

Single-Cell Genome Sequencing for Viral-Host Interactions

Vikram Dhillon and Xiaoman Li*

Burnett School of Biomedical Science, College of Medicine, University of Central Florida, Orlando, Florida 32816, USA

Abstract

Viruses are the most abundant biological entities and infectious agents present in almost every ecosystem on the planet. Yet our understanding of how viral-mediated gene transfer and metabolic reprogramming influence the evolutionary history of their hosts and microbial communities remains poor at best. At the same time, identifying and modeling the community dynamics of viruses from the environment through conventional plaque assays is complicated because less than one percent of microbial hosts have been cultivated in vitro. Computational methods in metagenomics and phage isolation techniques have limitations in identifying the uncultured hosts of most viruses. Moreover, the model system-based measurements derived from such techniques rarely reflect the network properties of natural microbial communities. To address these problems, development of high-throughput, massively parallel sequencing approaches that do not rely on cultivation to identify specific virus-host relations such as single-cell genomic sequencing (SCGS) has become critical. SCGS has advanced our capacity to understand the genomic and transcriptomic diversity that occurs during viral-host interactions in an individual uncultured host. Here, we review the major technological and biological challenges and the breakthroughs achieved, describe the remaining challenges, and provide a glimpse into the recent advancements.

Keywords: Metagenomics; Single-cell-genomics; Sequencing; Viral-host interactions; Next-generation sequencing

Introduction

Current State of Metagenomics

The heterogeneity of microbial communities has historically been ignored in metagenomic studies. Most of our current understanding about the dynamics of natural microbial communities has been derived from studies carried out on bulk population: Generally millions of cells are collected and analyzed in an ecological survey [1,2]. In the past, the outcome of such analyses answered the questions in the form, “What is there?” to address the “dark matter” [3-5] of microbial species found in a particular environment and sample their genetic composition. The data obtained from such population studies is processed using metagenomic tools which are not well-suited to answer questions about the individual components of a species. These questions entail topics such as the organization of the genes, evolutionary history of the organisms in the community and their metabolic exchange-repertoire [3,6]. Recently, the interest has shifted towards answering questions of the form “What does this particular part of the community do?” Consequently, these topics address the immense complexity of network-based interactions in a microbial community and the diversity within a species present in that environment [7-9]. The advent of single cell genomics (SCG) has improved our understanding and augmented our ability to answer these questions by making it accessible to compare the differences between individual cells. The practical advantages of single cell techniques were demonstrated with the recent applications of SCG in identifying copy-number variations in the human brain neurons and their consequences for neural cells in pathogenic states [10,11].

The success of SCG approaches resulted from major breakthroughs in two different areas: It was in part due to the advances in next-generation sequencing and another part was due to new extraction and isolation techniques of biological samples [12-14]. SCG provides access to sequences of all DNA in the analyzed cell, including chromosome material, plasmids, and pathogens. The genomic material of individual uncultured cells is amplified by techniques such as multiple displacement amplification (MDA) and whole-genome

sequencing is carried out to recover viral genomes for processing [6,12,15,16]. Further analysis of the data obtained from the sequencing of the cellular components has allowed for accurate and uncultured viral-host pairing. In parallel to sequencing, new flow cytometry and microfluidic manipulation techniques have allowed for single-cell isolation from most environments [14,17]. Despite these advances in SCG approaches, many challenges remain to be addressed before SCG becomes a standard technique for researchers.

The exponential growth in the amount of biological data being generated and curated from single-cell studies will require drastically new measures for data management, analysis and accessibility [6]. High-throughput platforms analyzing whole genome amplification products of single-cell isolates generate a tremendous amount of data. The interpretation of this data in an efficient and fast manner to distinguish true genomic variants from the background noise presents a major challenge for bioinformatics [18,19]. To investigate subtle biological heterogeneity and diversity between collected samples, precise methods with high sensitivity are required due to limiting amount of genetic material [20,21]. In addition, to enable single-cell genome sequencing of the uncultured microbial hosts or samples, a number of technical issues critical to the pipeline must be first addressed. These include problems such as the removal of background noise, amplification bias and errors, and the compatibility with currently available genome sequencing pipelines [2,20,22,23]. Although next-generation sequencing devices provide us with high-resolution genomic data-maps from single cells, identification of unique elements to viral predation and reprogramming such as chromosomal variants

*Corresponding author: Xiaoman Li, Burnett School of Biomedical Science, College of Medicine, University of Central Florida, Orlando, Florida 32816, USA, Tel: +1-407-823-4811; Fax: +1-407-823-5835; E-mail: xiaoman@mail.ucf.edu

Received March 14, 2015; Accepted April 02, 2015; Published April 04, 2015

Citation: Dhillon V, Li X (2015) Single-Cell Genome Sequencing for Viral-Host Interactions. J Comput Sci Syst Biol 8: 160-165. doi:10.4172/jcsb.1000183

Copyright: © 2015 Dhillon V, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

and alternative transcript processing in viral hosts is still in its infancy [17,24-26] (Figure 1).

Viral Host Responses

Host response-screening of a single-cell

Sequencing single cells was not possible until very recently because most bacteria contain a minuscule amount of genetic material that could not be properly extracted and processed [1,2,6]. The first breakthrough was the development of shotgun sequencing of DNA extracted from environmental samples which broadened the accessibility to sampling environments [6]. The second major development was the introduction of Multiple Displacement Amplification (MDA) which allowed for high-resolution amplification of the extracted genome, followed by eventual reconstruction of the target genome in the process through bioinformatics algorithms [6,12-14,27]. The development of single cell amplification came in the form of phi-29 (Φ 29) DNA-pol based replication of small circular DNA-templates [6,28]. The polymerase-based rolling-circle amplification allowed for hyper-branching of the DNA-strands which further allowed for thousands of copies to be generated. The early efforts had an increasing amount of amplification bias where random regions of the genome became under-represented in the resulting genome [6]. However, with gradually improving protocols, supplementation from computational corrections and cross-verification with the curated databases, better results with reduced bias are being obtained in recent studies [5,28,29].

The assembly of the target genome from the amplified genetic material still remains a pertinent challenge [4,9]. In most of the metagenomic studies performed on the single-cell genome, it has been incredibly difficult to assemble the genome of any individual species except for the most abundant ones [3,29,30]. Any accurately constructed assemblies obtained from the samples have been generated as a consensus genome from multiple fragments [31,32]. The genome-recovery process produces varying levels of success for different species which might result from a higher GC content or restrictive access to the genome because of the presence of DNA-bound proteins [5,16,25].

To better understand virus-host interactions in uncultured viruses and find biologically relevant genomic variants, accurate *de novo* sequence reconstruction of viral-genome is critical [12,13]. With improved genome-recovery and reconstruction techniques in near future, we will also gain the ability to precisely assign viral hosts ignoring the background noise and predict virus-related genomic elements with greater accuracy and efficiency [5,7,11,23,33].

Host-invasion record-viral-host interactions in the transcriptome

Pathogenicity involves multiple pathways and molecular mediators causing long-term changes in the genomics of the host, some of these changes have recently been observed via single-cell studies. Once the genetic material from a sample has been extracted and amplified, one approach for quantifying transcriptional complexity of gene expression in single-cells is through RNA-sequencing resulting in whole-transcriptome analysis [8,34]. The transcriptome extracted from a single-cell reflects the internal composition of the cell, the changes in gene expression in response to stimuli and more importantly a record of pathogenic invasion [8,24,25,35]. Viral markers can thus be identified from the transcripts made accessible from the host-genome amplification [8]. Single-cell transcriptomics has revealed complex patterns of heterogeneity within sub-population of cells by classification using computational techniques such as k-means clustering which have already show multimodal expression in some studies [7,36]. Currently, new viral diagnostic patterns are being investigated by uncovering similar multimodal expression peaks [34,36]. Recent studies have shown new and promising applications of transcriptome analysis in the context of viral predation [8,17,34]. Gene-element integration during the lysogenic conversion, for instance, in eukaryotic hosts can be identified from transcriptome instability patterns in differential-exon expression of the transcripts [8,9].

The transcriptome is only a small fraction of total genomic sequence therefore RNA-seq can be used to collect early insights into the relational phylogenics, early versus late gene transcription or

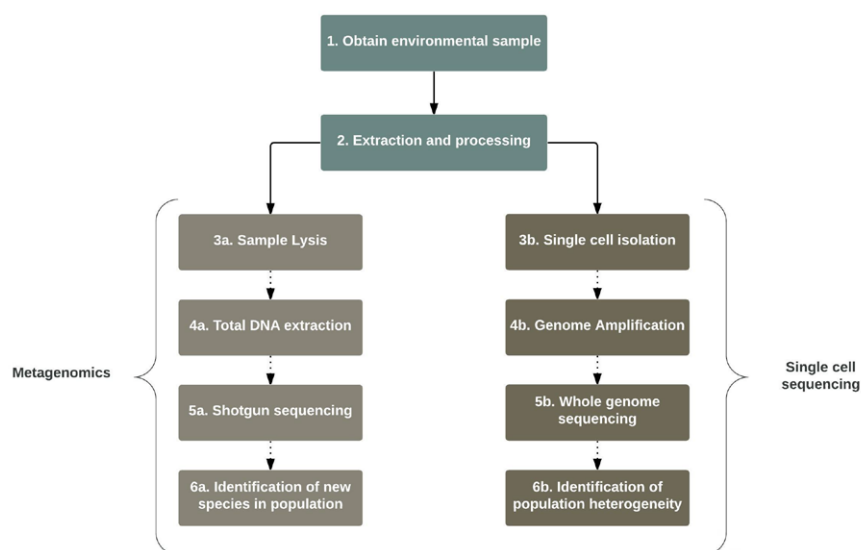


Figure 1: Branching of metagenomics and single-cell genomics. In the metagenomics, a sample is obtained from the environment, then lysed to extract total DNA and sequenced. In single cell sequencing however, a single cell is isolated from the environment and then amplified before its genomic contents are sequenced.

alignment [17,34,37]. There are two approaches to convert the RNA-seq data into transcript sequences: The first one involves the use of previously established model organisms and model-genomes and the second one is via *de novo* assembly [5]. The first approach has become standard for model organisms however it does not perform well for studying non-model organisms and this is especially true for viruses where model systems for viral-host interactions are limited [5,28]. One solution to this problem has recently emerged in the form of Trinity platform [27]. *De novo* assembly of RNA-seq data enables researchers to study transcriptomes without the need for complete genome coverage from the amplified fragments. Trinity platform also allows for *de novo* transcriptome assembly from RNA-seq data for non-model organisms [27,30]. The Trinity platform represents a unique opportunity for bioinformatics tools in single-cell genomics as a full-stack toolset for transcriptome analysis facilitating complementary information to metagenomic studies [4,15]. Assembly and reconstruction of the whole genome is still a very expensive endeavor however advances in nanofluidic and lab-on-chip approaches in near future will make single-cell experiments more affordable and accurate [13,26,30].

Modeling host responses to viral invasion

The bacterial genome stores a memory of the recent pathogenic invasions [21]. The identification of viral markers can be critical in reconstructing that response along with the key host-viral interactions taking place during the process. Single cell techniques can provide access to that level of detailed information from an infected cell and the data obtained can in turn be used to create simulations or models for the host response. Advanced simulation methods such as whole cell simulation [29] are very exhaustive in terms of computational resources and still only produce results in a limited context. However, next-generation sequencing technologies focusing on single-cell techniques along with the support from bioinformatics algorithms have greatly reduced the cost and increased the effectiveness of single-cell sequencing, priming it to tackle the problems of pathogenic-response reconstruction [3, 6, 16]. Algorithmic developments have also extended the previous uniform-genome coverage assumptions to be inclusive for non-uniform coverage from assembly and also to account for chimeric DNA segments accumulated during the MDA reaction [3,33].

A versatile algorithmic tool known as SPAdes, uses k-mers for creating the de-Bruijn graph upon which it performs graph-oriented operations based on factors such as sequence size and graph-shape [4]. With an increased coverage, SPAdes has been utilized in recent studies

to analyze viral-reconstruction of host-phenotypic response. For instance, Sen et al. [46] used SPAdes for clustering and construction of basal phenotype maps to understand how the Varicella-zoster virus alters tonsil T-cells. Furthermore, SPAdes is compatible with several of the commercially available sequencing platforms such as Ion Torrent and Illumina making it easy to integrate into existing pipelines [37,38, 47]. In broader application, single-cell cytometry in combination with bioinformatics tools like SPAdes can be used to provide new hypotheses on the reprogramming of host cells by intracellular parasites such as viruses [2,4]. In addition, the data can also provide insight into the local alteration in the genome for the differentiated cells as well as how those alterations differ from the changes happening in the microbial community to support pathogenesis [2,4]. The molecular mediators involved in such processes will be part of a complex interaction network and researchers have started to curate those networks as interactome databases [10,11].

Karr et al. [29] proposed a comprehensive computational model to understand how molecular interactions result in complex phenotypic responses for a cell. The whole-cell computational model is based on fundamental principles of signal transduction and curated components that integrate numerous biological pathways and cellular processes. Moreover, the simulation contains programmable modules which can be used to gain granular control over cell physiology in a manner that is not possible in a physical cell. The compartmentalization of the whole-cell model is being used to gain new insights on protein-DNA associations and biological discovery of new cellular mechanisms.

Through the whole cell simulation, Karr et al. hypothesized the emergence of a cell-cycle control mechanism as a result of the synchronous activity of the independent simulation modules that regulate various cellular behaviors. Similarly, the metabolome, the transcriptome, the genome and the proteome can lead to the emergence of interactions necessary for viral integration into the host as a “super” module. In near future, this would allow researchers to explore the capabilities of the modules in new directions to facilitate the understanding of a pathogenic episode.

Interactions at large (Viral-host Interactome)

The application of single cell genomics to viruses has only recently started to gain prominence but considerable progress is being made towards one of the broad goals of SCG: Compilation of vast protein-interaction networks between viruses and their hosts in the microbial community [34,41]. The rationale behind this goal is to elucidate the

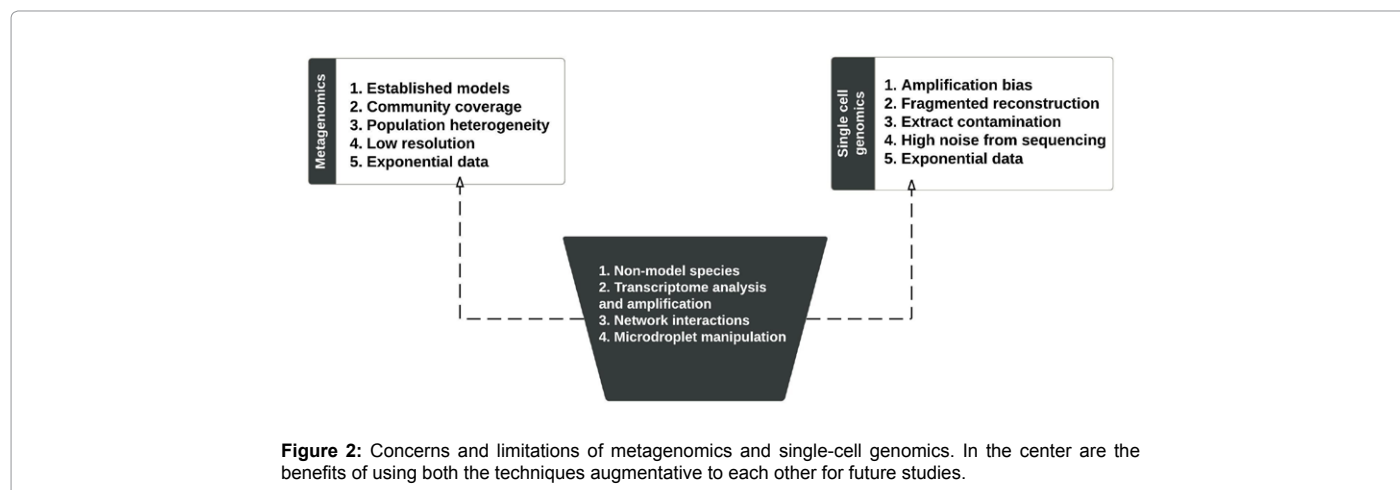


Figure 2: Concerns and limitations of metagenomics and single-cell genomics. In the center are the benefits of using both the techniques augmentative to each other for future studies.

broad-scale interactions that happen to the communities that are sampled in metagenomic studies and discover the function of each component from the sample-community.

A recent study exemplifies this process in the case of applying genome-wide screens to influenza virus to scan for the host factors that are required in viral replication [37]. The study identified 1,000 host factors in the virus-host interactome that are less likely to mutate under selective pressure from drugs and can therefore be used in a hit-to-lead progression for drug-development [39]. Similar studies using single cell techniques have identified biologically-relevant gene variants from interactome analysis and highlighted the clinical relevance of using host-targeting drugs [39,40,42]. The host-targeting drugs offer a reduced risk approach because it takes longer for the host to develop resistance. Unveiling the interactome and extending it to clinically relevant viruses as mentioned is being done in two approaches: A direct approach involving genetic screens and an indirect approach involving single-cell genomics [37,39,40]. The direct approach suffers from the limitation of using well-established model systems for virus-host systems [42]. However, the indirect approach bypasses this limitation and involves extraction of genetic material from individual cells followed by hybridization with the viral genome from the extracted genome. This allows us to select the positively hybridized viruses and

sequence them for assignment to the correct host-pair without ever culturing the virus-host system [28].

This methodology was adopted by Matrinez-Garcia et al. [5,28] in the general sense to identify viruses from the ‘microbial dark matter’ which comprises of the uncultured environmental viruses. Similar indirect approaches have been used in the past to study transmission and mutation of human immunodeficiency virus type I (HIV-1). Salazar-Gonzales et al. [43] used single-cell amplification to obtain intact HIV-1 envelope units from viron-RNA and demonstrate evidence for early mutation and diversification of the viral-envelope. In this case, conventional PCR-based techniques using Taq-polymerase were generating more noise from events such as recombination and nucleotide misincorporation [6]. To reduce the noise, single cell techniques such as amplification methods were applied to generate a neighbor-joining tree which was used to find an estimate for early mutation. The estimate was derived from the most recent common ancestor (MRCA) between the mutated viral strains which was obtained from the neighbor tree and ranged between 10-31 days [43]. The versatility of single-cell techniques augmenting metagenomic studies has become even more apparent in the case of virus-host systems and as the interactome expands, new indirect techniques will allow for complex networks to be generated to provide clinical insights [42,44].

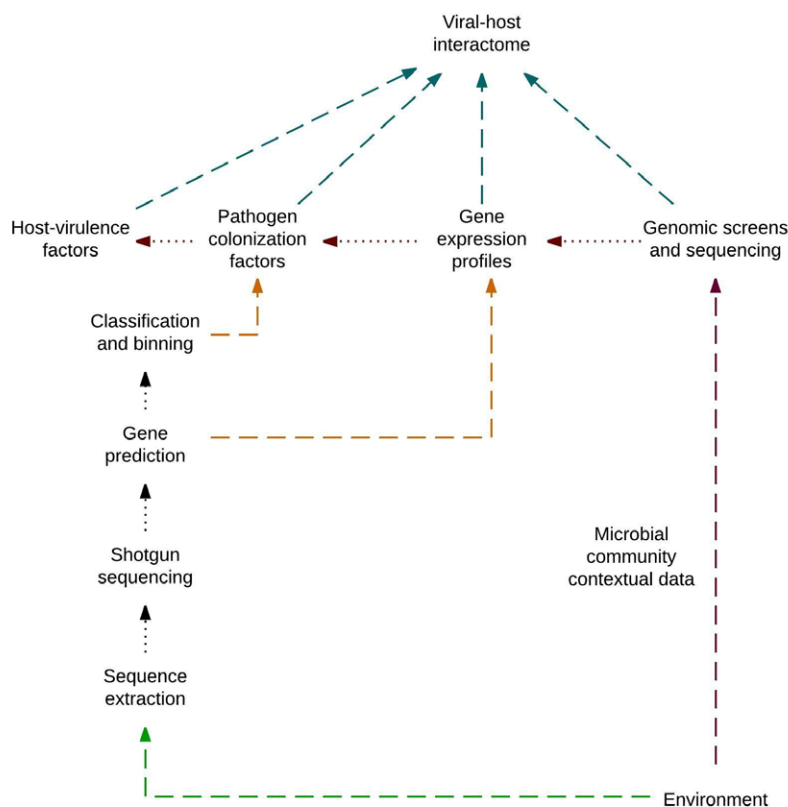


Figure 3: Construction of a viral-host interactome. An environmental sample is processed through various stages until various host-pathogen factors are identified. These factors are integrated to create a full viral-host interaction profile.

Most of the single-cell genomic techniques have been applied to identify patterns in the genetic code; however in near future it will become critical to also analyze the modifiers of the genetic code [15]. In eukaryotic systems, gaining a deep understanding of the modifiers such as histones and the code related to them, the so-called histone code will pave the way for complex system-level insights into the biology of viral-host interactions (Figure 3).

Viral-Host Interactions in Eukaryotic Cells

The dynamics of viral-host interactions in eukaryotic host-cells are collectively more complex and rich. Eukaryotic single-cell amplifications are more complex because they involve a greater number of regulatory elements and a sharp contrast in the genome involving the presence of histones and nucleosomes [45-47]. Histone-tail modifications have well-defined roles in gene and chromatin regulation however their long-term impact on cellular physiology during a pathogenic episode remains largely unclear [48]. Alongside the genetic code, in eukaryotic cells, the recently discovered histone-code is also known to play a major role in gene regulation.

The current single-cell sequencing studies have shown promise in distinguishing complex patterns of variance in the sequenced cells however interpreting the epigenomic code of a cell remains in its infancy [23,49]. The complexity of epigenetic regulations is immense and defining an underlying code to predict their function in parallel to transcriptome analysis will remain of great interest in near future [15]. Interestingly, protocols for detecting methylation status in single-cell analysis are becoming more sensitive and assisting in exploring the 'methylome' landscapes of individual cells during pathological processes such as tumorigenesis [45,48].

Detection of viral-host modifications to the epigenetic code are only beginning to be understood however single-cell techniques can replace traditional methodologies to supplement metagenomic and clinical studies. Furthermore, in near future, existent single-cell techniques would be applied to search for common overexpression changes such as hypermethylation [46,49]. For instance, a recent study by showed that hypermethylation of Somatostatin receptor-1 in the CpG islands assist in the progression of the Epstein Barr virus to gastric cancer [49,50]. Such changes in hypermethylation states are predictive of known pathogenic conditions. In the future, the hypermethylation modifications would become visible through the reconstruction process and also highlight the changes common to a viral community thriving inside a patient [49]. As single-cell techniques evolve to provide epigenetic information, a database of epigenetic code interactome parallel to the protein databases will emerge providing the complete picture of the connections between the transcript and the genetic modifications in the nucleosome [23,45].

Final Remarks

In this review, we discussed the use of single-cell genomic techniques applied to investigate viral-host relationships and the incredible challenges that stand in path of achieving that goal. We reviewed the advancements in next-generation sequencing and the bioinformatics algorithms that have allowed for reduced noise and better coverage. Next, we discussed the role of transcriptome constructed from the RNA-seq data which allows for quick and early diagnostics. Lastly, we described the application of single-cell techniques to creating the interactome what the future holds for deciphering the histone code. Presently, there are several limitations to the practical applications of single-cell techniques in regards to extraction, amplification, and

data-analysis [51-55]. Nanofluidic manipulation and Fluorescence-activated cell sorting (FACS) approaches required for the isolation of individual cells and the subsequent extraction of the genomic contents have had varying amounts of success with different environmental samples. Several groups [13,26,30] are working on improving those results to give reasonable success with most samples obtained. On the other hand, the multiple displacement reaction (MDA) required for amplification of the extracted genome suffers from bias wherein certain regions are repeated more often than others. Finally, the amount of data generated from the whole-cell sequencing of even a single-cell is enormous. One pertinent challenge in the area is to reduce the background noise and focus on the novel expression targets and new bioinformatics algorithms will improve detection hit-rates.

Single-cell genomic techniques have reduced the need for established and cultured model systems to study microbial communities. The interactions among viruses and their hosts have far-reaching consequences on the host genetics, and biochemistry. Single-cell genomics is also being used in drug-discovery pipelines along with advanced metagenomics and bioinformatics enabling fast identification of relevant sequences and "genome islands" [4,39]. The discovery of new drug targets through viral-host interactions follows a hit-to-lead progression. Next-generation sequencing technologies have made understanding genomics much more integrative: The high-throughput sequencing provides massive amounts of data which is analyzed by bioinformatics algorithms to identify potential hits [56-60]. The hits get processed to generate biologically relevant leads which are then understood in the larger context of -omics data such as the interactome for viral-host interactions discussed here. In addition, advances in metagenomics are creating new tools to act as supplements for standard techniques and provide additional information about the complexity emerging from a glimpse of the full landscape.

Acknowledgement

This work is supported by NSF award number: 1218275 and the YES Program-NSF Award Number: 0806931.

References

1. Amann R, Fuchs BM (2008) Single-cell identification in microbial communities by improved fluorescence in situ hybridization techniques. *Nat Rev Microbiol* 6: 339-348.
2. Blainey PC (2013) The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol Rev* 37: 407-427.
3. Macaulay IC, Voet T (2014) Single cell genomics: advances and future perspectives. *PLoS Genet* 10: e1004126.
4. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, et al. (2013) Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499: 431-437.
5. Santos F, Martínez-García M, Parro V, Antón J (2014) Microarray tools to unveil viral-microbe interactions in nature. *Front Ecol Evol* 2: 31.
6. Lasken RS (2012) Genomic sequencing of uncultured microorganisms from single cells. *Nat Rev Microbiol* 10: 631-640.
7. Irish JM, Hovland R, Krutzik PO, Perez OD, Bruserud Ø, et al. (2004) Single cell profiling of potentiated phospho-protein networks in cancer cells. *Cell* 118: 217-228.
8. Ramsköld D, Luo S, Wang YC, Li R, Deng Q, et al. (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol* 30: 777-782.
9. Sachs K, Perez O, Pe'er D, Lauffenburger DA, Nolan GP (2005) Causal protein-signaling networks derived from multiparameter single-cell data. *Science* 308: 523-529.

10. Cai X, Evrony GD, Lehmann HS, Elhosary PC, Mehta BK, et al. (2014) Single-cell, genome-wide sequencing identifies clonal somatic copy-number variation in the human brain. *Cell Rep* 8: 1280-1289.
11. Iourov IY, Vorsanova SG, Yurov YB (2012) Single cell genomics of the brain: focus on neuronal diversity and neuropsychiatric diseases. *Curr Genomics* 13: 477-488.
12. Kind J, Pagie L, Ortobozkoyun H, Boyle S, de Vries SS, et al. (2013) Single-cell dynamics of genome-nuclear lamina interactions. *Cell* 153: 178-192.
13. Lange BM (2005) Single-cell genomics. *Curr Opin Plant Biol* 8: 236-241.
14. Lecault V, White AK, Singhal A, Hansen CL (2012) Microfluidic single cell analysis: from promise to practice. *Curr Opin Chem Biol* 16: 381-390.
15. Mensaert K, Denil S, Trooskens G, Van Criekinge W, Thas O, et al. (2014) Next-generation technologies and data analytical approaches for epigenomics. *Environ Mol Mutagen* 55: 155-170.
16. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, et al. (2011) Tumour evolution inferred by single-cell sequencing. *Nature* 472: 90-94.
17. Tadmor AD, Ottesen EA, Leadbetter JR, Phillips R (2011) Probing individual environmental bacteria for viruses by using microfluidic digital PCR. *Science* 333: 58-62.
18. Allers E, Moraru C, Duhaime MB, Beneze E, Solonenko N, et al. (2013) Single-cell and population level viral infection dynamics revealed by phageFISH, a method to visualize intracellular and free viruses. *Environ Microbiol* 15: 2306-2318.
19. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, et al. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19: 455-477.
20. Friedman N (2013) Studying gene expression at the level of the single cell. *Brief Funct Genomics* 12: 73-74.
21. Kalisky T, Quake SR (2011) Single-cell genomics. *Nat Methods* 8: 311-314.
22. Hutchison CA 3rd, Venter JC (2006) Single-cell genomics. *Nat Biotechnol* 24: 657-658.
23. Hyun BR, McElwee JL, Soloway PD (2015) Single molecule and single cell epigenomics. *Methods* 72: 41-50.
24. Siegl A, Kamke J, Hochmuth T, Piel J, Richter M, et al. (2011) Single-cell genomics reveals the lifestyle of Poribacteria, a candidate phylum symbiotically associated with marine sponges. *ISME J* 5: 61-70.
25. Speicher MR (2013) Single-cell analysis: toward the clinic. *Genome Med* 5: 74.
26. Ståhlberg A, Rusnakova V, Kubista M (2013) The added value of single-cell gene expression profiling. *Brief Funct Genomics* 12: 81-89.
27. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, et al. (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 8: 1494-1512.
28. Martínez-García M, Santos F, Moreno-Paz M, Parro V, Antón J (2014) Unveiling viral-host interactions within the 'microbial dark matter'. *Nat Commun* 5: 4542.
29. Marcy Y, Ouverney C, Bik EM, Lösekann T, Ivanova N, et al. (2007) Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. *Proc Natl Acad Sci USA* 104: 11889-11894.
30. Shapiro E, Biezuner T, Linnarsson S (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet* 14: 618-630.
31. Kalisky T, Blainey P, Quake SR (2011) Genomic analysis at the single-cell level. *Annu Rev Genet* 45: 431-445.
32. Kamke J, Bayer K, Woyke T, Hentschel U (2012) Exploring symbioses by single-cell genomics. *Biol Bull* 223: 30-43.
33. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, et al. (2012) A whole-cell computational model predicts phenotype from genotype. *Cell* 150: 389-401.
34. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, et al. (2009) mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods* 6: 377-382.
35. Heler R, Samai P, Modell JW, Weiner C, Goldberg GW, et al. (2015) Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* 519: 199-202.
36. Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, et al. (2012) Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 150: 1209-1222.
37. Karlas A, Machuy N, Shin Y, Pleissner KP, Artarini A, et al. (2010) Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* 463: 818-822.
38. Sansom SN, Shikama-Dorn N, Zhanybekova S, Nusspaumer G, Macaulay IC, et al. (2014) Population and single-cell genomics reveal the Aire dependency, relief from Polycomb silencing, and distribution of self-antigen expression in thymic epithelia. *Genome Research* 24: 1918-1931.
39. de Chasse B, Meyniel-Schicklin L, Vonderscher J, André P, Lotteau V (2014) Virus-host interactomics: new insights and opportunities for antiviral drug discovery. *Genome Med* 6: 115.
40. Cherry S (2009) What have RNAi screens taught us about viral-host interactions? *Curr Opin Microbiol* 12: 446-452.
41. Weitz JS, Poisot T, Meyer JR, Flores CO, Valverde S, et al. (2013) Phage-bacteria infection networks. *Trends Microbiol* 21: 82-91.
42. Edwards RA, Rohwer F (2005) Viralmetagénomics. *Nat Rev Microbiol* 3: 504-510.
43. Sen N, Mukherjee G, Sen A, Bendall SC, Sung P, et al. (2014) Single-cell mass cytometry analysis of human tonsil T cell remodeling by varicella zoster virus. *Cell Rep* 8: 633-645.
44. Brandenburg B, Zhuang X (2007) Virus trafficking-learning from single-virus tracking. *Nat Rev Microbiol* 5: 197-208.
45. Guo H, Ping Z, Xinglong W, Xianlong L, Lu W, et al. (2013) Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res* 23: 2126-2135.
46. Hagarman JA, Soloway PD (2013) Epigenomics: one molecule at a time. *Cell Cycle* 12: 3451-3452.
47. Hattori N, Niwa T, Kimura K, Helin K, Ushijima T (2013) Visualization of multivalent histone modification in a single cell reveals highly concerted epigenetic changes on differentiation of embryonic stem cells. *Nucleic Acids Res* 4: 7231-7239.
48. Bheda P, Schneider R (2014) Epigenetics reloaded: the single-cell revolution. *Trends Cell Biol* 24: 712-723.
49. Li HP, Leu YW, Chang YS (2005) Epigenetic changes in virus-associated human cancers. *Cell Res* 15: 262-271.
50. Ryan JL, Jones RJ, Kenney SC, Rivenbark AG, Tang W, et al. (2010) Epstein-Barr virus-specific methylation of human genes in gastric cancer cells. *Infectious Agents and Cancer* 5: 27.
51. Allen LZ, Ishoey T, Novotny MA, McLean JS, Lasken RS, et al. (2011) Single virus genomics: a new tool for virus discovery. *PLoS One* 6: e17722.
52. Ding J, Dhillon V, Li X, Hu H (2014) Systematic discovery of cofactor motifs from ChIP-seq data by SIOMICs. *Methods*.
53. Ding J, Hu H, Li X (2014) SIOMICs: a novel approach for systematic identification of motifs in ChIP-seq data. *Nucleic Acids Res* 42: e35.
54. Ding J, Li X, Hu H (2014) MicroRNA modules prefer to bind weak and unconventional target sites. *Bioinformatics*.
55. Salazar-Gonzalez JF, Bailes E, Pham KT, Salazar MG, Guffey MB, et al. (2008) Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J Virol* 82: 3952-3970.
56. Stepanauskas R (2012) Single cell genomics: an individual look at microbes. *Curr Opin Microbiol* 15: 613-620.
57. Walker A, Parkhill J (2008) Single-cell genomics. *Nat Rev Microbiol* 6: 176-177.
58. Wang Q, Zhu X, Feng Y, Xue Z, Fan G, et al. (2013) Single-cell genomics: An overview. *Frontiers in Biology* 8: 569-576.
59. Yilmaz S, Singh AK (2012) Single cell genome sequencing. *Curr Opin Biotechnol* 23: 437-443.
60. Yoon HS, Price DC, Stepanauskas R, Rajah VD, Sieracki ME, et al. (2011) Single-cell genomics reveals organismal interactions in uncultivated marine protists. *Science* 332: 714-717.