Deciphering the role of phage in the cystic fibrosis airway

Dana Willner* and Mike Furlan
San Diego State University; Department of Biology; San Diego, CA USA

Cystic fibrosis (CF) is a fatal genetic disorder hallmarked by chronic and persistent microbial infections of the lungs and airways. Much attention has been paid to describing microbial communities and microbial pathogenesis in CF, however, viral communities have been largely ignored. We recently published a metagenomic study characterizing viral communities in the sputum of CF and Non-CF individuals for the first time. There was a striking difference in metabolic functions encoded by phage in CF versus Non-CF individuals. Regardless of which viral taxa were present, CF-associated phage shared a common core metabolism that reflected the disease state and aberrant airway physiology. Here, this finding is discussed further and its implications for the role of phage and the nature of phage-microbe interactions in the CF airway are explored.

Introduction

Cystic fibrosis (CF) is a lethal autosomal recessive genetic disorder caused by a mutation in the cystic fibrosis transmembrane regulator (CFTR) gene. Respiratory disease due to chronic and persistent microbial infections is the leading cause of morbidity and mortality in CF, as microbial colonization of the lungs and airways in CF leads to an innate immune response, extensive airway remodeling, and subsequent respiratory failure and death. Recently, we published a metagenomic study characterizing DNA viral communities in sputum from 5 CF and 5 Non-CF individuals. All study subjects were adults between the ages of 20 and 50. These communities were largely comprised of phage, and our results indicated that phage communities in CF and Non-CF individuals differed both functionally and taxonomically. Regardless of which phage taxa were present, phage metabolisms were dramatically different between the healthy and diseased states. Phage metabolic profiles in CF individuals reflected the aberrant physiology of the CF airway, as well as the metabolic states of host microbes.

CF-Associated Phage Share a Core Metabolic Profile

To further explore the core metabolism of CF-associated phage, a data set containing sequences both common and unique to the CF viromes was constructed, i.e., these sequences appeared in all CF viromes but in none of the Non-CF viromes. A second data set containing sequences common and unique to all Non-CF viromes was constructed for comparison. The data sets were generated using a sequential BLAST method. Two metagenomes from the same set were randomly selected (CF or Non-CF) and compared to each other using BLASTn. Sequences were considered to be shared if they were at least 98% similar. Shared sequences were then compared to a third randomly selected metagenome to identify a set of sequences common to all three metagenomes. This was repeated for the rest of the metagenomes in the set. The sequences common to each set of metagenomes were then compared to each metagenome in the

Key words: cystic fibrosis, phage, metagenomics, metabolic subsystems, top-down control

Submitted: 03/19/10
Revised: 04/14/10
Accepted: 04/15/10
Previously published online: www.landesbioscience.com/journals/virulence/article/12071

*Correspondence to: Dana Willner; Email: willner9@aol.com
other set to find unique shared sequences, i.e., the set of shared CF sequences were compared to each Non-CF metagenome individually and any common sequences were removed at each step, and vice versa. The CF viromes shared a unique set of 16,059 sequences (6.54% of total CF metagenomic sequences), while the Non-CF viromes had only 1,651 common sequences (0.12% of total Non-CF sequences). In accordance with prior results, this suggests that CF-associated metabolisms are being driven by common environmental factors.

The unique CF sequences had a distinct metabolic profile, while the small set of Non-CF shared sequences were evenly scattered across subsystem categories (Fig. 1A). Sequences were annotated using the MG-RAST server (e-value <10^-5), which assigns similarities to metabolic pathways, or subsystems, based on BLASTx similarities. The unique CF data set was enriched for two top-level metabolic pathways: carbohydrates and amino acids and derivatives (35% of all annotated sequences). Carbohydrates and amino acids have previously been shown to be the most abundant subsystems in both microbiomes and viromes in general. This suggests that there is a specific and well-defined core metabolic profile common to CF-associated viruses. The diabetic status of CF individuals used in this study is unknown, however, the prevalence of functions related to carbohydrate metabolism could represent viral and microbial adaptations to diabetic hosts. Within the amino acid subsystem, 6% of the unique CF sequences corresponded to aromatic amino acid biosynthesis pathways, while no sequences related to aromatic amino acids were found in the Non-CF sequence set. Many of the enzymes involved in aromatic amino acid biosynthesis are also utilized in the biosynthesis of the Pseudomonas quinolone signal (PQS) in *P. aeruginosa*, which is regarded as the most abundant microbe in the CF airway. Specifically, alginate metabolism and rhamnose synthesis. Alginate and rhamnose are thought to be major components of biofilms produced by CF-associated microbes, and rhamnolipids may play a significant role in evasion of host immune mechanisms. Annotations of sequences in the virulence subsystem identified similarities to many proteins commonly associated with antibiotic resistance including β-lactamases, multi-drug efflux pumps, and fluoroquinolone resistant topoisomerases. Iron-scavenging mechanisms such as pyoverdine biosynthesis were also detected as well as components of Type III secretion systems, which are associated with *P. aeruginosa* resistance and toxicity to host immune cells.

**CF-Associated Phage Reflect Core Microbial Metabolisms**

To compare the unique CF core metabolic profile to the core profile of *P. aeruginosa*, a dataset containing 3,484 core genes shared by four strains of *P. aeruginosa* was constructed by sequential BLAST as described above, using gene nucleotide sequences at the third (i.e., most specific) subsystem classification level using the MG-RAST heat map tool. The annotated *P. aeruginosa* core gene sequences were then compared to the annotated CF and Non-CF unique sequences at the third (i.e., most specific) subsystem classification level using the MG-RAST heat map tool.

All of the subsystems found in the Non-CF sequence set were shared with both *P. aeruginosa* and the unique CF sequences. The majority of subsystems (59%) identified in the *P. aeruginosa* and unique CF metabolisms were shared, including pathways for degradation of aromatic compounds and exopolysaccharide biosynthesis. The *P. aeruginosa* core contained 91 unique subsystems, which consisted of many regulatory elements and the complete phenazine biosynthesis pathway. There were 122 subsystems found in the unique CF virome which could not be identified in the *P. aeruginosa* data set, many of which were related to carbohydrate utilization.

**Implications for the Role of Phage in the CF Airway**

The CF phage core metabolic profile reflected host adaptations to the unique environment of the CF airway. CF-associated microbes have been shown to grow in biofilms, and are characterized by adaptations which allow evasion of the host immune system, antibiotic resistance, and improved fitness in anaerobic environments. While the majority of these adaptations have been observed in studies of *P. aeruginosa*, it is likely that they are shared by other microbes in the CF airway. Genes associated with PQS biosynthesis were detected in the core phage metabolism. Palmer et al. demonstrated using transcriptomics that genes related to PQS are upregulated during *P. aeruginosa* growth in CF sputum. PQS has been implicated in the regulation of virulence factors and biofilm formation in *P. aeruginosa*. The iron-scavenging mechanisms detected in the core phage metabolism also relate indirectly to biofilm formation, as biofilm production in *P. aeruginosa* is disrupted by iron limitation, especially in anaerobic conditions. Phage encoded genes for iron uptake would provide a selective advantage to their hosts in the anaerobic microenvironments of the CF airway. Functions related to synthesis of rhamnose, a biofilm structural component, were also detected. Rhamnose is also a component of rhamnolipids, which induce rapid necrosis of polymorphonuclear neutrophilic leukocytes (PMNs), allowing evasion of the host immune system.
Figure 1. Metabolic profile of the unique CF and non-CF core metagenomic sequences (A), and the Pseudomonas aeruginosa core (B). The percentage of best BLASTx (e-value < 10^-5) similarities to the top level subsystems in the SEED database as a percentage of the total number of metagenomic sequences are presented.
accessory genome.\textsuperscript{13-18} This accessory genome contains genes for antibiotic resistance as well as novel functions allowing rapid and efficient adaptation to growth in the respiratory tract, many of which are encoded in labile prophage elements.\textsuperscript{13-15,18} Antibiotics induce prophage in \textit{S. aureus} and most likely in other microbes, and may heighten the ability of phage to act as major agents of gene flow in CF microbial infections.\textsuperscript{13,19} Phage also have been shown to modulate phenotypic differences in \textit{P. aeruginosa} colonies and associated biofilms, leading to the emergence of highly virulent small colony variants.\textsuperscript{20} Comparison of the CF phage metabolic profile with the \textit{P. aeruginosa} core metabolism supports these previous findings. Our results suggest that CF-associated phage carry a critical set of functions derived from core host genes, while also providing an additional set of functions which hosts can incorporate through lateral gene transfer, thus leading to host diversification. Phage can also drive host diversity and control microbial populations through predation.\textsuperscript{21} Top-down control of microbes by phage has been demonstrated in aquatic and terrestrial environments, but the role of indigenous phage in controlling microbial populations in the CF lung is unknown.\textsuperscript{22,23} The exopolysaccharide layer surrounding microbial biofilms in the CF airway has been suggested as a significant barrier to phage infection, however, phage have adapted mechanisms to decrease polysaccharide viscosity, allowing for predation.\textsuperscript{24,25} In marine systems, protozoan grazing serves as a second significant mechanism of top-down control.\textsuperscript{26,27} PMNs and clinically administered antibiotics can be thought of as analogues to protozoan grazers in the CF airway. By providing genes which aid in biofilm formation, antibiotic resistance, and avoidance of PMNs as seen in the CF phage core metabolic profile, phage could ensure exclusive access to their susceptible hosts. Microbes are also subject to nutrient limitations (i.e., bottom-up controls), and adaptations which allow escape from phage predation may incur a metabolic cost or force microbes into novel metabolic states. For example, bacteria may downregulate or even abolish expression of specific cell surface receptors leading to a decrease in phage adsorption. The intended ligand of the receptor may be a specific carbon source or other nutrient, thus in gaining phage resistance, bacteria may shift their metabolic profiles.

The results of the original study and the core CF phage metabolic profile presented here indicate the importance of microbial-encoded genes in microbial adaptations to the CF airway. Phage metabolisms reflect essential host metabolisms and suggest that phage drive microbial diversification through lateral gene transfer and predation as observed in other environments. However, the nature of phage-host interactions remains largely unknown. Phage and microbes in natural systems often follow an oscillatory “Kill-the-Winner” model (Fig. 2). In this model, an increase in abundance of a host population is followed by a corresponding increase in the abundance of its phage predators, leading to increased killing of the successful host (the “winner”).\textsuperscript{28} This allows another host population to rise in abundance and continue the cycle, and also maintains less dominant hosts in the community.\textsuperscript{28} “Kill-the-Winner” dynamics have been shown to operate at the strain level in aquatic environments, thus maintaining a dynamic equilibrium where hosts do not change at the species level.\textsuperscript{29} Similar dynamics might be expected to operate in the CF airway where a small number of species are thought to dominate, and intra-species diversity is high.\textsuperscript{30}

\textbf{Figure 2.} Kill-the-Winner dynamics in phage-host systems. An increase in host population (Host 1) is followed by an increase in the population of its corresponding phage predator (Phage 1), resulting in the most abundant host being killed by its phage. This leaves a vacancy for a new host to become highly abundant (Host 2), and this new dominant host will subsequently be killed by its predator (Phage 2). These pairs continue to oscillate over time, with phage-host pairs blooming independently from each other.
12. O'May CY, Sanderson K, Roddam LF, Kirov SM, Reid DW. Iron-binding compounds impair *Pseudomonas aeruginosa* biofilm formation, especially under anaerobic conditions. J Med Microbiol 2009; 58:765-73.

13. Mathee K, Narasimhan G, Valdes C, Qu X, Matewish JM, Koehren M, et al. Dynamics of *Pseudomonas aeruginosa* genome evolution. Proc Nat Acad Sci USA 2008; 105:3100-5.

14. Winstanley C, Langille MGJ, Forthergill JL, Kukavica-Ibrulj I, Paradis-Beau C, Sanschagrin F, et al. Newly introduced genomic prophage islands are critical determinants of in vivo competitiveness in the Liverpool Epidemic Strain of *Pseudomonas aeruginosa*. Genome Res 2009; 19:12-25.

15. Rolain J, Franquinet P, Hernandez D, Birzar F, Richet H, Fournous G, et al. Genomic analysis of an emerging multiresistant *Staphylococcus aureus* strain rapidly spreading in cystic fibrosis patients revealed the presence of an antibiotic inducible bacteriophage. Biol Direct 2009; 4:1.

16. Holden MTG, Lindsay JA, Corron C, Quail MA, Cockfield JD, Parkh S, et al. Genome sequence of a recently emerged, highly transmissible, multi-antibiotic- and antiseptic-resistant variant of methillin-resistant *Staphylococcus aureus*, sequence type 239 (TW). J Bacteriol 2010; 192:888-92.

17. Stover CK, Pham QX, Erwin AL, Mizoguchi SD, Warriner P, Hickey MJ, et al. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature 2000; 406:959-64.

18. Shen K, Sayeed S, Antalis P, Gladitz J, Ahmed A, Dice B, et al. Extensive genomic plasticity in *Pseudomonas aeruginosa* revealed by identification and distribution studies of novel genes among clinical isolates. Infect Immun 2006; 74:5272-83.

19. Goerke C, Koller J, Wolz C. Gipsyfanacin and trimethoprism cause phage induction and virulence modulation in *Staphylococcus aureus*. Antimicrob Agents Chemother 2006; 50:171-7.

20. Webb JS, Lau M, Kjelleberg S. Bacteriophage and Phenotypic Variation in *Pseudomonas aeruginosa* Biofilm Development. J Bacteriol 2004; 186:4066-73.

21. Rodriguez-Valera F, Martin-Cuadrado A, Rodriguez-Brito B, Pasic L, Thingstad TF, Rohwer F, Mira A. Explaining microbial population genomics through phage predation. Nat Rev Micro 2009; 7:828-36.

22. Allen B, Willner D, Oechel WC, Lipson D. Top-down control of microbial activity and biomass in an Arctic soil ecosystem [Internet]. Environ Microbiol 2009; 11:940-52 Available from: http://www.ncbi.nlm.nih.gov/pubmed/19002186.

23. Fuhrman JA, Schwalbach M. Viral influence on aquatic bacterial communities. Biol Bull 2003; 204:192-5.

24. Hanlon GW, Denyer SP, Olliff CJ, Ibrahim LJ. Reduction in exopolysaccharide viscosity as an aid to bacteriophage penetration through *Pseudomonas aeruginosa* biofilms. Appl Environ Microbiol 2001; 67:2746-53.