Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Virology

Shopping for phages? Unpacking design rules for therapeutic phage cocktails

Cédric Lood[1,2], Pieter-Jan Haas[3], Vera van Noort[2,4] and Rob Lavigne[1]

In bacteriophage therapy, the combination of different phages into a single cocktail is of critical importance to overcome the narrow host range of single phage isolates. Today, the design of therapeutic cocktails is often akin to a black box and relies largely on intuition and (pre-)availability of isolates in local collections. Here we show that straightforward host range analysis can disclose design rules and we propose to apply/translate a data mining approach, historically applied in the field of marketing ('shopping cart analysis') to explore patterns in phage combinations. The technique is broadly applicable to host range datasets and can serve in combination with other molecular-based approaches to propose rationales for phage cocktail design.

Addresses
[1] Department of Biosystems, Laboratory of Gene Technology, KU Leuven, Leuven, Belgium
[2] Department of Microbial and Molecular Systems, Centre of Microbial and Plant Genetics, Laboratory of Computational Systems Biology, KU Leuven, Leuven, Belgium
[3] Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands
[4] Institute of Biology, Leiden University, Leiden, The Netherlands

Corresponding author: Lood, Cédric (cedric.lood@kuleuven.be)

Current Opinion in Virology 2022, **52**:236–243

This review comes from a themed issue on **Bacteriophage therapy**

Edited by **Joana Azeredo** and **Jean-Paul Pirnay**

For complete overview about the section, refer Bacteriophage therapy (2022)

Available online 28th December 2021

https://doi.org/10.1016/j.coviro.2021.12.011

Introduction

The application of bacterial viruses (phages) to treat bacterial infections came about at the beginning of the 20th century, shortly after the initial discovery of lytic phages as natural predators of bacteria [1,2]. The first documented phage therapy effort was led in 1919 by the co-discoverer of phages, Félix d'Hérelle, for the treatment of dysentery (See Ref. [3] for a historical review of the field). One hundred years later, a number of phase I/II clinical trials as well as multi-patients case series using phage products have been reported, but no trial has reached phase III to date and broad implementation of phage therapy in the West still eludes us [4••,5]. Nonetheless, successful cases relevant to human health have been reported, including (but not limited to) treatment of ESKAPE pathogens [6]. In addition, protocols outside the framework of clinical trials have been established to treat individual patients using phages [7••,8•].

Phage isolates used in therapy are typically evaluated on multiple characteristics, including safety parameters such as the absence of virulence factors, lysogenic and transduction potential, and their applicability to different strains [9]. This last parameter, known as the host range, is one of the key determinants for the therapeutic applications of a given phage, especially as inclusion criteria of patients for phage therapy often depend on the sensitivity of the infectious agents to the product [10]. Indeed, contrary to the broad spectrum of antibiotics, phages have a narrow host-range, typically infecting only a subset of strains within a single species. In addition, depending on the bacterial species targeted, phages having a broader host-range may not be available [11••]. When considering the ESKAPE pathogens, only *Staphylococcus aureus* appears to be susceptible to phages targeting almost all isolates of the population (>97% of MRSA strains for φMR003 [12]).

Bacterial species have extensive pangenomes, meaning that different strains within a species often have a unique assortment of accessory genes. The diversity in gene content together with strain-specific sequence reflect the intra-species diversity strains that make up their population [13]. Each strain within a species may be infected by a different subset of phages. To illustrate this last point, one may compare the phages that infect *Pseudomonas aeruginosa* [14]. There we find divergent phage clades, ranging from familiar *Caudovirales* isolates such as N4-like, T7-like, or jumbo phages like PA5oct and phiKZ [15,16] to filamentous phages and dsRNA phages [17,18]. Importantly, although these phages share the same host (*P. aeruginosa*), they are so diverse genetically that they represent disconnected viral kingdoms in the current classification models of the viral universe [19].

To overcome the host range limitations of single phage isolates, a common approach in phage therapy is to pool together different phages into a single cocktail product

[20]. To date, very few ground rules or analytical techniques for the design of such cocktails have been established [21]. As such, the choice of phages that will go in a cocktail typically depend on the pre-availability of products available via phage therapy centers with private collections of phage isolates, or at times recently isolated phages available in the lab connected to the therapeutic effort. With the increasing calls to establish centralized libraries of well-characterized phages [22••], the field could benefit from computational approaches that can infer ground rules for the design of cocktails.

In this opinion piece, we review current design rationales for phage cocktails that have been used in clinical trials and case series reports, and we propose to use a data mining technique, originally implemented in the field of marketing known as 'Association rules mining' that could be used to explore and unpack design rules. We illustrate the potential of this technique to generate design rules on a host range dataset of *P. aeruginosa* and its phages.

On the design of phage cocktails: current host-range driven principles

We surveyed the literature of phage clinical trials and case series reports for indications of phage cocktail products used and any design rules documented in the related publications and reports (Table 1) [4••,23]. Overall, out of the fourteen phase I/II clinical trials and nine case series selected, we can observe a few recurring products and strategies.

On the one hand, we note the frequent use of phage therapy products that have been in use in the Eastern world. These products include the Pyophage and Coli-Proteus cocktails from the Eliava Institute in Georgia and Microgen in Russia. These products have been explored using metagenomics and target multiple pathogen species [24•,25]. Their designs likely aggregate long term empirical experience and iterations of their composition, but we lack details on their exact makeup and design rules. This limits their implementation into Western medical practices and its legislative frameworks to some extent, including magistral preparations where individual phages need to be assessed [8•].

Other phage cocktails do have precise compositions and all their isolates have been characterized and sequenced (except studies with Id #8 and #15 in Table 1). However, they do not appear to be available commercially, except for PreforPro. These cocktails can be divided in two categories, i) single pathogen targeting (Biophage-PA, T4-like coliphage, PP1131, AB-SA01, or ii) multiple-species (WPP-201, BFC-1). In both cases, considerations on host-range are usually briefly discussed in the related (or adjunct) publications but the decisions pertaining to the final choice of composition are often not available. Notable exceptions are the BFC-1 cocktail, where considerations were given to select a single *S. aureus* phage with broad host range (possible in that species), as well as two *P. aeruginosa* phages with different receptors (LPS and type IV pili) to prevent resistance development. The other exception is the T4-like cocktail used in the Nestle/Bangladesh trial and that was designed to comprise isolates from different sub-clades of the T4-like phages.

Association rules mining: concept and translation to phage-bacteria interactions

Association rules mining, also known as frequent itemset mining [48] is a computational technique that enables the analysis of customer transactions. The purpose of such analysis is to extract rules that govern associations and patterns between items (e.g., customers may often buy both bread, butter, and jam together). The technique can also be readily extended to discover negative associations, that is, items that are mutually exclusive or never bought together [49]. When we translate this approach to the universe of interactions between phages and bacteria, the associations one may try to detect are whether some phage isolates tend to infect the same bacteria (positive association). We can also mine for negative associations, through which we may find phages that have mutually exclusive host ranges.

One key advantage of the approach is that it does not necessitate an *a priori* knowledge of the genomes of the bacteria or the phages. Indeed, the only required information are the labels of the host-range dataset and the results of the interaction assays. Mining for rules is essentially a combinatorial problem, and as such many rules can be inferred using these algorithms. However, not all rules are necessarily robust, and we typically associate to each rule a computed score or quality metric. The mining of rules stops by using criteria of computing time (for very large datasets) or depth of associations. A rule consists simply of one or more antecedents, and a consequent, for example, A and B could be antecedents and C a consequent and the rule could be 'A and B imply C'. After a set of rules has been inferred, it becomes possible to sort them using quality metrics and retrieve the top rules for closer inspection. For our application to host range analysis, the main workflow consists of the following steps (Figure 1):

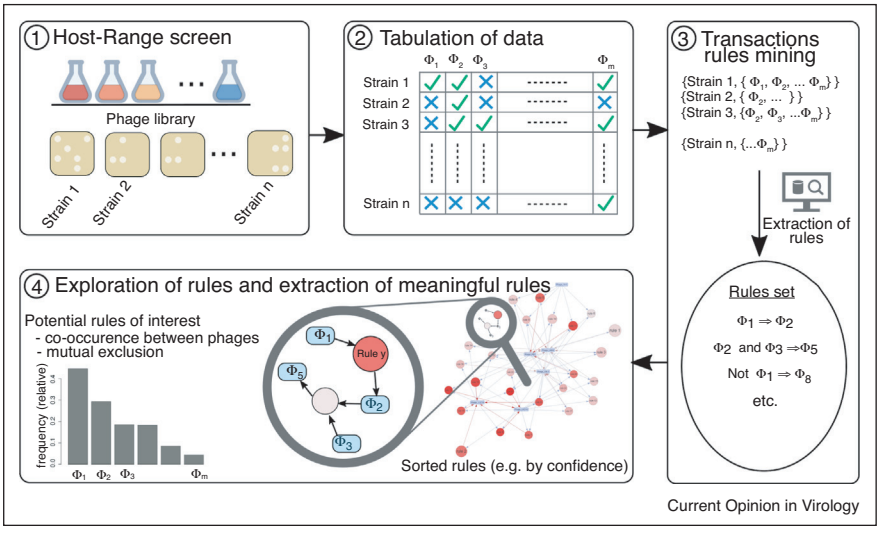
1 A host range experiment is conducted in which a series of phages is screened against a panel of bacteria to see which ones can productively infect (plaque) the strains of the panel.
2 The results of the host range are typically encoded in a table, with each row representing a bacterial strain and the columns represent the different phages. The values that are encoded are converted to binary values (productive infection or not).

Table 1

List of completed clinical trials studies augmented with multi patients case series reports (sorted by date). Information about the cocktail product used is provided when available.

Id	Study description	Target	Cocktail used, composition, and indication of design	Ref.
1	Chronic otitis, Phase I/II (EUDRACT 2004-001691-39)	*P. aeruginosa*	Biophage-PA cocktail, composed of 6 phages (*Pseudomonas* phage BC-BP-01 to 06). Patent US2007/0190033 describes the combination as 'maximizing host range on a panel of strains'.	[26]
2	Venous leg ulcers, Phase I/II (NCT00663091)	*S. aureus*, *P. aeruginosa*, *E. coli*	WPP-201, composed of 8 phages. No indication of design rules beyond targeting multiple bacterial species in the related publication.	[27]
3	Safety oral application (Nestle/Bangladesh Trial), Phase I	*E. coli*	T4-like coliphage cocktail composed of 9 phages, selected from a collection of 98 coliphages to maximize clade diversity (phylogeny).	[28]
4	Oral phage therapy diarrhea (Nestle/ Bangladesh trial), Phase I/II (NCT00937247)	*E. coli*	T4-like coliphage cocktail composed of 9 phages (see above Id #3) + ColiProteus cocktail (Microgen)	[29]
5	Safety oral application (Nestle/Bangladesh Trial), Phase I	*E. coli*	T4-like coliphage cocktail composed of 9 phages (see above Id #3) + ColiProteus phage cocktail (Microgen)	[30]
6	Safety trial *S. aureus* carriers (Phase I)	*S. aureus*	Pyophage cocktails from Eliava and Microgen + Monophage isolate with broad host range from Eliava collection.	[24•]
7	Gastrointestinal health safety, Phase I (NCT03269617)	*E. coli*	PreforPro commercial product composed of 4 coliphages (*Escherichia* phage LH01, LL4, T4D, and LL12). Indication of morphological diversity of the phages (*Myoviridae* and *Siphoviridae*)	[31]
8	PhagoBurn, Phase I/II (NCT02116010)	*P. aeruginosa*	P1131 phage cocktail composed of 12 *Pseudomonas* phages. Brief indications of morphological diversity (*Podoviridae* and *Myoviridae*). No sequences available. The cocktail is also discussed in Ref. [32] where they mention that it lyses 84% of their panel of 33 *P. aeruginosa* strains.	[33]
9	Effects of supplemental phage intake on inflammation and gut microbiota, Phase I (NCT03269617)	*E. coli*	PreforPro cocktail (see above Id #7)	[34]
10	Chronic rhinosinusitis, Phase I (ACTRN1261600000024)	*S. aureus*	AB-SA01 phage cocktail, composed of three phages closely related to *Staphylococcus* phage K. The cocktail is indicated to have a broad host range, see also Ref. [35]	[36]
11	Safety in serious infections by *S. aureus*, Phase I (NCT03395769)	*S. aureus*	AB-SA01 phage cocktail (see above Id #10)	[37•]
12	Gastrointestinal health, Phase I/II (NCT04511221)	*E. coli*	PreforPro cocktail (see above Id #7)	[38]
13	UTI treatment, Phase I/II (NCT03140085)	*Enterococcus spp.*, *E. coli*, *Streptococcus spp.*	Pyophage cocktail from Eliava Institute, extended with *Streptococcus* phages (no indication of strain/origin).	[39•]
14	Chronic prostatitis, case series (3 patients)	*E. faecalis*	Phages from the IIET collection, no further indication or sequencing available.	[40]
15	Safety trial oral coliphage, case series (10 patients)	*E. coli*	ColiProteus cocktail (Microgen).	[25]
16	Burn wounds, case series (9 patients)	*S. aureus*, *P. aeruginosa*	BFC-1 phage cocktail. The cocktail contains one broad host range *Staphylococcus* phage and two *Pseudomonas* phages with different receptor.	[41]
17	Diabetic foot ulcer, case series (6 patients)	*S. aureus*	Sb-1 single phage with broad host range.	[42]
18	UTI, case series (9 patients)	Multiple species	Pyophage cocktail	[43]
19	Musculoskeletal infections, case series (4 patients)	*S. aureus*, *P. aeruginosa, E. faecalis*	BFC-1 cocktail (see above Id #16) + Pyophage cocktail	[44••]
20	Lung transplant, case series (3 patients)	*P. aeruginosa*, *B. dolosa*	AB-SA01 (see above Id #16) + Navy phage cocktail (Five different *Pseudomonas* phages PaΦ1, PaSLWΦ17, PaSKWΦ22, PaATFΦ1, and PaATFΦ3 – all sequenced, no explicit design rules provided) + single *Burkholderia* phage BdPF16phi4281	[45]
21	Cardiothoracic surgery infections, case series (8 patients)	Multiple species	Multiple combinations of phages from the Gabrichevsky Institute. All phage sequenced and previously studied. No explicit design rules provided.	[46]
22	Prosthetic knee infection by *S. aureus*, case series (3 patients)	*S. aureus*	Three *Staphylococcus* phages (PP1493, PP1815, PP1957 from *Silviavirus* and *Rosenblumvirus*). Indicated to have been selected for their complementarity of host range.	[47]

Figure 1



Principles of applying association rules mining to host-range data.
1) Host-range data is collected on a panel of strains and a collection of phages (e.g., spoting assay of phage dilution series). 2) Productive infections are tabulated as binary values. 3) The host range data is converted into transactions, which converts each line in the host range dataset into a list of phages for which a productive infection was observed. This list of transaction is fed to the datamining algorithm which infers the set of rules observed in the dataset and attaches to each quality metrics such as confidence, support, and lift. 4) The rules can be sorted by the quality metrics and visualized graphically. This enables the researcher to browse through the 'top rules' in terms of the quality metrics to explore the dataset and underlying patterns.

3 Each row in the dataset is converted in a transaction and we use a software package such as *arules* via RStudio [50] to mine the set of rules (available software reviewed in Ref. [51]).
4 Multiple visualization interfaces are available to explore the set of rules. Typically, the rules can be filtered and sorted by their quality metrics and finally visualized as a network where one can see the rules antecedents and consequents.

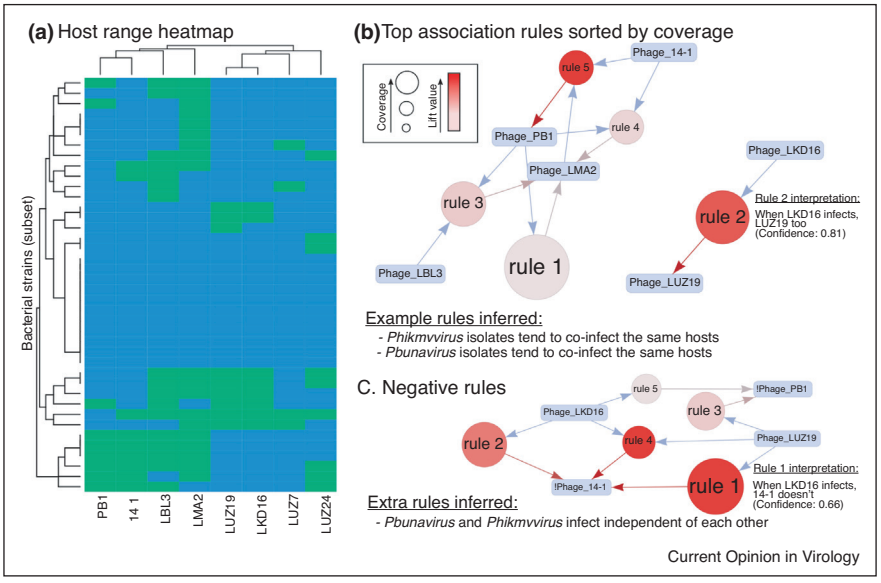
Two key metrics are typically consulted when filtering and inspecting the rules, i) the support for the rule A -> B, indicates how frequently the rule appears in the dataset, and ii) the confidence for a given rule A -> B, which is the frequency in which B is found with A relative to the frequency of A. Taken together, support and confidence relay how interesting rules are and one can set threshold to these values in order to filter down the number of rules to consider. Other common metrics are the so-called coverage and lift values, which represent measure of how often a rule can be applied, and how important the rule is, respectively. Other techniques to cluster rules and uncover unexpected patterns have also been proposed [52•].

Example application of association rules mining to *P. aeruginosa* and its phages

To illustrate the use of association rules mining, we applied the technique to a dataset of 579 *P. aeruginosa* strains of various environmental and clinical origins. All of the strains were tested against eight distinct phages, and we recorded whether individual plaques appeared in the dilution series of the phage, indicating a productive phage infection [53]. The other phenotypes, such a 'lysis from without' were ignored for this analysis. The dataset and the scripts used for the analysis are available in the supplementary material.

The original host range data was recorded in a spreadsheet software (579 lines by eight columns) and exported as a list of transactions that could be further processed using the R studio programming environment [54]. We used the *apriori* and *plot* functions available in the R package *arules* and *arulesviz* to generate the initial set of rules and explore it interactively as a graph network [50,55]. Multiple quality metrics are available for each rule, and we opted to sort the rules by their confidence and their support. Although no genomics information is necessary for this analysis, knowledge of the phage isolates from our previous work helped in the *a posteriori* interpretation of the network of rules. Indeed, we saw that *Pseudomonas* phage 14-1, PB1, LMA2, and LBL3 appeared together as antecedent and consequent of many inferred rules, and similarly for *Pseudomonas* phages LKD16 and LUZ19. Given that the phages from the first cluster all belong to the *Pbunavirus*, and the phages from the second cluster belong to the *Phikmvvirus* genus, we interpreted this as a clear signal that taxonomy position is

Figure 2



Application of association rules mining to a dataset of 550 *Pseudomonas aeruginosa* strains.
(a) Heatmap of the susceptibility of a subset of 50 bacteria to eight phages (in green = susceptible, in blue = no interactions). **(b)** Graphical exploration of positive association rules, sorted by confidence. The phage isolates are indicated in blue. Each circle represents a rule, and has one or more antecedent (incoming arrows) and a single consequent (outgoing arrow). **(c)** We look here at negative association rules, sorted by confidence.

an important predictor of infectivity. A finding previously highlighted in the study by Gencay *et al.* [56].

We extended the analysis by including a search for negative association rules in the host range dataset (supplementary material). One of the key results that we observed was the existence of exclusions between *Pbunavirus* and *Phikmvvirus* members. This is of particular interest as it indicates that these phages have complementary host ranges, and as such could be interesting to pool together into a phage cocktail which would have an expanded host range compared to both phages individually. Aside from that consideration, isolates from these two genera are known to use different phage receptors (LPS and Type IV pili respectively), meaning that this approach could also be used during large screens to selected phages on which to perform more in depth molecular characterization techniques (including omics) (Figure 2).

Potential, limitations, and outlook

There is a growing interest in applying phages to clear bacterial infections and mitigate the growing concerns and costs of antibiotic resistance. A recurrent work around to the limited host range of phage isolates is to pool together phages into cocktail products that have a larger host range than its individual components. As such, one may expect that some ground rules could be used to decide which phages to combine. However, as our survey of the literature of clinical trials and case series has revealed, there is room to improve and guide design decisions.

In this opinion piece, we proposed to implement a datamining technique known as 'association rules mining' as a strategy to unpack rules for the design of phage cocktails. The approach relies on the existence of host range datasets and does not necessitate *a priori* genomics data of either phages or bacteria. Using a dataset of *P. aeruginosa* and its phages, we showed that two types of questions relevant to cocktail design can be asked: i) do some phages appear redundant in terms of their host range and as such would not expand the cocktail's host range? and ii) are some phage isolates complementary in host ranges? If that should be the case, then having these phages in the cocktail could expand its host range.

As with other data-driven approaches, the quantity and quality of the input data is critical to yield statistically relevant insights. We look forward to the expansion of therapeutic phage libraries and representative panels of strains as a source of data to discover these patterns, but also envision it could be implemented on top of existing databases such as the Viral Host Range as a mean to explore existing datasets [57••]. An obvious limitation of the approach is that it is oblivious to the molecular details

about the phage and the bacteria, such as the receptors, and focuses only on the outcome of the infection cycle. Bacteria mutants can occur that have modified receptors or even cross-sensitization to other phages [58,59], and phages may also have activity against biofilms such as those found in prosthetic infections or endocarditis [60]. However, these limitations are balanced by the broad applicability of the method, going beyond the medical setting strictly. In general, it can be used as a first step to exploring host range data collected about new phage isolates, potentially guiding the researcher in narrowing down large collections of new phage isolates to characterize further. For instance, if the phages are candidate to enter the composition of a therapeutic cocktail, then genome sequencing would be critical in combination with microbiological studies such as stability, receptors, and detection of potential lysogenic behaviour. This approach is also compatible and adjunct to other strategies, such as the machine learning approach we have recently proposed to generate digital phagograms [61], and it could be extended in the future to include antibiotics-phage synergies.

Conflict of interest statement

Nothing declared.

Acknowledgements

CL is supported by the Research Foundation-Flanders (FWO) SB grants 1S64720N. RL and VvN are supported by the ID-N grant PHAGEFORCE from the KU Leuven (IDN/20/024). CL thanks Dr Ana Lechuga Mateo for her help with the illustrations and critical discussions.

The authors thank Lisette Scheepmaker at the UMC for the excellent technical assistance in collecting phage host range data

The research leading to these results has received support from the IMI Joint Undertaking under grant agreement no. 115721-2, resources of which are composed of financial contribution from the European Union's Seventh Framework Program [FP7/2007-2013] and EFPIA companies in kind contribution.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.coviro.2021.12.011.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Twort FW: **An investigation on the nature of ultra-microscopic viruses**. *Lancet* 1915, **186**:1241-1243.

2. d'Hérelle Félix: **On an invisible microbe antagonistic to dysentery bacilli**. *C R Acad Sci* 1917, **165**:373-375.

3. Summers WC: **The strange history of phage therapy**. *Bacteriophage* 2012, **2**:130.

4. •• McCallin S, Brüssow H: **Clinical trials of bacteriophage therapeutics**. *Bacteriophages* 2021:1099-1127 http://dx.doi.org/10.1007/978-3-319-41986-2_38

This is a recent and comprehensive review of all clinical trials that include phage products.

5. Pirnay J-P, Kutter E: **Bacteriophages: it's a medicine, Jim, but not as we know it**. *Lancet Infect Dis* 2021, **21**:309-311.

6. El Haddad L, Harb CP, Gebara MA, Stibich MA, Chemaly RF: **A systematic and critical review of bacteriophage therapy against multidrug-resistant ESKAPE organisms in humans**. *Clin Infect Dis* 2019, **69**:167-178.

7. •• Onsea J, Uyttebroek S, Chen B, Wagemans J, Lood C, Van Gerven L, Spriet I, Devolder D, Debaveye Y, Depypere M *et al.*: **Bacteriophage therapy for difficult-to-treat infections: the implementation of a multidisciplinary phage task force (the PHAGEFORCE study protocol)**. *Viruses* 2021, **13**:1543

In this protocol paper, the authors detail the streamlined procedure used to select and treat patients with phage cocktails products. They highlight the importance of assessing phage synergies and methods used to populate their phage banks with new isolates.

8. • Pirnay J-P, Verbeken G, Ceyssens P-J, Huys I, De Vos D, Ameloot C, Fauconnier A: **The magistral phage**. *Viruses* 2018, **10**:64

Details the approach of validating and using phage products within the operational framework of magistral preparation (or compound products).

9. Pirnay J-P, Blasdel BG, Bretaudeau L, Buckling A, Chanishvili N, Clark JR, Corte-Real S, Debarbieux L, Dublanchet A, De Vos D *et al.*: **Quality and safety requirements for sustainable phage therapy products**. *Pharm Res* 2015, **32**:2173-2179.

10. Ross A, Ward S, Hyman P: **More is better: selecting for broad host range bacteriophages**. *Front Microbiol* 2016, **7**:1352.

11. •• de Jonge PA, Nobrega FL, Brouns SJJ, Dutilh BE: **Molecular and evolutionary determinants of bacteriophage host range**. *Trends Microbiol* 2019, **27**:51-63

In this review, the authors review the factors influencing the host range of bacteriophages, with an eye towards phage therapy applications.

12. Peng C, Hanawa T, Azam AH, LeBlanc C, Ung P, Matsuda T, Onishi H, Miyanaga K, Tanji Y: **Silviavirus phage ΦMR003 displays a broad host range against methicillin-resistant *Staphylococcus aureus* of human origin**. *Appl Microbiol Biotechnol* 2019, **103**:7751-7765.

13. McInerney JO, McNally A, O'Connell MJ: **Why prokaryotes have pangenomes**. *Nat Microbiol* 2017, **2**:1-5.

14. Freschi L, Vincent AT, Jeukens J, Emond-Rheault J-G, Kukavica-Ibrulj I, Dupont M-J, Charette SJ, Boyle B, Levesque RC: **The *Pseudomonas aeruginosa* pan-genome provides new insights on its population structure, horizontal gene transfer, and pathogenicity**. *Genome Biol Evol* 2019, **11**:109-120.

15. De Smet J, Hendrix H, Blasdel BG, Danis-Wlodarczyk K, Lavigne R: **Pseudomonas predators: understanding and exploiting phage–host interactions**. *Nat Rev Microbiol* 2017, **15**:517-530.

16. Lood C, Danis-Wlodarczyk K, Blasdel BG, Jang HB, Vandenheuvel D, Briers Y, Noben J-P, van Noort V, Drulis-Kawa Z, Lavigne R: **Integrative omics analysis of *Pseudomonas aeruginosa* virus PA5oct highlights the molecular complexity of jumbo phages**. *Environ Microbiol* 2020, **22**:2165-2181.

17. Knezevic P, Voet M, Lavigne R: **Prevalence of Pf1-like (pro) phage genetic elements among *Pseudomonas aeruginosa* isolates**. *Virology* 2015, **483**:64-71.

18. Yang Y, Lu S, Shen W, Zhao X, Shen M, Tan Y, Li G, Li M, Wang J, Hu F *et al.*: **Characterization of the first double-stranded RNA bacteriophage infecting *Pseudomonas aeruginosa***. *Sci Rep* 2016, **6**:38795.

19. Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM, Brister JR, Kropinski AM, Krupovic M, Lavigne R *et al.*: **Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks**. *Nat Biotechnol* 2019, **37**:632-639.

20. Chan BK, Abedon ST, Loc-Carrillo C: **Phage cocktails and the future of phage therapy**. *Future Microbiol* 2013, **8**:769-783.

21. Molina F, Simancas A, Ramírez M, Tabla R, Roa I, Rebollo JE: **A new pipeline for designing phage cocktails based on phage-bacteria infection networks**. *Front Microbiol* 2021, **12**:149.

22. •• Lin RC, Sacher JC, Ceyssens P-J, Zheng J, Khalid A, Iredell JR: **Australian phage biobanking network: phage biobank: present challenges and future perspectives**. *Curr Opin Biotechnol* 2021, **68**:221-230

In this opinion, the authors detail the desirable features of establishing sustainable phage collections. With their focus on reproducibility and centralization, phage banks are poised to be an important source of data for analyses such as the one proposed here.

23. Liu D, Van Belleghem JD, de Vries CR, Burgener E, Chen Q, Manasherob R, Aronson JR, Amanatullah DF, Tamma PD, Suh GA: **The safety and toxicity of phage therapy: a review of animal and clinical studies**. *Viruses* 2021, **13**:1268.

24. • McCallin S, Sarker SA, Sultana S, Oechslin F, Brüssow H: **Metagenome analysis of Russian and Georgian Pyophage cocktails and a placebo-controlled safety trial of single phage versus phage cocktail in healthy *Staphylococcus aureus* carriers**. *Environ Microbiol* 2018, **20**:3278-3293

Detailed metagenomics analysis to uncover the composition of two traditional phage cocktails (Pyophage) commonly used but of unknown (phage isolates) composition.

25. McCallin S, Alam Sarker S, Barretto C, Sultana S, Berger B, Huq S, Krause L, Bibiloni R, Schmitt B, Reuteler G *et al.*: **Safety analysis of a Russian phage cocktail: from MetaGenomic analysis to oral application in healthy human subjects**. *Virology* 2013, **443**:187-196.

26. Wright A, Hawkins CH, Anggård EE, Harper DR: **A controlled clinical trial of a therapeutic bacteriophage preparation in chronic otitis due to antibiotic-resistant *Pseudomonas aeruginosa*; a preliminary report of efficacy**. *Clin Otolaryngol* 2009, **34**:349-357.

27. Rhoads DD, Wolcott RD, Kuskowski MA, Wolcott BM, Ward LS, Sulakvelidze A: **Bacteriophage therapy of venous leg ulcers in humans: results of a phase I safety trial**. *J Wound Care* 2009, **18**:237-238 240–243.

28. Sarker SA, McCallin S, Barretto C, Berger B, Pittet A-C, Sultana S, Krause L, Huq S, Bibiloni R, Bruttin A *et al.*: **Oral T4-like phage cocktail application to healthy adult volunteers from Bangladesh**. *Virology* 2012, **434**:222-232.

29. Sarker SA, Sultana S, Reuteler G, Moine D, Descombes P, Charton F, Bourdin G, McCallin S, Ngom-Bru C, Neville T *et al.*: **Oral phage therapy of acute bacterial diarrhea with two coliphage preparations: a randomized trial in children from Bangladesh**. *EBioMedicine* 2016, **4**:124-137.

30. Sarker SA, Berger B, Deng Y, Kieser S, Foata F, Moine D, Descombes P, Sultana S, Huq S, Bardhan PK *et al.*: **Oral application of *Escherichia coli* bacteriophage: safety tests in healthy and diarrheal children from Bangladesh**. *Environ Microbiol* 2017, **19**:237-250.

31. Gindin M, Febvre HP, Rao S, Wallace TC, Weir TL: **Bacteriophage for Gastrointestinal Health (PHAGE) study: evaluating the safety and tolerability of supplemental bacteriophage consumption**. *J Am Coll Nutr* 2019, **38**:68-75.

32. Oechslin F, Piccardi P, Mancini S, Gabard J, Moreillon P, Entenza JM, Resch G, Que Y-A: **Synergistic interaction between phage therapy and antibiotics clears *Pseudomonas aeruginosa* infection in endocarditis and reduces virulence**. *J Infect Dis* 2017, **215**:703-712.

33. Jault P, Leclerc T, Jennes S, Pirnay JP, Que Y-A, Resch G, Rousseau AF, Ravat F, Carsin H, Floch RL *et al.*: **Efficacy and tolerability of a cocktail of bacteriophages to treat burn wounds infected by *Pseudomonas aeruginosa* (PhagoBurn): a randomised, controlled, double-blind phase 1/2 trial**. *Lancet Infect Dis* 2019, **19**:35-45.

34. Febvre HP, Rao S, Gindin M, Goodwin NDM, Finer E, Vivanco JS, Lu S, Manter DK, Wallace TC, Weir TL: **PHAGE study: effects of supplemental bacteriophage intake on inflammation and gut microbiota in healthy adults**. *Nutrients* 2019, **11**:666.

35. Lehman SM, Mearns G, Rankin D, Cole RA, Smrekar F, Branston SD, Morales S: **design and preclinical development of a phage product for the treatment of antibiotic-resistant *Staphylococcus aureus* infections**. *Viruses* 2019, **11**:E88.

36. Ooi ML, Drilling AJ, Morales S, Fong S, Moraitis S, Macias-Valle L, Vreugde S, Psaltis AJ, Wormald P-J: **Safety and tolerability of bacteriophage therapy for chronic rhinosinusitis due to *Staphylococcus aureus***. *JAMA Otolaryngol Head Neck Surg* 2019, **145**:723-729.

37. • Petrovic Fabijan A, Lin RCY, Ho J, Maddocks S, Ben Zakour NL, Iredell JR: **Safety of bacteriophage therapy in severe *Staphylococcus aureus* infection**. *Nat Microbiol* 2020, **5**:465-472

Detailed report of a single-arm, non-comparative clinical trial of patients with severe *Staphylococcus aureus* infections treated intravenously with phage cocktail AB-SA01.

38. Grubb DS, Wrigley SD, Freedman KE, Wei Y, Vazquez AR, Trotter RE, Wallace TC, Johnson SA, Weir TL: **PHAGE-2 study: supplemental bacteriophages extend *Bifidobacterium animalis* subsp. lactis BL04 benefits on gut health and microbiota in healthy adults**. *Nutrients* 2020, **12**:E2474.

39. • Leitner L, Ujmajuridze A, Chanishvili N, Goderdzishvili M, Chkonia I, Rigvava S, Chkhotua A, Changashvili G, McCallin S, Schneider MP *et al.*: **Intravesical bacteriophages for treating urinary tract infections in patients undergoing transurethral resection of the prostate: a randomised, placebo-controlled, double-blind clinical trial**. *Lancet Infect Dis* 2021, **21**:427-436

An important and recent example of a large randomized, double-blind, and placebo-controlled clinical trial of a phage product (Pyophage) to treat urinary tract infections.

40. Letkiewicz S, Mie?dzybrodzki R, Fortuna W, Weber-Da? browska B, Górski A: **Eradication of *Enterococcus faecalis* by phage therapy in chronic bacterial prostatitis — case report**. *Folia Microbiol* 2009, **54**:457.

41. Rose T, Verbeken G, Vos DD, Merabishvili M, Vaneechoutte M, Lavigne R, Jennes S, Zizi M, Pirnay J-P: **Experimental phage therapy of burn wound infection: difficult first steps**. *Int J Burns Trauma* 2014, **4**:66-73.

42. Fish R, Kutter E, Wheat G, Blasdel B, Kutateladze M, Kuhl S: **Compassionate use of bacteriophage therapy for foot ulcer treatment as an effective step for moving toward clinical trials**. *Methods Mol Biol* 2018, **1693**:159-170.

43. Ujmajuridze A, Chanishvili N, Goderdzishvili M, Leitner L, Mehnert U, Chkhotua A, Kessler TM, Sybesma W: **Adapted bacteriophages for treating urinary tract infections**. *Front Microbiol* 2018, **9**:1832.

44. •• Onsea J, Soentjens P, Djebara S, Merabishvili M, Depypere M, Spriet I, De Munter P, Debaveye Y, Nijs S, Vanderschot P *et al.*: **Bacteriophage application for difficult-to-treat musculoskeletal infections: development of a standardized multidisciplinary treatment protocol**. *Viruses* 2019, **11**:891

Detailed case series report with standardized treatment and application guidelines towards the treatment of patients with severe musculoskeletal infections using the phage cocktail BFC-1.

45. Aslam S, Courtwright AM, Koval C, Lehman SM, Morales S, Furr C-LL, Rosas F, Brownstein MJ, Fackler JR, Sisson BM *et al.*: **Early clinical experience of bacteriophage therapy in 3 lung transplant recipients**. *Am J Transplant* 2019, **19**:2631-2639.

46. Rubalskii E, Ruemke S, Salmoukas C, Boyle EC, Warnecke G, Tudorache I, Shrestha M, Schmitto JD, Martens A, Rojas SV *et al.*: **Bacteriophage therapy for critical infections related to cardiothoracic surgery**. *Antibiotics (Basel)* 2020, **9**:E232.

47. Ferry T, Kolenda C, Batailler C, Gustave C-A, Lustig S, Malatray M, Fevre C, Josse J, Petitjean C, Chidiac C *et al.*: **Phage therapy as adjuvant to conservative surgery and antibiotics to salvage patients with relapsing *S. aureus* prosthetic knee infection**. *Front Med (Lausanne)* 2020, **7**:570572.

48. Agrawal R, Imieliński T, Swami A: **Mining association rules between sets of items in large databases**. In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*. 1993.

49. Antonie L, Li J, Zaiane O: **Negative association rules**. In *Frequent Pattern Mining*. Edited by Aggarwal CC, Han J. Springer International Publishing; 2014:135-145.

50. Hahsler M, Chelluboina S, Hornik K, Buchta C: **The arules R-package ecosystem: analyzing interesting patterns from large transaction data sets**. *J Mach Learn Res* 2011, **12**:2021-2025.

51. Naulaerts S, Meysman P, Bittremieux W, Vu TN, Vanden Berghe W, Goethals B, Laukens K: **A primer to frequent itemset mining for bioinformatics**. *Brief Bioinform* 2015, **16**:216-231.

52. • Danh B-T, Meysman P, Laukens K: **Clustering association rules to build beliefs and discover unexpected patterns**. *Appl Intell* 2020, **50**:1943-1954
An example of potential extension to the analysis proposed here that could help to further explore association patterns in host range data.

53. Ács N, Gambino M, Brøndsted L: **Bacteriophage enumeration and detection methods**. *Front Microbiol* 2020, **11**:594868.

54. R development Core Team: *R: A Language and Environment for Statistical Computing*. 2014.

55. Hahsler M, Grün B, Hornik K: **arules - a computational environment for mining association rules and frequent item sets**. *J Statist Softw* 2005, **14**:1-25.

56. Gencay YE, Gambino M, Prüssing TF, Brøndsted L: **The genera of bacteriophages and their receptors are the major determinants of host range**. *Environ Microbiol* 2019, **21**:2095-2111.

57. •• Lamy-Besnier Q, Brancotte B, Ménager H, Debarbieux L: **Viral host range database, an online tool for recording, analyzing and disseminating virus-host interactions**. *Bioinformatics* 2021, **37**:2798-2801 http://dx.doi.org/10.1093/bioinformatics/btab070
Web application that researchers can use to publish their host range data in a structured way. This represents an important effort towards standardizing the reporting of host range data outside of (supplementary) tables of publications and in a way propitious to the type of automation proposed in this paper.

58. Markwitz P, Lood C, Olszak T, van Noort V, Lavigne R, Drulis-Kawa Z: **Genome-driven elucidation of phage-host interplay and impact of phage resistance evolution on bacterial fitness**. *ISME J* 2021:1-10 http://dx.doi.org/10.1038/s41396-021-01096-5.

59. Majkowska-Skrobek G, Markwitz P, Sosnowska E, Lood C, Lavigne R, Drulis-Kawa Z: **The evolutionary trade-offs in phage-resistant *Klebsiella pneumoniae* entail cross-phage sensitization and loss of multidrug resistance**. *Environ Microbiol* 2021 http://dx.doi.org/10.1111/1462-2920.15476.

60. Ferriol-González C, Domingo-Calap P: **Phages for biofilm removal**. *Antibiotics* 2020, **9**:268.

61. Lood C, Haas P-J, van Noort V, Lavigne R: **Digital phagograms: predicting phage infectivity through a multilayer machine learning approach**. *Curr Opin Virol* 2022, **52**:174-181 http://dx.doi.org/10.1016/j.coviro.2021.12.004 In this issue.