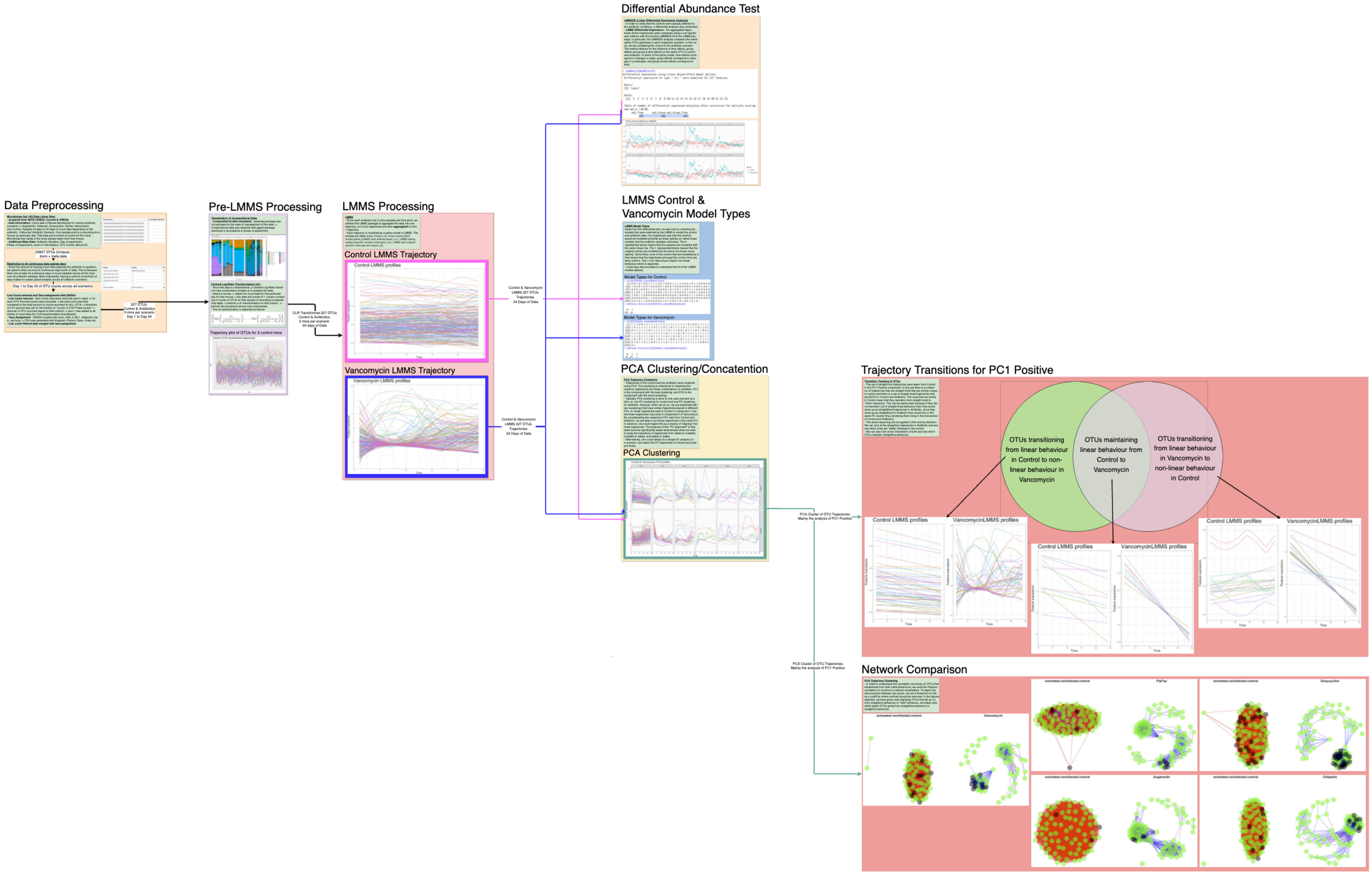
From 16S count data to dynamical network systems modelling.

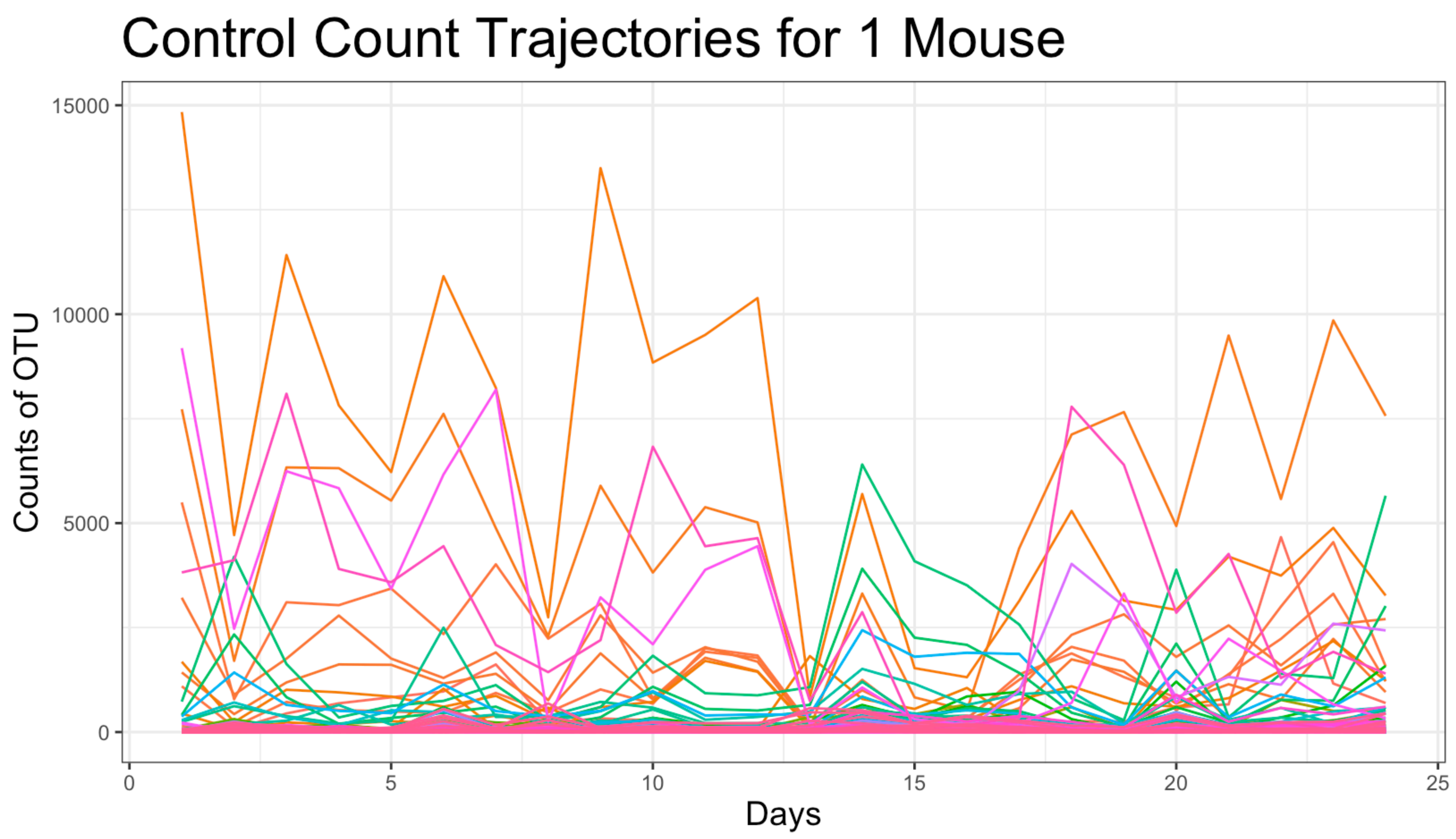
Paul Nguyen, Le Cao Lab.

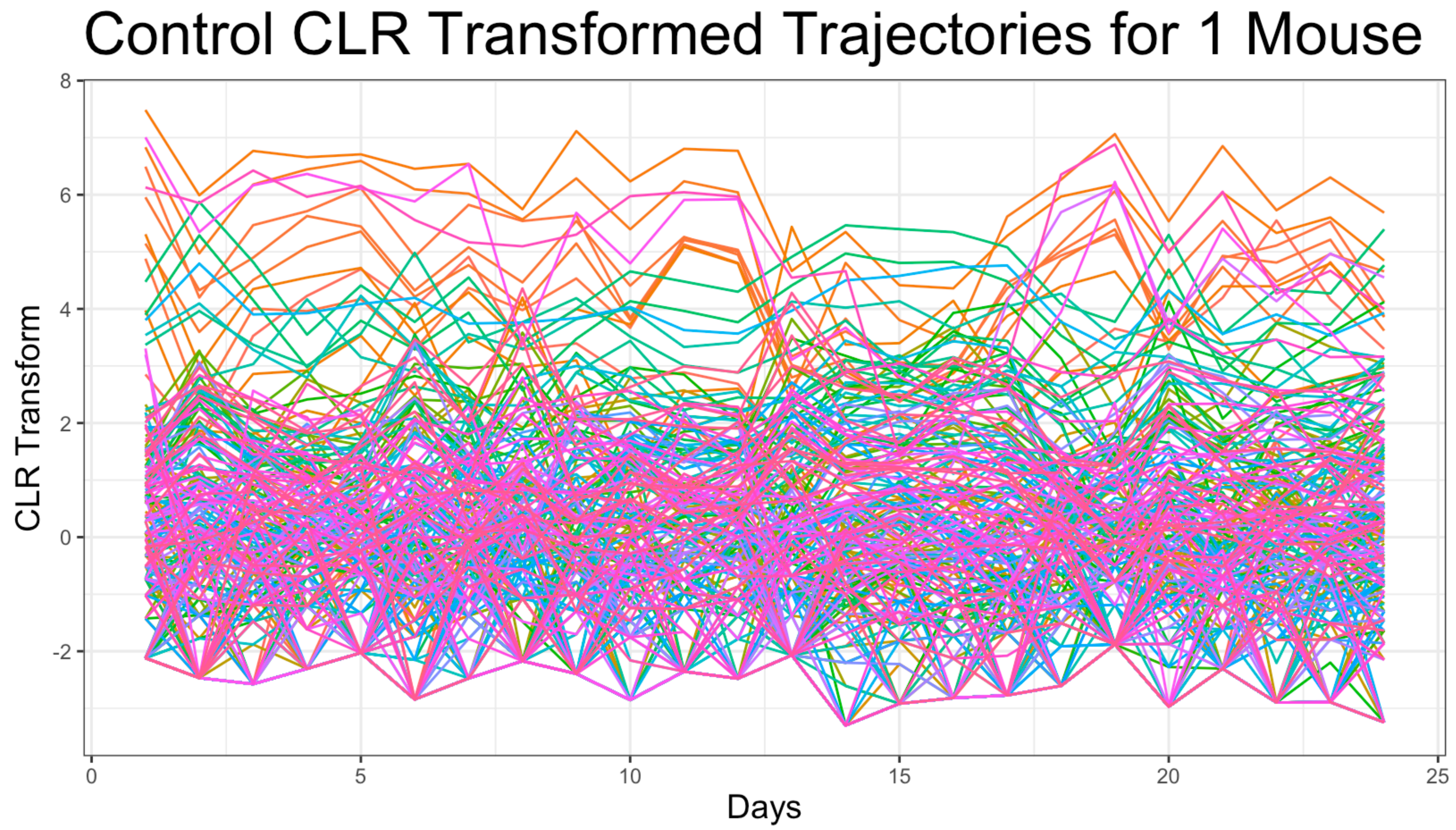


Data Preprocessing
Pre-LMMS Processing
LMMS Processing
Control LMMS Trajectory
Vancomycin LMMS Trajectory
Differential Abundance Test
LMMS Control &
Vancomycin Model Types
PCA Clustering/Concatention
PCA Clustering
Trajectory Transitions for PC1 Positive
OTUs transitioning from linear behaviour in Control to non-linear behaviour in Vancomycin
OTUs maintaining linear behaviour from Control to Vancomycin
OTUs transitioning from linear behaviour in Vancomycin to non-linear behaviour in Control
Control LMMS profiles
VancomycinLMMS profiles
Network Comparison

1. Data Preprocessing & Linear Mixed Model Spline Preprocessing

- **Data Information:**
- 16S Microbiota count data from antibiotic experiments. (Control, Vancomycin, Cefazolin, etc)
- 3 Experimental mice/3 Controls mice. 24 to 50 days of continuous data.
- **Restricted** analysis to 24 days of data across all experiments.
- **Low Count removal** results in 227 OTUs (from 24927 OTUs). 0.01 percent threshold
- **Central Log-Ratio Transformation (CLR)** for the compositional data.
- **Taxa Assignment** with DADA2.

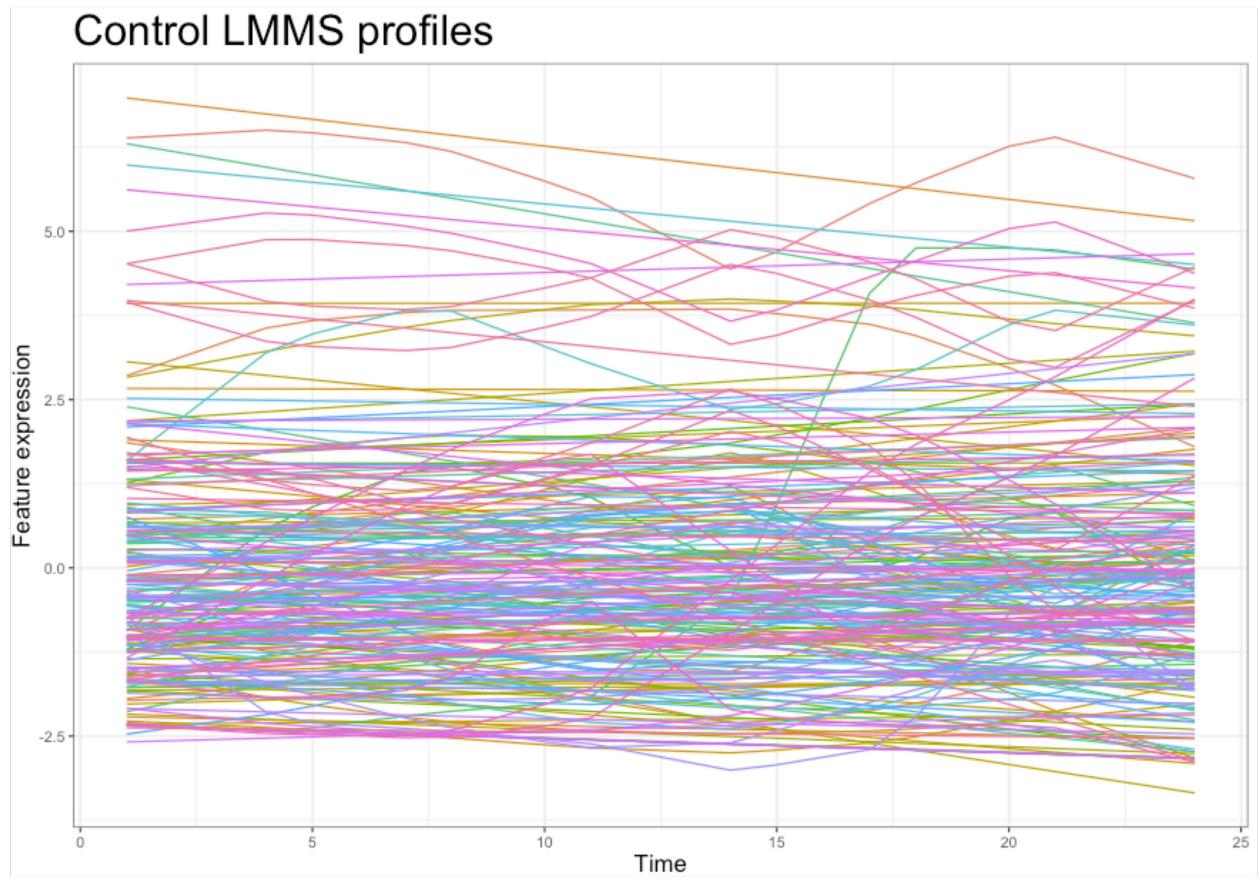


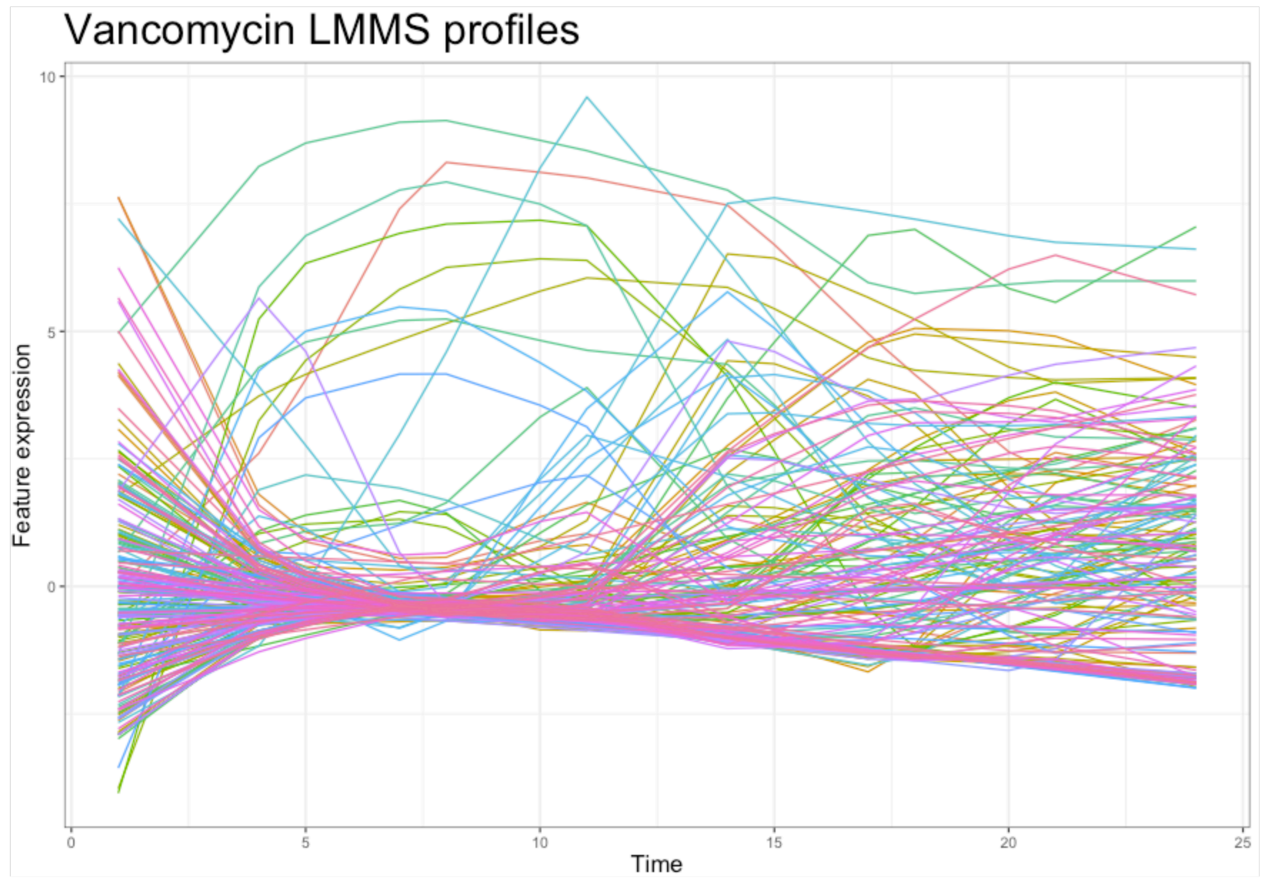


Output: 227 shared CLR Transformed OTUs from Day 1-24 for antibiotic and control mice groups (3 mice per group).

2.1 Linear Mixed Model Spline Processing

- **Aggregate** 3 trajectories into 1 with Linear Mixed Model Splines (LMMS R package).
- **Four types of splines are fitted:**
- Linear model (=0)
- Linear mixed effect spline model with p degree polynomial. (=1)
- Linear model with subject-specific intercepts (=2)
- Linear model with subject-specific slope and intercepts (=3)





Model Counts Control

Model Type	0	1	2
Counts of Model	163	63	0

Model Counts Vancomycin

Model Type	0	1	2
Counts of Model	49	176	1

Output: 1 OTU Spline summarising the time trajectory across Control or Antibiotic respectively. 1 OTU Spline x 227 OTUs = 227 OTU Splines.

Key message: Most of the trajectories in control are flat lines, indicating that the controls are well behaved.

2.2 Linear Mixed Model Spline Processing - Control & Vancomycin Model Types

- Differences between the antibiotic conditions and controls are determined by “Linear Differential Expression Analysis”.

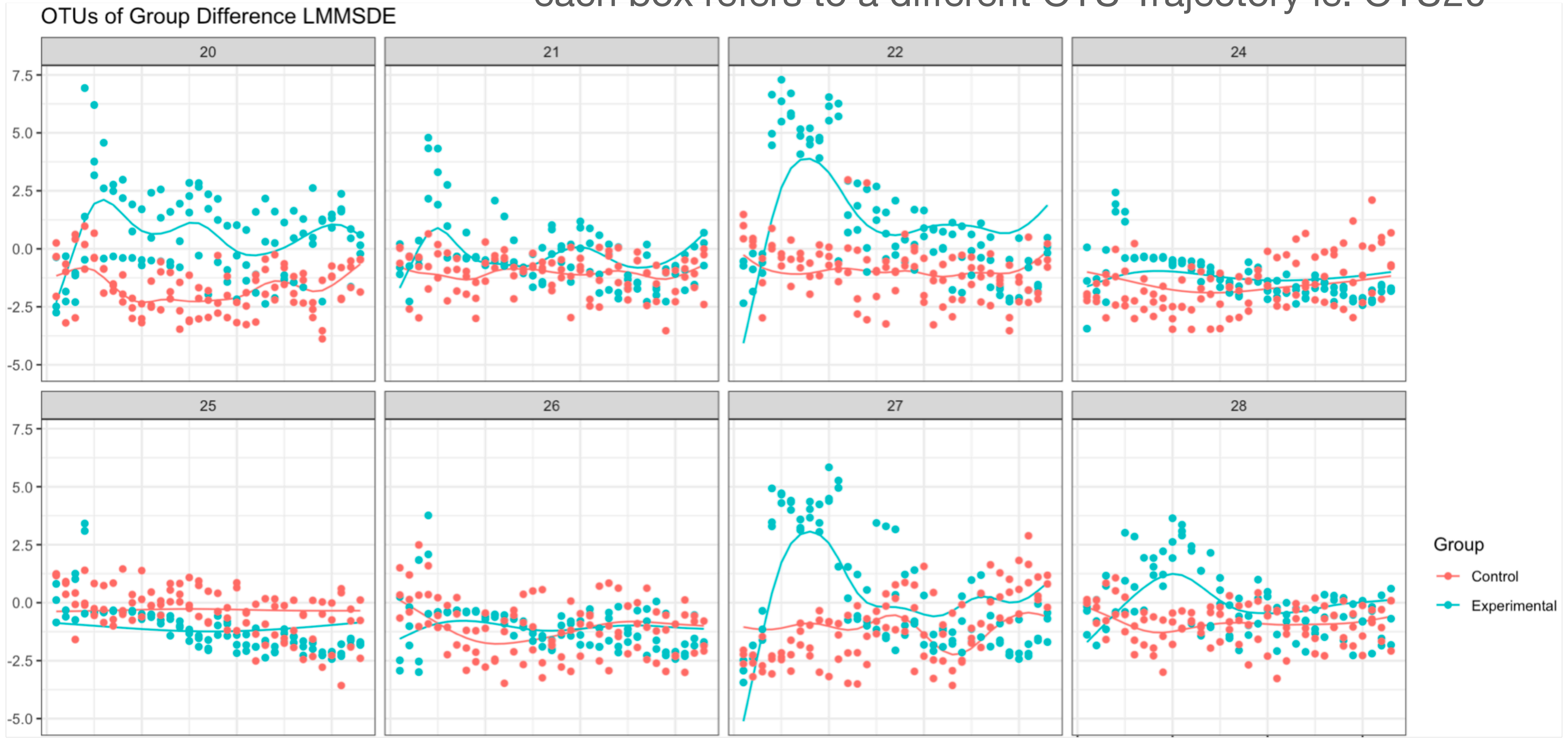
- **Linear Differential Expression Analysis** tests for:

- 1. Time - differences across time.

- 2. Group - differences between antibiotic group and control.

- 3. Interaction between Group & Time - combination of the two.

*each box refers to a different OTU Trajectory ie. OTU20



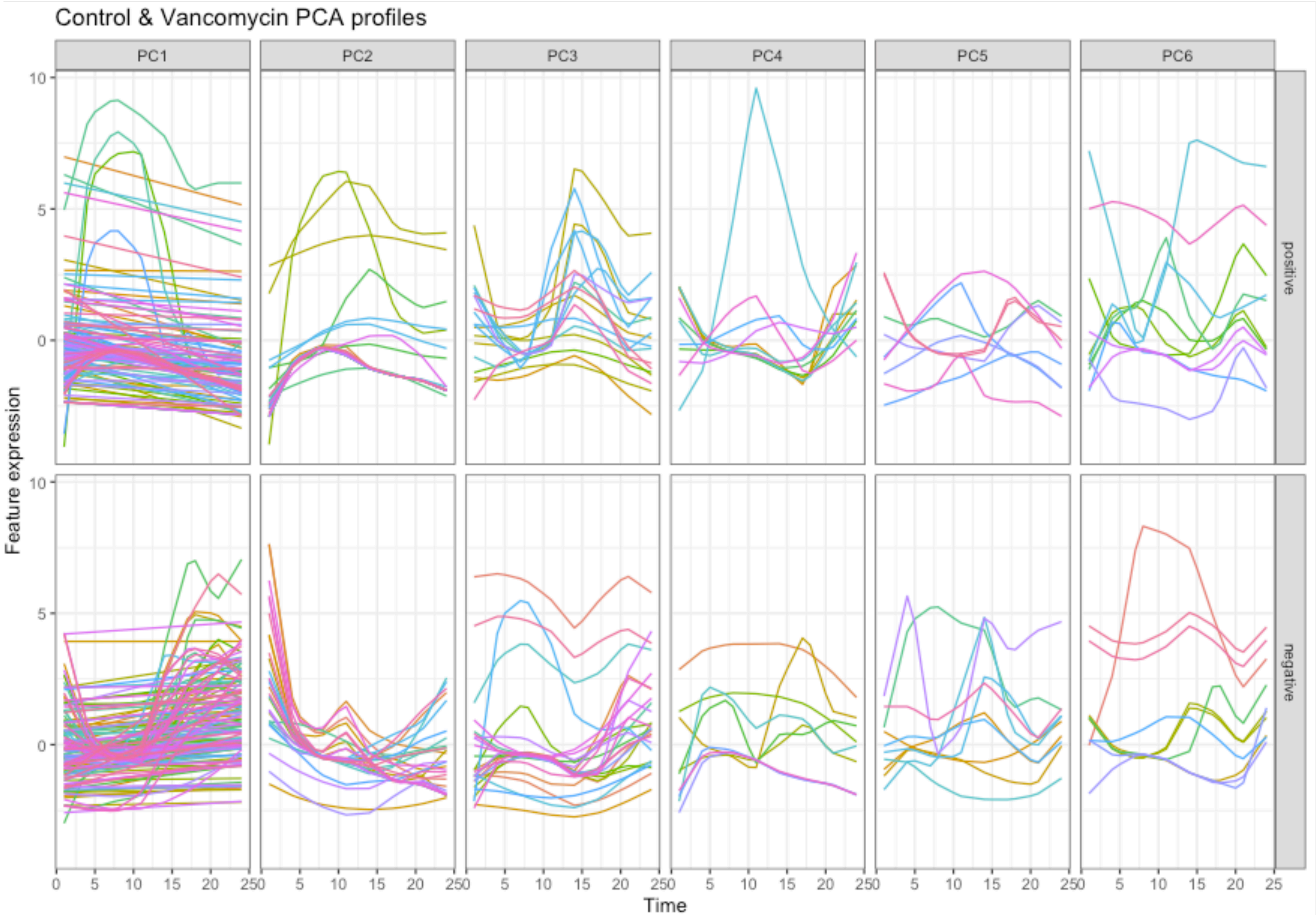
Model Difference Counts

	Time	Group	Group & Time
Number of OTUs significant for	185	206	203

Key message: Most OTUs are differentially expressed compared to control with time and/or group effects.

3. PCA Clustering

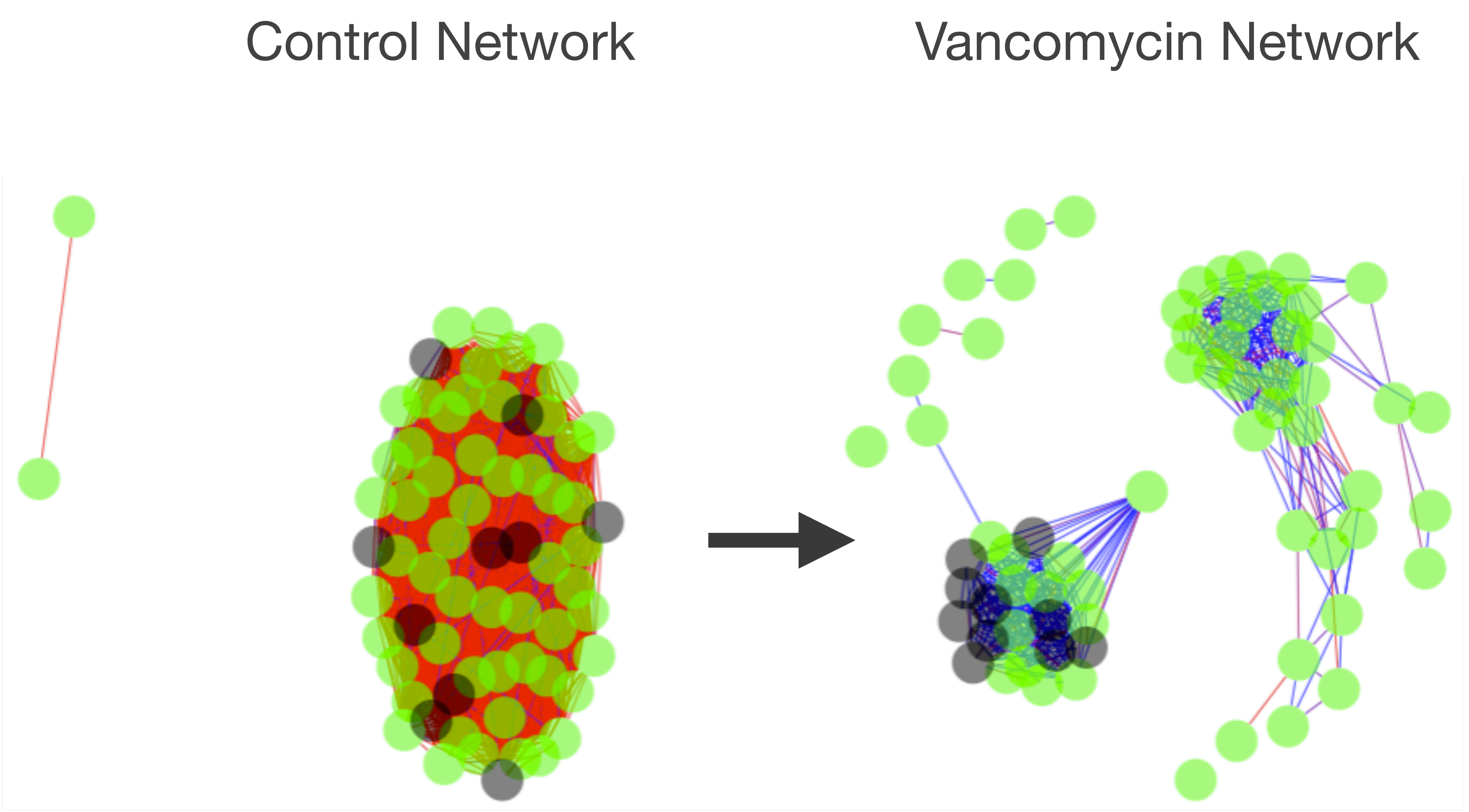
- Due to the high number of trajectories, it is easier to do analysis based on clusters.
- We cluster OTU trajectories from both control and antibiotics together in order to:
- 1. Cluster OTUs with similar trajectories.
- 2. Identify OTUs whose trajectories change between control and antibiotic groups.



Output: Select PC1 Positive Cluster trajectories for further analysis.

4.1 Network Comparison

- A correlation network was constructed using Pearson Correlation between OTU trajectories.
- Red lines - positive correlation.
- Blue lines - negative correlation.
- OTUs that do not change in behaviour during antibiotic exposure. Straight line to Straight line.
- OTUs that change in behaviour during antibiotic exposure. Straight line to linear splines.



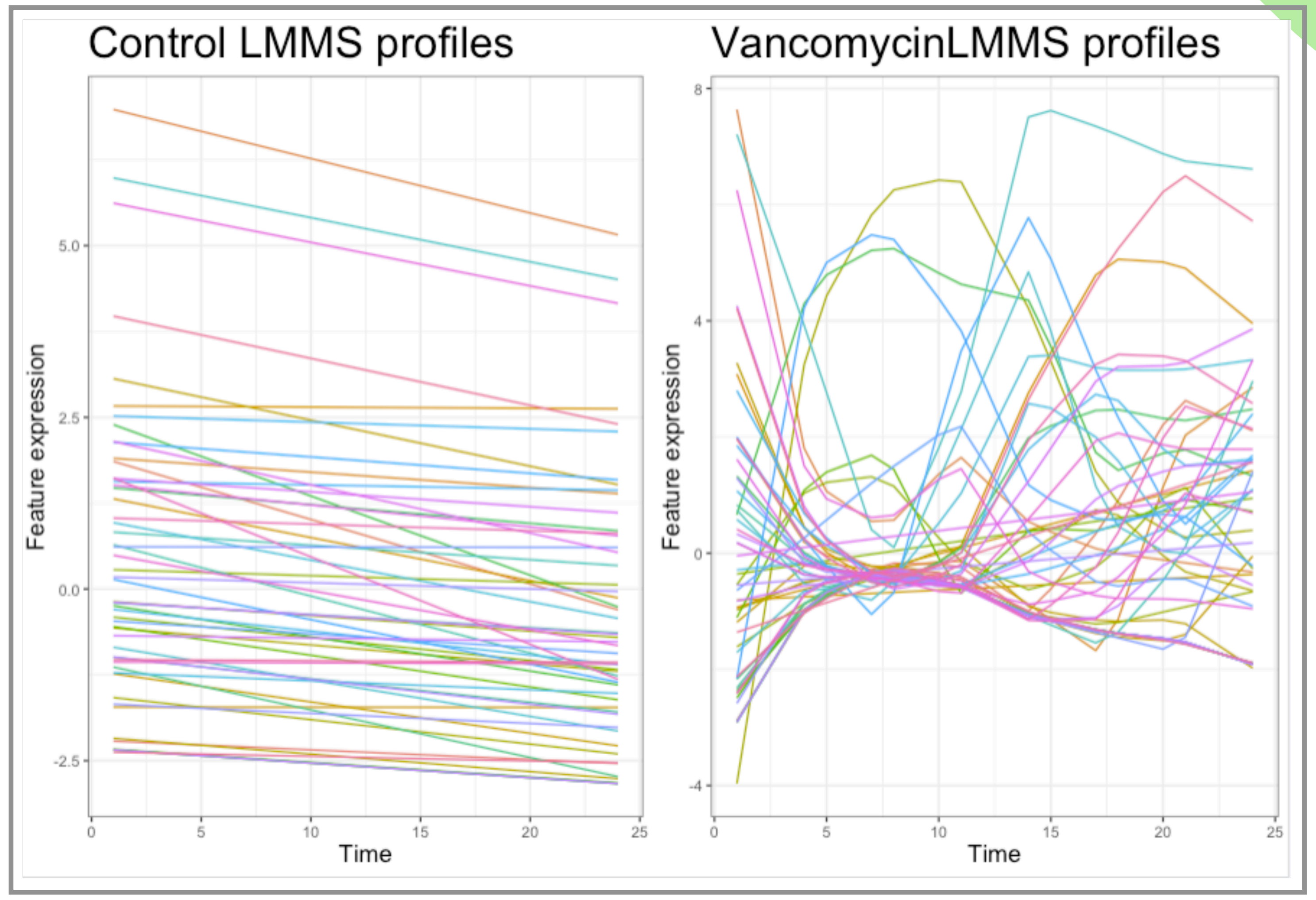
The trajectories here begin as straight lines in the Control group.

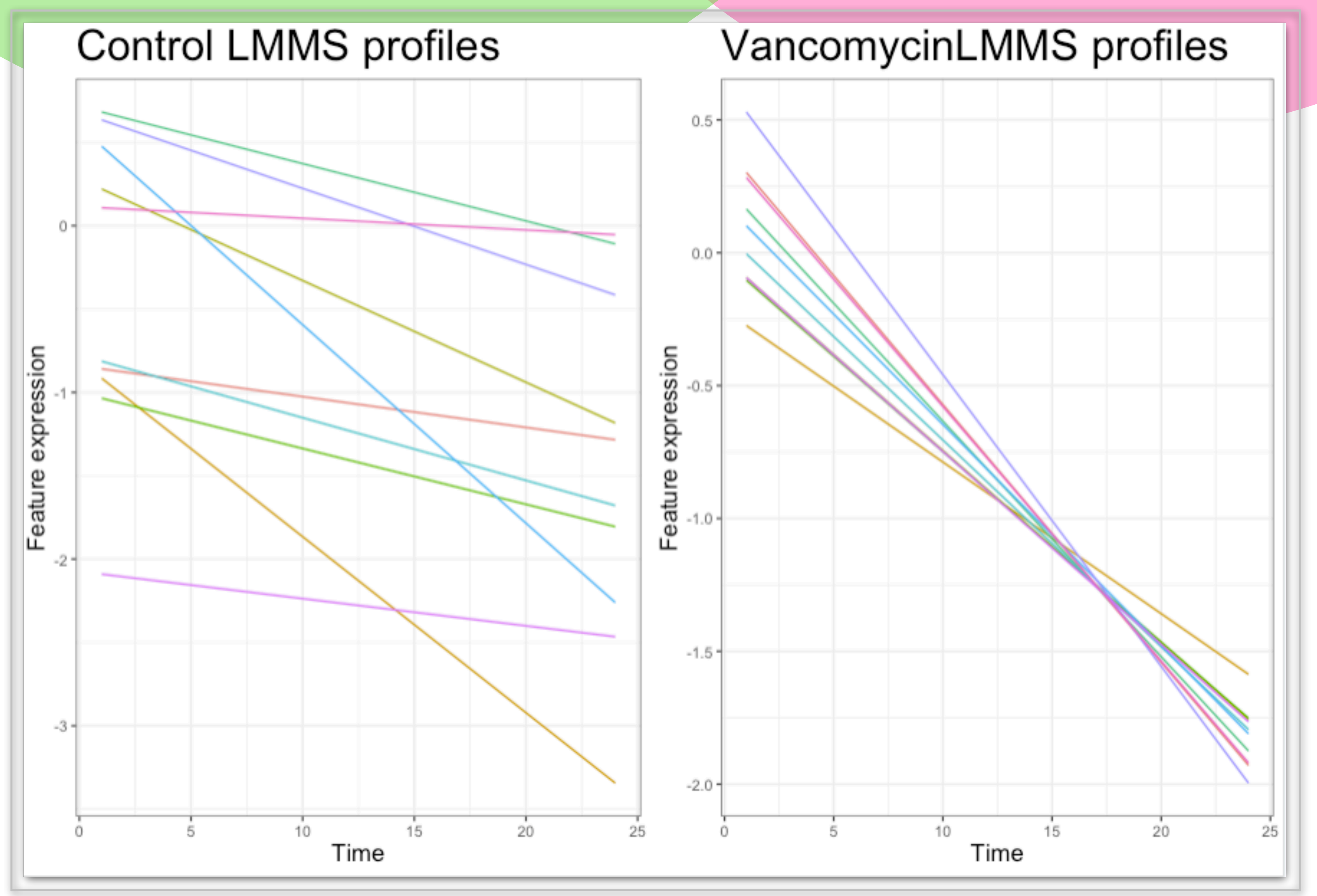
4.2 Trajectory Transitions

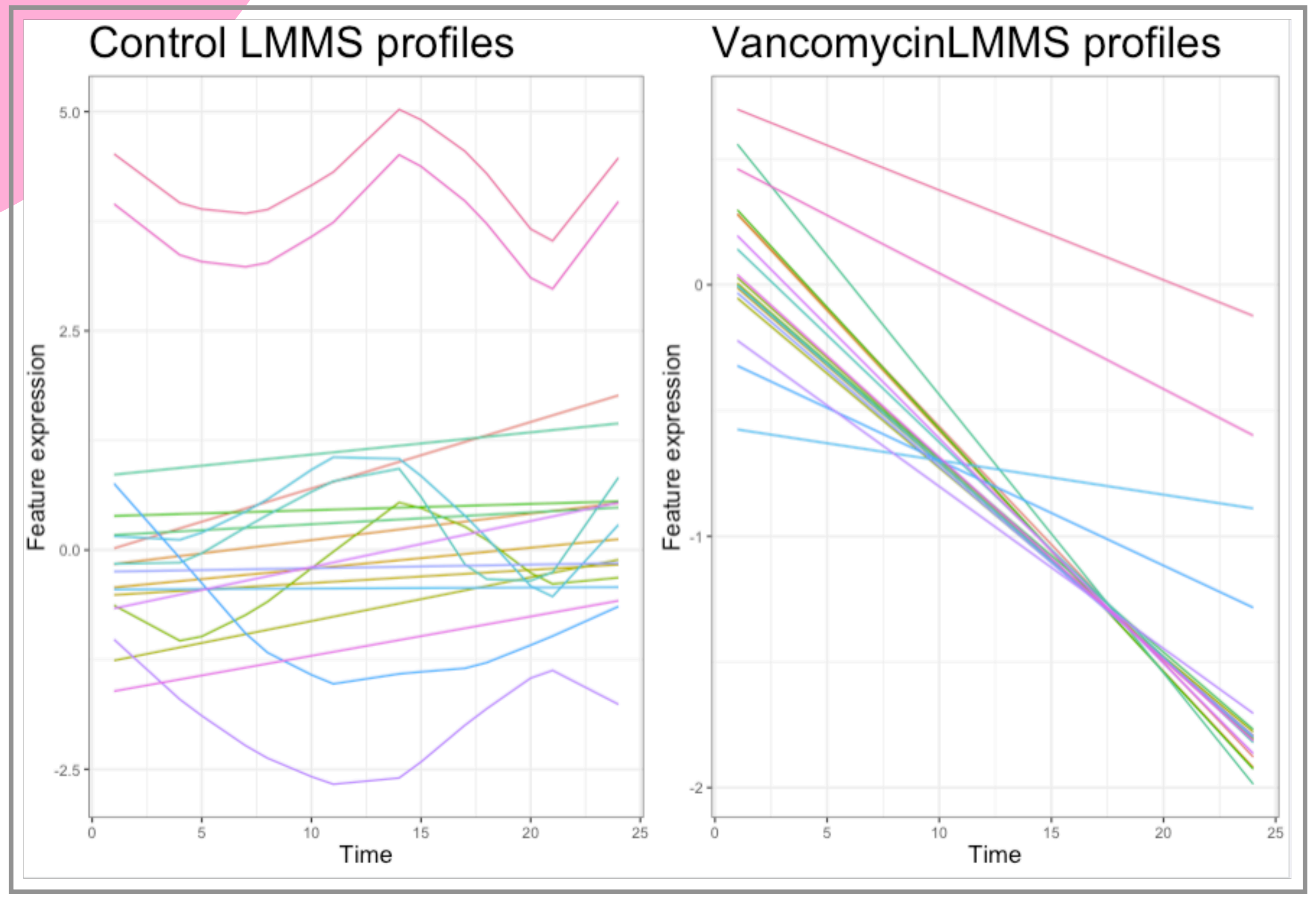
OTUs transitioning from straight lines in Control to linear splines in Vancomycin

OTUs maintaining straight line behaviour from Control to Vancomycin

OTUs transitioning from non-linear splines in Vancomycin to straight lines in Control







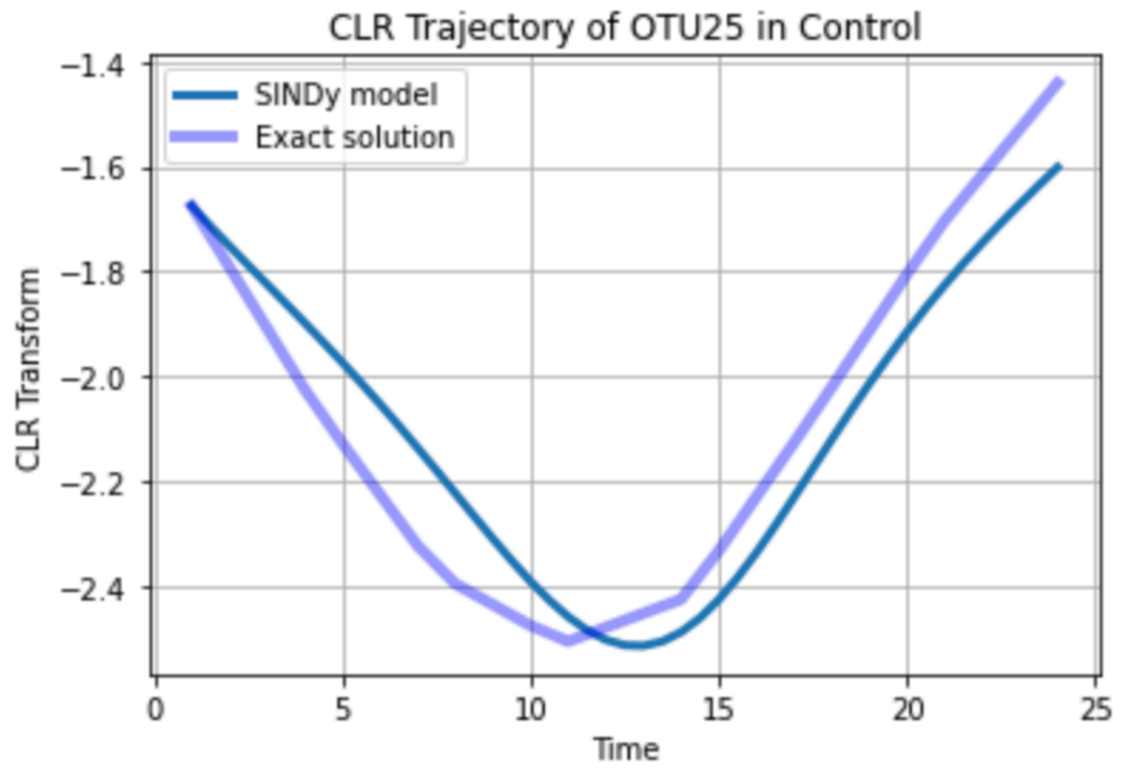
5. Dynamical Systems fitting & future directions

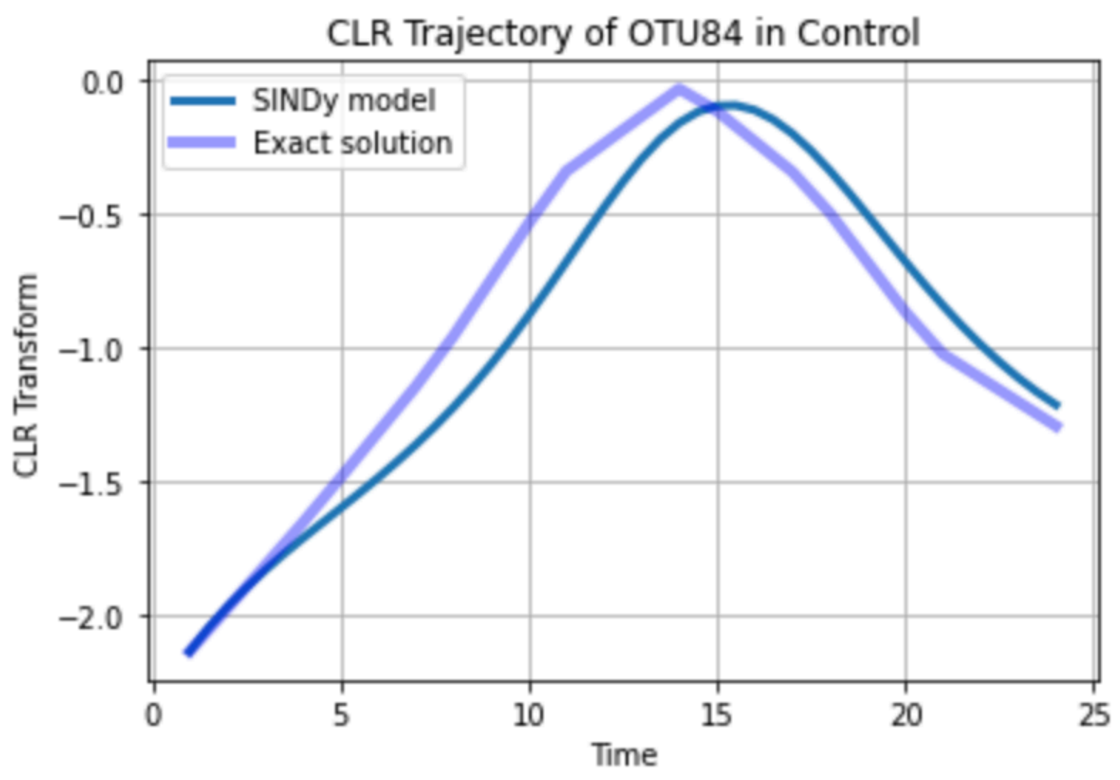
- Generalised Lotka-Volterra (predator-prey) equations can also be used to describe these dynamical systems.

$$\underbrace{\frac{dx_i}{dt}}_{\text{rate of growth of } i^{th} \text{ species}} = x_i\left(\underbrace{r_i}_{\text{intrinsic specific growth rate}} + \sum_{j=1}^{n} \underbrace{\propto_{ij}}_{\text{interaction coefficient}} x_j\right)$$

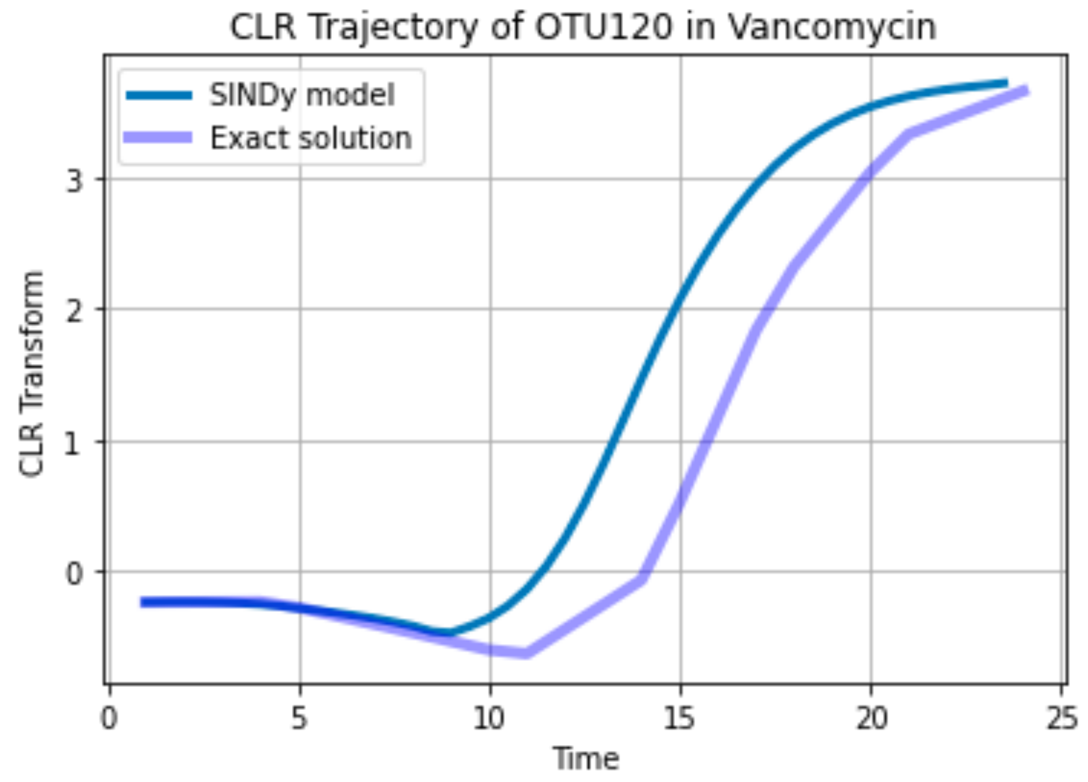
Consider OTUs trajectories for n OTUs where each OTU is governed by its own version of this equation. i=1…n.

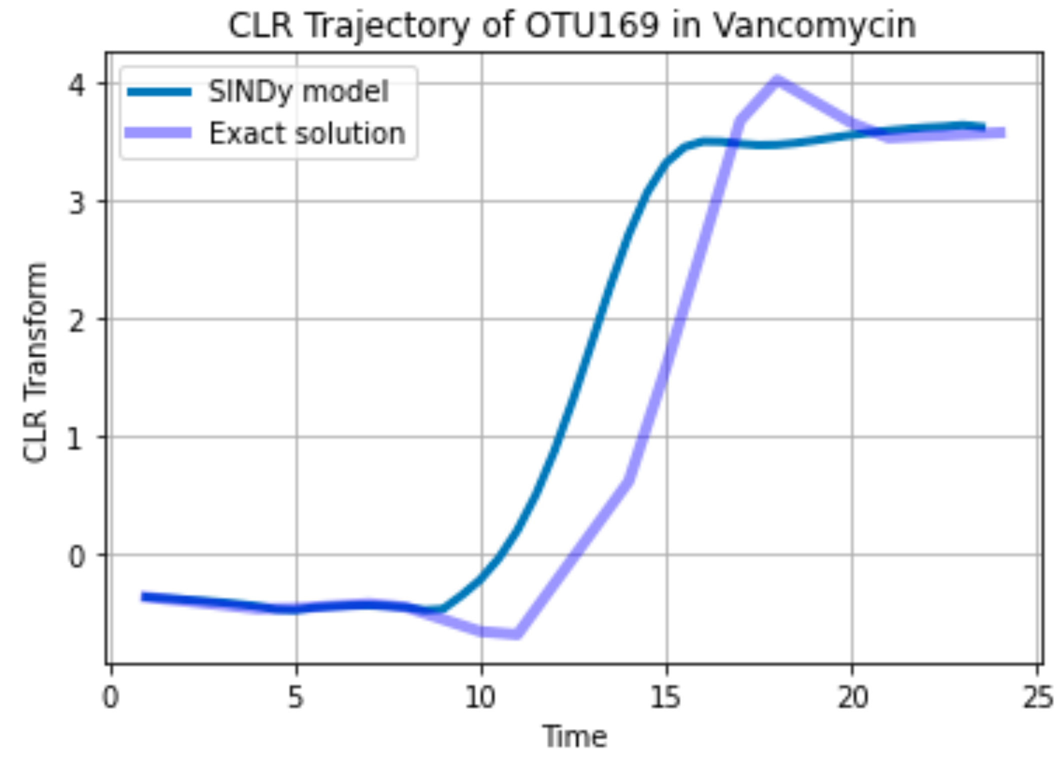
- **PySindy** was used to fit 2 OTUs trajectories within:
- 1. Control
- 2. Vancomycin - incorporated perturbations such as antibiotic exposure and VRE colonisation.






```
OTU25' = -0.033 1 + -0.394 OTU25 + -0.140 OTU25^2 + -0.087 OTU25 OTU84
OTU84' = 0.689 1 + 1.540 OTU25 + -0.372 OTU84 + 0.517 OTU25^2 + -0.048 OTU25 OTU84
```






```
OTU120' = 0.148 1 + 0.223 OTU120 + 0.224 OTU169 + 0.065 u0 + -0.084 OTU120^2 + -0.049 OTU120 u0 + -0.043 OTU169^2 +
0.065 u0^2
OTU169' = 0.195 1 + 0.696 OTU120 + 0.087 OTU169 + 0.178 u0 + 0.017 u1 + 0.538 OTU120^2 + -1.082 OTU120 OTU169 + -0.08
2 OTU120 u0 + -0.271 OTU120 u1 + 0.309 OTU169^2 + 0.178 u0^2 + 0.017 u1^2
```

*u0 and u1 refer to perturbation terms

Key Idea: Using functions like polynomials to construct simple equations that describe the dynamics of a collection of OTU trajectories.

Thank you and Q & A.