Revisiting Enterotypes: A viral perspective

Journal:	*Future Virology*
Manuscript ID	FVL-2017-0056
Manuscript Type:	Editorial
Keywords:	Human gut microbiome, human gut virome, bacteriophage

EDITORIAL

When first presented, the idea that all human gut microbiomes belong to one of three 'enterotypes' [1] hit a collective nerve and stimulated an important debate within the field. Taken at face value, the enterotypes model offers an intuitive and attractive way to stratify individuals based on microbial community composition. The ability to stratify individuals based on their gut microbiome composition is opening up intriguing options for applying personalised diagnostics, prognostics, and treatment approaches for a number of diseases as well as for general well-being.

Arumugam and co-workers [1] analysed the community composition of multiple existing and newly-generated human gut microbial metagenomes from diverse origins (European, Japanese and American), generated using a range of sequencing technologies. The relative abundance profiles of bacterial species within these datasets were clustered revealing that each metagenome could be attributed to one of three distinct groups, dominated by a different microbial genus, namely *Bacteroides, Prevotella*, or *Ruminococcus*. No association of these enterotypes with available characteristics of human hosts (from which datasets were derived) could be delineated, and they were found to be poor predictors of host properties such as age, gender and body mass index (BMI), as well as nationality. There was, however, some weak taxonomy-independent correlations between host features and function at the level of genes and pathways.

The study set in motion a flurry of activity to further define and validate these enterotypes, and provide the 'deeper and broader analysis' requested by the authors themselves. This encompassed microbial gut metagenomes originating from an extended demographic, ethnic and geographic range. Overall these studies have in broadly equal measures refuted and confirmed the original enterotypes theory and definitions, including contracting and redefining the major phylogenetic designations around these groups [2-6], potentially establishing associations with long-term dietary habits [3] and disease [7], as well as attempting to place the enterotype theory as a whole, within the ecological framework of the keystone species concept [8]. It has also been postulated that rather than a few completely distinct microbiome profiles (or enterotypes), there is instead an overall

biological gradient of microbiome population structures, with the discernible enterotypes initially described perhaps occupying the extremes of these gradients [9]. There also remains the caveat that to date such observations are derived from a highly restricted set of gut microbiomes, and a very small sample size, given the scale and diversity of the human population. Therefore, it remains to be seen if such discrete groupings will continue to be observed as new data accumulates.

However, the investigations conducted so far have, by and large, failed to incorporate the retinue of viruses intimately associated with the human gut ecosystem - the human gut virome. In particular, the dominant bacteriophage component, which is now increasingly recognised as an important facet of this ecosystem, with the potential to significantly influence microbial host microbiome structure and functioning [10-13]. In contrast to many other environments, the human gut virome is apparently dominated by temperate bacteriophage, which seem to have a temporally stable but highly variable structure, with a higher degree of inter-personal variation than we see within the associated gut microbiome [10-12]. The observed variation is most likely a result of the associated variability in underlying host microbiomes [14], the ability of bacteriophage to undergo rapid evolution to form new virotypes [12] and perhaps even the persistence of a core group of virotypes within individuals, i.e. the existence of a core 'phageome' [15]. Consideration of the human gut virome in defining any structural and/or functional groups within the human gut microbiome will therefore be of utility in elucidating the origins, consequences, and the nature of any community level structures such as enterotypes.

Our previous work [16] used a genome signature profiling approach to interrogate whole community human gut metagenomes, enabling us to access a subset of gut-specific *Bacteroidales*-like bacteriophage sequences that were poorly represented in publically available virus-like particle (VLP) metagenomes. Analysis of the inter-individual variation in the carriage of homologous sequences to these *Bacteroidales*-like bacteriophage, within 139 human gut metagenomes, revealed several variants of relative abundance profiles shared by multiple individuals. These profiles could be broadly assigned to four groups, which we termed putative 'viral-enterotypes'. Given that two out of the three microbial enterotypes presented by Arumugam *et al.* [1] were characterised by members of the *Bacteroidales*

(*Bacteroides* and *Prevotella*), it seems likely the associated bacteriophage may also follow a similar pattern. The bacteriophage-orientated groupings we observed were, however, less well-defined than those of Arumugam and co-workers, and we speculated that they may be more indicative of alternative models centred on inter-individual gradients in bacteriophage population structure rather than entirely discrete clusters [16]. As with alternative explanations for microbial enterotypes, the groupings we observed may only represent the extremes of a complex and multifaceted spectrum of viral community structures, as captured by the datasets available. In this context it is also important to note that investigation into the existence and nature of enterotypes (viral or bacterial), will need to also consider the potential temporal fluidity of gut communities [5,9,17,18], with datasets used by most studies providing only a fixed snap shot at a single time point.

Although our observations provided the first indication that the enterotypes concept could potentially be extended to the bacteriophage fraction of the gut microbiome, we were acutely aware that further work would need to be done to substantiate these initial observations. A broader analysis base is necessary for validating and determining what significance, if any, the existence of such viral-enterotypes would have for the structure and functioning of the human gut ecosystem. As with the follow-on studies triggered by the original enterotypes work [1], extended analyses are required that subjects the existence of putative viral groups, gradients, or 'community types' [19] to the same high level of scrutiny. Additional studies should incorporate both increased numbers of datasets of a broader geographical scale, as well as those that move beyond a single snap-shot of the human gut microbiome or virome, to those that consider the temporal dynamics of phage-host interactions [16]. Although intriguing and worthy of further study, the concept of "viral-enterotypes" should be treated with caution until more extensive research is conducted and further evidence presented.

As more large-scale datasets descriptive of the structure and functioning of the human gut virome and microbiome become available, the type of analysis methods applied will be pivotal for using and interpreting the data generated [20]. In the context of enterotypes, the statistical analysis techniques applied to identify potential community types have also been a source of debate and re-evaluation, highlighting potential artifacts generated when

analysing metagenomic datasets characterised by a high number of data points and a small sample size [5, 9, 18] and stimulating the emergence of alternative approaches to understand the origins and nature of enterotypes [19].

When taking a viral perspective, however, and exploring viral community structure, we face a number of unique challenges [20]. Much of our current understanding has been driven by alignment-based methodologies, e.g. BLAST, which use a range of algorithms to determine similarities between sequences at the nucleotide or amino acid levels. But the success of this approach is largely dependent on the existence of similar well-characterised, and accurately curated, sequences in databanks. For bacteriophage, in particular, we are still at the early stages of charting the diversity and ecological characteristics of the human gut, with a dearth of well-characterised gut-associated bacteriophage genomes available [20]. Moreover, ecological and evolutionary insight is hampered by a lack of a common phylogenetic anchor, similar to the 16S rRNA gene, used for many phylogenetic studies of the human gut microbiome.

Despite these challenges, progress in recent years is ushering us into a new era of viral ecogenomics, in which metagenomic sequencing and associated analysis techniques are enabling interrogation of the viral gene space with the same vigour as the microbiome. In particular, new viral–orientated analysis methods are allowing exploration of the bacteriophage-host relationship in unprecedented detail, with a view to unravelling the multi-partite interactions occurring within this ecosystem. Application of novel bioinformatics analysis approaches is highlighting unique ecological characteristics of the human gut virome, including the potential existence of a healthy core gut phageome and the conservation of virotypes across individuals of diverse geographical origins [15, 16, 21, 22]. However, as with study of the bacterial fraction of the gut microbiome, it must also be recognised that brute-force metagenomic sequencing alone (although powerful and undeniably useful) is not capable of delivering the required insights in isolation. More effort and emphasis needs to be placed on traditional culture-based approaches, and the provision of well-characterised reference bacteriophage (both in terms of genotype and phenotype). If we are to truly access and understand viral ecosystems, such as the gut microbiome, the synergistic application of both culture-based and metagenomics approaches will almost

certainly be required.

In summary, the ever growing knowledge base on the human gut microbiome (and virome) is providing many avenues of discovery, and importantly, initiating healthy and productive discussions regarding analysis and interpretation of metagenomic datasets in general. Further fine-scale analysis of existing and newly-generated datasets from human gut samples, applying viral-orientated techniques in combination with other microbiome targeted analysis (both culture dependent and independent), will provide a clearer view of the structure, gradients and groupings (if any) extant within the human gut virome.

REFERENCES

1. *Arumugam M, Raes J, Pelletier E* et al. Enterotypes of the human gut microbiome. *Nature* 473, 174–180 (2011).
2. *Huse SM, Ye Y, Zhou Y, Fodor AA*. A core human microbiome as viewed through 16S rRNA sequence clusters. PLoS One 7(6), e34242 (2012).
3. *Wu G, Chen J, Hoffmann C et al.* Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. Science 334(6052), 105-108 (2011).
4. *Lim MY, Rho M, Song YM* et al. Stability of gut enterotypes in Korean monozygotic twins and their association. Sci Rep 8, 7348 (2014).
5. *Gorvitovskaia A, Holmes SP, Huse SM*. Interpreting Prevotella and Bacteroides as biomarkers of diet and lifestyle. Microbiome 4, 15 (2016).
6. *Liang C, Tseng H-C, Chen H-M* et al. Diversity and enterotype in gut bacterial community of adults in Taiwan. BMC Genomics 18(Suppl 1), 932 (2017).
7. *de Moraes AC, Fernandes GR, da Silva* et al. Enterotype May Drive the Dietary-Associated Cardiometabolic Risk Factors. Front Cell Infect Microbiol 7, 47 (2017)
8. *Fisher CK, Mehta P.* Identifying Keystone Species in the Human Gut Microbiome from Metagenomic Time series Using Sparse Linear Regression PLoS One 9(7), e102451 (2014).
9. *Koren O, Knights D, Gonzalez A* et al. A Guide to Enterotypes across the Human Body: Meta-Analysis of Microbial Community Structures in Human Microbiome Datasets. PLoS Comput Biol. 9(1) (2013).
10. *Reyes A, Haynes M, Hanson N* et al. Viruses in the faecal microbiota of monozygotic twins and their mothers. Nature 466, 334–338 (2010).
11. *Minot S, Sinha R, Chen J* et al. The human gut virome, inter-individual variation and dynamic response to diet. *Genome Res.* 21, 1616–1625 (2011).
12. *Minot S, Grunberg S, Wu G* et al. Hypervariable loci in the human gut virome. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3962–3966 (2012)
13. *Modi SR, Lee HH, Spina CS* et al. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* 499 219–222 (2013).

14. *Huttenhower C, Gevers D, Knight R* et al. Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214 (2012).
15. *Manrique P, Bolduc B, Walk ST* et al. Healthy human gut phageome. Proc Natl Acad Sci USA 113 (37), 10400-5 (2016).
16. *Ogilvie L A, Bowler LD, Caplin J* et al. Genome signature-based dissection of human gut metagenomes to extract subliminal viral sequences. Nat. Commun. 4, 2420 (2013).
17. *Jeffery IB, Claesson MJ, O'Toole PW et al.* Categorization of the gut microbiota: enterotypes or gradients? Nat Rev Micro. 10(9), 591–592 (2012).
18. *Knights D, Ward TL, McKinlay CE* et al., Rethinking Enterotypes. Cell Host & Microbe 16(4), 433 - 437 (2014).
19. *Gibson TE, Bashan A, Cao HT et al.* On the Origins and Control of Community Types in the Human Microbiome. PLoS Comput Biol. 12 (2) e1004688 (2016).
20. *Ogilvie LA, Jones BV.* The human gut virome: a multifaceted majority. Front Microbiol 6, 918 (2015).
21. *Stern A, Mick E, Tirosh I, Sagy O* et al. CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome. Genome Res. 22, 1985–1994 (2012).
22. *Dutilh BE, Cassman N, McNair K* et al. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. Nat. Commun. 5, 4498 (2014).

FINANCIAL DISCLOSURE/ACKNOWLEDGEMENTS

This work was supported by funding from the University of Brighton (Research Challenges Grant awarded to LAO). The authors declare no competing interests.