



# Spatial heterogeneity analysis of the human virome with Taylor's power law



Zhanshan (Sam) Ma <sup>a,b,\*</sup>

<sup>a</sup> Computational Biology and Medical Ecology Lab, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

<sup>b</sup> Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

## ARTICLE INFO

### Article history:

Received 13 November 2020  
Received in revised form 27 April 2021  
Accepted 27 April 2021  
Available online 30 April 2021

### Keywords:

Spatial Heterogeneity  
Human Virome  
Human Microbiome  
Taylor's power law extensions (TPLE)  
Community Spatial Heterogeneity  
Mixed-Species Population Aggregation

## ABSTRACT

Spatial heterogeneity is a fundamental characteristic of organisms from viruses to humans. Measuring heterogeneity is challenging, especially for naked-eye invisible viruses, but of obvious importance. For example, spatial heterogeneity of virus distribution may strongly influence infection spreading and outbreaks in the case of pathogenic viruses; the spatial distribution (*i.e.*, the inter-subject heterogeneity) of commensal viruses within/on our bodies can influence the competition, coexistence, and dispersal of viruses within or between our bodies. Taylor's power law (TPL) was first discovered in the 1960s to describe the spatial distributions of plant and/or animal populations, and since then it has been verified by numerous experimental and theoretical studies. Recently, TPL has been extended from population to community level and applied to bacterial communities. Here we report the first comprehensive testing of the TPL fitted to human virome datasets. It was found that the human virome follows the TPL as bacterial communities do. Furthermore, the TPL heterogeneity scaling parameter of human virome is virtually the same as that of the human bacterial microbiome (1.916 vs. 1.926). We postulate that the extreme closeness of human viruses and bacteria in heterogeneity scaling coefficients could be attributed to the fact that most of the viruses that were annotated in this study actually belong to bacteriophages (86% viral OTUs) that "piggyback" on their bacterial hosts, and their distributions are likely host-dependent. The scaling parameter, which measures the inter-subject heterogeneity changes, should be an innate property of human microbiomes including both bacteria and viruses. It is similar to the acceleration coefficient of the gravity ( $g = 9.8$ ) as specified by Newton's law, which is invariant on the earth. Nevertheless, we caution that our postulation is contingent on an implicit assumption that the proportion of bacteriophages to total virome may not change significantly when more virus species can be identified in future.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Our body does not consist of our own cells alone. Instead, it is cohabited by trillions of microorganisms, collectively termed as human microbiome. For this reason, some scientists consider our body as the *holobiont*, consisting of the host cells and all of its symbiotic microbes [23,24]. It is estimated that the human body is inhabited by at least 38 trillion bacteria, which is about 10 times of the cell number of our body. However, the award for the most abundant microbes in the human microbiome cannot be awarded

to bacteria, and instead the award goes to viruses which are estimated to exceed 380 trillions, *i.e.*, approximately 10 times of the bacterial number, that are collectively termed as the human *virome*. As expected, absolute majority of the trillions of viruses are harmless to our hosts, but how they are *distributed* spatially and temporally certainly have significant impacts on our health and diseases. In the present study, we are interested in assessing and interpreting the spatial heterogeneity of the human virome in terms of the inter-subject scaling (changes) of viral OTU abundances.

The term "heterogeneity" literally represents the quality or state of being diverse in character or content, which appears to be the other side of evenness coin or a proxy of biodiversity. However, much of the formal and quantitative research of heterogeneity in ecology has been performed separately from diversity

\* Address: Computational Biology and Medical Ecology Lab, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China.

E-mail address: [ma@vandals.uidaho.edu](mailto:ma@vandals.uidaho.edu)

research [14,15,17,18]. For instance, in population ecology, the terms such as heterogeneity, aggregation, dispersion, and patchiness are often used interchangeably, and all of which can be used to characterize the *spatial distribution* of biological populations. Similarly, the terms such as *spatial heterogeneity*, *spatial dispersion*, and *spatial distribution* are often used interchangeably in population ecology. In these usages, the spatial information is referred to either implicitly or explicitly. Although *temporal heterogeneity* can also be defined and measured, it is essentially a proxy of (temporal) stability of population or community [9,11,20]. In this article, we focus on spatial heterogeneity exclusively.

If the concept of heterogeneity is constrained by the spatial distribution (arrangement) or temporal (variation), its difference with diversity or evenness becomes obvious. The traditional diversity concept and its various metrics do not usually deal with spatial distribution (arrangement) (e.g., [2]). We are only aware of two exceptions: One is beta-diversity, which may implicitly take into the spatial information but it is far less convenient in dealing with spatial information than the heterogeneity concept. Another exception is the concept of dominance [12,13], which has to do with its derivation from the concept of mean crowding [10]. Nevertheless, it is far less comprehensive than the power law parameter we use in this study because it is just an index, rather than a model that is built from multi-site abundance data.

*Spatial heterogeneity* is a fundamental property of any natural ecosystems including the human virome. In ecology, three most important scales that the spatial heterogeneity exhibits are the *population*, *community*, and *landscape*. At population scale, aggregation, dispersion and heterogeneity are often used interchangeably with each other, but the term aggregation is preferred. In population ecology, *aggregation* or heterogeneity is often measured quantitatively with Taylor's [31] power law, which achieved rare status of ecological law [3,28–30,32,33,3,21–22,5,18] as further explained below shortly.

At community scale, *community spatial heterogeneity* can be quantitatively measured with Taylor's power law extensions (TPLEs), which were proposed by Ma [11] via extending Taylor's [31] power law from population to community level. In addition, the heterogeneity of the assemblage of species, which falls between single-species population and community, can be measured with the so-termed mixed species population aggregation [11]. Several applications of the TPLE to human microbiome have been reported (e.g., [15–18,20]). In the present study, we propose and test the application of TPLE for measuring the spatial heterogeneity of human virome. The 'unit' of space in human virome study is usually *individual human subject*, and therefore the terms spatial heterogeneity and inter-individual heterogeneity are equivalent in this study.

The characterizations of human virome heterogeneity are of obvious importance. For example, in the case of pathogenic viruses, understanding the spatial heterogeneity (distribution) of viruses can be critical for investigating the epidemiology (such as outbreak thresholds) (Ma 2020). For another example, in the case of commensal viruses, the spatial distribution of commensal viruses within our bodies or the inter-subject heterogeneity (differences) can influence the competition, coexistence, dispersal, and spread of viruses within or between our bodies, and should have far reaching significance to our healthy and diseases. In fact, the term "heterogeneity" has been listed in the mission statements of both the HMP (Human Microbiome Project) and HMP2 or iHMP (integrative HMP) of the US-NIH, referring to various aspects of the human microbiome from inter-subject heterogeneity in microbiome composition to the heterogeneity in disease treatment responses [6,7]. These appearances highlighted the importance of the microbiome (virome) heterogeneity in human microbiome (virome) research. Nonetheless, assessing and interpreting the

heterogeneity of human microbiome (including virome) can be rather challenging. There have been several applications of the applications of TPL (TPLE) to the bacterial communities (e.g., Citations). To the best of our knowledge, the inter-individual heterogeneity of human virome has not been investigated comprehensively. We fill the gap in this article by applying the TPLE modeling for analyzing the big data of metagenomic sequencing reads of the human virome samples. Table S1 displayed the brief information of 287 virome samples from 4 research projects. Fig. 1 illustrated the strategy and computational procedures of our study.

## 2. Material and methods

### 2.1. Virome datasets and bioinformatics analysis for viral OTU tables

Four published datasets of human viromes were reanalyzed in this study, with a total sample size of 287 (see Table S1). A total of 147 bronchoalveolar lavage samples from lung donors and transplant recipients, including 61 control samples and 80 PGD (Primary Graft Dysfunction) samples, and 34 blood serum samples from recipients (16 control samples and 18 PGD samples) were collected in the United States, and DNA-sequenced for virome research in the dataset (1). Dataset (2) consists of 81 fecal samples of healthy children from four different locations in Venezuela, with sample sizes of Urban A, Village B, Village C and Village D being 20, 10, 16, 15, respectively. The colon samples of dataset (3) were taken from 5 healthy individuals, 6 CD (Crohn's disease) patients and 4 UC (Ulcerative Colitis) patients from Canada. Finally, a total of 30 stool samples of dataset (4) were collected from Tanzanian Hadza hunter-gatherers.

Since it was difficult to collect multiple datasets with the same or even very similar meta (environmental) factors, we decided simply to ignore the difference in meta factors such as host age, sex, etc. Instead, we only required that the datasets were collected for sequencing the reads of human viromes. In other words, we did not care the "heterogeneity" or lack of homogeneity among the treatments of the four datasets. Nevertheless, starting from the virome reads, we kept the exactly computational procedures and quality control measures to get the viral OTU (operational taxonomic unit) tables.

We adopted VirusSeeker, a BLAST-based NGS data analysis software pipeline [35], to reanalyze all of the virome reads with the same configurations and to ensure consistent computational procedures were applied to obtain the viral OTU tables. Fig. 1 illustrated the flowchart of VirusSeeker, as well as the follow-up procedures for performing the power law analysis with the viral OTU tables generated from the VirusSeeker pipeline. An advantage with the VirusSeeker pipeline is that it handles both eukaryotic virus and virome composition equally well, while many other pipelines usually focus on one or the other. This advantage is rather helpful for us to obtain consistent OTU tables across the four datasets.

### 2.2. Taylor's power law (TPL) and TPL extensions (TPLE)

Taylor [31] found that there is a rather robust power function relationship between population mean abundance ( $M$ ) and corresponding variance ( $V$ ) across regions of a species' distribution range, i.e.,

$$V = aM^b \quad (1)$$

For example,  $M_i$  can be the mean population abundance of an insect pest species in the  $i$ -th crop field, and  $V_i$  is the corresponding variance. The  $V_i$ - $M_i$  pairs across a series of crop fields then follow

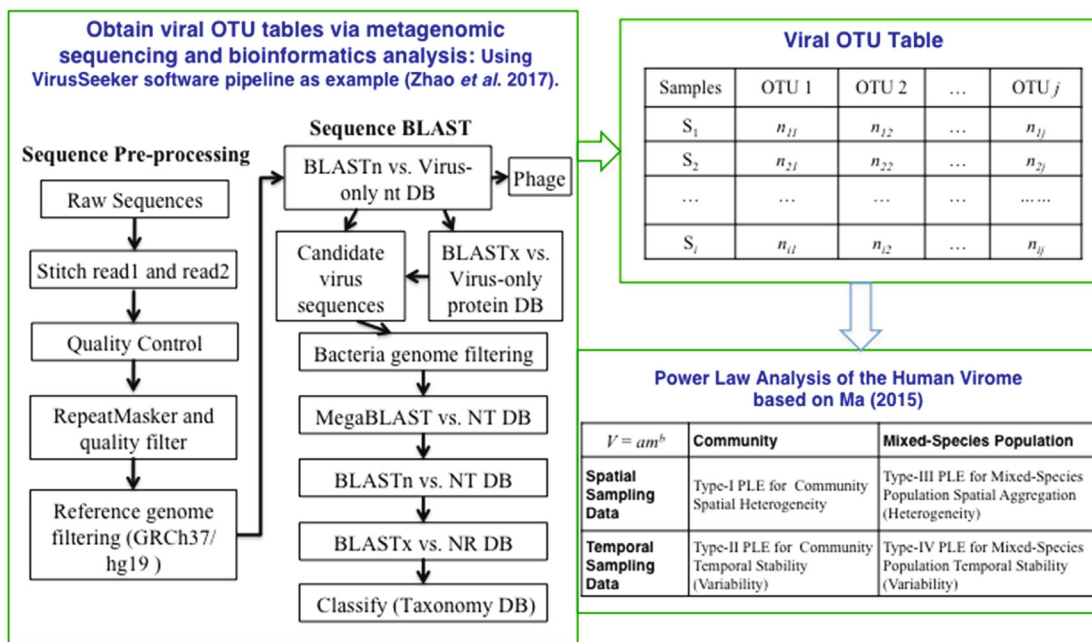


Fig. 1. A diagram illustrating the computational procedure for measuring the spatial heterogeneity of human virome with Taylor's power law extensions (TPLE).

the above power-function relationship. Since 1960s, the power function has been verified by numerous field observations and also theoretical analyses, and it has achieved a rare status of classic ecological law, known as Taylor's power law [4,28–33,3,21,22].

TPL is usually fitted by transforming Eq. (1) into the following log-linear model:

$$\ln(V) = \ln(a) + bM \tag{2}$$

TPL possesses two parameters  $a$  &  $b$ . Parameter  $a$  or  $\ln(a)$  is the intercept of the log-linear form of TPL, and is generally considered of little ecological significance; instead, it is influenced by the environmental factors such as sequencing platforms and/or sampling efforts (sample sizes) in the case of microbiome research. Parameter  $b$  is considered a species-specific parameter that is primarily determined by the species evolutionary history, therefore being invariant with environmental conditions. Note that this is not a consensus agreed upon by all ecologists. There are evidences supporting both invariance and non-invariance assumptions. We argue that whether or not parameter  $b$  is invariant should be determined by statistical tests. In this study, we rely on the permutation test (randomization test) to determine if  $b$  is invariant with environmental factors (or meta factors) such as host health status or sequencing platforms. An advantage of TPL is that parameter  $a$  may “absorb” the noise effects such as meta-factors and sampling efforts.

The classic TPL is a population level law (model), which is determined by the computation of  $V$ - $M$  pairs. Ma [11] extended classic TPL to community level by proposing four TPL extensions (TPLE): Type-I TPLE for measuring community spatial heterogeneity; Type-II for measuring community temporal stability (variability); Type-III for measuring mixed-species population spatial aggregation (heterogeneity); and Type-IV for measuring mixed-species population temporal stability (variability). The commonality between the classic TPL and TPLE is that they all share the same mathematical model (power function) [Eqn. (1) & (2)]. Nevertheless, the interpretations of the model parameter  $b$  in TPL and TPLE are different as explained previously. That is, the TPLEs (Type-I, II, III, & IV) possess four different interpretations, and all are different from the interpretation of  $b$  in the classic TPL.

Type-I and Type-III TPLEs are built with cross-sectional datasets or spatial series of  $V$ - $M$  pairs, and therefore they are used to measure spatial heterogeneity. Type-I TPLE parameter  $b$  is termed community spatial heterogeneity scaling parameter, measuring the change rate of heterogeneity with *mean species size* in viral community. Type-III TPLE parameter  $b$  is termed mixed-species population spatial aggregation scaling parameter, measuring the change rate of aggregation with the *mean population size* in viral community. The differences between Type-I TPLE and Type-III TPLE include: (i) Type-I is community level and Type-III is essentially species-level (mixed-species or averaged-species); (ii) The mean species size of Type-I is computed by averaging the species abundance of all species within a community, whereas the mean population size is computed by averaging the population abundance of a species across multiple communities (or community samples). For these differences, Type-I TPLE scaling parameter essentially measures the heterogeneity among species or community-level spatial heterogeneity, whereas Type-III TPLE scaling parameters essentially measures the aggregation (heterogeneity) of an averaged species across spatial series of communities, and hence is a species-level heterogeneity similar to the population aggregation in the classic TPL.

Type-II and Type-IV TPLEs are built with time-series data or temporal sampling of ecological communities, and therefore are used to measure temporal variability (stability) [11]. These two temporal versions of TPLE are not implicated in this article, but they have been applied to human microbiome and virome (e.g., [11,19,20]).

### 3. Results and discussion

We first fitted Type-I and III TPLEs to the four datasets of human viromes, which consisted of 14 different treatments as detailed in Table S1. The results from fitting the TPLEs to the 14 treatments were summarized in Table S2, and further illustrated in Figs. 2 and 3. These results demonstrated that the TPLEs, whether it is Type-I or Type-III, fitted to the human viromes significantly well ( $P$ -value < 0.0001).

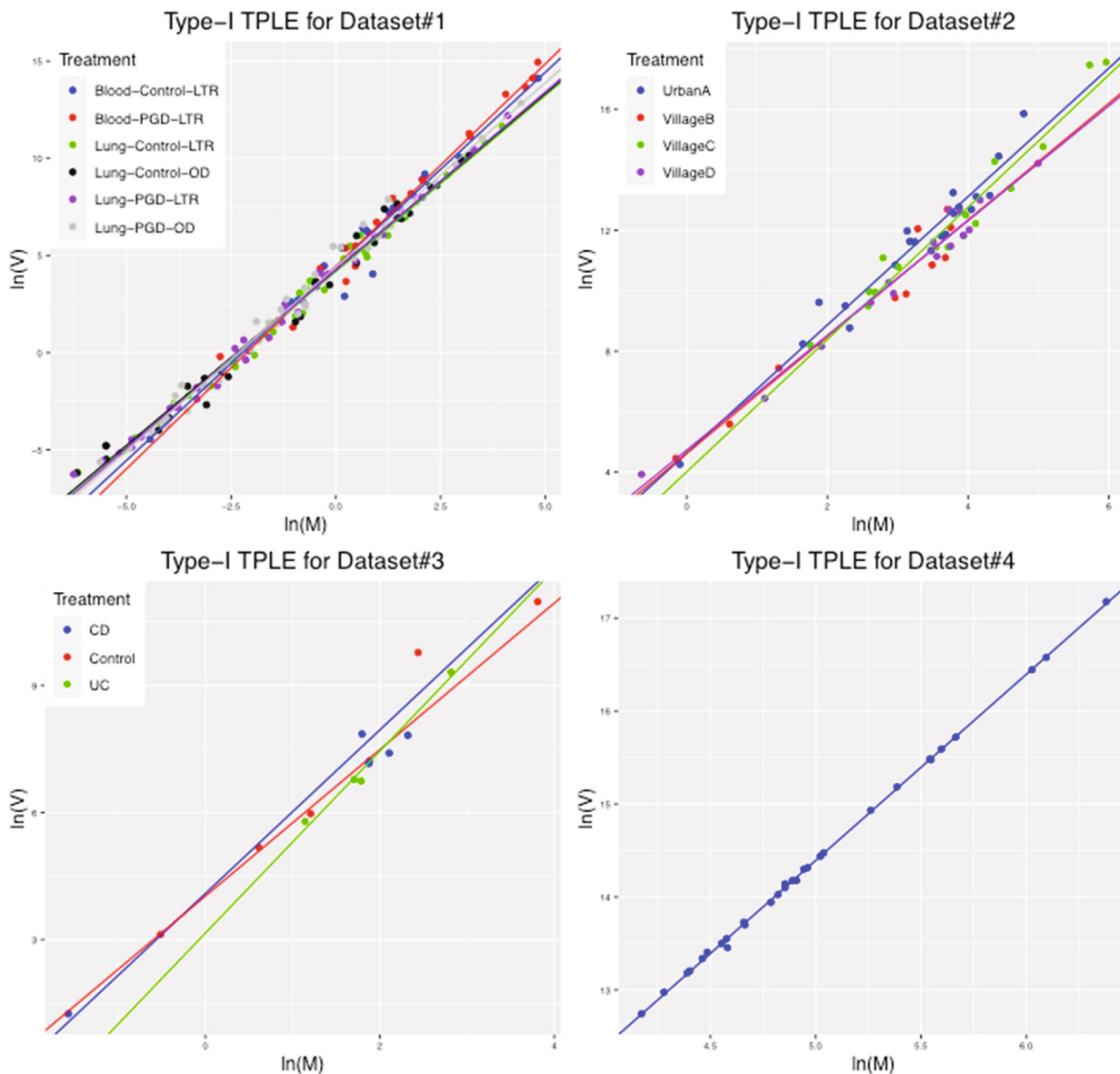


Fig. 2. Fitting Type-I TPLE (Taylor’s power law extension) for measuring community spatial heterogeneity to the 14 treatments of the four human virome datasets separately.

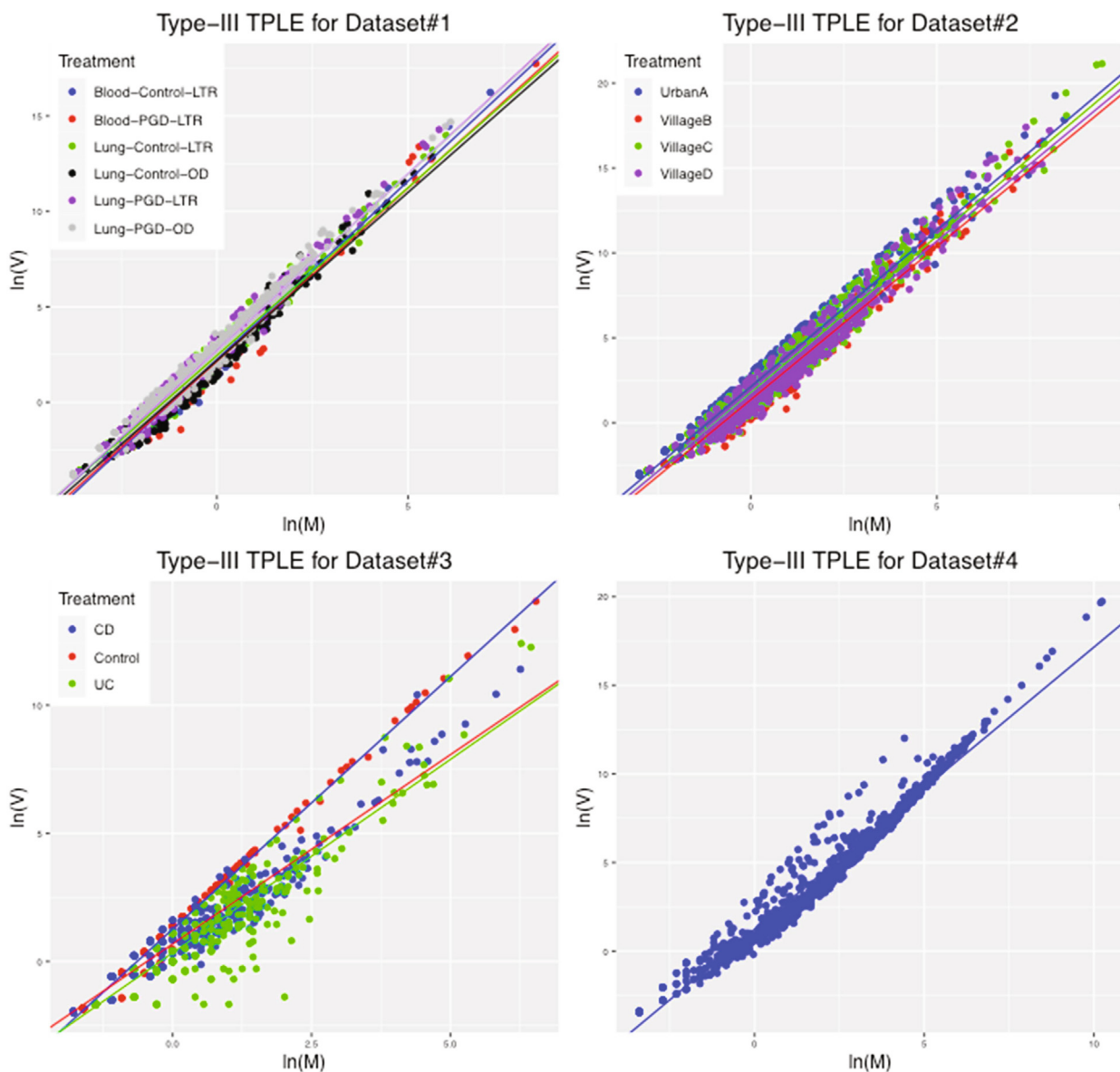
The heterogeneity scaling parameter ( $b$ ) for community scaling parameter ranged between 1.729 and 2.189 with average  $b = 1.956$  (Fig. 2). This range is similar to that of bacterial community of the human microbiome [11], and furthermore the average is extremely close to the mean of bacterial communities of the human microbiome (1.956 vs. 1.926).

The heterogeneity scaling parameter ( $b$ ) for mixed-species population aggregation ranged between 1.477 and 1.970 with an average  $b = 1.767$ , which is slightly higher than that of human bacterial communities (virome = 1.767 vs. bacteria = 1.545). Fig. 3 illustrated the fittings of Type-III TPLE to 14 human virome data groups.

We further tested the potential differences between the treatments of the human virome in the TPLE parameters and the results were listed in Table S3. It was shown that in majority cases (87.5%–100%), there were not any statistically significant differences among different treatments. This finding indicates that, despite potentially significant differences among the 14 treatments, the TPLE parameters are rather robust. This is consistent with a long-standing conjecture that the TPL scaling parameter ( $b$ ) is invariant

with environmental factors. Analogically, the heterogeneity scaling parameter is not unlike the acceleration coefficient of gravity ( $g = 9.8$ ) on the earth planet. The  $g$  is a characteristic (invariant) of the earth, although it may slightly vary with latitudes (longitudes), but the moon may have a different  $g$ . Here we conjecture that human virome should have a characteristic scaling value of community spatial heterogeneity,  $b = 1.956$  as averaged from Table S2. We further realize that this characteristic parameter is rather close to that of the bacterial communities of the human microbiome.

Given there is little significant difference between the treatments of the four datasets as shown in Table S3, we pooled together the samples from all four datasets and fitted the TPLE with the combined samples. Table 1 listed the TPLE parameters for the human virome from the pooled datasets, which we consider as the more realistic scaling parameters for the inter-individual (spatial) heterogeneity of the human virome. Fig. 4 further illustrated the TPLE model fittings. As shown in Table 1, the scaling parameter of community spatial heterogeneity  $b = 1.916$  is even



**Fig. 3.** Fitting Type-III TPLE (Taylor’s power law extension) for measuring mixed-species population spatial aggregation (heterogeneity) to the 14 treatments of the four human virome datasets separately.

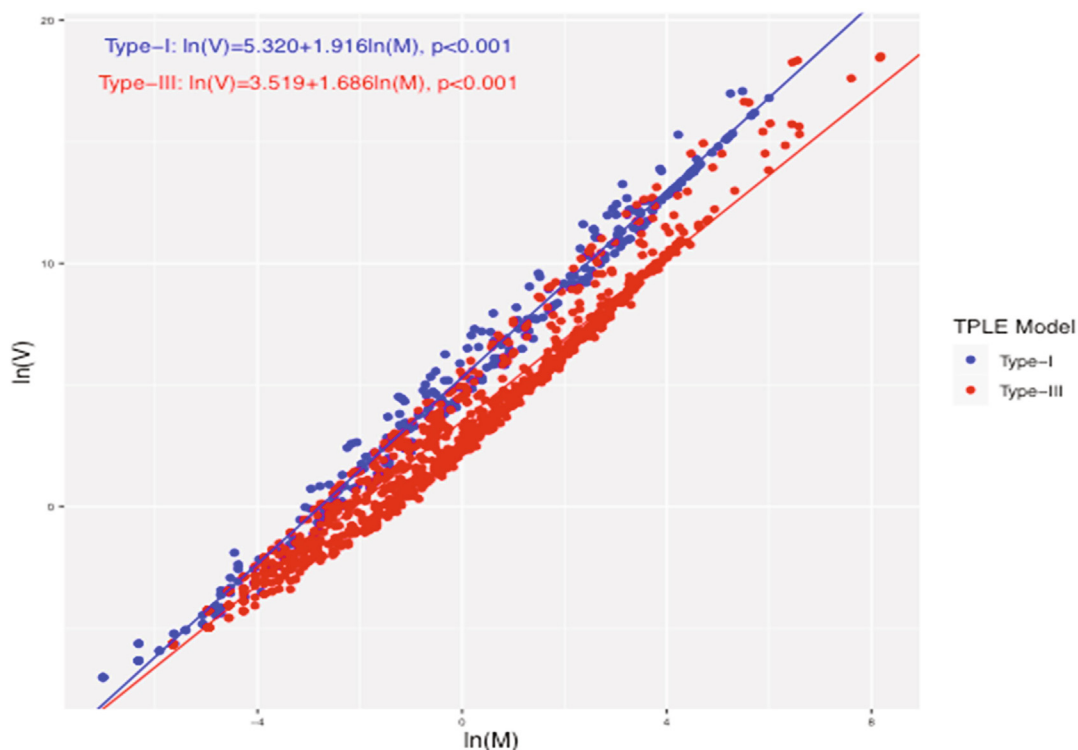
**Table 1**  
The TPLE (Taylor’s power law extension) parameters of the human virome\*.

TPLE Model ( $V = am^b$ )	$b$	$\ln(a)$	$M_0$	$R$	$P$ -value	$N$
Type-I TPLE for the Community Spatial Heterogeneity	1.916	5.320	0.003	0.994	0.000	287
Type-III TPLE for the mixed-species Population Aggregation	1.686	3.519	0.006	0.984	0.000	1115

more closer to the  $b$ -value = 1.926 of bacterial communities of the human microbiome, compared with the averaged  $b$ -value = 1.956 from four separately fitted TPL models in Table S2.

In summary, we conclude that the TPLEs successfully fitted the human viral communities just like they fitted the human bacterial communities. Furthermore, the degree of community spatial heterogeneity scaling of viral communities and bacterial communities are virtually the same ( $b = 1.926$  vs. 1.916) (see [11] for the bacterial community). Our explanation for the virtually same levels of community heterogeneities in both viral and bacterial communities is as follows: As shown in Table S1, approximately 78% of sequencing reads or 86% viral OTUs are actually bacterial phages,

whose distribution should depend on the bacteria because phages are the “parasites” of bacteria [25]. Nevertheless, we caution that the proportion of bacteriophages to total virome may change from dataset to dataset. The state-of-the-art technology in virome sequencing and identifications is actually rather limited given that only small percentages of virome sequences could be matched with known virome sequences in existing genomic databases. With more extensive and intensive virome studies, this limitation may be alleviated or even removed. It is unknown whether such technological advances will significantly change the proportion of bacteriophages to whole virome in future. If the proportion changes, then our previous explanation to the equality of spatial



**Fig. 4.** Fitting the TPLEs (Taylor's power law extension) for measuring community spatial heterogeneity (Type-I) and mixed species population aggregation (Type-III) with the pooled human virome datasets.

heterogeneity between bacterial microbiome and virome may be questioned. Nevertheless, the change does not necessarily invalidate our hypothesis.

As to the slightly difference in the Type-III TPLE scaling parameter ( $b$ ) between the viral communities and bacterial communities (virome = 1.686 vs. bacteria = 1.545), this should be to do with the nature of mixed-species population aggregation. The mixed-species population is similar to single-species population, or it is an averaged virtual population. It is essentially a population-level property. The aggregation levels of an average bacterial species and an average viral species are more likely to be different, as reflected by their different average  $b$ -values (virome = 1.686 vs. bacteria = 1.545).

In conclusions, the previous findings suggest that the spatial heterogeneity of the human virome is similar to that of the bacterial community in the human microbiome if measured at the community scale, and is slightly higher if measured at species level. The “homogeneity” between virome and bacteria is likely due to the parasite-host relationship between bacteria phages and bacteria (hosts), given that 80% of virome reads are actually reads of phages (Table S1).

As to existing similar studies, to the best of our knowledge, there is only one application of TPL (TPLE) to human virome, conducted by Martí [19], which analyzed the temporal fluctuation (stability) of human virome. The application belongs to Type-II TPLE [11] that measures the temporal stability of human virome. In the present study, our focus is to measure the spatial heterogeneity based on cross-sectional datasets of human virome, which utilize Type-I TPLE for measuring community spatial heterogeneity and Type-III TPLE for measuring mixed-species population aggregation, as introduced previously. At the population level, the applications of TPL have been relatively more. For example, Keeling and Grenfell [8] analyzed measles variability with classic single-species TPL. More recently, Ma (2020) applied single-species TPL to analyze the spatiotemporal fluctuations of COVID-19.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgements

This study received funding from National Science Foundation of China (Grant #31970116); Cloud-Ridge Industry Technology Leader Grant from Yunnan Development and Reform Commission; A grant for genomics/metagenomics big data from Yunnan Science and Technology Bureau. We appreciate the computational assistance from WM Xiao, LW Li, YT Qiao, and WD Li from Chinese Academy of Sciences.

#### Author Contributions

ZSM designed and performed the study, wrote the paper. All authors approved the submission.

#### Data availability and ethical approval

All datasets used in this study are available in public domain and their sources were noted in Table S1. No data related to human subjects are used in this study and no ethical approval is applicable.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2021.04.069>.

## References

- [2] Chao A, Gotelli NJ, Hsieh TC, et al. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecol Monogr* 2014;84:45–67.
- [3] Cohen JE, Saitoh T. Population dynamics, synchrony, and environmental quality of Hokkaido voles lead to temporal and spatial Taylor's laws. *Ecology* 2016;97(12):3402–13.
- [4] Cohen JE, Xu M. Random sampling of skewed distributions implies Taylor's power law of fluctuation scaling. *Proc Natl Acad Sci USA* 2015;112(25):7749–54.
- [5] Grilli J. Macroecological laws describe variation and diversity in microbial communities. *Nat Commun* 2020;11(1):4743.
- [6] Human Microbiome Project Consortium (HMP). Structure, function and diversity of the healthy human microbiome. *Nature* 2012;486:207–14.
- [7] Integrative HMP (iHMP) Research Network Consortium The integrative human microbiome project. (2019) *Nature*. 569:641–648.
- [8] Anderson RM, Keeling M, Grenfell B. Stochastic dynamics and a power law for measles variability. *Philos Trans R Soc Lond B Biol* 1999;354(1384):769–76.
- [9] Li LW, Ma ZS. Comparative power law analysis for the spatial heterogeneity scaling of the hot-spring and human microbiomes. *Mol Ecol* 2019;28(11):2932–43.
- [10] Lloyd M. Mean crowding. *J Anim Ecol* 1967;36(1):1. <https://doi.org/10.2307/3012>.
- [11] Ma ZS. Power law analysis of the human microbiome. *Mol Ecol* 2015;24(24):5428–45.
- [12] Ma ZS, Ellison AM. A unified concept of dominance applicable at both community and species scales. *Ecosphere* 2018;9(11):e02477. <https://doi.org/10.1002/ecs2.2018.9.issue-1110.1002/ecs2.2477>.
- [13] Ma ZS, Ellison AM. Dominance network analysis provides a new framework for studying the diversity–stability relationship. *Ecological Monograph* 2019;89(2):e01358. <https://doi.org/10.1002/ecm.2019.89.issue-210.1002/ecm.1358>.
- [14] Ma ZS, Li W. How and why men and women differ in their microbiomes: medical ecology and network analyses of the microgenderome. *Adv Sci* 2019;6(23):1902054.
- [15] Ma ZS. Assessing and interpreting the metagenome heterogeneity with power law. *Front Microbiol* 2020;11:648.
- [16] Ma ZS. Predicting the outbreak risks and inflection points of COVID-19 pandemic with classic ecological theories. *Adv Sci* 2020;7(21):2001530. <https://doi.org/10.1002/advs.v7.2110.1002/advs.202001530>.
- [17] Ma ZS. Heterogeneity-disease relationship in the human microbiome associated diseases. *FEMS Microbiol Ecol* 2020;96:faa093. <https://doi.org/10.1093/femsec/faa093>.
- [18] Ma ZS, Taylor RAJ. Human reproductive system microbiomes exhibited significantly different heterogeneity scaling with gut microbiome, but the intra-system scaling is invariant. *Oikos* 2020;129(6):903–11.
- [19] Martí JM (2018) Robust Analysis of Time Series in Virome Metagenomics. In: Moya A., Pérez Brocal V. (eds) *The Human Virome*. *Methods in Molecular Biology*, vol 1838. Humana Press, New York, NY. [https://doi.org/10.1007/978-1-4939-8682-8\\_17](https://doi.org/10.1007/978-1-4939-8682-8_17)
- [20] Oh J, Byrd A, Park M, Kong H, Segre J. Temporal stability of the human skin microbiome. *Cell* 2016;165(4):854–66.
- [21] Reuman DC, Gislason H, Barnes C, Mélin F, Jennings S, Webb T. The marine diversity spectrum. *Anim Ecol* 2014;83(4):963–79.
- [22] Reuman DC, Zhao L, Sheppard LW. (2017). Synchrony affects Taylor's law in theory and data. *Proc Natl Acad Sci USA*. 114:6788.
- [23] Rosenberg E, Sharon G, Zilber-Rosenberg I. The hologenome theory of evolution: a fusion of neo-Darwinism and Lamarckism. *Environ Microbiol* 2009;11:2959–62.
- [24] Rosenberg E, Zilber-Rosenberg I. The hologenome concept of evolution after 10 years. *Microbiome* 2018;6(1):78.
- [25] Secor PR, Dandekar AA (2020) More than simple parasites: the sociobiology of bacteriophages and their bacterial hosts. *mBio* 11:e00041-20. <https://doi.org/10.1128/mBio.00041-20>.
- [28] Stumpf MPH, Porter MA. Mathematics: critical truths about power laws. *Science (New York, N.Y.)* 2012;335:665–6.
- [29] Taylor LR, Perry JN, Woiwod IP, Taylor RAJ. Specificity of the spatial power-law exponent in ecology and agriculture. *Nature* 1988;332(6166):721–2.
- [30] Taylor LR, Taylor RAJ, Woiwod IP, Perry JN. Behavioural dynamics. *Nature* 1983;303(5920):801–4.
- [31] TAYLOR LR. Aggregation, variance and the mean. *Nature* 1961;189(4766):732–5.
- [32] Taylor LR. Assessing and interpreting the spatial distributions of insect populations. *Annu Rev Entomol* 1984;29(1):321–57.
- [33] Taylor R. *Taylor's power law: order and pattern in nature*. Cambridge, MA: Academic Press; 2019. p. 639.
- [35] Zhao G, Wu G, Lim ES, Droit L, Krishnamurthy S, Barouch DH, et al. VirusSeeker, a computational pipeline for virus discovery and virome composition analysis. *Virology* 2017;503:21–30.

## Further reading

- [1] Abbas AA, Diamond JM, Chehoud C, Chang B, Kotzin JJ, Young JC, et al. The perioperative lung transplant virome: torque teno viruses are elevated in donor lungs and show divergent dynamics in primary graft dysfunction. *Am J Transplant* 2017;17(5):1313–24.
- [26] Siqueira JD, Dominguez-Bello MG, Contreras M, et al. Complex virome in feces from Amerindian children in isolated Amazonian villages. *Nat Commun* 2018;9(1):4270.
- [27] Smits SA, Leach J, Sonnenburg ED, Gonzalez CG, Lichtman JS, Reid G, et al. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* 2017;357(6353):802–6.
- [34] Wang W, Jovel J, Halloran B, Wine E, Patterson J, Ford G, et al. Metagenomic analysis of microbiome in colon tissue from subjects with inflammatory bowel diseases reveals interplay of viruses and bacteria. *Inflamm Bowel Dis* 2015;1. <https://doi.org/10.1097/MIB.0000000000000344>.