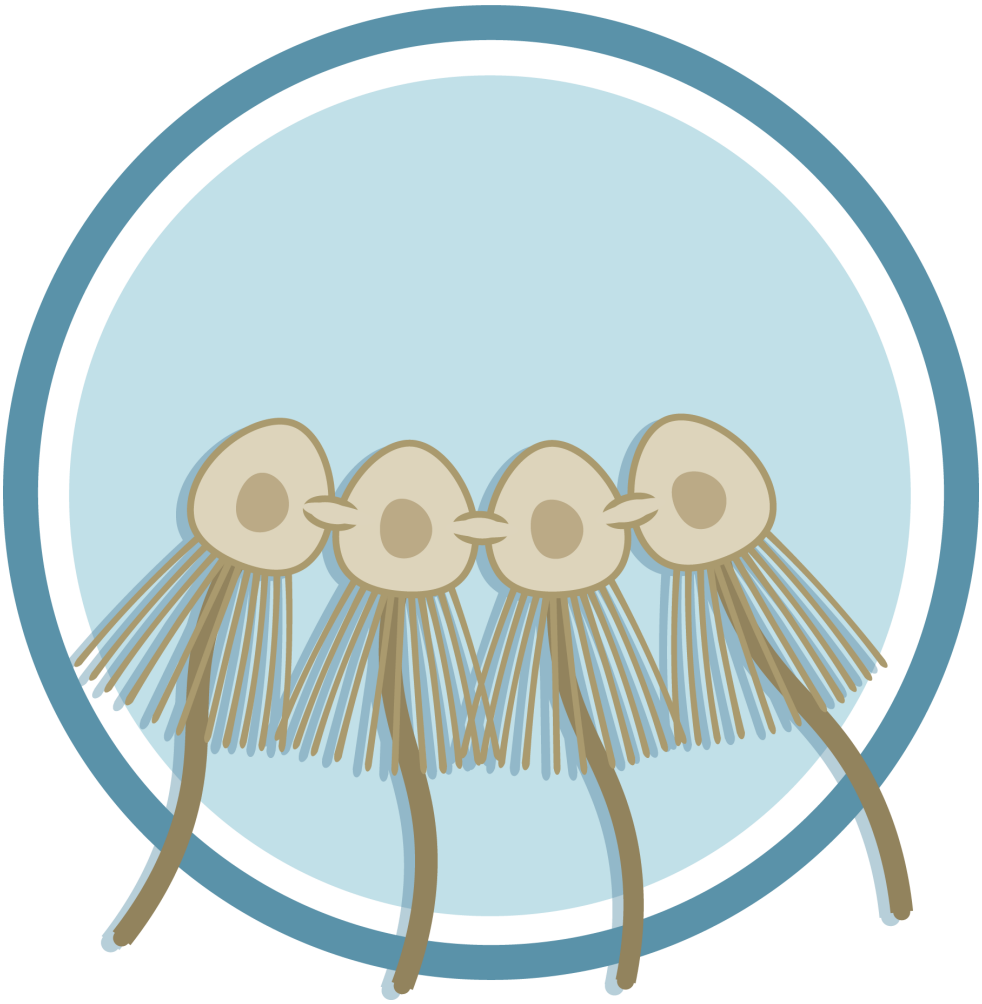
Utrecht University
UTRECHT UNIVERSITY

MAJOR RESEARCH PROJECT

Stochastic Modelling of *Salpingoeca rosetta* Chain Colonies

Anna Delhaas · Student ID: 6937438

Supervisor: Dr. Florian Berger
Department of Cell Biology, Neurobiology and Biophysics
Theoretical Biophysics Group
MSc Bioinformatics and Biocomplexity · October, 2024



Abstract

The transition to complex multicellularity from unicellular organisms was a major step in the life history of animals, how this transition happened is however, not fully known. To better understand the evolution of multicellularity, researchers study organisms that exhibit simple multicellularity, such as organisms that can form colonies. A defining feature of these colonies is their size. Although colony size has been modeled in various species, no model has yet focused specifically on the colonies of choanoflagellates, the closest living relatives of animals. To gain deeper insight into the origins of multicellularity in animals, this thesis presents a discrete stochastic model to simulate chain colonies of the choanoflagellate *Salpingoeca rosetta* (*S. rosetta*). The model focuses on the interplay of division and breaking of cytokinetic bridges and how their underlying waiting time distributions affect the distribution of colony sizes. These events are modeled using various probability distributions and, through fitting to experimental colony size data, we determined that a gamma distribution best represents the waiting times for division, while those for bridge breaks follow an exponential or gamma distribution. Additionally, the model revealed that *S. rosetta* chain colonies reach a steady state in the average number of cells per colony. We analyzed how the rates of cell division and bridge breakage influence this steady state, finding that the division rate is inversely related to colony size, whereas the bridge break rate follows a logarithmic relationship. Furthermore, a mathematical description of these steady states was found. These results offer new insights into the colonies of choanoflagellates, the closest living relatives of animals, and enhances our understanding of the origin of animal multicellularity.

Layman Summary

Choanoflagellates are single-celled eukaryotes found in aquatic environments. Even though they are unicellular, several species of choanoflagellates can form multicellular colonies. What makes these microorganisms unique is that they are the closest living relatives of animals, making them an interesting model organism to study various aspects of the origin of animals. In this thesis, *Salpingoeca rosetta (S. rosetta)*, a choanoflagellate species, is used to study the properties of simple multicellularity. *S. rosetta* cells are able to form chain colonies: linear chains of cells connected by intercellular bridges. One important aspect of these colonies is their size, i.e. how many cells are present per colony. This thesis presents a model that describes how *S. rosetta* chain colonies grow and split, revealing what factors determine colony sizes in simple multicellular colonies, leading to a better understanding of the development of more complex multicellular organisms. The model is built on two key processes that determine the colony sizes in chain colonies: cell division, which increases colony size, and bridge breakage, where connections between cells break, leading to new smaller colonies. The outputs of this model include the colony sizes after specific time points and the average number of cells per chain per time point. By doing various analyses with this model, we came to the following conclusions: The cell division in *S. rosetta* cells follows a specific pattern best described by a gamma distribution, which effectively means that division is age-dependent. We were not able to come to definitive conclusions on the lifetimes of bridges, but we were able to conclude that this process involves some randomness and is not deterministic. Furthermore, we found that after some time, the colonies in the simulation reach a steady state, where they stabilize in their size. The mean colony size at which they stabilize is dependent on the bridge break and division rate. These findings highlight how a simple model of two events can make accurate predictions on colony sizes. Our results contribute to the knowledge on the mechanics of colonies of choanoflagellates, the sister group to animals, which could in turn contribute to the knowledge on the transition from single-celled organisms to multicellular ones, which was a crucial step in the life history of animals.

Keywords: Choanoflagellate, *Salpingoeca rosetta*, Stochastic modelling, Multicellularity, Interdivision times, Cytokinetic bridges, Chain colonies, Steady state.

Contents

1 Introduction

The transition from unicellular to multicellular organisms was a crucial step in the evolution of animals [1–3]. It made the division of labor and increased specialization among cells possible, allowing for the development of more organized structures such as tissues and organs that form the basis of modern animals. To achieve the transition to complex multicellularity, several major steps were necessary, including cell differentiation, cell adhesion, and cell communication [4–7]. These features merged from organisms that portrayed simple multicellular features, like the ability to form colonies and basic cell differentiation. The driving forces behind these changes were likely evolutionary benefits such as predation avoidance, stress resistance, and increased motility [2, 8, 9].

In essence, multicellularity refers to organisms composed of multiple cells, making colony size one of the defining characteristics of multicellular systems. Studying colony sizes in organisms that exhibit simple multicellularity could therefore provide insight into one of the most fundamental aspects of multicellularity, thereby improving our understanding on the origin of multicellularity [9, 10].

The distribution of colony sizes and their underlying mechanisms have been studied within a variety of organisms including, phytoplankton, algae, and yeast. Yet, no studies have examined this in model organisms more closely related to animals [10–12]. To address this gap, it would be valuable to investigate the first animal. However, animals evolved 600 million years ago, and there are no fossil records of these organisms [13]. Therefore, choanoflagellates, as shown in Figure 1A, are frequently used as model organisms for animal multicellularity, as they are the closest living relative to animals [14–17].

Choanoflagellates are unicellular microbes that live in aquatic environments [14, 18]. These eukaryotes consist of a cell body, a collar made up of microvilli, and a flagellum, allowing them to swim and create currents to trap and consume bacteria [19, 20]. Due to their close evolutionary relationship with animals, choanoflagellates are often used as model organisms for animal multicellularity and development [14, 21, 22]. Additionally, many choanoflagellate species are able to form multicellular colonies, further establishing them as a valuable model for simple multicellularity [23]. Choanoflagellates have already proven to be useful in studying the origin of various animal behaviors and mechanisms, for example of collective contractility and generating action potentials [6, 24, 25]. Moreover, multiple genes have been identified in choanoflagellates, which were thought to be unique to animals, having led to a better understanding of the genetics underlying the origin of animals [5, 6, 22].

One of the most well-studied species of choanoflagellates is *Salpingoeca rosetta* [21]. *S. rosetta* makes for an interesting model organism due to its complex life cycle consisting of five main states: three unicellular states (fast swimmers, slow swimmers, and thecate cells), and two colonial states (chain colonies and rosette colonies) [14]. Additionally, they are able to sexually reproduce [6, 26]. *S. rosetta* can switch between these states in response to environmental factors, like food scarcity, with slow swimmers being able to transition to both colonial states. However, in order to form rosette colonies, bacterial factor RIF must be present in the medium [14, 27]. *S. rosetta* colonies form through serial cell division, which is one of two ways multicellular colonies can form, the other being through cell aggregation [14, 22, 28–30]. More specifically, in chain colonies, choanoflagellates divide and form a flat chain, whereas in rosette colonies, cells change their orientation to form a spherical colony.

In *S. rosetta* colonies, the cells are connected through cytoplasmic bridges and the extracellular matrix (ECM) [14]. Cytokinetic bridges are a result of incomplete cell division [28, 31]. Here, the final step of cell division, called abscission, which should cut the bridges, is not fully carried out, leaving the daughter cells connected [28]. These bridges can stabilize, allowing the cells to be connected for hours to days [31]. Cytokenetic bridges are conserved in many clades of life, where they mostly function in intercellular communication and providing structure [28, 32]. The internal structure of these bridges varies per clade [28]. In choanoflagellates, these bridges consist of two dense plates connecting the membranes of the cells, with one bridge connecting each pair of cells [14]. It is thought that cytokinetic bridges have played a role in the origin of multicellularity, where they likely contributed to providing stability and structure in simple colonies [28]. Therefore, studying cytokinetic bridges in choanoflagellate colonies has the potential to provide more insight into their role in the evolution of multicellularity and their dynamics in simple multicellular organisms.

To summarize, studying the choanoflagellate S. rosetta has proven to be useful for investigating the origin of multicellularity in animals [23]; Their ability to form multicellular colonies connected by cytokinetic bridges, together with the fact that choanoflagellates are the closest living relatives to animals, makes them a valuable model organism. In this thesis, we developed a simple discrete stochastic model to better understand the dynamics of S. rosetta chain colonies. Our model is driven by two events: cell division and bridge breakage. We analyzed experimental data, fitted our model to this data to obtain parameters, and assessed the model's performance with different underlying waiting time distributions for these events. Furthermore, we identified steady states in the average number of choanoflagellates per colony and derived a mathematical description of these states. This work aims to improve the understanding of the mechanisms underlying colony sizes in simple multicellular systems and specifically in choanoflagellates to further enhance our knowledge on the origin of animal multicellularity.

2 Materials and Methods

2.1 Model

For this thesis, we developed a discrete stochastic model using Python to study the effects of cell division and intercellular bridge breaks on *S. rosetta* chain colony formation. In this model, the chain colonies are represented as a dictionary: a data

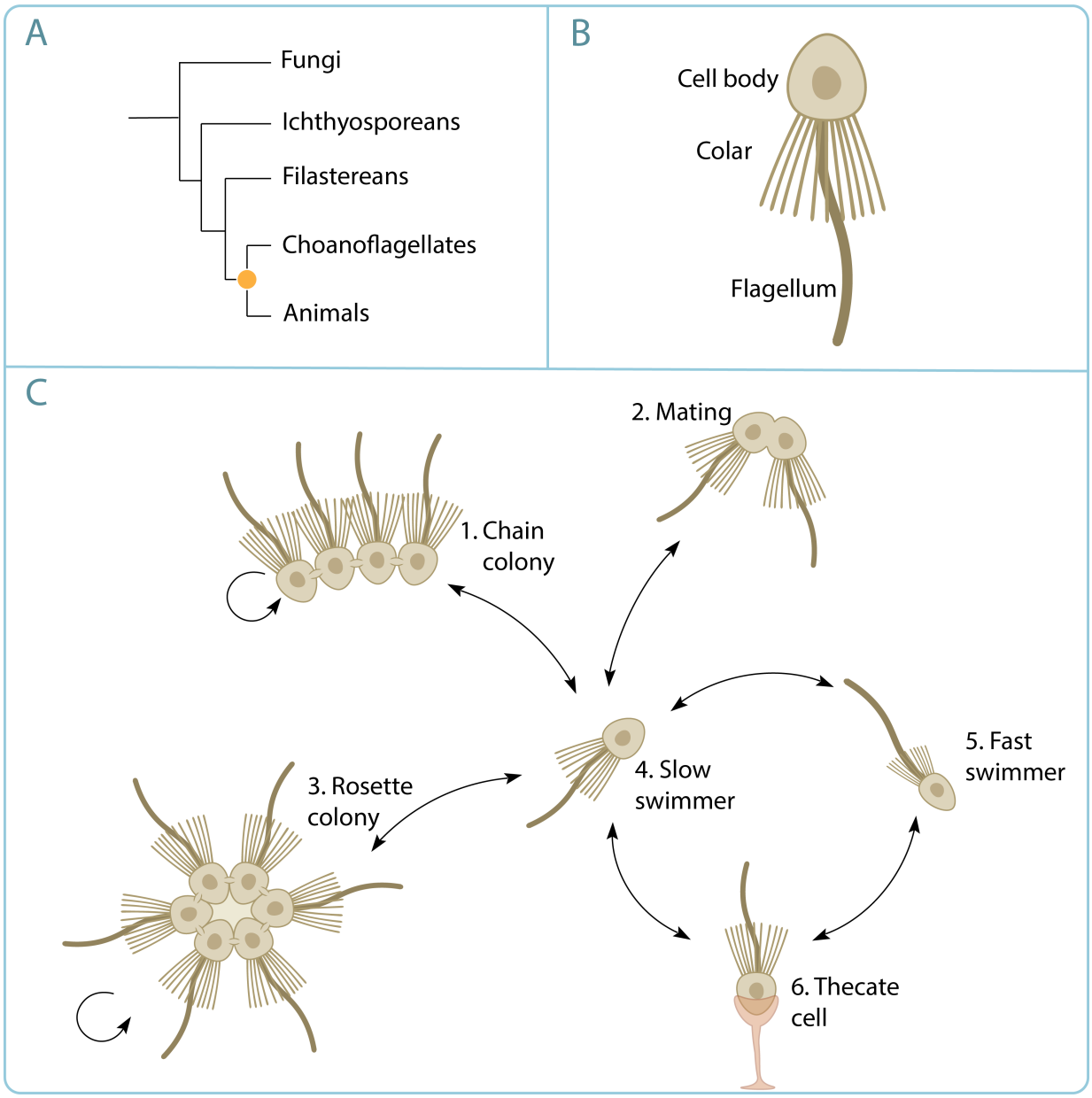


Figure 1: A) Phylogenic tree showing the evolutionary relationship between animals, choanoflagellates, filastereans, ichthyosporeans, and fungi. The orange dot represent the last common ancestor between animals and choanoflagellates highlighting their close evolutionary relationship. B) Simple representation of a choanoflagellate with its key structural components: the cell body, the collar made up of microvilli, and the flagellum. C) The life cycle of choanoflagellate S. rosetta, showing the six main states: (1) Chain-colony, (2) Mating, (3) Rosette-colony, (4) Slow swimmers, (5) Fast swimmers, and (6) Thecate cells, illustrating the possible transitions between the states [1].

type that stores data as keys and values. Here the keys correspond to the chains, with values representing the choanoflagellates within each chain. The cytokinetic bridges between the cells are not explicitly represented within this dictionary, but it is implied that each choanoflagellate in a chain has a bridge to its left as a result of a division event, except for the first one in the chain. This allows the bridges to be identified within the dictionary and handled accordingly. The dictionary is updated throughout the simulation in response to division and bridge break events. Our model is continuous-time event-driven, meaning that the model directly goes to the time where the next event happens, instead of going through every point in time to check if an event occurs. The timing of each event is determined by a list of waiting times, in which each entry specifies the time until an event occurs for each element in the chain dictionary.

If the object with the lowest waiting time in the list is a choanoflagellate, division occurs. During this event, the scheduled cell is replaced by two new choanoflagellates in the chain dictionary, along with a new bridge connecting them. For each

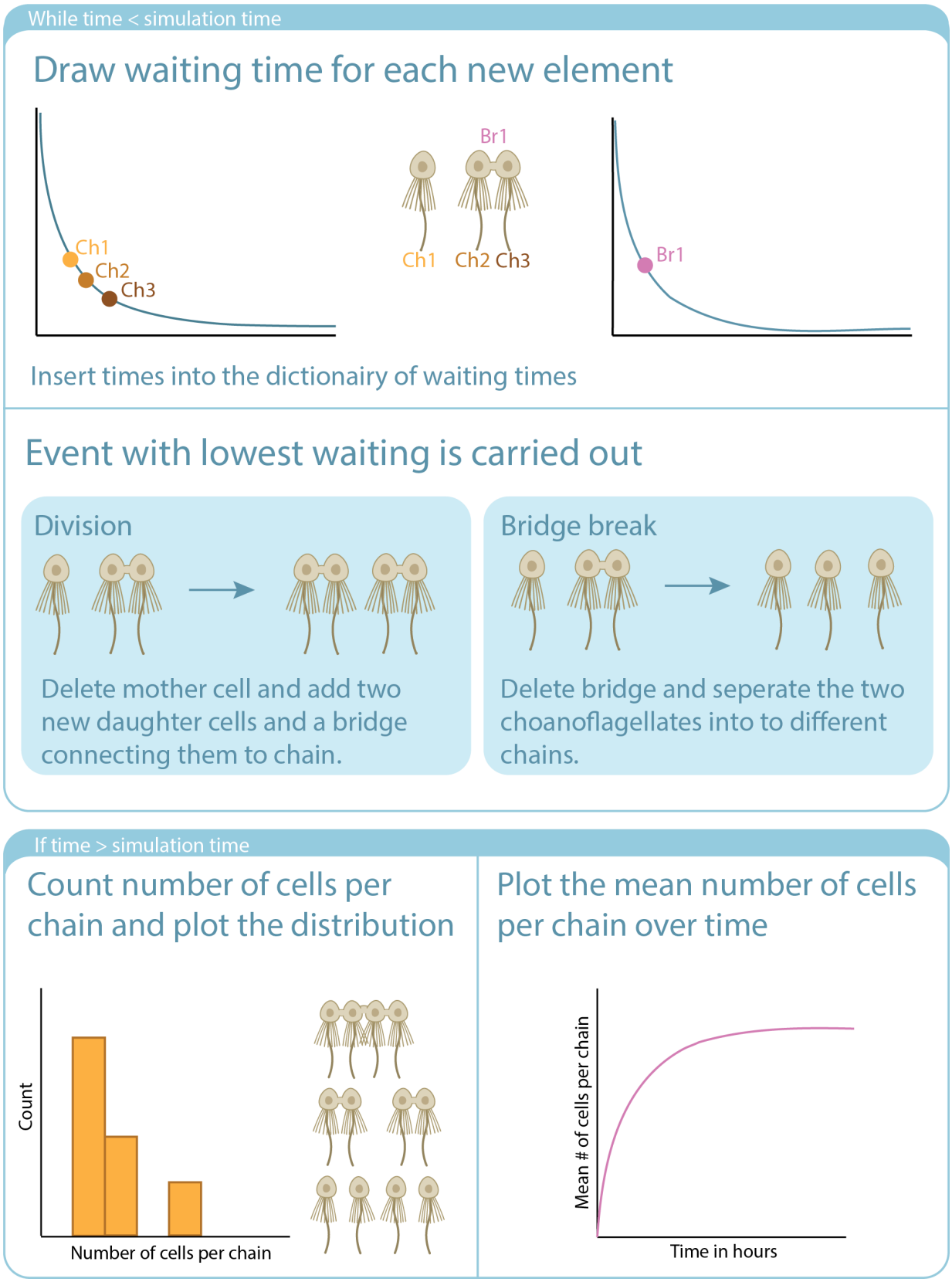


Figure 2: Graphical illustration of the steps of our simulation model for choanoflagellate chain colonies, consisting of two parts: before the time exceeds the simulation time and after the time exceeds the simulation time. The first part consists of drawing waiting times and carrying out events, while the second part consists of analyzing and plotting data obtained from the simulation.

new element, a waiting time is drawn and added to the waiting list.

If the object with the lowest waiting time is a bridge, then a bridge break event occurs. During this event, the chain containing the bridge splits into two separate chains at the choanoflagellate associated with the bridge. The model produces two outputs: the distribution of colony sizes at the end of the simulation, which can be compared to experimental distributions, and a plot of the mean number of choanoflagellates per chain colony over time, used to analyze steady states.

A graphical overview of the model and its steps are displayed in Figure 2.

2.1.1 Waiting times

As previously mentioned, a waiting time must be drawn for each new element at the moment it originates. This waiting time specifies the duration until the element either breaks (in case of a bridge) or divides (in case it is a cell). The final distribution is directly influenced by the timing of the events; therefore, the way these waiting times are drawn is very important. There are various methods for modeling the waiting times between events. Given the limited knowledge of the dynamics of bridge break events, various models for stochastic as well as deterministic processes were explored. For division, the best model was determined by comparing different models to available experimental cell cycle data.

To model waiting times stochastically, we draw values from a probability density function (PDF). A PDF maps a range of values to its associated probability, such that the area under the curve of the density function is one [33, 34]. There are several widely used PDFs, each with distinct properties. For our models, we incorporated exponential and gamma distributions.

Exponential distributions An exponential distribution is commonly used because of its memoryless property, meaning that the occurrence of the next event is independent of when the previous event occurred, i.e. the probability of an event taking place in the next interval does not vary through time [34]. This property is reflected in the PDF of the exponential distribution which has a long tail and is non-zero for all positive values (Fig 3). The distribution is characterized by a single parameter, the rate λ, which corresponds to the frequency that an event happens [34]. The inverse of this rate parameter is the scale parameter ϕ, which corresponds to the mean of the exponential distribution. Both λ and ϕ are real and positive numbers. The probability density function is defined by:

$$\lambda * \exp(-\lambda t), \quad t > 0 \tag{1}$$

or with the scale parameter as

$$\frac{1}{\phi} * \exp(-\frac{t}{\phi}). \tag{2}$$

with t representing time [34]. In Figure 3A, the PDF of the exponential distribution is plotted for different values of λ. The exponential distribution's simplicity, with only one parameter, makes it straightforward for sampling waiting times.

Gamma distributions The gamma distribution is a widely used distribution due to its versatility. This distribution can capture the chi-squared distribution and Erlang distribution as well as the exponential distribution as special cases [34]. The gamma distribution is defined by two parameters: the shape parameter α and the rate parameter β, both of which are real and positive numbers [34]. α defines the shape of PDF, with special cases for $\alpha = 1$ where the PDF is exponential, and for high values, it approximates the normal distribution [35]. The mean of the distribution is described by $\alpha * \beta$. The PDF function of the gamma distribution is given by :

$$f(t;\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t}, \quad t > 0 \tag{3}$$

[34] With $\Gamma(\alpha)$ as the gamma function, which is defined as

$$\Gamma(\alpha) = \int_0^{\infty} t^{(\alpha-1)} \exp(-t)\, dt \tag{4}$$

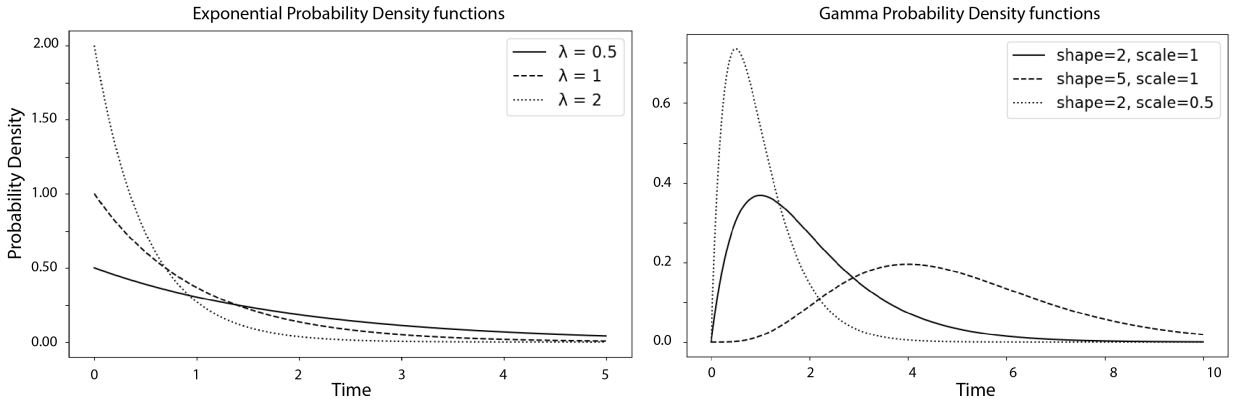


Figure 3: Exponential and Gamma Probability Density Functions (PDFs). The left graph shows exponential PDFs for different rates (λ = 0.5, 1, and 2). The right graph displays Gamma PDFs for various combinations of shape and scale parameters (α = 2 β=1, α = 5 β=1, and α = 2 β=0.5).

[36]. In Figure 3, several PDFs of the gamma distribution are shown for different values of α and β. The gamma distribution is more versatile and flexible than the exponential distribution due to its two parameters, making it another interesting distribution to sample waiting times from.

Deterministic The final method considered is to draw the waiting times deterministically. In this approach, the time between the origin of each element and its related event is a predetermined value.

2.1.2 Inputs

The input for the model depends on what method is used to draw the waiting times between events. However, there are parameters that are used in for each method:

- Initial population size: Initial chain colonies and single choanoflagellates in the population that the simulation will start with.
- Mean time until division: Average time until a choanoflagellate divides from the time of its formation.
- Mean time until bridge break: Average time until a bridge breaks from the time of its formation.
- Simulation time: The total duration (in hours) for which the simulation will model the system's dynamics.

All parameters can be adjusted separately, allowing to see the effect of each parameters on the outputs of the model.

2.2 Experimental data

Experimental data referenced in this thesis was obtained from the Chaigne lab at Utrecht University. For the experimental distributions of colony sizes, we received raw data of the number of choanoflagellates per colony, starting from a dilution of 1:10 and growing for 24 or 48 hours. The data were imported into Python and the distribution of the number of choanoflagellates per chain was plotted. B. van Amen from the Chaigne lab also provided the raw data on the cell cycle times of choanoflagellates, which were collected through live cell imaging and analyzed using ImageJ. These data were imported into Python and used to create a plot showing the distribution of cell cycle times.

2.3 Fitting

To quantify the goodness of fit between experimental data and model parameters, the sum of squared errors (SSE) was calculated. This approach measures the squared distance between experimental and predicted data [37];

$$\mathrm{SSE} = \sum_{i=1}^{N} (y_i - f(x_i, z_i))^2 \tag{5}$$

where y_i is the value of the experimental data at point i, and $f(x_i, z_i)$ is the predicted value at point i [37]. To fit the gamma distribution and the exponential distribution to the cell cycle data, the least squares error (LSE) approach was used [38]. This algorithm tries various parameter combinations and calculates the SSE for each of these. The distribution that best fits the data is the one with the least square error. To determine the optimal parameters, the algorithm starts with an initial guess. After calculating the LSE for the initial guess, the algorithm adjusts the parameters using the Nelder-Mead method [39], which does not require the derivative of the loss function. This method was chosen due to the complexity of the derivative of the loss function due to the derivative of the gamma function.

3 Results

3.1 Analyzing experimental data and fitting

The first step in analyzing colony sizes of *S. rosetta* chain colonies was to examine the experimental data provided by the Chaigne lab.

3.1.1 Gamma distribution provides the best fit to *S. rosetta* cell cycle durations

The Chaigne lab provided us with the cell cycle duration of *S. rosetta* cells, which can be found in Figure 4A. The purpose of using this data was to gain insight into the underlying mechanics of *S. rosetta* cell division. To do this, we rearranged the data into a cumulative probability distribution, which indicates the likelihood that a random variable will be less than or equal to a given value, accumulating probabilities from the lowest possible value up to the given value [40]. Next, we fitted the cumulative distribution functions (CDFs) of the gamma and the exponential distribution function to the experimental data. In addition, the step function was fitted, which represents a deterministic method. This fitting process was performed using the LSE approach, minimizing the SSD. As shown in Figure 4B, the gamma distribution, with a shape parameter of 3.37 and scale of 1.84, provided the best fit to the data. The exponential distribution, with a scale parameter of 7.57, produced an adequate fit, while the step function, with a median of 5.83, did not align well with the data.

Based on these results, we decided on the following combinations of waiting time distributions for division and bridge breakage, resulting in four models: Exponential - Exponential, Gamma - Exponential, Gamma - Gamma, and Gamma - Deterministic. In these models, the first distribution represents the time handling for division events, while the second corresponds to the time handling for bridge-breaking events. We chose to exclude the deterministic approach for division events, as it was not a good fit to the experimental data.

3.1.2 Gamma - Exponential and Gamma - Gamma models show the best fit to experimental colony size CDF

Next, we compared the CDF of the experimental distribution of colony sizes (recorded after 24 hours and 48 hours) with those produced by the simulations for each of the four models. This analysis aimed to determine good parameters for handling

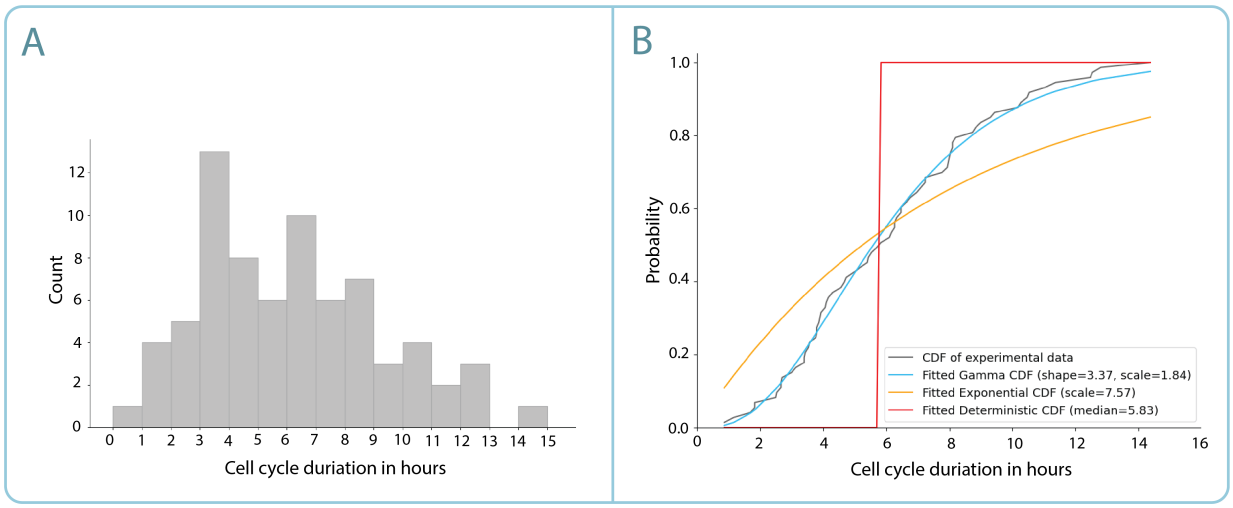


Figure 4: Distributions of and fitting of cell cycle durations over time. (A) Histogram of cell cycle durations in hours recorded during live cell imaging, made by Chaigne lab, n = 73 cell cycles. (B) CDF comparison of experimental cell cycle duration data in grey, with fitted gamma (shape = 3.37, scale = 1.84) in blue, exponential (scale = 7.57) in orange, and deterministic models (median = 5.03) in red, illustrating the fit of different models to the experimental data.

bridge break events, as well as evaluate the model outputs in comparison to the experimental data. The SSD was calculated between the experimental CDFs and the simulation-generated CDFs from the different models, where the CDFs from the models are averages of 1000 simulations. This procedure was repeated for various mean bridge lifetimes ranging from 1 to 40 with steps of 2. As can be seen in Figure 5, for each of the models, a clear minimum could be found, except for the Gamma - Gamma model. Interestingly, for this model, a range of different parameter combinations seem to lead to a low SSD, indicating that the Gamma - Gamma model is relatively robust to changes in parameters. This flexibility could be due to the two parameters that define the distribution, rather than one. We can see that the grid exhibits an L-shape at a mean lifetime of ten to fourteen hours and an alpha of one to two, where the lowest SSD values are centered around. Another observation from the plots in Figure5 is that the curves vary slightly for the 24-hour comparison and the 48-hour one for all models. The lowest SSD value and the best corresponding parameters for each model and each distribution can be found in Table 1. Here, we can find that the Gamma - Gamma model performs the best in each of the two time points, followed up by the Gamma - Exponential model. The Exponential - Exponential model performs relatively well in the 48-hour distribution, while the Gamma - Deterministic model does not fit well to the data compared to the other models. We chose to work with the values obtained from the fitting with the 48-hour experimental distributions, since for the remainder of this thesis we will mostly simulate the colonies for longer periods of time, making these parameters likely more reliable.

Next, the obtained parameters were used in the simulation to plot the CDFs and the colony size distributions, as can be seen in Figure 5B-D. We can see that the graphs overlaying the experimental data seem to match the results obtained from the fitting: the Gamma - Exponential and Gamma - Gamma resemble the experimental data the best. Interestingly, we can see that from the 24-hour CDFs, the models seem to have smaller colony sizes than the experimental distribution. In Figure 5C-D, it can again be seen that the Gamma - Exponential and the Gamma - Gamma models provide the most similar shapes to the experimental distributions, especially in the 48-hour case.

Model (Division-Bridge Break)	**SSD 24h**	**Best parameter(s) 24h**	**SSD 48h**	**Best parameter(s) 48h**
Exponential-Exponential	0.035	$\lambda = 1/30\ h^{-1}$	0.006	$\lambda = 1/22\ h^{-1}$
Gamma-Exponential	0.014	$\lambda = 1/32\ h^{-1}$	0.002	$\lambda = 1/24\ h^{-1}$
Gamma-Gamma	0.004	$\mu = 14h, \alpha = 6$	0.002	$\mu = 16h, \alpha = 2$
Gamma-Deterministic	0.037	t = 12 h	0.076	t = 11 h

Table 1: The sum of squared differences between the CDFs of experimental colony sizes and the CDFs of colony sizes for each model, using the best-fit parameter combinations.

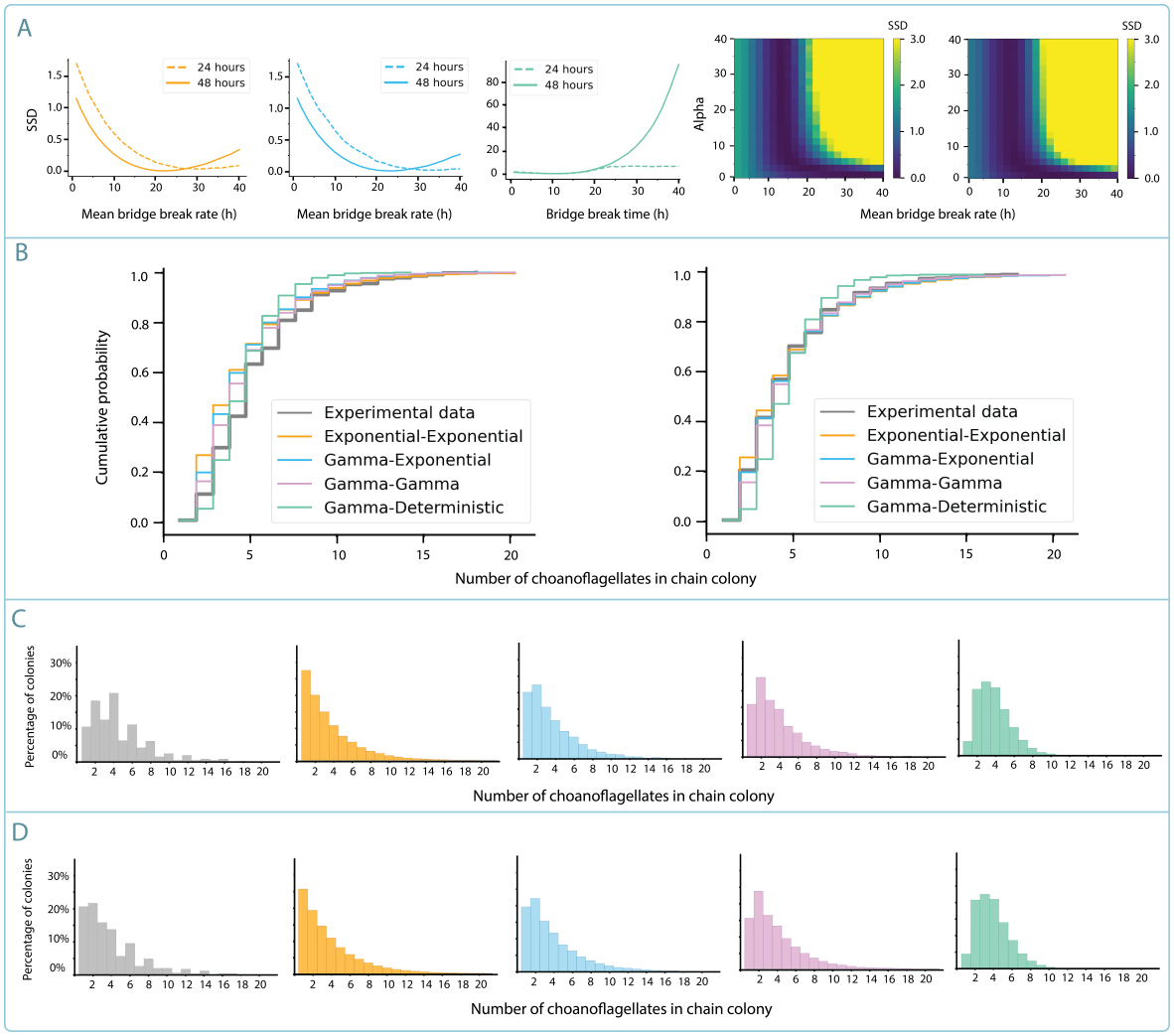


Figure 5: Comparing the distribution of *S. rosetta* chain colony sizes after 24 and 48 hours grown with a density of 1:10 made by Chaigne lab to simulation outputs for each model, averaging over 1000 simulation runs. (A) Plots of the SSD between the experimental colony size CDF after 24 and 48 hours and the simulations output performed for different values of parameters that determine the bridge break rate, done for each model: Exponential - Exponential in orange, Gamma - Exponential in blue, Gamma - Deterministic in sea green, and Gamma - Gamma as gradient map, where the 24 hours comparison is marked as a dotted line and the 48 hours with a continuous one. For the Gamma - Gamma model, the 24 hour comparison is on the left and 48 hours on the right, with SSD values exceeding 3.0 in yellow (B) Comparision of the colony size CDF after 24 hours (left) and 48 hours (right) and the simulations output for the four models using the bridge break parameters related to the minimal SSD for the 48 hours. (C-D) Comparison of experimental colony size distributions after (C) 24 hours and (D) 48 hours to the simulations output for each model: Exponential -Exponential in orange (Division: $\lambda = 1/7.57\ h^{-1}$, Bridge Breaks: $\lambda = 1/22\ h^{-1}$), Gamma - Exponential in blue (Division: $\alpha = 3.37$ $\mu = 6.2\ h$, Bridge Breaks: $\lambda = 1/24\ h^{-1}$), Gamma - Gamma in magenta (Division: $\alpha = 3.37$ $\mu = 6.2\ h$, Bridge Breaks: $\alpha = 2$ $\mu = 16\ h$), and Gamma - Deterministic in sea green (Division: $\alpha = 3.37$ $\mu = 6.2$, Bridge Breaks: $t = 11\ h$).

3.1.3 Modelled *S. rosetta* colonies stabilize in the average colony size and standard deviation

To gain further insight into the dynamics of the different models, we looked at the average and standard deviation of the number of cells per colony. This was done by comparing their trajectories over time. As Figure 6 shows, both the average and the standard deviation of the number of choanoflagellates per chain seem to stabilize over time in each model. However, the trajectories to this stabilization and the value at which the variable stabilizes vary per model. The three methods that utilize the gamma distribution to model the division times have similar trajectories at the start when mostly only division events occur but then start to diverge when bridge break events start to play a role, while in the trajectory of the Exponential - Exponential model diverges from the others from the start. For the average, we can see in Figure 6A-B that all models roughly converge towards the same value between 3.5 and 4.0 of choanoflagellates per colony. This agrees with the experimental data that reads that after 48 hours, the average number of colonies per chain is 3.83, with the mean of the Gamma - Gamma model being

the most similar. The trajectory of the Gamma - Deterministic model, the trajectory is very steep at the beginning, but stabilizes quickly, since then the deterministic bridge break events stabilize. For the stochastic bridge break models, it takes longer to stabilize as can be seen in Figure 6B. For the standard deviation, the Gamma - Deterministic stabilized value lies low relative to the other methods, while the other models roughly stabilize around the same value between 2.8 and 3.2 (Fig 6 C-D). This again aligns with the standard deviation of the experimental data of 2.9, with the standard deviation of the Gamma - Gamma model being the most similar. The observation that the Gamma - Deterministic model stabilizes at a smaller value compared to the other models, can be explained by the notion that deterministic bridge break events lead to more of the same colonies. This is also what was seen in the distributions of Figure 5C-D, where the distribution is dense, instead of spread out.

3.2 Analyzing steady states

In the previous section, we observed stabilization in the average number of cells per colony. In this section, this will be analyzed further.

3.2.1 The steady state is independent of the initial population for all models

First, we plotted the trajectories of the average number of choanoflagellates per colony for different initial populations to see if the same steady state is reached. This was done for each model. In Figure 7, we can see that for all four combinations, steady states in the number of cells per colony are reached independent of the initial population. Across all models, the time to reach a steady state varies. In the models where bridge breaks are not modeled with the exponential distribution, the steady states seem to be reached earlier. In the model where bridge

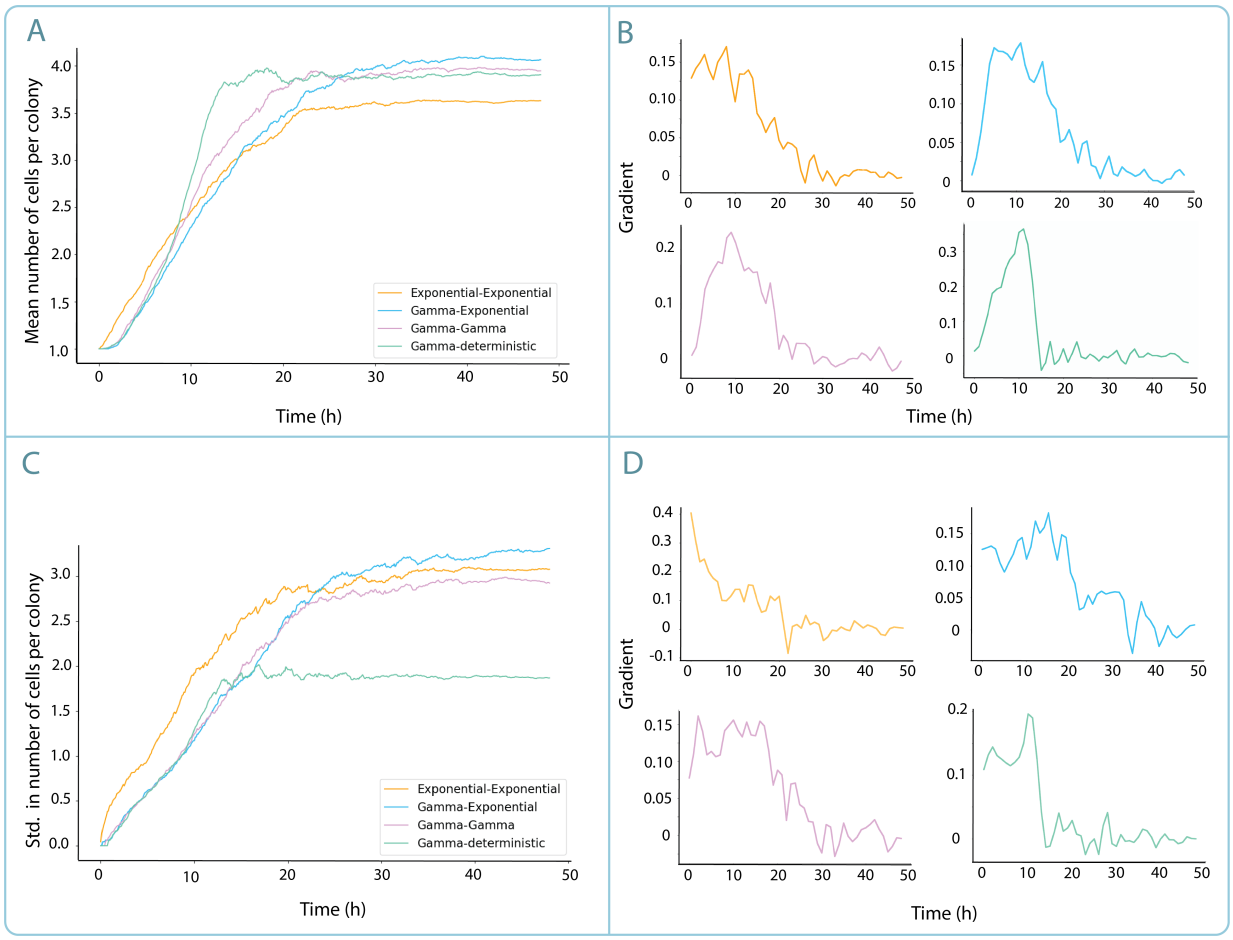


Figure 6: Analysis on the mean and standard deviation of colony sizes of simulated *S. rosetta* chain colonies, averaging over 500 simulation runs with Exponential -Exponential in orange (Division: $\lambda = 1/7.57\ h^{-1}$, Bridge Breaks: $\lambda = 1/22\ h^{-1}$), Gamma - Exponential in blue (Division: $\alpha = 3.37\ \mu = 6.2\ h$, Bridge Breaks: $\lambda = 1/24\ h^{-1}$), Gamma - Gamma in magenta (Division: $\alpha = 3.37\ \mu = 6.2\ h$, Bridge Breaks: $\alpha = 2\ \mu = 16\ h$), and Gamma - Deterministic in sea green (Division: $\alpha = 3.37\ \mu = 6.2$, Bridge Breaks: $t = 11\ h$). (A) Trajectories of the average number of cells per chain colony over time for each of the four models. (B) The gradients of the trajectories of the mean number of cells per colony for each of the four models. (C) Trajectories of the standard deviation in the number of cells per chain colony over time for each of the four models. (D) The gradients of the trajectories of the standard deviation for each of the four models.

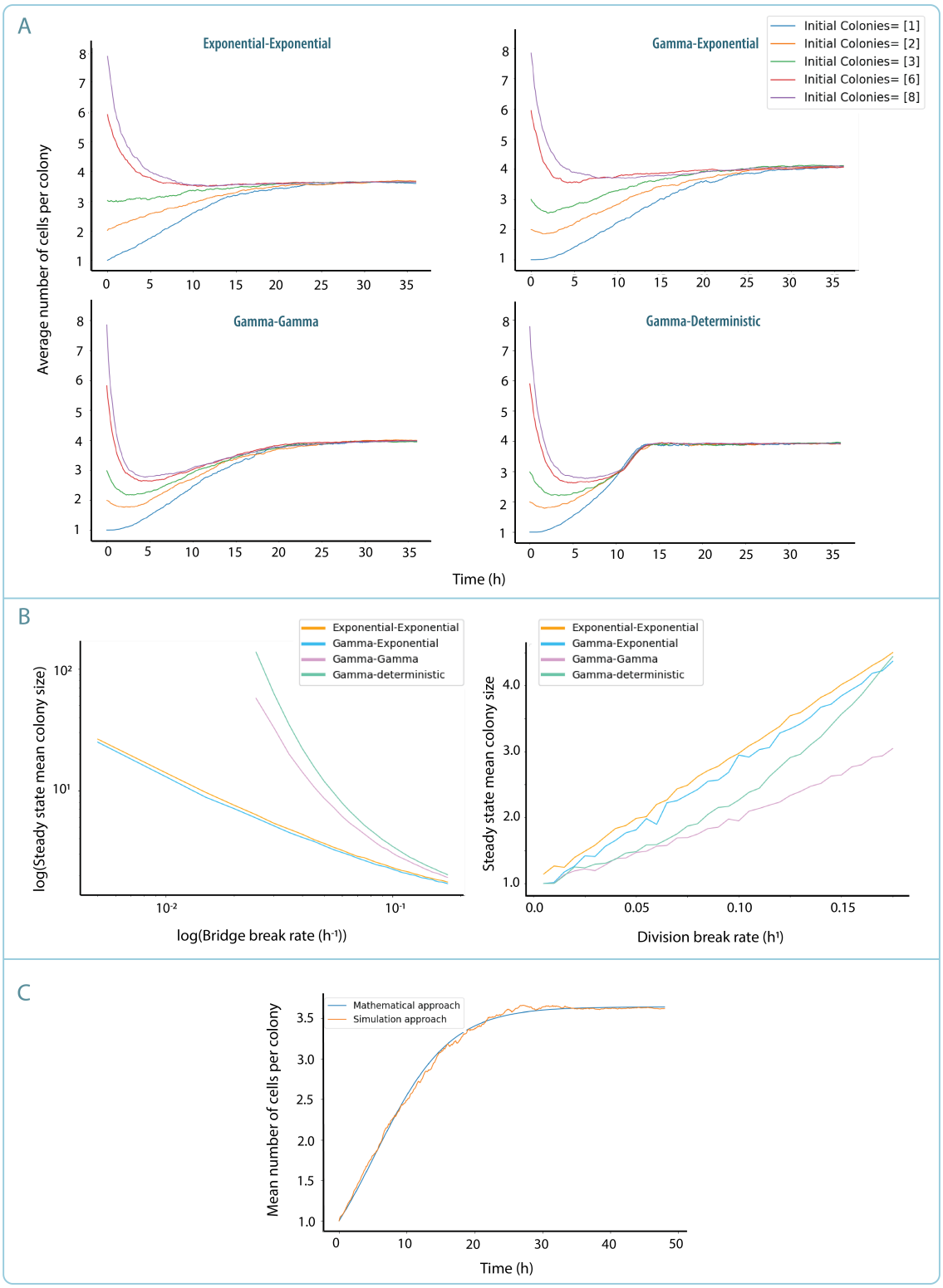


Figure 7: Analysis of the steady state in the average number of choanoflagellates per chain colony for each model, averaging over 500 simulations with Exponential -Exponential (Division: $\lambda = 1/7.57\ h^{-1}$, Bridge Breaks: $\lambda = 1/22\ h^{-1}$), Gamma - Exponential(Division: $\alpha = 3.37$ $\mu = 6.2\ h$, Bridge Breaks: $\lambda = 1/24\ h^{-1}$), Gamma - Gamma (Division: $\alpha = 3.37$ $\mu = 6.2\ h$, Bridge Breaks: $\alpha = 2$ $\mu = 16\ h$), and Gamma - Deterministic (Division: $\alpha = 3.37$ $\mu = 6.2$, Bridge Breaks: $t = 11\ h$) (A) Trajectories of the average number of choanoflagellates per chain colony for different initial colony sizes for the different models. (B) Parameter sweeps of the steady state value for various bridge break rates (left) and division rates (right) for each model. (C) Trajectory of the mean colony size over time from the Exponential - Exponential model compared to the trajectory of the mathematical description.

lifetimes are deterministic, the path to the steady state is very similar for all the initial population sizes, while for the other models, the paths are less predictable due to their more stochastic nature. Despite these differences, the steady state is reached in all cases, suggesting that simulated *S. rosetta* chain colonies grow towards a stable average size, independent of the initial population size.

3.2.2 Division and bridge break rate significantly influence the value of the steady state in *S. rosetta* chain colonies

Next, we analyzed how the rates affect the steady states. This was done by performing parameter sweeps for the bridge break rate and the division rate. Here, the parameters that determine one of the events were kept constant while varying the parameters of the other. Then, we plotted this variable against the steady state number of choanoflagellates per colony after 48 hours. The steady state is reached if there is no more than 5 percent change in the average over the past two hours. In cases where the event is modelled with the gamma distribution, where the distribution depends on two parameters, the ratio between the two parameters were kept constant while varying the mean value. The results can be seen in Figure 7. We can see that the bridge break is log-log related to the steady state while the division rate seems to be linear for all the models. Furthermore, it can be seen that for low bridge break rates, no steady state is reached within 48 hours, as indicated by the absence of plotted values. This is a result of the low to no number of bridge break events happening in 48 hours for these rates, leading to the unlimited growth of the colonies, while in cases where the bridge break is modeled stochastically, bridge breaks can still occur within these hours, reducing this unlimited growth.

3.2.3 Steady-state value is determined by the ratio of division and bridge break rate

Our final aim was to see if the trajectories of the average number of choanoflagellates per colony over time could be described mathematically and also find a description for the observed steady states. The average number of choanoflagellates per chain can be calculated by knowing the number of choanoflagellates and the number of bridges by:

$$M = \frac{c}{c-b} \tag{6}$$

with M being the average number of choanoflagellates per colony, c the total number of choanoflagellates, and b the total number of bridges. Here, $c-b$ describes the number of colonies, including single-celled ones. Furthermore, the change in c and b per time step can be described using:

$$\frac{dc}{dt} = ck_d \tag{7}$$

$$\frac{db}{dt} = ck_d - bk_b \tag{8}$$

with k_d as the division rate and k_b as the bridge break rate.

It is important to note that this description only holds when both events are handled with an exponential distribution. Now, by integrating these differential equations descriptions by using the initial values $c(0) = 1$ and $b(0) = 0$, M can be found as a function of time:

$$c = \exp(k_d t) \tag{9}$$

$$b = \frac{k_d}{k_b + k_d} * (\exp(-k_b t) * (-1 + \exp(k_b + k_d)t) \tag{10}$$

By plugging this in Equation 6, we obtain:

$$M(t) = \frac{k_b + k_d}{k_b + k_d \exp((k_b + k_d)t)} \tag{11}$$

In Figure 7C, the mathematical solution is plotted in the same graph as the simulation output for the Exponential - Exponential model, which seem to align well.

Furthermore, the steady states can be described mathematically. The steady state occurs when

$$\frac{dM}{dt} = 0. \tag{12}$$

We can find this by either using the chain rule with $\frac{dc}{dt}$ and $\frac{db}{dt}$ or by directly differentiating our description of M as a function of time. By using the chain rule, we can calculate $\frac{dM}{dt}$. Using:

$$\frac{dM}{dt} = \frac{dM}{dc}\frac{dc}{dt} + \frac{dM}{dt}\frac{db}{dt} \tag{13}$$

one ends up with

$$\frac{dM}{dt} = \frac{c^2 - c*b*k_d - n*b*k_b}{(c-b)^2} \tag{14}$$

Then with $\frac{dM}{dt} = 0$ to get the steady state we get:

$$M_{steady} = \frac{1}{1 - \frac{k_d}{k_d + k_b}} \tag{15}$$

This solution aligns with the results from the simulation.

4 Conclusion and Discussion

Multicellularity evolved many times in all clades of life, often starting with simple multicellular colonies. One main feature of simple multicellular colonies is their size. In this thesis, we presented a discrete stochastic model that simulates chain colonies of the choanoflagellate *S. rosetta*, the closest living relative to animals. This model is defined by two events: division and bridge breakage. Here, various probability distribution functions were tested to draw inter-division times and bridge lifetimes from. Our findings indicate that division follows an age-dependent distribution, as opposed to a memoryless exponential or deterministic pattern. This age-independent behavior is expected since the cell must progress through the cell cycle before it divides. For bridge lifetimes, the results seem to rule out the possibility that they follow a deterministic trend. Furthermore,

it was found that this system reaches steady states in the average number of cells per chain, where the bridge break rate has a log-log relationship with the steady-state value, and division is linearly related. Finally, it was observed that these trajectories and steady states can be mathematically described.

It has been described many times that there are variations in cell cycle times between individual cells, even in cells within the same colony [41–46]. This stochasticity is mostly attributed to variabilities in time spent in the G1 phase of the cycle [43, 47]. Especially in small populations of cells, it is of importance that cell-to-cell variability is accounted for to accurately describe their growth [48, 49]. In our model, we capture this variability in inter-division times by drawing these times from a PDF that describes a stochastic process. It was found through fitting experimental cell-cycle times to various PDFs, that they are best fitted by the gamma distribution ($\mu = 6.2$, $\alpha = 3.37$). This indicates that the division in *S. rosetta* is age-dependent, with variations around the mean age. It is important to note that due to the low availability of experimental data, the estimated parameters are not fully reliable, and should be reevaluated with more data. Various models agree that the cell cycle should be modeled with the gamma distribution, while there are also other methods described in the literature, such as with the delayed exponential model or an exponentially modified gamma distribution [45–47, 50, 51]. The results in this research indicate that the gamma distribution is the best option to describe the division of choanoflagellates until there is more experimental data to compare to.

Next, we found that the experimental distribution of colony sizes can be best fitted by the Gamma - Exponential and Gamma - Gamma models of our simulation. In further analysis of the mean and standard deviation of the mean colony size, we found that these values are more aligned with the Gamma - Gamma model, however, due to the limited experimental data and computing power, we can not make definitive conclusions about the best distribution. Moreover, it is reasonable that the gamma distribution provides a better fit for the data, as it includes more parameters compared to the other methods we tested. By gathering more experimental data, a conclusion can hopefully be drawn. If we find that the Gamma - Gamma model performs best, then this indicates that the lifetime of bridges is age-dependent, while if the Gamma - Exponential distribution fits best, then the bridges are likely memoryless and break through some other mechanisms or just randomly. Not much is known about intercellular bridges in choanoflagellates or the lifetime of intercellular bridges overall. It is known that stable intercellular bridges can persist for long periods, hours to days, but there has not been a clear description of their precise rate or the variabilities in this time [31]. In this thesis, we found estimates for the lifetime of bridges in *S. rosetta* chain colonies, which varied across models but generally ranged around twenty hours. Due to the limited knowledge of the dynamics of cytokinetic bridges, further validating these lifetimes and finding a PDF that describes them in choanoflagellates to broaden our understanding of this mechanism in a model organism close to animals could provide insight into these bridges in the origin of animals [14, 28].

Finally, the results indicate that the system seems to reach a steady state in the average number of cells per colony, indicating that the colonies seem to reach a balance in the division and the bridge break events resulting in stabilized growth. The time until a steady state was reached varied per model. Cases, where the bridge breaks are modeled deterministically, reach a steady state faster, likely due to less random fluctuations as opposed to the more stochastic models, which take a longer time to stabilize. We found that this steady state is reached regardless of the initial colony sizes, strengthening the generality of this steady state. Our mathematical description of this steady state shows that their value is solely dependent on the ratio of the bridge break rate and the division rate. Future research could focus on imaging *S. rosetta* colonies for a longer period to find whether these steady states are also observed in vitro.

In this model, colony sizes of chain colonies are determined by division, increasing colony sizes, and bridge breaks, which result in two separate colonies. This is a simple approach and might neglect processes that could also potentially affect colony size. In literature, various approaches are described to simulate colony growth, some taking a similar one to ours. A paper that lies close to our model is Nanda et al, where the model is described by: cell division, intercellular connection breakage, and the maximum number of connections a cell can have (the kissing number) [10]. This paper focused on modeling the colonies of budding yeast and found similar results to the ones described in this thesis where the system reaches the steady state determined by the ratio between the division rate and the connection breaking rate if the maximum number of links per cell is not reached. This is consistent with our finding that colony growth stabilizes due to a balance between division and bridge break events, leading to a steady state in the number of cells per colony. Despite these similarities, there are key differences between our models that are worth highlighting, such as that their model does not account for the variations in inter-division times and that their model has an extra layer of complexity by implementing the kissing number, while in the organism in our model, there is only one bridge connecting each pair of cells, which may limit the applicability of our model to other organisms since they usually do not grow in chains. Although the simplicity of our model has its benefits, such as making it straightforward to understand and easy to run, some models incorporate more events to describe colony sizes. So do some models focus on the events of quiescence, proliferation, and apoptosis to describe their sizes, which seem to perform well to model mammalian cells [47, 52]. Another model on phytoplankton describes its colonies by division, stomatocyst production, colony breakage, and colony loss, where colony loss describes all processes that could lead to colony loss such as sinking [53]. Incorporating other events such as quiescence, colony loss, and cell death into our model could improve our model.

The main limitation is that not enough experimental data was obtained to make claims on the exact rates and scale parameters that should be used to model choanoflagellate chain colonies. With more data, we can validate our findings and also obtain more reliable parameters for the waiting time distributions. Next, in the simulation, we set the initial colony size to one choanoflagellate for convenience, while in the experimental data, the colonies were grown from many single-celled colonies. This should not matter because the cumulative probability distribution functions were used; however, when comparing the model to classic distributions, the initial population parameter should be set to more single-celled colonies depending on the density of the cells measured. There are also limitations with the model: the simulation can be computationally demanding depending on the analysis; therefore, the division rate cannot be too high, and neither can the simulation time. This is because the number of cells increases rapidly, and at some point, there will be too many choanoflagellates for a normal computer to handle and the simulation will take a long time to run. Therefore, in our parameter sweeps, the division rate takes a maximum value of 0.20. It is unlikely that the division rate will be higher based on our fitting of the experimental data and the literature. If you wish to set a higher division rate, this issue can be solved by running the simulation using high-performance computing (HPC). Furthermore, the simulation has limitations in its complexity. This simulation is a highly simplified case of the experimental and natural environmental conditions in which *S. rosetta* lives. In the simulation, we only focused on two life-cycle stages and excluded the others. Furthermore, we did not consider environmental factors, such as food and predators.

In this thesis, we developed the first model that describes chain colonies of the choanoflagellate *S. rosetta*. The results provide a new understanding on the colonies of choanoflagellates and therefore could also contribute to the understanding of the origin of animal multicellularity and the possible mechanisms by which this happened [1]. Further developing this model by including other events could build upon this insight and also possibly make this model applicable to other organisms. The results obtained by this study should be revised when more experimental data is available to enhance them.

5 Acknowledgments

I want to thank Dr. Florian Berger for his supervision and guidance throughout this project, and also for the interesting conversations and refreshing perspectives on academia. I also want to thank everyone in the Chaigne lab for allowing me to attend their lab meetings, with special thanks to Babette van Amen and Dr. Agathe Chaigne for allowing me to use their data for my project. Last, I want to thank A. Elbakyan for her continued fight in making science accessible to everyone.

GenAI statement In this thesis, I used the free version of Quillbot (v15.421.10) as a grammar checker and to enhance the clarity and fluency of some of the sentences in this report [54].

References

[1] N. Ros-Rocher, A. Perez-Posada, M. Leger, and I. Ruiz-Trillo, "The origin of animals: An ancestral reconstruction of the unicellular-to-multicellular transition," *Open Biology*, vol. 11, p. 200359, 02 2021.

[2] K. Niklas and S. Newman, "The many roads to (and from) multicellularity," *Journal of experimental botany*, vol. 71, 12 2019.

[3] M. Maldonado, "Choanoflagellates, choanocytes, and animal multicellularity," *Invertebrate Biology*, vol. 123, no. 1, pp. 1–22, 2004.

[4] L. G. Nagy, G. M. Kovács, and K. Krizsán, "Complex multicellularity in fungi: evolutionary convergence, single origin, or both?" *Biological Reviews*, vol. 93, no. 4, pp. 1778–1794, 2018.

[5] S. R. Fairclough, Z. Chen, E. Kramer, Q. Zeng, S. Young, H. M. Robertson, E. Begovic, D. J. Richter, C. Russ, M. J. Westbrook *et al.*, "Premetazoan genome evolution and the regulation of cell differentiation in the choanoflagellate salpingoeca rosetta," *Genome biology*, vol. 14, pp. 1–15, 2013.

[6] T. Brunet, B. T. Larson, T. A. Linden, M. J. Vermeij, K. McDonald, and N. King, "Light-regulated collective contractility in a multicellular choanoflagellate," *Science*, vol. 366, no. 6463, pp. 326–334, 2019.

[7] L. Bich, T. Pradeu, and J.-F. Moreau, "Understanding multicellularity: The functional organization of the intercellular space," *Frontiers in Physiology*, vol. 10, 09 2019.

[8] K. J. Niklas and S. A. Newman, "The origins of multicellular organisms," *Evolution & development*, vol. 15, no. 1, pp. 41–52, 2013.

[9] K. Tong, G. O. Bozdag, and W. C. Ratcliff, "Selective drivers of simple multicellularity," *Current Opinion in Microbiology*, vol. 67, p. 102141, 2022.

[10] P. Nanda, J. Barrere, T. LaBar, and A. W. Murray, "A dynamic network model predicts the phenotypes of multicellular clusters from cellular properties," *Current Biology*, vol. 34, no. 12, pp. 2672–2683.e4, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960982224006122

[11] T. Tapics, H. M. Sosik, and Y. Huot, "A discrete, stochastic model of colonial phytoplankton population size structure: Development and application to in situ imaging-in-flow cytometer observations of dinobryon," *Journal of phycology*, vol. 59, no. 5, pp. 1005–1024, 2023.

[12] M. Boraas, D. Seale, and J. Boxhorn, "Phagotrophy by flagellate selects for colonial prey: A possible origin of multicellularity," *Evolutionary Ecology*, vol. 12, 02 1998.

[13] A. Sebé-Pedrós, B. M. Degnan, and I. Ruiz-Trillo, "The origin of metazoa: a unicellular perspective," *Nature Reviews Genetics*, vol. 18, no. 8, pp. 498–512, 2017.

[14] M. J. Dayel, R. A. Alegado, S. R. Fairclough, T. C. Levin, S. A. Nichols, K. McDonald, and N. King, "Cell differentiation and morphogenesis in the colony-forming choanoflagellate salpingoeca rosetta," *Developmental biology*, vol. 357, no. 1, pp. 73–82, 2011.

[15] N. King, "The unicellular ancestry of animal development," *Developmental cell*, vol. 7, no. 3, pp. 313–325, 2004.

[16] M. Carr, B. S. Leadbeater, R. Hassan, M. Nelson, and S. L. Baldauf, "Molecular phylogeny of choanoflagellates, the sister group to metazoa," *Proceedings of the National Academy of Sciences*, vol. 105, no. 43, pp. 16 641–16 646, 2008.

[17] I. Ruiz-Trillo, G. Burger, P. W. Holland, N. King, B. F. Lang, A. J. Roger, and M. W. Gray, "The origins of multicellularity:

a multi-taxon genome initiative," *TRENDS in Genetics*, vol. 23, no. 3, pp. 113–118, 2007.

[18] B. S. Leadbeater, *The choanoflagellates*. Cambridge University Press, 2015.

[19] N. King, "Choanoflagellates," *Current Biology*, vol. 15, no. 4, pp. R113–R114, 2005.

[20] J. M. Pinskey, A. Lagisetty, L. Gui, N. Phan, E. Reetz, A. Tavakoli, G. Fu, and D. Nicastro, "Three-dimensional flagella structures from animals' closest unicellular relatives, the choanoflagellates," *Elife*, vol. 11, p. e78133, 2022.

[21] T. T. Hoffmeyer and P. Burkhardt, "Choanoflagellate models—monosiga brevicollis and salpingoeca rosetta," *Current opinion in genetics & development*, vol. 39, pp. 42–47, 2016.

[22] D. Laundon, B. T. Larson, K. McDonald, N. King, and P. Burkhardt, "The architecture of cell differentiation in choanoflagellates and sponge choanocytes," *PLoS biology*, vol. 17, no. 4, p. e3000226, 2019.

[23] S. R. Fairclough, M. J. Dayel, and N. King, "Multicellular development in a choanoflagellate," *Current Biology*, vol. 20, no. 20, pp. R875–R876, 2010.

[24] J. Reyes-Rivera, Y. Wu, B. G. Guthrie, M. A. Marletta, N. King, and T. Brunet, "Nitric oxide signaling controls collective contractions in a colonial choanoflagellate," *Current Biology*, vol. 32, no. 11, pp. 2539–2547.e5, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960982222005863

[25] J. D. Spafford, A. Mehta, S. Waqar, P. S. Velayudhan, J. Fan, T. Sharma, A. Bhat, Z. Alsamman, C. Jeffery, V. Magdanz *et al.*, "Electrical and calcium signaling of nerve and muscle in unicellular choanoflagellates," *Biophysical Journal*, vol. 123, no. 3, p. 192a, 2024.

[26] T. C. Levin and N. King, "Evidence for sex and recombination in the choanoflagellate salpingoeca rosetta," *Current Biology*, vol. 23, no. 21, pp. 2176–2180, 2013.

[27] D. M. Needham, C. Poirier, C. Bachy, E. E. George, S. Wilken, C. C. Yung, A. J. Limardo, M. Morando, L. Sudek, R. R. Malmstrom *et al.*, "The microbiome of a bacterivorous marine choanoflagellate contains a resource-demanding obligate bacterial associate," *Nature microbiology*, vol. 7, no. 9, pp. 1466–1479, 2022.

[28] A. Chaigne and T. Brunet, "Incomplete abscission and cytoplasmic bridges in the evolution of eukaryotic multicellularity," *Current Biology*, vol. 32, no. 8, pp. R385–R397, 2022.

[29] A. H. Knoll, "The multiple origins of complex multicellularity," *Annual Review of Earth and Planetary Sciences*, vol. 39, no. 1, pp. 217–239, 2011.

[30] T. Brunet and N. King, "The origin of animal multicellularity and cell differentiation," *Developmental cell*, vol. 43, no. 2, pp. 124–140, 2017.

[31] J. Singh, J. Imran Alsous, K. Garikipati, and S. Y. Shvartsman, "Mechanics of stabilized intercellular bridges," *Biophysical Journal*, vol. 121, no. 16, pp. 3162–3171, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0006349522005434

[32] I. P. N. Kaisa Haglund and H. Stenmark, "Structure and functions of stable intercellular bridges formed by incomplete cytokinesis during development," *Communicative & Integrative Biology*, vol. 4, no. 1, pp. 1–9, 2011, pMID: 21509167. [Online]. Available: https://doi.org/10.4161/cib.13550

[33] N. van Kampen, *Stochastic Processes in Physics and Chemistry (Third Edition)*, third edition ed., ser. North-Holland Personal Library. Elsevier, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780444529657500064

[34] C. Forbes, M. Evans, N. Hastings, and B. Peacock, *Terms and Symbols*. John Wiley Sons, Ltd, 2010. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470627242.ch2

[35] K. B. Piao Chen and X. Zhao, "A comprehensive toolbox for the gamma distribution: The gammadist package," *Journal of Quality Technology*, vol. 55, no. 1, pp. 75–87, 2023. [Online]. Available: https://doi.org/10.1080/00224065.2022.2053794

[36] E. U., O. M.O.O., and F. C.E., "A study of properties and applications of gamma distribution," *African Journal of Mathematics and Statistics Studies*, vol. 4, pp. 52–65, 07 2021.

[37] D. L. Mohr, W. J. Wilson, and R. J. Freund, "Chapter 6 - inferences for two or more means," in *Statistical Methods (Fourth Edition)*, fourth edition ed., D. L. Mohr, W. J. Wilson, and R. J. Freund, Eds. Academic Press, 2022, pp. 243–299. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128230435000060

[38] S. J. Miller, *Chapter 24. The Method of Least Squares*. Princeton: Princeton University Press, 2017, pp. 625–635. [Online]. Available: https://doi.org/10.1515/9781400885381-026

[39] S. Singer and J. Nelder, "Nelder-Mead algorithm," *Scholarpedia*, vol. 4, no. 7, p. 2928, 2009, revision #91557.

[40] H. Pishro-Nik, *Introduction to Probability, Statistics, and Random Processes*. Kappa Research LLC, 2014, https://www.probabilitycourse.com.

[41] D. J. Kiviet, P. Nghe, N. Walker, S. Boulineau, V. Sunderlikova, and S. J. Tans, "Stochasticity of metabolism and growth at the single-cell level," *Nature*, vol. 514, no. 7522, p. 376—379, October 2014. [Online]. Available: https://doi.org/10.1038/nature13582

[42] A. Golubev, "Transition probability in cell proliferation, stochasticity in cell differentiation, and the restriction point of the cell cycle in one package," *Progress in Biophysics and Molecular Biology*, vol. 110, no. 1, pp. 87–96, 2012, special Issue: Chance at the heart of the cell. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0079610712000259

[43] J. A. Smith and L. Martin, "Do cells cycle?" *Proceedings of the National Academy of Sciences*, vol. 70, no. 4, pp. 1263–1267, 1973. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.70.4.1263

[44] F. Jafarpour, "Cell size regulation induces sustained oscillations in the population growth rate," 09 2018.

[45] T. S. Weber, I. Jaehnert, C. Schichor, M. Or-Guil, and J. Carneiro, "Quantifying the length and variance of the eukaryotic cell cycle phases by a stochastic model and dual nucleoside pulse labelling," *PLoS Computational Biology*, vol. 10, 2014. [Online]. Available: https://api.semanticscholar.org/CorpusID:2849218

[46] A. Golubev, "Applications and implications of the exponentially modified gamma distribution as a model for time variabilities related to cell proliferation and gene expression," *Journal of Theoretical Biology*, vol. 393, pp. 203–217, 2016.

[47] E. D. Hawkins, M. L. Turner, M. R. Dowling, C. Van Gend, and P. D. Hodgkin, "A model of immune regulation as a consequence of randomized lymphocyte division and death times," *Proceedings of the National Academy of Sciences*, vol. 104, no. 12, pp. 5032–5037, 2007.

[48] A. Barizien, M. Suryateja Jammalamadaka, G. Amselem, and C. N. Baroud, "Growing from a few cells: combined effects of initial stochasticity and cell-to-cell variability," *Journal of the Royal Society Interface*, vol. 16, no. 153, p. 20180935, 2019.

[49] J. J. Tyson and O. Diekmann, "Sloppy size control of the cell division cycle," *Journal of theoretical biology*, vol. 118, no. 4, pp. 405–426, 1986.

[50] A. Zilman, V. V. Ganusov, and A. S. Perelson, "Stochastic models of lymphocyte proliferation and death," *PLoS one*, vol. 5, no. 9, p. e12775, 2010.

[51] A. Maler and F. Lutscher, "Cell-cycle times and the tumour control probability," *Mathematical medicine and biology: a journal of the IMA*, vol. 27, no. 4, pp. 313–342, 2010.

[52] M. Kimmel and D. E. Axelrod, "Unequal cell division, growth regulation and colony size of mammalian cells: a mathematical model and analysis of experimental data," *Journal of theoretical biology*, vol. 153, no. 2, pp. 157–180, 1991.

[53] T. Tapics, H. M. Sosik, and Y. Huot, "A discrete, stochastic model of colonial phytoplankton population size structure: Development and application to in situ imaging-in-flow cytometer observations of dinobryon," *Journal of phycology*, vol. 59, no. 5, pp. 1005–1024, 2023.

[54] QuillBot, "Quillbot: Paraphrasing tool and grammar checker," n.d., accessed: 2024-10-01. [Online]. Available: https://www.quillbot.com