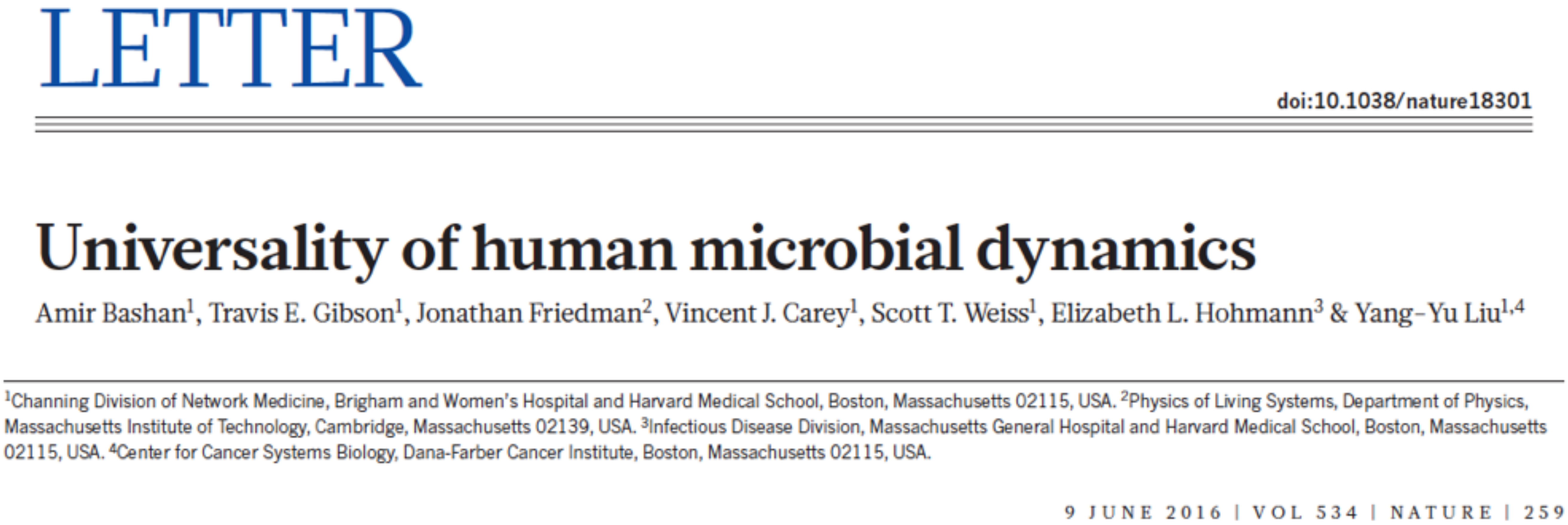


LETTER

doi:10.1038/nature18301

Universality of human microbial dynamics

Amir Bashan[1], Travis E. Gibson[1], Jonathan Friedman[2], Vincent J. Carey[1], Scott T. Weiss[1], Elizabeth L. Hohmann[3] & Yang-Yu Liu[1,4]

[1]Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. [2]Physics of Living Systems, Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [3]Infectious Disease Division, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. [4]Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA.

9 JUNE 2016 | VOL 534 | NATURE | 259

JC by KKY 9/1/2016

Human microbiome is a dynamical system

A microbial system can be represented via a population dynamic model:

$$\boldsymbol{x}^{(\nu)}(t) = (x_1^{(\nu)}(t), \ldots, x_N^{(\nu)}(t))$$

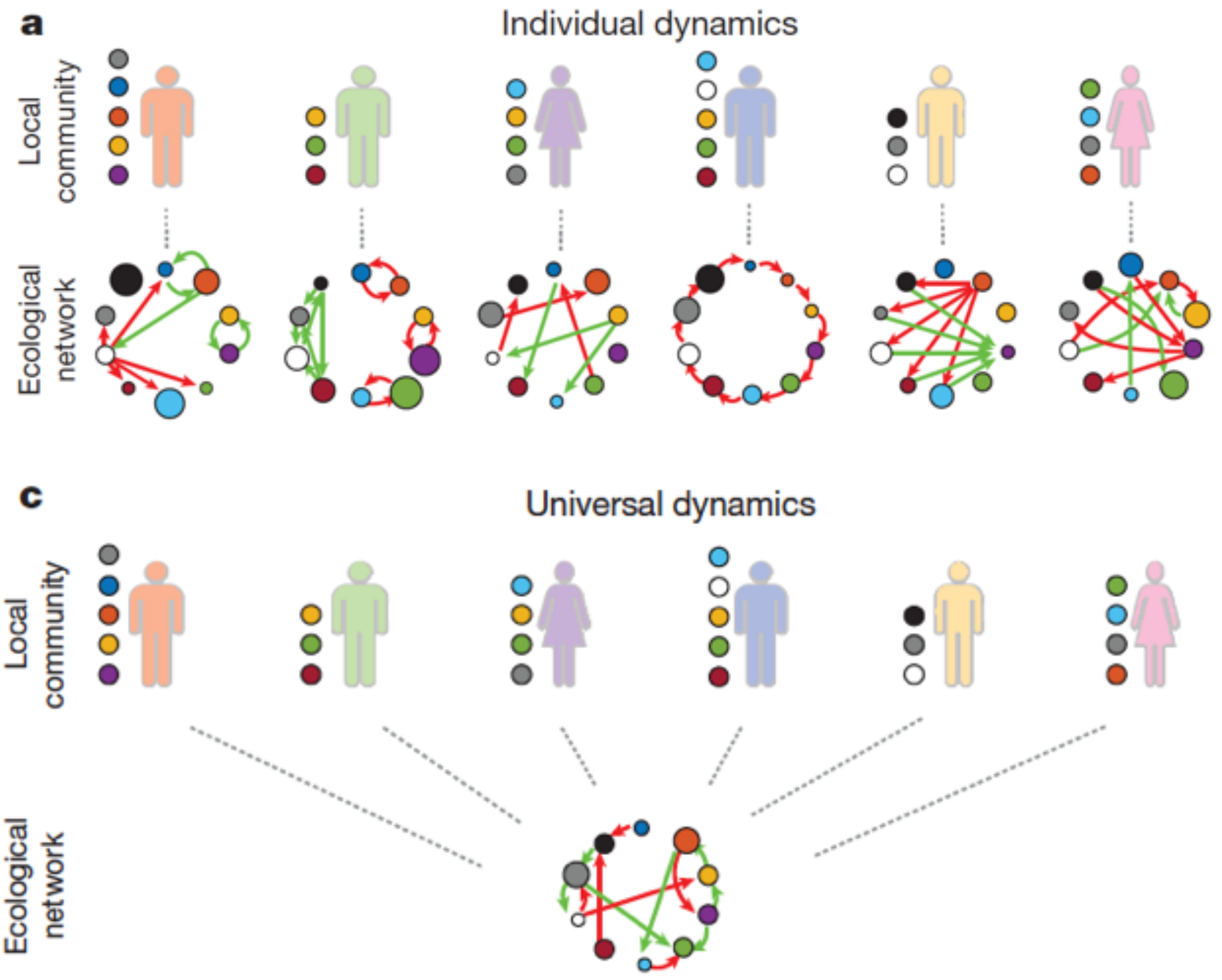
is a time-dependent abundance profile

$$\dot{\boldsymbol{x}}^{(\nu)} = \boldsymbol{f}(\boldsymbol{x}^{(\nu)}; \boldsymbol{\Theta}^{(\nu)})$$

non-linear function

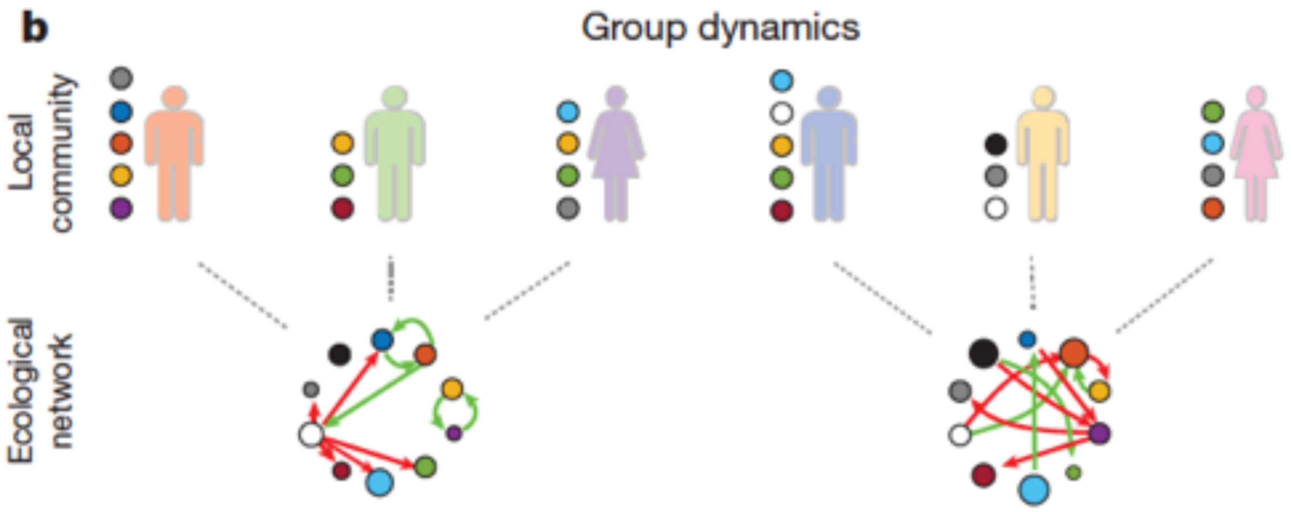
all ecological parameters: growth rates, intra-species and inter-species interactions

are these parameters host-independent?



Why should we care?

- Microbiome-based therapies:
  - A universal dynamics - general interventions to control microbial state
  - Host-specific - personalized interventions
  - middle ground: population specific dynamics



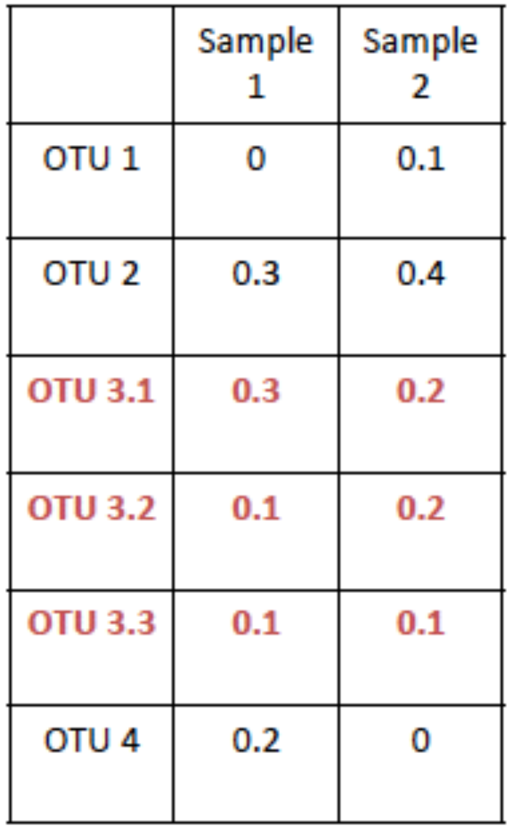
- Safety concern for faecal microbiota transplantation (FMT): Microbiota are good in donor’s gut may not be good in the recipient’s.

Computational challenges and novelties

- Naively speaking, there are 2 ways to infer the parameters in the dynamical model:
    - large number of healthy individuals from temporal metagenomics data
    - correlation-based analysis of cross-sectional data
- Rather than asking what the dynamics are, the paper wants to determine whether the dynamics is universal or not. The question is much easier.

How can we quantify variability between two samples?

- overlap between species assemblages



	Sample 1	Sample 2
OTU 1	0	0.1
OTU 2	0.3	0.4
OTU 3.1	0.3	0.2
OTU 3.2	0.1	0.2
OTU 3.3	0.1	0.1
OTU 4	0.2	0

relative abundances

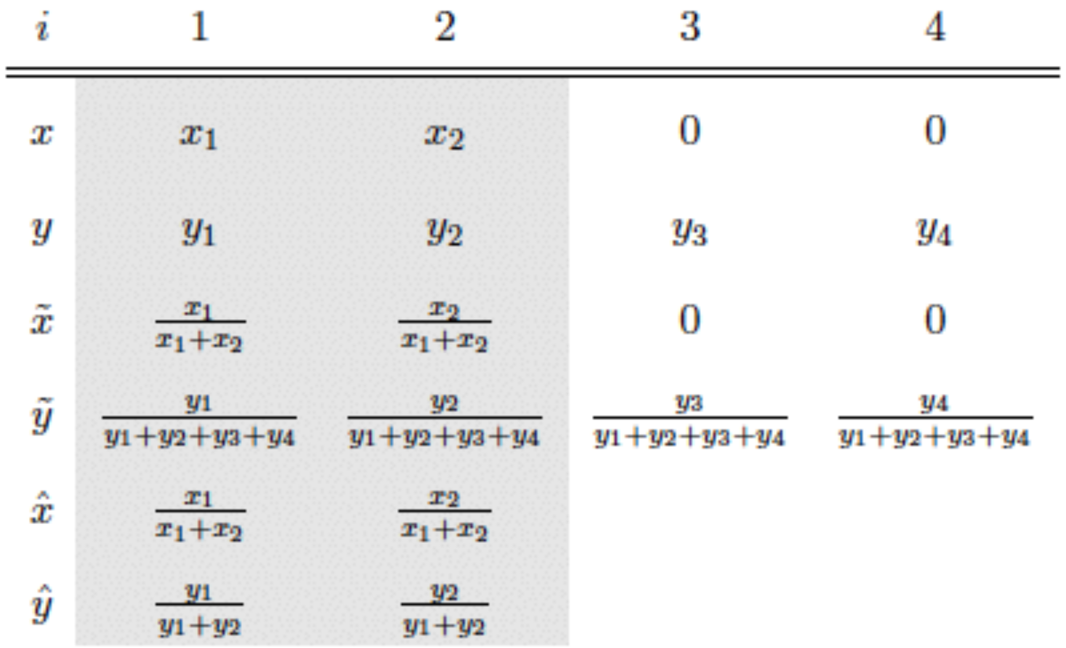
$$O(\tilde{x},\tilde{y}) \equiv \sum_{i\in S} \frac{\tilde{x}_i+\tilde{y}_i}{2}$$

S: set of common species

if S is empty, O=0
if 2 samples share all species, O=1
(no matter the difference in relative abundances)

Overlap = 0.85
Jaccard = 0.66

- dissimilarity between shared species



i	1	2	3	4
x	x_1	x_2	0	0
y	y_1	y_2	y_3	y_4
$\tilde{x}$	$\frac{x_1}{x_1+x_2}$	$\frac{x_2}{x_1+x_2}$	0	0
$\tilde{y}$	$\frac{y_1}{y_1+y_2+y_3+y_4}$	$\frac{y_2}{y_1+y_2+y_3+y_4}$	$\frac{y_3}{y_1+y_2+y_3+y_4}$	$\frac{y_4}{y_1+y_2+y_3+y_4}$
$\hat{x}$	$\frac{x_1}{x_1+x_2}$	$\frac{x_2}{x_1+x_2}$		
$\hat{y}$	$\frac{y_1}{y_1+y_2}$	$\frac{y_2}{y_1+y_2}$		

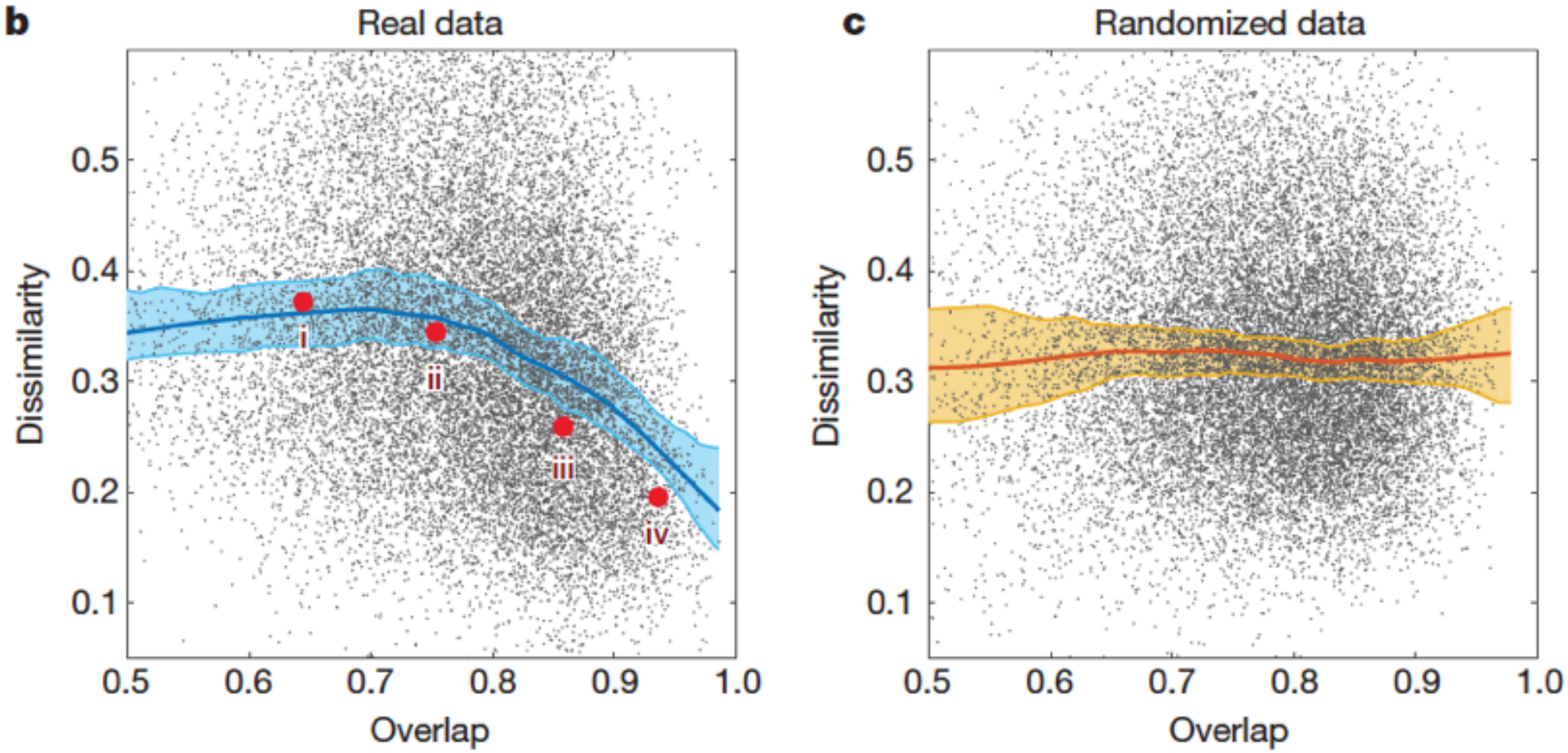
$$D(\hat{x},\ \hat{y}) = D_{\mathrm{rJSD}}(\hat{x},\ \hat{y}) \equiv \left[\frac{D_{\mathrm{KL}}(\hat{x},\boldsymbol{m}) + D_{\mathrm{KL}}(\hat{y},\boldsymbol{m})}{2}\right]^{\frac{1}{2}}$$

$$\boldsymbol{m} \equiv \frac{\hat{x}+\hat{y}}{2} \text{ and } D_{\mathrm{KL}}(\hat{x},\hat{y}) \equiv \textstyle\sum_{i\in S} \hat{x}_i \log\frac{\hat{x}_i}{\hat{y}_i}$$

Key observation:
Given a pair of samples, adding or removing unique species in whatever abundance will lead to a change in overlap, but NOT the dissimilarity

Dissimilarity-overlap curve (DOC)

HMP study: 190 gut samples. 17955 pairs

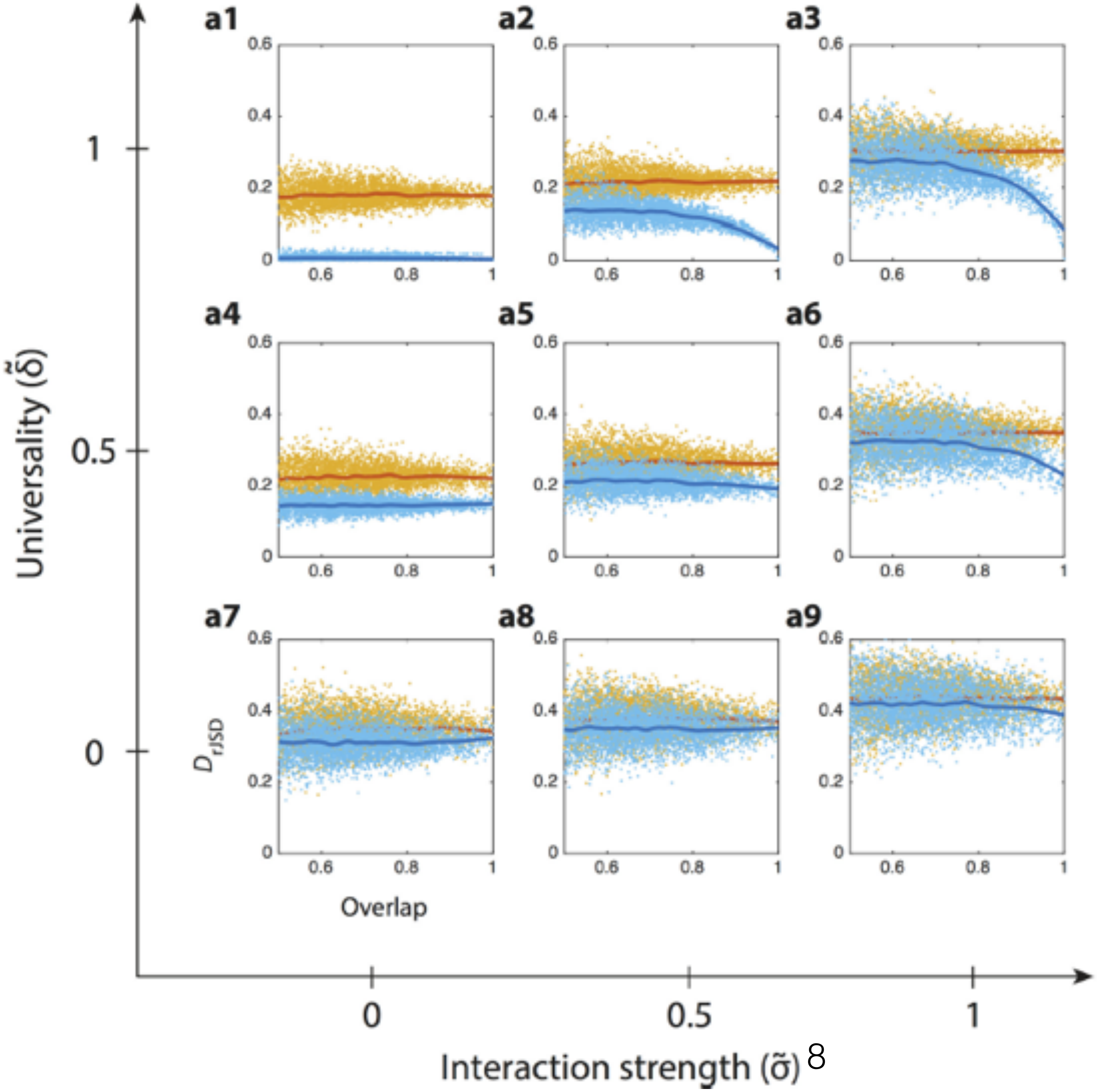


abundance profiles of sample pairs become
more similar as their overlap becomes higher

What’s the big deal of the negative slope?

- Claims:
    - A flat DOC is an indication of individual dynamics
    - A negative slope in the high-overlap regions indicates universal dynamics

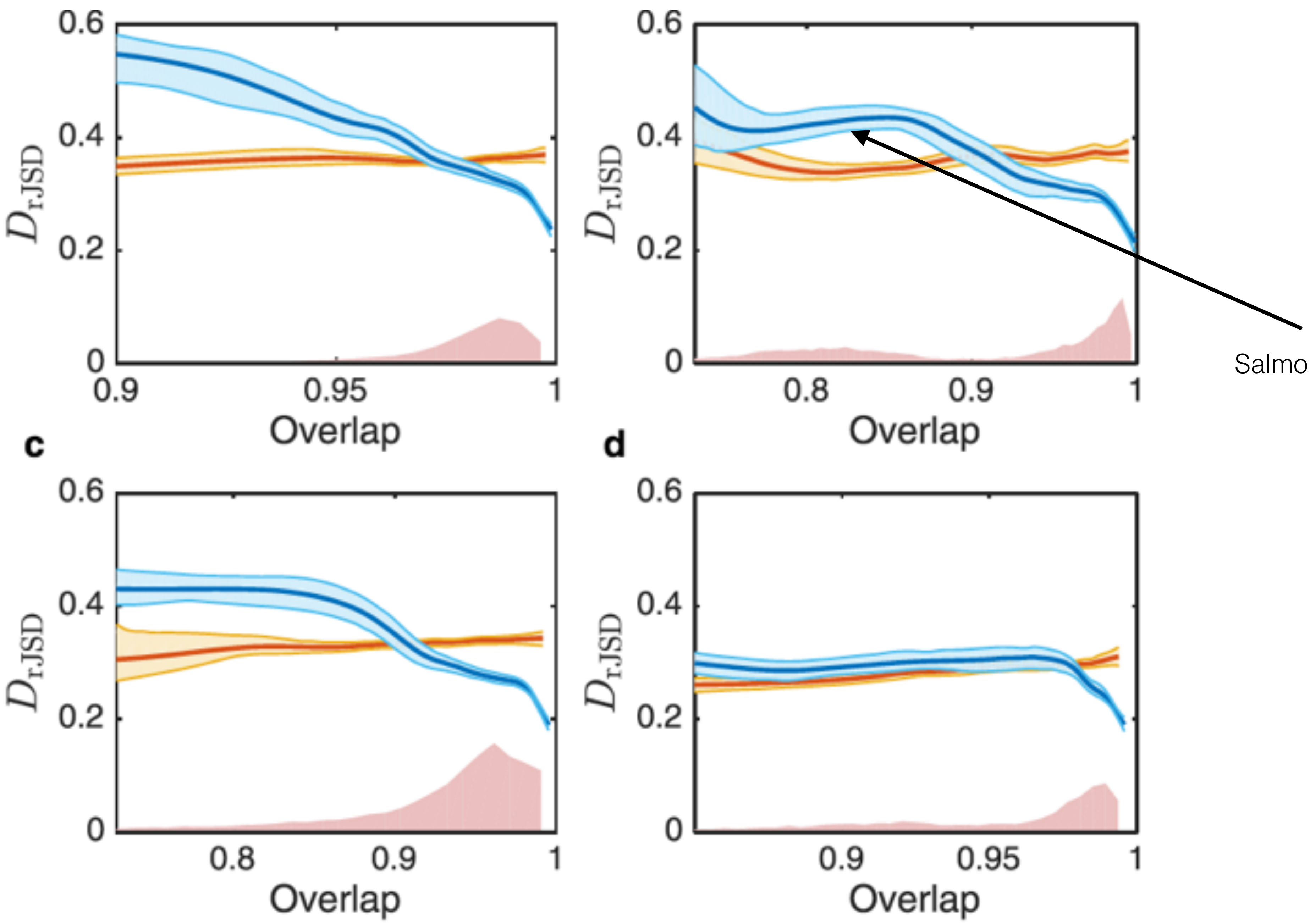
Support by simulated population dynamics models



Lotka-Volterra model

$$\frac{dx_i}{dt} = r_i x_i \left(1 - \frac{\sum_{j=1}^{N} \alpha_{ij} x_j}{K_i}\right)$$

DOC analysis of gut samples from longitudinal studies



Salmonella infection

time-invariant of healthy individuals

DOC analysis in different body sites

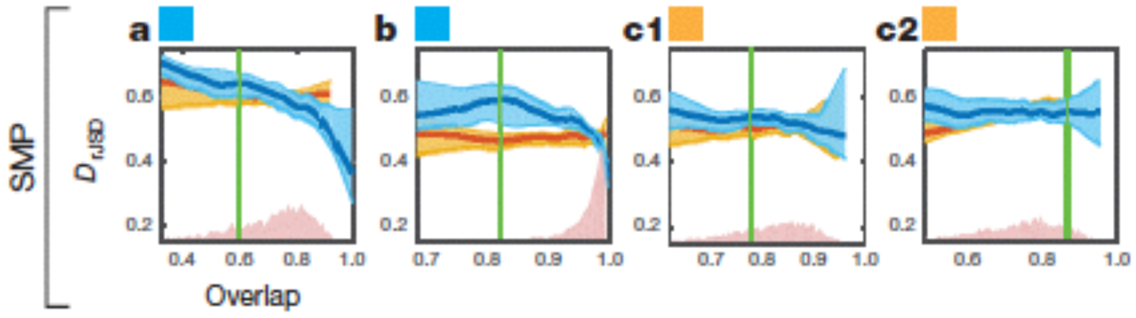
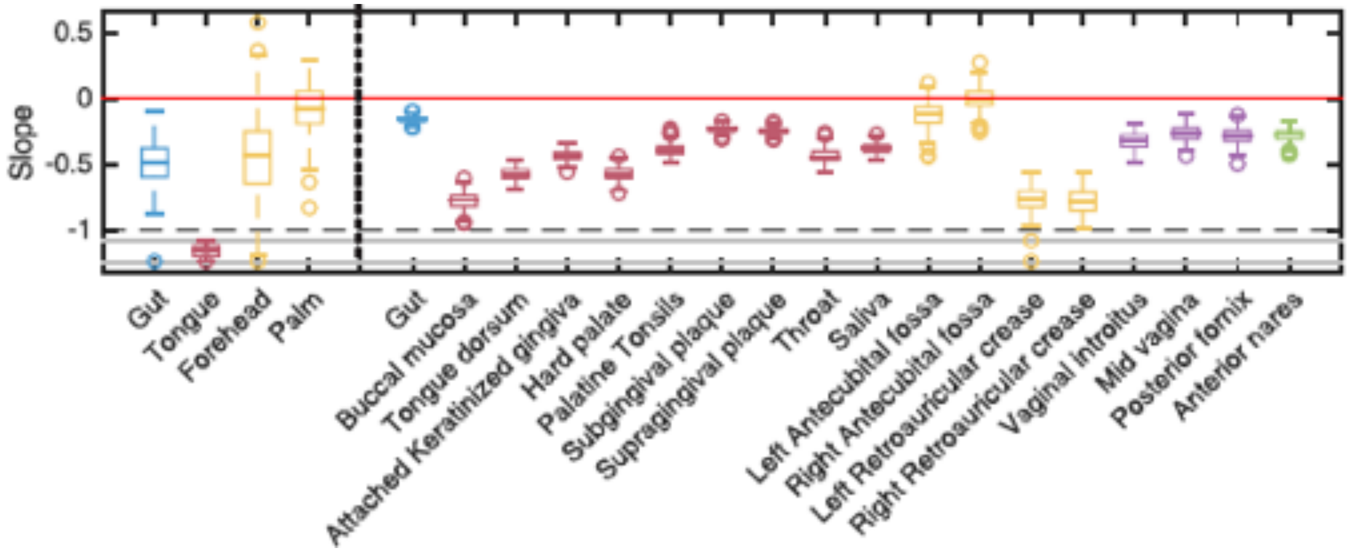
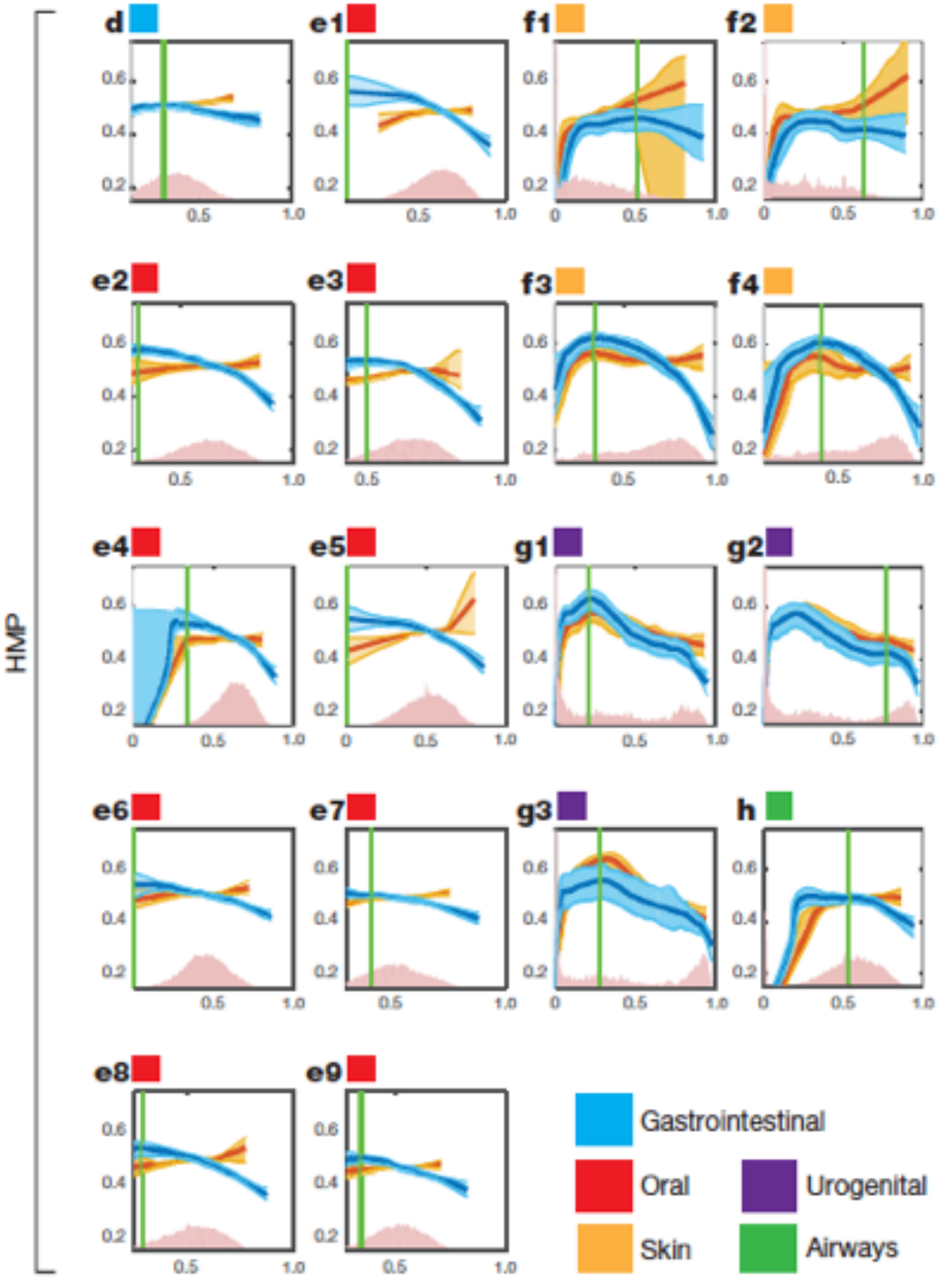
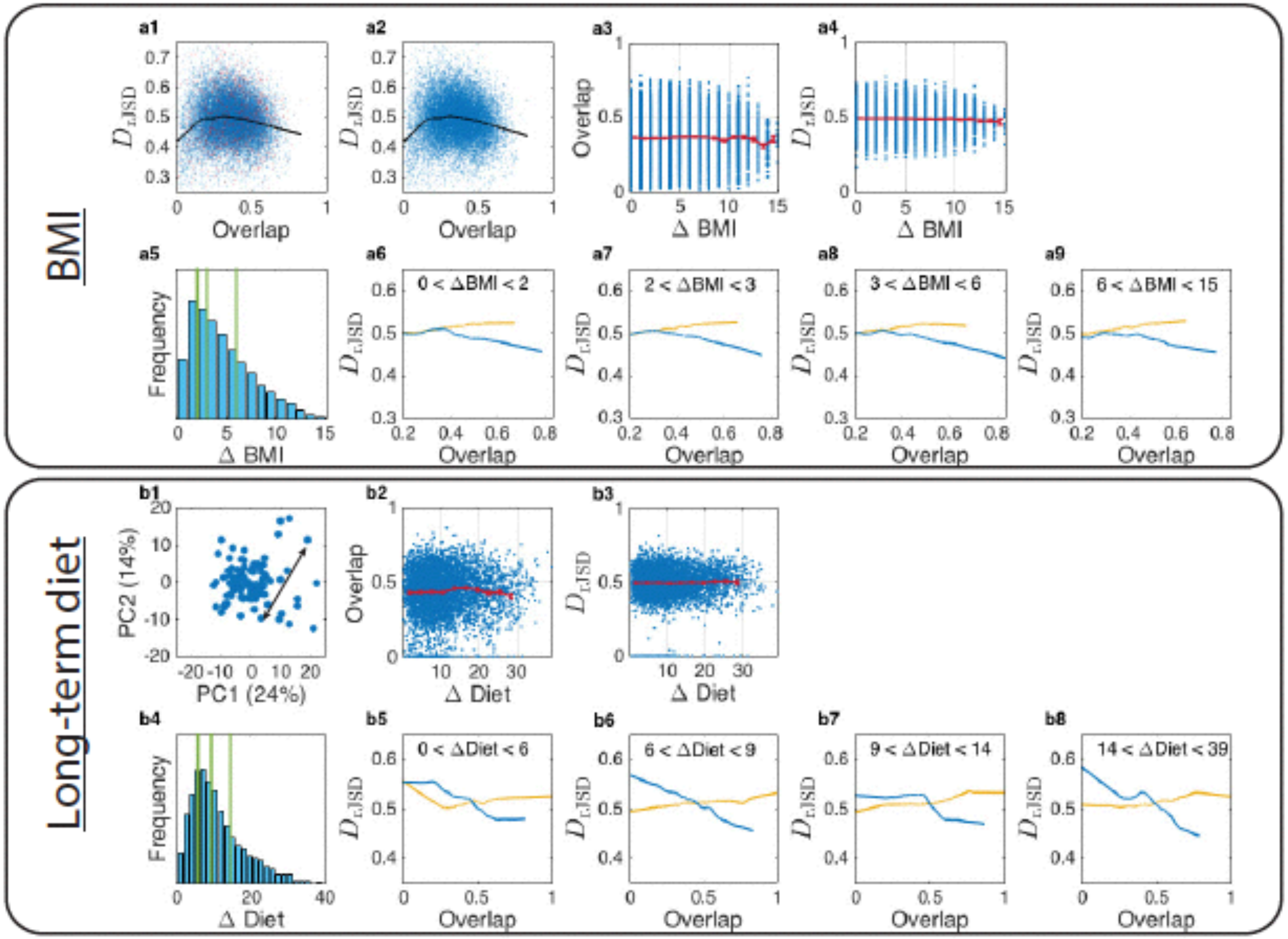


Figure 3 | Detecting universality of microbial dynamics in different body sites. a–h, We calculated DOCs for real (dark blue) and randomized (dark red) samples of two data sets: (1) SMP: gut (**a**), tongue (**b**), forehead skin (**c1**), palm skin (**c2**); (2) HMP: gut (**d**), tongue dorsam (**e1**), attached keratinized gingiva (**e2**), buccal mucosa (**e3**), hard palate (**e4**), palatine tonsils (**e5**), subgingival plaque (**e6**), supergingival plaque (**e7**), throat (**e8**), saliva (**e9**), left/right antecubital fossa (**f1**/**f2**), left/right retroauricular crease (**f3**/**f4**), vaginal introitus (**g1**), mid-vagina (**g2**), posterior fornix (**g3**), anterior nares (**h**). The overlap distributions of the real between-subjects sample pairs are shown in pink. The vertical green line represents the change point (see Methods).

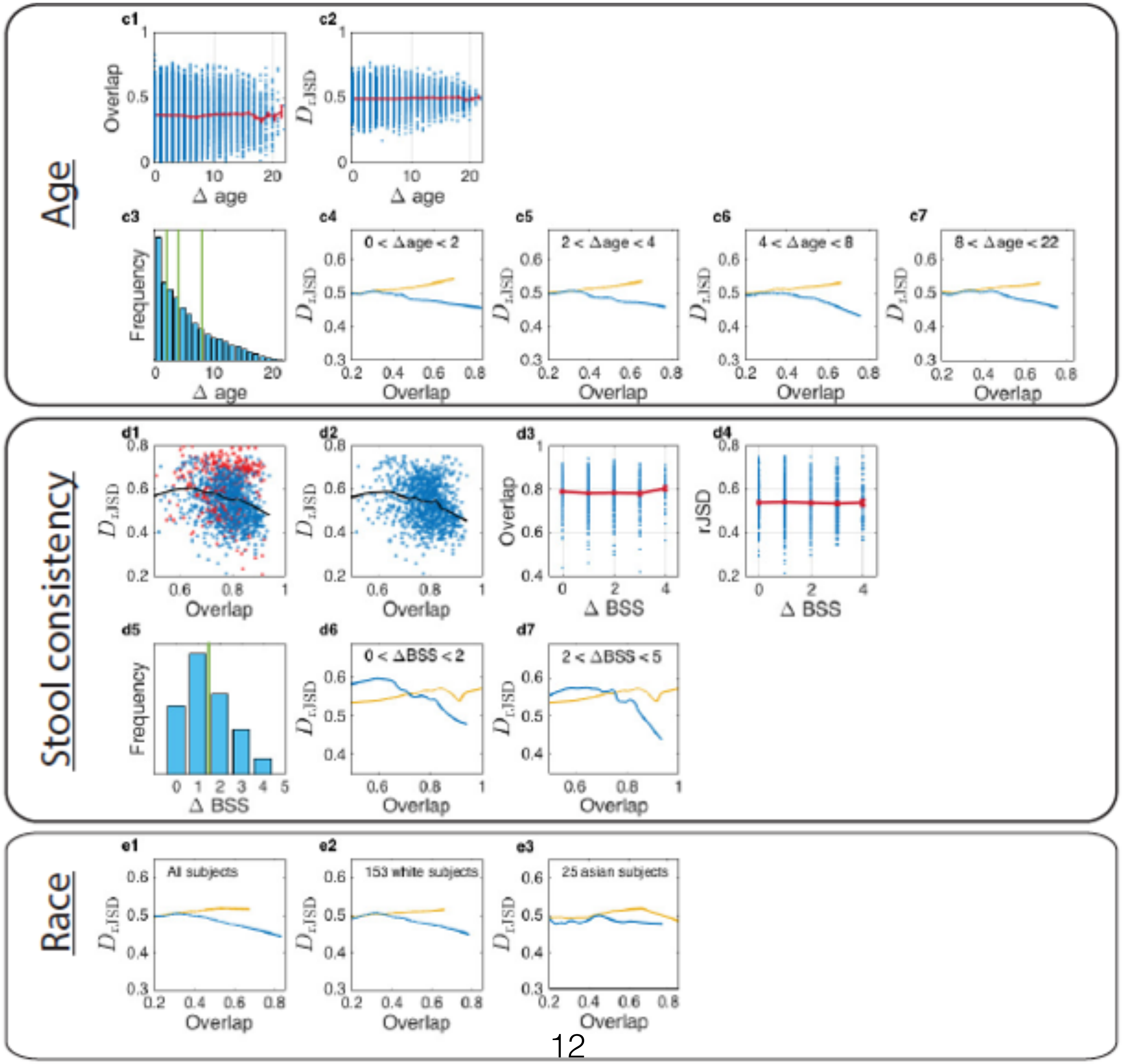




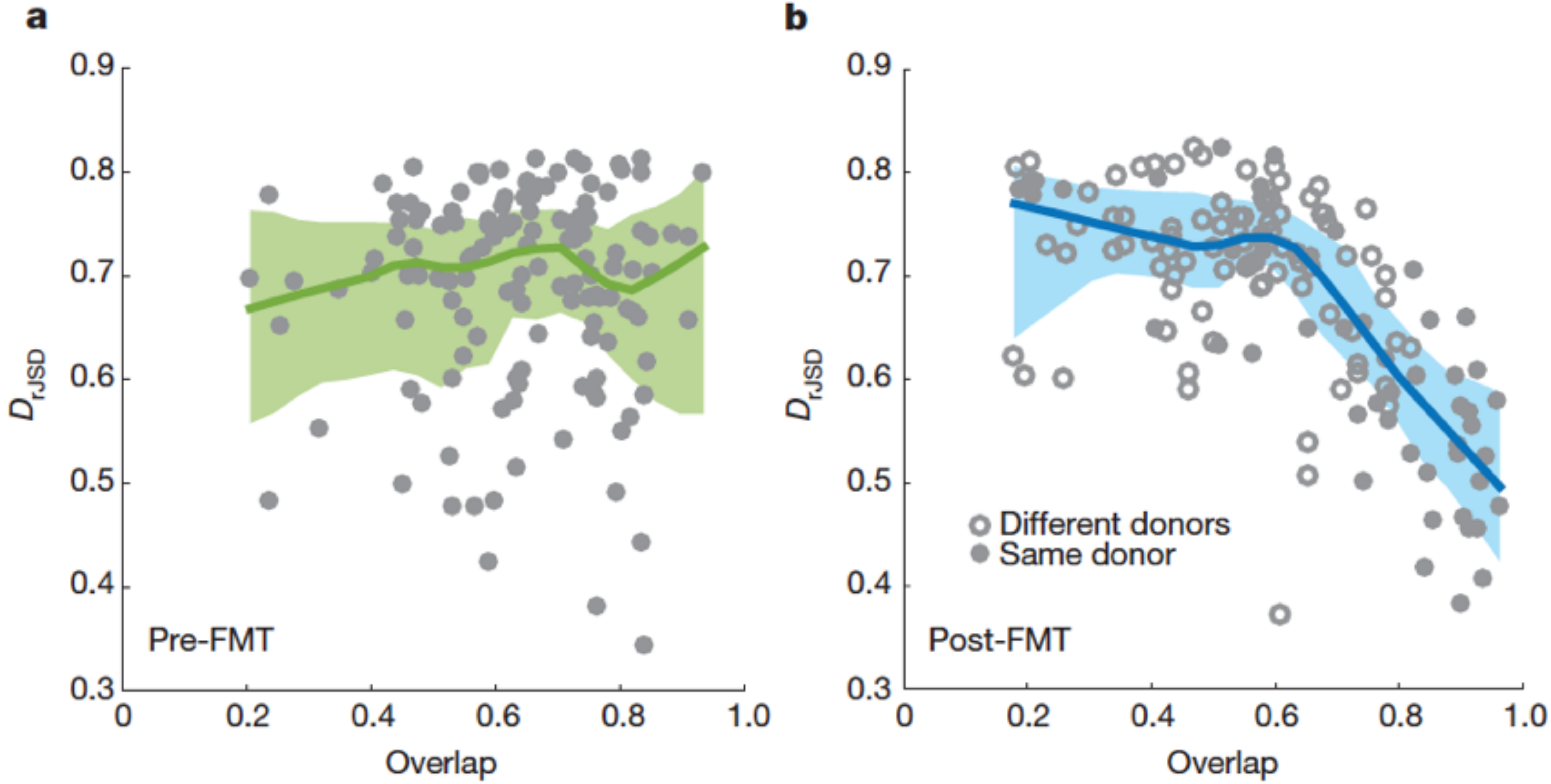
Effects of host factors



Effects of host factors



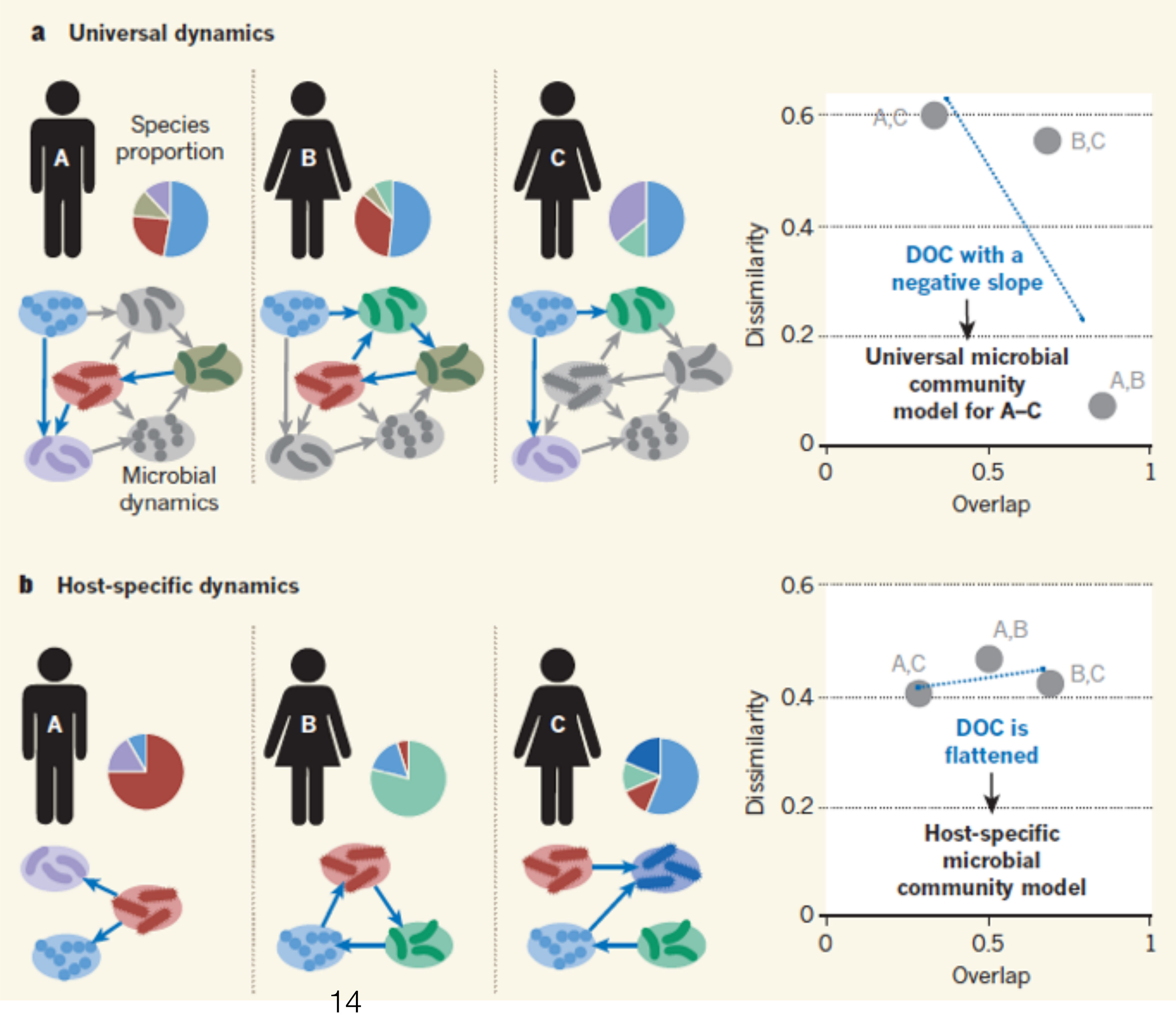
DOC analysis of human subjects with recurrent Clostridium difficile infection



Conclusion

A simple feature (DOC curve) interestingly distinguishes different microbial systems.

Steady state -> (??) dynamics



Faust and Raes Nature 2016