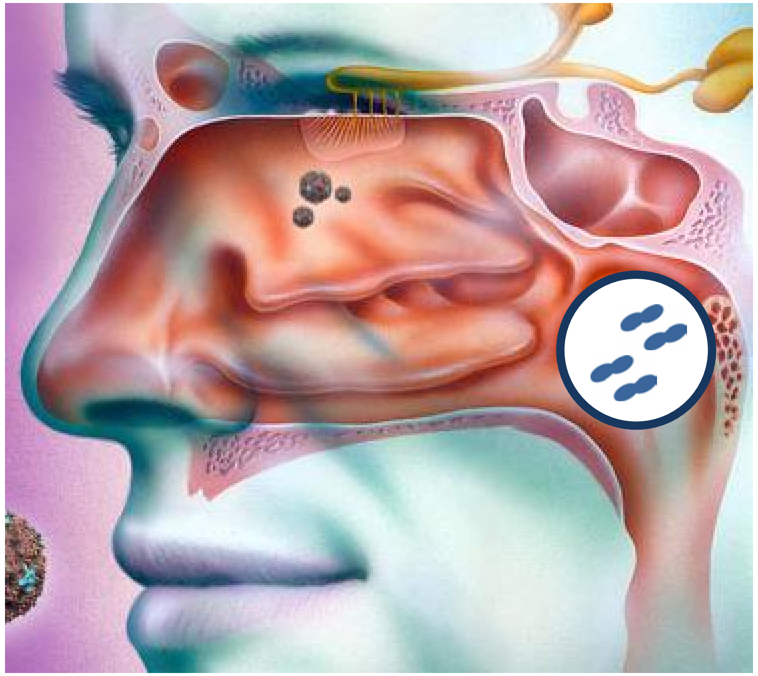
.. To diagnostics of companion animals and

Example 1:
Experimental Human Pneumococcal Carriage

Radboudumc

LSTM

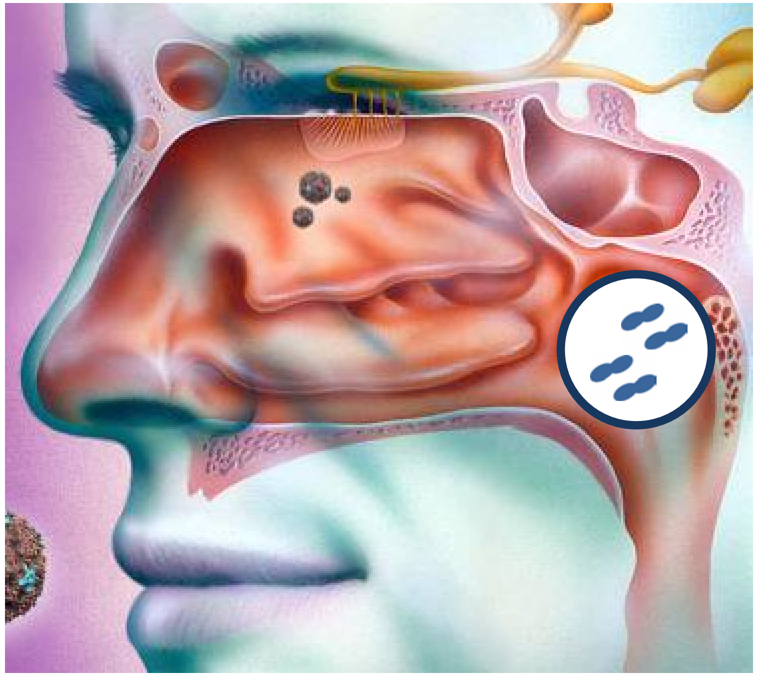
Pneumococcal disease



Pneumonia

Meningitis

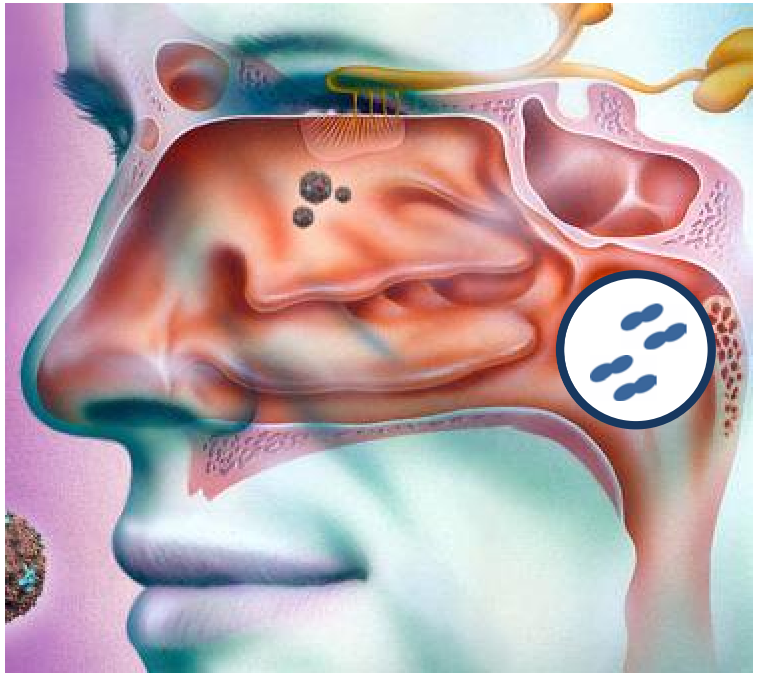
Determinants of pneumococcal carriage

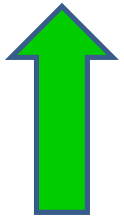


↑ Crowding
Smoke

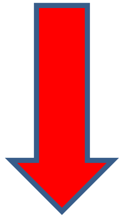
↓ Antibiotics
Immunity

Determinants of pneumococcal carriage

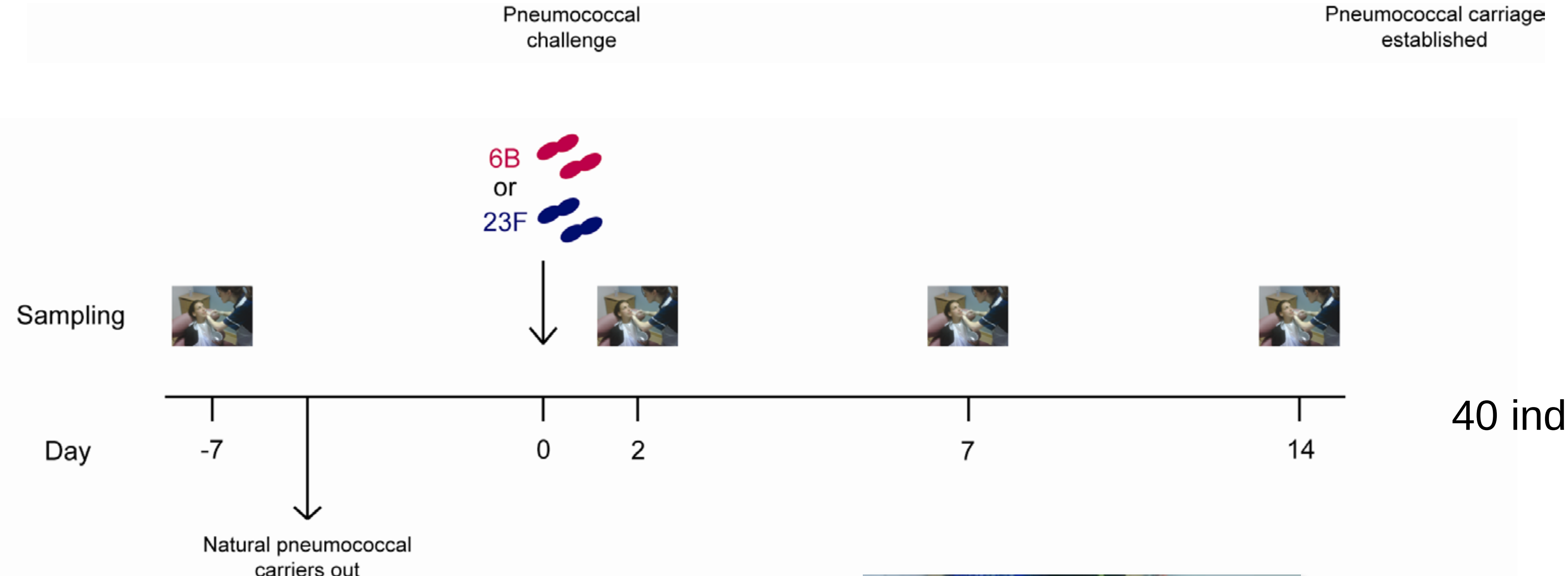




Crowding
Smoke
Other microbes in nasopharynx

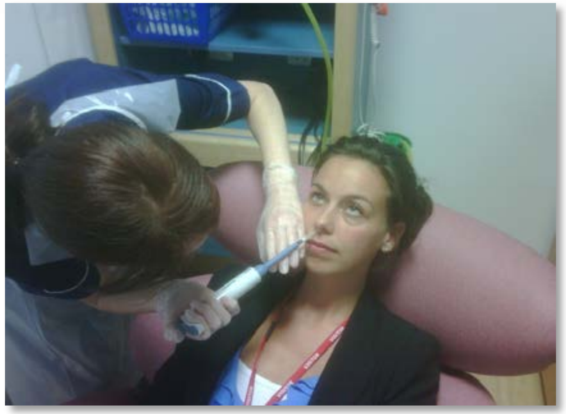


Antibiotics
Immunity
Other microbes in nasopharynx



40 individuals

Nasal wash sampling



Intranasal pneumococcal inoculation

LSTM

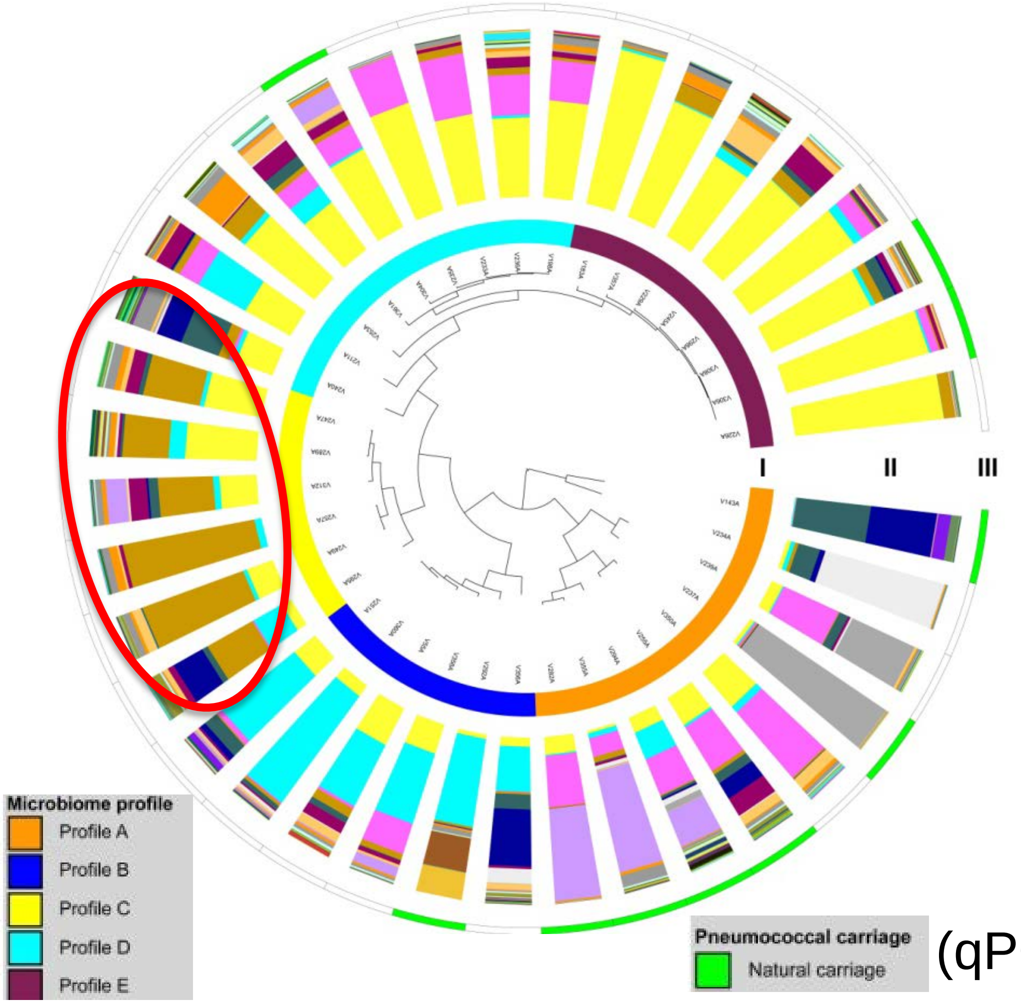
“Nasotypes” ?

At least 5 types observed

Staphyloccus dominated profile is negatively correlated with pneumococcus carriage

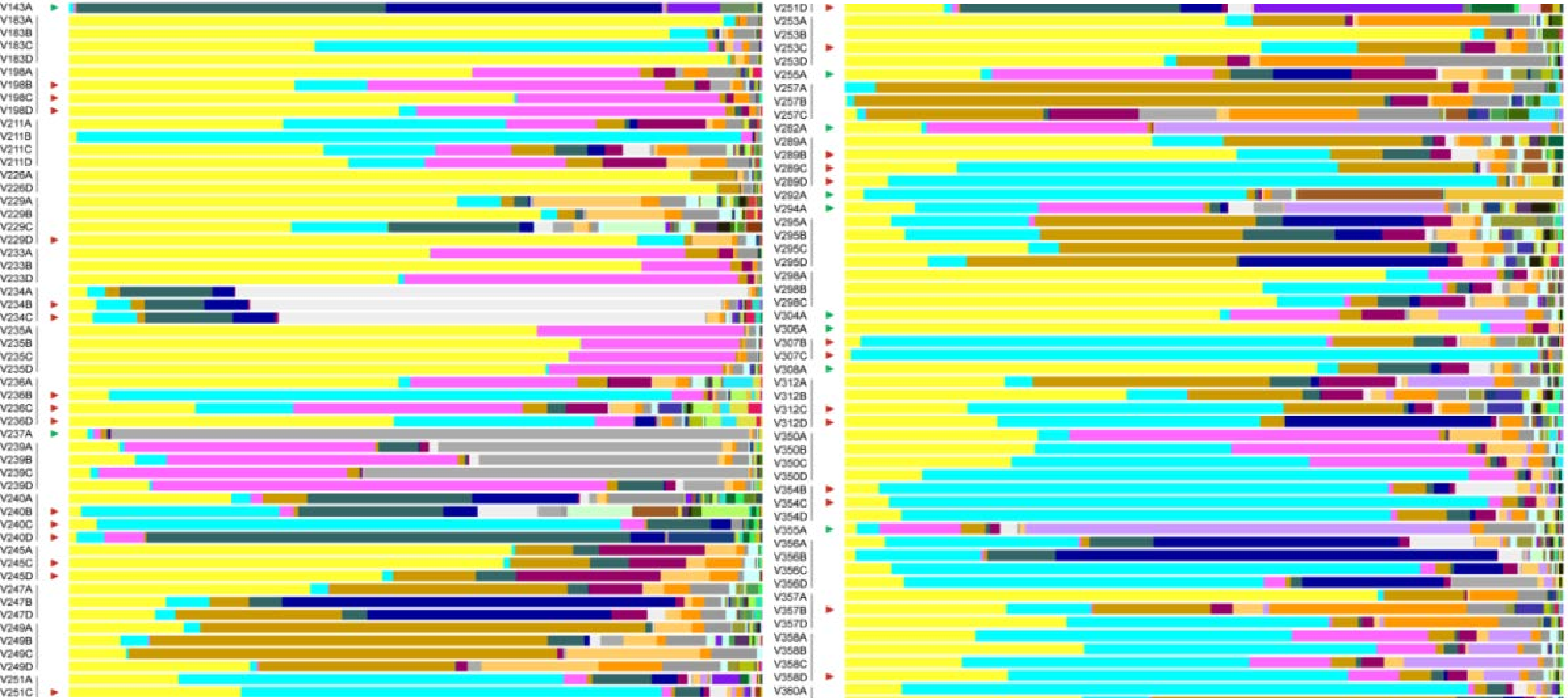
(also observed in larger qPCR study)

Pneumococcal carriers excluded and remaining subjects colonized



(qPCR)

Microbiome longitudinally sampled

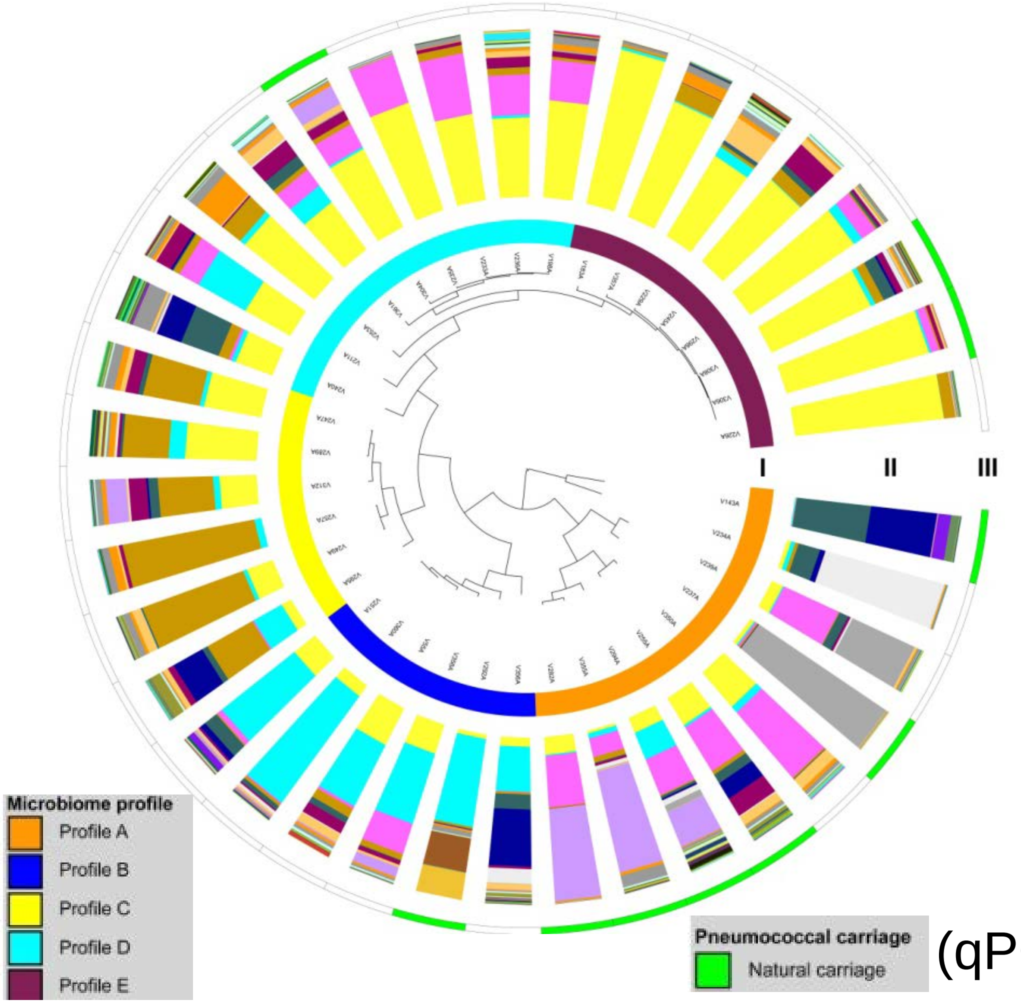


“Nasotypes” ?

Classes: ‘T0’ samples, nasotypes

Classifiers: the microbial density data – 200 of the most common ASVs

randomForest in R



(qPCR)

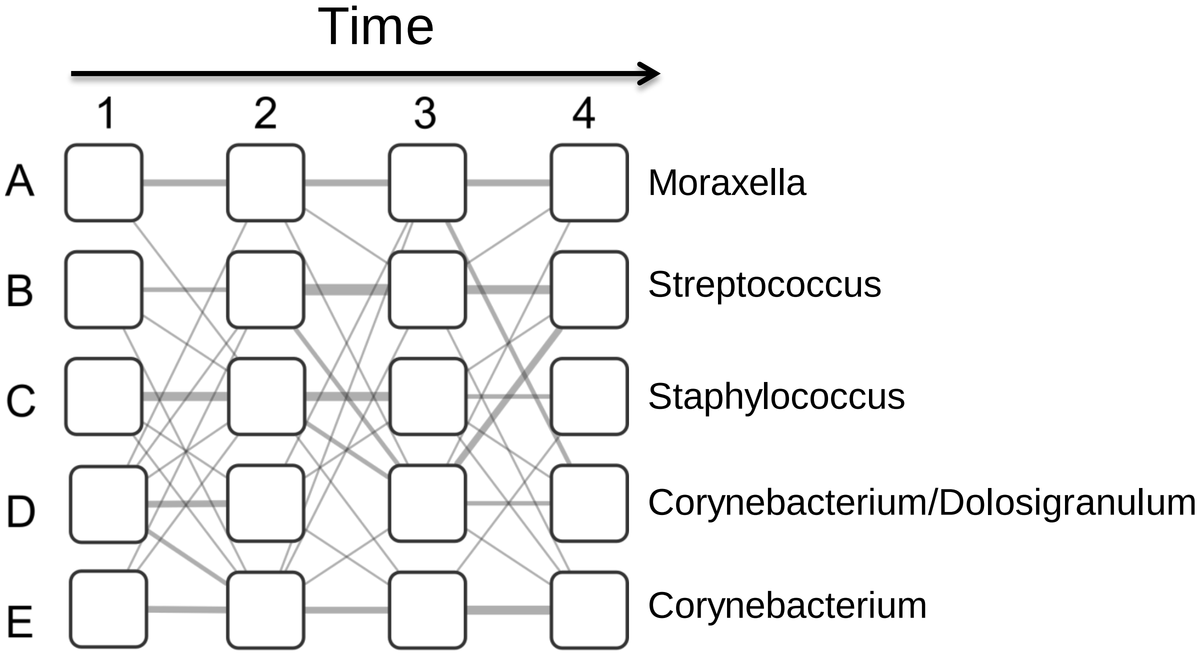
Does colonization with *S. pneumoniae* cause a nasotype shift?

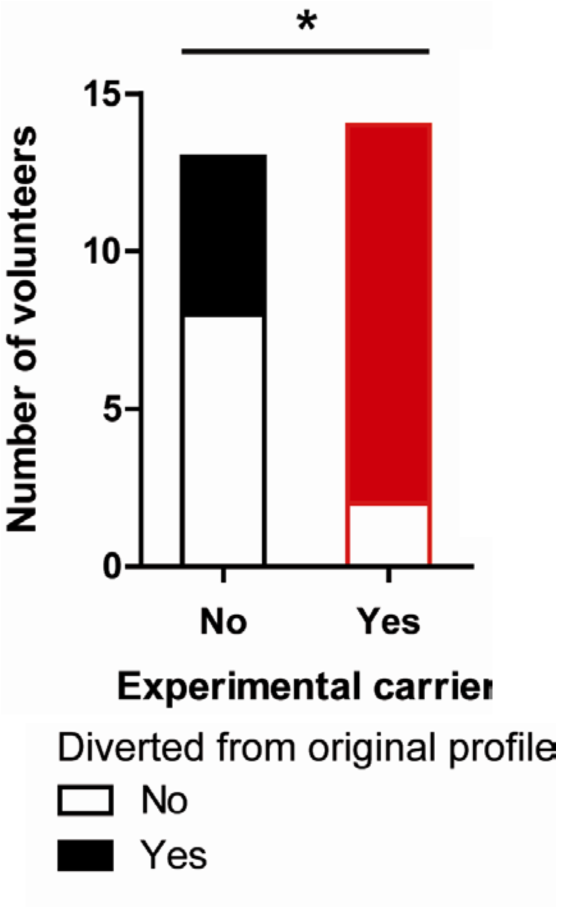
Classes: ‘T0’ samples, nasotypes

Classifiers: the microbial density data – 200 of the most common ASVs

randomForest in R

Pneumococcal colonization causes shifts away from original nasotype





- ➢ Not specifically to a streptococcus dominated profile
- ➢ Pneumococcal carriers never switched to Profile C

Part 1 - Questions?

Read more -> Cremers et al. 2016, Microbiome

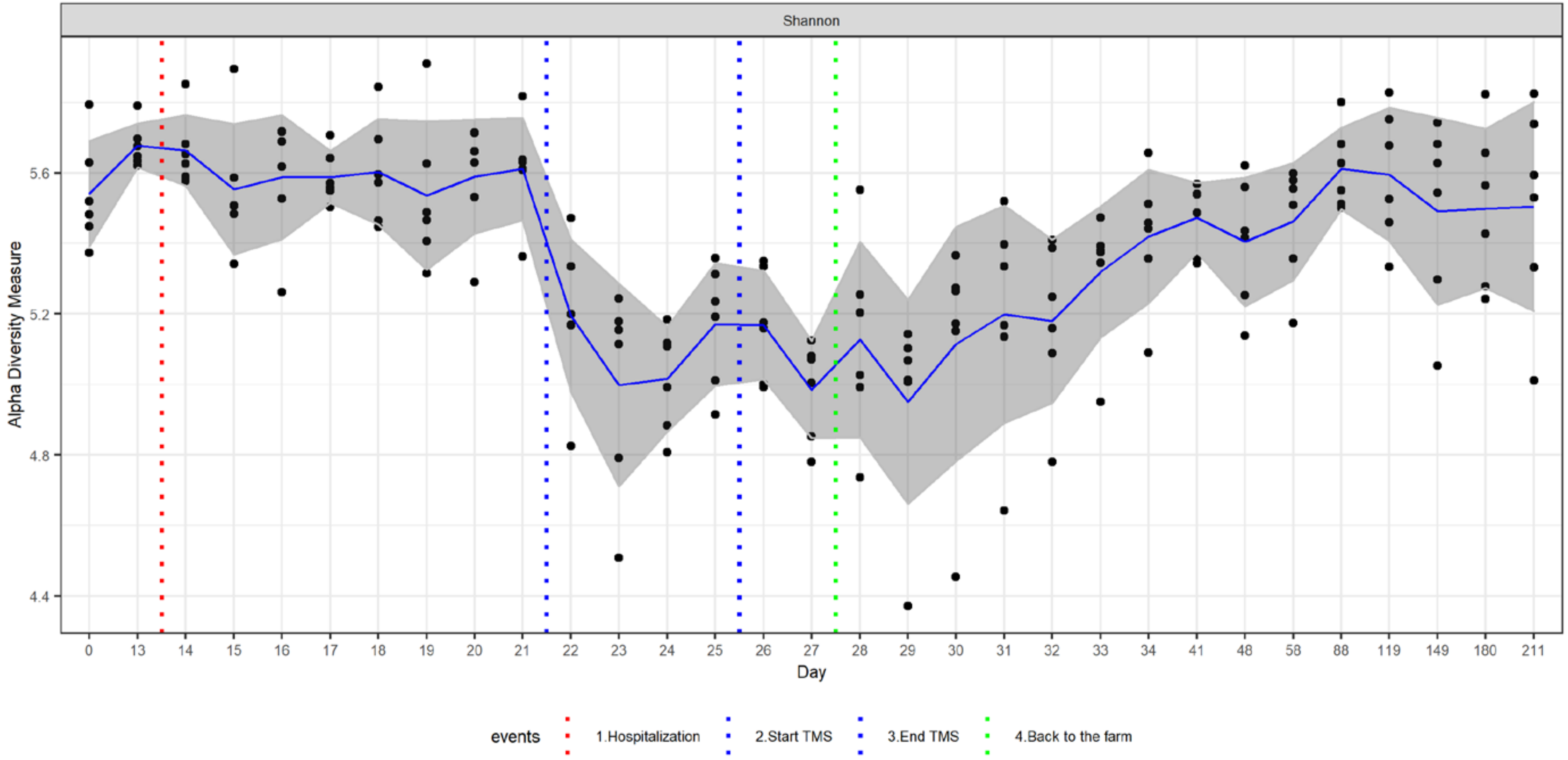
Short- and long-term effect of hospitalization and oral antimicrobial treatment on the equine faecal microbiome and resistome. By Mathijs Theelen /Roosmarijn Luiken

- Longitudinal sampling of 6 healthy ponies
- 16S amplicon sequencing
- Shotgun metagenomics selected samples

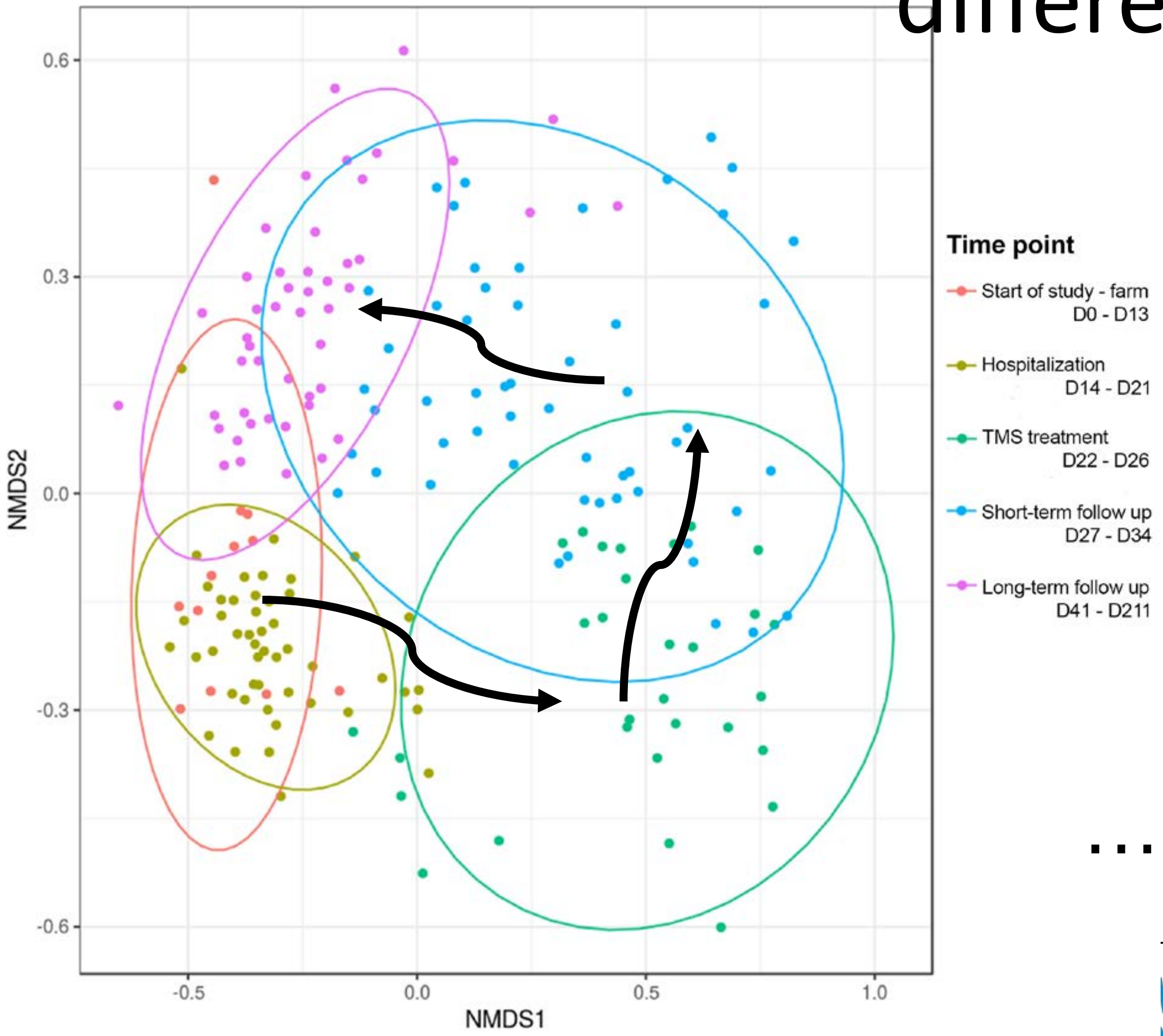
at the farm - in the clinic - TMS treatment - back to the farm

D0-13 D14-21 D22-26 D27-211

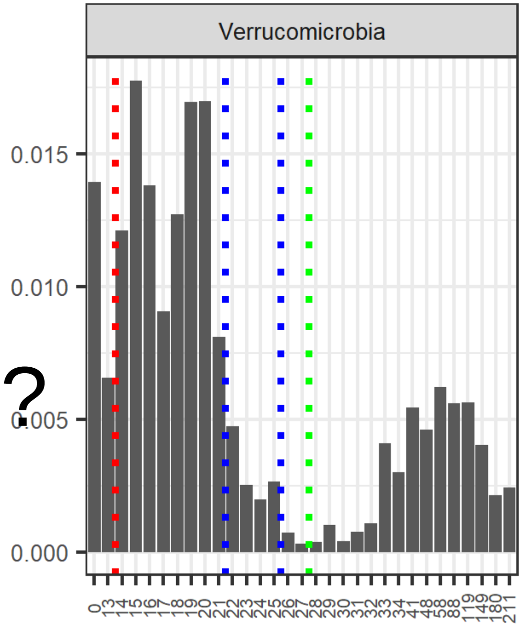
Shannon diversity (= alpha div.) decreases during treatment and seems to recover after several weeks



The microbiome composition (= beta div.) changes after treatment and is after 6 months still significantly different



…but how different?



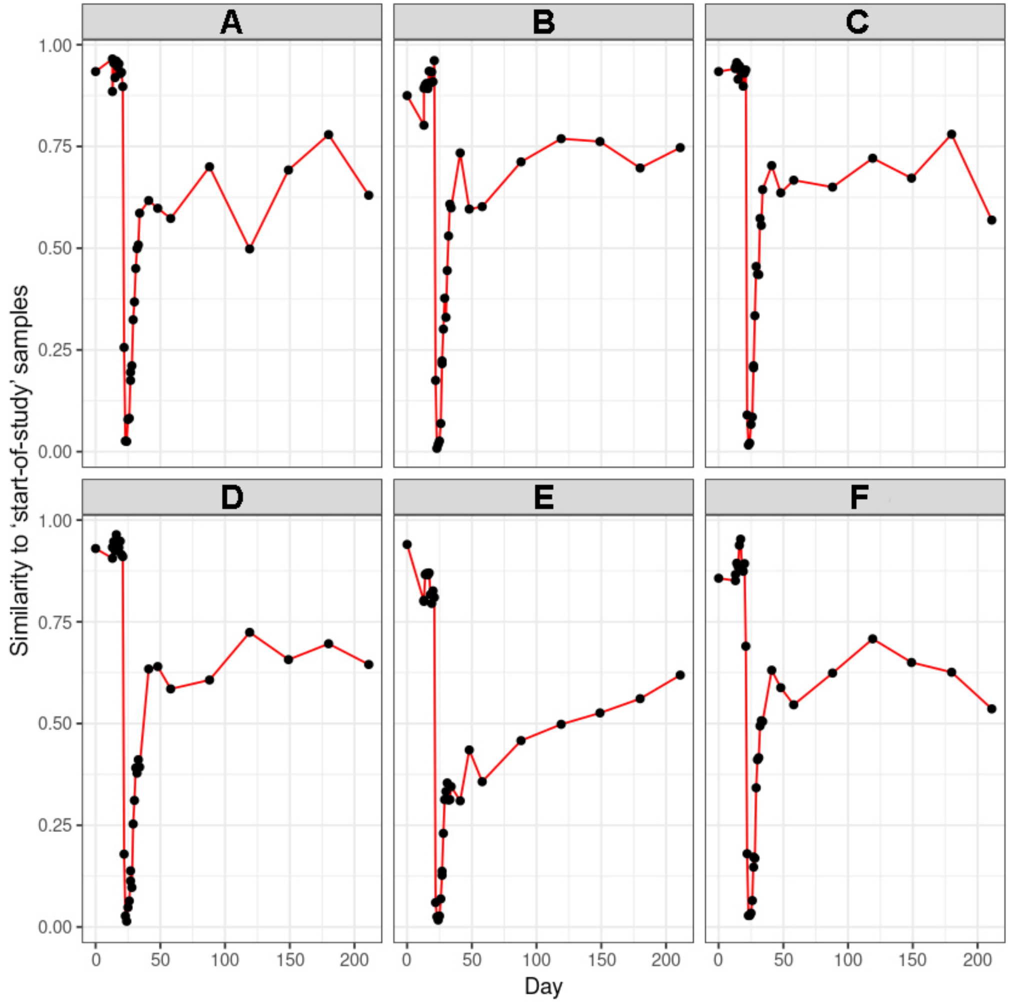
Random Forest prediction: after 6 months a different microbiome than at the start of the study

Classes: ‘start of study’ and ‘TMS treatment’ samples

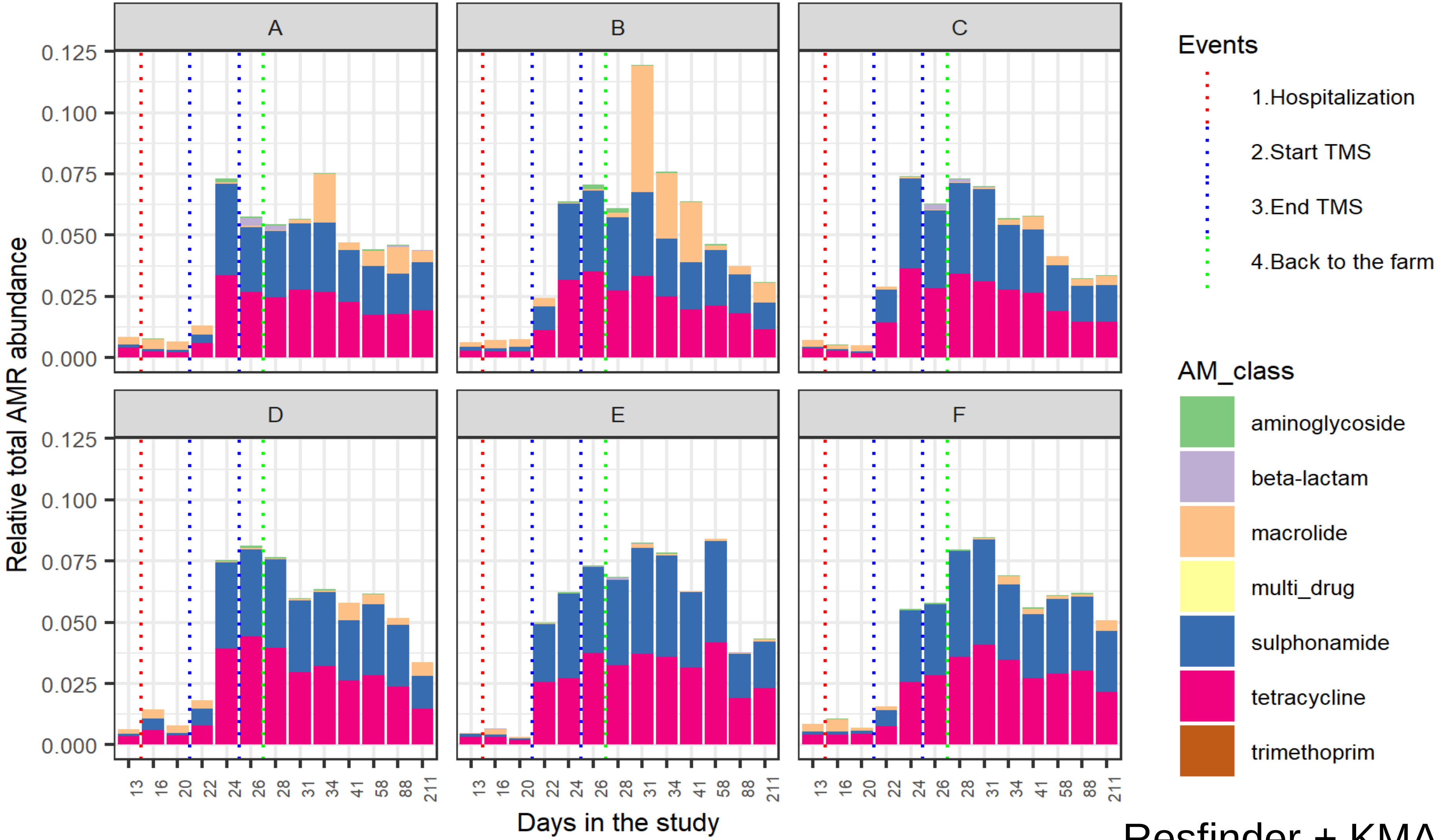
Classifiers: the microbial density data – 200 of the most common ASVs

Aim of RF: predict if microbiota composition after 6 months was more similar to microbiota composition at the start of the study or during TMS treatment

More similar to pre-treatment than treatment after approx. 3 weeks (except E), however only 75% of votes reached at most in all horses



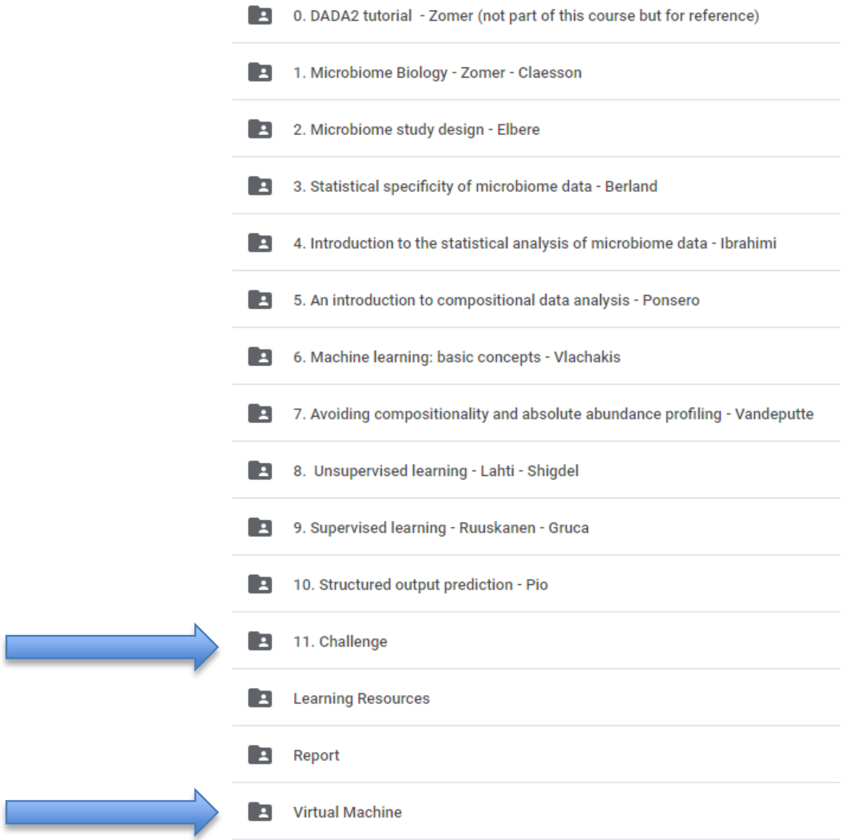
The resistome changes and increases after treatment starts and is still significantly different after 6 months



Resfinder + KMA

Code to try (at your own leisure)

- ML4Microbiome online course material
- Folder “11. Challenge”
- Either run Jupyter notebook on own machine, on Google Colab (free) or on the Utrecht University server (ask for access)
- VM image available on request (100gb) https://drive.google.com/drive/folders/1RKWnKEDRF7ZkVlbltO8FADF_yhBuXglB



0. DADA2 tutorial - Zomer (not part of this course but for reference)
1. Microbiome Biology - Zomer - Claesson
2. Microbiome study design - Elbere
3. Statistical specificity of microbiome data - Berland
4. Introduction to the statistical analysis of microbiome data - Ibrahimi
5. An introduction to compositional data analysis - Ponsero
6. Machine learning: basic concepts - Vlachakis
7. Avoiding compositionality and absolute abundance profiling - Vandeputte
8. Unsupervised learning - Lahti - Shigdel
9. Supervised learning - Ruuskanen - Gruca
10. Structured output prediction - Pio
11. Challenge

Learning Resources

Report

Virtual Machine

Questions – part 2

- Unpublished, but submitted soon

Thank you!

- Roosmarijn Luiken
- Mathijs Theelen
- Jaap Wagenaar
- Amelieke Cremers
- Daniela Ferreira

Utrecht University

Radboudumc

LSTM