Title:

Health and disease imprinted in the time variability of the human microbiome

Running title:

Microbiota, are you sick?

Jose Manuel Martí[1,2,*], Daniel Martínez-Martínez[1,2,3,*], Manuel Peña[2], César Gracia[1,2], Teresa Rubio[1], Amparo Latorre[1,3,4,5], Andrés Moya[1,3,4,5,#] & Carlos P. Garay[1,2,#]

[1]Institute for Integrative Systems Biology (I2SysBio), 46980, Spain.

[2]Instituto de Fisica Corpuscular, CSIC-UVEG, P.O. 22085, 46071, Valencia, Spain.

[3]FISABIO, Avda de Catalunya, 21, 46020, Valencia, Spain.

[4]Cavanilles Institute of Biodiversity and Evolutionary Biology, UVEG, 46980, Spain.

[5]CIBER en Epidemiología y Salud Pública (CIBEResp), Madrid, Spain

Words count for the Abstract section: 154 of 250 max

Words count for the Importance section: 105 of 150 max

Words count for the rest of text (inc.refs): 3996 of 5,000 max

Number of floats: 8

Number of supplementary figures: 5

[*] Equally contributed

[#] Corresponding authors: andres.moya@uv.es, penagaray@gmail.com

Abstract

Animal microbiota (human included) plays an important role keeping healthy the physiological status of the host. Increasing research activity is dedicated to understand how changes in composition and function of the microbiota are associated to disease or not. We analyze 16S rRNA and shotgun metagenomic sequencing (SMS) published data from the gut microbiota of 97 individuals monitored in time. Temporal fluctuations in the microbial composition reveal significant differences due to factors such us dietary changes, antibiotic intake, age or disease. Here we show that a fluctuation scaling law describes the temporal changes in the gut microbiota. This law allows to estimate the temporal variability of the microbial population and quantitatively characterizes the path toward disease by a noise-induced phase transition. The estimation of the systemic parameters for follow-up studies may have clinical use and, more generally, applications in other fields where it is important to know if a given community is stable or not.

Importance

Human microbiota is tightly associated to the health status of a person. Here we analyse the microbial composition of several subjects under different conditions, over a time span that ranges from days to months. Using the Langevin equation as the basis of our mathematical framework in order to evaluate microbial temporal stability, we prove that we are capable to distinguish stable from unstable microbiotas. This first step will help us to determine how microbiota temporal stability is related to the healthiness of the people, and it will allow the development of a more complete framework in order to deepen the knowledge of this complex system.

Keywords— microbiome, systems biology, ecological modeling, metagenomics, stability

Introduction

The desire to understand the factors that influence human health and cause diseases has always been one of the major driving forces of biological research. As evidence of new concepts 'holobiont' and 'hologenome' is increasing each day (*1*, *2*), research not only focus on the human physiology but also on the microbial population that surround ourselves. However, these concepts are still in debate (*3*). We are populated by a myriad of microorganisms that are interacting with us in several physiological processes such as metabolism of the bile acids (*4*), of the choline (*5*) or key-route metabolites as short-chain fatty acids (*6*, *7*) which are also involved in immune system maturation (*8*, *9*). Human microbiota has been suggested to be closely related to diseases like type 2 diabetes (*10*), cardiovascular disease (CVD) (*11*), irritable bowel syndrome (*12*), Crohn's disease (*13*), some affections as obesity (*14*, *15*), malnutrition (*16*) among other multiple diseases (*17*). Current studies reveal that gut microbiota also influences brain function and behaviour and is related with neurological disorders like Alzheimer's disease through the brain-gut-microbiome axis (*18*, *19*). Recently, even a mystifying and elusive diagnosis condition as chronic fatigue syndrome, which has often been suggested to be a psychosomatic disease, has been closely related to reduced diversity and altered composition of the gut microbiome (*20*).

High throughput methods for microbial 16S ribosomal RNA gene and SMS (shotgun metagenomic sequencing) have now begun to reveal the composition of archaeal, bacterial, fungal and viral communities located both, in and on the human body. Modern high-throughput sequencing and bioinformatics tools provide a powerful means of understanding how the human microbiome contributes to health and its potential as a target for therapeutic interventions (*21*). To define normal host-gut microbe interactions and how the microbiota compositional changes can origin some diseases are important issues still in need for scientific answers (*22–24*).

Biology has recently acquired new technological and conceptual tools to investigate, model and understand living organisms at the system level, thanks to the spectacular progress in quantitative techniques, large-scale measurement methods and the integration of experimental and computational approaches. In particular, Systems Biology has placed a great effort to unveil the general laws governing the complex behaviour of microbial communities (*25–27*), even proposing that they have universal dynamics (*28*). Microbiota can be approached under the light of ecological theory where we can find, for instance, general principles as the Taylor's law (*29*), which relates spatial or temporal variability of the population with its mean. This law, also known as fluctuation scale law, is ubiquitous in the natural world and can be found in several systems as random walks (*30*), stock markets (*31, 32*), tree (*33*) and animal populations (*29, 34, 35*), gene expression (*36*), or in the human genome (*37*). Taylor's law has been applied to microbiota in a spatial way in the work of Zhang *et al.*, (2014) (*38*), where they show that this population tend to be in an aggregated way rather than in a random distribution. Despite its ubiquity, it has been studied only in experimental settings (*39, 40*) but never been applied in follow-up studies from microbiota even that a great effort has been made to infer the community structure from a dynamical point of view (*41–43*)

Here we present the imprints of health status (healthy or disease) in macroscopic properties of microbiota, by studying its temporal variability. We have analyzed more than 35000 time series of taxa from the gut microbiome of 97 individuals obtained from publicly available high throughput sequencing data on different conditions: diseases, diets, obese status, antibiotic therapy and healthy individuals. Having seen that all cases follows Taylor's law, we use this empirical fact to model how the relative abundances of taxa evolves toward time thanks to the Langevin equation, in a similar way as it was applied recently by Blumm *et al.* (*44*). We use this mathematical framework to explore the temporal stability of the microbiota in different conditions in order to understand how this affects the healthy status of the subjects.

Results

We have analyzed the microbiome temporal variability to extract global properties of the system. As fluctuations in total counts are plagued by systematic errors we worked on temporal variability of relative abundances for each taxon. Our first finding was that, in all cases, changes in relative abundances of taxa follow a ubiquitous pattern known as the fluctuation scaling law (*45*) or Taylor's power law (*29*), i.e., microbiota of all detected taxa follows $\sigma_i = V \cdot x_i^{\beta}$, a power law dependence between mean relative abundance x_i and dispersion σ_i. The law seem to be ubiquitous, spanning even to six orders of magnitude in the observed relative abundances. As can be seen in Figure 1, the most abundant species are less volatile in relative terms than the less abundant. The fitting to the power law is always robust ($R^2 > 0.88$) and does not depend on the microbiome condition. The power law (or scaling) index β and the variability V (hereafter Taylor parameters) appear to be correlated with the stability of the community. . The variability V is a direct estimator of amplitude of fluctuations over time, while the scaling index give us information about the statistic properties of the ecosystem. The fact that it is less than the unity (when looking at the standard deviation) tell us that the most abundant taxa are less susceptible to any perturbation than the less abundant elements of the population. The Taylor parameters are related with the health status of the host, which we consider the main finding exposed in this article.

Taylor parameters describing the temporal variability of the gut microbiome in our sampled individuals are shown in Supplementary Tables S1 to S7. Our results hint at an ubiquitous behavior. On the first hand, the variability (which corresponds to the maximum amplitude of fluctuations) is large, which suggests resilient capacity of the microbiota. On the other hand, the scaling index is always smaller than one, which means that more abundant taxa are less volatile than less abundant ones. In addition, Taylor parameters for the microbiome of healthy individuals in different studies are compatible within estimated errors. This enables us to define an area in the Taylor parameter space that we called the *healthy zone*.

In order to jointly visualize and compare the results of individuals from different studies (*12*, *46*–*51*), their Taylor parameters have been standardized, where standardization means that each parameter is subtracted by the mean value and divided by the standard deviation of the group of healthy individuals for each study (for details of the procedure, please see Standardization subsection in Material and Methods). The healthy zone and the standardized Taylor parameters for individuals whose gut microbiota is compromised (i.e., suffering from kwashiorkor, a Salmonella infection, travel abroad, altered diet, antibiotics or IBS) is shown in Figure 2. Children developing kwashiorkor show smaller variability than their healthy twins. A meat/fish-based diet increases the variability significantly when compared to a plant-based diet. All other cases presented increased variability, which is particularly severe, and statistically significant at more than 95% CL, for obese patients grade III on a diet, individuals taking antibiotics, the subject who had a Salmonella infection, the subject who did a travel abroad or the IBS–diagnosed patients. A global property emerges from all worldwide data collected: Taylor parameters characterize the statistical behavior of microbiome changes. Furthermore, we have verified that our conclusions are robust to systematic errors due to taxonomic assignment (see Taxa level selection in Material and Methods).

Taylor's power law has been explained in terms of various effects, all without general consensus. It can be shown to have its origin in a mathematical convergence similar to the central limit theorem, so virtually any statistical model designed to produce a Taylor law converge to a Tweedie distribution (*52*), providing a mechanistic explanation based on the statistical theory of errors (*53*–*55*). To unveil the generic mechanisms that drive different scenarios in the β–V space, we model the system by assuming that taxon relative abundance follows a Langevin equation with, on the one hand, a deterministic term that captures the fitness of each taxon and, on the other hand, a randomness term associated with Gaussian random noise (*44*). Both terms are modeled by power laws, with coefficients that can be interpreted as the taxon fitness F_i and the variability V (see Model under Material and Methods). In this

model, when V is sufficiently low, abundances are stable in time. Differences in variability V can induce a noise-induced phase transition in relative abundances of taxa. The temporal evolution of the probability of a taxon having abundance x_i given its fitness is governed by the Fokker–Planck equation. The results of solving this equation show that the stability is best captured by a phase space determined by fitness F and amplitude of fluctuations V (see Figure 3).

The model predicts two phases for the gut microbiome: a stable phase with large variability that permits some changes in the relative abundances of taxa; and an unstable phase with larger variability, above the phase transition, where the order of abundant taxa varies significantly with time. The microbiome of all healthy individuals was found to be in the stable phase, while the microbiome of several other individuals was shown to be in the unstable phase. In particular, individuals taking antibiotics and IBS–diagnosed patient P2 had the most severe symptoms. In this phase diagram, each microbiota state is represented by a point at its measured variability V and inferred fitness F. The model predicts high average fitness for all taxa, i.e., taxa are narrowly distributed in F. The fitness parameter has been chosen with different values for demonstrative purposes. Fitness is larger for the healthiest subjects and smaller for the IBS–diagnosed patients.

Rank stability of the taxa

The rank dynamics and stability plots in Figure 4 and 5 show the variation in the rank with time for the most dominant taxa and their calculated Rank Stability Index (RSI, as discussed in Material and Methods) for the taxa of a healthy subject (individual *A*, Figure 4) and from a subject diagnosed with IBS (patient *P2*, Figure 5) of the IBS study (*12*). The taxa are listed ordered by the accumulated frequency along the time series, so y-axis is an overall dominance axis for each sample set. Generally speaking, we observe that the most dominant taxa are

the most rank stable.

Nevertheless, in the particular case of the healthy individual in Figure 4, *Burkholderiales* and *Betaproteobacteria* (taxa ordered as 18th and 25th in the dominance axis) show comparatively very low rank stability regarding similar dominant taxa while, on the other hand, *Comamonadaceae*, *Lactobacillaceae*, *Fusobacteriaceae*, *Aerococcaceae* and *Carnobacteriaceae* show higher stability than other more dominant taxa, forming a kind of *rank stability island* for medium-ranked taxa around position 40 in the dominance axis, and thus colored in orange, following the color criteria shown in the table included in Figure 4, since they show a moderately stable RSI.

In the IBS diagnosed patient of Figure 5, beyond the differences in dominance for the particular taxa, we still observe that the most dominant are the most rank stable. However, as opposed to the healthy individual results, far from presenting a *rank stability island*, the medium-ranked taxa are very rank unstable, mostly due to transient (often one or two consecutive samples) but deep drops in their relative abundance, which are usually happening more than twice along the time series. That is, for instance, the case of *Sphingobacteriales* with two non-consecutive samples dropping to 111th rank position. In other cases, the high rank instability comes from a rank fluctuation over all the time series, as for *Streptococcaceae* and *Burkholderiales*, which are ranking 26th and 29th respectively in the overall dominance axis but show very low RSI, and thus colored in black attending to the color criteria shown in the table included in Figure 5.

We found the presence of such of *rank stability island* for medium-ranked taxa in the other healthy subjects (*B* and *C*) of the IBS study (*12*) together with its total absence for the other IBS diagnosed patient (patient *P1*), which also presents very high rank instability in its medium-ranked taxa. See Supplementary Figure 1 and Supplementary Figure 2 for further examples of rank plots for a healthy and light IBS diagnosed patients, respectively.

Time dependence of model parameters

Finally, we have studied the time dependence of the variability V and power law index β (see Model under Material and Methods) by using a sliding window approach. The total number of time points are divided in subsets of five points, where the next subset is defined by adding the next time sampling and by eliminating the earliest one. Both parameters were calculated for each subset against the average time lapse. Figure 6 shows the variability V as a function of time for the largest sampling: two individuals in the Caporaso's study (*46*) corresponding to the gut microbiota of a male (upper plot) and a female (lower plot). Both samples shows changes in the variability V with quasi–periodic behavior peaked at about 10 days. Variability grows more for the gut microbiota of the male and share a minimal value around 0.1 with the gut microbiota of the female.

Figure 7 shows the time evolution of V for patient P2 of the IBS study (*12*) (upper plot) and patient D in the antibiotics study (*47*) (lower plot). The variability of the gut microbiota of P2 decreases from above 0.3 to below 0.2, showing a slow tendency to increase the order of the system. Antibiotic intake leaks to a quick increase of variability which lasts for a few days to recover ordering. The second antibiotic treatment shows some memory (lower increase of variability) with a slower recovery.

Discussion

One of the mains features of this work is to have shown that, independently of its condition, the microbiota follows the Taylor's law. We have seen that the value of the scaling index in each case is always less than the unity (using standard deviation as the measurement for dispersion), which is informing us about the community structure. This means that, in relative terms, the most abundant elements in the population are less volatile to perturbations than

the less abundant ones. The explanation for this universal pattern is not clear although some hypothesis have been tested in other studies, as the presence of negative interactions in the population (*56*), or the demonstration that it may depend on reproductive correlation (*57*). Nevertheless, none of these explanations are enough when we are talking about microbiota as the reproduction term is diffuse, the interactions between its components are not only based on competition (*58–60*) and that even that kind of negative interaction may not effectively yield in values less than the unity when referring to a bacterial species (*40*). Anyhow, the values obtained in all cases are very similar among them, which could be suggesting that the community structure is preserved throughout the different scenarios that we have studied.

The second parameter is informing about the noise and can be directly related with the variability or the fluctuation amplitude of the population over time. It is a direct estimator of the stability of the system under study. As we showed above, the healthy subset of each study have lower variability than the non-healthy subset when dealing with adult individuals. Interestingly, the variability parameter was higher in the healthy subset for the study of the discordant twins suffering from kwashiorkor disease (*49*). In this regard it has been shown that the infant microbiota needs to develop toward a definite, adult state (*61*). This implies that the temporal variability would be greater in children respect to a healthy adult state, which should be temporally stable. Thus, our results could be pointing out the need of this variability in order to reach that adult state. Furthermore, as we wanted to see how this variability behaved over time, we calculated the evolution of this parameter for the samples which had enough time sampling. As can be seen in Figure 6, the variability of the microbiota has some fluctuations over time. It is interesting to note in Figure 7 how this parameter can capture the two antibiotic intakes in one of the patients from the study of Dethlefsen and Relman (*47*), especially that it seems to be some resilience process in the microbiota due to the lower variability increase in the second antibiotic intake.

The primary hypothesis of this work is that, in adults, having a healthy microbiota means that

population is stable in time and does not travel into a state where it is highly susceptible to external or internal perturbations, causing in such a case a dysbiotic state in the microbiota. In order to use the valuable information which gives us the empirical law of Taylor's work, we propose the use of Langevin equation to model how the ranking stability evolves in time. While we can measure directly the component of the noise of the system as their variability, the other main term needs to be inferred from the model. This term, which we have named as 'fitness', is the one that gives the ability to the system to be stable to potential perturbations. In ecological terms, this could mean the nature of interactions that are present among the bacteria, between bacteria and other minority populations as fungi or archaea, between bacteria and the viral component in the microbiota, and the interactions between host and the whole microbiota. Being this a first step to model the temporal stability of the microbiota and due to its complicated nature, we have calculated the fitness term using the Fluctuation Dissipation Theorem as a first approximation (*62*). Thus, the fitness of the microbiota still remains to be modelled in future works in order to make the model more accurate and with a higher predictive power.

By solving the Langevin differential equation, we can obtain a phase diagram where each microbiota sample can be placed according to its fitness and variability in one of two phases according to the ranking stability of the system. As we can see in the phase-space in Figure 3, we are showing three different conditions that could happen. First, we can have a healthy microbiota which could have some fluctuations as showed by one of the subjects of Caporasso et al study (*46*). Because the fitness of this cases will be high enough, the temporal variability will not place the microbiota in the unstable phase of the diagram. Second, we have a subject from the study of Dethlefsen and Relman (*47*) which is perturbed twice by an antibiotic intake. His microbiota is altered enough to lose its stability, and hence be placed in the unstable part. So located, it is more sensitive to potential perturbations as, for example, opportunist infections. Third and last condition, the subject is already in the unstable phase

due to some healthy issue as IBS. This can be observed in one of the patients from Durban et al study (*12*). In addition, it was shown that this subject improved its healthy status in the time when the experiment was done, implying that his microbiota also recovered the lost stability. It is interesting to notice that in the subject from the study of David et al. (*51*), who suffered a Salmonella infection during the experiment, we can observe a huge shift in the variability and a final recovery from the perturbed state (see 3).

Specifically, the analysis of the rank stability of the samples of healthy and IBS diagnosed patients studied in our lab (*12*), suggests that the presence of *rank stability islands* among medium-ranked taxa is a interesting feature. The higher stability of these taxa goes against the global meaning of the scaling index. Interestingly, that stability disappears when we look at the IBS patients. From the literature, it seems that some genera from the families Comamonadaceae, Neisseriaceae and Carnobacteriaceae have been reported to lower their abundance in IBS patients against healthy controls (*63*). In our case, we see that these families that are present in the rank stability island of the healthy patient decrease their rank stability index or even disappear in the IBS patient. However, we also see contradictory results in other families as Lactobacillaceae or Fusobacteriaceae, which seems to increase their abundance in IBS patients (*64*), while we observe an increase of the former and a decrease of the latter. The presence of members of the Lactobacillae family have been reported to have positive effects against gut inflammation and visceral hyperalgesia (*65*), usual symptoms of gastrointestinal disorders. It could be happening that a disorder in the stability of this particular group may help to arise the onset of symptoms associated to gastrointestinal disorders. The Aerococacceae family is also enriched in rats with IBS symptoms who have been treated with immunomodulators (*66*). Inside this island of stability, we also have families as Fusobacteriaceae o Hallomonadaceae which include pathogenic genera in them, but that are not present in the IBS patient. It could be brought into question the rule of these taxa as key players in the phase transition of the microbiota, or whether they are more susceptible

to perturbations than the most abundant. The types of interactions that could be sustaining this particular behavior are not clear, as these non-abundant taxa are not usually included in dynamical studies in order to get the community matrix. Further experiments and data analysis is needed to clarify if this is not a unique event, or it is a widespread feature of stable microbiotas.

However, we have to be aware that the hypothesis above is too simplistic to be directly related with reality. It has been demonstrated that the situation is more complex than the outlook provided separating healthy people from non-healthy people just by compositional terms, as Moya and Ferrer underline in their recent review (*17*). There are several different feasible scenarios in which we can consider the microbiota as stable independently of their compositional evolution over time. For example, depending on their ability to recover the initial composition (resilience), or whether it can recover the original function despite the composition (functional redundancy). What we have shown in this work could be explained as the transitions of a stable microbiota into a dysbiosis state.

As a first step toward understanding the microbiota stability, the model presents some limitations and there is still work to do. From the biological perspective, many questions arise from this work. We have observed the same pattern in Taylor's parameters in all the different conditions we have studied, but a pertinent question is whether it is really a universal feature in the huge diversity of microbial niches. Furthermore, another relevant question is which mechanisms are involved in maintaining the population structure. The nature of the interactions among the elements of the community is surely of great importance in this matter, and it is related to the fitness of the community as has been commented above. How we should address the community fitness is not clear, but works as Tikhonov's (*67*) could help us to aim at the correct direction toward unraveling the complexity of the microbiota.

Materials and Methods

Model

We model the microbial abundances across time along the lines of Blumm *et al.* (*44*). The dynamics of taxon relative abundances is described by the Langevin equation:

$$\dot{x_i} = F_i \cdot x_i^{\alpha} + V \cdot x_i^{\beta} \xi_i(t) - \phi(t) \cdot x_i,$$

where F_i captures the fitness of the taxon i, V corresponds to the noise amplitude and ξ_i(t) is a Gaussian random noise with zero mean $< \xi_i(t) > = 0$ and variance uncorrelated in time, $< \xi_i(t)\xi_i(t') > = \delta(t'-t)$, . The function $\phi(t)$ ensures the normalization at all times, $\sum x_i(t) = 1$, and corresponds to $\phi(t) = \sum F_i x_i^{\alpha} + \sum V x_i^{\beta} \xi_i(t)$. The temporal evolution of the probability that a taxon i has a relative abundance $x_i(t)$, P(x_i,t), is determined by the Fokker-Planck equation:

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x_i}[(F_i \cdot x_i^{\alpha} - \phi(t) \cdot x_i) \cdot P] + \frac{1}{2}\frac{\partial^2}{\partial x_i^2}(V^2 \cdot x_i^{2\beta} \cdot P).$$

The microbiota evolves towards a steady-state with a time-independent probability depending on the values of α, β, F_i and V. For $\alpha < 1$ (otherwise, systems are always unstable), the steady-state probability may be localized in a region around a preferred value or broadly distributed over a wide range, depending on whether the fitness F_i dominates or is overwhelmed by the noise amplitude V. The steady-state solution of the Fokker-Planck equation is given by:

$$P_0(x_i) = C_{ne}(\alpha, \beta, F_i, V) \cdot x_i^{-2\beta} \cdot \exp\Big[\frac{2F_i}{V^2}\frac{x_i^{1+\alpha-2\beta}}{1+\alpha-2\beta} - \frac{\phi_0}{V^2}\frac{x_i^{2-2\beta}}{1-\beta}\Big] \quad \text{if} \quad 2\beta \neq 1+\alpha,$$

$$P_0(x_i) = C_e(\alpha, \beta, F_i, V) \cdot x_i^{\frac{2F_i}{V^2}-2\beta} \cdot \exp\Big[\frac{\phi_0}{V^2}\frac{x_i^{2-2\beta}}{1-\beta}\Big] \quad \text{if} \quad 2\beta = 1+\alpha,$$

where $\phi_0 = (\sum_i F_i^{1/(1-\alpha)})^{1-\alpha}$ and C_{ne} and C_e are integrals that should be solved numerically for the parameters of interest. The ordered phase happens when the solution has a maximum in the physical interval ($0 < x_i < 1$). For larger V, the transition to a disordered phase happens when the maximum shifts to the unphysical region $x_i < 0$, which sets the phase transition region $V(\alpha, \beta, F_i)$. The phase transition region can be calculated analytically in particular cases:

$$\begin{aligned} F_i^2 &= 4\beta\phi_0 V^2 \quad \text{if} \quad \beta = \alpha \neq 1, \\ F_i &= \beta V^2 \quad \text{if} \quad 2\beta = 1 + \alpha, \end{aligned}$$

where the first case, simplifies to $F = 3V^2$ if $\beta = 0.75$ and the fitness of this taxon dominates in ϕ_0. In many physical systems (Brownian motion is the classical example (*68*)), the two terms of the Langevin equation are related. The *Fluctuation Dissipation Theorem* states a general relationship between the response to an external disturbance and the internal fluctuations of the system (*62*). The theorem can be used as the basic formula to derive the fitness from the analysis of fluctuations of the microbiota, assuming that it is in equilibrium (the ordered phase).

Standardization

In order to properly show all the studies under common axes, we decided to standardize the Taylor parameters using the group of healthy individuals for each study. With this approach, all the studies can be visualized in a shared plot with units of Taylor-parameters standard-deviation on their axes.

For a Taylor parameter, e.g. V, the estimate of the mean ($\widehat{V}$) for the healthy subpopulation,

composed of h individuals, is:

$$\widehat{V} = \frac{1}{W_1}\sum_{i=1}^{h} V_i\omega_i = \sum_{i=1}^{h} V_i\omega_i$$

as $W_1 = \sum_i^h \omega_i = 1$, since ω_i are normalized weights calculated as:

$$\omega_i = \frac{\frac{1}{\sigma^2_{V_i}}}{\sum_i^h \frac{1}{\sigma^2_{V_i}}}$$

being σ_{V_i} the estimation of the uncertainty in V_i obtained together with V_i from the X-weighted power-law fit described in Section , for healthy individuals.

Likewise, the estimation of the standard deviation for the healthy population ($\widehat{\sigma}_V$) is:

$$\widehat{\sigma}_V = \sqrt{\frac{1}{W_1 - \frac{W_2}{W_1}}\sum_{i=1}^{h}\left[\omega_i\left(V_i - \hat{V}\right)^2\right]}$$

being $W_2 = \sum_i^h \omega_i^2$, which finally yields to:

$$\widehat{\sigma}_V = \sqrt{\frac{1}{1 - \sum_i^h \omega_i^2}\sum_{i=1}^{h}\left[\omega_i\left(V_i - \hat{V}\right)^2\right]}$$

Selection and Methods

The bacteria and archaea taxonomic assignations were obtained by analysing 16S rRNA sequences, which were clustered into operational taxonomic units (OTUs) sharing 97 % sequence identity using QIIME (*69*). Shotgun metagenomic sequencing (SMS) data (*49*) were analysed and assigned at strain level by the Livermore Metagenomic Analysis Toolkit (LMAT) (*70*), according to their default quality threshold. Genus, with best balance between error assignment and number of taxa, was chosen as our reference taxonomic level. We have

verified that our conclusions are not significantly affected by selecting family or species as the reference taxonomic level (see Supplementary Figure 4).

Sample selection

We have chosen studies about relevant pathologies containing metagenomic sequencing time data series of bacterial populations from humans in different healthy and non-healthy states. We have selected only those individuals who had three or more time points of data available in databases. Caporaso et al. study (*46*) was selected as it has two healthy individuals measured over a very large timespan, with almost daily sampling. Faith et al. study (*48*) was selected due to the BMI differences between subjects, moreover some of them had diets which can be treated as system perturbations. We considered healthy only those individuals who had normal or overweight BMI. Smith et al., study (*49*) was selected for both the age of the patients and the rare disease. We only worked with the discordant twins, and considered healthy those who were not affected by kwashiorkor in each pair of patients. David et al., study (*50*) was selected for its differential diets. We considered as the healthy part the first time samples of each individual before the diet, and the rest of time points as perturbations. Dethlefsen and Relman work (*47*) was selected due to the interesting treatment of two antibiotic intakes of the same antibiotic by three different subjects. We considered as the healthy part only those times before any antibiotic treatment, and as perturbations the time of antibiotic intakes and the period after that. David et al. work (*51*) was selected due to the great longitudinal data that it provides plus its complete metadata and the interesting events that happened to both subjects (and infection and a travel). We considered as a healthy part the previous time points before each event. And finally, we also considered a study made in our group carried by Durban et al., (*12*) in which we considered as healthy subjects those who didn't suffer from irritable bowel disease, and as perturbation the patients who had this disease.

Metadata of each study is provided in Supplementary Tables S1 to S7. All used 16S rRNA gene sequencing except for the study of the discordant kwashiorkor twins (*49*), where both SMS and 16S rRNA data were used. In the latter case we selected to work with SMS data to show that our method is valid regardless of the source of taxonomic information. Each one of the datasets was treated as follows:

16rRNA sequences processing

Reads from the selected studies were first quality filtered using the FastX toolkit (*71*), allowing only those reads which had more than 25 of quality along the 75% of the complete sequence. 16S rRNA reads were then clustered at 97% nucleotide sequence identity (97% ID) into operational taxonomic units (OTUs) using QIIME package software (*69*) (version 1.8) We followed open reference OTU picking workflow in all cases. The clustering method used was uclust, and the OTUs were matched against Silva database (*72*) (version 111, July 2012) and were assigned to taxonomy with an uclust-based consensus taxonomy assigner. The parameters used in this step were: similarity 0.97, prefilter percent id 0.6, max accepts 20, max rejects 500.

Metagenomic sequences processing

Shotgun metagenomic (and 16S too) sequences were analyzed with LMAT (Livermore Metagenomics Analysis Toolkit) software package (*70*) (version 1.2.4, with Feb'15 release of data base *LMAT-Grand*). LMAT was run using a Bull shared-memory node belonging to the team's HPC (high performance computing) cluster. It is equipped with 32 cores (64 threads available using Intel Hyper-threading technology) as it has two Haswell-based Xeons (22 nm technology), the E5-2698v3@2.3 GHz, sharing half a tebibyte (0.5 TiB, that is, 512 gibibytes) of DRAM memory. This node is also provided with a PCIe SSD card as NVRAM, the Micron

P420m HHHL, with 1.4 TB, and 750000 reading IOPS, 4 KB, achieving 3.3 GB/s. The computing node was supplied with a RAID-0 (striping) scratch disk area. We used the "Grand" database (*73*), release Feb'15, provided by the LMAT team, where "Grand" refers to a huge database that contains k-mers from all viral, prokaryote, fungal and protist genomes present in the NCBI database, plus Human reference genome (hg19), plus GenBank Human, plus the 1000 Human Genomes Project (HGP) (this represent about 31.75 billion k-mers occupying 457.62 GB) (*73*). Previously to any calculation, the full database was loaded in the NVRAM. With this configuration the observed LMAT sustained sequence classification rate was 20 kpb/s/core. Finally, it is worth mentioning that a complete set of Python scripts have been developed as back-end and front-end of the LMAT pipeline in order to manage the added complexity of time series analysis.

Taxa level robustness

We selected genus as taxonomic level for the subsequent steps of our work. In order to ensure that, between adjacent taxonomic levels, there were not crucial differences which could still be of relevance after standardization (see last subsection of Material and Methods), we tested two different data sets. In the former, the antibiotics study (*47*) with 16S data, we tested the differences between genus and family levels. The latter dataset tested was the kwashiorkor discordant twins study (*49*) for both genus and species taxonomic levels. The Supplementary Figures 4 (overview) and 5 (detail) plot the comparison between studies (and so, 16S and SMS) and between adjacent taxonomic levels.

X-weighted power-law fit

When fitting the power-law of std vs. mean, we can take into account that every mean has uncertainty and estimate it for a sample size n by the SEM (*Standard Error of the Mean*). Here,

the uncertainties affect the independent variable, so the fit is not so trivial as a Y-weighted fit, where the uncertainties affect the dependent variable. A standard approach to do this fit is: a) invert your variables before applying the weights, b) then perform the weighted fit, and finally, c) revert the inversion. This method is deterministic, but the approximate solution worsens with smaller coefficients of determination. To overcome this limitation, we developed a stochastic method by using a bootstrapping-like strategy that avoids the inversion and is applicable regardless of the coefficient of determination.

The basic idea of bootstrapping is that inference about a population from sample data (sample → population) can be modeled by resampling the sample data and performing inference on (resample → sample) (*74*). To adapt this general idea to our problem, we resample the x-data array using its errors array. That is, for each replicate, a new x-data array is computed based on:

$$x_i^* = x_i + v_i$$

where v_i is a Gaussian random variable with mean $\mu_i = 0$ and standard deviation $\sigma_i = \mathrm{SEM}_i$, as defined previously. For each replicate a complete un-weighted power-law fit is performed, where to choose between fitting power laws ($y = Vx^{\beta}$) using linear regression on log-transformed (LLR) data versus non-linear regression (NLR) we mainly follow *General Guidelines for the Analysis of Biological Power Laws* (*75*). The parameters of the X-weighted fit are then estimated by averaging through all the replicate fits performed, and their errors are estimated by computing the standard deviation also for all the fits. At the end of each step, the relative error is calculated by comparing the fit parameters estimation in the last step with the previous one. Finally, both the coefficient of determination of the fit and the coefficient of correlation between the fit parameters are estimated by averaging.

Rank Stability Index

The Rank Stability Index (RSI) is shown as a percentage in a separate bar on the right of the rank matrix plot shown in Figures 4 and 5. The RSI is strictly 1 for an element whose range never changes over time, and is strictly 0 for an element whose rank oscillates between the extremes from time to time. So, RSI is calculated, per element, as 1 less the quotient of the number of true rank hops taken between the number of maximum possible rank hops, all powered to p:

$$\text{RSI} = \left(1 - \frac{\text{true rank hops}}{\text{possible rank hops}}\right)^p = \left(1 - \frac{D}{(N-1)(t-1)}\right)^p$$

where D is the total of rank hops taken by the studied element, N is the number of elements that have been ranked, and t is the number of time samples. The power index $p = 4$ is arbitrarily chosen to increase the resolution in the stable region.

Acknowledgments

Authors declare that there are no competing financial interests in relation to the work described here. We thereby express our acknowledgement to Bull/Atos and Micron Technology for providing us with the PCIe SSD card Micron P420m HHHL as a free-of-charge sample for high performance throughput database testing purposes.

Funding Information

This work was supported by grants to AM from the Spanish Ministry of Science and Competitivity (projects SAF2012-31187, SAF2013-49788-EXP, SAF2015-65878-R), Carlos III Institute of Health (projects PIE14/00045 and AC15/00022), Generalitat Valenciana (project

PrometeoII/2014/065) and co-financed by FEDER, and grants to CPG from the Generalitat Valenciana Prometeo Grants II/2014/050, II/2014/065, 419 by the Spanish Grants FPA2011-29678, BFU2012-39816-C02-01 of MINECO and by PITN-GA-420 2011-289442-INVISIBLES. From the Ministry of Economy and Competitiveness (grants FPI BES-2012-052900 and FPI BES-2013-062767).

References

1. **Rosenberg E, Zilber-Rosenberg I.** 2016. Microbes Drive Evolution of Animals and Plants: the Hologenome Concept. MBio **7**:e01395–15–.

2. **Bordenstein SR, Theis KR.** 2015. Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes. PLOS Biol **13**:e1002226.

3. **Moran NA, Sloan DB.** 2015. The Hologenome Concept: Helpful or Hollow? PLoS Biol **13**:1–10.

4. **Swann JR, Want EJ, Geier FM, Spagou K, Wilson ID, Sidaway JE, Nicholson JK, Holmes E.** 2011. Systemic gut microbial modulation of bile acid metabolism in host tissue compartments. Proc Natl Acad Sci **108**:4523–4530.

5. **Spencer MD, Hamp TJ, Reid RW, Fischer LM, Zeisel SH, Fodor AA.** 2011. Association between composition of the human gastrointestinal microbiome and development of fatty liver with choline deficiency. Gastroenterology **140**:976–986.

6. **Samuel BS, Shaito A, Motoike T, Rey FE, Backhed F, Manchester JK, Hammer RE, Williams SC, Crowley J, Yanagisawa M, Gordon JI**. 2008. Effects of the gut microbiota on host adiposity are modulated by the short-chain fatty-acid binding G protein-coupled receptor, Gpr41. Proc Natl Acad Sci **105**:16767–16772.

7. **Smith PM, Howitt MR, Panikov N, Michaud M, Gallini CA, Bohlooly-Y M, Glickman JN, Garrett WS.** 2013. The Microbial Metabolites, Short-Chain Fatty Acids, Regulate Colonic Treg Cell Homeostasis. Science (80-) **341**:569–573.

8. **Kimura I, Ozawa K, Inoue D, Imamura T, Kimura K, Maeda T, Terasawa K, Kashihara D, Hirano K, Tani T, Takahashi T, Miyauchi S, Shioi G, Inoue H, Tsujimoto G.** 2013. The gut microbiota suppresses insulin-mediated fat accumulation via the short-chain fatty acid receptor GPR43. Nat Commun **4**:1829.

9. **Maslowski KM, Vieira AT, Ng A, Kranich J, Sierro F, Di Yu, Schilter HC, Rolph MS, Mackay F, Artis D, Xavier RJ, Teixeira MM, Mackay CR.** 2009. Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. Nature **461**:1282–1286.

10. **Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D, Peng Y, Zhang D, Jie Z, Wu W, Qin Y, Xue W, Li J, Han L, Lu D, Wu P, Dai Y, Sun X, Li Z, Tang A, Zhong S, Li X, Chen W, Xu R, Wang M, Feng Q, Gong M, Yu J, Zhang Y, Zhang M, Hansen T, Sanchez G, Raes J, Falony G, Okuda S, Almeida M, LeChatelier E, Renault P, Pons N, Batto J-M, Zhang Z, Chen H, Yang R, Zheng W, Li S, Yang H, Wang J, Ehrlich SD, Nielsen R, Pedersen O, Kristiansen K, Wang J.** 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature **490**:55–60.

11. **Brown JM, Hazen SL.** 2015. The Gut Microbial Endocrine Organ: Bacterially Derived Signals Driving Cardiometabolic Diseases. Annu Rev Med **66**:343–359.

12. **Durbán A, Abellán JJ, Jiménez-Hernández N, Artacho A, Garrigues V, Ortiz V, Ponce J, Latorre A, Moya A.** 2013. Instability of the faecal microbiota in diarrhoea-predominant irritable bowel syndrome. FEMS Microbiol Ecol **86**:581–589.

13. **Gevers D, Kugathasan S, Denson LA, Vázquez-Baeza Y, Van Treuren W, Ren B, Schwager E, Knights D, Song SJ, Yassour M, Morgan XC, Kostic AD, Luo C,**

González A, McDonald D, Haberman Y, Walters T, Baker S, Rosh J, Stephens M, Heyman M, Markowitz J, Baldassano R, Griffiths A, Sylvester F, Mack D, Kim S, Crandall W, Hyams J, Huttenhower C, Knight R, Xavier RJ. 2014. The treatment-naive microbiome in new-onset Crohn's disease. Cell Host Microbe **15**:382–392.

14. **Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau L, Griffi NW, Lombard V, Henrissat B, Bain JR, Michael J, Ilkayeva O, Semenkovich CF, Funai K, Hayashi DK, Lyle J, Martini MC, Ursell LK, Clemente JC, Treuren W Van, William A, Knight R, Newgard CB, Heath AC, Gordon JI, Kau AL, Griffin NW, Muehlbauer MJ.** 2013. Gut Microbiota from Twins Discordant for Obesity Modulate Metabolism in Mice Gut Microbiota from Twins Metabolism in Mice. Science **341**:1241214.

15. **Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI.** 2009. LETTERS A core gut microbiome in obese and lean twins. Nature **457**:480–484.

16. **Subramanian S, Huq S, Yatsunenko T, Haque R, Mahfuz M, Alam MA, Benezra A, DeStefano J, Meier MF, Muegge BD, Barratt MJ, VanArendonk LG, Zhang Q, Province MA, Petri WA, Ahmed T, Gordon JI.** 2014. Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature **510**:417–21.

17. **Moya A, Ferrer M.** 2016. Functional Redundancy-Induced Stability of Gut Microbiota Subjected to Disturbance. Trends Microbiol **24**:402–413.

18. **Cryan JF, Dinan TG**. 2012. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. Nat Rev Neurosci **13**:701–712

19. **Xu R, Wang Q**. 2016. Towards understanding brain-gut-microbiome connections in Alzheimer's disease. BMC Systems Biology **10**:277-285

20. **Giloteaux L, Goodrich JK, Walters WA, Levine SM, Ley RE, Hanson MR**. 2016. Reduced diversity and altered composition of the gut microbiome in individuals with myalgic encephalomyelitis/chronic fatigue syndrome. Microbiome. **4**:30

21. **Marchesi JR, Adams DH, Fava F, Hermes GD a, Hirschfield GM, Hold G, Quraishi MN, Kinross J, Smidt H, Tuohy KM, Thomas L V, Zoetendal EG, Hart A.** 2015. The gut microbiota and host health: a new clinical frontier. Gut 1–10.

22. **Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, Kurilshikov A, Bonder MJ, Valles-Colomer M, Vandeputte D, Tito RY, Chaffron S, Rymenans L, Verspecht C, De Sutter L, Lima-Mendez G, Dhoe K, Jonckheere K, Homola D, Garcia R, Tigchelaar EF, Eeckhaudt L, Fu J, Henckaerts L, Zhernakova A, Wijmenga C, Raes J.** 2016. Population-level analysis of gut microbiome variation. Science (80-) **352**:560–564.

23. **Zhernakova A, Kurilshikov A, Bonder MJ, Tigchelaar EF, Schirmer M, Vatanen T, Mujagic Z, Vila AV, Falony G, Vieira-Silva S, Wang J, Imhann F, Brandsma E, Jankipersadsing SA, Joossens M, Cenit MC, Deelen P, Swertz MA, Weersma RK, Feskens EJM, Netea MG, Gevers D, Jonkers D, Franke L, Aulchenko YS, Huttenhower C, Raes J, Hofker MH, Xavier RJ, Wijmenga C, Fu J.** 2016. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. Science (80-) **352**:565–569.

24. **Amato KR** 2016. Incorporating the Gut Microbiota Into Models of Human and Non-Human Primate Ecology and Evolution. Yearbook Of Physical Anthropology **157**:S196–S215.

25. **Wu H, Tremaroli V, Bäckhed F.** 2015. Linking Microbiota to Human Diseases: A Systems Biology Perspective. Trends Endocrinol Metab **26**:758–770.

26. **Noecker C, Eng A, Srinivasan S, Theriot CM, Young VB, Jansson JK, Fredricks**

DN, Borenstein E. 2016. Metabolic Model-Based Integration of Microbiome Taxonomic and Metabolomic Profiles Elucidates Mechanistic Links between Ecological and Metabolic Variation. mSystems **1**:e00013–15.

27. **Greenblum S, Turnbaugh PJ, Borenstein E.** 2012. Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. Proc Natl Acad Sci **109**:594–599.

28. **Bashan A, Gibson TE, Friedman J, Carey VJ, Weiss ST, Hohmann EL, Liu Y-Y.** 2016. Universality of human microbial dynamics. Nature **534**:259–262.

29. **Taylor, L.R.** 1961. Aggregation, Variance and the mean. Nature **189,** 732-35.

30. **de Menezes MA, Barabási A-L.** 2004. Fluctuations in network dynamics. Phys Rev Lett **92**:1–4.

31. **Mantegna RN, Stanley HE.** 1995. Scaling behaviour in the dynamics of an economic index. Nature **376**:46–49.

32. **Eisler Z, Kertesz J, Yook SH, Barabasi AL.** 2005. Multiscaling and non-universality in fluctuations of driven complex systems. Europhys Lett **69**:664–670.

33. **Cohen JE, Xu M, Schuster WSF**. 2013. Stochastic multiplicative population growth predicts and interprets Taylor's power law of fluctuation scaling. Proc R Soc B Biol Sci **280**:20122955.

34. **Reed DH, Hobbs GR.** 2004. The relationship between population size and temporal variability in population size. Anim Conserv **7**:1–8.

35. **Anderson RM, Gordon DM, Crawley MJ, Hassell MP.** 1982. Variability in the abundance of animal and plant species. Nature **18**: 245–248

36. **Živković J, Tadić B, Wick N, Thurner S.** 2006. Statistical indicators of collective behavior and functional clusters in gene networks of yeast. Eur Phys J B **50**:255–258.

37. **Kendal WS.** 2003. An Exponential Dispersion Model for the Distribution of Human Single Nucleotide Polymorphisms. Mol Biol Evol **20**:579–590.

38. **Zhang Z, Geng J, Tang X, Fan H, Xu J, Wen X, Ma ZS, Shi P.** 2014. Spatial heterogeneity and co-occurrence patterns of human mucosal-associated intestinal microbiota. ISME J **8**:881–93.

39. **Kaltz O, Escobar-Paramo P, Hochberg M, Cohen** JE. 2012. Bacterial microcosmos obey Taylor's law: Effects of abiotic and biotic stress and genetics on mean and variance of population density. Ecol Process **1**:5.

40. **Ramsayer J, Fellous S, Cohen JE, Hochberg ME.** 2012. Taylor's Law holds in experimental bacterial populations but competition does not influence the slope. Biol Lett **8**:316–319.

41. **Pérez-Cobas AE, Artacho A, Ott SJ, Moya A, Gosalbes MJ, Latorre A.** 2014. Structural and functional changes in the gut microbiota associated to Clostridium difficile infection. Front Microbiol **5**:1–15.

42. **Ding T, Schloss PD.** 2014. Dynamics and associations of microbial community types across the human body. Nature **509**:357–360.

43. **Gajer P, Brotman RM, Bai G, Sakamoto J, Schütte UME, Zhong X, Koenig SSK, Fu L, Ma ZS, Zhou X, Abdo Z, Forney LJ, Ravel J.** 2012. Temporal dynamics of the human vaginal microbiota. Sci Transl Med **4**:132ra52.

44. **Blumm N, Ghoshal G, Forró Z, Schich M, Bianconi G, Bouchaud J-P, Barabási A-L.** 2012. Dynamics of Ranking Processes in Complex Systems. Phys Rev Lett **109**:128701.

45. **Eisler Z, Bartos I, Kertész J.** 2008. Fluctuation scaling in complex systems: Taylor's law and beyond1. Adv Phys **57**:89–142.

46. **Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, Knights D, Gajer P, Ravel J, Fierer N, Gordon JI, Knight R.** 2011. Moving pictures of the human microbiome. Genome Biol **12**:R50.

47. **Dethlefsen L, Relman DA.** 2011. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. Proc Natl Acad Sci **108**:4554–61.

48. **Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, Clemente JC, Knight R, Heath AC, Leibel RL, Rosenbaum M, Gordon JI.** 2013. The long-term stability of the human gut microbiota. Science **341**:1237439.

49. **Smith MI, Yatsunenko T, Manary MJ, Trehan I, Mkakosya R, Cheng J, Kau AL, Rich SS, Concannon P, Mychaleckyj JC, Liu J, Houpt E, Li J V, Holmes E, Nicholson J, Knights D, Ursell LK, Knight R, Gordon JI.** 2013. Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. Science **339**:548–54.

50. **David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling A V, Devlin AS, Varma Y, Fischbach MA, Biddinger SB, Dutton RJ, Turnbaugh PJ.** 2014. Diet rapidly and reproducibly alters the human gut microbiome. Nature **505**:559–63.

51. **David LA, Materna AC, Friedman J, Campos-Baptista MI, Blackburn MC, Perrotta A, Erdman SE, Alm EJ.** 2014. Host lifestyle affects human microbiota on daily timescales. Genome Biol **15**:R89.

52. **Jørgensen B, Martinez JR, Tsao M.** 1994. Asymptotic behaviour of the variance function. Scand J Stat **21**:223–243.

53. **Fronczak A, Fronczak P.** 2010. Origins of Taylor's power law for fluctuation scaling in complex systems. Phys Rev E **81**:066112.

54. **Kendal, W.S., Jorgensen,B.** Taylor's power law and fluctuation scaling explained by a central-limit-like convergence. Phys. Rev. E **83** 066115.

55. **Kendal WS, Jørgensen B.** 2011. Tweedie convergence: A mathematical basis for Taylor's power law, 1/f noise, and multifractality. Phys Rev E **84**:066120.

56. **Kilpatrick a M, Ives a R.** 2003. Species interactions can explain Taylor's power law for ecological time series. Nature **422**:65–68.

57. **Ballantyne IV F, J. Kerkhoff A.** 2007. The observed range for temporal mean-variance scaling exponents can be explained by reproductive correlation. Oikos **116**:174–180.

58. **Stein RR, Bucci V, Toussaint NC, Buffie CG, Rätsch G, Pamer EG, Sander C, Xavier JB.** 2013. Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota. PLoS Comput Biol **9**:e1003388.

59. **Fisher CK, Mehta P.** 2014. Identifying keystone species in the human gut microbiome from metagenomic timeseries using sparse linear regression. PLoS One **9**:e102451.

60. **Bucci V, Tzen B, Li N, Simmons M, Tanoue T, Bogart E, Deng L, Yeliseyev V, Delaney ML, Liu Q, Olle B, Stein RR, Honda K, Bry L, Gerber GK.** 2016. MDSINE: Microbial Dynamical Systems INference Engine for microbiome time-series analyses. Genome Biol **17**:121.

61. **Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE.** 2011. Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci **108**:4578–4585.

62. **Weber, J.** 1956. Fluctuation Dissipation Theorem. Phys. Rev. **101**:1620-6

63. **Ng SC, Lam EFC, Lam TTY, Chan Y, Law W, Tse PCH, Kamm M a, Sung JJY, Chan FKL, Wu JCY.** 2013. Effect of probiotic bacteria on the intestinal microbiota in irritable bowel syndrome. J Gastroenterol Hepatol **28**:1624–31.

64. **Carroll IM, Ringel-Kulka T, Siddle JP, Ringel Y.** 2012. Alterations in composition and diversity of the intestinal microbiota in patients with diarrhea-predominant irritable bowel syndrome. Neurogastroenterol Motil **24**:521-30

65. **Xu D, Gao J, Gillilland M, Wu X, Song I, Kao JY, Owyang C.** 2014. Rifaximin alters intestinal bacteria and prevents stress-induced gut inflammation and visceral hyperalgesia in rats. Gastroenterology **146**:484–496.e4.

66. **Rooks MG, Veiga P, Wardwell-scott LH, Tickle T, Segata N, Michaud M, Gallini CA, Beal C, van Hylckama-Vlieg JET, Ballal SA, Morgan XC, Glickman JN, Gevers D, Huttenhower C, Garrett WS, Hylckama-vlieg JET Van, Ballal SA, Morgan XC, Glickman JN, Gevers D, Huttenhower C, Garrett WS.** 2014. Gut microbiome composition and function in experimental colitis during active disease and treatment-induced remission. ISME J **8**:1403–17.

67. **Tikhonov M.** 2016. Community-level cohesion without cooperation. Elife **5**.

68. **Einstein A.** 1905. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Annalen der Physik, **322**:549-560.

69. **Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley G a, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone C a, Mcdonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters W a, Widmann J, Yatsunenko T, Zaneveld J, Knight R.** 2010. correspondence QIIME allows analysis of high- throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. Nat Publ Gr **7**:335–336.

70. **Ames SK, Hysom DA, Gardner SN, Lloyd GS, Gokhale MB, Allen JE.** 2013. Scalable metagenomic taxonomy classification using a reference genome database. Bioinfor-

matics **29**:2253-2260.

71. **Gordon, A, Hannon, GJ.** 2010. FASTX-Toolkit. FASTQ/A shortreads pre-processing tools v0.0.13. http://hannonlab.cshl.edu/fastx_toolkit/ (last accessed 26 Jul 2016).

72. **Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO**. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Acids Res. **41**:D590-D596

73. **Ames SK, Gardner SN, Marti JM, Slezak TR, Gokhale MB, Allen JE**. 2015. Using populations of human and microbial genomes for organism detection in metagenomes. Genome Res. **25**:1056-67.

74. **Wu, C.F.J.** 1986. Jackknife, bootstrap and other resampling methods in regression analysis. (with discussions) *The Annals of Statistics* **14**:1261-1350

75. **Xiao X, White EP, Hooten MB, Durham SL**. 2011. On the use of log-transformation vs. nonlinear regression for analyzing biological power laws. Ecology **92**:1887-1894

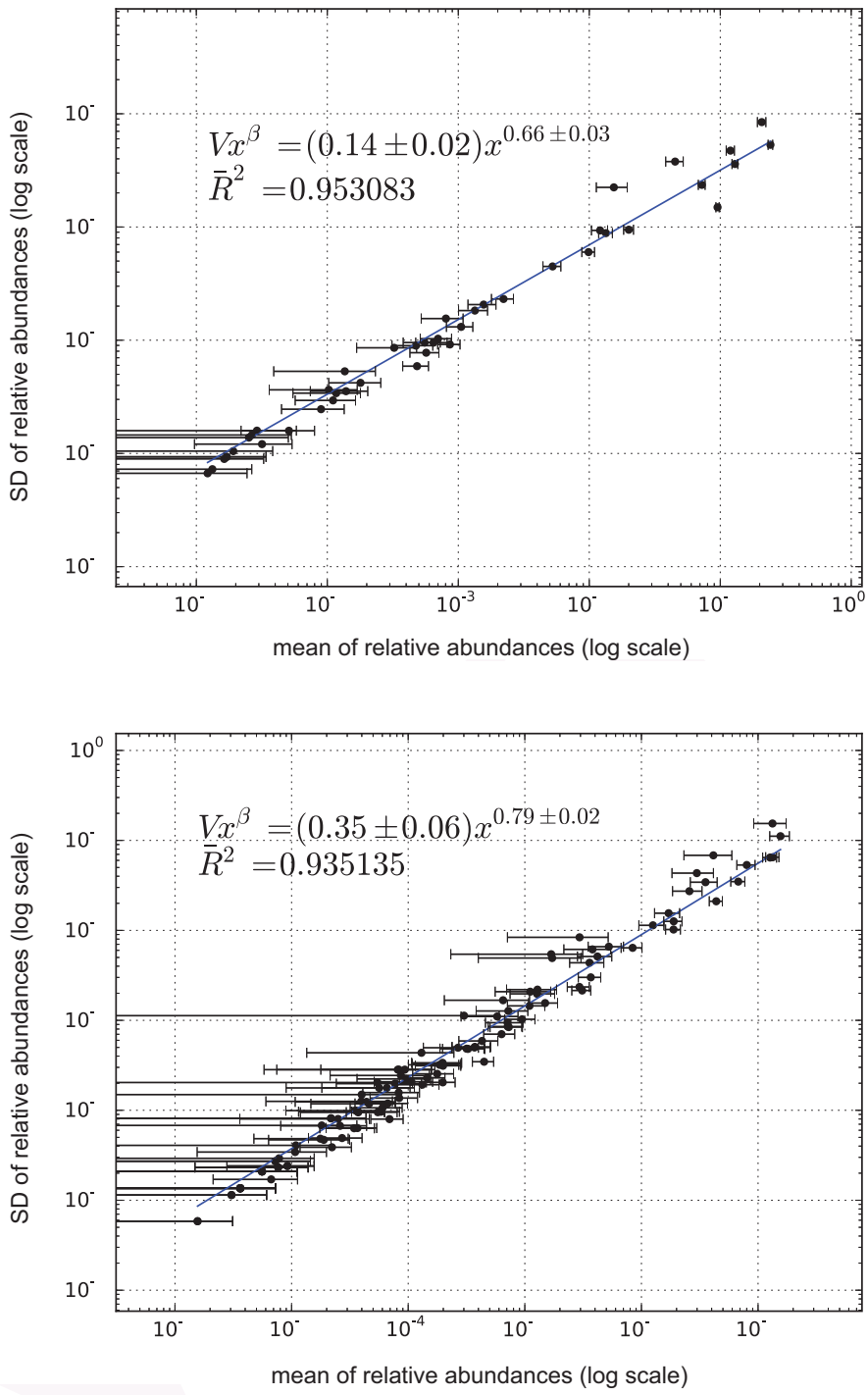


Figure 1. X-weighted power-law fits of the standard deviations versus the mean values for each bacterial genus monitored in time. We show the fit for samples from a healthy subject (top) and from a subject diagnosed with irritable bowel syndrome (bottom), studied in our lab (*12*). Taylor's power law seems to be ubiquitous, spanning to six orders of magnitude.

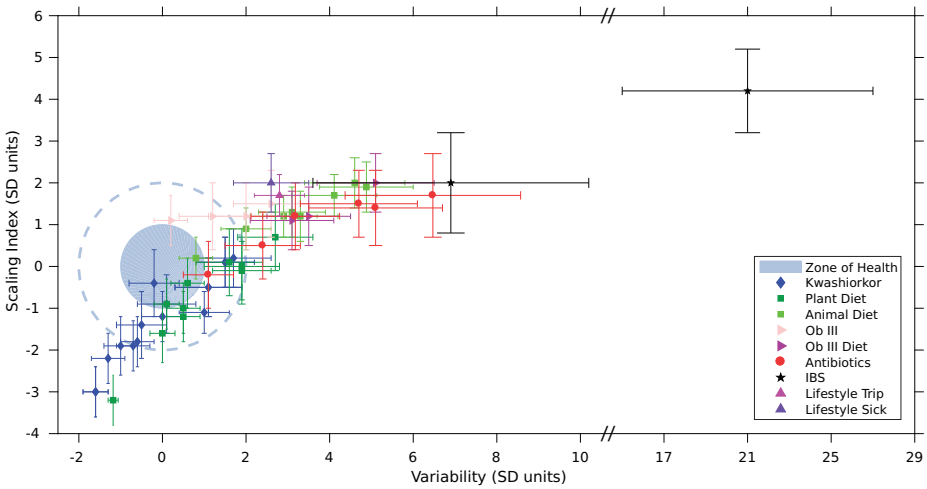


Figure 2. Taylor's law parameter space. We have compiled here all the data studied in this work. The coloured circle corresponds to 68% confidence level (CL) region of healthy individuals in the Taylor parameter space, while dashed line delimites the 98% CL region. Points with errors place gut microbiome in the Taylor space, for each individual whose microbiota was compromised. Note that the parameters have been standardized (standard deviation units) to the healthy group in each study for demonstrative and comparative purposes.

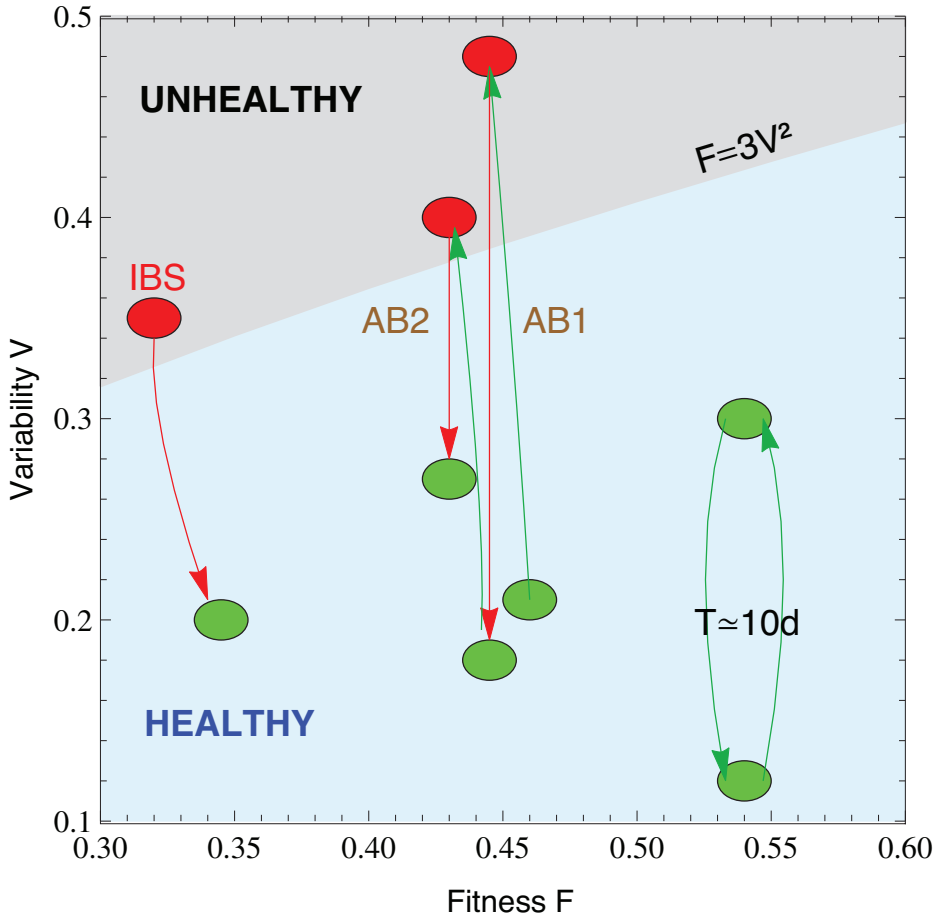


Figure 3. Microbiota states can be placed in the phase space F–V. The light blue shaded region corresponds to the stable phase, while the grey shaded region is the unstable phase (the phase transition line is calculated for $\alpha = \beta = 0.75$). We place healthy individuals (green) and individuals whose gut microbiota is threatened (antibiotics, IBS) in the phase space fitness–variability. Gut microbiota of healthy individuals over a long term span show a quasi–periodical variability (central period is ten days). We show that taking antibiotics (AB1 and AB2 correspond to first and second treatment respectively) induces a phase transition in the gut microbiota, which impacts its future changes. We also show an IBS–diagnosed patient transiting from the unstable to the stable phase.

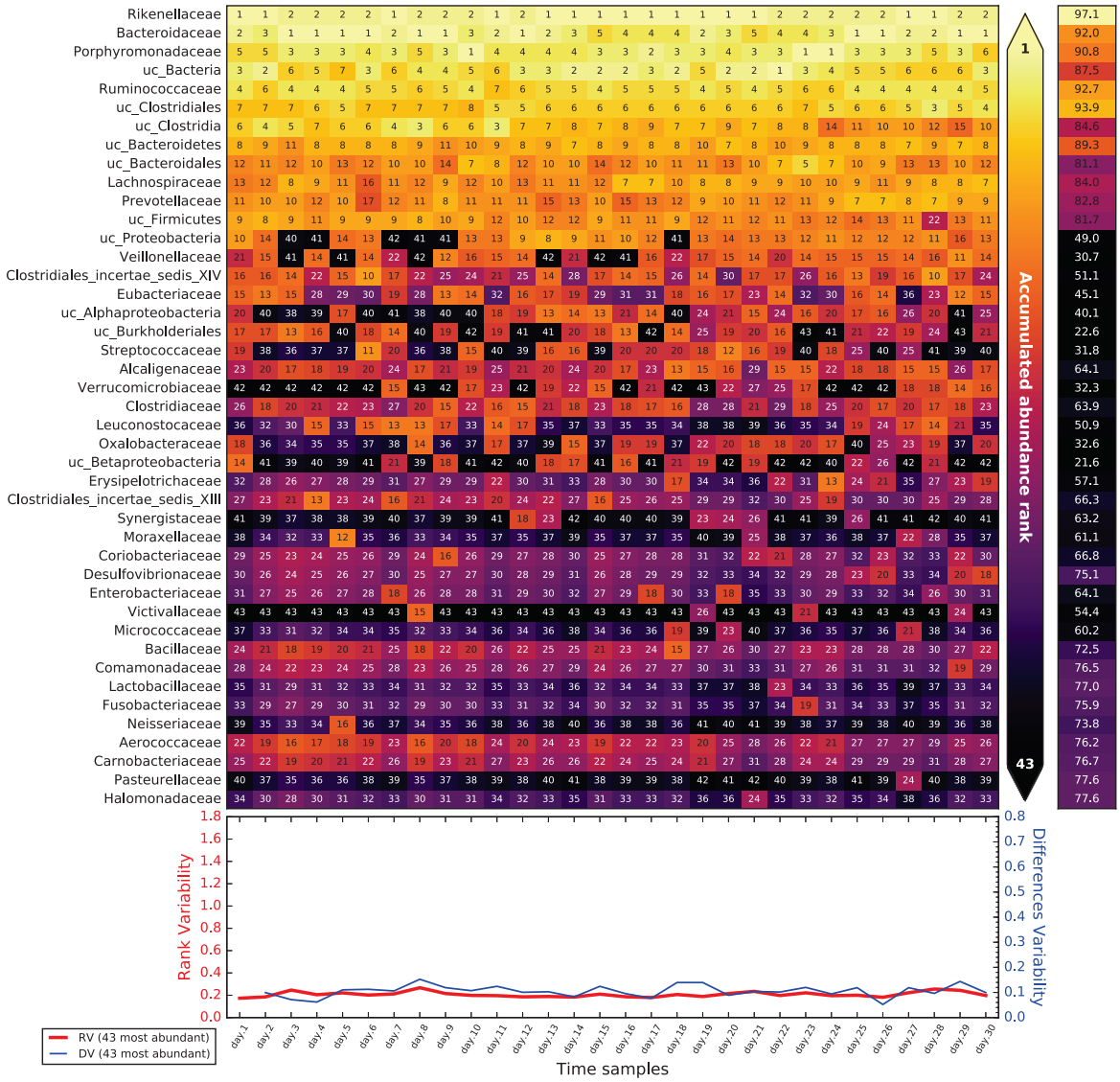


Figure 4. Rank variation throughout time for the most dominant elements (taxa) and their calculated Rank Stability Index (as shown in Material and Methods) for samples from a healthy subject studied in our lab (*12*).

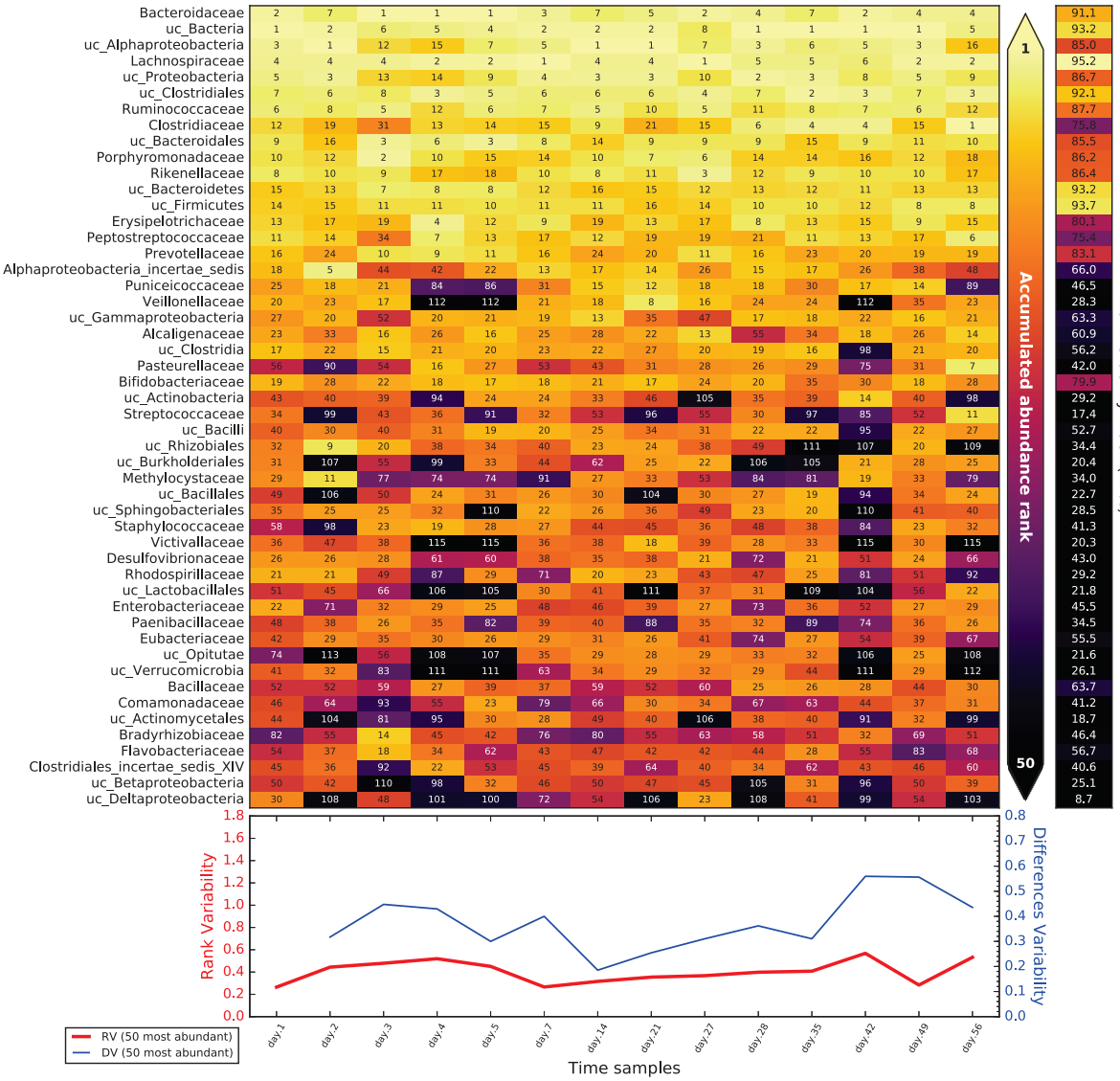


Figure 5. Rank variation throughout time for the most dominant elements (taxa) and their calculated Rank Stability Index for samples from a subject diagnosed with irritable bowel syndrome studied in our lab (*12*).

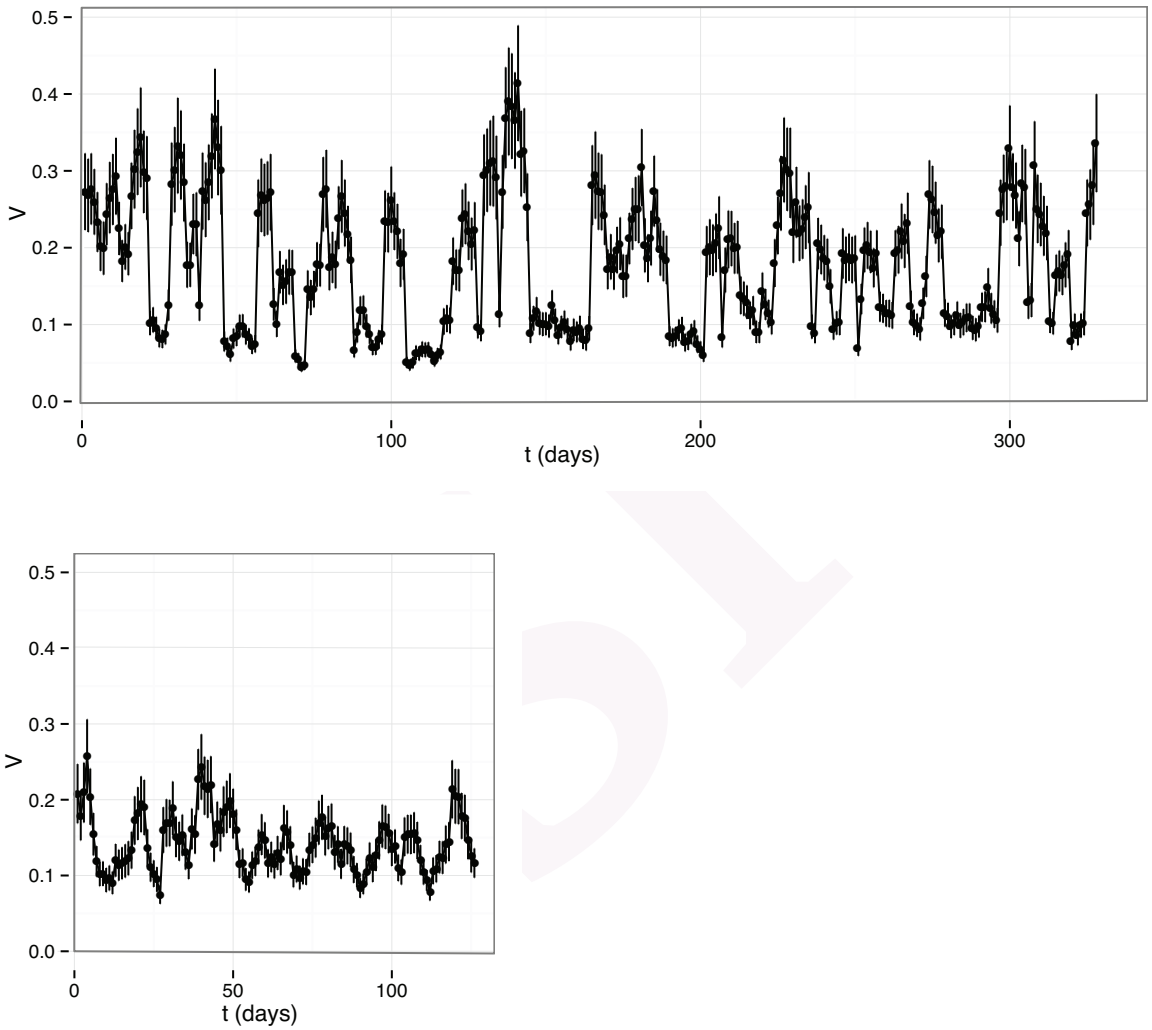


Figure 6. V as a function of time for the two individuals in the Caporaso's study (*46*): samples of gut microbiome of a male (upper plot) and a female (lower plot).

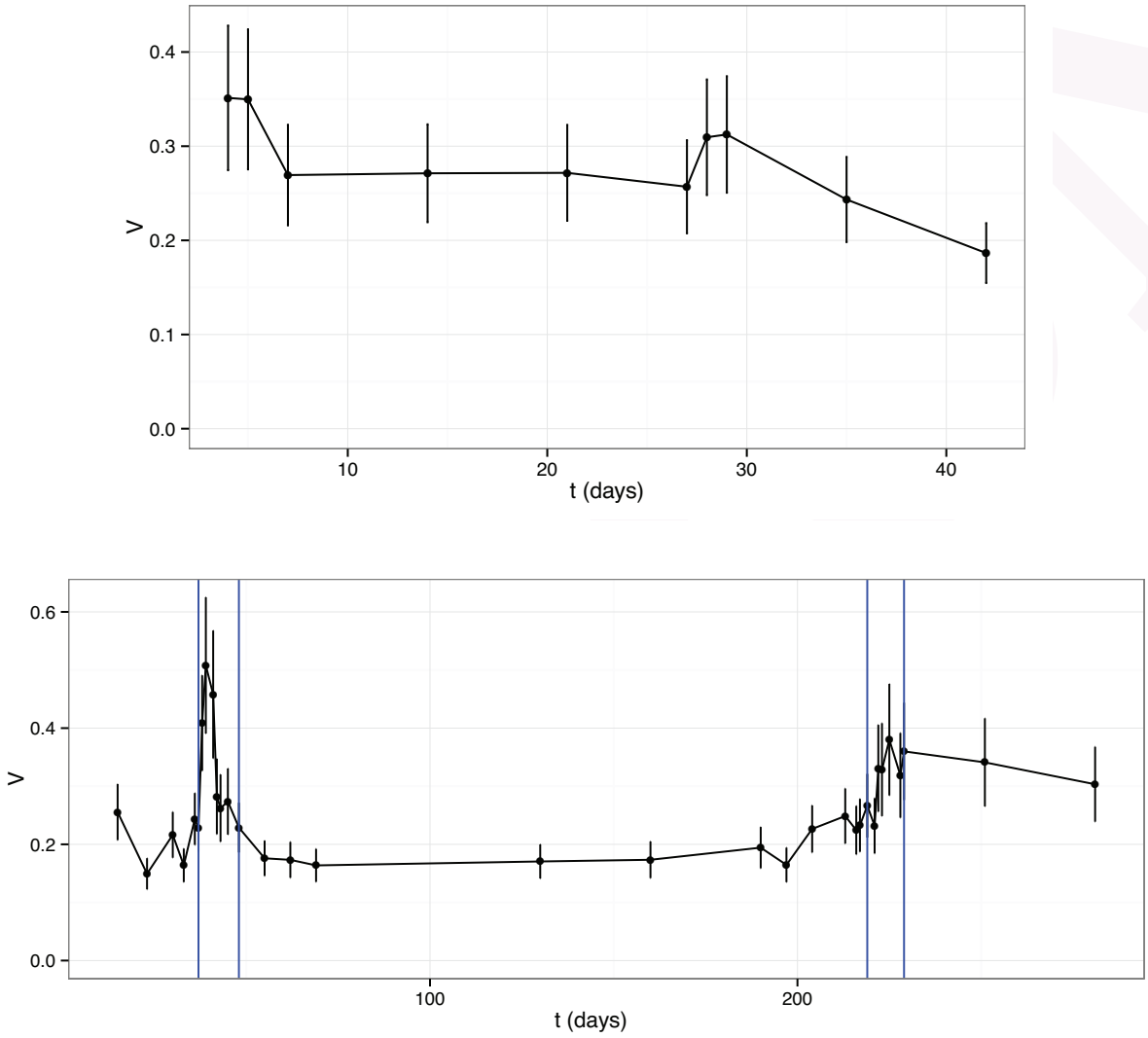


Figure 7. *V* as a function of time for patient P2 of the IBS study (*12*) (upper plot) and patient D in the antibiotics study (*47*) (lower plot). The blue vertical lines in the lower plot are showing the periods of antibiotic treatment.

Supplementary Table S1. Taylor parameters. Individuals with either animal-based (A) or plant-based (P) diets (*50*). Previous to diet, the population sampled is described by $\bar{V} = 0.09 \pm 0.05, \bar{\beta} = 0.77 \pm 0.04$.

Supplementary Table S2. Taylor parameters for individuals taking antibiotics (*47*). Prior to antibiotics intake, the population sampled is described by $\bar{V} = 0.12 \pm 0.05, \bar{\beta} = 0.75 \pm 0.04$.

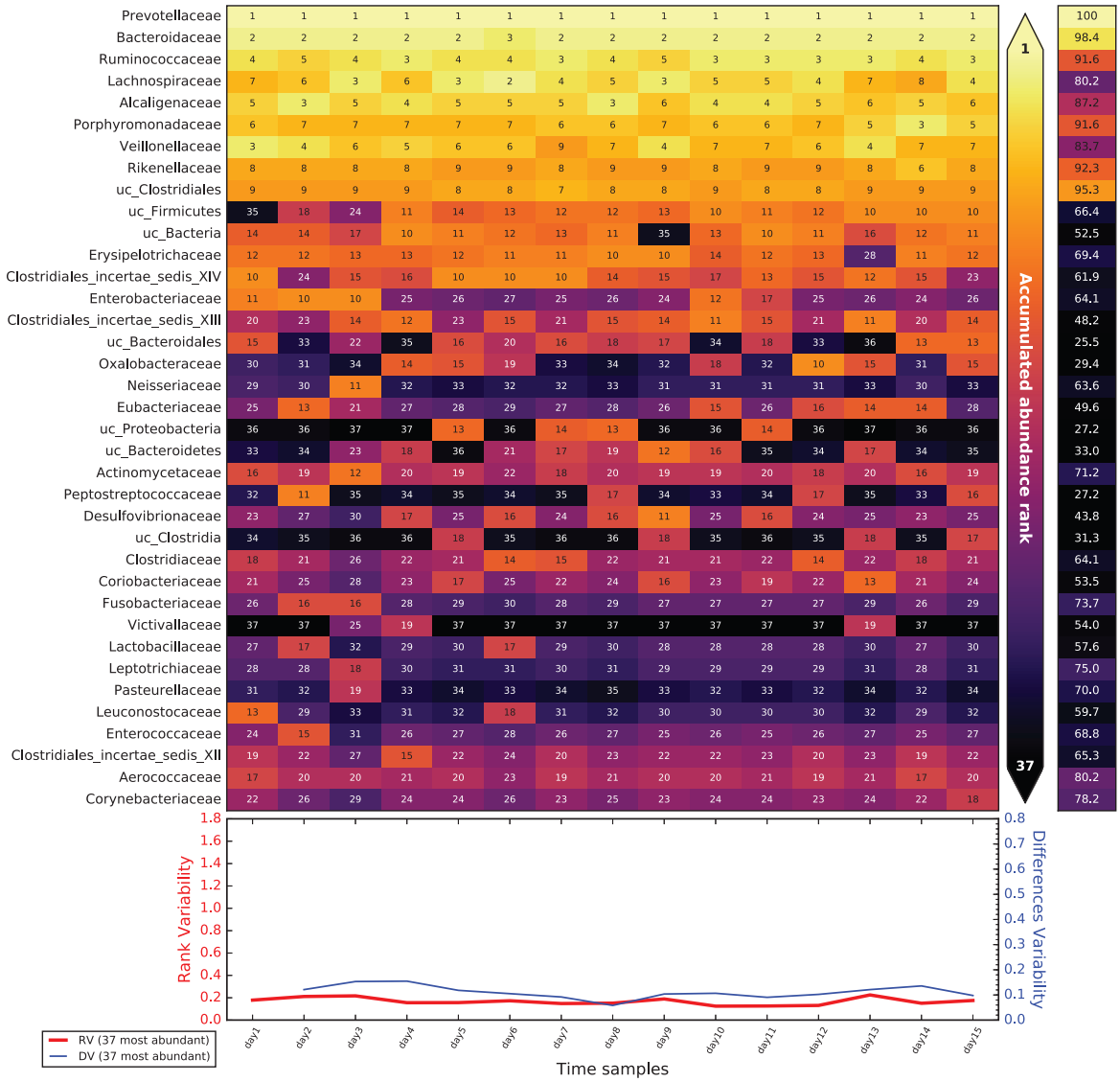
Supplementary Table S3. Taylor parameters for persons diagnosed with irritable bowel syndrome (IBS) (*12*). Healthy individuals sampled in this study are characterized by $\bar{V} = 0.135 \pm 0.010, \bar{\beta} = 0.692 \pm 0.024$.

Supplementary Table S4. Taylor parameters for the healthy subject of the discordant twins (*49*). This table continues in Supplementary Table S5. The population of healthy twins is characterized by $\bar{V} = 0.25 \pm 0.10, \bar{\beta} = 0.863 \pm 0.028$.

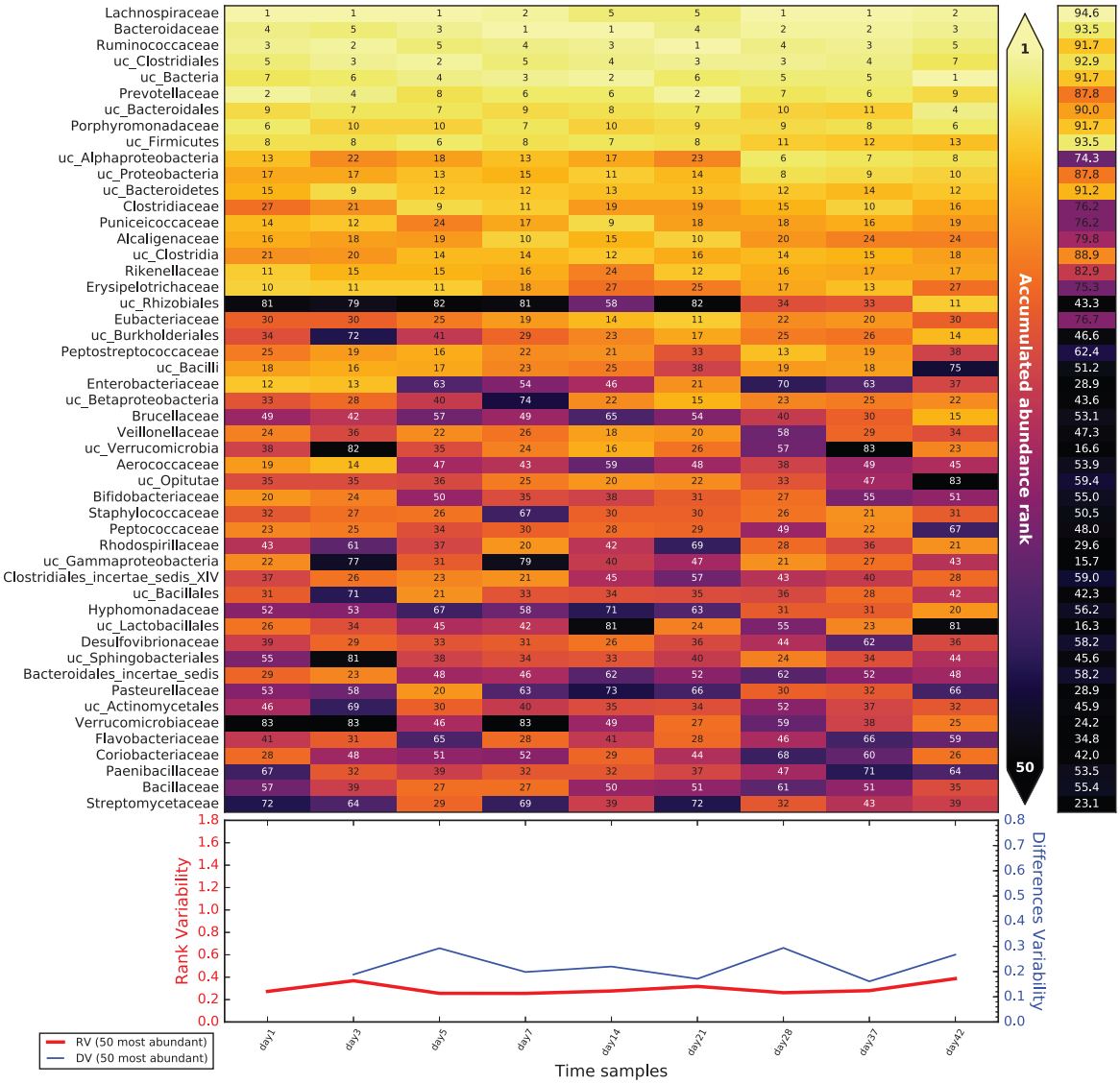
Supplementary Table S5. Taylor parameters for the kwashiorkor part of the discordant twins (*49*). This is a continuation of Supplementary Table S4. The population of healthy twins is characterized by $\bar{V} = 0.25 \pm 0.10, \bar{\beta} = 0.863 \pm 0.028$.

Supplementary Table S6. Taylor parameters for individuals with different degrees of overweight and obesity (*48*). Healthy people in this study, whom were not obese, are characterized by $\bar{V} = 0.19 \pm 0.06, \bar{\beta} = 0.806 \pm 0.034$.

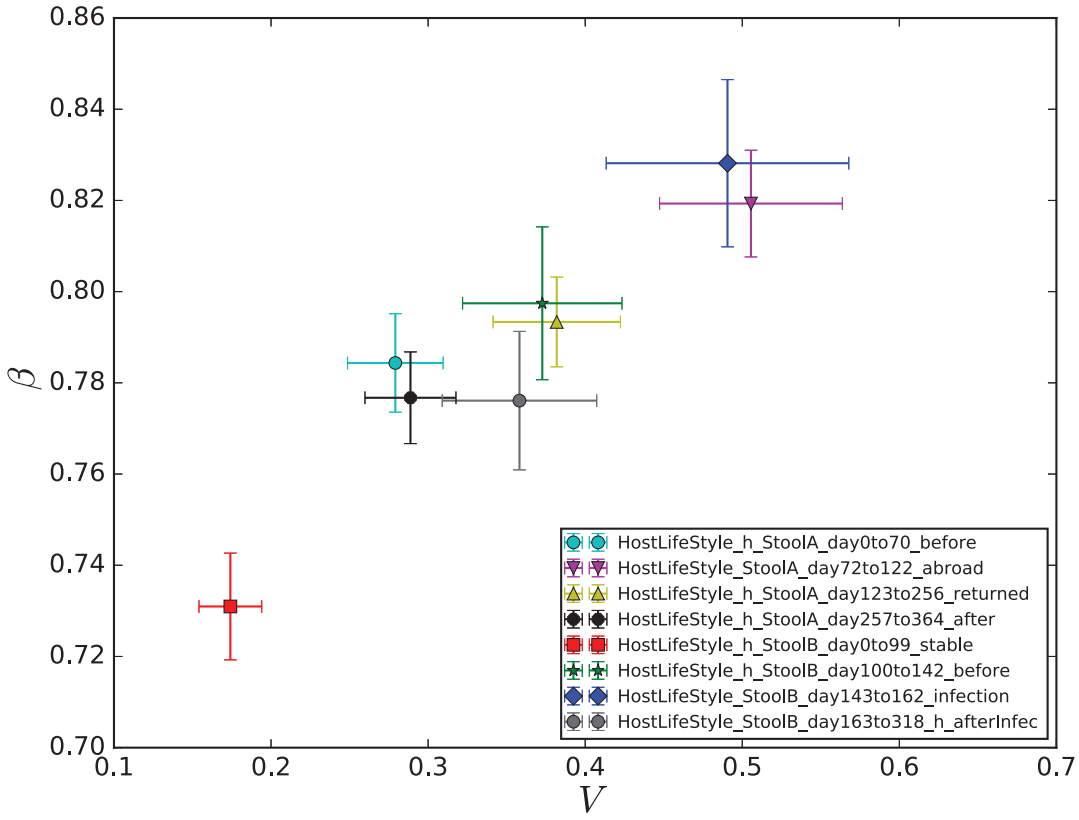
Supplementary Table S7. Taylor parameters for special intervals concerning gut microbiota in the host lifestyle study (*51*). The healthy and quotidian periods are characterized by $\bar{V} = 0.25 \pm 0.09, \bar{\beta} = 0.777 \pm 0.025$.



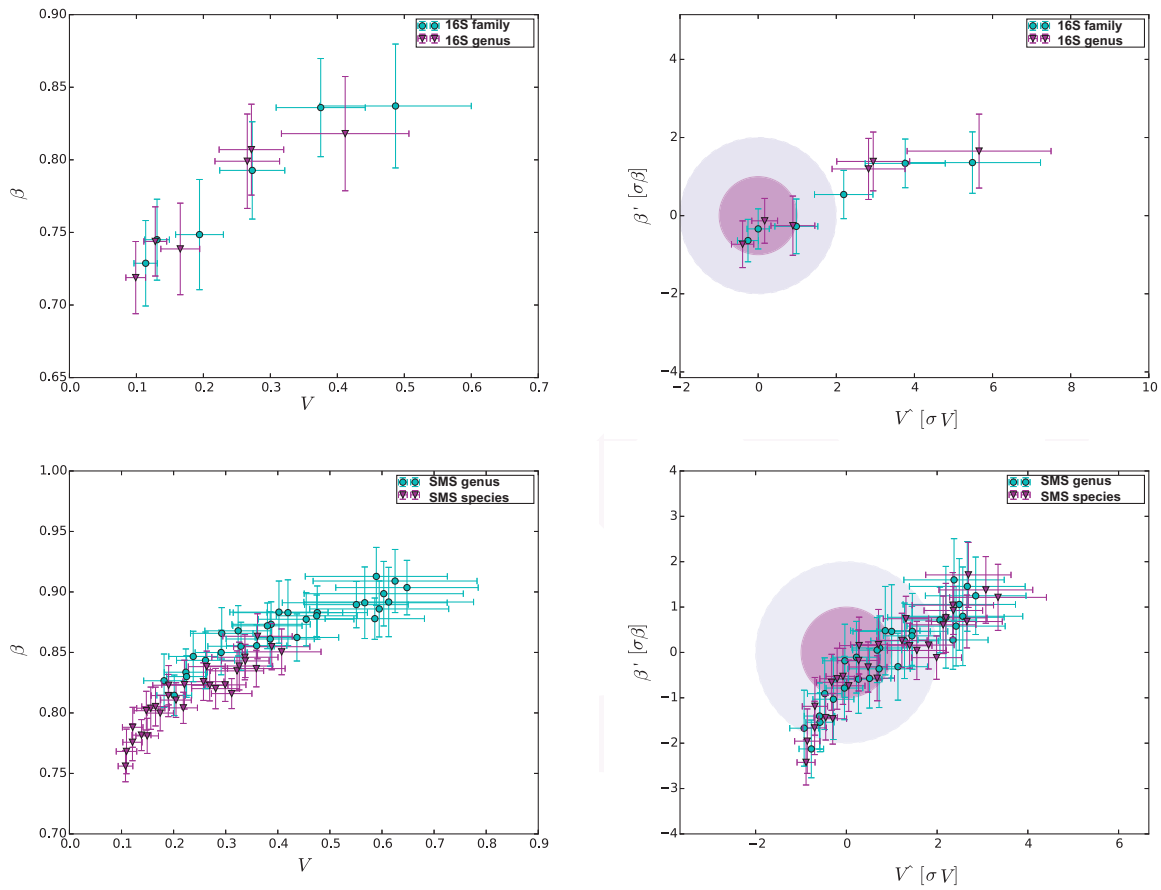
Supplementary Figure 1. Rank variation throughout time for the most dominant elements (taxa) and their calculated RSI for samples from another healthy subject studied in our lab (*12*).



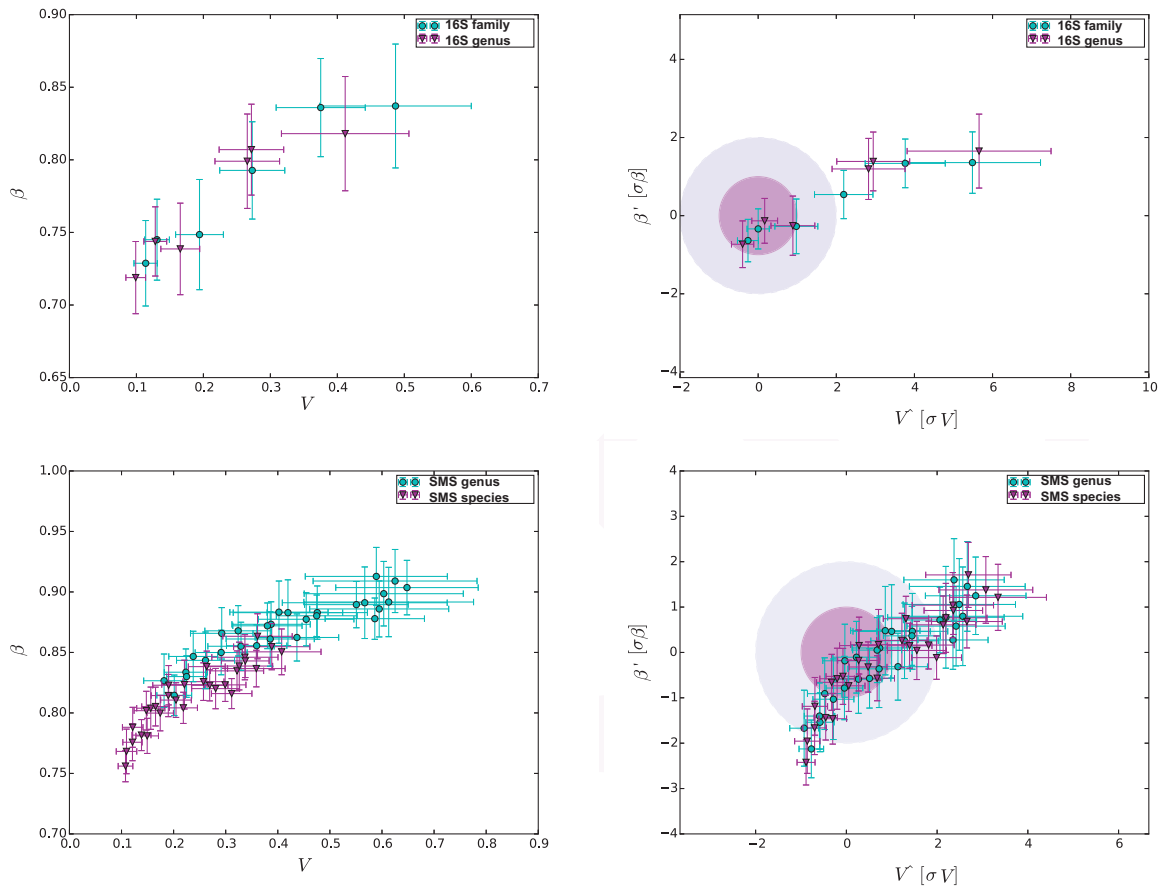
Supplementary Figure 2. Rank variation throughout time for the most dominant elements (taxa) and their calculated RSI for samples from a subject diagnosed with light IBS and studied in our lab (*12*).



Supplementary Figure 3. Taylor's law parameter space for intervals concerning gut microbiota in the host lifestyle study (*51*).



Supplementary Figure 4. Overview of comparison of different approaches based on adjacent taxonomic levels using plots in the Taylor-parameters space. The former row of subfigures is for 16S, where levels are family (blue circles) vs. genus (purple triangles). Whereas the latter row of subfigures is for SMS, where levels are genus (blue circles) vs. species (purple triangles). The left column shows the raw results and the right column plots the standardized results (see Standardization under Material and Methods).



Supplementary Figure 5. Detail of comparison of different approaches based on adjacent taxonomic levels using plots of X-weighted power-law fits (see Material and Methods). The former row of subfigures shows examples for 16S, whereas the latter row of subfigures plots examples for SMS. The left column shows results for the superior taxonomic level (family for 16S, genus for SMS), while the right column shows results for the inferior level (genus for 16S, specie for SMS).