



# Intriguing Interaction of Bacteriophage-Host Association: An Understanding in the Era of Omics

Krupa M. Parmar<sup>1\*</sup>, Saurabh L. Gaikwad<sup>1</sup>, Prashant K. Dhakephalkar<sup>1</sup>, Ramesh Kothari<sup>2</sup> and Ravindra Pal Singh<sup>2\*†</sup>

<sup>1</sup> Bioenergy Group, Agharkar Research Institute, Pune, India, <sup>2</sup> Department of Biosciences, Saurashtra University, Rajkot, India

## OPEN ACCESS

### Edited by:

Florence Abram,  
NUI Galway, Ireland

### Reviewed by:

Steven Ripp,  
University of Tennessee, Knoxville,  
USA  
Sanna Sillankorva,  
University of Minho, Portugal

### \*Correspondence:

Krupa M. Parmar  
krupa\_11091@yahoo.co.in  
Ravindra Pal Singh  
ravindrapal.1441@gmail.com

### † Present Address:

Ravindra Pal Singh,  
Department of Biological Chemistry,  
John Innes Centre, Norwich Research  
Park, Norwich, UK

### Specialty section:

This article was submitted to  
Microbiotechnology, Ecotoxicology  
and Bioremediation,  
a section of the journal  
Frontiers in Microbiology

**Received:** 03 January 2017

**Accepted:** 16 March 2017

**Published:** 07 April 2017

### Citation:

Parmar KM, Gaikwad SL,  
Dhakephalkar PK, Kothari R and  
Singh RP (2017) Intriguing Interaction  
of Bacteriophage-Host Association:  
An Understanding in the Era of Omics.  
*Front. Microbiol.* 8:559.  
doi: 10.3389/fmicb.2017.00559

Innovations in next-generation sequencing technology have introduced new avenues in microbial studies through “omics” approaches. This technology has considerably augmented the knowledge of the microbial world without isolation prior to their identification. With an enormous volume of bacterial “omics” data, considerable attempts have been recently invested to improve an insight into virosphere. The interplay between bacteriophages and their host has created a significant influence on the biogeochemical cycles, microbial diversity, and bacterial population regulation. This review highlights various concepts such as genomics, transcriptomics, proteomics, and metabolomics to infer the phylogenetic affiliation and function of bacteriophages and their impact on diverse microbial communities. Omics technologies illuminate the role of bacteriophage in an environment, the influences of phage proteins on the bacterial host and provide information about the genes important for interaction with bacteria. These investigations will reveal some of bio-molecules and biomarkers of the novel phage which demand to be unveiled.

**Keywords:** bacteriophage, genomics, next-generation sequencing, transcriptomics

## INTRODUCTION

Innovations in next-generation sequencing (NGS) technology and the decline in the sequencing cost have triggered a revolution to gain an understanding into the diversity, structure, and function of complex microbial processes (Vlahou and Fountoulakis, 2005). NGS has enhanced our concept of various influences of microbes in maintaining equilibrium in the environment and accentuating the function of diverse hosts including humans (Li et al., 2008). NGS has expedited the interpretation of microbiome using techniques, such as metagenomics, metatranscriptomics, metabolomics, proteomics, and single cell genomics. Apart from sequencing, bioinformatics and statistical tools also represent a significant role in sequence assembly, alignment, binning, and annotation. Software of bioinformatic assists in decoding the identity, abundance profile, genetic composition and functional channels of an organism or for a microbial community (Meyer et al., 2008; Glass et al., 2010). Genomics deals with sequencing of the whole genome of a distinct organism whilst metagenomics study a pool of genomes of a community of different populations (Handelsman, 2004). According to the central dogma, the flow of genetic information in a cell is from DNA (genome) to RNA (transcriptome) and then it is translated into proteins (proteome) (Crick, 1970). Genomics elucidates presence of the gene in an organism, while transcriptomics provides the information about the genes that are actively expressed in an organism and proteomics

study the structure and function of every protein in an organism. A novel technique called single cell genomics takes in record information of the sequence from an individual cell which procures a better degree of accuracy in cellular differences and a finer understanding into the function of a particular cell in an ecosystem (Eberwine et al., 2014). However, metabolomics includes the analysis of metabolites of an organism and these results may vary from genomics and transcriptomics data as they are influenced by surrounding environments (Apel and Hirt, 2004). Apart from sequencing, bioinformatics and statistical tools assist in sequence assembly, alignment, binning and annotation.

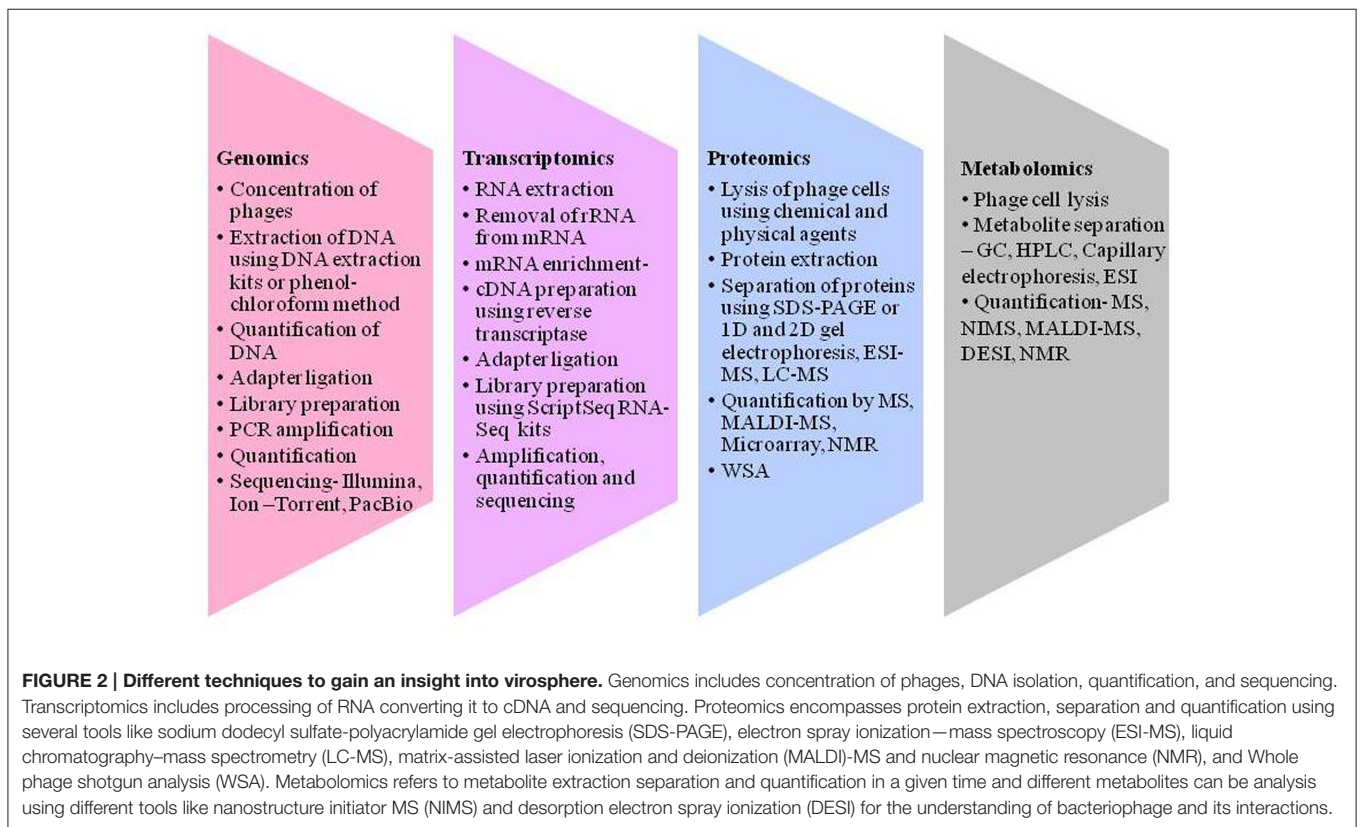
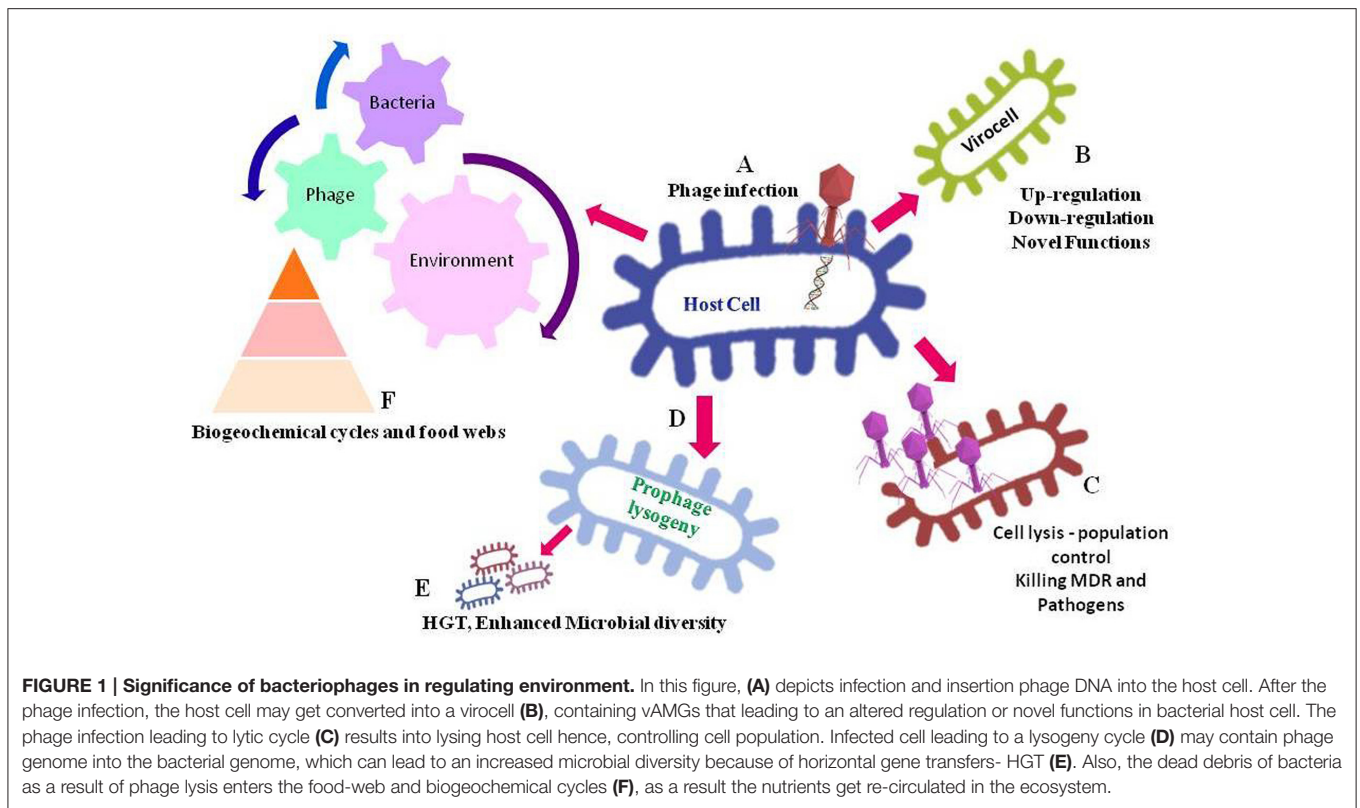
Whilst, the information about bacteria present on earth is better understood, data regarding viruses particularly bacteriophages (henceforth called phages) is still in its infancy. Phages are the most abundant and diverse group of viruses found on Earth (Short and Suttle, 2005). Interactions between the bacterial host and phage have significantly played a role in biogeochemical cycles, regulation of the microbial community structures and governing the microbial populations (**Figure 1**). Recent surveys have documented the capacity of phages in maintaining the stability of microflora in the human gut (Minot et al., 2011) and regulation of pathogen and multidrug resistance (MDR) in the environment (Parmar et al., 2017). In bacteria, 16S rRNA genes and several house-keeping genes serve as a biomarker which facilitate their identification, whereas, there is an absence of biomarker gene among phages, which poses as a hindrance for identification of phages and hence the phage database is quite insufficient (Edwards and Rohwer, 2005). Addressing the challenge to design a biomarker for phages may uncover new avenues in better understanding the virosphere. Plaque assay being a culture-dependent technique isolates a specific phage against a bacterial host. Subsequently, their identification can be worked out using phenotypic characteristics and sequencing (Sanchez et al., 2015). In an approach to improve a comprehensive insight into uncultivable phage and their interaction with the bacterial host, this review summarizes different NGS techniques and the bioinformatics tools that are applied to evaluate such data.

## BACTERIOPHAGES AND THEIR INTERACTION WITH BACTERIAL HOST

Phages are in 10:1 ratio with bacteria on Earth, though viral DNA corresponds to only 0.1% of total DNA among microbial communities (Qin et al., 2010). Phages are ubiquitous in the environment and are found abundantly where bacterial hosts thrive. Mostly phages flourish in oceans, soil, wastewater treatment plants, hot-water springs and animal gut (Wommack and Colwell, 2000; Prigent et al., 2005; Prestel et al., 2008; Srinivasiah et al., 2008). Phages are classified on the basis of their size, structural composition, genome organization and on the host it infects (Ackermann, 2009). Electron microscopy assists in studying phenotypic characteristics of phage on the basis of the size and shape of head, tail and tail fibers. The most abundant phages in the environment are dsDNA belonging to order *Caudovirales* (Weinbauer and Rassoulzadegan, 2004).

*Caudovirales* are furthermore classified into *Podoviridae* having a short tail, *Siphoviridae* with long non-contractile tail and *Myoviridae* possessing a long contractile tail. Every phage is specific toward a particular bacterial host and may have a narrow or broad host range depending on its infection capability. Hosts provide the enzymatic machinery for the phages to replicate and multiply by infecting the most active (exponentially growing) bacteria as implied by “kill the winner” hypothesis (Rodrigue et al., 2009). Phages undergo two types of life cycles, (1) in the lytic lifecycle, phage injects own DNA into a host cell and multiply by manipulating host replicating machinery. After replication, phage cleaves the host bacterial cells releasing progeny virus particles into the environment. Whereas, in (2) lysogeny lifecycle, phage DNA integrates into the bacterial genome, replicates their DNA along the bacterial genome and transfer on to the progeny of host cells (Bertani, 1951).

The interplay between host and phage particle initiates as soon as the phage recognizes specific receptors on the bacterial cell wall. The tail proteins of phage particle recognize the receptor protein(s) of bacteria and inject own DNA into host cytoplasm choosing either lytic or lysogenic lifecycle. Once the phage DNA is inserted into the bacterial cell, the cell is termed as a “virocell” carrying virus auxiliary metabolic genes (vAMGs), which are believed to augment the metabolic potential of the host during infection process as shown in **Figure 1** (Rosenwasser et al., 2016). The phages acquire new genes into their genomes by interactions with the host genome in order to replicate in the host cells. The bacterial genes that attach near to the prophage attachment site, suggests the genes were acquired by inaccurate prophage excision. Some novel genes can similarly be transmitted into the interior part of the genome by some unexplained mechanism (Juhala et al., 2000). However, these genes may be autonomous transcripts or repressed prophages that provide benefit to hosts (Brüssow et al., 2004). Horizontal gene transfer (HGT) by these phage particles from one host to another host genomes, results in an increased microbial diversity (Dutta and Pan, 2002; Weinbauer and Rassoulzadegan, 2004). Thus, the interaction between phage and host chiefly emphasizes the structure of microbial communities (Rohwer and Thurber, 2009). Some genes derived by phage also aid in nutrient cycling and gear up the biogeochemical cycles on Earth. Furthermore, phages have a crucial aspect in host mortality, carbon cycling (Breitbart et al., 2004) and nutrient cycling (Suttle, 2007). Also, microbial lysis by phage infection has significance in bacterial population control and the debris of these dead microbes act as a food source in the food web of the environment (Sime-Ngando and Colombet, 2009), thus involved the cycling nutrients (**Figure 1**). Phages are thus accounted as an application to limit bacterial pathogens and multi-drug resistant organisms in the environment by the mechanism of specifically lysing the bacterial hosts (Parmar et al., 2017). Despite an immense abundance and diversity of phages and their reimbursement in the global webs, molecular knowledge of phage-host interactions is missing. In the era of NGS, employing genomics, single cell genomics, transcriptomics, proteomics, and metabolomics can be a smart attempt to understand the interaction among the phages and their bacterial hosts (**Figure 2**). A review of the literature has been



solicited to confer the claims of omics approach in phage research (Table 1).

## APPLICATION OF GENOMICS TO REVEAL PHAGE DIVERSITY

Owing to the insufficiency of viral database, there is more than 90% viral dark matter (Hurwitz et al., 2013, 2015). Additionally, the absence of a biomarker gene among phages leads to sequence the whole phage genome for its understanding (Thurber et al., 2009). The genomics of phages would elucidate the genetic composition and putative functional role in the environment. Phage metagenomics would furthermore assist in determining the diversity of phages in a community and reveal novel genes demonstrating phages to be the most diverse beings on the globe (Edwards and Rohwer, 2005). Subsequently, interpretation of functional channels of bacterial viruses would illuminate host-phage interactions (Brum et al., 2015).

A typical genomic experiment begins with isolation of genomic DNA of virus particles (Figure 3). The primary step is filtration of the sample through 0.22  $\mu\text{m}$  filters for elimination of bacterial constituents and other contaminations. Samples are then concentrated by ultracentrifugation or polyethylene glycol precipitation (Helms et al., 1985) and subjected to DNase and RNase treatments to exclude residual genomic material from any contaminant bacteria that may pass through 0.22  $\mu\text{m}$  filter. This treated sample would include only virus particles which can be cleaved and their genome can be extracted using kits or standard methods (Adhikary et al., 2014). In order to examine with NGS platforms, DNA is fragmented, ends are repaired and are ligated with adaptors (Holmfeldt et al., 2013). Finally, fragmented DNA library is cleaned and amplified through PCR as well as is quantified and sequenced. Several sequencing platforms are available such as Ion Torrent, Illumina, PacBio which are preferred as per the requirement of their read length, coverage, paired reads, insert size, accuracy, error rates, sequencing yield, run time and sequencing cost (Quail et al., 2012). To reduce chances of bacterial contamination in the library, a section of DNA is PCR amplified for 16S rRNA genes, and if bands are detected, it conveys the presence of host contamination. In the instance of lesser viral DNA yields, amplification using multiple displacement amplification (MDA) can be performed, but it may generate chimeras (Yilmaz et al., 2010). Apparently, amplifications using linkers may depict impartial viromes (Duhaime et al., 2012; Hurwitz et al., 2015). The sequences acquired by sequencing are developed for data filtering and the sequence reads that passes quality check, is mapped to reference genomes or assembled *de novo*. If the sequence of contaminating host is furthermore present in reads even after purifying the sample, it can be distinguished by comparing reads to reference bacterial genome or 16S rRNA database (Hurwitz et al., 2013). For annotation of viral genomes, a database such as NCBI non-redundant nucleotides can be used. ORFs can be determined and annotated using CyVerse (Goff et al., 2011) in the PCPipe application through the iVirus project (Hurwitz et al., 2014).

Bioinformatic tools mine enormous volume of sequence data to determine common patterns that govern microbes in an ecosystem. Viral diversity can be estimated using PHACCS toolkit (Angly et al., 2005). To decode the correlation between virus community and environmental factors, an application called Fizkin by CyVerse cyber infrastructure iVirus project selects 300K reads arbitrarily from viromes and examines it using Jellyfish that generates a matrix of shared sequence counts between each virome pair. This matrix uses an input file for Bayesian network analysis resulting in a table of the relevance of environmental factor that determines the diversity of virus and a social network graph (Hurwitz et al., 2014, 2016). This will assist in ecological profiling of viral communities without requiring assembly and annotation. To elucidate sequence matches with a reference database, BLAST is regularly employed along with MG-RAST (Glass et al., 2010), MetaPhyler (Liu et al., 2010) or CARMA (Gerlach et al., 2009). For a taxonomy of viruses, MEGAN (Huson et al., 2007) software can be used whereas Hidden Markov Models, e.g., HMMER (Finn et al., 2011) are applied to match Pfam or KEGG domains. To find specific viral species present in metagenome, k-mer based algorithms such as CLARK (CLAssifier based on Reduced K-mers) (Ounit et al., 2015), USEARCH (Edgar, 2010), KRAKEN (Wood and Salzberg, 2014), and NBC (Naïve Bayes Classifier) (Rosen et al., 2011) have been applied. Sometimes, whole host genome can be observed in viromes when gene transfer agents (GTAs cluster) have filtered along with virus-like particles (Roux et al., 2013b). GTAs and sporadic contaminations can likewise be recognized using software CLARK (Ounit et al., 2015). Alignment of sequences with reference bacteria genome may reveal a prophage viral element using “recruitment plot” in the bacterial genome. Some of the bioinformatics tools adapted for prophage detection include ACLAME, Prophinder, PHAST and PhiSpy which can serve in confirming phage annotation (Akhter et al., 2012).

Along with elucidating diversity and taxonomy of phages, establishing the origin of genes (bacterial or viral) is vague. This ambiguity occur because of vAMGs which incorporates (enhancing cell metabolism in the host) into host cells or some viruses may also uptake some bacterial genes near the prophage excision site. Nevertheless, during integrating into host tRNAs, phages carry an attachment site (attP) which denote a definite match of a host tRNA gene. Example, integrase gene and an attP site (53 bp) of the *Prochlorococcus* phage P-SS2 is a precise analogue of the host tRNA (attB, 36 bp) of *Prochlorococcus* MIT9313 (Sullivan et al., 2009). Such phages that display a putative attP site and an integrase identical to a host tRNA gene fragment are suggestive of a host-phage association (Mizuno et al., 2013). Metagenomics serves to find diversity among phages but knowledge about interaction among phage and host is relatively scanty. By analyzing the spacers in CRISPR to phage metagenomes, the bacterial host of phage (Dutilh et al., 2014) and phage-host interactions can be deduced (Anderson et al., 2011; Berg Miller et al., 2012; Edwards et al., 2016). Characterizing these constraints is requisite to develop our insight about bacterial-phage co-evolution.



**TABLE 1 | Applications of omics approach for understanding bacteriophage-bacterial host interaction.**

Sr. No.	Bacteriophage	Bacterial Host	Omics	NGS platform and tools	Applications	References
1	<i>E. coli</i> bacteriophages lambda and T4	<i>E. coli</i>	Single virus genomics	FACS, MDA, and 454 GS FLX	Viral diversity, evolution, adaptation, and ecology	Allen et al., 2011
2	Rumen bacteriophages	Rumen bacteria	Metagenomics	Roche GS FLX	Phage-bacterial relationship	Berg Miller et al., 2012
3	Ø11b	<i>Flavobacterium</i>	Genomics and Proteomics	Shotgun sequencing, PAGE and MALDI-TOF-MS	Genome and proteome analysis	Borriss et al., 2007
4	Marine viruses	Marine bacteria	Metaproteomics	HiSeq-2000	Protein profile study	Brum et al., 2016
5	ØVB_CoM-IBB_35	<i>Campylobacter coli</i>	Genomics and Proteomics	Pyrosequencing, SDS-PAGE and MS	Genomics and protein genes	Carvalho et al., 2012
6	Six different phages	<i>P. aeruginosa</i>	Metabolomics	TOF-MS	Alterations in host metabolism during infection	De Smet et al., 2016
7	Ø 812	<i>Staphylococcus</i>	Proteomics	1-DE, 2-DE, and MALDI-TOF-MS	Proteome analysis	Eyer et al., 2007
8	A4 Mycobacteriophage Kampy	<i>M. smegmatis</i>	Transcriptomics	RNA-Seq and MS	Profiling active genes	Halleran et al., 2015
9	Mycobacteriophage	<i>Mycobacteria</i>	Metaproteomics	ABI 3730 or ABI3100	Protein family studies	Hatfull et al., 2006
10	Skin virome	-	Metagenomics	Computational approaches	Exploring the viral dark matter	Hurwitz et al., 2016
11	MS2 bacteriophage	<i>E. coli</i>	Metabolomics	Flux balance analysis	Host-pathogen interaction	Jain and Srivastava, 2009
12	vB_EcoP_SU10	<i>E. coli</i>	Genomics and proteomics	Roche/454 and Nano LC-MS/MS analysis	Genomics and protein cataloging	Khan Mirzaei et al., 2014
13	RNA Bacteriophages		Metagenomics and transcriptomics	Various based on subjects studied	Study novel RNA phages and their transcripts	Krishnamurthy et al., 2016
14	ØRV5	<i>E. coli</i> O157:H7	Genomics and proteomics	Clone library and primer walking and MS/MS	Using phages as analytical tools	Kropinski et al., 2013
15	Marine surface viruses	Marine surface bacterioplankton	Single cell genomics	Various	Host-virus interaction	Labonté et al., 2015
16	ØKVM	<i>P. aeruginosa</i>	Proteomics	SDS-PAGE, WSA and LC-MS	Structural proteins	Lavigne et al., 2006
17	Mycoviruses	Plant pathogenic fungi	Metatranscriptomics	RNAseq- illumina	Diversity of mycoviruses	Marzano et al., 2016
18	ØVB_YecM_ØR1-37 (ØR1-37)	<i>Yersinia enterocolitica</i>	Transcriptomics	Whole genome transcription	Host-virus interaction	Leskinen et al., 2016
19	Cyanomyo-virus P-SSM2	marine <i>Cyanobacterium</i>	Transcriptomics	whole genome transcription	Host transcriptional response on infection	Lin et al., 2016
20	ØDp-1 and ØCp-1	<i>S. pneumoniae</i>	Proteomics	AP/MS	Interactions between bacteria and phage	Mariano et al., 2016
21	T5-like Bacteriophage	<i>E. coli</i> O157:H7	Genomics, Proteomics	454 Technology by GS FLX, SDS-PAGE, and MALDI TOF MS	Check for potential as biocontrol agent	Niu et al., 2012
22	ØMSP	<i>S. aureus</i>	Proteomics	SDS-PAGE and 2DE	Proteomic characterization	Sangha et al., 2014
23	Halophilic viral communities	Halophiles	Metatranscriptomics	Microarray, DGGE, and ABI PRISM 310 DNA Sequencer	Metatranscriptome analysis of halophilic viruses	Santos et al., 2011
24	ØCP390 and ØCP26F	<i>C. perfringens</i>	Genomics and proteomics	Pyrosequencing, 2DE, and MALDI-TOF-MS	Prediction of genome and proteome	Seal et al., 2011
25	ØTMA	<i>Thermusthermophilus</i>	Genomics and proteomics	ABI 3700 Sequencer, PFGE, and SDS-PAGE	Genomic and proteomic characterization	Tamakoshi et al., 2011
26	Different phages against <i>P. aeruginosa</i>	<i>P. aeruginosa</i>	Proteomics	Affinity purification with MS	Identification of hypothetical proteins	Van den Bossche et al., 2014

(Continued)

TABLE 1 | Continued

Sr. No.	Bacteriophage	Bacterial Host	Omics	NGS platform and tools	Applications	References
27	MC1061 (φ24B)	<i>E. coli</i>	Transcriptomics	Whole-genome shotgun pyrosequencing	Altered functions of host after phage infection	Veses-Garcia et al., 2015
28	Øwv8	<i>E. coli</i> O157:H7	Genomics and proteomics	Pyrosequencing, MALDI-MS, and QqTOF	Genomic and proteomic analysis	Villegas et al., 2009
29	ØPHV1 type A and ØPHV1 type B	<i>Cyanobacteria</i>	Metagenomics and metatranscriptomics	HiSeq-2000	Interactions between bacteria and phage	Voorhies et al., 2016
30	ØvB_BpuM_BpSp	<i>Bacillus</i>	Proteomic analysis	HPLC-ESI-MS/MS	Check a particular activity of protein	Yuan and Gao, 2016
31	ØPaP3	<i>P. aeruginosa</i>	Transcriptomics	Microarray, RT-qPCR	Interactions among bacteria and phage	Zhao et al., 2016

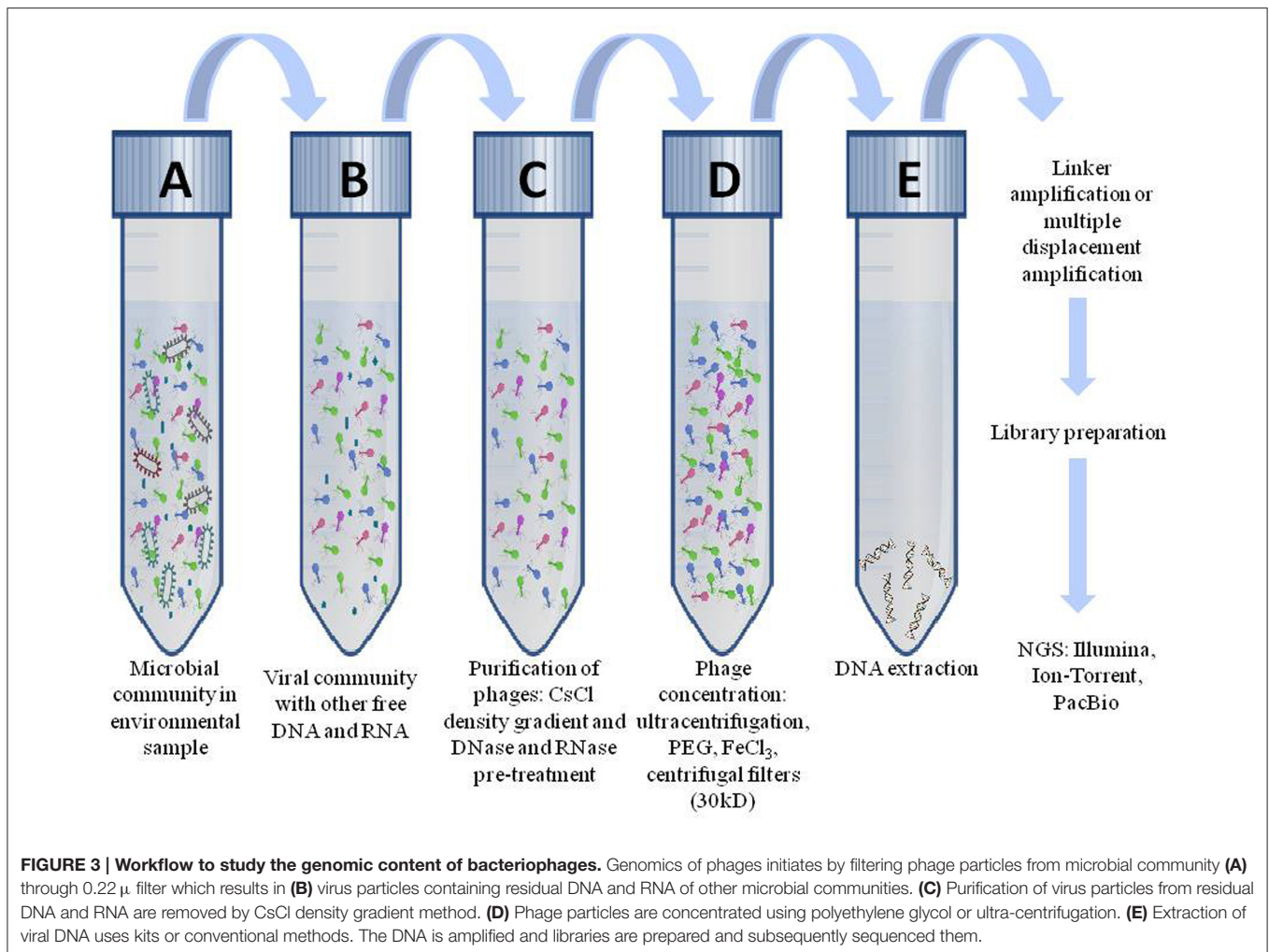
1-DE, 1 dimensional electrophoresis; DGGE, denaturing gradient gel electrophoresis; FACS, fluorescence-activated cell sorting; LC-MS, Liquid chromatography-mass spectrometry; MDA, multiple displacement amplification; MALDI-TOF-MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; PAGE, polyacrylamide gel electrophoresis; PFGE, pulsed field gel electrophoresis; RT-qPCR, Real-time quantitative polymerase chain reaction; RNA-Seq, RNA-Sequencing.

## Single Cell Genomics

Apart from metagenomics, attempts have been instigated to investigate only an individual isolate in detail reinforcing our perception of the mechanisms of a specific cell rather than the influences of the entire population. Recently, single cell genomics (SCG) has been promoted to infer the phage genome which is present in or on the surface of host cells in a particular niche without culturing (Lasken and McLean, 2014) and facilitates in assuming sole genetic and metabolic profiles of uncultivable microbes. SCG helps in understanding the interplay between the phage and their host and can ascertain phage genome in a bacterial host cell. To isolate a single cell from an environment, techniques such as flow cytometry (Podar et al., 2007) and micromanipulation (Ishøy et al., 2006) have demonstrated to be advantageous. To sort a single cell precisely, a fluorescence-activated cell sorter (FACSaria™) with a forward scatter photomultiplier tube (PMT) has been adopted to simplify accurate detection and high-resolution entry of single cell (Picot et al., 2012). Confocal laser scanning microscopy has been applied to support a single phage separation stained by fluorescent dyes lodged into agarose (Luef et al., 2009). Multiple displacement amplification (MDA) (Hosono et al., 2003) utilizes an advanced properties of phi29 DNA polymerase which intensifies a microbial genome at million-fold, sufficient for sequencing using any of the available sequencing platforms.

Interpretation of viral diversity has become easier after the expansion of single virus genomics (Allen et al., 2011) while attending one virus at a time. New computational challenges to analyze the outcomes of SCG using bioinformatics tools have emerged, reflecting the vast opportunity to figure out the *in-situ* phage-host communications. Several bioinformatics tools have been in practice for the classification of prophages-pathogenicity islands such as PIPS (Soares et al., 2012) and HGT- using Alien Hunter (Vernikos and Parkhill, 2006), but these tools seem weak when studying novel phages because of a deficiency of genomic sequences in the viral public database. Because of this constraint, semi-continuous and partial SCG sequences in the database do not allow the accurate identity of isolates (Kalisky and Quake, 2011). However, SCG provides cytoplasmic insights during various interactions with phages like lysogeny, lytic infections, chronic infections and unspecific attachments (Allen et al., 2011). For distinguishing between these synergies, sequences have been examined for integration of phage into the host DNA, portion of phage and host DNA was measured for the speed of single cell MDA reactions and comparisons have been made between the coverage depth of phage and host contigs (Labonté et al., 2015). Phages infecting previously unknown hosts have been discovered in the marine environment using this technique (Roux et al., 2013a, 2014; Labonté et al., 2015).

Non-specific amplification or distortions in the single genome may be a reason for the loss of data, but approximately 90% of DNA can be retrieved using SCG (Rodrigue et al., 2009). A newly developed technology called Hi-C sequencing determines closely arranged genome sequences, like virus-bacterial host genomes within an individual cell (Beitel et al., 2014). The concept of this facility includes genome cross-linking using formaldehyde



and a restriction excision, followed by re-ligation of sequences using ligases in a dilute condition that support ligation events between cross-linked DNA fragments, conforming the pairs to each other that were originally in close contiguity (van Berkum et al., 2010). This technique can be adapted to phage-bacterial host communities to figure out close entity while they have been successfully operated for various microbial studies (Beitel et al., 2014; Burton et al., 2014).

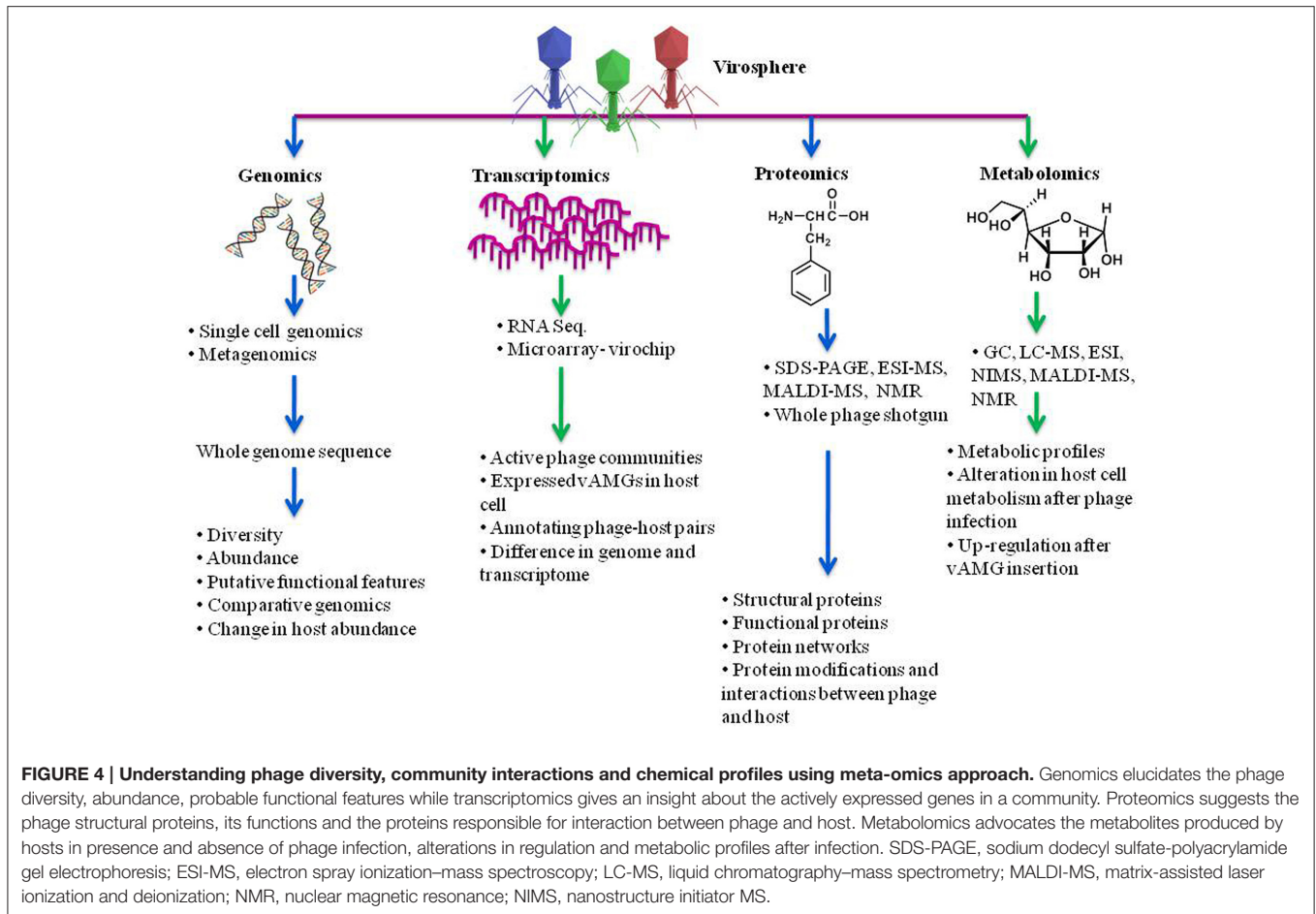
Oxford Nanopore sequencing has been utilized where individual DNA molecule is directly sequenced without amplifying or labeling genome with chemical or using visualization tool to recognize the chemical label (Branton et al., 2008). Nanopore sequencing works on the principle that when a voltage is applied to a nanopore imbedded in a conducting liquid, electric current can pass through the nanopore. This electric current is highly responsive to nanopore size and shape such that indeed a single passage of DNA nucleotide pass through nanopore could affect an alteration in the current. The magnitude of current differs based on the type of nucleotides (A, T, G, or C) passing through the nanopore. Thus changing in current corresponds to the precise sequence of a DNA stretch. Viral pathogens have been examined using Nanopore technology

(Greninger et al., 2015). Concurrently, MinION sequencer has similarly been recommended which is incredibly rapid, smaller in size, produce 200 kb long reads with high accuracy as well as it has been used to study lambda phage DNA (Mikheyev and Tin, 2014).

Metagenomics and SCG technologies can be strongly adapted to illustrate the exact identity and diversity of phages in an environment which can guide across the dark matter of viral ecosystem (**Figure 4**). Along with investigating the diversity of phages in an environment, to succumb with a coherent outline of their functional aspect in an ecosystem, it becomes imperative to deduce mechanisms underlying transcriptions of phages. Hence, transcriptomic studies provide knowledge about functions of active genes in given condition.

## EMPLOYING TRANSCRIPTOMICS TO STUDY ACTIVE PHAGE FUNCTION

Transcriptomics provides a measure to investigate the active microorganisms within a community at a specific time and under a definite array of conditions. Study of the transcriptome



is critical to analyze molecular constituent of phages and to understand genome function during a distinct period or situation such as development or infection state. The principle objectives of transcriptomics include recording transcripts of all species including mRNAs, non-coding RNAs and small RNAs, to estimate transcriptional organization of genes in terms of 5' end and 3' end, gene splicing and post-transcriptional modification as well as to demonstrate varying activity of each gene under different conditions (Bikel et al., 2015). Several technologies have been developed to inspect and determine transcriptomes, such as DNA hybridization technique, DNA microarray, cDNA-amplified fragment length polymorphism (cDNA-AFLP), expressed sequence tag (EST) sequencing, serial analysis of gene expression (SAGE), massive parallel signature sequencing (MPSS), and RNA-seq (Mutz et al., 2013). DNA hybridization employs fluorescently labeled cDNA to hybridize with DNA templates on microarray chips. However, this tool possesses some reserves as it relies on the hitherto studied genome sequences, high background levels for cross-hybridization and a smaller detection range. Phage meta-transcriptomics may present few difficulties as phages are incredibly diverse and their database is considerably less. Also, the availability of RNA, especially mRNA may be in a rather less volume because of inactive phase when phage is not in association with bacterial hosts. Thus, this

may lead to challenges in isolating and enriching mRNA for sequencing.

A typical transcriptomics analysis using NGS includes isolation of the total RNA from the virus particles, depending on the RNA to be sequenced (mRNA, lincRNA or microRNA). Initially, the bacterial fractions are separated and purified from the phage particles. DNase and RNase treatments are administered to filter phage particles from any free DNA or RNA of bacteria. The RNA is thus extracted from virus particle using RNA extraction kits such as RNeasy mini kit. Selective elimination of rRNAs can be achieved using rRNA removing kits or using probes complementary to the rRNA region that is attached to magnetic beads. The mRNA can be enriched by magnetic bead capture method of rRNA, preferential polyadenylation of mRNA or preferential digestion of rRNA through enzymes. The cDNA are synthesized using random hexamers or oligo (dT) primers or priming with poly dT primers after polyadenylation. For amplification, RNA polymerase (Ozsolak and Milos, 2011) or MDA (Gonzalez et al., 2005) or emulsion PCR is/are performed (Tang et al., 2009). The 5' and/or 3' ends of the cDNA are then repaired along with adapter ligation, following library cleanup, amplification, quantification and sequencing of the library. Single-end or paired-end libraries can be prepared using kits like ScriptSeq



RNA-Seq library preparation kit (Illumina, San Diego, CA) and can be sequenced on platforms such as Illumina HiSeq2500. Sometimes, conversion of RNA into cDNA introduces bias into the quantification of transcripts, thus a semi-direct sequencing of RNA by-passing the synthesis of cDNA has been established (Hickman et al., 2013).

The bioinformatics analysis of raw data retrieved by transcriptome sequencing uses reference genes and genomes to map against the raw reads or performing a *de novo* assembly for unreported transcriptomes. Mapping of transcripts against reference genome would confer taxonomy and function of active phages. Mapping the active functional pathways would expound the up-regulated, down-regulated or unaffected genes of phage during development or infection cycle. The same can be advised for the bacterial host during infection by a phage. The transcriptome reads which are short can be assembled *de novo* using several softwares such as Trinity. Efficiency and sensitivity of the software are exceptionally promising in procuring full-length transcripts (Ghaffari et al., 2014). The assembled contigs that are obtained by *de novo* or reference-based assembly can be equated with the NCBI viral reference amino acid sequence database using USEARCH (Edgar, 2010). Moreover, the virus annotated hits can be compared with NCBI non-redundant database using BLASTX. Bowtie software can be used to calculate sequence read number and coverage depth (Langmead et al., 2009). Alignment of reads can be prepared using software called MUSCLE (Edgar, 2004) and for constructing neighbor-joining tree, MEGA software can be employed as uses for the bacterial analyses (Tamura et al., 2013).

Several phage meta-transcriptomics studies have been conducted in an attempt to analyze active phage communities. Studies have revealed the effectiveness of phage metagenome for constructing templates in the microarray (Virochip) to annotate and identify the sample (Santos et al., 2011). The total RNA extracted from the sample can be converted to cDNA, labeled and allowed to hybridize with the virochip (Santos et al., 2012). A combination of metagenomics and meta-transcriptomics study would specifically determine active phages in an environment in which phage transcripts may vary as compared to their genomic abundance (Lim et al., 2013). There can be a case when a particular set of family of genes remain less abundant in metagenomic analyses whilst those genes may be remarkably active in meta-transcriptomics dataset and/or vice versa (Franzosa et al., 2014). This insinuates that performing only a metagenomic study may not be a perfect snapshot of functional active genes in a metagenome. To overcome the tedious isolation of viral mRNA from total mRNA, SCG can be conducted along with microarrays to designate the phage-host systems without cultivating them (Santos I. M. et al., 2015). In another study, phage-host pairs have been investigated in which a fosmid viral metagenomic library was constructed and immobilized on microarray “virochip,” along with them. The genomes of uncultured bacterial host cells can be sorted by fluorescence-activated cell sorting (FACS) followed by amplification via MDA. Single host cells were hybridized on virochip, and the host cells and immobilized phages with positive results were sequenced (Martínez-García et al., 2014).

With this new technique, advancements toward the discovery of phage-host interactions arise in current decade (Santos F. et al., 2015). Moreover, meta-transcriptomics based enzyme discovery from phages can assist in utilizing novel enzymes with specific enzymatic characteristics for the industries and scientific communities (Schoenfeld et al., 2010).

Transcriptomics can be employed to analyze the influencing of a phage on the bacterial host after the phage infection. One such response includes induction of Shiga toxin production and acid resistance in *E. coli* by Shiga toxigenic phages (Veses-García et al., 2015). Studies confirm the fact that host genes get differentially expressed after the phage infection such as a phage “PaP3” had a down regulatory impact on host transcriptional regulators and it proved early genes of phage affected strongly by regulation of hosts (Zhao et al., 2016). This feature of phages can be promoted for formulating a phage therapy. Transcriptomics studies of phage during infection of the host can serve an insight of sequence of transcriptional events, such as initial phase consisting of gene metabolism, DNA synthesis, and regulation genes, is accompanied by a prolonged phase of structural and lysis genes (Halleran et al., 2015). These views can yield information about vAMGs which alters metabolic functions of bacterial host after phage infection. During late phase of phage infection, several up-regulating mechanisms have been observed in bacterial gene expression including stress response and stability of membranes (Leskinen et al., 2016). Additionally, enrichment of ATP synthase and ribosomal protein genes have been revealed during phage infection of phosphorous starved *Cyanobacterium* host (Lin et al., 2016). With further advancement and a few drawbacks, transcriptomics gives an insight of the phage-host interactions and evaluates the regulatory mechanism in bacterial hosts by phage and/or vice versa (effect of host interactions on phage regulation) which are noteworthy for developing phage therapy and comprehend novel phage antimicrobial compounds.

## UNDERSTANDING THE PROTEOMIC PROFILE OF PHAGES

A proteome can be represented as a set of all expressed proteins in a cell, tissue or an organism (Theodorescu and Mischak, 2007). Proteomics is a methodology for the characterization of genetic data in a cell or an organism via protein pathways and networks (Petricoin et al., 2002) and for distinguishing the functional implication of proteins (Vlahou and Fountoulakis, 2005). It focuses at cataloging protein expression profiles at a specific period, in a definite location of the cell and as a response to foreign stimulations. It is applied to design a plot of protein networks which can be used to demonstrate interaction among protein in an organism (Corpillo et al., 2004). It provides an estimate of occurrence, quantity and modified state of proteins in an environment in a significant-throughput method.

Genome and transcriptome analysis evaluate the indirect functional profile of a cell or a community whereas proteomic reveal a direct estimate of functional activity of a cell (Schwanhäusser et al., 2011). Abundance profiles of proteins

can be plotted using comparative metaproteomics, while the reduction or increase in the quantity of some proteins may signify a distinct purpose in an organism or during particular situations of phage infection on the bacterial host (Sangha et al., 2014). The post-infection protein expression changes can be classified as (1) function which alters rapidly on phage-infection, but can get reverted back (2) variations that develop gradually and persist consistent or cannot revert back, and (3) alterations that appear abruptly and are maintained for a longer term.

Developments in next generation tools have drastically enhanced quantification and identification of proteins (Schleicher and Wieland, 1978). The proteomic analysis commences with phage concentration accompanied by lysing phage using physical and chemical agents, consequently releasing phage proteins (Figure 2). The concentration of proteins can be measured using Bradford's method (Bradford, 1976) or can be denatured using urea (Lavigne et al., 2006) or can be digested by trypsin (Borriss et al., 2007). Several approaches and facilities have been in practice for proteomic studies (Chandramouli and Qian, 2009), however, employing some tools such as a mass spectrometer (MS) and protein-chips (microarray) have significantly contributed in the field (Horgan and Kenny, 2011). Proteins have earlier been detected and quantified using enzyme-linked immunosorbent assay (ELISA) and Western blot where proteins were initially separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Lavigne et al., 2006). Studies have been performed to understand phage proteins using MS after separating by 1D and 2D PAGE (Clement et al., 2013). Additionally, mass-spectrometry-based techniques such as matrix-assisted laser desorption/ionization (MALDI-MS) (Borriss et al., 2007) and electron spray ionization (ESI) (Carvalho et al., 2012) have been established for analyzing various proteins of phages. Recently, fluorescence 2D differential gel electrophoresis has been employed to distinguish between amounts of human lymph and plasma proteins (Clement et al., 2013). Structural proteomics can interpret the structure of proteins thereby determining the functions of novel genes. Lavigne et al. (2006) described structural proteome of phiKMV, a lytic bacteriophage of *Pseudomonas aeruginosa* using SDS-PAGE, LC-ESI-MS/MS, and GC-MS. Nuclear magnetic resonance (NMR) (Horgan and Kenny, 2011) and X-ray crystallography (Drulis-Kawa et al., 2012) can be employed to investigate the interaction between phage-binding protein and receptor site on the bacterial host (Sundell and Ivarsson, 2014).

Protein analysis using MS requires a prior separation of the sample either by 2D-gel electrophoresis (Renesto et al., 2006) or isotope-coded affinity tag (ICAT) labeling (Weston and Hood, 2004), accompanied by digestion into peptides and separating peptides using LC. Microarrays can be applied for assorting protein interaction with DNA, protein or ligands. Protein microarray technique can be exploited in the analytical study to check for presence/absence of a distinct protein in a sample (biomarker detection during phage infections) or for defining function (Uzoma and Zhu, 2013). When phage proteins are immobilized on a microarray chip, it can be applied to probe for complementary bacterial host receptors that bind with

phage recognition proteins (Santos F. et al., 2015). Reverse-phase protein microarray can serve as a comparative protein profile in case of phage-infected and uninfected bacterial host (Haider and Pal, 2013). Thus, correlative examination of proteome and genome provide an interpretation of the post-translational modifications.

Functional identification of hypothetical phage proteins is performed using MS analysis after affinity purification of host protein mixtures (Van den Bossche et al., 2014). MS/MS spectra can be interpreted using SEQUEST (<http://fields.scripps.edu/sequest/>) or Mascot (Matrix Sciences) and classifying using DTASelect and Contrast softwares (Tabb et al., 2002). Proteomic phage display techniques are similarly employed to identify target proteins and consensus motifs (Sundell and Ivarsson, 2014). Whole phage shotgun analysis (WSA) is a recently developed technique for protein analysis using NGS platform. It is a culture-independent technique which offers annotation of proteins associated with phages. WSA combines all structural proteins separated on the basis of mass and charge before identification (Lavigne et al., 2006). After separation, the data can be annotated to open reading frames (ORFs) by aligning with reference protein sequences using BLASTP. HHpred is another tool for assigning the protein structure (Hildebrand et al., 2009). The function and evolution of identified proteins can be determined by program COGNITOR ([www.ncbi.nlm.nih.gov/COG](http://www.ncbi.nlm.nih.gov/COG)) and InterProScan to find conserved domains (Eyer et al., 2007). When a predicted protein does not match along known proteins from the database, protein clustering can be developed for the comparative analyses to assess the protein diversity (Hurwitz et al., 2013; Brum et al., 2016). Some software can extract the data from MS and microarray and decipher protein identification using databases such as UniProt (<http://www.uniprot.org/>), PROSITE (<http://prosite.expasy.org/>), Pfam, Conserved Domain and PDB databases. Thus, with an advent in high-throughput proteomic technology, analytical tools, bioinformatics software and database, research on proteins have emerged as an easy task to elucidate protein matter in an environment.

## CATALOGING THE METABOLOME OF VIROSPHERE

The breakdown products of metabolism or intermediates involved in the process of metabolism are termed as metabolites. Metabolites can be (1) primary- which are precisely involved in process of metabolism or (2) secondary- which may not directly take part in the growth of an organism. The metabolome of an organism corresponds to a set of metabolites including hormones, intermediates, signaling and secondary molecules in a particular cell, tissue, organ or an organism (Griffin and Vidal-Puig, 2008; Jordan et al., 2009). To explain the physiology of a particular cell, the study of metabolites is very substantial as every cell possesses a specific metabolic catalog which can influence the accurate implication of function of a cell or an organism (Nicholson and Wilson, 2003; Zhang et al., 2016). These are results of gene transcriptional and

translational mechanisms which remain exceptionally complex, hence variations in metabolites intensify as compared to variations among transcriptome and proteome.

Various approaches have been established for separation and detection of metabolites, chiefly when metabolites are of higher molecular mass. The segregations of metabolites can be carried out using gas chromatography (GC) and high performance liquid chromatography (HPLC), capillary electrophoresis, electron spray ionization (ESI) accompanied by GC, atmospheric-pressure chemical ionization (APCI) on the ground of characteristics of metabolite to be processed (Alonso et al., 2015). Detection of separated metabolites have furthermore been attainable by using nanostructure-initiator MS (NIMS), MALDI-MS, secondary ion mass spectrometry (SIMS), desorption electron spray ionization (DESI), and NMR (Drexler et al., 2007; Cornett et al., 2008; Wiseman et al., 2008; Greer et al., 2011). Statistical tools are additionally applicable for the evaluation of elicited data such as XCMS (Patti et al., 2013), MZmine (Katajamaa et al., 2006), MetAlign (Lommen, 2009), MathDAMP (Baran et al., 2006), and LCMStats<sup>1</sup> (Gahlaut et al., 2013). The metabolic database is available in form of METLIN (Smith et al., 2005).

Metabolomics would serve in interpreting the significance of active phage community on the environment in real time. Based on distinct phases of the phage infection or metabolic profile of the host infected with a phage, the gene markers can be inferred. The modification in host-cell metabolism by phage-encoded genes (vAMGs) into the host genome, is described as a virocell amendment (Rosenwasser et al., 2016). Studying highly specific metabolic profiles of a virocell can improve in interpreting metabolic profile of vAMGs. Such comparisons were conducted to recognize host-viral interactions (Vardi et al., 2009, 2012; Fulton et al., 2014). Metabolomic analysis of phage interprets the influence of the vAMGs which is responsible for enhancing nucleotide biosynthesis (De Smet et al., 2016) via degrading host macromolecules such as DNA through catabolic pathways. The vAMGs encoded nucleases can generally degrade host DNA and encoded triglyceride lipase can degrade host triacylglycerols which yield energy and ultimately engages in the formation of virus membrane (Malitsky et al., 2016). Example, ceramidase in *Mimivirus* helps in the catabolism of sphingolipids (Arslan et al., 2011). Thus, the vAMGs develop the metabolic potential of virocell through triggering novel enzymes which were not present in host machinery prior to phage infection (DeAngelis et al., 1997; Graves et al., 1999). The vAMGs can serve as a shunt between phage and their host by imparting several functional genes from one another especially

assisting during stress conditions (Rosenwasser et al., 2016). These mechanisms illustrate unique attributes of gene products of phage that can mediate dynamics phage-host interaction, as an effect, shaping the microbial communities in an environment. Thus, biochemical composition and metabolic profile of bacterial hosts are greatly governed by phages and released metabolites in the environment influence the microbial food web (Miki et al., 2008). Cataloging the metabolome of phages can elucidate special phage-derived metabolites which usually act as decision making between lytic or lysogenic lifecycle in virocell. Study of metagenomic and metabolomic profiles can simultaneously determine whether the metabolites are encoded by the phage or the host. Furthermore, metabolic profiles of phage can aid in tracing a novel biomarker to recognize the nutrient source in biogeochemical cycles. Thus, the advents in the omics approaches utilizing NGS techniques bear a tremendous potential in exploring virosphere and thus the microbial world.

## CONCLUSIONS

Advancements in the field of NGS have facilitated the discoveries on the verge of a revolution in the course of microbial research. There has been a tremendous microbial data generated about the microbes present on Earth and their diversity and functional roles in regulating the ecosystem. Progress in interpreting the phage diversity and functions as well as the interactions among phages and their hosts are promising using the “omics” concepts. This would illuminate the function of phages in regulating microbial diversity by HGTs, governing the biogeochemical cycles, host population controls and determining the novel biomarkers. NGS will also nurture the upcoming phage therapy research for limiting MDR pathogens. With strong prospects in the field and developments in phage database, “omics” approach is witnessing a remarkable motive for a transformation in the yet uncultivable microbial research.

## AUTHOR CONTRIBUTIONS

KP, PD, and RS conceived and designed the work. KP, SG, and RS wrote the manuscript. PD, RK, and RS carefully checked the manuscript and corrected. All of the authors contributed to the discussion and approved the final manuscript.

## ACKNOWLEDGMENTS

The authors are thankful to Agharkar Research Institute, John Innes Center and Saurashtra University for support and encouragement.

<sup>1</sup>Lcmstats: LCMStats: an [R (programming language)] package for detailed analysis of LCMS data. <http://sourceforge.net/projects/lcmstats.html>.

## REFERENCES

Ackermann, H. W. (2009). Phage classification and characterization. *Methods Mol. Biol.* 501, 127–140. doi: 10.1007/978-1-60327-164-6\_13

Adhikary, A. K., Hanaoka, N., and Fujimoto, T. (2014). Simple and cost-effective restriction endonuclease analysis of human adenoviruses. *Biomed. Res. Int.* 2014:363790. doi: 10.1155/2014/363790



- Akhter, S., Aziz, R. K., and Edwards, R. A. (2012). PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic Acids Res.* 40:e126. doi: 10.1093/nar/gks406
- Allen, L. Z., Ishoey, T., Novotny, M. A., McLean, J. S., Lasken, R. S., and Williamson, S. J. (2011). Single virus genomics: a new tool for virus discovery. *PLoS ONE* 6:e17722. doi: 10.1371/journal.pone.0017722
- Alonso, A., Marsal, S., and Julia, A. (2015). Analytical methods in untargeted metabolomics: state of the art in 2015. *Front. Bioeng. Biotechnol.* 3:23. doi: 10.3389/fbioe.2015.00023
- Anderson, R. E., Brazelton, W. J., and Baross, J. A. (2011). Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *FEMS Microbiol. Ecol.* 77, 120–133. doi: 10.1111/j.1574-6941.2011.01090.x
- Angly, F., Rodriguez-Brito, B., Bangor, D., McNairn, P., Breitbart, M., Salamon, P., et al. (2005). PHACCS, an online tool for estimating the structure and diversity of uncultured viral communities using metagenomic information. *BMC Bioinformatics* 6:41. doi: 10.1186/1471-2105-6-41
- Apel, K., and Hirt, H. (2004). Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu. Rev. Plant Biol.* 55, 373–399. doi: 10.1146/annurev.arplant.55.031903.141701
- Arslan, D., Legendre, M., Seltzer, V., Abergel, C., and Claverie, J. M. (2011). Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proc. Natl. Acad. Sci. U.S.A.* 108, 17486–17491. doi: 10.1073/pnas.1110889108
- Baran, R., Kochi, H., Saito, N., Suematsu, M., Soga, T., Nishioka, T., et al. (2006). MathDAMP: a package for differential analysis of metabolite profiles. *BMC Bioinformatics* 7:530. doi: 10.1186/1471-2105-7-530
- Beitel, C. W., Froenicke, L., Lang, J. M., Korf, I. F., Micheltore, R. W., Eisen, J. A., et al. (2014). Strain- and plasmid-level deconvolution of a synthetic metagenome by sequencing proximity ligation products. *PeerJ* 2:e415. doi: 10.7717/peerj.415
- Berg Miller, M. E., Yeoman, C. J., Chia, N., Tringe, S. G., Angly, F. E., Edwards, R. A., et al. (2012). Phage-bacteria relationships and CRISPR elements revealed by a metagenomic survey of the rumen microbiome. *Environ. Microbiol.* 14, 207–227. doi: 10.1111/j.1462-2920.2011.02593.x
- Bertani, G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* 62, 293–300.
- Bikel, S., Valdez-Lara, A., Cornejo-Granados, F., Rico, K., Canizales-Quinteros, S., Soberon, X., et al. (2015). Combining metagenomics, metatranscriptomics and viromics to explore novel microbial interactions: towards a systems-level understanding of human microbiome. *Comput. Struct. Biotechnol. J.* 13, 390–401. doi: 10.1016/j.csbj.2015.06.001
- Borriss, M., Lombardot, T., Glockner, F. O., Becher, D., Albrecht, D., and Schweder, T. (2007). Genome and proteome characterization of the psychrophilic Flavobacterium bacteriophage 11b. *Extremophiles* 11, 95–104. doi: 10.1007/s00792-006-0014-5
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3
- Branton, D., Deamer, D. W., Marziali, A., Bayley, H., Benner, S. A., Butler, T., et al. (2008). The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* 26, 1146–1153. doi: 10.1038/nbt.1495
- Breitbart, M., Felts, B., Kelley, S., Mahaffy, J. M., Nulton, J., Salamon, P., et al. (2004). Diversity and population structure of a near-shore marine-sediment viral community. *Proc. Biol. Sci.* 271, 565–574. doi: 10.1098/rspb.2003.2628
- Brum, J. R., Ignacio-Espinoza, J. C., Kim, E. H., Trubl, G., Jones, R. M., Roux, S., et al. (2016). Illuminating structural proteins in viral “dark matter” with metaproteomics. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2436–2441. doi: 10.1073/pnas.1525139113
- Brum, J. R., Ignacio-Espinoza, J. C., Roux, S., Doulier, G., Acinas, S. G., Alberti, A., et al. (2015). Ocean plankton. Patterns and ecological drivers of ocean viral communities. *Science* 348:1261498. doi: 10.1126/science.1261498
- Brüssow, H., Canchaya, C., and Hardt, W. D. (2004). Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* 68, 560–602. doi: 10.1128/MMBR.68.3.560-602.2004
- Burton, J. N., Liachko, I., Dunham, M. J., and Shendure, J. (2014). Species-level deconvolution of metagenome assemblies with Hi-C-based contact probability maps. *G3 (Bethesda)* 4, 1339–1346. doi: 10.1534/g3.114.011825
- Carvalho, C. M., Kropinski, A. M., Lingohr, E. J., Santos, S. B., King, J., and Azeredo, J. (2012). The genome and proteome of a *Campylobacter coli* bacteriophage vB\_CcoM-IBB\_35 reveal unusual features. *Virology* 435, 9–35. doi: 10.1016/j.virus.2012.09.009
- Chandramouli, K., and Qian, P.-Y. (2009). Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. *Hum. Genomics Proteomics* 2009:239204. doi: 10.4061/2009/239204
- Clement, C. C., Aphkhasava, D., Nieves, E., Callaway, M., Olszewski, W., Rotzschke, O., et al. (2013). Protein expression profiles of human lymph and plasma mapped by 2D-DIGE and 1D SDS-PAGE coupled with nanoLC-ESI-MS/MS bottom-up proteomics. *J. Proteomics* 78, 172–187. doi: 10.1016/j.jprot.2012.11.013
- Cornett, D. S., Frappier, S. L., and Caprioli, R. M. (2008). MALDI-FTICR imaging mass spectrometry of drugs and metabolites in tissue. *Anal. Chem.* 80, 5648–5653. doi: 10.1021/ac800617s
- Corpillo, D., Gardini, G., Vaira, A. M., Basso, M., Aime, S., Accotto, G. P., et al. (2004). Proteomics as a tool to improve investigation of substantial equivalence in genetically modified organisms: the case of a virus-resistant tomato. *Proteomics* 4, 193–200. doi: 10.1002/pmic.200300540
- Crick, F. (1970). Central dogma of molecular biology. *Nature* 227, 561–563. doi: 10.1038/227561a0
- DeAngelis, P. L., Jing, W., Graves, M. V., Burbank, D. E., and Van Etten, J. L. (1997). Hyaluronan synthase of chlorella virus PBCV-1. *Science* 278, 1800–1803. doi: 10.1126/science.278.5344.1800
- De Smet, J., Zimmermann, M., Kogadeeva, M., Ceysens, P. J., Vermaelen, W., Blasdel, B., et al. (2016). High coverage metabolomics analysis reveals phage-specific alterations to *Pseudomonas aeruginosa* physiology during infection. *ISME J.* 10, 1823–1835. doi: 10.1038/ismej.2016.3
- Drexler, D. M., Garrett, T. J., Cantone, J. L., Ditters, R. W., Mitroka, J. G., Prieto Conaway, M. C., et al. (2007). Utility of imaging mass spectrometry (IMS) by matrix-assisted laser desorption/ionization (MALDI) on an ion trap mass spectrometer in the analysis of drugs and metabolites in biological tissues. *J. Pharmacol. Toxicol. Methods* 55, 279–288. doi: 10.1016/j.vascn.2006.11.004
- Drulis-Kawa, Z., Majkowska-Skrobek, G., Maciejewska, B., Delattre, A. S., and Lavigne, R. (2012). Learning from bacteriophages—advantages and limitations of phage and phage-encoded protein applications. *Curr. Protein Pept. Sci.* 13, 699–722. doi: 10.2174/138920312804871193
- Duhaime, M. B., Deng, L., Poulos, B. T., and Sullivan, M. B. (2012). Towards quantitative metagenomics of wild viruses and other ultra-low concentration DNA samples: a rigorous assessment and optimization of the linker amplification method. *Environ. Microbiol.* 14, 2526–2537. doi: 10.1111/j.1462-2920.2012.02791.x
- Dutilh, B. E., Cassman, N., McNair, K., Sanchez, S. E., Silva, G. G., Boling, L., et al. (2014). A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nat. Commun.* 5:4498. doi: 10.1038/ncomms5498
- Dutta, C., and Pan, A. (2002). Horizontal gene transfer and bacterial diversity. *J. Biosci.* 27, 27–33. doi: 10.1007/BF02703681
- Eberwine, J., Sul, J. Y., Bartfai, T., and Kim, J. (2014). The promise of single-cell sequencing. *Nat. Methods* 11, 25–27. doi: 10.1038/nmeth.2769
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461
- Edwards, R. A., McNair, K., Faust, K., Raes, J., and Dutilh, B. E. (2016). Computational approaches to predict bacteriophage-host relationships. *FEMS Microbiol. Rev.* 40, 258–272. doi: 10.1093/femsre/fuv048
- Edwards, R. A., and Rohwer, F. (2005). Viral metagenomics. *Nat. Rev. Microbiol.* 3, 504–510. doi: 10.1038/nrmicro1163
- Eyer, L., Pantucek, R., Zdrahal, Z., Konecna, H., Kasperek, P., Ruzickova, V., et al. (2007). Structural protein analysis of the polyvalent staphylococcal bacteriophage 812. *Proteomics* 7, 64–72. doi: 10.1002/pmic.2006.00280



- Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367
- Franzosa, E. A., Morgan, X. C., Segata, N., Waldron, L., Reyes, J., Earl, A. M., et al. (2014). Relating the metatranscriptome and metagenome of the human gut. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2329–E2338. doi: 10.1073/pnas.1319284111
- Fulton, J. M., Fredricks, H. F., Bidle, K. D., Vardi, A., Kendrick, B. J., Ditullio, G. R., et al. (2014). Novel molecular determinants of viral susceptibility and resistance in the lipidome of *Emiliania huxleyi*. *Environ. Microbiol.* 16, 1137–1149. doi: 10.1111/1462-2920.12358
- Gahlaut, A., Vikas, D., Dahiya, M., Gothwal, A., Kulharia, M., Chhillar, A. K., et al. (2013). Proteomics & metabolomics: mapping biochemical regulations. *Drug Invention Today* 5, 321–326. doi: 10.1016/j.dit.2013.08.007
- Gerlach, W., Junemann, S., Tille, F., Goesmann, A., and Stoye, J. (2009). WebCARMA: a web application for the functional and taxonomic classification of unassembled metagenomic reads. *BMC Bioinformatics* 10:430. doi: 10.1186/1471-2105-10-430
- Ghaffari, N., Sanchez-Flores, A., Doan, R., Garcia-Orozco, K. D., Chen, P. L., Ochoa-Leyva, A., et al. (2014). Novel transcriptome assembly and improved annotation of the whiteleg shrimp (*Litopenaeus vannamei*), a dominant crustacean in global seafood mariculture. *Sci. Rep.* 4:7081. doi: 10.1038/srep07081
- Glass, E. M., Wilkening, J., Wilke, A., Antonopoulos, D., and Meyer, F. (2010). Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. *Cold Spring Harb. Protoc.* 2010:pdb.prot5368. doi: 10.1101/pdb.prot5368
- Goff, S. A., Vaughn, M., McKay, S., Lyons, E., Stapleton, A. E., Gessler, D., et al. (2011). The iPlant collaborative: cyberinfrastructure for plant biology. *Front. Plant Sci.* 2:34. doi: 10.3389/fpls.2011.00034
- Gonzalez, J. M., Portillo, M. C., and Saiz-Jimenez, C. (2005). Multiple displacement amplification as a pre-polymerase chain reaction (pre-PCR) to process difficult to amplify samples and low copy number sequences from natural environments. *Environ. Microbiol.* 7, 1024–1028. doi: 10.1111/j.1462-2920.2005.00779.x
- Graves, M. V., Burbank, D. E., Roth, R., Heuser, J., Deangelis, P. L., and Van Etten, J. L. (1999). Hyaluronan synthesis in virus PBCV-1-infected chlorella-like green algae. *Virology* 257, 15–23. doi: 10.1006/viro.1999.9628
- Greer, T., Sturm, R., and Li, L. (2011). Mass spectrometry imaging for drugs and metabolites. *J. Proteomics* 74, 2617–2631. doi: 10.1016/j.jprot.2011.03.032
- Greninger, A. L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., et al. (2015). Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Med.* 7:99. doi: 10.1186/s13073-015-0220-9
- Griffin, J. L., and Vidal-Puig, A. (2008). Current challenges in metabolomics for diabetes research: a vital functional genomic tool or just a ploy for gaining funding? *Physiol. Genomics* 34, 1–5. doi: 10.1152/physiolgenomics.0009.2008
- Haider, S., and Pal, R. (2013). Integrated analysis of transcriptomic and proteomic data. *Curr. Genomics* 14, 91–110. doi: 10.2174/1389202911314020003
- Halleran, A., Clamons, S., and Saha, M. (2015). Transcriptomic characterization of an infection of *Mycobacterium smegmatis* by the cluster a4 mycobacteriophage kamy. *PLoS ONE* 10:e0141100. doi: 10.1371/journal.pone.0141100
- Handelsman, J. (2004). Metagenomics: application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* 68, 669–685. doi: 10.1128/MMBR.68.4.669-685.2004
- Hatfull, G. F., Pedulla, M. L., Jacobs-Sera, D., Cichon, P. M., Foley, A., Ford, M. E., et al. (2006). Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genet.* 2:e92. doi: 10.1371/journal.pgen.0020092
- Helms, C., Graham, M. Y., Dutchik, J. E., and Olson, M. V. (1985). A new method for purifying lambda DNA from phage lysates. *DNA* 4, 39–49. doi: 10.1089/dna.1985.4.39
- Hickman, S. E., Kingery, N. D., Ohsumi, T. K., Borowsky, M. L., Wang, L. C., Means, T. K., et al. (2013). The microglial sensome revealed by direct RNA sequencing. *Nat. Neurosci.* 16, 1896–1905. doi: 10.1038/nn.3554
- Hildebrand, A., Rimmert, M., Biegert, A., and Söding, J. (2009). Fast and accurate automatic structure prediction with HHpred. *Proteins* 77 (Suppl. 9), 128–132. doi: 10.1002/prot.22499
- Holmfeldt, K., Solonenko, N., Shah, M., Corrier, K., Riemann, L., Verberkmoes, N. C., et al. (2013). Twelve previously unknown phage genera are ubiquitous in global oceans. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12798–12803. doi: 10.1073/pnas.1305956110
- Horgan, R. P., and Kenny, L. C. (2011). ‘Omic’ technologies: genomics, transcriptomics, proteomics and metabolomics. *J. Obstet. Gynaecol.* 13, 189–195. doi: 10.1576/toag.13.3.189.27672
- Hosono, S., Faruqi, A. F., Dean, F. B., Du, Y., Sun, Z., Wu, X., et al. (2003). Unbiased whole-genome amplification directly from clinical samples. *Genome Res.* 13, 954–964. doi: 10.1101/gr.816903
- Hurwitz, B. L., Brum, J. R., and Sullivan, M. B. (2015). Depth-stratified functional and taxonomic niche specialization in the ‘core’ and ‘flexible’ Pacific Ocean virome. *ISME J.* 9, 472–484. doi: 10.1038/ismej.2014.143
- Hurwitz, B. L., Hallam, S. J., and Sullivan, M. B. (2013). Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biol.* 14:R123. doi: 10.1186/gb-2013-14-11-r123
- Hurwitz, B. L., Uren, J. M., and Youens-Clark, K. (2016). Computational prospecting the great viral unknown. *FEMS Microbiol. Lett.* 363:fnw077. doi: 10.1093/femsle/fnw077
- Hurwitz, B. L., Westveld, A. H., Brum, J. R., and Sullivan, M. B. (2014). Modeling ecological drivers in marine viral communities using comparative metagenomics and network analyses. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10714–10719. doi: 10.1073/pnas.1319778111
- Huson, D. H., Auch, A. F., Qi, J., and Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Res.* 17, 377–386. doi: 10.1101/gr.5969107
- Ishoy, T., Kvist, T., Westermann, P., and Ahning, B. K. (2006). An improved method for single cell isolation of prokaryotes from meso-, thermo- and hyperthermophilic environments using micromanipulation. *Appl. Microbiol. Biotechnol.* 69, 510–514. doi: 10.1007/s00253-005-0014-x
- Jain, R., and Srivastava, R. (2009). Metabolic investigation of host/pathogen interaction using MS2-infected *Escherichia coli*. *BMC Syst. Biol.* 3:121. doi: 10.1186/1752-0509-3-121
- Jordan, K. W., Nordenstam, J., Lauwers, G. Y., Rothenberger, D. A., Alavi, K., Garwood, M., et al. (2009). Metabolomic characterization of human rectal adenocarcinoma with intact tissue magnetic resonance spectroscopy. *Dis. Colon Rectum* 52, 520–525. doi: 10.1007/DCR.0b013e31819c9a2c
- Juhala, R. J., Ford, M. E., Duda, R. L., Youlton, A., Hatfull, G. F., and Hendrix, R. W. (2000). Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdaoid bacteriophages. *J. Mol. Biol.* 299, 27–51. doi: 10.1006/jmbi.2000.3729
- Kalisky, T., and Quake, S. R. (2011). Single-cell genomics. *Nat. Methods* 8, 311–314. doi: 10.1038/nmeth0411-311
- Katajamaa, M., Miettinen, J., and Oresic, M. (2006). MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics* 22, 634–636. doi: 10.1093/bioinformatics/btk039
- Khan Mirzaei, M., Eriksson, H., Kasuga, K., Haggard-Ljungquist, E., and Nilsson, A. S. (2014). Genomic, proteomic, morphological, and phylogenetic analyses of vB\_EcoP\_SU10, a podoviridae phage with C3 morphology. *PLoS ONE* 9:e116294. doi: 10.1371/journal.pone.0116294
- Krishnamurthy, S. R., Janowski, A. B., Zhao, G., Barouch, D., and Wang, D. (2016). Hyperexpansion of RNA bacteriophage diversity. *PLoS Biol.* 14:e1002409. doi: 10.1371/journal.pbio.1002409
- Kropinski, A. M., Waddell, T., Meng, J., Franklin, K., Ackermann, H. W., Ahmed, R., et al. (2013). The host-range, genomics and proteomics of *Escherichia coli* O157:H7 bacteriophage rV5. *Virol. J.* 10:76. doi: 10.1186/1743-422X-10-76
- Labonté, J. M., Swan, B. K., Poulos, B., Luo, H., Koren, S., Hallam, S. J., et al. (2015). Single-cell genomics-based analysis of virus-host interactions in marine surface bacterioplankton. *ISME J.* 9, 2386–2399. doi: 10.1038/ismej.2015.48

- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25
- Lasken, R. S., and McLean, J. S. (2014). Recent advances in genomic DNA sequencing of microbial species from single cells. *Nat. Rev. Genet.* 15, 577–584. doi: 10.1038/nrg3785
- Lavigne, R., Noben, J. P., Hertveldt, K., Ceysens, P. J., Briers, Y., Dumont, D., et al. (2006). The structural proteome of *Pseudomonas aeruginosa* bacteriophage phiKMV. *Microbiology* 152, 529–534. doi: 10.1099/mic.0.28431-0
- Leskinen, K., Blasdel, B. G., Lavigne, R., and Skurnik, M. (2016). RNA-sequencing reveals the progression of phage-host interactions between phir1-37 and *Yersinia enterocolitica*. *Viruses* 8:111. doi: 10.3390/v8040111
- Li, M., Wang, B., Zhang, M., Rantalainen, M., Wang, S., Zhou, H., et al. (2008). Symbiotic gut microbes modulate human metabolic phenotypes. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2117–2122. doi: 10.1073/pnas.0712038105
- Lim, Y. W., Schmieder, R., Haynes, M., Willner, D., Furlan, M., Youle, M., et al. (2013). Metagenomics and metatranscriptomics: windows on CF-associated viral and microbial communities. *J. Cyst. Fibros.* 12, 154–164. doi: 10.1016/j.jcf.2012.07.009
- Lin, X., Ding, H., and Zeng, Q. (2016). Transcriptomic response during phage infection of a marine cyanobacterium under phosphorus-limited conditions. *Environ. Microbiol.* 18, 450–460. doi: 10.1111/1462-2920.13104
- Liu, B., Gibbons, T., Ghodsi, M., and Pop, M. (2010). “MetaPhyler: taxonomic profiling for metagenomic sequences,” in *IEEE International Conference on Bioinformatics and Biomedicine* (Hong-Kong: IEEE), 95–100.
- Lommen, A. (2009). MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. *Anal. Chem.* 81, 3079–3086. doi: 10.1021/ac900036d
- Luef, B., Neu, T. R., and Peduzzi, P. (2009). Imaging and quantifying virus fluorescence signals on aquatic aggregates: a new method and its implication for aquatic microbial ecology. *FEMS Microbiol. Ecol.* 68, 372–380. doi: 10.1111/j.1574-6941.2009.00675.x
- Malitsky, S., Ziv, C., Rosenwasser, S., Zheng, S., Schatz, D., Porat, Z., et al. (2016). Viral infection of the marine alga *Emiliania huxleyi* triggers lipidome remodeling and induces the production of highly saturated triacylglycerol. *New Phytol.* 210, 88–96. doi: 10.1111/nph.13852
- Mariano, R., Wuchty, S., Vizoso-Pinto, M. G., Hauser, R., and Uetz, P. (2016). The interactome of *Streptococcus pneumoniae* and its bacteriophages show highly specific patterns of interactions among bacteria and their phages. *Sci. Rep.* 6:24597. doi: 10.1038/srep24597
- Martinez-Garcia, M., Santos, F., Moreno-Paz, M., Parro, V., and Antón, J. (2014). Unveiling viral-host interactions within the ‘microbial dark matter’. *Nat. Commun.* 5:4542. doi: 10.1038/ncomms5542
- Marzano, S. Y., Nelson, B. D., Ajayi-Oyetunde, O., Bradley, C. A., Hughes, T. J., Hartman, G. L., et al. (2016). Identification of diverse mycoviruses through metatranscriptomics characterization of the viromes of five major fungal plant pathogens. *J. Virol.* 90, 6846–6863. doi: 10.1128/JVI.00357-16
- Meyer, F., Paarmann, D., D’Souza, M., Olson, R., Glass, E. M., Kubal, M., et al. (2008). The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9:386. doi: 10.1186/1471-2105-9-386
- Mikheyev, A. S., and Tin, M. M. (2014). A first look at the Oxford Nanopore MinION sequencer. *Mol. Ecol. Resour.* 14, 1097–1102. doi: 10.1111/1755-0998.12324
- Miki, T., Nakazawa, T., Yokokawa, T., and Nagata, T. (2008). Functional consequences of viral impacts on bacterial communities: a food-web model analysis. *Freshw. Biol.* 53, 1142–1153. doi: 10.1111/j.1365-2427.2007.01934.x
- Minot, S., Sinha, R., Chen, J., Li, H., Keilbaugh, S. A., Wu, G. D., et al. (2011). The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res.* 21, 1616–1625. doi: 10.1101/gr.122705.111
- Mizuno, C. M., Rodriguez-Valera, F., Kimes, N. E., and Ghai, R. (2013). Expanding the marine virosphere using metagenomics. *PLoS Genet.* 9:e1003987. doi: 10.1371/journal.pgen.1003987
- Mutz, K. O., Heilkenbrinker, A., Lonne, M., Walter, J. G., and Stahl, F. (2013). Transcriptome analysis using next-generation sequencing. *Curr. Opin. Biotechnol.* 24, 22–30. doi: 10.1016/j.copbio.2012.09.004
- Nicholson, J. K., and Wilson, I. D. (2003). Opinion: understanding ‘global’ systems biology: metabonomics and the continuum of metabolism. *Nat. Rev. Drug Discov.* 2, 668–676. doi: 10.1038/nrd1157
- Niu, Y. D., Stanford, K., Kropinski, A. M., Ackermann, H. W., Johnson, R. P., She, Y. M., et al. (2012). Genomic, proteomic and physiological characterization of a T5-like bacteriophage for control of Shiga toxin-producing *Escherichia coli* O157:H7. *PLoS ONE* 7:e34585. doi: 10.1371/journal.pone.0034585
- Ounit, R., Wanamaker, S., Close, T. J., and Lonardi, S. (2015). CLARK: fast and accurate classification of metagenomic and genomic sequences using discriminative k-mers. *BMC Genomics* 16:236. doi: 10.1186/s12864-015-1419-2
- Ozsolak, F., and Milos, P. M. (2011). RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet.* 12, 87–98. doi: 10.1038/nrg2934
- Parmar, K. M., Hathi, Z. J., and Dafale, N. A. (2017). Control of multidrug-resistant gene flow in the environment through bacteriophage intervention. *Appl Biochem Biotechnol.* 181, 1007–1029. doi: 10.1007/s12010-016-2265-7
- Patti, G. J., Tautenhahn, R., Rinehart, D., Cho, K., Shriver, L. P., Manchester, M., et al. (2013). A view from above: cloud plots to visualize global metabolomic data. *Anal. Chem.* 85, 798–804. doi: 10.1021/ac3029745
- Petricoin, E. F., Zoon, K. C., Kohn, E. C., Barrett, J. C., and Liotta, L. A. (2002). Clinical proteomics: translating benchside promise into bedside reality. *Nat. Rev. Drug Discov.* 1, 683–695. doi: 10.1038/nrd891
- Picot, J., Guerin, C. L., Le Van Kim, C., and Boulanger, C. M. (2012). Flow cytometry: retrospective, fundamentals and recent instrumentation. *Cytotechnology* 64, 109–130. doi: 10.1007/s10616-011-9415-0
- Podar, M., Abulencia, C. B., Walcher, M., Hutchison, D., Zengler, K., Garcia, J. A., et al. (2007). Targeted access to the genomes of low-abundance organisms in complex microbial communities. *Appl. Environ. Microbiol.* 73, 3205–3214. doi: 10.1128/AEM.02985-06
- Prestel, E., Salamitou, S., and Dubow, M. S. (2008). An examination of the bacteriophages and bacteria of the Namib desert. *J. Microbiol.* 46, 364–372. doi: 10.1007/s12275-008-0007-4
- Prigent, M., Leroy, M., Confalonieri, F., Dutertre, M., and Dubow, M. S. (2005). A diversity of bacteriophage forms and genomes can be isolated from the surface sands of the Sahara Desert. *Extremophiles* 9, 289–296. doi: 10.1007/s00792-005-0444-5
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. doi: 10.1038/nature08821
- Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., et al. (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13:341. doi: 10.1186/1471-2164-13-341
- Renesto, P., Abergel, C., Decloquement, P., Moinier, D., Azza, S., Ogata, H., et al. (2006). Mimivirus giant particles incorporate a large fraction of anonymous and unique gene products. *J. Virol.* 80, 11678–11685. doi: 10.1128/JVI.00940-06
- Rodrigue, S., Malmstrom, R. R., Berlin, A. M., Birren, B. W., Henn, M. R., and Chisholm, S. W. (2009). Whole genome amplification and *de novo* assembly of single bacterial cells. *PLoS ONE* 4:e6864. doi: 10.1371/journal.pone.0006864
- Rohwer, F., and Thurber, R. V. (2009). Viruses manipulate the marine environment. *Nature* 459, 207–212. doi: 10.1038/nature08060
- Rosen, G. L., Reichenberger, E. R., and Rosenfeld, A. M. (2011). NBC: the Naive Bayes Classification tool webserver for taxonomic classification of metagenomic reads. *Bioinformatics* 27, 127–129. doi: 10.1093/bioinformatics/btq619
- Rosenwasser, S., Ziv, C., Creveld, S. G., and Vardi, A. (2016). Virocell metabolism: metabolic innovations during host-virus interactions in the ocean. *Trends Microbiol.* 24, 821–832. doi: 10.1016/j.tim.2016.06.006
- Roux, S., Enault, F., Bronner, G., Vault, D., Forterre, P., and Krupovic, M. (2013a). Chimeric viruses blur the borders between the major groups of eukaryotic single-stranded DNA viruses. *Nat. Commun.* 4:2700. doi: 10.1038/ncomms3700
- Roux, S., Hawley, A. K., Torres Beltran, M., Scofield, M., Schwientek, P., Stepanauskas, R., et al. (2014). Ecology and evolution of viruses infecting uncultivated SUP05 bacteria as revealed by single-cell- and meta-genomics. *eLife* 3:e03125. doi: 10.7554/eLife.03125
- Roux, S., Krupovic, M., Debroas, D., Forterre, P., and Enault, F. (2013b). Assessment of viral community functional potential from viral metagenomes

- may be hampered by contamination with cellular sequences. *Open Biol.* 3:130160. doi: 10.1098/rsob.130160
- Sanchez, S. E., Cuevas, D. A., Rostron, J. E., Liang, T. Y., Pivaroff, C. G., Haynes, M. R., et al. (2015). Phage phenomics: physiological approaches to characterize novel viral proteins. *J. Vis. Exp.* e52854. doi: 10.3791/52854
- Sangha, K. K., Kumar, B. V., Agrawal, R. K., Deka, D., and Verma, R. (2014). Proteomic characterization of lytic bacteriophages of *Staphylococcus aureus* Isolated from Sewage Effluent of India. *Int. Sch. Res. Notices* 2014:265298. doi: 10.1155/2014/265298
- Santos, F., Martínez-García, M., Parro, V., and Antón, J. (2015). Microarray tools to unveil viral-microbe interactions in nature. *Front. Ecol. Evol.* 3:31. doi: 10.3389/fevo.2014.00031
- Santos, F., Moreno-Paz, M., Meseguer, I., Lopez, C., Rossello-Mora, R., Parro, V., et al. (2011). Metatranscriptomic analysis of extremely halophilic viral communities. *ISME J.* 5, 1621–1633. doi: 10.1038/ismej.2011.34
- Santos, F., Yarza, P., Parro, V., Meseguer, I., Rossello-Mora, R., and Anton, J. (2012). Culture-independent approaches for studying viruses from hypersaline environments. *Appl. Environ. Microbiol.* 78, 1635–1643. doi: 10.1128/AEM.07175-11
- Santos, Í. M., Da Rosa, E. A., Graf, T., Ferreira, L. G., Petry, A., Cavalheiro, F., et al. (2015). Analysis of immunological, viral, genetic, and environmental factors that might be associated with decreased susceptibility to HIV infection in serodiscordant couples in Florianópolis, Southern Brazil. *AIDS Res. Hum. Retroviruses* 31, 1116–1125. doi: 10.1089/aid.2015.0168
- Schleicher, E., and Wieland, O. H. (1978). Evaluation of the Bradford method for protein determination in body fluids. *J. Clin. Chem. Clin. Biochem.* 16, 533–534.
- Schoenfeld, T., Liles, M., Wommack, K. E., Polson, S. W., Godiska, R., and Mead, D. (2010). Functional viral metagenomics and the next generation of molecular tools. *Trends Microbiol.* 18, 20–29. doi: 10.1016/j.tim.2009.10.001
- Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., et al. (2011). Global quantification of mammalian gene expression control. *Nature* 473, 337–342. doi: 10.1038/nature10098
- Seal, B. S., Fouts, D. E., Simmons, M., Garrish, J. K., Kuntz, R. L., Woolsey, R., et al. (2011). Clostridium perfringens bacteriophages PhiCP39O and PhiCP26F: genomic organization and proteomic analysis of the virions. *Arch. Virol.* 156, 25–35. doi: 10.1007/s00705-010-0812-z
- Short, C. M., and Suttle, C. A. (2005). Nearly identical bacteriophage structural gene sequences are widely distributed in both marine and freshwater environments. *Appl. Environ. Microbiol.* 71, 480–486. doi: 10.1128/AEM.71.1.480-486.2005
- Sime-Ngando, T., and Colombet, J. (2009). [Virus and prophages in aquatic ecosystems]. *Can. J. Microbiol.* 55, 95–109. doi: 10.1139/W08-099
- Smith, C. A., O'maille, G., Want, E. J., Qin, C., Trauger, S. A., Brandon, T. R., et al. (2005). METLIN: a metabolite mass spectral database. *Ther. Drug Monit.* 27, 747–751. doi: 10.1097/01.ftd.0000179845.53213.39
- Soares, S. C., Abreu, V. A., Ramos, R. T., Cerdeira, L., Silva, A., Baumbach, J., et al. (2012). PIPS: pathogenicity island prediction software. *PLoS ONE* 7:e30848. doi: 10.1371/journal.pone.0030848
- Srinivasiah, S., Bhavsar, J., Thapar, K., Liles, M., Schoenfeld, T., and Wommack, K. E. (2008). Phages across the biosphere: contrasts of viruses in soil and aquatic environments. *Res. Microbiol.* 159, 349–357. doi: 10.1016/j.resmic.2008.04.010
- Sullivan, M. B., Krastins, B., Hughes, J. L., Kelly, L., Chase, M., Sarracino, D., et al. (2009). The genome and structural proteome of an ocean siphovirus: a new window into the cyanobacterial 'mobilome'. *Environ. Microbiol.* 11, 2935–2951. doi: 10.1111/j.1462-2920.2009.02081.x
- Sundell, G. N., and Ivarsson, Y. (2014). Interaction analysis through proteomic phage display. *Biomed. Res. Int.* 2014:176172. doi: 10.1155/2014/176172
- Suttle, C. A. (2007). Marine viruses—major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi: 10.1038/nrmicro1750
- Tabb, D. L., McDonald, W. H., and Yates, J. R. III. (2002). DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* 1, 21–26. doi: 10.1021/pr015504q
- Tamakoshi, M., Murakami, A., Sugisawa, M., Tsuneizumi, K., Takeda, S., Saheki, T., et al. (2011). Genomic and proteomic characterization of the large Myoviridae bacteriophage varphiTMA of the extreme thermophile *Thermus thermophilus*. *Bacteriophage* 1, 152–164. doi: 10.4161/bact.1.3.16712
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst1197
- Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* 6, 377–382. doi: 10.1038/nmeth.1315
- Theodorescu, D., and Mischak, H. (2007). Mass spectrometry based proteomics in urine biomarker discovery. *World J. Urol.* 25, 435–443. doi: 10.1007/s00345-007-0206-3
- Thurber, R. V., Haynes, M., Breitbart, M., Wegley, L., and Rohwer, F. (2009). Laboratory procedures to generate viral metagenomes. *Nat. Protoc.* 4, 470–483. doi: 10.1038/nprot.2009.10
- Uzoma, I., and Zhu, H. (2013). Interactome mapping: using protein microarray technology to reconstruct diverse protein networks. *Genomics Proteomics Bioinformatics* 11, 18–28. doi: 10.1016/j.gpb.2012.12.005
- van Berkum, N. L., Lieberman-Aiden, E., Williams, L., Imakaev, M., Gnirke, A., Mirny, L. A., et al. (2010). Hi-C: a method to study the three-dimensional architecture of genomes. *J. Vis. Exp.* 39:e1869 doi: 10.3791/1869
- Van den Bossche, A., Ceyssens, P.-J., De Smet, J., Hendrix, H., Bellon, H., Leimer, N., et al. (2014). Systematic identification of hypothetical bacteriophage proteins targeting key protein complexes of *Pseudomonas aeruginosa*. *J. Proteome Res.* 13, 4446–4456. doi: 10.1021/pr500796n
- Vardi, A., Haramaty, L., Van Mooy, B. A., Fredricks, H. F., Kimmance, S. A., Larsen, A., et al. (2012). Host-virus dynamics and subcellular controls of cell fate in a natural coccolithophore population. *Proc. Natl. Acad. Sci. U.S.A.* 109, 19327–19332. doi: 10.1073/pnas.1208895109
- Vardi, A., Van Mooy, B. A., Fredricks, H. F., Pependorf, K. J., Ossolinski, J. E., Haramaty, L., et al. (2009). Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton. *Science* 326, 861–865. doi: 10.1126/science.1177322
- Vernikos, G. S., and Parkhill, J. (2006). Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the Salmonella pathogenicity islands. *Bioinformatics* 22, 2196–2203. doi: 10.1093/bioinformatics/btl369
- Veses-Garcia, M., Liu, X., Rigden, D. J., Kenny, J. G., McCarthy, A. J., and Allison, H. E. (2015). Transcriptomic analysis of Shiga-toxicogenic bacteriophage carriage reveals a profound regulatory effect on acid resistance in *Escherichia coli*. *Appl. Environ. Microbiol.* 81, 8118–8125. doi: 10.1128/AEM.02034-15
- Villegas, A., She, Y. M., Kropinski, A. M., Lingohr, E. J., Mazzocco, A., Ojha, S., et al. (2009). The genome and proteome of a virulent *Escherichia coli* O157:H7 bacteriophage closely resembling Salmonella phage Felix O1. *Virol. J.* 6:41. doi: 10.1186/1743-422X-6-41
- Vlahou, A., and Fountoulakis, M. (2005). Proteomic approaches in the search for disease biomarkers. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* 814, 11–19. doi: 10.1016/j.jchromb.2004.10.024
- Voorhies, A. A., Eisenlord, S. D., Marcus, D. N., Duhaime, M. B., Biddanda, B. A., Cavalcoli, J. D., et al. (2016). Ecological and genetic interactions between cyanobacteria and viruses in a low-oxygen mat community inferred through metagenomics and metatranscriptomics. *Environ. Microbiol.* 18, 358–371. doi: 10.1111/1462-2920.12756
- Weinbauer, M. G., and Rassoulzadegan, F. (2004). Are viruses driving microbial diversification and diversity? *Environ. Microbiol.* 6, 1–11. doi: 10.1046/j.1462-2920.2003.00539.x
- Weston, A. D., and Hood, L. (2004). Systems biology, proteomics, and the future of health care: toward predictive, preventative, and personalized medicine. *J. Proteome Res.* 3, 179–196. doi: 10.1021/pr0499693
- Wiseman, J. M., Ifa, D. R., Zhu, Y., Kissinger, C. B., Manicke, N. E., Kissinger, P. T., et al. (2008). Desorption electrospray ionization mass spectrometry: imaging drugs and metabolites in tissues. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18120–18125. doi: 10.1073/pnas.0801066105
- Wommack, K. E., and Colwell, R. R. (2000). Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69–114. doi: 10.1128/MMBR.64.1.69-114.2000
- Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15:R46. doi: 10.1186/gb-2014-15-3-r46
- Yilmaz, S., Allgaier, M., and Hugenoltz, P. (2010). Multiple displacement amplification compromises quantitative analysis of

- metagenomes. *Nat. Methods* 7, 943–944. doi: 10.1038/nmeth1210-943
- Yuan, Y., and Gao, M. (2016). Proteomic analysis of a novel bacillus jumbo phage revealing glycoside hydrolase as structural component. *Front. Microbiol.* 7:745. doi: 10.3389/fmicb.2016.00745
- Zhang, A., Sun, H., Yan, G., Wang, P., and Wang, X. (2016). Mass spectrometry-based metabolomics: applications to biomarker and metabolic pathway research. *Biomed. Chromatogr.* 30, 7–12. doi: 10.1002/bmc.3453
- Zhao, X., Chen, C., Shen, W., Huang, G., Le, S., Lu, S., et al. (2016). Global transcriptomic analysis of interactions between *Pseudomonas aeruginosa* and Bacteriophage PaP3. *Sci. Rep.* 6:19237. doi: 10.1038/srep19237

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Parmar, Gaikwad, Dhakephalkar, Kothari and Singh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.