Onset of virus systemic infection in plants is determined by speed of cell-to-cell movement and number of primary infection foci

Guillermo Rodrigo1,2, Mark P. Zwart2 and Santiago F. Elena2,3
1Institute of Systems and Synthetic Biology, CNRS—Université d’Évry Val d’Essonne—Genopole, Évry 91030, France
2Instituto de Biología Molecular y Celular de Plantas, CSIC—Universidad Politécnica de Valencia, Valencia 46022, Spain
3The Santa Fe Institute, Santa Fe, NM 87501, USA

The cornerstone of today’s plant virology consists of deciphering the molecular and mechanistic basis of host–pathogen interactions. Among these interactions, the onset of systemic infection is a fundamental variable in studying both within- and between-host infection dynamics, with implications in epidemiology. Here, we developed a mechanistic model using probabilistic and spatio-temporal concepts to explain dynamic signatures of virus systemic infection. The model dealt with the inherent characteristic of plant viruses to use two different and sequential stages for their within-host propagation: cell-to-cell movement from the initial infected cell and systemic spread by reaching the vascular system. We identified the speed of cell-to-cell movement and the number of primary infection foci in the inoculated leaf as the key factors governing this dynamic process. Our results allowed us to quantitatively understand the timing of the onset of systemic infection, describing this global process as a consequence of local spread of viral populations. Finally, we considered the significance of our predictions for the evolution of plant RNA viruses.

1. Introduction

Virus colonization of a multicellular host is a highly complex process that involves interactions within the infecting viral population as well as interactions between the pathogen and the host organism [1]. Unravelling and quantitatively describing these interactions is essential for understanding the infection process in any depth and, moreover, for successfully predicting disease dynamics and to a more efficient design of antiviral interventions. Biochemistry has been harnessed to elucidate mechanisms underlying infection at the cellular and molecular levels, rendering spectacular results and powerful applications [2–4]. Similarly, the application of computational methods to study spread at the population level has led to striking successes in both understanding of infectious diseases and mitigating their impact [5–7]. However, infection dynamics at the within-host level are, in many ways, poorly understood, in particular with respect to providing quantitative descriptions of how infection progresses through time and space. Barriers to scientific progress include the difficulty—using non-destructive methods and in real time of infection—to target tissues where the virus replicates, and the complexity of the host immune system.

Plant RNA viruses are highly suitable model systems for studying within-host infection dynamics, and RNA viruses the main targets of replication—the mesophyll cells—are located in the largely planar leaves. This allows infection to be followed using a wide array of non-destructive approaches, including the expression of marker proteins [8–10]. Moreover, leaves can be easily and cleanly removed, effectively allowing for the removal of host organs at any point during infection and for their further study [11]. Second, plant immune responses to
RNA viruses are mainly composed of the hypersensitive response, leading to the formation of visibly discernible local lesions [12,13]. RNA interference, where the presence of double-stranded RNA induces the local and systemic destruction of viral RNA [2,14], hormone-mediated (e.g. salicylic acid) defence pathways [15] and, as recently shown, pattern-triggered immunity [16]. This effective set of host immune responses have contributed to a rich tradition of quantitative research on infection dynamics using plant viruses [9,13,17–19]. Moreover, this array of plant immune mechanisms probably contributes to the relatively low between-host variation typically found in experimental settings [20]. Finally, although plants are invaluable model systems, they are also of great agricultural importance, as are the many diseases thwarting crop production.

Here, we consider the dynamics of a critical step during the plant RNA virus infection cycle: the initiation of systemic infection [11]. Plant viruses expand within the host by (i) cell-to-cell movement, the local spread of infection by virions or ribonuclear complexes, and by (ii) systemic movement, whereby the virus accesses the vascular tissue and is transported within and between leaves [8]. Plant viruses are usually phloem transported, meaning that they will spread towards tissues that import photo-assimilates, typically including apical tissues. Therefore, upon the establishment of a primary infection focus—a cell infected following inoculation by a viruliferous vector or mechanical means—a virus can only colonize a small region of the inoculated leaf by cell-to-cell movement [20]. The probability of between-host transmission will be highest if a virus infects the majority of host tissues, including rapidly growing apical tissues. To initiate such a systemic infection of the plant, the virus needs to (i) reach the host vasculature, (ii) be loaded into the phloem to then be transported out of the inoculated leaf and (iii) unload elsewhere in the host [1,21].

Here, we attempt (i) to better understand the timing of the onset of systemic infection, (ii) to identify what are the key factors governing this dynamic process and finally, (iii) to consider its significance for the evolution of plant RNA viruses.

We develop a simple theory that is able to explain differences in dynamic signatures of viral systemic infection, based on the speed of cell-to-cell movement and number of primary infection foci in the inoculated leaf (figure 1). Moreover, we confront these models with experimental data in order to test their validity. Our work is of relevance as it contributes to tackle mathematically questions about the relationship between local and global virus movement, the determinants of the variance of infected individual plants in a population and the interaction between virus genetic factors (e.g. speed of cell-to-cell movement) and environmental conditions (e.g. dose of the inoculum) to get systemic infection.

2. Results

2.1. Model definition

Primary infection foci tend to be quite circular in form and of approximately the same size [8,9,20]. To model the process of radial expansion of primary infection foci, we introduced the effective diffusion coefficient of the local infection (D). Although viral cell-to-cell movement is complex and there may be stochastic effects, for our model we considered isotropic diffusion. Thus, we could write

\[ D = \frac{dA(t)}{dt}, \]  

where \( A(t) \) is the circular area covered by the virus at a given time \( t \)—usually measured in hours post-inoculation (hpi). Typical values of \( D \) are 0.01–0.03 mm\(^2\) h\(^{-1}\) [22,23]. This reaction–diffusion system, where viral particles replicate within cells and then move to neighbour cells to proliferate, constitutes the local infection dynamics—the first step for long-distance infection—before the virus reaches the host vasculature and is loaded into the phloem. We then assumed that the area required to reach the vascular system follows a normal distribution, with mean \( \mu \) and standard deviation \( \sigma \). This is justified by the regular spacing observed between veins and the...
way the particular developmental programmes operate [24], which could reflect a given evolutionary contingency linking leaf shape and vascular patterning. Of note, \( \mu \) and \( \sigma \) are, at least in theory, host-dependent parameters that vary with the density of the vascular system in the leaf, being independent of the virus.

The time to reach systemic infection needs also to account, among other factors, for latency periods \( (T_0) \)—e.g. time for replication in the first infected cell—and vascular movements of the virus \( T_1 \). Following the seminal work by Samuel [25], there is slow cell-to-cell movement of the virus in the infection foci, combined with rapid dissemination throughout the plant via the vascular system. According to recent experimental data, the time required to complete the infection cycle of Tobacco mosaic virus in the primary inoculated cell of Nicotiana tabacum is \( T_0 = 18–20 \) hpi [26]. This is in tune with model predictions showing latency periods of about 1 day [17]. In addition, virions move through the veins and midrib at rates of 20–50 mm h\(^{-1} \) [22], which gives 2–5 h to cover a space of 100 mm (typical rough length to reach the stem from the primary infection foci). \( T_1 \) also accounts for systemic movements through further vascular tissues to infect the whole plant, a process that may take additional hours, depending on the host molecular infrastructure [1], and that we did not model in this work. For simplicity, we joined together these terms in a single parameter \( (T = T_0 + T_1) \), which may depend on the whole plant pathosystem. An upper bound of \( T \) can be obtained as the minimal time measured when systemic infection appears—e.g. about 40 hpi for Tobacco etch virus (TEV) infecting \( N. \) tabacum [11].

Therefore, by integrating equation (2.1), we had

\[
A(t) = D\int_{T_1}^{t-T_1} dt = D(t-T),
\]

where \( t-T \) represents the time for which the virus moves locally in the primary infection foci, being \( t \) the absolute time. Because, on average, systemic infection is produced when the infected area in the inoculated leaf is \( \mu \), we obtained

\[
\mu = D(t_{sys}) - T,
\]

where \( t_{sys} \) is the average time of systemic infection—the time at which 50% of the individuals in the population present systemic infection. Moreover, simply by squaring and averaging equation (2.2), we obtained

\[
\sigma = D\Delta t_{sys},
\]

where \( \Delta t_{sys} \) is the standard deviation of systemic infection time.

To calculate the probability of systemic infection at a given time, \( P_{sys}(t) \), as \( A(t) \) is assumed Normally distributed, we followed the corresponding cumulative distribution function (where erf is the error function) to derive

\[
P_{sys}(t) = \begin{cases} 
0, & t \leq T \\
\frac{1}{2} \left[ 1 + \text{erf} \left( \frac{A(t) - \mu}{\sigma \sqrt{2}} \right) \right], & t > T
\end{cases}
\]

This expression can be rewritten in terms of time, instead of area, by using equations (2.2)–(2.4). In particular, we derived the following equality:

\[
\frac{A(t) - \mu}{\sigma} = \frac{t - t_{sys}}{\Delta t_{sys}},
\]

which shows the equivalence between standardized areas and times.

Importantly, this model can explain differences in dynamic signatures of systemic infection between viruses. Indeed, the speed of cell-to-cell movement \( (D) \) appears as a fundamental parameter. For instance, if we denote two viruses infecting a common host organism as \( \Lambda \) and \( \beta \), and given that \( \mu \) and \( \sigma \) are parameters that only depend on the host organism, it turns out

\[
\delta = \frac{D^\Lambda}{D^\beta} \frac{(t_{sys}^\beta - T)}{(t_{sys}^\Lambda - T)},
\]

having defined \( \delta \) as the ratio of speed of cell-to-cell movement between viruses, and

\[
\frac{D^\Lambda}{D^\beta} = \frac{\Delta t_{sys}^\beta}{\Delta t_{sys}^\Lambda}.
\]

Therefore, we can predict relative changes in the average time of systemic infection and the associated variance—magnitudes that quantify the infection at the global level—according to differences in the speed of cell-to-cell movement, which describes infection dynamics at the local level.

In addition, it has been argued that the number of primary infection foci in the inoculated leaf plays an important role in determining when the plant will become systemically infected [11]. We therefore generalized this theoretical framework by also considering the effect of the number of primary infection foci in the inoculated leaf \( (N) \). We derived an analytical expression to calculate the effective probability of systemic infection for \( N \) different viral populations spreading at the same time, \( P_{sys}(t)_N \). Lafforgue et al. [11] showed that systemic infection by multiple viral populations does not depend on their interactions, thus it is justified to assume that \( P_{sys}(t)_N = 1 - (1 - P_{sys}(t))^N \), and therefore

\[
P_{sys}(t)_N = \begin{cases} 
0, & t \leq T \\
1 - \left[ 1 - \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{A(t) - \mu}{\sigma \sqrt{2}} \right) \right] \right]^N, & t > T
\end{cases}
\]

where for \( N = 1 \), equations (2.5) and (2.9) are the same. According to equation (2.5), \( P_{sys}(t_{sys}) = 0.5 \), by definition. However, this does not hold for the effective probability, \( P_{sys}(t_{sys})_N = 1 - 1/2^N \). Indeed, the effective average time of systemic infection, \( t_{sys,N} \), decreases with \( N \), \( t_{sys,N} < t_{sys} \), as systemic infection is produced once the vascular system is invaded by the first viral population. By imposing \( P_{sys}(t_{sys})_N = 0.5 \), the effective average time of systemic infection was

\[
t_{sys,N} = T + \max \left\{ \frac{\mu}{D} + \frac{\sigma}{D} \sqrt{2} \text{erf}^{-1}(1 - 2^{-1/N}), 0 \right\}
\]

\[
= \max \left\{ t_{sys} + \Delta t_{sys} \sqrt{2} \text{erf}^{-1}(1 - 2^{-1/N}), T \right\}.
\]

The term \( \text{erf}^{-1}(1 - 2^{-1/N}) \) is negative, describing the decrease with \( N \). Moreover, by imposing \( (\text{d}/\text{d}t)P_{sys}(t_{sys})_N = 1/\Delta t_{sys,N} \sqrt{2\pi} \), the effective standard deviation of systemic infection time was given by

\[
\Delta t_{sys,N} = \frac{\Delta t_{sys}}{N} \exp \left[ \frac{\text{erf}^{-1}(1 - 2^{-1/N})^2}{N - \frac{1}{N} \ln(2)} \right].
\]
where this expression is only valid for $\langle t_{sys} \rangle_N > T$. This value also decreases with $N$.

### 2.2. Experimental validation

To test this theory, we considered an experimental system with the plant pathogen *Turnip mosaic virus* (TuMV) and the host plant *Nicotiana benthamiana* (see [23]). To perform infections, we used a TuMV clone tagged with the enhanced green fluorescent protein (eGFP) [27] and a large population of four-week-old *N. benthamiana* plants. Five micrograms of infectious TuMV plasmid (GenBank no. AF530055.2) were mixed with 5 ml of inoculation buffer (carborundum 100 mg ml$^{-1}$, 50 mM potassium phosphate, pH 8) on one leaf per plant and gently rubbed. Inoculated plants were maintained in a growth chamber (16 h light at 25°C, 8 h dark at 22°C). In order to consider two pathogens with biologically relevant differences, we used viruses carrying two alleles of the Pretty Interesting Potyviridae ORF (*pipo*) cistron of different lengths [23]. *Pipo* is expressed as a C-terminal fusion to the P3 protein, referred to as P3N-PIPO. One viral strain carries an allele that is 70 amino acids long (*pipo$^{70}$*), whereas the second strain has a 61 amino acids long allele (*pipo$^{61}$*) [23]. Therefore, as P3N-PIPO is involved in viral movement [28], we expected differences in the speed of cell-to-cell and systemic movements. To obtain empirical estimates of $D$ for each viral strain, we analysed data for the formation of primary infection foci. Plants were mechanically inoculated with virions, and the inoculated leaf was examined at 48, 72 and 96 hpi. In figure 2a, we represent $A(t)$ for these two viral strains. Using equation (2.1), we estimated $D_{pipo^{70}} = 0.0527$ mm$^2$ h$^{-1}$ and $D_{pipo^{61}} = 0.0099$ mm$^2$ h$^{-1}$. This gave a ratio $\delta = D_{pipo^{70}}/D_{pipo^{61}} = 5.32$, and according to our model we predicted $\langle t_{sys} \rangle_{pipo^{70}} < \langle t_{sys} \rangle_{pipo^{61}}$ and $\Delta t_{sys, pipo^{70}} < \Delta t_{sys, pipo^{61}}$. We calculated $P_{sys}(t)$ as the fraction of systemically infected plants at a given time, examining plants at 24, 40, 48, 72 and 96 hpi. In this plant pathosystem, the number of primary infection foci was very limited—perhaps because of a low dose for inoculation—so, for simplicity, we considered $N = 1$. In fact, this is more a reformulation than a simplification (see equations (2.10) and (2.11)). In figure 2b, we represent $P_{sys}(t)$ for the two viral strains, showing significantly different dynamic signatures. We considered $T \approx 40$ hpi. By nonlinear regression of equation (2.5) to the data (figure 2b), we inferred firstly $\langle t_{sys} \rangle_{pipo^{70}} = 47.32$ hpi (with 95% CI [46.96, 47.67]) and $\Delta t_{sys, pipo^{70}} \approx 1.58$ h (with 95% CI [0.76, 2.40]) ($R^2 = 0.996$), and secondly $\langle t_{sys} \rangle_{pipo^{61}} = 80.35$ hpi (with 95% CI [73.01, 87.70])

**Figure 2.** Experimental results of two strains of TuMV with different *pipo* alleles (*pipo$^{61}$* and *pipo$^{70}$*). (a) Dynamics of expansion of infection foci. Error bars correspond to standard deviations. We estimated $D = 0.0527$ mm$^2$ h$^{-1}$ for *pipo$^{70}$* and $D = 0.0099$ mm$^2$ h$^{-1}$ for *pipo$^{61}$*. (b) Dynamics of systemic infection (circles and squares). Error bars correspond to standard deviations calculated according to a binomial distribution. Solid lines correspond to model predictions (equation (2.5)).
Figure 3. Experimental results of systemic infection at different times (hpi) as a function of the number of primary infection foci \((N)\). Solid lines correspond to model predictions. Equation (2.9), in combination with equation (2.6), was fitted to \(t_{Sys} = 51.95 \text{ hpi} \pm 5.81 \text{ hpi} \pm 4.45 \text{ hpi} \) (nonlinear regression).

and \(\Delta t_{Sys} = 21.14 \text{ hpi} \) (with 95% CI [10.71, 31.57]) \((R^2 = 0.998)\). The fit of the empirical data to a binary logistic regression model also shows a significant heterogeneity among the two dynamic signatures of systemic infection (Pearson goodness-of-fit test: \(\chi^2 = 86.084, 44 \text{ d.f.}, p < 0.001\)). As the model—equation (2.9)—fits the data well, it is safe to assume that there is no evidence for interactions between primary infection foci in terms of the onset of systemic infection. Moreover, using equation (2.10), we predicted for \(N = 100\), as illustrative case, an effective average time of systemic infection of \(t_{Sys}(t = 40)_{N=100} = 40.99 \text{ hpi}\), which agrees with the experimental data as \(P_{Sys}(t = 40)_{N=100} \approx 0.5\) in this plant pathosystem.

To better study the effect of the number of primary infection foci, we considered further experimental data on systemic infection (reported in [11]). The experimental system consisted of TEV, another potyvirus closely related to TuMV, and \(N. \ tabacum\) as the host plant. To perform infections, a TEV clone tagged with eGFP [27] and a population of 10 four-week-old \(N. \ tabacum\) plants was used for each dose, with variation in dose resulting in different numbers of primary infection foci. Plants were inoculated by abrasion of the third true leaf with 15 \(\mu\)L of the corresponding TEV dose and 5 \(\mu\)L of inoculation buffer. Inoculated plants were maintained in a BSL2 greenhouse at 25°C and 16 h light. In this case, \(D\) was fixed (one virus), and the distribution of foci ranged from \(N = 2–224\). Plants were examined at 40, 44, 46, 50 and 54 hpi. By nonlinear regression of equation (2.9) to data shown in figure 3, we obtained \(t_{Sys} = 51.95 \text{ hpi} \pm 5.81 \text{ hpi} \pm 4.45 \text{ hpi} \) (with 95% CI [3.95, 4.95]) \((R^2 = 0.917)\) (figure 3). Once again, these estimates are in good agreement with those obtained by fitting a binary logistic model to the data (Pearson goodness-of-fit test: \(\chi^2 = 86.084, 44 \text{ d.f.}, p < 0.001\)). As the model—equation (2.9)—fits the data well, it is safe to assume that there is no evidence for interactions between primary infection foci in terms of the onset of systemic infection. Moreover, using equation (2.10), we predicted for \(N = 100\), as illustrative case, an effective average time of systemic infection of \(t_{Sys}(t = 40)_{N=100} = 40.99 \text{ hpi}\), which agrees with the experimental data as \(P_{Sys}(t = 40)_{N=100} \approx 0.5\) in this plant pathosystem.

3. Discussion

The model presented here is useful because it helps to distill the two key processes governing the timing of the onset of systemic infection of plant viruses, namely, the cell-to-cell movement rate in relation to the density of host vasculature \((D)\) and the number of primary infection foci in the inoculated leaf \((N)\). As the movement rate and number of foci increase, the onset of systemic infection will tend to be quicker and its variability will be reduced. However, the model also predicts that for large numbers of primary infection foci, the onset of systemic infection is limited by the latency period \((T)\) rather than the cell-to-cell movement rate (figure 4). These model predictions have therefore implications for the evolution of plant viruses.

Whereas some plant viruses can be transmitted vertically, horizontal transmission depends on vectors [29] or contact between plants [30]. The intensity of between-host transmission
will probably vary greatly in the field, strongly depending on the density of both plants and vectors. Plant viruses typically exclude superinfection by conspecific or closely related viruses [31,32], a phenomenon that even extends to simultaneous coinoculation [9]. Therefore, a virus that causes full-blown systemic infection is likely to exclude its competitors, making it plausible that the onset of systemic infection is an important fitness component under some conditions. Therefore, when high levels of transmission result in a large number of primary infection foci, our model suggests the key characteristic that will be under selection is latency (i.e. viruses with a faster replication in the first infected cell will be selected). The rate of cell-to-cell spread may then still be important for colonization of systemically infected tissues, but it likely will not affect the onset of systemic infection (figure 4). On the other hand, single primary infection focus typically results in low levels of transmission and systemic infections [9]. In this situation, the onset of systemic infection may not be an important fitness component either, because low infections levels will limit co-infection and therefore minimize the importance of within-host competition. Consequently, our results suggest that intraspecific competition will impose selection for rapid cell-to-cell movement only for a limited range of primary infection foci values (1 \( N \), N \( \leq 30 \)). Moreover, the need to overcome host immune responses could, by itself, still drive selection for rapid cell-to-cell movement. Indeed, in some cases, the host plant can restrict the virus spread in the inoculated leaf (hypersensitive response), thus rapid cell-to-cell movement would help to surmount this barrier [33].

**Funding statement.** This work was supported by the grant no. BFU2012-30805 from Spain Ministerio de Economía y Competitividad (MINECO) to S.F.E. G.R. was supported by an EMBO long-term fellowship co-funded by Marie Curie actions (ALTF-1177-2011) and an AXA post-doctoral fellowship, and M.P.Z. by a Juan de la Cierva post-doctoral contract (JCI-2011-10379) from MINECO.