Towards an integrative view of virus phenotypes

John P. DeLong[1] ✉, Maitham A. Al-Sammak[2,3], Zeina T. Al-Ameeli[3,4], David D. Dunigan[3,5], Kyle F. Edwards[6], Jeffry J. Fuhrmann[7,8], Jason P. Gleghorn[8,9], Hanqun Li[8,10], Kona Haramoto[7,10], Amelia O. Harrison[7,10], Marcia F. Marston[11], Ryan M. Moore[10,12], Shawn W. Polson[7,8,10,12], Barbra D. Ferrell[10,12], Miranda E. Salsbery[1], Christopher R. Schvarcz[6], Jasmine Shirazi[9], Grieg F. Steward[6], James L. Van Etten[3,5] and K. Eric Wommack[7,8,10] ✉

Abstract | Understanding how phenotypes emerge from genotypes is a foundational goal in biology. As challenging as this task is when considering cellular life, it is further complicated in the case of viruses. During replication, a virus as a discrete entity (the virion) disappears and manifests itself as a metabolic amalgam between the virus and the host (the virocell). Identifying traits that unambiguously constitute a virus's phenotype is straightforward for the virion, less so for the virocell. Here, we present a framework for categorizing virus phenotypes that encompasses both virion and virocell stages and considers functional and performance traits of viruses in the context of fitness. Such an integrated view of virus phenotype is necessary for comprehensive interpretation of viral genome sequences and will advance our understanding of viral evolution and ecology.

Viruses are ubiquitous, and virus particles (virions) in most environments outnumber the hosts on which they depend for replication[1–3]. Concentrations of free virions can reach up to approximately 10^8 per millilitre in aquatic systems[4–6], 10^9 per gram of soil[7] and 10^{10} per cubic centimetre of sediments[8,9]. These high numbers arise from cycles of infection and lysis of cells that influence food web dynamics, community structure and nutrient cycling[1,4,10–12]. Viral infections can also dramatically alter host physiology as a result of viral gene expression and drive evolutionary innovation through virus-mediated horizontal gene transfer[13–15].

Understanding how viruses shape life requires grappling with their exceptional diversity. All extant cellular life forms derive from a common ancestor and share the ancestral trait of a double-stranded DNA genome and a set of core metabolic genes. Viruses, however, have arisen multiple times[16–18] as molecular symbionts that must replicate inside a cell with some dependence on host resources[1]. With only lifestyle, rather than a common ancestor, uniting viruses as a group, there is no universally shared gene among them. Viruses differ widely in morphological features as well as genome content and nature. The genome may be single-stranded or double-stranded and composed of DNA or RNA, or even some hybrid of these options[16–18]. Regardless of its biochemical composition, the genome of a virus, like that of a cell, encodes the information that translates into the observable traits constituting the phenotype of that virus.

A phenotype is an observable characteristic of an individual. 'Phenotype' is often used synonymously with 'traits', such as morphological structures, behaviours, developmental or physiological processes, molecules or emergent properties of networks (BOX 1). Selection acts on phenotypes because they determine the functional interactions between organisms and their environment and between biological entities such as viruses and their hosts. These interactions ultimately determine the survival and reproduction of individuals, which are the core components of fitness and the drivers of virus and host community dynamics[19].

Viruses may evolve rapidly in response to host defences or environmental conditions[20–22], which influences infection dynamics[23,24]. The genetic basis of this evolution is detectable as shifts in allelic composition, divergence of genes between virus lineages, and the acquisition and exchange of genetic material[15,16]. However, because natural selection acts on phenotypes and not genotypes, our understanding of the mechanisms of viral evolution and population biology depends on clearly defining and quantifying virus phenotypes and linking them to their genetic underpinnings (genotype-to-phenotype map; BOX 1).

✉e-mail: *jpdelong@unl.edu*; *wommack@dbi.udel.edu*

https://doi.org/10.1038/s41579-021-00612-w

Reassortment
When genome segments from different infecting viruses combine to form a new viral genome.

Episome
A length of viral genome occurring within a host cell.

T number
The number of sides of a virion.

Identifying and measuring virus phenotypes is challenging, in part because of the poorly delineated genotype-to-phenotype transition for molecular phenotypes (BOX 1), and in part because virions are microscopic and can be observed only with specialized instruments. Perhaps the biggest obstacle, though, is that many viral traits are expressed during replication in a host cell, at which point the metabolisms of the host and the virus are inexorably entangled[25]. As a result, aside from the basic features of the virion (for example, size and shape) and broad classification based on genome structure and replicative strategy (for example, Baltimore classification scheme[26]), virologists frequently use the genotype as a lens to the phenotype. This reflects the relative ease of genome sequencing compared with the experiments required to quantify viral infection or determine morphology and macromolecular structure. With the accumulation of new viral genome sequences rapidly outpacing our capacity to isolate and characterize the phenotypes of the associated viruses, there is a push to allow taxonomic assignments based entirely on sequence data[3,27]. However, many viral genes have no known function that can be inferred from sequence homology to a known gene, suggesting that challenges remain in identifying virus phenotypes through sequence data alone.

References to virus phenotype frequently focus on the physical properties of the virion[28,29]. Newly produced virions can differ from the parent virion (for example, through genetic reassortment[30]), but virions generally inherit copies of the original genomic information and recapitulate the phenotype of the 'parent', or originally infecting virion. However, when a virion disintegrates to release its genomic material in a host cell, does the virus no longer have a phenotype? Or can a phenotype be identified as something beyond the traits of the discrete particle that initiated the infection? The latter case would contradict the standard view that traits must be ascribable to and/or measured on an individual[31], because the virus is no longer an individual but a chimeric entity often referred to as a 'virocell'[29]. Thus, an expanded and integrative view of virus phenotypes — a view that can accommodate the full diversity of viruses — is needed. In this Review, we present an integrative view of the virus phenotype with the aim of facilitating a more mechanistic understanding of viral evolution, improving efforts at comparative virology by helping to establish a common language about virus phenotypes[32] and enabling the genotype–phenotype mapping essential to understand the molecular basis of virus functions.

Author addresses

[1]School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE, USA.
[2]Tropical Biological Research Unit, College of Science, University of Baghdad, Baghdad, Iraq.
[3]Nebraska Center for Virology, University of Nebraska-Lincoln, Lincoln, NE, USA.
[4]Medical Technical Institutes, Middle Technical University, Baghdad, Iraq.
[5]Department of Plant Pathology, University of Nebraska-Lincoln, Lincoln, NE, USA.
[6]Department of Oceanography, University of Hawai'i at Mānoa, Honolulu, HI, USA.
[7]Department of Plant and Soil Sciences, University of Delaware, Newark, DE, USA.
[8]Department of Biological Sciences, University of Delaware, Newark, DE, USA.
[9]Department of Biomedical Engineering, University of Delaware, Newark, DE, USA.
[10]Delaware Biotechnology Institute, University of Delaware, Newark, DE, USA.
[11]Department of Biology and Marine Biology, Roger Williams University, Bristol, RI, USA.
[12]Center for Bioinformatics and Computational Biology, University of Delaware, Newark, DE, USA.

Virus phenotypes

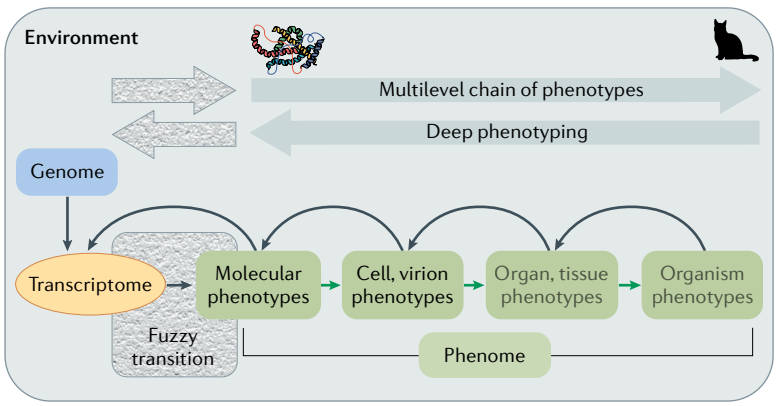
Generally, the virus life cycle alternates between two stages. In the virion stage, a virus exists as an encapsidated genome, usually spending some time outside a cell as a means of dispersal. In the infection stage[29], the virus exists as unencapsidated nucleic acid within a host cell participating in genome transcription, translation, replication, and assembly and packaging[33] (FIG. 1). Virion production may start immediately or after a quiescent period. In the latter case, viral genomic information persists in the cell either as an episome or integrated into the host genome. Although there are exceptions to this generalization (for example, some fungal viruses have no extracellular stage[34]), these two stages are widespread among viruses and are physically and functionally distinct. Furthermore, the types of phenotypic traits that could be ascribed to the virus are mutually exclusive between the stages. As a result, two key categories of virus phenotypes are virion and virocell traits. This division of phenotypes is loosely analogous to the phenotypic division of flowering plants and seeds — seeds are a distinct life stage with phenotypes that differ from those of the plant that produced them, while the adult plants have their own unique set of phenotypic traits, including those involved in producing the seed[29].

Virion phenotype. Identifying the phenotype of a virion is relatively straightforward, as a virion can be recognized as an individual with measurable features (TABLE 1). Some morphological traits of virions include capsid size (for example, mass, volume or diameter), capsid protein arrangement (for example, *T* number), number and type of surface-embedded proteins, presence or absence of an external or internal membrane, presence and type of attachment structures, and capsid shape (for example, helical or non-helical). Some non-morphological phenotypes include tolerance to environmental stressors and whether there is diffusion of virion contents to the surrounding environment, and what those molecules are[35].

Virocell phenotype. By contrast, the physical identity of a virus is lost after it infects a cell, even if a virion capsid is still attached to the host cell membrane or wall. Given that a phenotype is defined with respect to an individual, the loss of the discrete entity that can be identified as the virus poses some problems for defining the phenotype in the infection stage. With disintegration of the virion, the virus no longer has mass, size or morphology. With the virus not being an individual, how can we define the phenotype of the virus itself in the infection stage? We suggest that the concept of phenotype has to be altered to accommodate the infection stage of viruses. Virocells are where replication occurs in viruses, and they are neither just a host cell nor an individual virus. The virocell is a chimeric individual composed of the host and the virus. As an individual, the virocell may have a phenotype of its own, with traits that differ

Box 1 | **Multilevel phenotypes and the fuzzy transition between the genome and the phenome**

One of the main challenges of linking genomes to phenomes is that the phenome is composed of numerous features distributed across hierarchical levels of organization from molecules to whole organisms (see the figure). Starting with the molecular building blocks, phenotypes emerge across levels through interactions between molecules, cells and tissues, leading to the complex phenotypes of metazoans (a phenotype chain; green rectangles in the figure). Measuring phenotypes across levels down to the molecular level is sometimes referred to as 'deep phenotyping'. Predicting high-level phenotypes is complicated by the many transformations involved between components. Identifying and classifying lower-level phenotypes such as molecular building blocks or machinery present another type of challenge. Although some molecules, such as proteins, are clearly phenotypic, other molecules may be involved in transcription and translation or regulation of gene expression and are thus part of the machinery making phenotypes. It may not always be readily apparent which molecules make phenotypes and which are phenotypes themselves, generating a fuzzy transition between the two bins (mottled grey rectangle). The implication is that understanding the genome-to-phenome transition can be hampered by lack of clarity about where the phenotype starts and thus how far down deep phenotyping needs to go (mottled arrow). This problem is particularly relevant for the phenotypes of viruses, as they do not contain the higher-level phenotypes (indicated by grey text) and therefore there is considerable focus on molecular traits. In this Review, we do not consider the higher-level phenotypes associated with metazoans (two far-right green boxes).



from either component individually (host and infecting virion), but it is challenging to call these traits those of the virus, unless the virus is (re)defined as a chimera, which we do not recommend because this obscures the crucial role of the host phenotype in driving the virocell phenotype. Thus, we proceed by discussing the phenotypes of virocells, recognizing that the identity of the host and the identity of the virus have both changed in the infection stage, and do not refer to these traits as virus phenotypes.

Our definition of a virocell deviates from the original in which the virocell is the virus[29]. We argue that for clarity in defining virus phenotypes and thus the nature of viruses, the virocell must be understood as an amalgam rather than the virus itself for two key reasons. First, in the original view, if an infected cell is 'the virus', then its phenotypes must map onto the virion's genotypes, but this is impossible, as much of the virocell phenotype depends on the host genome. Second, in the original virocell concept, identical virions would produce consistently different 'viruses' (both genetically and phenotypically) depending on the type of cell infected. In our description, identical virions may result in different phenotypes if they are infecting different cells (or even the same cell type under different environmental conditions), but at least the 'virus' itself is a genetically consistent entity, because it is not synonymous with the virocell.

At some point after a virus enters a host cell, viral gene expression begins (FIG. 1), and we suggest it is at this time that a core feature of the virocell phenotype emerges. This feature is the set of proteins that are being synthesized on behalf of the viral genome (that is, its proteome). The proteome includes building blocks of future virions and proteins involved in replication, virion assembly or auxiliary metabolic functions[36,37], as well as proteins that are secreted to perform other functions outside the cell or prevent further infection by additional virions[38]. The proteome of the virion is necessarily a subset of the virocell proteome, which includes both structural and non-structural proteins linked to the eventual capsid as well as all the proteins, such as viral genome replication proteins, involved in the virocell stage. These virocell proteins are a temporary physical manifestation of the virus within the host, and the distribution and biochemistry of these proteins thus are a main feature of the virocell phenotype during infection. One advantage of recognizing that the virocell proteome is part of the phenotype of a virocell is that protein information can be directly connected with genomic information from either the virus or the host[39].

Other phenotypic traits. In addition to traits associated with the virion or virocell stages, some aspects of a virus's genome may be phenotypic. Viruses show variation with respect to whether their genetic information is stored in DNA or RNA; whether their genome is single-stranded or double-stranded; if single-stranded and RNA, whether the genome is positive sense or negative sense; and whether the genome is segmented[40]. Such features of the genome are generally not viewed as phenotypic[31], but the way genetic information is organized might have fitness consequences. If the genome structure has fitness consequences, then logically it must be phenotypic because selection acts on phenotypes. Some evidence for genotype structure as phenotypic includes the following: shorter genomes tend to have a reduced degradation rate, which favours genome segmentation[41]; selection to minimize genome size appears to be common among many viruses[42,43]; an evolutionary shift from monopartite to bipartite is likely if the multiplicity of infection during evolution is high enough, because more smaller segments increases stability, but having multiple segments packaged into independent viral capsids requires a higher multiplicity of infection to function[44]; and high G+C content may confer protection from damage caused by ultraviolet light in marine phages[45]. These examples suggest that at least some structural aspects of the genome are part of a virus's phenotype.

Although the taxonomy of viruses curated by the International Committee on Taxonomy of Viruses likely reflects the evolutionary history and the *Bauplan* of viruses in a useful way, there is tremendous additional variation in viruses within these high-level viral taxonomic categories. This finer-scale phenotypic variation reflects functionally important variation between viruses

Segmented
A genome separated into different parts and not physically connected.

Monopartite
Referring to a virus with a non-segmented genome.

Bipartite
Referring to a virus with a genome segmented into two parts.

Multiplicity of infection
The ratio of infecting viruses to hosts.

G+C content
The proportion of all the nucleotides that are guanine or cytosine.

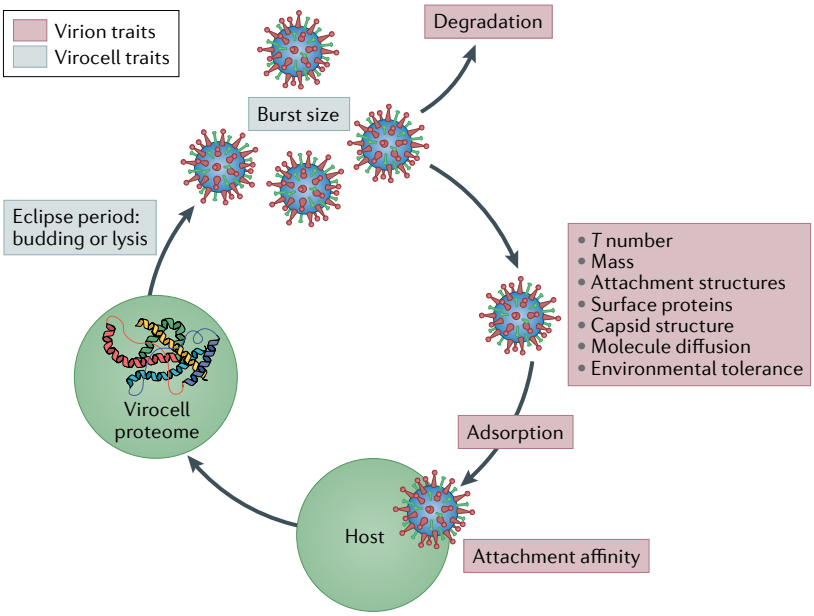


Fig. 1 | **Phenotypes and traits during the virus life cycle.** Virus phenotypes are associated with either the virion phase between infections or the infection and replication phase within a host cell (virocell). Different phenotypes are expressed and detectable at different times of the life cycle. A more detailed description of phenotypes and traits is given in TABLE 1.

that is informative for virus–host dynamics, evolutionary trajectories and the community ecology of viruses, including their impact on ecosystem processes. It is at this scale of phenotypic variation, at which traits differ between genomically defined 'species' of viruses, where research will improve linking genotype to phenotype.

Phenotypes and classification of traits

The connection between genotype and fitness may be seen as a chain of dependencies. Specifically, genotypes give rise to macromolecules that lead to morphological, physiological or behavioural traits that influence an organism's survival, growth and reproduction, which combine to determine an individual's fitness (FIG. 2). Partly borrowing from the plant ecology literature[46], this chain might be described as follows: genome → proteome → functional trait → performance trait → fitness. This organization is useful for categorizing traits on the basis of their position in the chain and maps onto all manner of microscopic to macroscopic organisms, including viruses. We tentatively classify viral traits according to this scheme in TABLE 1.

Performance traits directly determine reproduction and survival, the two core components of fitness. By setting rates of reproduction, changes in state (that is, virion to virocell) and mortality, performance traits include phenotypes often referred to as life history traits. Some virus life history traits are directly analogous to multicellular organism life history traits such as burst size (analogous to the number of seeds), virion decay rate (seed survival rate) and latent period (age at reproduction). Performance traits can often be identified as parameters of epidemiological (or other population dynamic) models. For example, with use of a model of the population growth of a lytic virus, it has been shown[47] that the fitness of a virus is linked to how well the viruses suppress its host cell population. Therefore, virus fitness is proportional to

$$w = \frac{k(be^{-mL} - 1)}{d}, \quad (1)$$

where w is the inverse of host cell density (unit of volume per cell), b is the burst size (unit of virions per cell), k is the adsorption rate (volume per virion per time unit), m is virocell mortality (death rate of the host cell; per time unit), L is the latent period (time) and d is the decay rate of free virions (per time unit). This quotient represents the net virion production from an infected cell ($be^{-mL} - 1$) multiplied by the mean virion lifespan ($1/d$) multiplied by the effective volume over which a virus can encounter and infect a host cell per time unit (k). In other words, Eq. 1 shows how performance traits determine fitness, and in this simple model, virus fitness increases with burst size and adsorption rate, and decreases with the latent period and the death rates of both hosts and virions. Thus, births and deaths in the model are driven by performance traits. This connection between fitness, model parameters and traits is the foundation of trait-based approaches that facilitate an understanding of population dynamics and epidemiology beyond species identity[19], meaning that it is the traits that are the core ecological drivers of virus populations.

Performance traits, in turn, can be written as functions of functional traits[32,48]. Functional traits are phenotypes that are involved in processes that contribute to performance traits. A key functional trait that may influence several viral performance traits is virion size, which is linked to burst size, survival and other aspects of the virus life cycle[47]. These links between virion size and performance may arise because larger virion size is often linked to a larger genome, and larger genomes create a cost (for example, requiring more resources per virion) for virion production and potentially a benefit (for example, augmenting metabolic pathways) within the host cell, altering both the production and the packaging of new virions. As a more specialized example, chloroviruses infect *Chlorella*-like green algae that live as symbionts ('zoochlorellae') within small aquatic organisms such as *Paramecium bursaria*[43]. The ecological challenge chloroviruses face is that their hosts are safely sequestered as a symbiont inside another organism and cannot be infected. How does the chlorovirus infect its host then, and what phenotypes have emerged to facilitate infection? Chloroviruses can attach to the outside of *P. bursaria* cells by the hundreds through fibres that extend from the surface of the virus[49]. The viruses are therefore physically near zoochlorellae, albeit attached to the outside of the paramecium. If a predator disrupts the paramecium, the attached chloroviruses are well positioned for infection of their chlorella host[50]. Described in terms of the genotype to fitness chain, the fibres of chloroviruses are a functional trait that influences a performance trait such as adsorption rate[35].

As another example of how phenotypes and traits are linked to fitness, we consider the determinants of burst

size (FIG. 2). Burst size is a highly variable performance trait of virocells. Burst size should increase if the host contains more resources, if the virions are smaller[19] and if the latency period or lysis timing is longer (more time to make virions)[51]. These traits can combine in a way analogous to the classic offspring size–number life history trade-off, wherein for a given amount of resource available, an organism can produce many small offspring or a few large offspring[52]. In the burst size example, the host produces the resource available to a virus, determined by the host's genome and the environment; consequently, burst size is also determined by the amount of limiting resource needed for building capsids or copies of the genome and dividing that by the per progeny virion resource requirements[53]. Typically, once the genetic cascade towards viral assembly and cell lysis begins, virocells continue making virion parts until the cell runs out of resources or lyses. Thus, the production of virions can continue through time, increasing burst size with lysis timing. Each of these components — virion size, host resources and lysis timing — is a functional trait that is in turn influenced by other functional traits, such

Table 1 | **A non-exhaustive list of virus phenotypes**

Phenotype	Description	Type	Genetic origin	Phenotypic plasticity?
Genome traits				
Nucleic acid type	DNA or RNA	Functional	Virus	No
Strandedness	Double-stranded or single-stranded, positive sense or negative sense	Functional	Virus	No
Genome segmentation	Number of nucleic acid segments	Functional	Virus	No
G+C content	Proportion of nucleotides that are guanine or cytosine	Functional	Virus	No
Genome size	Total length of genome and number of genes	Functional	Virus	No
Virion traits				
Entry mechanism	Membrane penetration, fusion, endocytosis	Functional	Virus and host	Possibly
Virion size	Mass, volume, diameter	Functional	Virus	Possibly
Capsid structure and protein arrangement	For example, *T* number	Functional	Virus	No
Nucleocapsid shape	For example, helical or icosahedral	Functional	Virus	Possibly
Surface proteins	Number and type of surface-embedded protein	Functional	Virus	Possibly
Virion membrane	Presence or absence	Functional	Virus and host	Possibly
Presence and type of attachment structures	For example, phage tail fibres	Functional	Virus	Possibly
Virion molecular contents	Diffusible molecules	Functional	Virus	Yes
Virion proteome	Proteins present in the virion	Functional	Virus	Yes
Virion stability	How long a virion can remain infectious in the environment	Performance	Virus	Yes
Host range	Number (range) of cell types a virion could infect	Performance	Virus and host	Yes
Attachment affinity	Likelihood of attaching to host cell	Performance	Virus and host	Yes
Virocell traits				
Trigger molecules	Mechanism of host cell lysis	Functional	Virus	No
Virocell proteome	Type and number of viral proteins being made in infected cell	Functional	Virus and host	Yes
Eclipse period	Time between initial infection and appearance of first progeny within the host cell	Performance	Virus and host	Yes
Escape period (latent period, lysis timing)	Time between initial infection and release of first progeny outside the host cell	Performance	Virus and host	Yes
Budding rate	Rate of emergence of individual virions from the host cell	Performance	Virus and host	Yes
Burst size	The number of progeny produced from a single infected cell	Performance	Virus and host	Yes

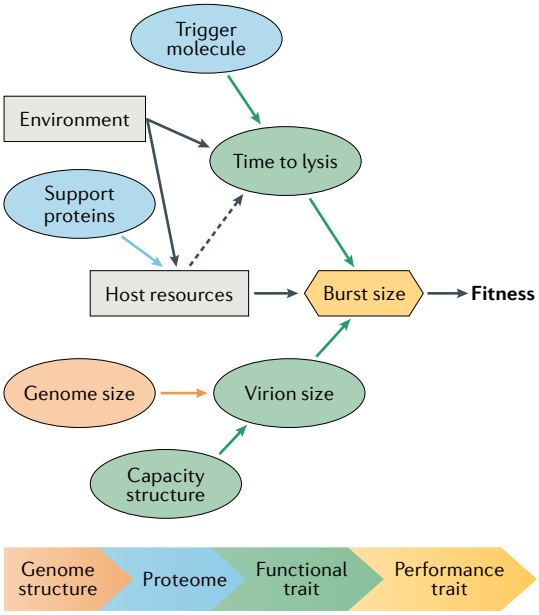


Fig. 2 | **How the performance trait of burst size may arise from the interactions of multiple functional traits and the environment.** In this example, genome size and capsid structure influence virion size, which interacts with host resources to determine burst size. Host resources in turn may be impacted by the environment or by auxiliary metabolic genes within the viral genome that, once expressed, provide support proteins altering some aspect of the host's metabolism. Resources may further influence burst size directly as well as indirectly through their effect on the time to lysis. Finally, each virocell must respond to some type of trigger molecule that initiates escape, which itself is potentially influenced by molecular precursors. The dashed line suggests a possible but unconfirmed connection between host resources and the time to lysis.

as genome length (analogous to genome streamlining and decreased G+C content in bacteria[54]). All else being equal, there can be a straightforward sequence of traits leading to fitness. However, all else is often not equal. Some viruses, such as the cyanophage P-SSP7 infecting *Prochlorococcus*, have genomes that code for proteins that bolster host cell resource acquisition, simultaneously increasing resource levels and genome size, potentially having antagonistic effects on burst size[55]. Thus, functional traits interact with each other and the environment to determine performance traits, which combine to determine fitness.

Phenotypes and their genetic origin

Some traits may be mostly hard-coded by the virus's genome, including functional traits such as capsid morphology (size and structure), which is a direct outcome of protein assembly. Performance traits that are linked to the virus's genome could include virion stability, which influences how long a virion remains viable in the environment. By contrast, some traits may not be linked strictly to the virus's genome. For example, the performance trait of burst size is a function of both the size of the virus's genome and the capacity of the virocell to produce copies of the virus's genome (perhaps reflected in the host growth rate[56]). Therefore, burst size — which can be clearly identified as a performance trait of a virocell — must also be understood to be a trait co-defined by the virus and its host (that is, the two genomes in the virocell). We refer to this type of trait as an 'interaction-mediated trait'. That many virocell traits are interaction mediated is underscored by the fact that host cell proteome responses to different viruses, or responses of different hosts to the same virus, may differ[57,58]. From a genotype-to-phenotype perspective, understanding how interaction-mediated traits arise requires mapping both viral and host genomes to the virocell phenotype.

Host range is an interaction-mediated performance trait with substantial ecological and epidemiological impacts. It is defined as the set of hosts that an individual virion can productively infect. In practice, host range is measured as the set of hosts a population of a particular type of virus can infect, which means that host range generally reflects the set of hosts that at least some members of a particular viral strain can successfully infect. A series of molecular interactions between the virus and the host mediate a successful infection: virion attachment followed by virion (or the viral genome) entry into the host cell and subsequent use of the host's cellular machinery for producing new virus particles. For example, the interactions between host cell receptors and viral attachment molecules (themselves phenotypes) facilitate recognition, attachment and entry into the cell[59]. The outcome of viral infection in a cell is also influenced by active resistance mechanisms in host cells meant to limit productive infection (for example, antiviral cytokines, such as interferons[60]; restriction–modification systems[61]; and CRISPR–Cas immunity[62]) and the evasion or countering of these defences by the virus[63,64]. At very fine scales, such co-evolutionary dynamics can lead to host ranges (visualized as infection matrices[65]) that are nested when one is observing pairwise interactions, with some viruses capable of infecting many host genotypes and other viruses capable of infecting only a few. Furthermore, all viruses infect cells, regardless of whether those cells are organized as tissues within an organism, within the physicochemical strata of microbial mats or as free-living entities. Yet 'host' can also refer to a multicellular organism that contains different cell types that are hosts at the level of infection. Thus, viruses infecting multicellular hosts often have the additional phenotype of infecting a particular cell or tissue type within a multicellular host organism (tropism)[60]. For example, the human immunodeficiency virus infects only particular lineages of white blood cells but can, during the course of infection in an individual, evolve to infect additional cell types[66]. The molecular interactions that influence cell and tissue tropism during the infection of multicellular organisms is analogous to the interactions that determine host range in viruses infecting free-living cells. Thus, the processes and phenotypes that we consider here apply equally to viruses infecting free-living cells and viruses infecting cells within multicellular organisms.

The potential for phenotypic plasticity in viruses

Many traits of multicellular and unicellular organisms vary with abiotic or biotic factors. Examples include a smaller body size at higher temperatures

Interferons
Host proteins that can inhibit virus reproduction.

Restriction–modification systems
A tool for breaking up foreign DNA within host cells.

CRISPR–Cas immunity
Genetic sequences that can be used to identify and destroy foreign genomes.

(temperature–size rule[67]), photoacclimation responses of phytoplankton in which cells adjust their pigment content on the basis of light intensity[68], shifts in cell size and/or shape with the addition or loss of specific endosymbionts[69], and smaller cell size under low-nutrient conditions[70]. Because viruses rely on the host phenotype for expression of their genotype, interaction-mediated viral traits can be influenced by their host's phenotypic plasticity and may vary systematically along an environmental gradient (for example, temperature or resources) as a consequence of host phenotypic plasticity, leading to reaction norms (environment–phenotype relationships) for viruses.

For example, because of the temperature–size rule, we might predict that the burst size of viruses would decrease with increasing temperatures, all else being equal, because smaller host cells have fewer resources available for virus replication (FIG. 3). The full effect of temperature on burst size, however, would emerge from the net effect of temperature on the many steps and processes that lead up to the generation of burst size (FIG. 2). Similarly, phosphorus can limit both cell proliferation in *Micromonas pusilla* and the burst size of *M. pusilla* virus[71], and light levels influence the burst size of cyanophage N-1 infecting the cyanobacterium *Nostoc muscorum*[72]. With T4 phage infecting *Escherichia coli*, several interaction-mediated traits varied with the carbon source for *E. coli* growth, including burst size and lysis timing[73]. Other interactions, such as co-infection with other viruses or interactions with predators or mutualists, could alter the host in such a way as to alter the phenotype of the virocell or the virion.

Abiotic factors can influence host-dependent viral plasticity in other ways. For example, the temperature dependence of interactions between prasinoviruses and their phytoplankton hosts (*Micromonas* sp.) indicates a combination of direct and host plasticity-induced changes in virus phenotype that have opposing effects[74]. In this study[74], increasing temperature reduced both infectivity and virion survival across several host and virus genotypes (plastic performance traits). However, increasing temperature increased host growth rate (host plasticity), thereby indirectly increasing the overall rate of virus production and decreasing the latent period (an interaction-mediated performance trait). In another study, bacteriophage φ29-infected *Bacillus amyloliquefaciens* showed faster rates of DNA synthesis but smaller burst size when grown at 45 °C rather than 37 °C (REF.[75]). Thus, aspects of the virus phenotype emerge as a complex set of responses to the environment and the host phenotype, in addition to the virus's own genotype-to-phenotype map.

Phenotypic plasticity in viruses or their hosts has important implications for viral evolution, persistence and epidemiology. Phenotypic plasticity can alter host–virus interactions because traits mediate those interactions (for example, burst size, encounter rates and lysis timing), with consequences for disease dynamics and the persistence of viruses in the environment[76]. For example, the dependence of viral burst size on light can generate diel cycles in cyanophage production[77,78]. Likewise, host phenotypic plasticity can influence virus transmission, for example, through an increase in host activity and contacts, even if virus phenotypes have not changed[79]. Finally, phenotypic plasticity itself is a trait that can evolve, meaning the expression of a different phenotype in different environments is under selection because it is sometimes advantageous and sometimes not advantageous to express different traits in different environments[80]. Plasticity appears to have evolved

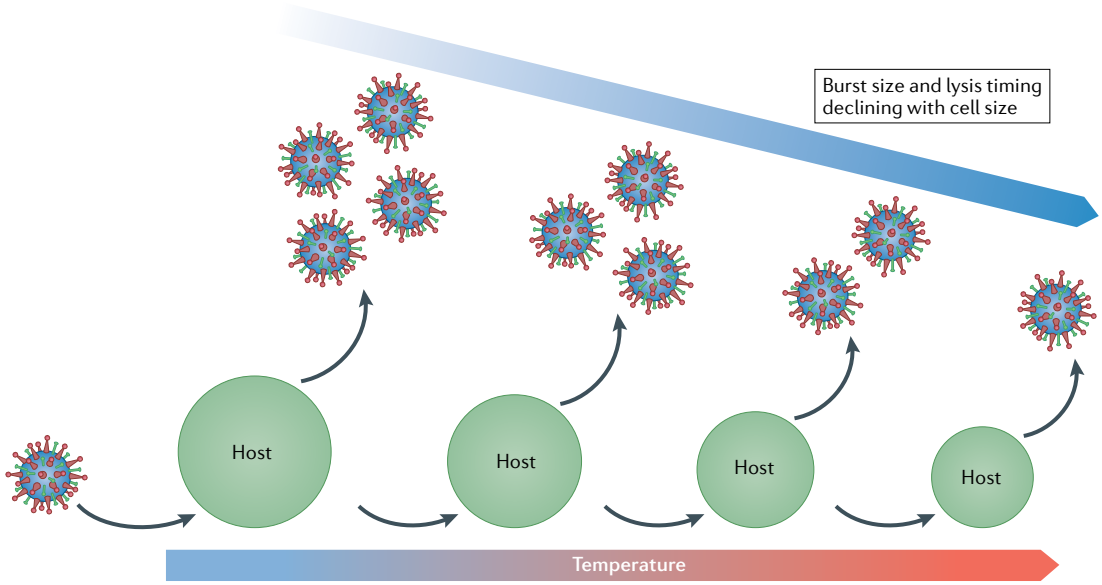


Fig. 3 | **Phenotypic plasticity of a virus due to temperature.** Increasing temperature can lead to smaller host cells[67] and shifts in the total and relative amounts of cellular supplies of key nutrients such as nitrogen and carbon. In addition, increased temperature may increase the rates of biological reactions, which applies to both host-controlled and virus-controlled reactions. Therefore, infection at higher temperatures may lead to shorter times to lysis, smaller burst sizes or shifts in the chemical content of the virion.

Plasmodesmata
Cytoplasmic connections between neighbouring plant cells.

in bacteriophages during co-infection, leading to faster lysis in co-infected hosts than in host cells infected by a single virus[81]. It is therefore clear that host plasticity plays a role in setting phenotypes and their functional consequences.

Although viruses may not generally be thought to have behaviour, behavioural plasticity may arise in viruses when the likelihood of certain viral actions changes with host or environmental factors. For example, the decision by a temperate phage to integrate into a host genome as a prophage rather than initiate a lytic infection is governed by a molecular switch that depends on cell conditions, such as the number of co-infecting viruses and host nutritional state[82]. This decision may be made on the basis of signalling molecules produced by the virocell (that is, part of the virocell proteome)[83]. We may understand this as a form of behavioural phenotypic plasticity because the phenotype itself (lysis timing) arises from a decision process that is influenced by the environment.

Virus reproduction strategies

Viruses may be further categorized by their reproduction strategies within the cell, which describe the path that viruses take from entry of a virus (or a viral genome)

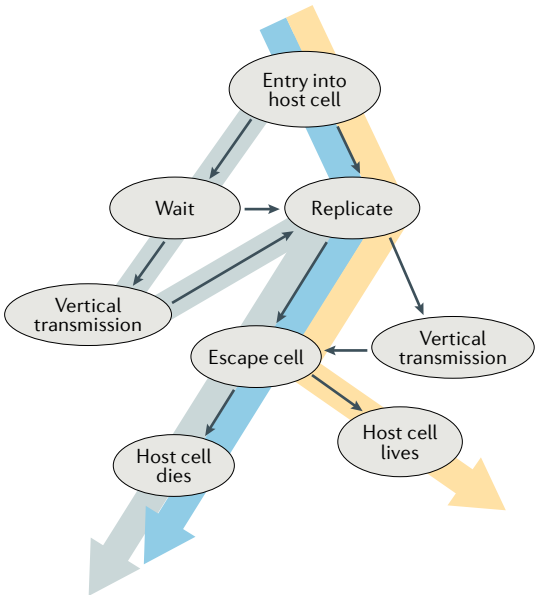


Fig. 4 | **Reproduction strategies and pathways of replication inside host cells.** The reproduction strategies and pathways themselves are not phenotypes, but phenotypes determine both which steps occur and how fast the transitions occur. The pathways shown are not exhaustive of all possible pathways, and some pathways may be flexible depending on the conditions. Three main pathways are shown. The grey pathway reflects the typical action of temperate phages and retroviruses that enter cells, spend some time in a dormant phase, begin to replicate and then kill the host cell upon escape. In the blue pathway, lytic viruses such as chloroviruses and influenza viruses infect host cells, replicate and then escape with concomitant host cell death. By contrast, the yellow pathway describes the route of most plant viruses and chronic viruses, such as hepatitis B virus, with infection, replication and escape happening without their killing the host cell.

to release of virions from the cell (FIG. 4). Reproduction strategies are not phenotypes per se, although they clearly reflect numerous phenotypic traits that govern the replication steps within the cell. Reproduction strategy also dictates what performance traits are relevant for a given virus. For example, chronic viruses have a release rate and lytic viruses have a burst size. As a result, reproduction strategies are reflected in different model structures used to describe the population dynamics of viruses[84].

Several potential reproduction strategies are shown in FIG. 4. The grey pathway reflects the typical strategy of temperate phages and retroviruses, which enter cells, spend some time in a dormant phase, begin to replicate and then kill the host cell upon escape. In the blue pathway, lytic viruses such as chloroviruses and influenza viruses infect the host, replicate and then escape with concomitant host cell death. By contrast, the yellow pathway describes the route for most plant viruses and chronic viruses such as hepatitis B virus, with infection, replication and escape happening without killing the host cell. Host cells may survive because some routes of escape, such as by budding or passing through the plasmodesmata for most plant viruses, do not rupture the cell. We expect that most viruses follow one of these paths, with variations in the exact mechanisms governing the transitions from step to step. Some viruses may have flexibility in the path taken, such as temperate phages, which may or may not have a waiting (latent) period depending on the conditions. We also stress that the patterns of infection used to characterize the progress of disease in multicellular organisms operate on longer timescales and often across cell types and organs, whereas the reproduction strategies we describe are narrowly focused on the reproductive steps of a virus from virion entry to exit from a host, or in some cases, a sequence of hosts through vertical transmission.

Phenotypes, selection and evolution

Clearly defining the virus phenotype at different life cycle stages will facilitate mapping from genetic variation to phenotypic variation, from phenotypic variation to fitness differences, and from fitness differences to evolutionary trajectories and the ecological dynamics that depend upon the virus phenotype. However, measuring phenotypic variation and selection on phenotypes in virus populations is challenging owing to the minute size of viruses. Technologies with single-cell resolution, such as flow cytometry and microfluidics (see BOX 2), could enable trait measurements of single virocells, and in combination with single-cell sequencing this could advance genotype-to-phenotype mapping[85]. Some insights into selection on virus phenotypes have been gained from experimental evolution, where the experimenter imposes selection directly or indirectly and the average phenotype of the whole virus population evolves in response. For example, microbial hosts often evolve resistance to lytic viruses in culture, and the virus population may co-evolve to overcome this resistance[86]. Isolation and sequencing of virus and host strains during co-evolution of phage λ and its *Escherichia coli* host identified a series of mutations that switched the virus

Box 2 | **Facing the challenge of measuring virus phenotypes with microfluidic technologies**

Phenotypes are central to our understanding of viruses but remain challenging to observe and measure. Current approaches, such as electron microscopy and one-step growth curves, will remain important for measuring phenotypes but are likely insufficient for a fuller view that matches the scale of genomic information currently being produced. A promising approach to increase the capacity to measure virus phenotypes is to shrink measuring devices to viral length scales using microfluidic technologies. Recently, microfluidic assays have focused on microbial hosts (reviewed in[115]) as they are easier to visualize and measure, but applications to viruses are expanding. For example, microfluidic devices were used to generate a concentration gradient of GFP-expressing baculoviruses across the width of a fluidic channel with host cells attached along the channel floor. As the fluid progressed along the channel length, the 'lawn' of cells within the channel was spatially exposed to different virus concentrations, enabling high-throughput measurement of the effect of multiplicity of infection (infectious units per cell) quantified by the number of GFP-expressing hosts[116,117]. Several such proof-of-concept microfluidic studies have demonstrated measurements of virus phenotypes similar to those obtained with traditional laboratory methods for viral infection studies, including a one-step growth curve in a microfluidic channel using a GFP-expressing pseudorabies virus[118]. Microfluidic devices also have been developed that directly enumerate fluorescently labelled viruses using a dielectrophoretic field that separates and traps virus particles flowing through a channel, allowing epifluorescent imaging[119]. A more recent strategy for virus–host quantification has been the adaptation of droplet microfluidics, wherein tens of thousands of aqueous droplets are created and kept isolated from each other in an oil phase, thus creating individual 'reaction' chambers that are hundreds of microns in size. These microfluidic technologies improve our ability to 'see' viruses and investigate virus–host interaction[120–122].

host-recognition protein from the LamB receptor to the OmpF receptor after the host evolved lower expression of the LamB receptor[87]. In general, such studies have been performed on well-known phage–host model systems, but similar approaches may help establish genotype–phenotype links and trait evolution dynamics in microorganisms with key ecosystem roles, such as marine cyanobacteria[88] or freshwater chlorophytes[89]. One challenge is to determine whether genetic and phenotypic changes that occur in the laboratory are representative of how virus populations evolve in the wild, where host densities are often far lower than those used in laboratory experiments[90].

A focus on quantitative trait variation between viruses will enable testing of theoretical predictions from models of trait evolution as well as predictions from trade-offs that may constrain phenotypic variation. For example, a lytic virus should evolve a shorter latent period in dense host populations, even though shorter latent periods reduce burst size, because the virus population will have a greater overall multiplication rate if host cells are lysed sooner, enabling progeny viruses to infect new cells[91]. Experimental evolution with *E. coli* phage RB69 confirmed this prediction, as higher host cell densities selected for a mutant phage strain with a shorter latent period and lower burst size, which outcompeted the ancestral strain[92]. Importantly, predictions from such models may depend on how ecological dynamics interact with evolutionary trait change (eco-evolutionary dynamics)[93]. Viruses can suppress the density of their host population, and a lower host density affects the relative fitness of different life history strategies. In an eco-evolutionary model of phytoplankton viruses, the optimal latent period is often the one that maximizes burst size, under the constraints set by host cell resources for virus replication, because virus-induced host mortality has made host cells scarce[32]. This prediction is consistent with double-stranded DNA phytoplankton viruses often having burst sizes that appear to maximize recycling of host nucleotides[32,94]. Important directions for future work include understanding the genetic basis of latent period evolution in these diverse viruses, and how evolutionary dynamics depend on interaction-mediated virocell traits such as burst size and latent period.

Effects on host environments

Virus phenotypes are key to understanding virus fitness, population dynamics of viruses and their hosts, and viral evolution. Phenotypes, however, also influence the consequences of viral infection and replication for the fate of host resources as well as the composition of the overall community in which the host cell resides, be that community a microbial one in the ocean or a mammalian gut or the different cell types within a metazoan. Connecting virus phenotypes to host resource utilization, population dynamics and the characteristics of host physiology during infection is therefore central to understanding the impacts of viruses on the larger ecosystems that viruses and hosts inhabit[19]. An example of how phenotypic variation could substantially alter ecosystem impacts of microorganisms is the viral shunt[95–97]. In short, viral lysis releases material from cells into particulate and dissolved organic carbon pools, efficiently shunting carbon (and other elements) away from biomass that can be consumed by higher trophic levels. For example, release of specific elements such as iron[97] and nitrogen[98,99] in highly bioavailable forms can substantially enhance the growth of non-infected microbial populations, and the pools of particulate and dissolved organic carbon also fuel further growth of microorganisms. Thus, any trait influencing lysis timing, how efficiently the virus exhausts (or sustains) host resources and the specific elemental building blocks required to make new virions[100] will influence the type and amount of materials released from the host cell upon lysis and passing through the viral shunt, whereas burst size, escape period and viral stability impact further infection of other cells, altering population densities and community structure[96].

Some viruses, and many cellular organisms, have emergent properties that influence their environment and that may be thought of as 'extended phenotypes'[101–103]. The term is typically used in reference to behaviours of animals, but the concept might be extended to encompass certain context-dependent effects that derive from the molecular properties of a virus. For example, a Tupanvirus that infects *Acanthamoeba*, causes cytotoxic effects in *Tetrahymena*, a ciliate that readily ingests the virus but is not itself a host[104]. If this reduces predation losses for the virus, then the virion phenotype (that is, its molecular composition) could be under selection to be toxic in some cells. The cytotoxicity, and its cascading effects on the food web, is not a virus phenotype per se, but can be viewed as an extended molecular phenotype that enhances survival of the virus.

Not all emergent properties, however, should necessarily be categorized as an extended phenotype.

An example is plaque morphology. A common test for the presence and infectivity of some viruses is the plaque assay, in which a virus suspension is combined with an excess of host cells. The cells are immobilized in a thin layer in a culture dish, but can grow to form a dense 'lawn'. If a cell in the lawn is productively infected by a virus, the viruses released infect neighbouring cells. After successive rounds of viral infection and lysis, a zone of dead cells creates a clearing in the lawn, called a plaque. Depending on the virus, host and growth conditions, plaques may differ in size, shape, colour and texture[105]. Plaque morphology is not a feature of an individual virion or the virocell. It is thus not itself a virus phenotype, but is an emergent property influenced by underlying virion and virocell phenotypes, such as burst size, and infectivity[106]. Plaque morphology can be a useful diagnostic of changes in an underlying phenotype, but, unlike virion toxicity mentioned in the previous example, plaque morphology has no inherent consequence that would influence selection outside the laboratory context. Similarly, diseases (symptoms) in hosts (for example, necrotic lesions in the leaves of plants infected by viruses, and immune responses) are also not virus phenotypes but instead represent phenotypic responses of the host to infection rather than traits of individual virions or virocells.

Genome-to-phenome connections

At this time, we have a poor understanding of how the genomic features of a virus give rise to its phenotypic characteristics, leaving a large fraction of the ever-expanding pool of viral sequence data with no functional interpretation. Although we are beginning to see efforts to predict the reproduction strategies of phages based on sequence information[78,107,108], prediction of virus performance traits such as host range, infectivity or virion decay is not possible on the basis of genome data alone. However, recent work examining hypothesized links between the biochemical characteristics of viral genome replication proteins and the predicted infection dynamics of unknown viruses illustrates the possibilities for discerning virocell phenotypes from genotype information. For example, the performance trait of lysis timing may be predictable from a group of interacting functional traits[109] linked to genes coding for family A DNA polymerases (PolA), helicases and ribonucleotide reductase (RNR) in phages. In particular, viruses having a more efficient PolA and other associated replication proteins could enable more rapid assembly of virions, which in turn could influence burst size and/or latent period[78,110,111].

Another example of genomic clues to virus phenotypes that have implications for virus fitness comes from cellular metabolic genes carried in viral genomes that, at first glance, do not appear to directly link with a virus functional trait[107,112,113]. Such genes have been termed 'auxiliary metabolic genes', and in our model (FIG. 2) they can be viewed as giving rise to phenotypes (support proteins) that aid in the acquisition of host resources. These support proteins likely influence virus fitness under resource-limited environmental conditions. One of the longer-known examples is the occurrence of photosynthesis genes within cyanophages[37], where the number of photosynthesis genes carried by a cyanophage aligned with its degree of host specificity[15]. Cyanophages containing both *psb*A and *psb*D photosynthetic genes had broader host ranges than cyanophages having only *psb*A. Because host range is an interaction-mediated performance trait, cyanophages carrying both photosynthetic genes rely less on allosteric interactions between viral and host proteins to sustain photosynthesis during infection than those carrying only one gene, leading to an increase in fitness on a broad range of hosts[15,114]. Other cellular metabolic genes involved in energy pathways, such as the gene encoding dissimilatory sulfite reductase[36], have also been observed in unknown viruses, indicating that viruses sustain other key metabolic pathways during the virocell stage, presumably with other impacts on virus performance traits; however, outside the better studied photosynthetic genes in cyanophages, the impact of most auxiliary metabolic genes or infection support proteins on infection dynamic phenotypes has yet to be determined.

Conclusion

We propose that renewed emphasis on virus phenotypes is essential for progress in virology in all its forms. To this end, we have presented a framework that provides a common language to describe the nature and different categories of virus phenotypes and illustrates how they form a chain of interactions linking genotype to fitness. By describing how the key performance traits that determine virus fitness (burst size, adsorption rate and so on) are affected by many lower-level traits, which themselves are affected by many genes in the viral and host genomes, we hope the path leading from viral genome to phenome, although complicated, is made conceptually clear. The ultimate goal of better predicting the phenotypic characteristics of an unknown virus from its genome will remain a challenge, however, until many more viruses across the broad spectrum of diversity have been comprehensively characterized. The pace at which genomic information is being produced continues to accelerate. Commensurate efforts to generate phenotypic information, and develop new technologies to do so, are urgently needed.

These efforts will dramatically improve our understanding of the ecological consequences of viruses in their environments, from ecosystems to gut microbiotas, and advance our understanding of viral evolution and ecology.

Published online 14 September 2021

1. Suttle, C. A. Marine viruses — major players in the global ecosystem. *Nat. Rev. Microbiol.* **5**, 801–812 (2007).
2. Rohwer, F. & Thurber, R. V. Viruses manipulate the marine environment. *Nature* **459**, 207–212 (2009).
3. Simmonds, P. et al. Virus taxonomy in the age of metagenomics. *Nat. Rev. Microbiol.* **15**, 161–168 (2017).
4. Fuhrman, J. A. Marine viruses and their biogeochemical and ecological effects. *Nature* **399**, 541–548 (1999).
5. Suttle, C. A. Viruses in the sea. *Nature* **437**, 356–361 (2005).
6. Jiang, S., Steward, G., Jellison, R., Chu, W. & Choi, S. Abundance, distribution, and diversity of viruses in alkaline, hypersaline Mono Lake, California. *Microb. Ecol.* **47**, 9–17 (2004).

7. Williamson, K. E., Fuhrmann, J. J., Wommack, K. E. & Radosevich, M. Viruses in soil ecosystems: an unknown quantity within an unexplored territory. *Annu. Rev. Virol.* **4**, 201–219 (2017).
8. Cai, L. et al. Active and diverse viruses persist in the deep sub-seafloor sediments over thousands of years. *ISME J.* **13**, 1857–1864 (2019).
9. Wei, M. & Xu, K. New insights into the virus-to-prokaryote ratio (VPR) in marine sediments. *Front. Microbiol.* **11**, 1102 (2020).
10. Wilhelm, S. W. & Suttle, C. A. Viruses and nutrient cycles in the sea. *BioScience* **49**, 781–788 (1999).
11. Brussaard, C. P. D. et al. Global-scale processes with a nanoscale drive: the role of marine viruses. *ISME J.* **2**, 575–578 (2008).
12. Howard-Varona, C. et al. Phage-specific metabolic reprogramming of virocells. *ISME J.* **14**, 881–895 (2020).
13. Nee, S. & Maynard Smith, J. The evolutionary biology of molecular parasites. *Parasitology* **100**, S5–S18 (1990).
14. Hambly, E. & Suttle, C. A. The viriosphere, diversity, and genetic exchange within phage communities. *Curr. Opin. Microbiol.* **8**, 444–450 (2005).
15. Sullivan, M. B. et al. Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PLoS Biol.* **4**, e234 (2006).
16. Holmes, E. C. What does virus evolution tell us about virus origins? *J. Virol.* **85**, 5247–5251 (2011).
17. Wolf, Y. I. et al. Origins and evolution of the global RNA virome. *mBio* **9**, e02329-18 (2018).
18. Kuhn, J. H. et al. Classify viruses-the gain is worth the pain. *Nature* **566**, 318–320 (2019).
19. Record, N. R., Talmy, D. & Våge, S. Quantifying tradeoffs for marine viruses. *Front. Mar. Sci.* https://doi.org/10.3389/fmars.2016.00251 (2016).
Investigates trade-offs in phenotypes of marine viruses that may influence virus population dynamics and biogeography.
20. Domingo, E. et al. Basic concepts in RNA virus evolution. *FASEB J.* **10**, 859–864 (1996).
21. Solé, R. V., Ferrer, R., González-Garcïa, I., Quer, J. & Domingo, E. Red queen dynamics, competition and critical points in a model of RNA virus quasispecies. *J. Theor. Biol.* **198**, 47–59 (1999).
22. Stern, A. & Sorek, R. The phage-host arms race: shaping the evolution of microbes. *Bioessays* **33**, 43–51 (2011).
23. Daugherty, M. D. & Malik, H. S. Rules of engagement: molecular insights from host-virus arms races. *Annu. Rev. Genet.* **46**, 677–700 (2012).
24. Tegally, H. et al. Sixteen novel lineages of SARS-CoV-2 in South Africa. *Nat. Med.* **27**, 440–446 (2021).
25. Lederberg, J. in *Emerging Viruses* (ed. Morse, S. S.) 3–9 (Oxford University Press, 1993).
26. Baltimore, D. Expression of animal virus genomes. *Microbiol. Mol. Biol. Rev.* **35**, 235–241 (1971).
27. Coutinho, F. H., Edwards, R. A. & Rodrïguez-Valera, F. Charting the diversity of uncultured viruses of archaea and bacteria. *BMC Biol.* **17**, 109 (2019).
28. King, A. M. Q., Adams, M. J., Carstens, E. B. & Lefkowitz, E. J. (eds) *Virus Taxonomy*. 163–173 (Elsevier, 2012).
29. Forterre, P. The virocell concept and environmental microbiology. *ISME J.* **7**, 233–236 (2013).
Among the first reports articulating the viewpoint that infected cells undergoing active virus replication should be recognized as the 'living form' of a virus known as a virocell.
30. Lowen, A. C. Constraints, drivers, and implications of influenza A virus reassortment. *Annu. Rev. Virol.* **4**, 105–121 (2017).
31. Mahner, M. & Kary, M. What exactly are genomes, genotypes and phenotypes? And what about phenomes? *J. Theor. Biol.* **186**, 55–63 (1997).
32. Edwards, K. F. & Steward, G. F. Host traits drive viral life histories across phytoplankton viruses. *Am. Nat.* **191**, 566–581 (2018).
Examines the inter-relationships between virus traits and their consequences for population dynamics and the evolution of burst size.
33. Flint, S. J., Racaniello, V. R., Rall, G. F., Skalka, A. M. & Enquist, L. W. *Principles of Virology* 4th Edn (Wiley, 2015).
34. Ghabrial, S. A., Castõn, J. R., Jiang, D., Nibert, M. L. & Suzuki, N. 50-plus years of fungal viruses. *Virology* **479–480**, 356–368 (2015).
35. Dunigan, D. D. et al. Chloroviruses lure hosts through long-distance chemical signaling. *J. Virol.* **93**, e01688-18 (2019).
36. Anantharaman, K. et al. Sulfur oxidation genes in diverse deep-sea viruses. *Science* **344**, 757–760 (2014).
37. Mann, N. H., Cook, A., Millard, A., Bailey, S. & Clokie, M. Bacterial photosynthesis genes in a virus. *Nature* **424**, 741 (2003).
Shows how the virus genome interacts with the host to facilitate virus reproduction.
38. Mavrich, T. N. & Hatfull, G. F. Evolution of superinfection immunity in cluster A mycobacteriophages. *mBio* **10**, e00971-19 (2019).
39. Marine, R. L., Nasko, D. J., Wray, J., Polson, S. W. & Wommack, K. E. Novel chaperonins are prevalent in the virioplankton and demonstrate links to viral biology and ecology. *ISME J.* **11**, 2479–2491 (2017).
40. ICTV. Virus Taxonomy: The ICTV Report on Virus Classification and Taxon Nomenclature. https://talk.ictvonline.org/ictv-reports/ictv_9th_report/ (2019).
41. Ojosnegros, S. et al. Viral genome segmentation can result from a trade-off between genetic content and particle stability. *PLoS Genet* **7**, e1001344 (2011).
42. Belshaw, R., Pybus, O. G. & Rambaut, A. The evolution of genome compression and genomic novelty in RNA viruses. *Genome Res.* **17**, 1496–1504 (2007).
43. Van Etten, J. L., Agarkova, I. V. & Dunigan, D. D. Chloroviruses. *Viruses* **12**, 20 (2020).
44. Iranzo, J. & Manrubia, S. C. Evolutionary dynamics of genome segmentation in multipartite viruses. *Proc. Biol. Sci.* **279**, 3812–3819 (2012).
45. Kellogg, C. A. & Paul, J. H. Degree of ultraviolet radiation damage and repair capabilities are related to G+C content in marine vibriophages. *Aquat. Microb. Ecol.* **27**, 13–20 (2002).
46. Violle, C. et al. Let the concept of trait be functional! *Oikos* **116**, 882–892 (2007).
47. Edwards, K. F., Steward, G. F. & Schvarcz, C. R. Making sense of virus size and the tradeoffs shaping viral fitness. *Ecol. Lett.* **24**, 363–373 (2021).
48. Bonachela, J. A. & Levin, S. A. Evolutionary comparison between viral lysis rate and latent period. *J. Theor. Biol.* **345**, 32–42 (2014).
49. Yashchenko, V. V., Gavrilova, O. V., Rautian, M. S. & Jakobsen, K. S. Association of *Paramecium bursaria Chlorella* viruses with *Paramecium bursaria* cells: ultrastructural studies. *Eur. J. Protistol.* **48**, 149–159 (2012).
50. DeLong, J. P., Al-Ameeli, Z., Duncan, G., Van Etten, J. L. & Dunigan, D. D. Predators catalyze an increase in chloroviruses by foraging on the symbiotic hosts of zoochlorellae. *Proc. Natl Acad. Sci. USA* **113**, 13780–13784 (2016).
51. Wang, I.-N. Lysis timing and bacteriophage fitness. *Genetics* **172**, 17–26 (2006).
52. Smith, C. & Fretwell, S. The optimal balance between size and number of offspring. *Am. Nat.* **108**, 499–506 (1974).
53. You, L., Suthers, P. F. & Yin, J. Effects of *Escherichia coli* physiology on growth of phage T7 In vivo and in silico. *J. Bacteriol.* **184**, 1888–1894 (2002).
54. Swan, B. K. et al. Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proc. Natl Acad. Sci. USA* **110**, 11463–11468 (2013).
55. Hellweger, F. L. Carrying photosynthesis genes increases ecological fitness of cyanophage in silico. *Environ. Microbiol.* **11**, 1386–1394 (2009).
56. Schenk, H. & Sieber, M. Bacteriophage can promote the emergence of physiologically sub-optimal host phenotypes. *bioRxiv* https://doi.org/10.1101/621524 (2019).
57. Howard-Varona, C. et al. Multiple mechanisms drive phage infection efficiency in nearly identical hosts. *ISME J.* **12**, 1605–1618 (2018).
58. Zimmerman, A. E. et al. Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nat. Rev. Microbiol.* **18**, 21–34 (2020).
59. Grove, J. & Marsh, M. The cell biology of receptor-mediated virus entry. *J. Cell Biol.* **195**, 1071–1082 (2011).
60. McFadden, G., Mohamed, M. R., Rahman, M. M. & Bartee, E. Cytokine determinants of viral tropism. *Nat. Rev. Immunol.* **9**, 645–655 (2009).
61. Bernheim, A. & Sorek, R. The pan-immune system of bacteria: antiviral defence as a community resource. *Nat. Rev. Microbiol.* **18**, 113–119 (2020).
62. Nussenzweig, P. M. & Marraffini, L. A. Molecular mechanisms of CRISPR-Cas immunity in bacteria. *Annu. Rev. Genet.* **54**, 93–120 (2020).
63. Hampton, H. G., Watson, B. N. J. & Fineran, P. C. The arms race between bacteria and their phage foes. *Nature* **577**, 327–336 (2020).
An overview of the mechanisms and phenotypes related to phage infection and host defence mechanisms.
64. Samson, J. E., Magadán, A. H., Sabri, M. & Moineau, S. Revenge of the phages: defeating bacterial defences. *Nat. Rev. Microbiol.* **11**, 675–687 (2013).
65. Flores, C. O., Meyer, J. R., Valverde, S., Farr, L. & Weitz, J. S. Statistical structure of host–phage interactions. *Proc. Natl Acad. Sci. USA* **108**, E288–E297 (2011).
Demonstrates the role of virus host range in generating community-wide patterns of host–phage interactions.
66. Regoes, R. R. & Bonhoeffer, S. The HIV coreceptor switch: a population dynamical perspective. *Trends Microbiol.* **13**, 269–277 (2005).
67. Atkinson, D., Ciotti, B. J. & Montagnes, D. J. Protists decrease in size linearly with temperature: ca. 2.5% C^{-1}. *Proc. R. Soc. Lond. B* **270**, 2605–2611 (2003).
68. Falkowski, P. G. in *Primary Productivity in the Sea* (ed. Falkowski, P. G.) 99–119 (Springer, 1980).
69. Salsbery, M. E. & DeLong, J. P. The benefit of algae endosymbionts in *Paramecium bursaria*is temperature dependent. *Evol. Ecol. Res.* **19**, 669–678 (2018).
70. Kimmance, S. A., Atkinson, D. & Montagnes, D. J. S. Do temperature–food interactions matter? Responses of production and its components in the model heterotrophic flagellate *Oxyrrhis marina*. *Aquat. Microb. Ecol.* **42**, 63–73 (2006).
71. Maat, D. S., van Bleijswijk, J. D. L., Witte, H. J. & Brussaard, C. P. D. Virus production in phosphorus-limited *Micromonas pusilla* stimulated by a supply of naturally low concentrations of different phosphorus sources, far into the lytic cycle. *FEMS Microbiol. Ecol.* **92**, fiw136 (2016).
72. Amla, D. V., Rowell, P. & Stewart, W. D. P. Metabolic changes associated with cyanophage N-1 infection of the cyanobacterium *Nostoc muscorum*. *Arch. Microbiol.* **148**, 321–327 (1987).
73. Hadas, H., Einav, M., Fishov, I. & Zaritsky, A. Bacteriophage T4 development depends on the physiology of its host *Escherichia coli*. *Microbiology* **143**, 179–185 (1997).
74. Demory, D. et al. Temperature is a key factor in *Micromonas*–virus interactions. *ISME J.* **11**, 601–612 (2017).
Shows the effect of temperature on the kinetics, phenotypes and life history strategies of prasinoviruses.
75. Schachtele, C. F., Oman, R. W. & Anderson, D. L. Effect of elevated temperature on deoxyribonucleic acid synthesis in bacteriophage φ29-infected *Bacillus amyloliquefaciens*. *J. Virol.* **6**, 430–437 (1970).
76. Choua, M., Heath, M. R., Speirs, D. C. & Bonachela, J. A. The effect of viral plasticity on the persistence of host-virus systems. *J. Theor. Biol.* **498**, 110263 (2020).
77. Ni, T. & Zeng, Q. Diel infection of cyanobacteria by cyanophages. *Front. Mar. Sci.* https://doi.org/10.3389/fmars.2015.00123 (2016).
78. Sakowski, E. G. et al. Ribonucleotide reductases reveal novel viral diversity and predict biological and ecological features of unknown marine viruses. *Proc. Natl Acad. Sci. USA* **111**, 15786–15791 (2014).
Demonstrates that genomic features in the viral replicon (that is, module of genes responsible for viral genome replication) may predict the biogeographical distribution of viruses.
79. Reeson, A. F. et al. Effects of phenotypic plasticity on pathogen transmission in the field in a Lepidoptera-NPV system. *Oecologia* **124**, 373–380 (2000).
80. Stearns, S. C. The evolutionary significance of phenotypic plasticity. *BioScience* **39**, 436–445 (1989).
81. Leggett, H. C., Benmayor, R., Hodgson, D. J. & Buckling, A. Experimental evolution of adaptive phenotypic plasticity in a parasite. *Curr. Biol.* **23**, 139–142 (2013).
82. Oppenheim, A. B., Kobiler, O., Stavans, J., Court, D. L. & Adhya, S. Switches in bacteriophage lambda development. *Annu. Rev. Genet.* **39**, 409–429 (2005).
83. Erez, Z. et al. Communication between viruses guides lysis–lysogeny decisions. *Nature* **541**, 488–493 (2017).
Demonstrates the use of communication peptides that determine lysogeny in temperate phages.
84. Weitz, J. S., Li, G., Gulbudak, H., Cortez, M. H. & Whitaker, R. J. Viral invasion fitness across a continuum from lysis to latency. *Virus Evol.* **5**, vez006 (2019).

85. Labonté, J. M. et al. Single cell genomics indicates horizontal gene transfer and viral infections in a deep subsurface Firmicutes population. *Front. Microbiol.* **6**, 349 (2015).
86. Koskella, B. & Brockhurst, M. A. Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* **38**, 916–931 (2014).
87. Meyer, J. R. et al. Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* **335**, 428–432 (2012).
88. Marston, M. F. et al. Rapid diversification of coevolving marine *Synechococcus* and a virus. *Proc. Natl Acad. Sci. USA* **109**, 4544–4549 (2012).
Demonstrates the rapid co-evolution of virus and host but highlights the challenge of identifying the critical phenotypes mediating the interaction.
89. Frickel, J., Feulner, P. G. D., Karakoc, E. & Becks, L. Population size changes and selection drive patterns of parallel evolution in a host–virus system. *Nat. Commun.* **9**, 1706 (2018).
90. Knowles, B. et al. Temperate infection in a virus–host system previously known for virulent dynamics. *Nat. Commun.* **11**, 4626 (2020).
91. Wang, I.-N., Dykhuizen, D. E. & Slobodkin, L. B. The evolution of phage lysis timing. *Evol. Ecol.* **10**, 545–558 (1996).
92. Abedon, S. T., Hyman, P. & Thomas, C. Experimental examination of bacteriophage latent-period evolution as a response to bacterial availability. *Appl. Environ. Microbiol.* **69**, 7499–7506 (2003).
93. Palkovacs, E. P. & Hendry, A. P. Eco-evolutionary dynamics: intertwining ecological and evolutionary processes in contemporary time. *F1000 Biol. Rep.* **2**, 1 (2010).
94. Brown, C. M., Lawrence, J. E. & Campbell, D. A. Are phytoplankton population density maxima predictable through analysis of host and viral genomic DNA content? *J. Mar. Biol. Assoc. UK* **86**, 491–498 (2006).
95. Wommack, K. E. & Colwell, R. R. Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* **64**, 69–114 (2000).
96. Weitz, J. S. et al. A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes. *ISME J.* **9**, 1352–1364 (2015).
97. Poorvin, L., Rinta-Kanto, J. M., Hutchins, D. A. & Wilhelm, S. W. Viral release of iron and its bioavailability to marine plankton. *Limnol. Oceanogr.* **49**, 1734–1741 (2004).
98. Shelford, E. J., Middelboe, M., Møller, E. F. & Suttle, C. A. Virus-driven nitrogen cycling enhances phytoplankton growth. *Aquat. Microb. Ecol.* **66**, 41–46 (2012).
99. Ankrah, N. Y. D. et al. Phage infection of an environmentally relevant marine bacterium alters host metabolism and lysate composition. *ISME J.* **8**, 1089–1100 (2014).
100. Jover, L. F., Effler, T. C., Buchan, A., Wilhelm, S. W. & Weitz, J. S. The elemental composition of virus particles: implications for marine biogeochemical cycles. *Nat. Rev. Microbiol.* **12**, 519–528 (2014).
101. Dawkins, R. *The Extended Phenotype: The Long Reach of the Gene* (Oxford University Press, 1999).
102. Dawkins, R. Extended phenotype–but not too extended. A reply to Laland, Turner and Jablonka. *Biol. Philosophy* **19**, 377–396 (2004).
103. Ogata, H. Habitat alterations by viruses: strategies by Tupanviruses and others. *Microbes Environ.* **33**, 117–119 (2018).
104. Abrahão, J. et al. Tailed giant Tupanvirus possesses the most complete translational apparatus of the known virosphere. *Nat. Commun.* **9**, 749 (2018).
105. Clark, H. F. & Wiktor, T. J. Plasticity of phenotypic characters of rabies-related viroses: spontaneous variation in the plaque morphology, virulence, and temperature-sensitivity characters of serially propagated Lagos bat and Mokola viruses. *J. Infect. Dis.* **130**, 608–618 (1974).
106. Abedon, S. T. & Culler, R. R. Optimizing bacteriophage plaque fecundity. *J. Theor. Biol.* **249**, 582–592 (2007).
107. Luo, E., Eppley, J. M., Romano, A. E., Mende, D. R. & DeLong, E. F. Double-stranded DNA virioplankton dynamics and reproductive strategies in the oligotrophic open ocean water column. *ISME J.* **14**, 1304–1315 (2020).
108. Bidle, K. D. Elucidating marine virus ecology through a unified heartbeat. *Proc. Natl Acad. Sci. USA* **111**, 15606–15607 (2014).
109. Schmidt, H. F., Sakowski, E. G., Williamson, S. J., Polson, S. W. & Wommack, K. E. Shotgun metagenomics indicates novel family A DNA polymerases predominate within marine virioplankton. *ISME J.* **8**, 103–114 (2014).
110. Nasko, D. J. et al. Family A DNA polymerase phylogeny uncovers diversity and replication gene organization in the virioplankton. *Front. Microbiol.* **9**, 3053 (2018).
111. Harrison, A. O., Moore, R. M., Polson, S. W. & Wommack, K. E. Reannotation of the ribonucleotide reductase in a cyanophage reveals life history strategies within the virioplankton. *Front. Microbiol.* **10**, 134 (2019).
112. Breitbart, M. Marine viruses: truth or dare. *Annu. Rev. Mar. Sci.* **4**, 425–448 (2012).
113. Hurwitz, B. L. & U'Ren, J. M. Viral metabolic reprogramming in marine ecosystems. *Curr. Opin. Microbiol.* **31**, 161–168 (2016).
114. Lindell, D., Jaffe, J. D., Johnson, Z. I., Church, G. M. & Chisholm, S. W. Photosynthesis genes in marine viruses yield proteins during host infection. *Nature* **438**, 86–89 (2005).
115. Rusconi, R., Garren, M. & Stocker, R. Microfluidics expanding the frontiers of microbial ecology. *Annu. Rev. Biophys.* **43**, 65–91 (2014).
116. Walker, G. M., Ozers, M. S. & Beebe, D. J. Cell infection within a microfluidic device using virus gradients. *Sens. Actuators B Chem.* **98**, 347–355 (2004).
117. Cimetta, E. et al. Microfluidic-driven viral infection on cell cultures: theoretical and experimental study. *Biomicrofluidics* **6**, 024127 (2012).
118. Xu, N. et al. A microfluidic platform for real-time and in situ monitoring of virus infection process. *Biomicrofluidics* **6**, 034122 (2012).
119. Akin, D., Li, H. & Bashir, R. Real-time virus trapping and fluorescent imaging in microfluidic devices. *Nano Lett.* **4**, 257–259 (2004).
120. Yu, J. Q. et al. Droplet optofluidic imaging for λ-bacteriophage detection via co-culture with host cell Escherichia coli. *Lab. Chip* **14**, 3519–3524 (2014).
121. Mashaghi, S. & van Oijen, A. M. Droplet microfluidics for kinetic studies of viral fusion. *Biomicrofluidics* **10**, 024102 (2016).
122. Fischer, A. E. et al. A high-throughput drop microfluidic system for virus culture and analysis. *J. Virol. Methods* **213**, 111–117 (2015).

Acknowledgements
This Review was inspired by discussions at a workshop supported by the US National Science Foundation under grant no. 1736030, and by the Delaware INBRE programme, with a grant from the US National Institute of General Medical Sciences (P20 GM103446) from the US National Institutes of Health and the State of Delaware.

Author contributions
Z.T.A.-A., M.A.A.-S., D.D.D., J.P.D., K.F.E., B.D.F., J.J.F., J.P.G., A.O.H., K.H., H.L., M.F.M., R.M.M., S.W.P., M.E.S., J.S. and K.E.W. researched data for the article, D.D.D., J.P.D., K.F.E., J.P.G., M.F.M., C.R.S. and K.E.W. contributed substantially to discussion of the content, D.D.D., J.P.D., K.F.E., J.J.F., J.P.G., M.F.M., S.W.P., C.R.S., G.F.S., J.L.V.E. and K.E.W. wrote the manuscript and D.D.D., J.P.D., K.F.E., J.J.F., J.P.G., M.F.M., S.W.P., C.R.S. and K.E.W. reviewed or edited the manuscript before submission.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Microbiology thanks H. Ogata, who co-reviewed with H. Endo, S. Roux and the other, anonymous, reviewer for their contribution to the peer review of this work.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.