HYPE-IT: Cas9 and gRNA complex formation.

Valencia UPV iGEM Team

Complex Cas9:gRNA formation

Since Cas9 and gRNA constructions are delivered in the host, they are transcripted and translated by the cell's machinery. From their interaction, it is created the Cas9:gRNA complex, which will perform the knockout in the Testing System construction. The light signal registered in the luciferase assays will vary according to the number of knockouts performed by the complex. Therefore, there is a direct relation between the light signal measured and the eficiency in the production of the complex.
Studying this step, our aim was to analyze the system stability in function of CRISPR/Cas9 components under different conditions. Using Mass Action Kinetics, Quasi Steady State Analysis and Direct Liapunov Method, we determined conditions to reach the steady state conditions that provide the optimum amount of Cas9:gRNA complex. Firstly we analyze the interaction between Cas9 and gRNA, following with a study of agroinfiltration times in the second section.

Cas9 and gRNA interaction

Once Cas9 and gRNA constructions are delivered in the host, they will be produced by the cell's machinery. From their interaction, it is created the Cas9:gRNA complex, which will perform the knockout in the Testing System construction.
The dynamic of these species can be represented by the following reactions:

$$Cas9 + gRNA \underset{k_{-f}}{\overset{k_f}{\rightleftarrows}} (Cas9 + gRNA)$$

$$(Cas9 + gRNA) \xrightarrow{k_I} Cas9 : gRNA$$

$$Cas9 \xrightarrow{d_{Cas9}} \emptyset$$

$$gRNA \xrightarrow{d_{gRNA}} \emptyset$$

$$(Cas9 + gRNA) \xrightarrow{d_{(Cas9:gRNA)}} \emptyset$$

$$(Cas9 : gRNA) \xrightarrow{d_{(Cas9:gRNA)}} \emptyset$$

Cas9 and gRNA find each other in the plant nucleus. They bind each other leading to an Intermediate complex which is unstable, being this reaction governed by the rate k_f. This union will produce a conformational rearrangement in the Cas9 structure, represented by the constant k_I.
Elements from reactions above (Cas9, gRNA and the Intermediate complex) are usually expressed in terms of concentration ([Cas9], [gRNA], [complex]). We rename those concentrations using the variables $x_1 = [Cas9]$, $x_2 = [gRNA]$, $x_3 = [Cas9{+}gRNA]$ intermediate, and $x_4 = [Cas9 : gRNA]$ complex. Then, using mass action kinetics, we can build the system of Ordinary Differential Equations:

$$\begin{array}{rcl} [Cas9] & \rightarrow & x_1 \\ [gRNA] & \rightarrow & x_2 \\ [Intermediate] & \rightarrow & x_3 \\ [Cas9 : gRNA] & \rightarrow & x_4 \end{array}$$

$$\dot{x_1} = k_{Cas9} - k_f \cdot x_1 \cdot x_2 - d_{Cas9} \cdot x_1 \tag{1}$$

$$\dot{x}_2 = k_{gRNA} - k_f \cdot x_1 \cdot x_2 - d_{gRNA} \cdot x_2 \tag{2}$$

$$\dot{x}_3 = k_f \cdot x_1 \cdot x_2 - k_I \cdot x_3 - d_{Cas9 \cdot gRNA} \cdot x_3 \tag{3}$$

$$\dot{x}_4 = k_I \cdot x_3 - d_{Cas9 \cdot gRNA} \cdot x_4 \tag{4}$$

Steady state conditions of the system will be determined by the stability of the interaction between Cas9 and gRNA. Thus, equations (1) and (2) expressing the variation of these elements, must be analyzed in order to characterize the behavior of the system. They fit to the general expression of the following model:

$$\dot{x_1} = \alpha_1 - \alpha_2 \cdot x_1 \cdot x_2 - \alpha_3 \cdot x_1 \tag{5}$$

$$\dot{x_2} = \beta_1 - \beta_2 \cdot x_1 \cdot x_2 - \beta_3 \cdot x_2 \tag{6}$$

The equivalence between parameters used in the general expression (equations (6) and (6)) and their particularization in our model (equations (1) to (4)) is represented in Table 1, with their corresponding numerical values:

General model	**Particular value**	**Numeric value** (min^-1)
α_1	k_{Cas9}	0.000374737
α_2	k_f	0.00006
α_3	δ_{Cas9}	0.0000552
β_1	k_{gRNA}	0.0025284
β_2	k_f	0.00006
β_3	δ_{gRNA}	0.000252

Table 1: Parameters used in the model and values extracted from R.Moore et al.

This model represents the dynamics of two elements which interact affecting the steady-state condition of each other. The elements x_1 and x_2 are created with values α_1 and β_2. Constants α_1 and β_1 reflect a continuous input of x_1 and x_2.
In the biological context this is implies that the model is accounting a single cell. Initially, Cas9 and gRNA constructions get to the cell by viral infection. Thus, if the subject of study was a cell population, the flux of Cas9 and gRNA would vary according to the spread of the viral infection among cells.
However, from an individual perspective, each infected cell produces constitutively Cas9 and gRNA. Thus, production of both elements in a single cell, fits the typical behavior of constitutive production, i.e. reaches the steady-state after certain time. This permanent production is equivalent to the constant input of Cas9 and gRNA in the cell, reflected in terms α_1 and β_1, respectively.
Terms $\alpha_2 \cdot x_1 \cdot x_2$ and $\beta_2 \cdot x_1 \cdot x_2$ describe the exchange of x_1 and x_2 when they interact. Finally, $\alpha_3 \cdot x_1$ and $\beta_3 \cdot x_2$ represent the disappearance of x_1 and x_2 from the system.

Two solutions $(\tilde{x}_{1_i}, \tilde{x}_{2_i})$ can be provided to this system, with $i = 1, 2$. The equilibrium values will be obtained when (1) is equal to zero, i.e.:

$$(1) \to 0$$

$$\tilde{x}_{1_i} = \frac{\alpha_1}{\alpha_2 \cdot \tilde{x}_{2_i} + \alpha_3} \tag{7}$$

Considerig now that $\tilde{x}_2$ reaches its steady state, we get:

$$(2) \to 0$$

Assuming that both x_1 and x_2 are in their equilibrium points, then the expression for x_1 at equilibrium ($\tilde{x}_{1_i}$ from eq.(7)) can be pulggin in the previous (2), getting:

$$0 = \beta_1 - \beta_2 \cdot \tilde{x}_{1_i} \cdot \tilde{x}_{2_i} - \beta_3 \cdot \tilde{x}_{2_i}$$

$$0 = \beta_1 - \beta_2 \cdot \frac{\alpha_1}{\alpha_2 \cdot \tilde{x}_{2_i} + \alpha_3} \cdot \tilde{x}_{2_i} - \beta_3 \cdot \tilde{x}_{2_i}$$

$$\alpha_2\beta_1\tilde{x}_{2_i} + \alpha_3\beta_1 = \alpha_1\beta_2 \cdot \tilde{x}_{2_i} + \alpha_2\beta_3\tilde{x^2}_{2_i} + \alpha_3\beta_3\tilde{x}_{2_i}$$

$$0 = \alpha_2\beta_3 \cdot \tilde{x}_{2_i} + (\alpha_1\beta_2 - \alpha_2\beta_1 + \alpha_3\beta_3) \cdot \tilde{x}_{2_i} - \alpha_3\beta_1$$

$$\tilde{x}_{2_i} = \frac{-\alpha_1\beta_2 + \alpha_2\beta_1 - \alpha_3\beta_3 \pm \sqrt{(\alpha_1\beta_2 - \alpha_2\beta_1 + \alpha_3\beta_3)^2 + 4 \cdot \alpha_2\beta_3\alpha_3\beta_1}}{2 \cdot \alpha_2\beta_3} \tag{8}$$

Stability conditions are necessary to determine if a solution is a critical point. Using the Direct Liapunov Method, we can get two expressions which can be used to know about the stability of the system given a solution. Assuming that a system of differential equations has a critical point $(\tilde{x}_{1_i}, \tilde{x}_{2_i})$, we can say that this point of the function $E(x_1, x_2)$ will be asymptotically stable when:

1. $E(x_1, x_2)$ is a positive-definite function.
2. Its derivative $E'(x_1, x_2)$ is negative-semidefinite.

These conditions are reflected in expressions (9) and (10):

$$\Delta = \frac{\alpha_3\beta_1}{\tilde{x}_{2_i} + \alpha_3\beta_3\tilde{x}_{2_i}} \rightarrow \Delta > 0 \tag{9}$$

The second expression of the system stability is:

$$\sigma = \frac{\partial}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial}{\partial x_2}\frac{dx_2}{dt} = -\alpha_2 \cdot \tilde{x}_2 - \alpha_3 - \beta_2 \cdot \tilde{x_1} - \beta_3 \rightarrow \sigma < 0 \tag{10}$$

From equations (9) and (10), we can assume that there will be one of the equilibrium solutions (x_{1_1}, x_{2_1}) that will be always stable, whereas the second one (x_{1_2}, x_{2_2}) is an unstable one. As α_3 is the degradation rate of Cas9, we can assume that $\alpha_3 \rightarrow 0$. Therefore, the model is simplified, resulting in the equilibrium positions and stability conditions:

$$\tilde{x}_{1_i} = \frac{\alpha_1}{\alpha_2 \cdot \tilde{x}_{2_i}} \tag{11}$$

$$\tilde{x}_{2_i} = \frac{C_1\beta_1 - \alpha_1}{\beta_3 C_1} \tag{12}$$

$$\sigma = -\alpha_2 \cdot \tilde{x}_{2_i} - \beta_2 \cdot \tilde{x}_{1_i} - \beta_3 \tag{13}$$

$$\Delta = \alpha_2\beta_3\tilde{x}_{2_i} \tag{14}$$

Under this conditions, both solutions previously considered become one solution which will vary according to the value of C_1. If conditions $\sigma < 0$ and $\Delta > 0$ are provided, then the constant C_1 will be big enough to let the system achieve a local stability, being:

$$C_1 = \frac{\alpha_2}{\beta_2} \tag{15}$$

Replacing each model parameter by its value in our particular case, we get the following results: $[\tilde{Cas9}]$,$[\tilde{gRNA}]$ and restrictions σ and Δ, are:

$$C_1 = \frac{\alpha_2}{\beta_2} = \frac{k_f}{k_f} = 1 \tag{16}$$

Using (16) in (12):

$$[g\tilde{RN}A] = \frac{k_{gRNA} - k_{Cas9}}{\delta_{gRNA}} \approx 8.5 \tag{17}$$

Thus, (11) is:

$$[\tilde{Cas9}] = \frac{k_{Cas9}}{k_f \cdot [g\tilde{RN}A]} \approx 0.7 \tag{18}$$

And stability conditions of the system (13) and (14) are achieved:

$$\sigma = -k_f \cdot [g\tilde{RN}A] - k_f \cdot [\tilde{Cas9}] - \delta_{gRNA} \approx -0.0008 \rightarrow \sigma < 0 \tag{19}$$

$$\Delta = k_f \cdot \delta_{gRNA}[g\tilde{RN}A] \approx 1.3 \cdot 10^{-}7 \rightarrow \Delta > 0 \tag{20}$$

Therefore, we can assume that our system is locally stable around the equilibrium point $(\tilde{x}_{1_i}, \tilde{x}_{2_i})$. Moreover, steady-state values obtained analytically are mostly the same as equilibrium values obtained simulating in Matlab:

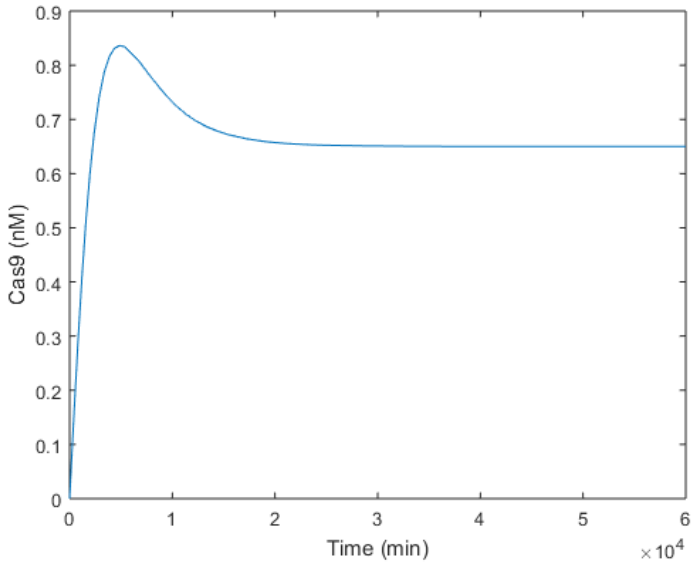

Figure 1: Simulations of [Cas9] evolution using parameters from Table 1.

Simulations performed let us check the equilibrium values obtained analytically. The limiting reagent in this case is the Cas9. The Cas9 endonuclease is produced reaching its maximum at the time of peak $t_p = 4877$ minutes, i.e. 4 days approx. From day 4, there is enough of both compounds in order to form the Cas9:gRNA complex, explaining the decay in Cas9 after 4 days from infiltration. Four days later, at $t \approx 10$ days (15000 minutes), Cas9 reaches its steady-state near the 0.7 nM predicted analytically. This behavior is similar to a Second Order System response after a step input. It reaches the equilibrium value with a time slope of $t_s = 2380$ minutes, i.e. 1 day and 15 hours.

Regarding the gRNA, it is produced without any oscillation. The maximum production (8.6 nM)is reached between around the 13th day, and the 63% of this value at $\tau \approx 3400$ minutes. As gRNA has a dynamic considerably faster than Cas9, it keeps growing until the endonuclease is produced. As soon as this happens, Cas9 and gRNA are used to form the Cas9:gRNA complex. Thus, the peak in Figure 1 makes reference to this complex formation. In Figure 2 this consumption is not appreciated because there is plenty more gRNA disposable than gRNA used to form the complex with Cas9 produced.

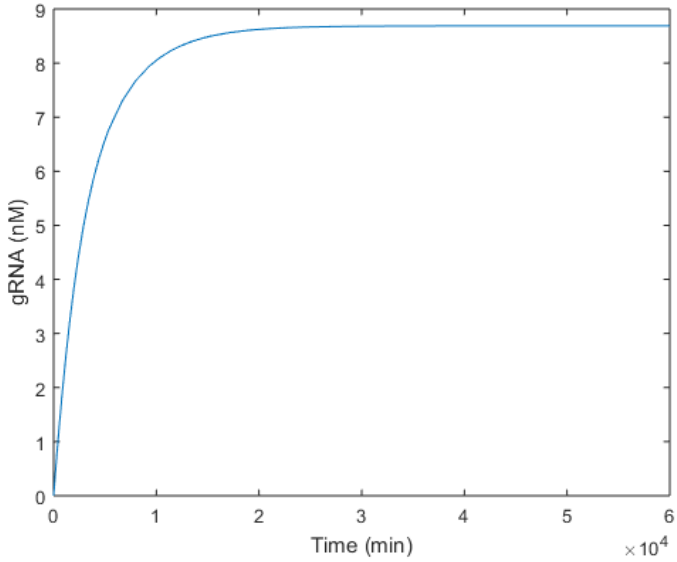

Figure 2: Simulations of [gRNA] evolution using parameters from Table 1.

However, using one gene copy number is not a current laboratory condition. As production of Cas9 and gRNA is directly related to the number of gene copies, higher amounts of them are a more realistic subject of study. Since terms k_gRNA and k_Cas9 are in function of the number of copies, changing this parameter the equilibrium values will be affected. The following graphics represent different results varying the gene copy number only for Cas9, only for gRNA and finally for both of them.

In the following simulations the gene copy number of the gRNA was increased gradually from 1 to 45. Results obtained each 5 copies were plotted to analyze the stability of Cas9 when the gRNA is produced faster and with higher quantities.

From simulations above, it is clear that Cas9 efficacy is highly sensible to the presence of gRNA in the nucleus. The biggest difference among concentrations of Cas9 in the graphic above, is produce between the result with 1 copy and the result with 5. This means that even though Cas9 production is not increased, the endonuclease will be able to find the gRNA and form complexes with it.

There is an indirect relation between the maximum Cas9 presence, the time of peak t_p at which this maximum is reached, and the amount of gRNA. It seems quite logical, as the more gRNA present, the more likely it will be that all Cas9 produced is consumed to form the complex with those gRNAs.

From the picture below, it can be confirmed that an increment in the Cas9 gene copy number, affects homogeneously to the quantity of gRNA. Qualitatively is has the same behavior as in the scenario with one copy. The only difference is reflected in a gain of the steady-state gRNA value, which increases in 5-fold from the previous simulation with 5 gene copies less.

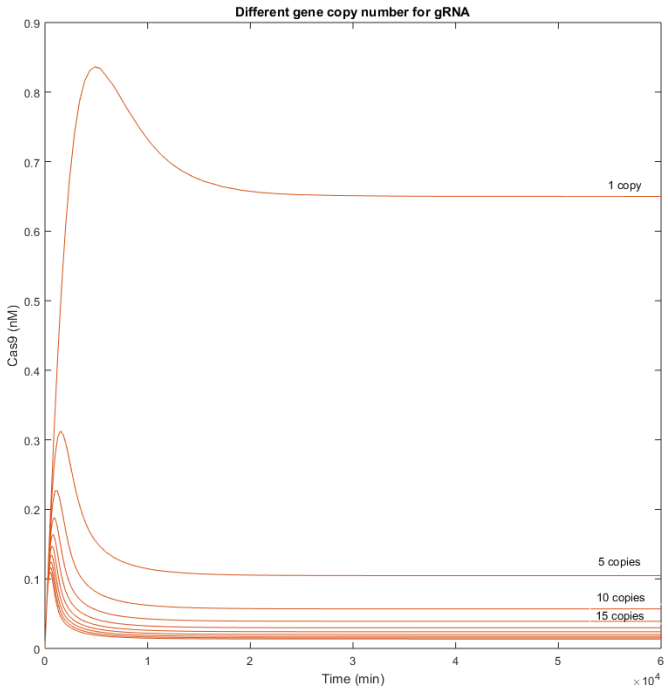

Figure 3: Simulations of [Cas9] with different gRNA gene copy numbers.

When the gene copy number for Cas9 is increased (Figure 5), the production of the endonuclease behaves similarly to the gRNA (Figure 4). However, the slope is considerably less. This difference in the dynamics was expected, as it is known that RNA reactions are several orders faster than protein ones.

There is no decrease in the Cas9 evolution which let us know when does the interaction with the gRNA begin. However, the beginning of this reaction is well appreciated in the graphic with the gRNA evolution.

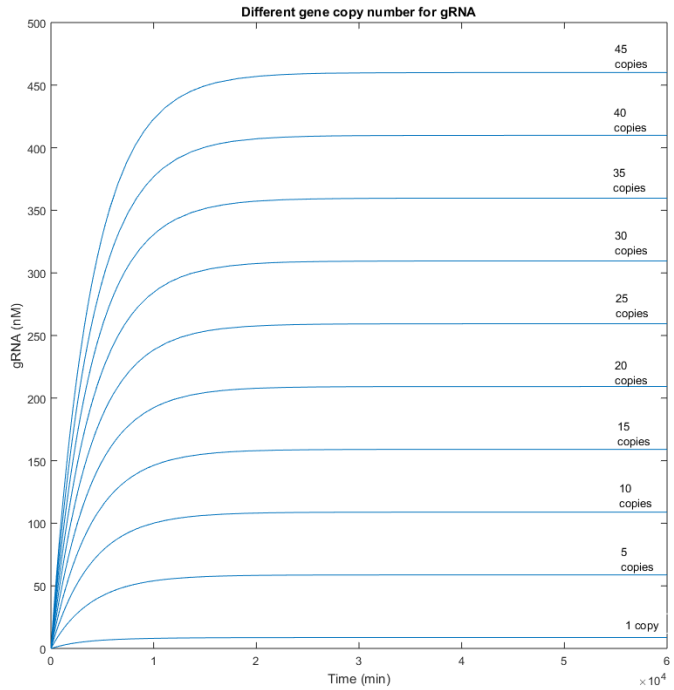

Figure 4: Simulations of [gRNA] with different gRNA gene copy numbers.

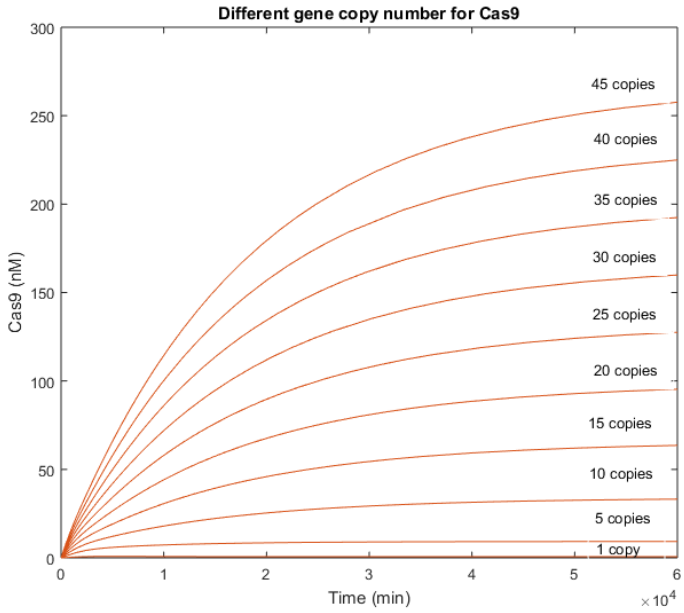

Figure 5: Simulations of [Cas9] with different Cas9 gene copy numbers.

As it is mentioned above, the initial slope of gRNA is considerably faster than Cas9s, as gRNA only means transcription but Cas9 must be translated as well, delaying its growing. This behavior is comparable to Cas9 variations when the gRNA copy number was increased. However, the difference between simulations with one copy number and 5 copies, are less remarkable in this case, reaching the half of the one-copy steady state value in this case. Again, this difference is smaller due to the fast production of gRNA.

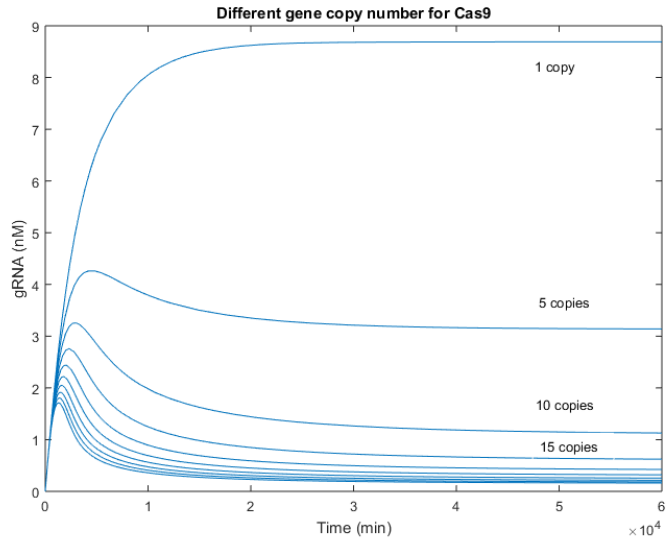

Figure 6: Simulations of [gRNA] with different Cas9 gene copy numbers.

The time of peak changes as well among simulations with different number of Cas9 gene copies, being the maximum at approx. 5000 minutes, i.e. 3 and a half days. The minimum time peak is produced near 2 days after production.

Several simulations varying simultaneously both gene copy numbers were performed. This time was calculated the amount of intermediate complex produced, which is shown in the following graphic.

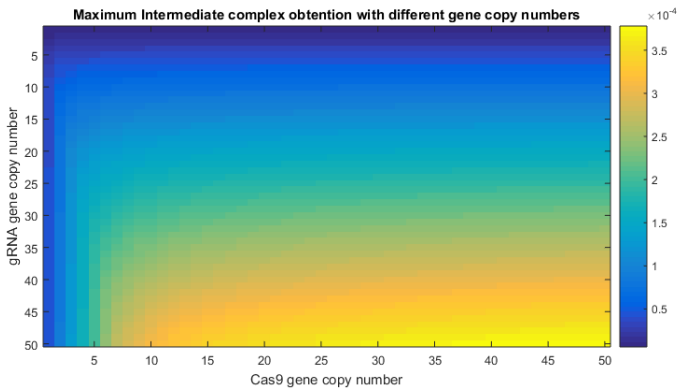

Figure 7: Simulations of [Cas9:gRNA] intermediate complex obtention with different number of copies for Cas9 and gRNA

From these results (Figure 7) it can be inferred that the amount of complex being produced is very sensible to the amount of Cas9 in the nucleus. This is reflected in the higher values of complex when Cas9 gene copy number has reached only the value of 5. Necessary amounts of gRNA to increase the intermediate presence, are higher than those needed for Cas9. This means that gRNA will be the limiting reagent of this process. Though this may refutes the fact that gRNA is produced faster by cells machinery, its fast degradation explains why this element is limiting the Cas9:gRNA complex ready to perform the knockout.

Therefore, less amount of Cas9 will not be as determining as less amount of gRNA. It will be critical to provide the system with the necessary gRNA in order to let enough complex be formed. Otherwise, less knockouts would be performed in the Testing System and light signal recorded during Luciferase assays would be less significant.

Time intervals between Cas9 and gRNA agroinfiltration

Until this point, all simulations have been subjected to null initial conditions. Experimentally this is equivalent to assume that both Cas9 and gRNA are infiltrated at the same time in the plant. However, the difference between Cas9 and gRNA dynamics is known and it has been affecting as well previous simulations. Next scenarios will assume that Cas9 has been introduced few days earlier than gRNA, letting time enough to be produced until reaching a certain concentration.

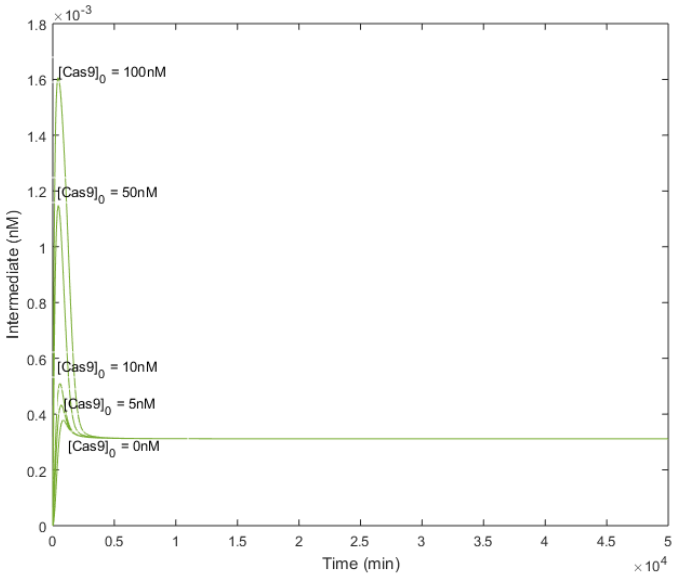

Figure 8: Simulations of [Cas9:gRNA] intermediate complex obtention with different initial conditions relying [Cas9] in the host nucleus.

As expected, infecting plants previously with Cas9 to let it be produced, improves considerably the amount of complex obtained. Morever, the t_p does not change drastically in comparison to null [Cas9] initial conditions, being around 1000 minutes. The steady state is achieved as well without compromising the stability of the system. Therefore, greater amounts of the complex will be produced in less time, or in other words, the efficiency of the system will be increased if the Cas9 is already in the cell when the gRNA is infiltrated.

Furthermore, with more presence of Cas9, the limiting factor of the gRNA is reduced, as it can be appreciated in graphics below:

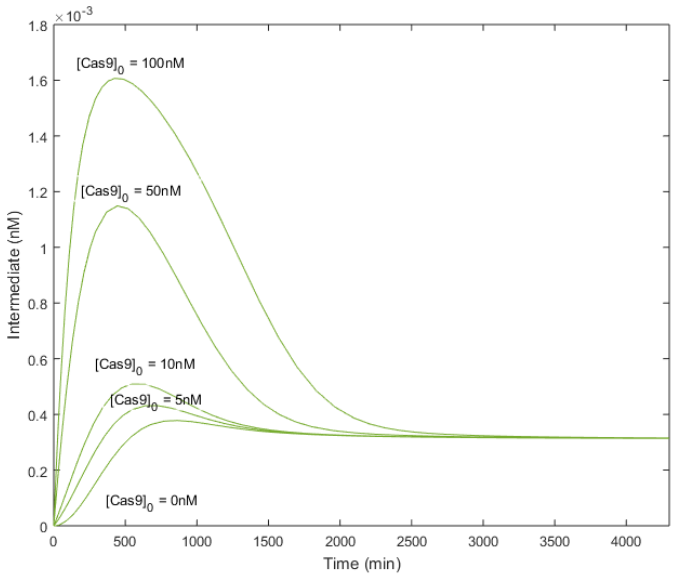

Figure 9: Simulations of [Cas9:gRNA] intermediate complex obtention with different initial conditions relying [Cas9] in the host nucleus.

In graphics below it can be appreciated how the system becomes more robust as more Cas9 is available in the cell nucleus. The advantage of reducing the importance of the gRNA presence, is that the amount of light signal will not be perturbed by a low gRNA concentration. Thus, the Testing System output will vary according to the efficiency of the gRNA to find the target, and not relying on other external factors such as the cell's efficiency in the gRNA transcription.

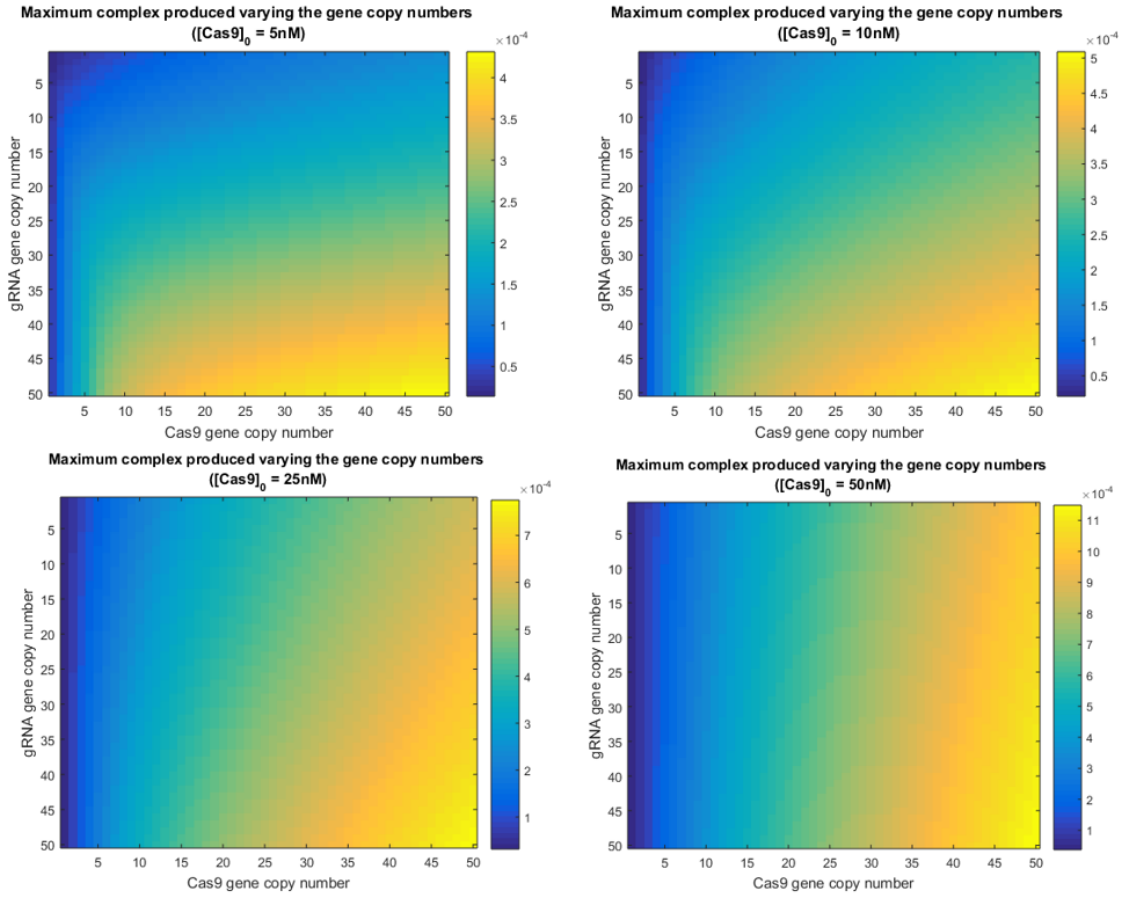

Figure 10: Simulations of [Cas9:gRNA] complex obtention with different initial concentrations of Cas9.