

Defective subviral particles modify ecological equilibria and enhance viral coexistence

Adriana Lucía-Sanz¹, Jacobo Aguirre^{2,3}, Aurora Fraile^{4,5}, Fernando García-Arenal^{4,5} and Susanna Manrubia^{3,6,*}

¹ School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, USA

² Centro de Astrobiología (CSIC-INTA), Madrid, Spain

³ Grupo Interdisciplinar de Sistemas Complejos (GISC), Madrid, Spain

⁴ Centro de Biotecnología y Genómica de Plantas (CBGP), UPM-INIA, Madrid, Spain

⁵ Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid, Madrid, Spain

⁶ Systems Biology Department, National Centre for Biotechnology, CSIC, Madrid, Spain

Correspondence*:

Susanna Manrubia

National Centre for Biotechnology (CSIC), c/ Darwin 3, 28049 Madrid, Spain.

smanrubia@cnb.csic.es

2 ABSTRACT

Cooperation is a main driver of biological complexity at all levels. In the viral world, gene sharing among viral genomes, complementation between genomes or interactions within quasispecies are frequently observed. In this contribution, we explore the advantages that flexible associations between fully fledged viruses and subviral entities, such as virus satellites, might yield. We devise a mathematical model to compare different situations of competition between two viruses and to quantify how the association with a satellite qualitatively modifies dynamical equilibria. The relevant parameter is the invasion fitness of each virus or of the virus-satellite tandem, which in the model depends on the transmission rate of viruses and on their effect on host survival. While in a virus-virus competition one of the viruses becomes eventually extinct, an association with a satellite might change the outcome of the competition to favor the less competitive virus (regardless of whether it is the helper virus or not) or to allow for the stable coexistence of the two viruses and the satellite. We hypothesize that the latter scenario, in particular, constitutes a parsimonious evolutionary pathway towards more stable cooperative associations, such as bipartite viral forms.

Keywords: viral dynamics model, cooperation, ecological competition, subviral agents, viral coexistence

1 INTRODUCTION

There are many instances in the Virosphere of associations of viruses with kin or with subviral entities (Elena et al., 2014). The involved agents propagate independently, and engage in transient associations that can be contingent or necessary regarding successful completion of the replication cycle. Virus satellites, for instance, are subviral particles that require the assistance of a specific helper virus for its replication and/or encapsidation (Kassanis, 1962), see Fig. 1. Their association with the helper virus is therefore necessary

for the satellite, but contingent in principle for the virus (though some viruses also require the satellite for encapsidation). Bipartite viruses, in turn, have their genome fragmented into two pieces that encapsidate separately but are essential in the viral cycle, so their mutual cooperation is necessary Sicard et al. (2016); Lucía-Sanz and Manrubia (2017).

Virus-satellite associations are ubiquitous in plants, and less frequent in other hosts. Examples infecting animals are the *Hepatitis δ virus* (HDV) (Makino et al., 1987), the genus *Dependoparvovirus* (Cotmore et al., 2014; Krupovic et al., 2016) that infects vertebrates, and virus-satellite associations that have bees (Ribi  re et al., 2010) and planthoppers (Nakashima et al., 2006) as hosts. There are also two cases of dsRNA satellites associated to the *Totiviridae* family that infect unicellular eukaryotes (Khoshnan and Alderete, 1995; Schmitt and Breining, 2002). In several cases, co-infection of animal hosts with the corresponding satellite results in an attenuation of viral symptoms (Nakashima et al., 2006; King et al., 2012; Yau et al., 2011).

Infections by multiple viruses and a variety of subviral entities are common in wild plants (Roossinck, 2005), opening up a high potential for mutual interaction. It has been observed that unrelated viruses within these mixtures do not seem to compete, but rather to cooperate, a fact that might explain their ubiquity (Leeks et al., 2018). Co-infection of a host with subviral particles usually modifies the pathogenicity of the helper virus (Roossinck, 2005; Garc  a-Arenal and Fraile, 2017). The spectrum of phenotypic modifications elicited in plants through such interactions is actually remarkably broad. Some viral species within the *Tombusviridae* family, for instance, take part in associations that include a variety of satellites coding for the capsid protein (Syller, 2003). *In vitro*, Tombusviruses spontaneously generate defective interfering particles (DIPs) that systematically lessen the symptoms of the infection by interfering with the replication process Havelda et al. (2005). However, co-infections with heterologous viruses, e.g. that of Potyvirus with Tombusvirus species as *maize chlorotic mottle*, often increase the virulence of Tombusvirus (Scheets, 1998), as it also happens in other Potyvirus co-infections with heterologous viruses (Pruss et al., 1997). In Bromoviruses, virus-satellite associations can both attenuate or enhance the pathogenicity of the phenotype (Palukaitis and Roossinck, 1996). *Geminiviridae* is the family with the largest number of virus-satellite associations, often modifying the virulence and host-range of the virus (Rehman and Fauquet, 2009). Overall, while short-lived interactions may just involve an increase or decrease in viral accumulation levels (a common effect in infections where DIPs or satellites have spontaneously emerged), long-term interactions correlate with a variety of changes in transmission (Elena, 2011), host range (McLeish et al., 2019a), or cell tropism. Sustained interactions are in all likelihood a prerequisite to eventually develop a necessary, mutual interdependence between any two contingently interacting elements.

From an ecological perspective, phenotypic changes caused by transient associations with other viral and subviral agents affect viral dynamics, change viral epidemiology (Alcaide et al., 2020), and redefine the overall ecological role of the association (Betancourt et al., 2011; McLeish et al., 2019b; Syller, 2020), as compared to infections caused by a single virus. As such, they have to be subject to strong selection, since they may play a main role in virus survival by varying the cost that infection impinges on the host, changing adaptive strategies or, as we show in this contribution, qualitatively modifying the outcome of competition for hosts with co-circulating viral species. An integration of quantitative epidemiology at various levels, including in-host evolution and diverse adaptive strategies, is essential to better understand molecular associations and their ecological effects (Jeger, 2020).

Previous modelling studies have addressed the effect of co-infection by viruses and subviral particles unable to replicate on their own. Specific models have addressed the evolution of virulence in *Cucumber*

67 *mosaic virus* over two hosts (Betancourt et al., 2013) and have shown that highly virulent phenotypes
68 occur in mixed infections, in agreement with field observations. Dynamics of virus-DIPs have been also
69 formally studied based on results of *in vitro* observations, revealing that dynamics in virus cultures can be
70 oscillatory (Frank, 2000) but also intrinsically unpredictable under very general conditions (Kirkwood and
71 Bangham, 1994); among others, DIPs can cause the extinction of the parental virus (Grande-Pérez et al.,
72 2005). General models exploring properties of virus-satellite associations have often focused on conditions
73 for coexistence of the partners, showing that it can be favored under structured demes (Szathmáry, 1992)
74 but is in general difficult under metapopulation dynamics (Nee, 2000). A general result for infections
75 within a single host states that beneficial co-infection suffices to allow for the stable coexistence of different
76 variants (Leeks et al., 2018). Still, the variety of models, and therefore mechanisms, identified for the stable
77 coexistence of heterologous viral associations is reduced, probably as a result of our limited knowledge of
78 the intricacies of such associations in nature, and to the small number of cases described (García-Arenal
79 and Fraile, 2017). The state-of-the-art is similar in the case of necessary associations, such as in multipartite
80 viruses, where specific mechanisms that may counteract the deleterious effects of independent propagation
81 are still unclear (Lucía-Sanz et al., 2018; Zwart et al., 2021).

82 In this work, we present a simple model of virus-satellite association and evaluate its epidemiological
83 effects when two monopartite viruses and a satellite that uses one of them as helper virus co-circulate in a
84 host population. Our main result is that the virus-satellite association changes the possible stable equilibria
85 and, in particular, promotes the long-term stable coexistence of both viruses and the satellite in a host
86 population when, as most natural associations do, the satellite modulates the viral phenotype. In contrast,
87 stable co-circulation of two monopartite viruses is not possible in the studied scenario, since one of the
88 species unavoidable invades the host population in finite time.

2 MODELS

89 We address the epidemic propagation of viruses in a well-mixed host population through a set of differential
90 equations representing the amount of susceptible hosts, H , of hosts infected by a virus, X , of hosts infected
91 by the helper virus, Y , and of hosts simultaneously infected by the helper virus and the satellite, S (see
92 Fig. 2). Capital letters, therefore, describe the abundance of hosts in each of four states: susceptible H
93 or infected hosts in the infected states X , Y , and S ; where needed, we will use lower-case letters to refer
94 to virus x , helper virus y , and satellite s . The model, therefore, does not explicitly consider free viral
95 populations. In our models, infections are persistent and there is no class of recovered hosts. This is the
96 rule for plant viruses, but not necessarily so for animal viruses. Therefore, and although we generically
97 speak of hosts all through the paper, our scenarios better apply to plant-infecting viruses and, as it has been
98 defined, to any system where infections are persistent. Finally, note that we refer to competition for hosts
99 in all instances, not to within-host competition.

100 We keep the model intendedly simple so as to derive general principles arising from competition for
101 hosts between the association of a helper virus and a satellite and a second monopartite virus. We do so to
102 emphasize that relevant ecological dynamics do depend on the success of a specific strategy in front of
103 alternative others. In order to highlight the main processes involved in such competition, we have made
104 some simplifying assumptions. For instance, we assume superinfection exclusion between the two viruses,
105 such that hosts cannot be simultaneously infected by the two monopartite viruses. This is a convenient
106 simplification to obtain exact results; as a consequence, coexistence in the context of our model means
107 "ecological coexistence", in the sense that the two viral strains and the satellite stably co-circulate in the host
108 population under conditions that will be made explicit later. A second simplification is that of a mean-field

approximation (common in paradigmatic epidemic models (Hethcote, 2000)), where the dynamics are described through differential equations and space is not explicitly modeled, therefore assuming that hosts interact homogeneously through averaged values.

2.1 Competition between two viruses: Model V

The general dynamics of the process are as follows. Healthy hosts appear at a constant rate g and decay or die at a natural rate d . The amount of healthy, susceptible hosts at time t is $H(t)$ and that of hosts infected by either virus is $X(t)$ and $Y(t)$. If no confusion arises, we will obviate the explicit dependence on t . Contacts between susceptible and infected hosts cause infection of susceptible hosts at rates p_x and p_y , respectively, that stand for the transmission rates of either virus. Infection by x or y increases the death rate of the host in amounts d_x and d_y . This scenario corresponds to a simple situation where two fully competent viruses compete for hosts (see Fig. 2A), and is described by the following equations:

$$\frac{dH}{dt} = g - dH - p_x XH - p_y YH \quad (1)$$

$$\frac{dX}{dt} = p_x XH - (d + d_x)X \quad (2)$$

$$\frac{dY}{dt} = p_y YH - (d + d_y)Y. \quad (3)$$

In the absence of viruses ($X = 0, Y = 0$), the abundance of hosts stabilizes at $H^* = g/d$, as obtained by solving $dH/dt = 0$. Therefore, Eq. 1 entails a maximum carrying capacity of hosts in the absence of parasites. From now on, and given its relevance in determining the existence and stability of the various solutions, as it will be shown, the quantity d/g will be called host turnover.

It is a consequence of the symmetry between eqs. (2) and (3) that one of the viruses always displaces the other—with rare marginal exceptions, as we show in the results. Alternatively, for a small enough host replacement, both viruses become extinct: infected hosts die before new susceptible hosts are available to maintain the epidemic.

2.2 Competition between a virus and a helper virus with a satellite: Model S

In the former scenario of competition between two viruses, we are now introducing a possible association with a satellite that can only multiply in presence of its helper virus. Due to the symmetry between the equations describing the dynamics of each virus, Eqs. (2) and (3), the choice of the helper virus does not affect the results, so we selected virus y without loss of generality. Class S , corresponding to hosts infected by the tandem $y - s$, is affected by an increase d_s in mortality.

Due to the dependence of the satellite on its helper virus for replication, s can only be transmitted (to classes H or Y) through contacts with class S , either to healthy hosts, at rate p_{sy} , or to hosts already infected with y , at rate p_s . In agreement with our superinfection exclusion, hosts in class X cannot be infected by classes Y or S . Finally, hosts in class S can also transmit only the helper virus under contacts with class H , this process occurring at rate p_{-y} . A scheme of the interactions of this model can be seen in Fig. 2B, and the set of equations considering all of the previous processes reads:

$$\frac{dH}{dt} = g - dH - p_x XH - p_y YH - p_{sy} SH - p_{-y} SH \quad (4)$$

$$\frac{dX}{dt} = p_x XH - (d + d_x)X \quad (5)$$

$$\frac{dY}{dt} = p_y YH - (d + d_y)Y - p_s YS + p_{-y} SH \quad (6)$$

$$\frac{dS}{dt} = p_s YS - (d + d_s)S + p_{sy} SH. \quad (7)$$

Note that, if we set $S(t = 0) = 0$ as an initial condition, this system is in practice equivalent to Model V. This nonetheless, the formal inclusion of the satellite modifies the study of the fixed points and stability of the system, as we show in the Supplementary Information.

The satellite non-trivially breaks the symmetry between the dynamics of hosts in classes X and Y , and therefore between the two corresponding viruses. As we will show through analytical and computational analyses, the satellite can benefit or impair its helper virus (thus impairing or benefiting virus x), or can promote the stable coexistence of all parts.

3 RESULTS

In order to clarify the effects of the introduction of a satellite in the outcome of competition between two monopartite viruses infecting the same host (model S), we begin by analyzing the stable solutions of the dynamics in the absence of the satellite (model V). All parameters in the models (death and transmission rates and host growth) are positive defined. The results obtained for model V are not conceptually new, but serve to properly compare the dynamics of the more complex model S with a baseline situation. Fixed points of the dynamics are the solutions to $d\vec{R}/dt = \vec{0}$, where $\vec{R} = (H, X, Y)$ (model V) or $\vec{R} = (H, X, Y, S)$ (model S). The values of the variables satisfying the fixed point equations are denoted H^* , X^* , Y^* , and S^* in the following. These values have to be equal to or larger than zero for the solution to have biological meaning: this results into conditions for existence and non-negativity that the model parameters have to fulfill. The study of stability of the solutions has been carried out using standard methodology of dynamical systems, as detailed in the Supplementary Information. Stability of the solutions also yields conditions on formal relationships that the model parameters have to satisfy. In the next two sections we summarize the solutions of models V and S and their stability properties, together with their ecological interpretation.

3.1 Coexistence is unstable in a system of two viruses in competition for a host population

Depending on model parameters, model V has four different, positive and stable solutions: i) none of the virus is able to invade the population of hosts and both get extinct; ii) and iii) one of the two viruses stably coexists with the host and the other virus becomes extinct; iv) the two viruses coexist in the population. These four solutions map on three main extended regions in the space of parameters.

An important quantity arising in the formal analysis of the model is the ratio between the transmission rate of a virus, p_i , and its overall effect on host death rate, $d + d_i$, where $i = x, y$. This ratio is an instance of the *invasion fitness* $F_i = p_i/(d + d_i)$ of each virus, and determines which strategy is uninventable (Lehmann

et al., 2016). The ratio d/g is a measure of the replacement of healthy hosts. Smaller d/g means shorter times are needed for the appearance of susceptible healthy hosts.

Invasion fitness F_i and host turnover d/g are the two main quantities determining the relevant equilibria of the system. A summary of such solutions and their stability is reported in Table 1 (see Supplementary Material for further details). Figure 3 shows a representative example of the localization in parameter space of the four possible solutions for model V. In that particular example, we have fixed $d = 0.05$, $d_y = 0.25$, $p_y = 0.1$, and $p_x = 0.1$ and explored the $\{g, d_x\}$ plane. Changes in parameter values shift the boundaries between different solutions but do not cause qualitative changes. We observe large regions where both viruses become extinct or where either x or y coexist with the host.

3.1.1 Extinction of both viral populations

If host turnover is larger than invasion fitness for both viruses (solution **V.1**), the viral populations becomes extinct. This region corresponds to the blue area in Fig. 3A, and describes a solution where the rate of appearance of new susceptible hosts is slower than the time needed by any of the viruses to invade the population: in absence of sufficient susceptible hosts, any parasite becomes extinct.

3.1.2 Survival of one viral population

This situation holds whenever the invasion fitness of the two viruses is different and at least one of them is higher than host turnover: $F_i > F_j$ and $F_i > d/g$ for at least one $i, j \in \{x, y\}$. The solution is symmetric under the exchange $i \leftrightarrow j$. These situations correspond to solutions **V.2** and **V.3**, represented as green and yellow regions in **Figure 3.A**. The abundances of susceptible and infected hosts depends on all model parameters.

3.1.3 Coexistence of both viruses

This is a marginal solution sitting at the boundary separating the two solutions where one of the viruses invades and the other becomes extinct. Coexistence only occurs when the two viruses have identical invasion fitness, and both are higher than host turnover, $F_y = F_x > d/g$. This condition maps to a single line in parameter space. Further, since the equilibrium values of X^* and Y^* are degenerated, a second condition establishes that none of them be zero (in which case the solution would correspond to cases **V.2** or **V.3**). This means that, at this boundary, the amount of hosts infected by x or y is not unique, as illustrated by the vertical line in Fig. 3B, where the amounts of X and Y appear superimposed. The degeneration of the solution is also reflected in its analytical form **V.4**, since Y^* is given as a function of X^* (see Table 1 and Supplementary Information).

It is not to be expected that any natural system strictly meets the conditions for coexistence here derived. In the example in Fig. 3, any minor inaccuracy in the value of d_x would kick the system out of coexistence and cause the extinction of one of the viruses. In general, arbitrary changes in d_x , d_y , p_x or p_y would cause the system to depart from the (quasi)stable solution. Therefore, in more realistic situations, stochasticity, noise, or heterogeneity of any kind would push the system away from this marginal solution and into one of the extended regions where either one or both viruses become extinct. Therefore, co-circulating different viral species are, in general, not to be found in a host population under the conditions of model V.

3.2 Dynamics with a co-infecting satellite

The introduction of the satellite breaks the symmetry of equations describing the dynamics of the two viruses. In this new situation, it is not straightforward to predict which of the viruses or associations is

going to succeed in invading the host population. A variety of outcomes is possible depending on the new parameters.

We first extend our definition of invasion fitness to the virus-satellite association,

$$F_i = \frac{p_i}{d + d_i}, \quad i \in \{x, y, sy, c\}, \quad (8)$$

where for convenience we define $p_c = p_{sy} + p_{-y}$ (with $d_c = d_{sy} = d_s$). Note that, since the likelihood of independent transmission is larger than joint transmission, $F_c > F_{sy}$.

The system of equations (4-7) has seven different fixed point solutions (for all parameters taking positive values), summarized in Table 2. Four of them coincide with the solutions for model V, namely: extinction of the two viruses (S.1), invasion of virus x (S.2) or virus y (S.3) and (marginal) coexistence of hosts with viruses x and y (S.4). Three new solutions appear where the satellite co-circulates with one or two of the viruses. The first of them (S.5) corresponds to the extinction of virus x and the stable coexistence of the tandem virus y -satellite with the host. In two more solutions (embraced under S.6) the three agents (virus x , virus y and satellite) stably co-circulate in the population of hosts. In the following, we examine in detail solutions S.5 and S.6 to discuss how the introduction of the satellite modifies the different equilibria depending on the phenotypic effects of its association with the helper virus.

3.2.1 Neutral association with a satellite does not modify the outcome of competition

If the satellite behaves as a commensal parasite, co-infection with its helper virus does neither modify the regions of different equilibria nor their stability. Formally, this occurs when $F_c = F_y$. Fig. 4 depicts the regions for the different equilibria and shows that the region corresponding to solution V.3 is now split into two subregions embracing solutions S.3 and S.5: in the former, the satellite becomes extinct; in the latter, it co-circulates with virus y .

Invasion of virus y without the satellite (S.3) holds when $F_y(1 - \gamma) > F_{sy}$, where $\gamma = (g/(d + d_y) - d/p_y)p_s/(d + d_s)$. This condition is shown as a vertical, blue dashed line in Fig. 4. As the invasion fitness F_{sy} of the virus-satellite association increases, it becomes harder for the helper virus to survive on its own until, eventually, solution S.5 becomes stable and substitutes S.3: in S.5, $Y^* \neq 0$ and, simultaneously, $S^* \neq 0$. In general, this means that satellites are persistent in helper virus populations, but the virus still infects a subpopulation of hosts in the absence of the satellite. The size of this region depends on the ability of virus y to infect susceptible hosts alone when jumping from joint virus-satellite infections. Therefore, it progressively narrows if p_{-y} decreases. A successful strategy from the viewpoint of the satellite, therefore, could be to evolve towards associations that make the y -virus more dependent on s for transmission, progressively selecting increases in p_s that could eventually turn p_{-y} negligible. That is, the association would evolve from commensal to mutualistic.

3.2.2 A mutualistic satellite can prevent extinction of the helper virus

Consider the situation in model V where $F_x > F_y$, with $F_x > d/g$ (solution V.2): virus y is at a disadvantage and therefore unable to invade the population of hosts. In that situation, cooperation with a satellite can prevent its extinction if the satellite provides a sufficiently high increase in invasion fitness to its association with y . This occurs when the invasion fitness of the virus-satellite association is slightly larger than F_x . A region of parameters that was previously occupied only by solution S.2 becomes now bistable, such that depending on the initial conditions it can lead to the elimination of y (and the concomitant

disappearance of the satellite), or to the extinction of x in favor of the virus-satellite association (solution **S.5**), see Fig. 5.

At the boundary that separates the region of bistability (where both **S.2** and **S.5** are possible) from solution **S.2** as a unique solution, the values of H^* for each solution coincide and equal $H^* = (d + d_x)/p_x$, as Fig. 5B shows. By using the explicit expression for H^* obtained for **S.5** (see Supplementary Information) one can obtain an implicit analytical expression for the lower bistability boundary.

From an evolutionary viewpoint, satellites that confer their helper virus an increase in invasion fitness may be a low-cost solution to guarantee rapid adaptation to new hosts, since such an adaptive strategy, in principle, does not require adaptive mutations.

3.2.3 A parasitic satellite can promote viral coexistence

Consider now the situation in model V where $F_y > F_x$, with $F_y > d/g$ (solution **V.3**): virus x is at a disadvantage and virus y has invaded the host population. An association of y with a parasitic satellite can be a burden to the helper virus and diminish its relative advantage while, at the same time, it could potentially benefit x , the virus that was initially eliminated in the competition between the two monopartite viruses. It turns out that the association with a satellite under the conditions above does not necessarily entail, in contrast to previous equilibria, the elimination of either x or the y -satellite tandem. Actually, if the invasion fitness of the association virus y -satellite is such that $F_c < F_x < F_y$, then virus x can simultaneously coexist with virus y and the satellite in regions where it became extinct when the satellite was absent. The conditions above correspond to the relevant solution embraced under **S.6**, that we label **S.6.a**, see SI. There is a second degenerate solution that requires $F_c = F_x = F_y$ (solution **S.6.b**) that has not biological relevance, in the same sense as discussed above for other degenerate solutions.

Figure 6 illustrates the parameter space for the case of a parasitic satellite and shows the appearance of an extended region of coexistence. The boundary limiting the coexistence region can be analytically calculated by noting that, at that boundary, the values of X^* for solutions **S.5** and **S.6.a** coincide and are equal to zero. See Supplementary Information for an explicit expression of the latter solution.

The case of coexistence is an interesting and ecologically relevant situation where the association with a parasitic satellite is only partly detrimental for virus y , while it actually benefits both virus x and the diversity of circulating viral species (and subviral entities).

3.3 Increased dependence between helper virus and satellite

We have shown in the previous sections how the asymmetric relationship between virus y and the satellite (the virus can propagate and replicate on its own, while the satellite depends on the helper virus) breaks the symmetry of the simple model V in a non-trivial way. This causes in particular the appearance of regions of bistability and coexistence that depend on the phenotype of the association, as compared to that of virus y alone.

There are two other situations implicit in model S and in the solutions analyzed so far that are worth discussing: the case where the helper virus cannot propagate if the host is not co-infected with the satellite, and the limit case where the "helper virus" needs the "satellite" both for propagation and replication. The former case (section 3.3.1) still distinguishes between the two categories of helper virus and satellite, while in the latter (section 3.3.2) the association has turned necessary: the two partners that can no longer be identified as helper virus and satellite, since their interdependence is mutual and alike. Actually, these

two cases might represent intermediate steps in a possible parsimonious evolutionary pathway towards bipartitism.

3.3.1 A satellite virus coding for genes essential in transmission

Consider a situation where virus y can replicate but not propagate without the concurrence of the satellite. Formally, this situation is implemented by setting $p_y = 0$, meaning that hosts in class Y are not infective. This scenario agrees with various common natural situations (see Fig. 1) where a satellite virus codes for the capsid needed for the encapsidation of the helper virus or is essential for vector transmission. Examples of the latter are the requirement of *groundnut rosette satellite* for the transmission of *groundnut rosette virus*, of RNA4 for transmission of *beet necrotic yellow vein virus* (Palukaitis et al., 2008) or the dependence of umbraviruses on luteoviruses for transmission (Taliensky and Robinson, 2003).

In our formal context, this restriction implies that only hosts in class S can generate virions with virus y able to infect other susceptible hosts. This is the only way in which the Y class is fed, since the term $p_y Y H$, corresponding to infections caused by contacts between Y and H hosts, is now absent. Taking an evolutionary perspective, the association between a subviral particle with multiplicative ability and a satellite that endows it with propagative ability appears beneficial for the particle. From the viewpoint of the satellite, it might be also convenient, in order to ensure survival, to carry a highly-valued genome piece, as a capsid or others involved in transmission, that can be likely used by a variety of replicative particles. From this perspective, evolutionary associations between replicative entities and satellites coding for essential genes could be highly dynamic, successful adaptive strategies.

3.3.2 The helper virus can neither replicate nor propagate without the satellite

The subsequent case corresponds to a situation where the partners become mutually dependent both for multiplication and propagation. The formal representation of this situation corresponds to setting $Y = 0$, since the "helper" virus cannot replicate in the absence of the "satellite", and therefore class Y is empty—in the same sense that a class infected only by satellites is irrelevant for the dynamics, and therefore dismissed in our models from scratch. When the set of Eqs. (4)–(7) is reduced to this case, one can see that they are equivalent to model V under the correspondence $Y \rightarrow S$, $d_y \rightarrow d_s$, $p_{sy} \rightarrow p_y$, and $p_{-y} = 0$, since class Y is absent by definition. Still, there is an important difference due to the fact that virus y and the satellite do not necessarily propagate as a unit, but as two independent particles. As a result, p_{sy} in the equivalent formulation may take low effective values, meaning that the probability of simultaneous infection by two independently propagating particles is always lower than the probability of infection by a single propagating particle. Therefore, this situation better describes a bipartite virus, where the genome is divided into two segments that are independently transmitted, where both fragments are needed to successfully complete the viral cycle, and where there is a cost due to independent propagation.

4 DISCUSSION

We have introduced a simple epidemic model to analyze stable states in a system formed by two fully fledged viruses and an associated satellite. This model contains the minimum set of players with symmetric (between the two viruses) and asymmetric interactions (between a virus and a satellite) of different nature (commensal, parasitic or mutualistic) able to display a range of ecological equilibria when competing for a common host. Note that our results only refer to long-time equilibria (i.e. to stable coexistence in a mathematical sense), and therefore do not apply to transient coexistence. In natural systems, the time needed for one viral species to cause the elimination of a second one (corresponding, for example, to

solutions **V.2** or **V.3**) can be so large that transient co-circulation of various viral species could be the most common situation: this transient coexistence should not be mistaken by solution **V.4**, which strictly refers to stable, long-term coexistence. A detailed analysis of how long it takes to attain ecological equilibrium and a quantification of typical evolutionary times for the emergence of new viral species infecting a given host would be needed, among others, to discriminate between transient and actual coexistence. Such an analysis is out of the scope of the present contribution, but it is important to keep in mind that, due to the long times involved in ecological dynamics, natural systems can be out-of-equilibrium most of the time.

In the light of our results, it is worth discussing the major assumption we have made, that is superinfection exclusion. Our second assumption, the use of a deterministic description of the dynamics, mostly affects the stability of marginal solutions (i.e., those solutions demanding a precise coincidence of certain parameters). The incorporation of stochasticity of any kind (in the form of space, which entails probabilistic contagion, finite populations, heterogeneous contacts, noise or any form of disorder), as it would occur in any natural system, turn solutions **V.4**, **S.4**, and **S.6.b** irrelevant, since they become unstable under any minor modification of most parameter values.

Assuming superinfection exclusion is useful to obtain exact analytical results, and therefore to gain insight into major ecological effects of virus-satellite associations. If this simplification is lifted, further interaction terms between viruses x and y , between x and the satellite, or even among the three entities, and new equations accounting for different states with various combinations of co-infecting entities have to be considered. Both the number of equations involved and their complexity would increase, preventing the achievement of exact expressions for equilibria and their stability, and therefore masking the interpretation of generic results. This nonetheless, and since we are arguing that coexistence of different viruses and subviral particles in the same host population might be a stepping stone towards stable mutually dependent associations (as bipartitism), let us be more explicit about this point and sketch how the relaxation of that assumption could affect the results here obtained. The simplest scenario corresponds to commensal co-infection, in which case the qualitative results here presented remain unchanged: propagation of different entities in the host population could be independently analyzed, and the state of each individual host at a given time would be the product of probabilities of being infected with any of the viruses or sets of partners considered in the model. If co-infection causes changes in phenotype, however, the nature of the different equilibria might change, but such changes would be qualitatively analogous to these explored here.

Take the following example as illustration. Suppose we include in our equations hosts co-infected with viruses x and y , and let us call Z this new class. The dynamical equation for Z would read

$$\frac{dZ}{dt} = p_x XY + p_y YX + p_{-y} XS + p_z HZ - (d + d_z)Z \quad (9)$$

The terms in the right-hand side of the equation indicate, from left to right, the rate of infection of hosts in class Y under a contact with class X (virus x is transmitted), of hosts in class X under a contact with Y (virus y is transmitted), of hosts in class X under a contact with S when only y is transmitted, new infections of susceptible hosts under contacts with Z when both viruses are simultaneously transmitted (this introduces the new rate p_z) and, finally, the modification of the death rate due to the co-infection with both viruses (new rate d_z). Equations for dH/dt , dX/dt and dY/dt would also acquire new terms coming from the new transitions to class Z (equation for dS/dt would not be modified since there are no possible transitions between S and Z , since the latter does not contain the satellite). Note that this enlarged model Z would have all fixed points we have described for model S (when $Z^* = 0$) and some new

ones corresponding to $Z^* \neq 0$. The analysis would proceed along similar lines, and the actual challenge would then be to characterize the new conditions under which coexistence occurs, since these could take much more involved mathematical expressions. Still, the mechanisms underlying coexistence, in particular, would not differ from the ones identified in this study as long as interaction terms take mathematical forms analogous to those considered here, that is, if interactions occur in pairs. This is not a limiting assumption, since all general contagion processes involve only pair contacts, regardless of whether multiple particles of various types are transmitted through the contact. For example, if another class W infected by x , y and s is considered, it can be easily shown that its dynamical equation only involves pair interactions again, even if the three entities can be simultaneously transmitted (for example through contacts between H and W). We therefore conclude that the basic model explored in this work contains essential mechanisms for coexistence resulting from asymmetrical interactions among the involved partners, and its results can be conceptually extrapolated to scenarios with a larger number of players.

This problem of identifying coexistence conditions and the involved partners also admits a complementary viewpoint. Instead of considering an enlarged dynamical system with many interacting viral and subviral entities, and struggle to identify the conditions ensuring coexistence, we could take a more pragmatic approach and ask which is the subset of viruses and subviral agents able to coexist when pooled from a large ensemble of viruses and subviral particles that infect a given host. This subset would be, by definition, an ensemble fulfilling the conditions for coexistence in that host, and it could vary from host to host. From this viewpoint, the result of evolving towards regions of coexistence —should this be a situation under selection— is equivalent to drawing the subset of viral species and subviral entities able to coexist. This latter scenario conceptually agrees with common observations of multiple viral species (and possible subviral particles) persistently co-infecting the same host, especially in plants.

Indeed, some of the equilibrium solutions and the overall scenarios we have found in our models and their extensions find their counterpart in natural situations, and do share some of their difficulties. While co-infection by multiple viral species or strains is common, detection and identification of involved parties and the nature of their interactions presents major problems in etiological and epidemiological studies (Jeger, 2020). In our modelling approach, we can just describe regions of parameter space where coexistence is or is not possible, but only a quantification of interactions and parameters can allow for a more specific identification of partners and their interactions. For example, in our solution **S.6.a**, we show that due to synergy between virus y and the satellite, x can stably infect the host, while it became extinct in the absence of s . This would be the case of the obligatory mutualism between *Pea enation mosaic virus 1* (PEMV1) and *Pea enation mosaic virus 2* (PEMV2), where PEMV1 provides a capsid required for PEMV2 transmission while PEMV2 provides a movement protein required for systemic movement of PEMV within the infected plant (Doumayrou et al., 2016).

Other studies have shown that the level of satellite abundance in virus-satellite associations is highly variable, with up to 100% in some regions for helper virus population in *Rice yellow mottle virus* satRNA (Pinel et al., 2003) to very low, as in *Cucumber mosaic virus* satRNA (Palukaitis and García-Arenal, 2003). These observations are compatible with solutions **S.5** and **S.6.a**, where the level of co-infected hosts in class S depends on several different model parameters. In particular, it grows when the rate of joint y - s transmission, p_s increases, and *vice versa*, reaching 100% in situations where, in the context of our model, both the replication and the propagation of y become fully dependent on the presence of s . Still, we cannot tell the specific mechanism (i.e. parameter value, or set of), allowing for coexistence.

Our main qualitative result is that stable coexistence of all players is possible under a broad range of parameters. Though biological and epidemiological outcomes of interactions among co-infecting viruses

411 seem to be hardly predictable (Syller, 2020), long-lived loose interactions as those arising between the
412 agents in our model create favorable circumstances for evolution towards more permanent associations
413 and to promote recombination between dissimilar viral and subviral particles (Elena, 2011; García-Arenal
414 et al., 2003). For example, recombination has led to the emergence of new species, both in DNA viruses as
415 begomoviruses (Chakraborty et al., 2008), and in RNA viruses, as is the case of *Watermelon mosaic virus*,
416 which arose through recombination of two legume-infecting viruses (Desbiez et al., 2011).

417 All examples of satellite viruses summarized in Fig. 1 represent necessary, asymmetric associations that
418 have likely evolved from more loose interactions between helper viruses and satellites. In a parsimonious
419 evolutionary scenario, it is sensible to assume that a monopartite virus coding for all functions needed
420 to complete the viral cycle modified at some point its genome through the association with a satellite
421 coding for genes providing new or different functions, either improving its packaging efficiency or its
422 transmissibility, or modifying its infecting phenotype. A colonization of a new host where the genes
423 provided by the satellite was instrumental could have also been a plausible evolutionary pathway, followed
424 by the elimination of the old, now useless, equivalent gene in the viral genome. This sort of reasoning
425 agrees with hypotheses relating high viral diversity to advantages at early phases of host colonization
426 (Gonzalez-Jara et al., 2009), where loose associations could quickly modify the viral phenotype without
427 the need of incorporating genetic changes (Lucía-Sanz and Manrubia, 2017).

428 Coexistence of viral and subviral agents in a single host may act as well as stepping stones towards
429 mutually dependent necessary associations, as bipartite viral forms. Symmetric associations characterized
430 by mandatory complementation between the parties could arise through different evolutionary pathways.
431 They could evolve from asymmetric associations, as these described in the previous paragraph, when the
432 survival of the helper virus becomes more dependent on the satellite. Also, they could be generated *de*
433 *novo* through associations between complementary, defective particles. Finally, fully fledged viruses of
434 dissimilar origin could produce a new phenotype and, eventually, streamline their genomes to minimize
435 functional redundancy, in a sort of viral Black Queen association (Morris et al., 2012). The three pathways
436 above could eventually yield what is usually described as a bipartite virus. Further particles could be added
437 to the system through similar processes, causing in a parsimonious way the emergence of multipartite viral
438 forms.

CONFLICT OF INTEREST STATEMENT

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

AUTHOR CONTRIBUTIONS

AL, JA and SM participated in the overall conception and design of the study; SM designed the models; AL implemented the models and performed the simulations; AL and JA carried out the analytical calculations. AF and FG made substantial contributions to the discussion and interpretation of the results. AL and SM wrote the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

Grant PID2020-113284GB-C21 (SM) funded by MCIN/AEI/10.13039/501100011033; grant RTI2018-094302-B-I00 (AF, FG) funded by MINECO. The Spanish MICINN has also funded the “Severo Ochoa” Centers of Excellence to CNB, SVP-2014-068581 (AL), SEV 2017-0712 (AL, JA, SM) and to CBGP, SEV 2016-0672 (AF, FG). JA received support from the Spanish State Research Agency (AEI) through project MDM-2017-0737 Unidad de Excelencia “María de Maeztu”—Centro de Astrobiología (CSIC-INTA). AL is funded by the Centers for Disease Control and Prevention (grant 75D30121P10600) and the Army Research Office (grant W911NF1910384).

ACKNOWLEDGMENTS

SUPPLEMENTAL DATA

A Supplementary Information text file includes details on calculations retrieving the fixed points of the dynamics and conditions for stability.

DATA AVAILABILITY STATEMENT

The computational code required to generate all the results in this study is publicly available at <https://github.com/aluciasanz/satcompet.git>

REFERENCES

- Alcaide, C., Rabadán, M. P., Moreno-Pérez, M. G., and Gómez, P. (2020). Chapter five - implications of mixed viral infections on plant disease ecology and evolution. In *Advances in Virus Research*, eds. M. Kielian, T. C. Mettenleiter, and M. J. Roossinck (Academic Press), vol. 106. 145–169
- Betancourt, M., Escriu, F., Fraile, A., and García-Arenal, F. (2013). Virulence evolution of a generalist plant virus in a heterogeneous host system. *Evolutionary Applications* 6, 875–890
- Betancourt, M., Fraile, A., and García-Arenal, F. (2011). Cucumber mosaic virus satellite RNAs that induce similar symptoms in melon plants show large differences in fitness. *J. Gen. Virol.* 92, 1930–1938
- Chakraborty, S., Vanitharani, R., Chattopadhyay, B., and Fauquet, C. M. (2008). Supervirulent pseudorecombination and asymmetric synergism between genomic components of two distinct species

- of begomovirus associated with severe tomato leaf curl disease in india. *Journal of General Virology* 89, 818–828. doi:<https://doi.org/10.1099/vir.0.82873-0>
- Cotmore, S. F., Agbandje-McKenna, M., Chiorini, J. A., Mukha, D. V., Pinteland, D. J., Qiu, J., et al. (2014). The family *Parvoviridae*. *Arch Virol* 159, 1239–47
- Desbiez, C., Joannon, B., Wipf-Scheibel, C., Chandeysson, C., and Lecoq, H. (2011). Recombination in natural populations of watermelon mosaic virus: new agronomic threat or damp squib? *Journal of General Virology* 92, 1939–1948. doi:<https://doi.org/10.1099/vir.0.031401-0>
- Doumayrou, J., Sheber, M., Bonning, B. C., and Miller, W. A. (2016). Role of pea enation mosaic virus coat protein in the host plant and aphid vector. *Viruses* 8. doi:10.3390/v8110312
- Elena, S. (2011). Evolutionary constraints on emergence of plant RNA viruses. *Recent Advances in Plant Virology* , 283–300
- Elena, S. F., Bernet, G. P., and Carrasco, J. L. (2014). The games plant viruses play. *Current Opinion in Virology* 8, 62–67
- Frank, S. A. (2000). Within-host spatial dynamics of viruses and defective interfering particles. *Journal of Theoretical Biology* 206, 279–290
- García-Arenal, F. and Fraile, A. (2017). Chapter 56 - origin and evolution of satellites. In *Viroids and Satellites*, eds. A. Hadidi, R. Flores, J. W. Randles, and P. Palukaitis (Boston: Academic Press). 605–614. doi:<https://doi.org/10.1016/B978-0-12-801498-1.00056-5>
- García-Arenal, F., Fraile, A., and Malpica, J. (2003). Variation and evolution of plant virus populations. *International microbiology : the official journal of the Spanish Society for Microbiology* 6, 225–32. doi:10.1007/s10123-003-0142-z
- Gonzalez-Jara, P., Fraile, A., Canto, T., and García-Arenal, F. (2009). The multiplicity of infection of a plant virus varies during colonization of its eukaryotic host. *J. Virol.* 83, 7487–7494
- Grande-Pérez, A., Lázaro, E., Domingo, E., and Manrubia, S. C. (2005). Suppression of viral infectivity through lethal defection. *Proc. Natl. Acad. Sci. USA* 102, 4448–4452
- Hadid, A., Flores, R., Randles, J. W., and Palukaitis, P. (2017). *Viroids and Satellites* (Elsevier Science), 1st edn.
- Havelda, Z., Hornyik, C., Válczi, A., and Burgyán, J. (2005). Defective interfering RNA hinders the activity of a tombusvirus-encoded posttranscriptional gene silencing suppressor. *Journal of Virology* 79, 450 – 457
- Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review* 42, 599–653
- Jeger, M. J. (2020). The epidemiology of plant virus disease: Towards a new synthesis. *Plants* 9
- Kassanis, B. (1962). Properties and behaviour of a virus depending for its multiplication on another. *J. Gen. Microbiol.* 27, 477–488
- Khoshnan, A. and Alderete, J. F. (1995). Characterization of double-stranded RNA satellites associated with the *Trichomonas vaginalis* virus. *J Virol* 69, 6892–97
- King, A. M., Adams, M. J., Carstens, E. B., and Lefkowitz, E. J. (eds.) (2012). *The Subviral Agents* (San Diego: Elsevier)
- Kirkwood, T. B. L. and Bangham, C. R. (1994). Cycles, chaos, and evolution in virus cultures: a model of defective interfering particles. *Proc Natl Acad Sci USA* 91, 8685–89
- Krupovic, M. and Cvirkaite-Krupovic, V. (2012). Towards a more comprehensive classification of satellite viruses. *NRM* 234
- Krupovic, M., Kuhn, J. H., and Fisher, M. G. . (2016). A classification system for virophages and satellite virus. *Arch. Virol.* 161, 233–247

- 510 Leeks, A., Segredo-Otero, E. A., Sanjuán, R., and West, S. A. (2018). Beneficial coinfection can promote
511 within-host viral diversity. *Virus evolution* 4, vey028–vey028
- 512 Lehmann, L., Mullan, C., Akçay, E., and Van Cleve, J. (2016). Invasion fitness, inclusive fitness, and
513 reproductive numbers in heterogeneous populations. *Evolution* 70, 1689–1702
- 514 Lucía-Sanz, A. and Manrubia, S. (2017). Multipartite viruses: adaptive trick or evolutionary treat? *npj*
515 *Systems Biology and Applications* 3, 34. doi:10.1038/s41540-017-0035-y
- 516 Lucía-Sanz, A., Aguirre, J., and Manrubia, S. (2018). Theoretical approaches to disclosing the emergence
517 and adaptive advantages of multipartite viruses. *Current Opinion in Virology* 33, 89–95. doi:https:
518 //doi.org/10.1016/j.coviro.2018.07.018. Virus vector interactions • Special Section: Multicomponent
519 viral systems
- 520 Makino, S., Chang, M.-F., Shieh, C.-K., Kamahora, T., Vannier, D. M., Govindarajan, S., et al. (1987).
521 Molecular cloning and sequencing of a human hepatitis delta (δ) virus RNA. *Nature* 329, 343–346
- 522 McLeish, M., Fraile, A., and García-Arenal, F. (2019a). Evolution of plant–virus interactions: host range
523 and virus emergence. *Current Opinion in Virology* 34, 50–55. doi:10.1016/j.coviro.2018.12.003
- 524 McLeish, M., Sacristán, S., Fraile, A., and García-Arenal, F. (2019b). Coinfection organizes
525 epidemiological networks of viruses and hosts and reveals hubs of transmission. *Phytopathology*
526 109, 1003–1010. doi:10.1094/PHYTO-08-18-0293-R
- 527 Morris, J., Lenski, R., and Zinser, E. (2012). Morris jj, lenski re, zinser er.. the black queen hypothesis:
528 evolution of dependencies through adaptive gene loss. *mbio* 3: e00036-12. *mBio* 3. doi:10.1128/mBio.
529 00036-12
- 530 Nakashima, N., Kawahara, N., Omura, T., and Noda, H. (2006). Characterization of a novel satellite
531 virus and a strain of Himetobi P virus (*Dicistroviridae*) from the brown planthopper, *Nilaparvata lugens*.
532 *Journal of Invertebrate Pathology* 91, 53–56
- 533 Nee, S. (2000). Mutualism, parasitism and competition in the evolution of coviruses. *Phil. Trans. R. Soc.*
534 *Lond. B* 355, 1607–1613
- 535 Palukaitis, P. and García-Arenal, F. (2003). Cucumoviruses. *Advances in virus research* 62, 241–323.
536 doi:10.1016/S0065-3527(03)62005-1
- 537 Palukaitis, P., Rezaian, A., and García-Arenal, F. (2008). Satellite nucleic acids and viruses. In *Encyclopedia*
538 *of Virology (Third Edition)*, eds. B. W. Mahy and M. H. Van Regenmortel (Oxford: Academic Press).
539 Third edition edn., 526–535
- 540 Palukaitis, P. and Roossinck, M. J. (1996). Spontaneous change of a benign satellite RNA of cucumber
541 mosaic virus to a pathogenic variant. *Nature Biotechnology* 14, 1264–1268
- 542 Pinel, A., Abubakar, Z., Traore, O., Konaté, G., and Fargette, D. (2003). Molecular epidemiology of
543 the rna satellite of rice yellow mottle virus in africa. *Archives of Virology* 148, 1721–33. doi:10.1007/
544 s00705-003-0138-1
- 545 Pruss, G., Ge, X., Shi, X. M., Carrington, J. C., and Bowman Vance, V. (1997). Plant viral synergism:
546 the potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of
547 heterologous viruses. *The Plant Cell* 9, 859–868
- 548 Rehman, M. S. N. u. and Fauquet, C. M. (2009). Evolution of geminiviruses and their satellites. *FEBS Lett.*
549 583, 1825–1832
- 550 Ribière, M., Olivier, V., and Blanchard, P. (2010). Chronic bee paralysis: A disease and a virus like no
551 other? *J Invert Pathol* 103, S120 – S131
- 552 Roossinck, M. . J. ., Sleat, D., and Palukaitis, P. . (1992). Satellite RNAs of plant viruses: structures and
553 biological effects. *Microbiol Revs* 56, 256–279

- 554 Roossinck, M. J. (2005). Symbiosis versus competition in plant virus evolution. *Nat. Revs. Microbiol.* 3,
555 917–924
- 556 Scheets, K. (1998). Maize chlorotic mottle machlomovirus and wheat streak mosaic rymovirus
557 concentrations increase in the synergistic disease corn lethal necrosis. *Virology* 242, 28 – 38
- 558 Schmitt, M. J. and Breining, F. (2002). The viral killer system in yeast: from molecular biology to
559 application. *FEMS Microbiol Rev* 26, 257–275
- 560 Sicard, A., Michalakakis, Y., Gutiérrez, S., and Blanc, S. (2016). The strange lifestyle of multipartite viruses.
561 *PLoS Path.* 12, e1005819
- 562 Stanley, J., Saunders, K., Pinner, M. S., and Wong, S. M. (1997). Novel defective interfering dnas
563 associated with ageratum yellow vein geminivirus infection of ageratum conyzoides. *Virology* 239, 87 –
564 96
- 565 Syller, J. (2003). Molecular and biological features of umbraviruses, the unusual plant viruses lacking
566 genetic information for a capsid protein. *Physiological and Molecular Plant Pathology* 63, 35–46
- 567 Syller, J. (2020). Chapter 30 - interspecific and intraspecific interactions among plant viruses in mixed
568 infections. In *Applied Plant Virology*, ed. L. Awasthi (Academic Press). 437–453
- 569 Szathmáry, E. (1992). Viral sex, levels of selection, and the origin of life. *J. Theor. Biol.* 159, 99–109
- 570 Taliansky, M. E. and Robinson, D. J. (2003). Molecular biology of umbraviruses: phantom warriors.
571 *Journal of General Virology* 84, 1951–1960. doi:<https://doi.org/10.1099/vir.0.19219-0>
- 572 Yau, S., Lauro, F. M., DeMaere, M. Z., Brown, M. V., Thomas, T., Raftery, M. J., et al. (2011). Virophage
573 control of antarctic algal host–virus dynamics. *Proc Natl Acad Sci USA* 108, 6163–6168
- 574 Zwart, M. P., Blanc, S., Johnson, M., Manrubia, S., Michalakakis, Y., and Sofonea, M. T. (2021). Unresolved
575 advantages of multipartitism in spatially structured environments. *Virus Evolution* 7. doi:10.1093/ve/
576 veab004. Veab004

FIGURE CAPTIONS

Satellite virus	ssRNA	Tombusviridae, Virgaviridae, Nodaviridae Satellite CBPV, Albetovirus, Aumivavirus, Papanivirus, Virtovirus, Macronovirus Nilaparvata lugens commensal X virus	
	dsDNA (Viriophage)	Mimiviridae, Phycodnaviridae Sputnikvirus, Mavirus, Organic lake satellite	
	ssDNA	Adenoviridae Herpesviridae	Dependoparvovirus
Satellite Nucleic Acid	ssDNA	Geminiviridae	Alphasatellite
		Nanoviridae	Betasatellite
		Geminiviridae	Deltasatellite
	dsRNA	Totiviridae M virus satellite, T1 virus satellite	
	ssRNA	Large linear	Secoviridae, Alphaflexividae
		Small linear	Tombusviridae, Bromoviridae
		Small circular	Secoviridae, Luteoviridae
Small circular (Virusoid)		Solemoviridae Hepadnaviridae Deltavirus	
DI			

Figure 1. Summary of classes of satellite viruses. There is no straightforward classification of satellites, although distinguishable groups come to light when looking at the genetic material, helper virus family, host, or coded proteins. Here we show a possible classification into two groups, depending on whether they code for a capsid protein (satellite viruses) or not (satellite nucleic acid). Satellite viruses comprise plant satellites with jelly roll capsid proteins, dependoparvovirus and viriophages (Krupovic and Cvrkaite-Krupovic, 2012; Krupovic et al., 2016). Satellite nucleic acids include α/β -satellites and *Secoviridae* satellites coding for a replicase or a replication helper protein; M virus satellite expresses a toxin and HDV codes for an antigen. This group includes non-coding —circular— ssDNAs (Stanley et al., 1997) and RNAs with a compact folded structure with ribozyme activity (Hadid et al., 2017; Roossinck et al., 1992). Colors on the right bars show the protein each satellite codes for: capsid (blue), replicase (green), other (yellow). Grey stands for no coded protein. Helper virus families are shown in bold font; DI stands for defective interfering genomes (e.g., coated like defective interfering particles or uncoated).

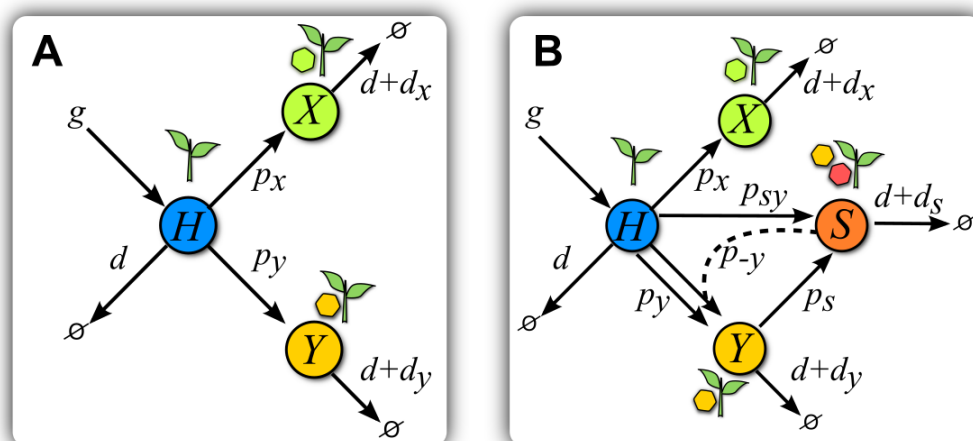


Figure 2. Scheme of models. (A) Scheme of Model V. Healthy hosts H are seeded at a constant rate g and die with basal rate d . Transitions to infected states happen at rates p_x and p_y for each virus. Infected hosts X and Y suffer from an increase d_i in basal mortality. (B) Scheme of Model S. The introduction of a satellite with y as helper virus adds new transitions with respect to (A). Hosts in class Y can get infected at a rate p_s by a satellite upon contact with hosts in class S . Hosts in this latter class are simultaneously infected by virus y and its satellite, and have an increase d_s in their mortality rate. Infection of H plants by the helper virus and the satellite simultaneously occurs at a rate p_{sy} upon contact with class S . A co-infected plant S can also infect H plants only with virus y at a rate p_{-y} . The color code is maintained all through the paper.

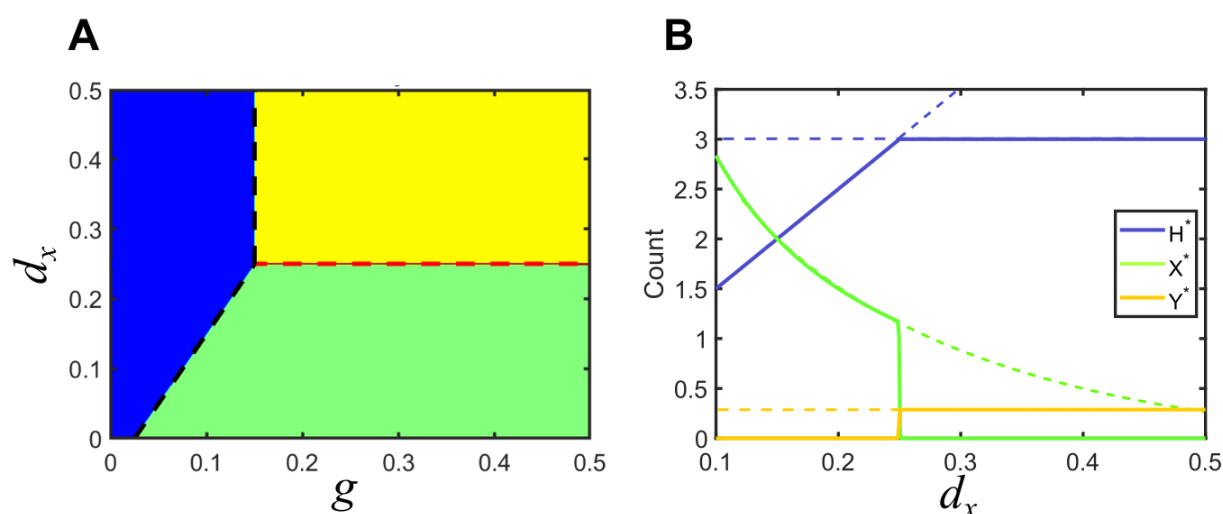


Figure 3. Solutions for Model V in parameter space. (A) Phase space of solutions in the plane $\{g, d_x\}$. In the blue region, none of the viruses can invade the population of hosts, $d/g > F_i$, $i = x, y$. The black dashed line indicates the boundary where $d/g = F_i$. Green and yellow regions correspond to solutions that satisfy $F_x > F_y$ and $F_y > F_x$, respectively, where one of the viruses invades the population of hosts and displaces the other. The red dashed line corresponds to the marginal coexistence solution. (B) Bifurcation diagram for a cross section of the phase space in A corresponding to $g = 0.5$. Lines indicate the amount of each class at the fixed point, as shown in the legend. Dashed lines represent unstable solutions, while solid lines correspond to stable solutions. Green and yellow regions in A correspond to $d_x < 0.25$ and $d_x > 0.25$. Vertical lines show the degenerate solution of coexistence of the two viruses for $d_x = 0.25$. Note that the values of H^* and Y^* become independent of d_x in the region where y is the invading virus (and thus x becomes extinct). Parameters used: $d = 0.05$, $p_x = 0.1$, $p_y = 0.1$, $d_y = 0.25$.

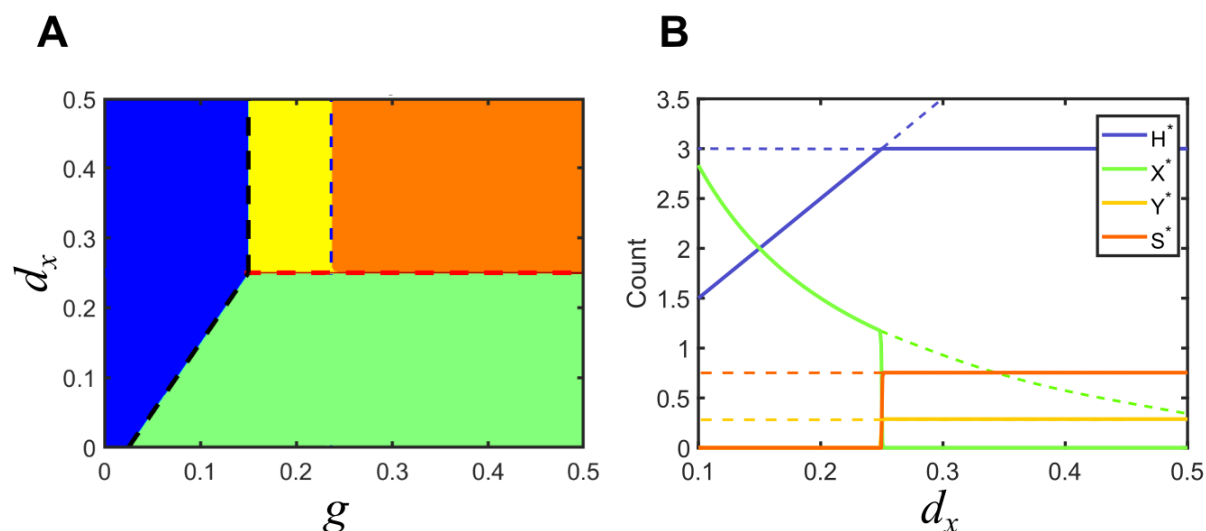


Figure 4. Solutions for model S in parameter space: commensal satellite. (A) Phase space of solutions in the plane $\{g, d_x\}$ for a commensal satellite ($F_c = F_y$). The description of regions corresponding to solutions **S.1** (in blue, viruses cannot infect hosts), **S.2** (in green, virus x invades and virus y becomes extinct) and **S.3** (in yellow, virus y invades and virus x becomes extinct) are analogous to those described for Model V. In the new orange region, the satellite stably coexists with its helper virus, **S.5**. Red and black dashed lines are as in Fig. 3. The blue dashed line signals the boundary separating the regions with and without co-circulating satellite. Note that the absence of virus y prevents the stable circulation of the satellite (below $d_x = 0.25$). (B) Bifurcation diagram of a cross section of A, for $g = 0.5$. As above, lines stand for the amount of each class at equilibrium. Dashed lines correspond to unstable solutions, while solid lines correspond to stable solutions. Solutions collide with the degenerated solution **S.4** at $d_x = 0.25$. Parameters used: $d = 0.05, p_x = 0.1, p_y = 0.1, d_y = 0.25, p_s = 0.8, p_{sy} = 0.04, d_s = 0.3, p_{-y} = 0.0766$.

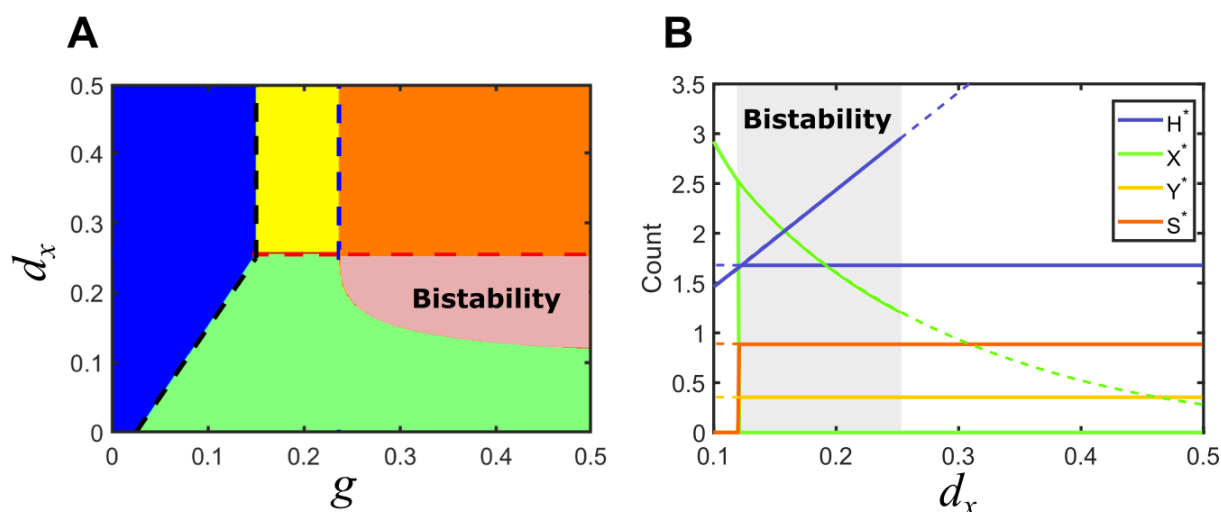


Figure 5. Solutions for model S in parameter space: mutualistic satellite. (A) Phase space of solutions in the plane $\{g, d_x\}$ for a mutualistic satellite. Here, $F_c > F_y$. Some of the previous equilibrium solutions are maintained, as long as the inequalities that define them remain unchanged. However, a region of bistability appears, where solutions **S.2** and **S.5** are possible for the same parameters, depending on the initial conditions. (B) Bifurcation diagram of a cross section of A, for $g = 0.5$. As above, lines stand for the amount of each class at equilibrium. Dashed lines correspond to unstable solutions, while solid lines correspond to stable solutions. A region of bistability for $0.12 < d_x < 0.25$ is shown in gray. A bifurcation where two solutions appear occurs at $d_x = 0.12$; solutions collide with the degenerated solution **S.4** and stability is modified at $d_x = 0.25$. Parameters used: $d = 0.05, p_x = 0.1, p_y = 0.1, d_y = 0.25, p_s = 0.8, p_{sy} = 0.04, d_s = 0.3, p_{-y} = 0.2$.

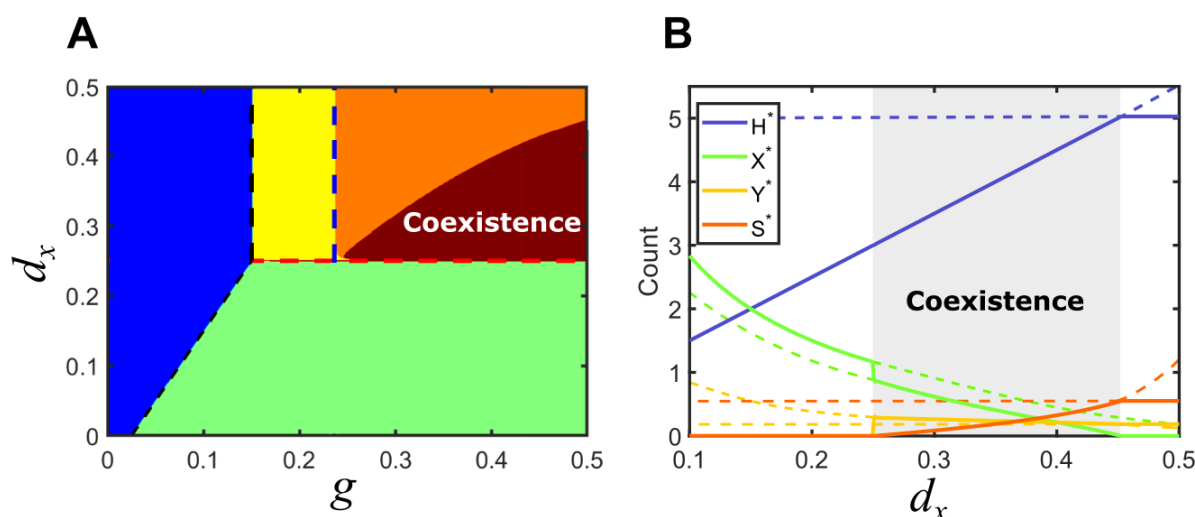


Figure 6. Solutions for model S in parameter space: parasitic satellite. (A) Phase space of solutions in the plane $\{g, d_x\}$ for a parasitic satellite in the particular case $F_c < F_y$. A region of co-existence occupies part of the area where the virus y -satellite eliminated virus x if the satellite was commensal or mutualistic. (B) Bifurcation diagram of a cross-section of A, for $g = 0.5$. A region of co-existence of all the populations is shown in gray for $0.25 < d_x < 0.4529$, corresponding to solution **S.6**. Solutions collide with the degenerated solution **S.4** at $d_x = 0.25$; a bifurcation occurs at $d_x = 0.4529$. Parameters used: $d = 0.05, p_x = 0.1, p_y = 0.1, d_y = 0.25, p_s = 0.8, p_{sy} = 0.04, d_s = 0.3, p_{-y} = 0.016$.

Fixed point	Classes at equilibrium	Existence, non negativity and stability
V.1: $\left(\frac{g}{d}, 0, 0\right)$	  	$F_x < d/g$ and $F_y < d/g$
V.2: $\left(\frac{d+d_x}{p_x}, \frac{g}{d+d_x} - \frac{d}{p_x}, 0\right)$	  	$F_x > d/g$ and $F_x > F_y$
V.3: $\left(\frac{d+d_y}{p_y}, 0, \frac{g}{d+d_y} - \frac{d}{p_y}\right)$	  	$F_y > d/g$ and $F_y > F_x$
V.4: $\left(\frac{d+d_x}{p_x}, X^*, \frac{g}{d+d_y} - \frac{d}{p_y} - \frac{p_x}{p_y} X^*\right)$	  	$F_x = F_y > \frac{d}{g}$ and $\frac{g}{d+d_x} - \frac{d}{p_x} > X^* > 0$

Table 1. Properties of the solutions of model V. The left column shows the analytical solution (fixed point of the dynamics, (H^*, X^*, Y^*)); the central column symbolically illustrates the states with positive population at equilibrium (colored circles), gray circles are for classes with zero population. The right column specifies the conditions for existence, non-negativity and stability of each solution.

TABLES

Fixed point	Classes at equilibrium	Existence, non negativity and stability
S.1: $\left(\frac{g}{d}, 0, 0, 0\right)$	   	$F_x < d/g$, $F_y < d/g$ and $F_{sy} < d/g$
S.2: $\left(\frac{d+d_x}{p_x}, \frac{g}{d+d_x} - \frac{d}{p_x}, 0, 0\right)$	   	$F_x > d/g$, $F_x > F_y$ and $F_x > F_{sy}$
S.3: $\left(\frac{d+d_y}{p_y}, 0, \frac{g}{d+d_y} - \frac{d}{p_y}, 0\right)$	   	$F_y > d/g$, $F_y > F_x$ and $F_y(1 - \gamma) > F_{sy}$
S.4: $\left(\frac{d+d_x}{p_x}, X^*, \frac{g}{d+d_y} - \frac{d}{p_y} - \frac{p_x}{p_y} X^*, 0\right)$	   	$F_y = F_x > d/g$ and three others, see SI
S.5: $(H^*, 0, Y^*(H^*), S^*(H^*))$	   	See SI
S.6: $\left(\frac{d+d_x}{p_x}, X^*(H^*), Y^*(H^*), S^*(H^*)\right)$	   	See SI

Table 2. Properties of the solutions of model S. The left column shows the analytical solution (fixed point of the dynamics, (H^*, X^*, Y^*, S^*)); the right column summarizes all conditions for existence, non-negativity and stability of that solution. We have defined $\gamma = (g/(d + d_y) - d/p_y)p_s/(d + d_s)$, see main text. Solution **S.4** is degenerate in the same sense than **V.4** was: there is a continuum of values of X^* and Y^* that satisfy the fixed point equations. Further, it is a marginally stable solution, since parameter values have to be fine-tuned for it to persist. This is the reason it requires strict conditions that we detail in the Supplementary Information. Coexistence solution **S.6** takes two different forms, therefore embracing two different fixed points (see main text and SI for further details).