Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Virology

Closing the gap: the challenges in converging theoretical, computational, experimental and real-life studies in virus evolution

Editorial overview

Marco Vignuzzi and Raul Andino

Current Opinion in Virology 2012, **2**:515–518

For a complete overview see the Issue

Available online 29th September 2012

http://dx.doi.org/10.1016/j.coviro.2012.09.008

Marco Vignuzzi

Institut Pasteur, 28 rue du Dr Roux, Paris, Cedex 15, 75725, France
e-mail: marco.vignuzzi@pasteur.fr

Dr. Marco Vignuzzi is a virologist within the Department of Virology at the Institut Pasteur in Paris. His laboratory studies the link between replication fidelity, mutation rates and virus fitness *in vivo*.

Raul Andino

Department of Microbiology and Immunology, School of Medicine, University of California, 600 16th Street, San Francisco, CA, USA
e-mail: Raul.Andino@ucsf.edu

Dr. Raul Andino is a virologist at the Department of Microbiology and Immunology of the School of Medicine at the University of California in San Francisco. He has with long-standing interest in RNA virus replication mechanisms, host antiviral functions, evolution and pathogenesis. Dr. Andino has worked on viral genomic plasticity a major determinant of viral pathogenesis. He has demonstrated that interfering with this property of viruses can yield both antiviral drugs as well as safe and effective vaccine strains.

Opening remarks

The recent study of virus evolution has relied mostly on phylogenetic comparison and computational analysis of consensus sequences isolated over time (epidemics) and space (geographic distance). These powerful computational approaches have helped formulate hypotheses on the role of intrinsic and extrinsic factors shaping virus evolutionary trajectories. These initial studies inspired the design of well-controlled experimental evolutionary studies in tissue culture, which were elegant albeit simple in nature. The question has been whether or not these cell culture experiments reflect the complex, real-world, virus evolution processes. Indeed, even a decade ago, characterizing a viral population by isolating individual variants to explore whole genome evolution at numerous time-points and growth conditions was a laborious and ambitious task. But as novel cell culture and analytical technologies and approaches are developed, the possibility to design experimental studies with high temporal resolution and large sequencing and fitness data sets has became a reality. The lag between experimental and computational evolution is beginning to close.

As experimental, computational and theoretical approaches are converging, it will be possible to revisit and revise fundamental concepts in virus evolution. What are the intrinsic and extrinsic mechanisms that determine evolutionary paths and modulate the mutation-selection balance? What complexities exist within organism models that are not reproduced in tissue culture? How do population bottlenecks and fluctuating virus population sizes, altered intracellular environments, host immune systems, distinct virus reservoirs and niches modulate the ability of viruses to adapt and survive? To what extent should the diversity and ecology of the host within its own environment be taken into account, when studying the virus in its host environment? How should experimental studies be designed and quantified to best reflect phylogenetic data and what can be learned from natural and artificial infection models? Is it possible to describe the fitness landscape of an evolving virus, link it to its genetic structure and predict potential evolutionary trajectories?

In this issue, we have selected topics that reflect on each of these questions and demonstrate the significant progress made in recent years.

Optimal mutation rates and robustness

RNA viruses have the highest mutation rates in nature, errors that are generated, and not corrected, by the viral RNA polymerases that lack classic

proofreading mechanisms. Recent analyses helped shape the notion that RNA genomes are limited to relatively short lengths, at least partly due to their extreme error rates [1]: the longer the genome, the more likely detrimental mutations will accumulate to inactivate the virus. Indeed, most RNA virus genomes average 7–12 kb in length. A length that, given the error rates of RNA dependent RNA polymerases, is expected to receive 1–2 mutations per nucleotide site per replication. RNA viruses rely on high mutation rates for efficient adaptation to the dynamic environment of infected individuals. Indeed, higher replication fidelity variants demonstrate reduced fitness in animal models [2–4]. More recently, low fidelity RNA viruses, with a higher mutation load of per replication cycle, have been isolated. These viruses present compromised fitness *in vivo* [5], suggesting that an optimal mutation rate has been selected through evolution to generate sufficient diversity for adaptation but limiting the accumulation of too many 'bad' mutations. These recent works suggest that virus mutation rates are not necessarily fixed or constant and might be modulated depending on the growth conditions, by altering viral polymerase fidelity.

However, the limits to genome length have some notable exceptions among the larger nidoviruses that have puzzled virologists, such as the coronaviruses that reach up to 32 kb. Rather than being a question of polymerase fidelity, the answer may lie in the RNA modifying activities of other virally encoded enzymes. Smith and Denison [http://dx.doi.org/10.1016/j.coviro.2012.07.005] cover recent evidence that the non structural protein nsp14, which carries an exoribonuclease domain, acts as an RNA editing, proofreading enzyme that may 'lighten' the burden of erroneous replication on such a large genome [6,7]. Thus, it is tempting to speculate that in acquiring an accessory proofreading mechanism, a dramatic jump in genome length could be tolerated, and with it, the acquisition of a larger repertoire of virally encoded functions.

But is proofreading and fidelity the only response possible to the burden and risk of extreme mutation frequencies? Theoretically, the specific nucleotide and codon sequences of a genome determine the mutational robustness of that sequence. For example, for amino acids encoded by multiple codons, changing nucleotides at the same position within a codon can result in different amino acid substitutions with different impacts on protein structure and function. Alternatively, regions of complex RNA folding and structure may be differently affected depending on which nucleotide is substituted into a given site. In his review, Santiago Elena [http://dx.doi.org/10.1016/j.coviro.2012.06.008] provides a thorough discussion of this aspect of virus evolution that may benefit from recent technological advances, such as the ability to describe mutation distribution within the virus population by sequencing technologies, more controlled means of manipulating mutation rates and measuring error frequencies. Elena describes how in addition to fine-tuning replication fidelity and proofreading, RNA viruses could potentially evolve more or less robust genomes in response to the mutational burden and demands they encounter while replicating in their hosts [8]. As Elena describes, virus populations could evolve as clusters of genomes that localize within a broad neutral network (that is, where many mutations would be neutral with little or no impact on phenotype and fitness). In this context then, the evolvability of the population depends on what regions of sequence space is covered by the neutral network and whether overlap exists between the current network and a new network that would emerge as the fitness landscape changes with environment.

Challenges in quantifying, defining and describing virus evolution

Although new deep sequencing technologies promise to facilitate the description of mutation distributions within a given virus population, quantifying and assigning their relative fitness and incorporating these values into mathematical and computational formulations will be a formidable task. In her review, Manrubia [http://dx.doi.org/10.1016/j.coviro.2012.06.006] recounts the advances, but also the limitations confronted by genomic and computational biologists, in determining how viral and host parameters are employed in developing mathematical models [9]. As she illustrates, although current, simplified models exclude key features in virus evolution, they are helpful in elaborating new experiments that can validate or refine them. The important goal for future research is to reach a satisfactory and realistic model to describe fitness landscapes, which are surely not as random or as smooth as we tend to draw them out in 2-D illustrations.

Most virology studies employ the word 'fitness' at some point, but the exact definition and implications are often ill-defined and imprecise. In essence, fitness means 'better than' and it is up to the author, although more often the reader, to figure out exactly to which property one is referring. As Wargo and Kurath [http://dx.doi.org/10.1016/j.coviro.2012.07.007] highlight, virus fitness frequently refers to replicative fitness, the ability of one virus to outcompete another in co-infections in tissue culture or *in vivo* [10]. In this context the winners are sprinters, the fastest at replicating the genome and packaging the genetic material into virion. However, the authors make the important point that, as many evolutionary studies are moving into more complex infection models and natural hosts, overall fitness becomes the ability to outperform in numerous tasks: entry, replication, dissemination, transmission, colonization. . .somewhat analogous to the decathlon event of the Olympics, it is no longer just a sprinter's race. Indeed, the fastest 'replicator' might kill its host too quickly and fail at transmission; the best 'disseminator' might alert the immune system too soon,

and so on. In addition to covering a very large body of recent work addressing fitness, the authors raise the current challenges on how to define it, how to properly measure it and how to consolidate evidence from tissue culture and *in vivo* experiments in different hosts.

Viral evolution and complexity of the host

It is important to keep in mind, when speaking of fitness that the values assigned to any genotype are extrinsically dependent on the environment in which the genotype currently exists. Thus, fitness landscapes can only be assigned to very narrow experimental conditions that often do not reflect the dynamic nature of the infected individual. A genotype that could be considered of high fitness in one tissue (fitness peak) may represent a lower fitness genotype in another tissue (plateau or valley in the new fitness landscape). Indeed, a substantial difficulty in measuring and monitoring fitness in the context of a more complex infection model, particularly in whole host organisms, is disentangling the population dynamics that occur when a virus is colonizing new tissues or transmitting between hosts. In these situations, the relative fitness of individual viruses or the overall virus population will be influenced by the relative population size as it fluctuates through any number of population bottlenecks: selection acts intensely on large populations, while small populations are more significantly impacted by stochastic events in genetic drift. As Gutiérrez, Michalakis and Blanc [http://dx.doi.org/10.1016/j.coviro.2012.08.001] report, this requires determining the effective size of the *replicating* population at different times and in different places within the host system [11]. Covering recent, important contributions in the literature of plant, veterinary and medical virology, the authors reveal that quantitative information is still sparse on how a fluctuating multiplicity of infection impacts the overall fitness of a virus population, the selection of the fittest variants and ultimately, viral evolution.

In addition to understanding how virus population size and structure fluctuate in the host, a challenge to studying virus evolution and fitness *in vivo* lies in identifying to what extent fitness increases in the virus come at a cost, or in some cases a benefit, to the fitness of the host. This leads to the question of evolution of virulence and the trade-off hypothesis where a balance is struck between competing to more rapidly colonize and dominate a host environment, and minimizing the potential harm imposed to the host which could reduce the virus' success in transmission. For a virus evolving in permissive cell culture models where resources and infection conditions are kept constant, evolution may favor selection of the fastest (sometimes the most 'virulent') replicator. But in more natural settings, the results of such experimental studies lose relevance. For tripartite relationships in which a virus infects a microbial host, which in turns colonizes a macrobial host, the evolutionary effects of fitness trade-off could be particularly strong. Marquez and Roossinck [http://dx.doi.org/10.1016/j.coviro.2012.06.010] present recent works on viruses infecting unicellular eukaryotes (e.g. Leishmania, Trypanosoma, Trichomonas that infect higher eukaryotes), in which the presence of virus increases survival or spread of the microbial parasite [12]. Equally interesting is their own work on thermal tolerance in which the presence of a virus in a fungus infecting tropical panic grass (results that were reproduced in tomato plants) permitted the plant (and thus, both the fungus and virus) to thrive at higher temperatures that would otherwise be lethal [13].

Finally, there is a clear and logical trend in experimental evolution to move into more natural infection models, whether it be *in vitro* or *in vivo*. The reasons are obvious, to characterize virus population dynamics *in vivo* one should take in to consideration the ongoing virus–host arms race, the evolutionary tug-of-war. Sawyer and Elde [http://dx.doi.org/10.1016/j.coviro.2012.07.003] remind us of how these virus evolution studies reveal important information on host evolution as well and interestingly, they also uncover positive spins to virus–host 'unnatural' mismatches [14]. The authors bring forth several points that we should all bear in mind when designing experiments. For example, using clonal hosts and clonal viruses in the laboratory may not always reveal the spectrum of possible host–virus interactions that may exist in nature, particularly for viruses infecting different species or crossing over into a new species. Furthermore, virologists often hunt for cell lines or hosts that are highly permissive for the replication of a certain virus in order to recover high viral titers. It is important to consider that, by ignoring other less permissive hosts, we may be casting aside opportunities to discover new restriction factors that in the situation where the natural host is used, the virus may have already evolved evasion mechanisms. Thus, coupling these studies with other non-host species may not necessarily be a bad thing, and could help identify the specific mechanism involved.

The road ahead

Whereas virus evolution may have been relegated in the past as the curiosity of a few dedicated teams, its postulated theories and confirmed hypotheses have since reached well into the primary research themes of many virologists focused on antivirals and vaccines, epidemiology, (re)-emerging diseases, ecology, virus–host interactions, virulence and pathogenesis. The resources available to us to study virus evolution are unparalleled in history: worldwide sentinel and surveillance networks for field samples, a plethora of natural and transgenic animal models, increasing molecular precision (such as infectious clones and single genome amplification), ultra deep sequencing technology, and bioinformatic tools. However, the challenge in merging the theoretical, computational, experimental and real-world research paths being forged is daunting and will require a solid, collaborative commitment by all of us to

reach beyond our individual comfort zones. The topics covered in this issue, we believe, layout some of the groundwork that is needed to get us closer to the next horizon in virus evolution: the move from the descriptive, through the mechanistic, towards the predictive.

References and recommended reading

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

- • of special interest
- •• of outstanding interest

1. Sanjuan R, Nebot MR, Chirico N, Mansky LM, Belshaw R: **Viral mutation rates**. *Journal of Virology* 2010, **84**:9733-9748.
2. Pfeiffer JK, Kirkegaard K: **Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice**. *PLoS Pathogens* 2005, **1**:e11.
3. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R: **Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population**. *Nature* 2006, **439**:344-348.
4. Coffey LL, Beeharry Y, Borderia AV, Blanc H, Vignuzzi M: **Arbovirus high fidelity variant loses fitness in mosquitoes and mice**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**:16038-16043.
5. Gnadig NF, Beaucourt S, Campagnola G, Borderia AV, Sanz-Ramos M, Gong P, Blanc H, Peersen OB, Vignuzzi M: **Coxsackievirus B3 mutator strains are attenuated in vivo**. *Proceedings of the National Academy of Sciences of the United States of America* 2012, **109**:E2294-E2303.
6. Eckerle LD, Becker MM, Halpin RA, Li K, Venter E, Lu X, Scherbakova S, Graham RL, Baric RS, Stockwell TB *et al.*: **Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing**. *PLoS Pathogens* 2010, **6**:e1000896.
7. Smith EC, Denison MR: **Implications of altered replication fidelity on the evolution and pathogenesis of coronaviruses**. *Current Opinion in Virology* 2012, **2**:519-524.
8. Elena SF: **RNA virus genetic robustness: possible causes and some consequences**. *Current Opinion in Virology* 2012, **2**:525-530.
9. Manrubia SC: **Modelling viral evolution and adaptation: challenges and rewards**. *Current Opinion in Virology* 2012, **2**:531-537.
10. Wargo AR, Kurath G: **Viral fitness: definitions, measurement, and current insights**. *Current Opinion in Virology* 2012, **2**:538-545.
11. Gutierrez S, Michalakis Y, Blanc S: **Virus population bottlenecks during within-host progression and host-to-host transmission**. *Current Opinion in Virology* 2012, **2**:546-555.
12. Marquez LM, Roossinck MJ: **Do persistent RNA viruses fit the trade-off hypothesis of virulence evolution?** *Current Opinion in Virology* 2012, **2**:556-560.
13. Marquez LM, Redman RS, Rodriguez RJ, Roossinck MJ: **A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance**. *Science* 2007, **315**:513-515.
14. Sawyer SL, Elde NC: **A cross-species view on viruses**. *Current Opinion in Virology* 2012, **2**:561-568.