PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

The role of intracellular signaling in the stripe formation in engineered *Escherichia coli* populations

Xiaoru Xue[1], Chuan Xue[2]*, Min Tang[1]*

1 School of Mathematics and Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai, China, **2** Department of Mathematics, Ohio State University, Columbus, Ohio, United States of America

* tangmin@sjtu.edu.cn (MT); cxue@math.osu.edu (CX)

OPEN ACCESS

Citation: Xue X, Xue C, Tang M (2018) The role of intracellular signaling in the stripe formation in engineered *Escherichia coli* populations. PLoS Comput Biol 14(6): e1006178. https://doi.org/10.1371/journal.pcbi.1006178

Editor: Kevin Painter, UNITED KINGDOM

Received: January 1, 2018

Accepted: May 5, 2018

Published: June 4, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: MT was supported by National Science Foundation of China 91330203 (http://www.nsfc.gov.cn) and the RGC of the HKSAR under grant 202713 during her visits to HKBU(http://www.ugc.edu.hk/eng/rgc/). CX was supported in part by US National Science Foundation grant No. DMS 1312966 and National Science Foundation CAREER Award 1553637 (https://www.nsf.gov). The funders had no role in study design, data

Abstract

Recent experiments showed that engineered *Escherichia coli* colonies grow and self-organize into periodic stripes with high and low cell densities in semi-solid agar. The stripes develop sequentially behind a radially propagating colony front, similar to the formation of many other periodic patterns in nature. These bacteria were created by genetically coupling the intracellular chemotaxis pathway of wild-type cells with a quorum sensing module through the protein CheZ. In this paper, we develop multiscale models to investigate how this intracellular pathway affects stripe formation. We first develop a detailed hybrid model that treats each cell as an individual particle and incorporates intracellular signaling via an internal ODE system. To overcome the computational cost of the hybrid model caused by the large number of cells involved, we next derive a mean-field PDE model from the hybrid model using asymptotic analysis. We show that this analysis is justified by the tight agreement between the PDE model and the hybrid model in 1D simulations. Numerical simulations of the PDE model in 2D with radial symmetry agree with experimental data semi-quantitatively. Finally, we use the PDE model to make a number of testable predictions on how the stripe patterns depend on cell-level parameters, including cell speed, cell doubling time and the turnover rate of intracellular CheZ.

Author summary

One of the central problems in biology is to understand the underlying mechanisms responsible for spatial pattern formation in complex systems. This is a difficult task because the essential mechanisms for pattern formation often involve multiple space and time scales and are often buried in overwhelmingly complex physiological details. Recently, synthetic biology has made it possible to investigate strategies of pattern formation in relatively simpler, but still complex, systems. Here we develop multiscale models to help explain the role of intracellular signaling in the formation of stripe patterns in engineered *E. coli* colonies.

collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Understanding the formation of regularly spaced structures, such as vertebrate segments, hair follicles, fish pigmentation and animal coats, is a fundamental problem in developmental biology [1–7]. These patterns involve the complex interaction of intracellular signaling, cell-cell communication, cell growth and cell migration. The overwhelmingly complex physiological context usually makes it difficult to uncover the interplay of these mechanisms. Synthetic biology has recently been used to extract essential components of complex biological systems and examine potential strategies for pattern formation [8–11].

One of these problems relate to the bacterium *Escherichia coli*. Recently in [12], the chemotaxis signaling pathway of *E. coli* has been engineered and coupled with a quorum sensing module, leading to cell-density suppressed cell motility. When a suspension of the engineered cells is inoculated at the center of a petri dish with semi-solid agar and rich nutrient, the colony grows, moves outward and sequentially establishes rings or "stripes" with a high density of cells behind the colony front (Fig 1A). These spatial patterns form in a strikingly similar way as many periodic patterns in other biological systems. When the maximum density of the motile front reaches a threshold, an immotile zone is nucleated. The immotile zone then absorbs bacteria from its neighborhood to expand, forming alternating high and low density zones. These patterns do not form when using wild-type *E. coli*; instead, the colony simply expands outward and forms a uniform lawn. The goal of this paper is to use mathematical models to elucidate the underlying mechanisms for this pattern formation, with a special focus on the roles of intracellular signaling.

E. coli is an enteric gram-negative bacterium that moves by alternating forward-moving "runs" and reorienting "tumbles". It has 6-8 flagella on its surface that can rotate either clockwise (CW) or counterclockwise (CCW) (Fig 1B). If the majority of its flagella rotate CCW they form a bundle and push the cell to run forward with a speed $\sim 10 - 30\mu$m/s. If some flagella

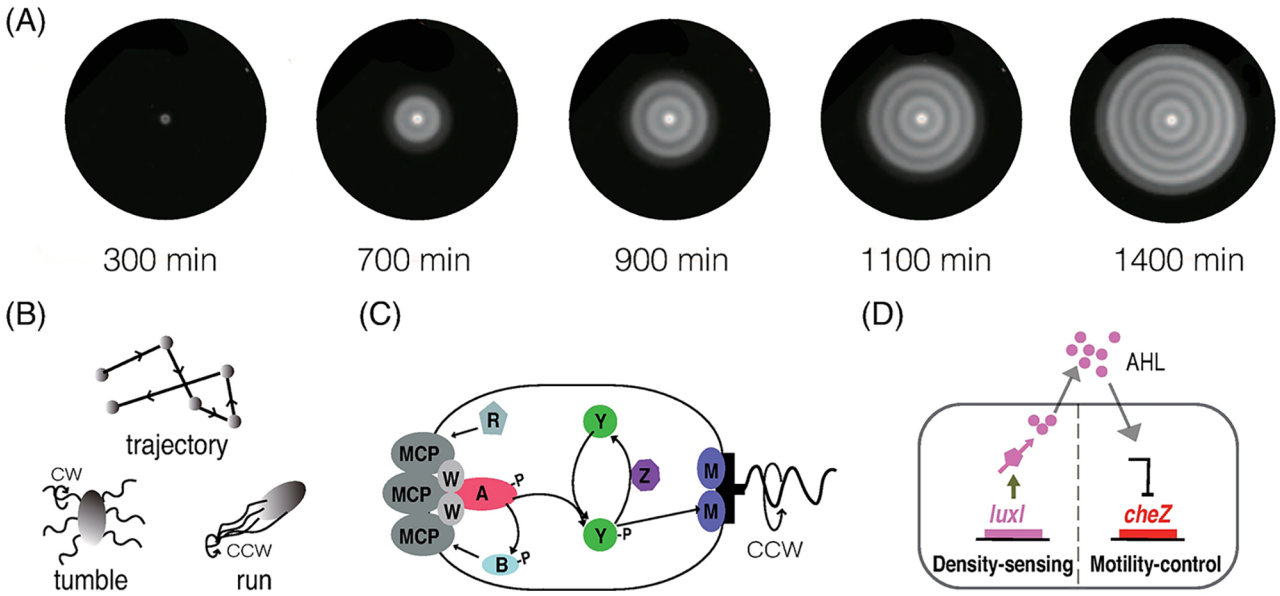

Fig 1. Sequential stripe formation in an engineered *E. coli* colony. (A) Concentric stripe patterns formed in experiments. Scale bar: 1 cm. (B) Run-and-tumble movement. (C)The intracellular chemotaxis pathway of *E. coli*. MCP is the transmembrane receptors. The letters in the figure represents the corresponding proteins involved in chemotaxis, e.g., A represents CheA. (D) The quorum-sensing module. (A), (D) Reproduced from Fig. 1 in Liu et al, Science, Vol 334, 238–241, 2011 [12].

https://doi.org/10.1371/journal.pcbi.1006178.g001

rotate CW they fly apart and the cell tumbles in place. *E. coli* can bias its movement in response to external chemical signals, e.g, towards locations with higher concentration of chemoattractant or lower concentration of repellent, which is called chemotaxis. The molecular mechanism of *E. coli* chemotaxis is summarized in Fig 1C. The transmembrane chemoreceptors (denoted as MCP) form stable ternary complexes with the intracellular signaling proteins CheA and CheW. CheA is an auto-kinase and also a kinase for the response regulators CheY and CheB. The activity of CheA depends on the ligand-binding state of the receptor complex as well as its methylation level: attractant-binding reduces CheA activity and methylation increases it. The phosphorylated form CheYp binds to the flagella motor and increases the probability of clockwise (CW) rotation. On the other hand, CheBp and CheR change the methylation state of the receptor at a slower rate: CheR methylates it and CheBp demethylates it. Upon attractant binding, CheA activity is reduced immediately, leading to lower CheYp and CheBp. Then a shift of the methylation-demethylation cycle gradually restores CheA activity on a slower time scale.

In [12], the quorum-sensing module of bacterium *Vibrio fischeri* was embedded into *E. coli* and used to control the transcription of *cheZ* (Fig 1D). The engineered cell synthesizes and secretes acyl-homoserine lactone (AHL), a small molecule that is freely diffusible across the cell membrane and degrades rapidly. At high concentrations, AHL suppresses the transcription of *cheZ* in an ultra-sensitive manner. If *cheZ* is suppressed, CheZ protein becomes diluted as the cell grows and divides. Because CheZ is a dephosphorylation kinase of CheYp, a reduction of CheZ protein can immediately lead to higher CheYp concentration and thus more persistent tumbles of the cell. This, in turn, causes changes to the chemoreceptors as well as to other proteins involved in chemotaxis, and triggers a non-classic chemotactic cellular response. To quantify the effect of AHL in single cell movement, one must take into account the whole chemotaxis pathway as well as CheZ turnover.

A phenomenological PDE model was used to explain the pattern formation process in [12] and a simplified version was analyzed in [13]. The model consists of a system of reaction-diffusion equations for the cell density, AHL and nutrient concentrations. The diffusion rate of the cell population is assumed to be a switch-like function of the local AHL concentration. Since the whole chemotaxis pathway is involved in the pattern formation process, it is unclear how cell movement can be reduced to an anisotropic (or cross) diffusion process. Moreover, the model does not address the role of intracellular signaling in stripe formation and cannot be used to understand how the spatial structure of the high-density and low-density regions depends on cell-level parameters.

To address these questions, we first developed a hybrid model for the stripe formation that accounts for the behavior of individual cells. The model starts with a detailed description of intracellular signaling, single cell movement and cell division. This individual-based component is then coupled with reaction-diffusion equations for AHL and nutrient concentrations. The multiscale nature of this model allows us to explore the relations between cellular processes on a time scale of seconds to minutes and population dynamics on a time scale of hours. Simulations of our hybrid model showed the same stripe patterns as observed in experiments, but they are very time-consuming due to the large number of cells involved in the pattern formation process. If cells double every 30 minutes, then during a typical time period for pattern formation, e.g. 10 hours, the population size can grow $2^{20} \approx 10^6$ times.

To overcome this computational challenge, we then derived a macroscopic PDE for the cell density from the hybrid model, using asymptotic analysis and moment closure methods. Parameters of the PDE model are fully determined using parameters of the hybrid model. Numerical comparisons of the hybrid model and the PDE model showed quantitative agreement in 1D under biologically-relevant parameter regimes. This justifies using the PDE model

as a quantitative and predictive tool to understand the relation between population patterning and cellular dynamics.

We then used our PDE model to investigate how concentric stripe patterns change when cells are subject to other chemicals or mutations as discussed in [12]. Numerical simulations of our PDE model in 2D with radial symmetry agree with experimental data semi-quantitatively. Finally, we used our PDE model to make a number of predictions on how stripe formation depends on cell-level parameters. Specifically, we investigated how the colony front speed, the wavelength of the spatial pattern and the structure within a single spatial element depend on the individual cell speed, cell doubling time as well as the rate of CheZ turnover. Our simulations suggested that the individual cell speed and the cell doubling time primarily affect the colony front speed and the pattern wavelength, while the the turnover rate of CheZ mainly affects the spatial structure of each stripe.

Methods

We describe two mathematical models: hybrid model and PDE model in this section. The structure of these models and their relation are shown in Fig 2. The initial and boundary conditions are described in Results.

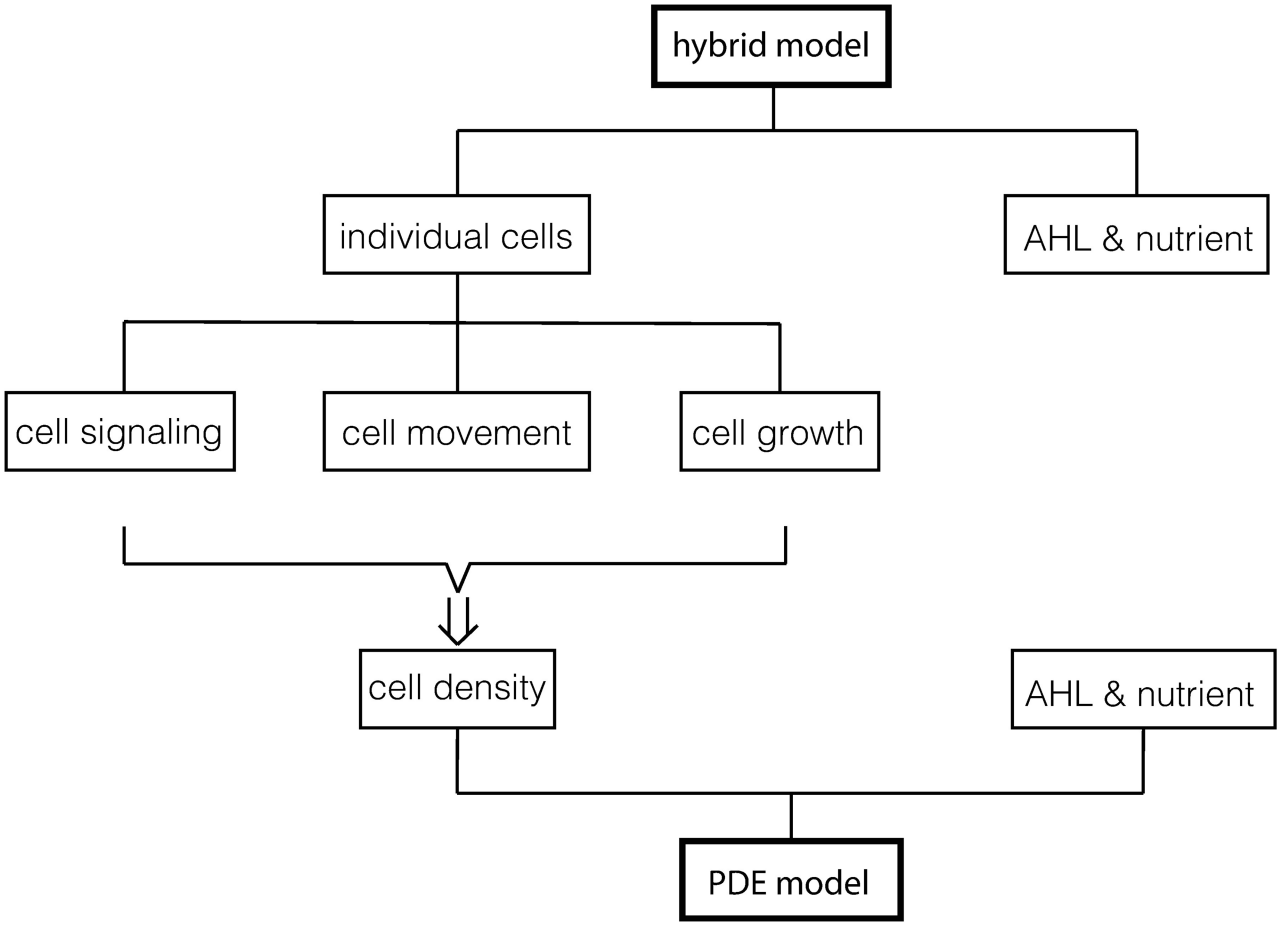

Fig 2. The structure of the hybrid model and the PDE model. The hybrid model includes an individual-based description for the cell dynamics, including intracellular signaling, cell movement and cell growth. The cell density equation (Eq 17) in the PDE model is derived from the probability density distribution for individual cells. The derivation is included in S2 Text. Both models use reaction-diffusion equations for the concentrations of the quorum-sensing signal AHL denoted by $h(\mathbf{x}, t)$ and nutrient denoted by $n(\mathbf{x}, t)$.

https://doi.org/10.1371/journal.pcbi.1006178.g002

Hybrid model

Each cell is described as an individual particle with location $\mathbf{x}^i$, velocity $\mathbf{v}^i$, and internal state $\mathbf{y}^i$. The superscript i is the index for the cell. Cell signaling is modeled by an internal ODE system for $\mathbf{y}^i$. Cell movement is modeled by a velocity jump process in which transition rates are functions of $\mathbf{y}^i$. Cell growth is implemented by random creation of new daughter cells from mother cells. The cell dynamics is then coupled with reaction-diffusion equations for $h(\mathbf{x}, t)$ and $n(\mathbf{x}, t)$. A similar type of model was used to model pattern formation in the slime mold *Dictyostelium discoideum* in [14]. Details of each component are given below. For simplicity of notation, we omitted the superindex i below.

Intracellular signaling described by a system of ODEs. We adopt the mathematical model for *E. coli* chemotactic signaling described in [15], and extend it to include the temporal dynamics of total *CheZ* protein in response to extracellular *AHL*. The model in [15] is a simplified form of the model derived in [16]. These models are based on the detailed biochemistry of the signaling network and are derived rigorously from mass action kinetics and asymptotic analysis.

We denote the total concentration of CheZ protein, i.e., the sum of CheZ and $\mathrm{CheZ_p}$, by $z(t)$. The dynamics of z is governed by protein production due to transcription and translation as well as dilution due to cell growth. Let $V(t)$ be the volume of a cell and k_V be its growth rate. Then between cell divisions we have $V' = k_V V$. If there is no production of CheZ protein, Vz remains constant and

$$z' = -V'z/V = -k_V z.$$

If we assume that the production rate of *CheZ* is constant in wild-type cells, then

$$z' = k_V(Z_w - z),$$

where Z_w is the steady state of z. For the stripe-forming cells created in [12], high concentration of *AHL* can indirectly suppress the transcription of *cheZ*, resulting in a sharp decrease of *CheZ* (Fig 1C). To model this effect, we take

$$z' = g(z,h) = \begin{cases} k_V(Z_w - z), & \text{if} \quad h < h_0, \\ -k_V z, & \text{if} \quad h \geq h_0, \end{cases} \tag{1}$$

where h_0 is the threshold AHL level for the suppression of *cheZ*. Our simulations show similar results when a smooth interpolation of $g(z, h)$ was used instead, but the transition of the two states has to be sharp.

We next couple *CheZ* dynamics with the rest of the chemotactic signaling pathway. Denote the mean methylation level of the chemoreceptors by m. Based on [15], the equation of m is governed by the methylation and demethylation reactions mediated by CheR (R) and CheB_p (B_p) as

$$\frac{\mathrm{dm}}{\mathrm{dt}} = f(m,z) = k_R R(1 - A(m)) - k_{Bp} B_p A(m). \tag{2}$$

Here $A(m)$ is the mean receptor activity:

$$A(m) = \frac{1}{1 + \exp\left[N_r \alpha_0 (m_0 - m)\right]}, \tag{3}$$

where $m_0 = 1$ is the reference methylation level, $\alpha_0 = 1.7$ measures how the free energy of the receptor complex depends on m, and $N_r = 6$ is the average number of nearest neighbors of the receptor functioning units. Note that $A(m)$ increases with m and has a sharp transition near $m = m_0$. In general, A is a function of both m and the ligand binding state of the receptors

(see Eqn. 5.12 in [15]). The function A reduces to Eq (3) if there is no ligand-binding signal involved as in this context. *CheR* concentration R is given by

$$R = \frac{R_t}{1 + K_R T_t (1 - A(m))}. \tag{4}$$

CheB$_p$ concentration B_p is implicitly given by a system of algebraic equations of B_p, Y_p (concentration of CheY$_p$) and T_p (concentration of CheAp-associated receptors), as

$$\begin{aligned} k_A(T_t A(m) - T_p) - k_Y Y T_p - k_B B T_p &= 0, \\ k_Y Y T_p - \mu_Y Y_p - k_Z Z Y_p &= 0, \\ k_B B T_p - \mu_B B_p &= 0, \end{aligned} \tag{5}$$

with

$$Y = \frac{Y_t - (1 + K_Z Z) Y_p}{1 + K_Y T_p}, \quad Z = \frac{z}{1 + K_Z Y_p}, \quad B = \frac{B_t - (1 + K_{B_p} T_t A(m)) B_p}{1 + K_B T_p}. \tag{6}$$

Here the constants T_t, Y_t, B_t and R_t are the total concentrations of the corresponding proteins; $z \equiv Z_t$ is total CheZ given by Eq (1); and the k's, K's, and μ's are reaction rate constants. Eq (5) are derived from quasi-steady state approximations of B_p, Y_p and T_p respectively, based on the fact that the chemical reactions involving B_p, Y_p and T_p occur on a time scale much faster than the methylation and demethylation of the receptors. Eq (6) are derived from the conservation conditions [15, 16].

In summary, the ODE model for each cell is a pair of differential equations for the internal states $y = (z, m)$ (CheZ concentration and receptor methylation level) coupled with several nonlinear equations for the intracellular chemotaxis protein concentrations. The parameter values for the model are summarized in Table 1. The parameters involved in chemotactic signaling are taken from [15, 16], please see references therein.

Cell movement as a velocity-jump process with moving and resting states. *E. coli* cells move by alternating between two movement states, running and tumbling. The speed of running is about $10 - 30\mu$m/s, and during tumbling they stop immediately with almost no displacement. For wild-type cells, the average tumbling time ($0.1s$) is much shorter than the average running time ($1s$). For this reason, the tumbling time is frequently ignored in previous models [17, 18]. However, for the engineered cells we consider here, intracellular CheZ can be significantly lower than that of wild-type cells, leading to significantly longer tumbling. In this case cell tumbling cannot be naively ignored.

Based on the above considerations, we describe the movement of each cell as an independent velocity jump process with a moving state and a resting state. We assume cells can move in any directions with constant speed $s_0 = 20\mu$m/s. We denote the rate for a running bacterium to stop and tumble by λ and the rate for a tumbling cell to start running by μ, i.e.,

$$\text{Run} \underset{\mu}{\overset{\lambda}{\rightleftharpoons}} \text{Tumble}.$$

We further assume that a cell chooses a new direction randomly with equal probability after tumbling. Adding a slight directional persistence as observed in [19, 20] does not alter the main results of the paper.

The turning rates λ and μ depends on the intracellular CheYp level (Y_p). We determine these rates using the following method that involves a voting process [21].

Table 1. Parameters for intracellular signaling. See [15, 16] for references.

Param.	Description	Values
k_R	Methylation rate mediated by CheR	$3.82 \times 10^{-2}\, s^{-1}$
k_{B_p}	Demethylation rate mediated by $CheB_p$	$3.25 s^{-1}$
k_A	Phosphorylation rate of T_p mediated by CheA	$100 s^{-1}$
k_Y	Dephosphorylation rate of T_p mediated by CheY	$130 \mu M^{-1}\, s^{-1}$
k_B	Dephosphorylation rate of T_p mediated by CheB	$7.5 \mu M^{-1}\, s^{-1}$
k_Z	Dephosphorylation rate of Y_p mediated by CheZ	$8.45 \mu M^{-1}\, s^{-1}$
μ_Y	Degradation rate of Y_p	$0.1 s^{-1}$
μ_B	Degradation rate of B_p	$1 s^{-1}$
K_B	Association constant for CheB phosphorylation	$0.25 \mu M^{-1}$
K_{B_p}	Association constant for receptor demethylation	$6.5 \mu M^{-1}$
K_R	Association constant for receptor methylation	$0.15 \mu M^{-1}$
K_Y	Association constant for CheY phosphorylation	$0.65 \mu M^{-1}$
K_Z	Association constant for CheYp dephosphorylation	$1 \mu M^{-1}$
B_t	total concentration of CheB	$2 \mu M$
R_t	total concentration of CheR	$0.3 \mu M$
T_t	total concentration of CheT	$5/3 \mu M$
Y_t	total concentration of CheY	$18 \mu M$
Z_w	CheZ concentration of wild type *E. coli*	$1.23 \mu M$
h_0	AHL threshold (nondimensional)	0.25

https://doi.org/10.1371/journal.pcbi.1006178.t001

An *E. coli* cell has several flagella and each flagellar motor can rotate either clockwise (CW) or counter-clockwise (CCW). We denote the switching rates from CCW to CW by λ_f and from CW to CCW by μ_f, i.e.,

$$\mathrm{CCW} \underset{\mu_f}{\overset{\lambda_f}{\rightleftharpoons}} \mathrm{CW}.$$

These rates have been estimated using experimental data [21, 22],

$$\lambda_f = a_1 \exp(b_1 Y_p), \tag{7}$$

$$\mu_f = a_2 \exp(-(b_2 - Y_p)^4 / c), \tag{8}$$

where a_1, b_1, a_2, b_2 and c are constants specified in Table 2. The fitting is replotted in S1 Text for readers' convenience. We note that given z and m for each cell, $Y_p = Y_p(m, z)$ can be solved from (5) and used to determine λ_f and μ_f.

We assume that each cell has n_f flagella that rotate independently. If at least w flagella rotate CCW simultaneously then the cell runs forward; otherwise it tumbles in place. In reality different flagella may interact with each other through the surrounding fluid, but we ignore this effect for simplicity.

The probability of having exactly i flagella rotating CCW is given by

$$P^i_{CCW} = \binom{n_f}{i} \left(\frac{\mu_f}{\lambda_f + \mu_f}\right)^i \left(\frac{\lambda_f}{\lambda_f + \mu_f}\right)^{n_f - i}. \tag{9}$$

Table 2. Parameters for cell movement and cell growth.

Param.	Description	Value
s_0	cell speed	$0.02\text{mm} \cdot \text{s}^{-1}$
a_1	coefficient in Eq (11)	$0.0174001\ s^{-1}$
b_1	coefficient in Eq (11)	$1.32887\ \mu M^{-1}$
a_2	coefficient in Eq (12)	$12.0809\ s^{-1}$
b_2	coefficient in Eq (12)	$-5.83762\ \mu M$
c	coefficient in Eq (12)	2892.12
n_f	total flagella number	8
w	minimum number of CCW flagella needed to run	6
r	cell growth rate	$3.85 \times 10^{-4}\text{s}^{-1}$

https://doi.org/10.1371/journal.pcbi.1006178.t002

The probability for the cell to be in the run and tumble states are given by

$$P_{run} = \sum_{i=w}^{n_f} P_{CCW}^{i}, \quad P_{tumble} = 1 - P_{run}. \tag{10}$$

The probability for two flagella switching rotation simultaneously is very small; therefore the switch from run to tumble primarily occurs when the cell has exactly w flagella rotating CCW and one of them switches to CW direction. Based on these observations, λ can be estimated as

$$\lambda(m, z) = w\lambda_f \cdot \frac{P_{CCW}^{w}}{P_{run}}. \tag{11}$$

Similarly, a cell switches from tumble to run primarily when there are exactly $w - 1$ flagella rotating CCW at that moment and one of the rest switches to CCW. This argument leads to

$$\mu(m, z) = (n_f - w + 1)\mu_f \cdot \frac{P_{CCW}^{w-1}}{P_{tumble}}. \tag{12}$$

Using $n_f = 8$, $w = 6$ and total CheZ concentration $1.23\mu M$, we obtain $\lambda = 0.594s^{-1}$ and $\mu = 6.1143s^{-1}$ at basal CheYp level, which is consistent with experimental data for wild-type cells (See S1 Text).

Cell growth. We assume that the growth rate of the cells is a linear function of the local nutrient concentration $n(\mathbf{x}, t)$, i.e.,

$$k_V(\mathbf{x}, t) = rn(\mathbf{x}, t). \tag{13}$$

Furthermore we model cell proliferation as a Poisson process, i.e., during the time interval $[t, t + dt)$, the probability for a cell to divide into two daughter cells is $k_V(\mathbf{x}, t)dt$. An alternative approach is to introduce a cell cycle variable for each cell and divide it when it doubles in size [23]. We tested both approaches numerically and found no visible difference. For this reason, we used the Poisson process approach in this paper for the ease of mathematical analysis in S2 Text.

We assume that the cell doubling time is approximately 30 minutes at maximum nutrient level and regard $n(\mathbf{x}, t)$ as the nutrient concentration normalized by its initial value. Therefore we have $r = ln2/(30\text{min}) \approx 3.85 \times 10^{-4}\text{s}^{-1}$ (Table 2).

AHL and nutrient dynamics. Denote the total number of bacteria at time t as n_b. We assume that AHL is secreted by each cell with a constant rate α_d and degrades naturally. We further assume that the nutrient is consumed by each cell at a rate proportional to the nutrient

Table 3. Parameters for AHL and the nutrient.

Param.	Description	Value
D_h	diffusion coefficient of *AHL*	$5 \times 10^{-4}\text{mm}^2 \cdot \text{s}^{-1}$
D_n	diffusion coefficient of the nutrient	$7.7 \times 10^{-4}\text{mm}^2 \cdot \text{s}^{-1}$
α_d	production rate of AHL per cell	$10^{-6}\, s^{-1}$
β	degradation rate of AHL	$10^{-3}\, s^{-1}$
γ_d	consumption rate of nutrient per cell	$1.155 \times 10^{-6}\, s^{-1}$
n_0	Initial concentration of nutrient (non-dimensional)	1

https://doi.org/10.1371/journal.pcbi.1006178.t003

concentration. Based on these assumptions, we have

$$\begin{aligned} \partial_t h(\mathbf{x}, t) &= D_h \Delta h(\mathbf{x}, t) + \alpha_d \sum_{i=1}^{n_b} \delta(\mathbf{x} - \mathbf{x}_i) - \beta h(\mathbf{x}, t), \\ \partial_t n(\mathbf{x}, t) &= D_n \Delta n(\mathbf{x}, t) - \gamma_d \sum_{i=1}^{n_b} \delta(\mathbf{x} - \mathbf{x}_i) n(\mathbf{x}, t). \end{aligned} \tag{14}$$

The parameters for AHL and nutrient dynamics are listed in Table 3. Since different experimental conditions can lead to different parameter values which were not reported in [12], they effect the patterns and we will explore the parameters more later on.

PDE model

To reduce computational cost, we derived a PDE model from the hybrid model using moment closure methods and asymptotic analysis. Let $p(\boldsymbol{x}, \boldsymbol{v}, m, z, t)$ be the density of cells at position $\boldsymbol{x}$, with velocity v, internal states m and z, and at time t. Let $p_0(\boldsymbol{x}, m, z, t)$ be the density of cells resting at position $\boldsymbol{x}$ with internal states m and z. According to the hybrid model we have

$$\begin{aligned} \partial_t p + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} p + \partial_z (g(z, h) p) + \partial_m (f(m, z) p) &= Q(p, p_0), \\ \partial_t p_0 + \partial_z (g(z, h) p_0) + \partial_m (f(m, z) p_0) &= Q_0(p, p_0). \end{aligned} \tag{15}$$

Here $g(z, h)$ and $f(m, z)$ are the right-hand sides of Eqs (1) and (2), and

$$\begin{aligned} Q(p, p_0) &= -\lambda(m, z) p + \mu(m, z) p_0 / |V| + rnp, \\ Q_0(p, p_0) &= \lambda(m, z) \int_V p d\boldsymbol{v} - \mu(m, z) p_0 + rnp_0, \end{aligned} \tag{16}$$

where $V = s_0 \partial B_0^1$, $\lambda(m, z)$, $\mu(m, z)$ are given by Eqs (11) and (12), and $n = n(\mathbf{x}, t)$ is the local nutrient concentration. The first two terms in $Q(p, p_0)$ and $Q_0(p, p_0)$ represent the density change due to velocity jumps and the third terms are due to cell growth.

Let $\rho^z(\mathbf{x}, z, t)$ be the density of cells at position $\mathbf{x}$ with internal state z, then

$$\rho^z = \int_{\mathbb{R}} \left(p_0 + \int_V p \mathrm{d}\mathbf{v} \right) \mathrm{dm}.$$

We derived the following approximating equation for $\rho^z(\mathbf{x}, z, t)$ from (15) and (16) (see S2 Text),

$$\partial_t \rho^z = \nabla_{\mathbf{x}} \cdot (D(z) \nabla_{\mathbf{x}} \rho^z) - \partial_z (g(z, h(\mathbf{x}, t)) \rho^z) + rn\rho^z. \tag{17}$$

Here h is the AHL concentration and

$$D(z) = \frac{s_0^2 \mu_0(z)}{d\lambda_0(z)[\mu_0(z) + \lambda_0(z)]}, \tag{18}$$

where d is the space dimension, and $\lambda_0(z)$ and $\mu_0(z)$ are the switching frequencies when m equals its quasi-steady state. We note that the intracellular chemotactic signaling enters into Eq (17) through the quasi-steady state of m only. This is because the methylation time scale is much smaller than the time scale for the change of z.

The derivation was based on time scale separation of the intrinsic biological processes: the time scale for chemotactic signaling is seconds to minutes, the time scale for CheZ dynamics is tens of minutes, and the time scale for the stripe formation is several hours. The derivation involves moment closure methods and asymptotic analysis, similar to our previous works [15, 17, 18, 24].

The PDE model is formed by coupling Eq (17) with the continuous version of (14), namely,

$$\begin{aligned} \partial_t h(\mathbf{x}, t) &= D_h \Delta h(\mathbf{x}, t) + \alpha \rho(\mathbf{x}, t) - \beta h(\mathbf{x}, t), \\ \partial_t n(\mathbf{x}, t) &= D_n \Delta n(\mathbf{x}, t) - \gamma \rho(\mathbf{x}, t) n(\mathbf{x}, t), \end{aligned} \tag{19}$$

where

$$\rho(\mathbf{x}, t) = \int \rho^z(\mathbf{x}, z, t)\, \mathrm{d}z.$$

The parameters of the PDE model are fully determined by those of the hybrid model. In our simulations we choose the cell density scale to be $\rho_s = 1000\text{cells} \cdot \text{mm}^{-1}$. As a consequence, α and γ can be calculated as $\alpha = \alpha_d \rho_s$ and $\gamma = \gamma_d \rho_s$.

Results

How does single cell dynamics depend on total CheZ

We first investigated how intracellular signaling and cell movement depend on the total concentration of CheZ protein (denoted as Z_t in this section).

Dependence of intracellular signaling on total CheZ. CheYp is the intracellular protein that binds to a cell's flagellar motor and changes its rotation (Fig 1B). CheYp concentration (Y_p) depends on total CheZ (Z_t) and the methylation state of the receptors (m) in a ultra-sensitive manner. The relation was solved from (5) and (6) and plotted in Fig 3A. Total CheZ in a cell changes on a much slower time scale (tens of minutes) than the methylation and demethylation of the cell receptors (several seconds to minutes). As a result, in the absence of external receptor-binding signals, m and Y_p are close to their steady states m^* and Y_p^*. Figs 3B and 2C plot m^* and Y_p^* obtained using the ODE model (1)–(6). We note that these variables vary significantly with Z_t. This implies that cells can demonstrate chemotactic-like behavior if Z_t changes along their trajectories.

Dependence of cell movement on total CheZ. Assuming $m = m^*$, we calculated the turning rates for a single flagellum (λ_f, μ_f) as well as for the whole cell (λ, μ) as a function of Z_t. Fig 4A shows that λ_f and λ decrease with Z_t, while μ_f and μ increase with Z_t. The mean time fraction that a single flagella spends in the CCW rotation is given by $\mu_f/(\lambda_f + \mu_f)$, and the mean time fraction that a cell spend in the running state is $\mu/(\lambda + \mu)$. Fig 4B shows that these quantities are very sensitive to Z_t and decrease significantly if Z_t is reduced: cells spend 90% of their time running if $Z_t = Z_w \approx 1.23\mu M$ but only 25% if Z_t is reduced to 1.11 μM. Moreover, the velocity jump of a cell is more sensitive to Z_t than the rotation direction change of a single

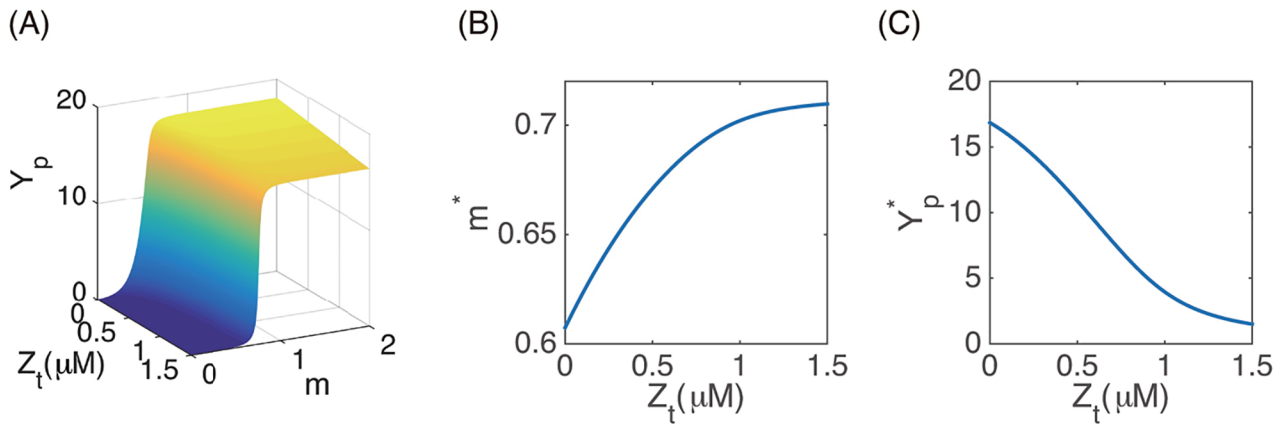

Fig 3. Dependence of intracellular signaling on total CheZ. (A) Y_p as a function of Z_t and m. (B) the stationary methylation level m^* as a function of Z_t; (C) the stationary CheYp concentration Y_p^* given different Z_t. Parameter values are the same as in Table 1.

https://doi.org/10.1371/journal.pcbi.1006178.g003

flagellum due to the cooperative behavior of different flagella. These results reconfirm the significance of the tumbling state in periodic stripe formation.

Fig 4C plots the effective diffusion coefficient for cell movement in 1D calculated using Eq (18). Experimental measurements in [12] showed that the cell diffusion rate decreases sharply with increasing cell density. Our calculation is consistent with experimental data and gives an explanation of this relation at the molecular level: high cell density is associated with high AHL concentration, which shuts off the production of CheZ inside the cells. This in turn causes Z_t to decrease and cells tumble in place extensively.

Quantitative agreement between hybrid and PDE models in 1D

If cells are initially seeded on a horizontal line in an agar plate, they will grow, spread out laterally and form straight stripes of equal spacing (see Fig. S4 of [12]). Motivated by these experiments, we first investigated the population pattern formation on a 1D domain $[-L, L]$, representing a cross-section of the stripe patterns, using the hybrid model and the PDE model. Simulations suggest that both models predict the same spatial-temporal population dynamics

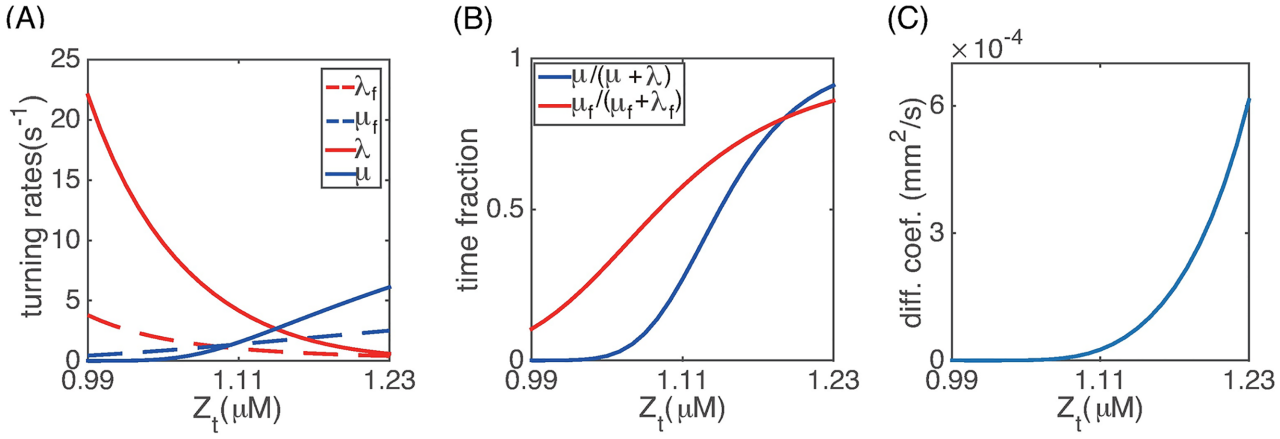

Fig 4. (A) Dependence of the turning rates on Z_t; (B) Mean time fraction of CCW rotation and cell running as a function of Z_t. (C) Effective cell diffusion coefficient calculated using Eq (18).

https://doi.org/10.1371/journal.pcbi.1006178.g004

for the engineered stripe-forming mutants as well as wild-type cells as in experiments. Moreover, the derived PDE model agrees with the hybrid model quantitatively in biologically-relevant parameter regimes.

To mimic the experimental setup, we assumed that all cells initially cluster near the center ($x = 0$) with internal states at equilibrium, i.e., $z = Z_w$ and $m = m_0$. Specifically, for the hybrid model, we randomly put 500 cells in the domain according to the distribution

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

with $\sigma = 2$mm at $t = 0$. Correspondingly, for the PDE model, we took

$$\rho^z(x, z, 0) = \rho_0 P(x)\delta(z - Z_w),$$

where $\rho_0 = 500\text{cells} \cdot \text{mm}^{-1}/\rho_s = 0.5$ ($\rho_s = 1000\text{cells} \cdot \text{mm}^{-1}$ is the cell density scale). For both models, we took the initial nutrient concentration to be a constant everywhere and assumed that there was no AHL added in the domain, i.e.,

$$n(x, 0) = 1, \quad h(x, 0) = 0.$$

We used no-flux boundary conditions throughout the paper. For AHL and nutrient concentrations, we imposed $\nabla h \cdot \mathbf{n} = \nabla n \cdot \mathbf{n} = 0$ at the boundary of the spatial domain, where $\mathbf{n}$ is the outward normal vector. For individual cell movement, we assumed that once a cell reaches the boundary, it bounces back with its velocity reflected by the boundary. In 1D, the cell direction simply reverses. For Eq (17), we chose the computational range $z \in [z_{min}, z_{max}]$ to be large enough to include all possible CheZ concentrations such that $\rho^z(\mathbf{x}, z_{min}, t) = \rho^z(\mathbf{x}, z_{max}, t) = 0$. In the $\mathbf{x}$ direction, we imposed that $\nabla\rho^z \cdot \mathbf{n} = 0$ for all z.

We first simulated the cell population dynamics for the stripe-forming mutants with parameters specified in Tables 1–3. Fig 5 presents the time course data of the cell density as well as the distribution of the internal variable Z_t. Panels A and B are the heat maps of the cell density as a function of space and time. The normalized cell density for the PDE model was obtained by integrating ρ^z over z. The normalized cell density for the hybrid model was calculated using histograms of the cell positions. Panels C-F present the detailed comparisons of the normalized cell density in space (top) as well as the Z_t distribution (bottom) given by the two approaches at different time points. The Z_t distribution was obtained by normalizing the cell number in each rectangular grid with size 0.1mm × 0.03μM by 100 cells. In these simulations, the AHL concentration h also shows the same stripe pattern as the cell density ρ, with peaks and valleys coinsiding with those of ρ; while the nutrient forms a wave front at the colony front, increasing from 0 to the initial normalized state.

Fig 5 shows that the total cell number grows significantly as cells divide. As the colony grows, it propagates outward continuously with a more or less constant front speed. Meanwhile cells produce AHL continuously. As extracellular AHL concentration becomes high and reaches the threshold h_0 locally, the intracellular Z_t at these locations start to drop. As a result, cells at these locations spend more time in the tumbling stage and become less mobile. In contrast, cells in nearby regions with low AHL move more persistently until they migrate into a high AHL region. The existence of high and low mobility regions leads to the sequential establishment of high-density stripes behind the colony front, similar to experiments.

As a comparison, we then simulated the population dynamics for wild-type cells that do not secrete AHL, i.e., $\alpha_d = \alpha = 0$ (Fig 6). In this case, cells grow, consume nutrients, and the colony propagates outward with a constant wave front speed. However, stripes do not appear behind the colony front.

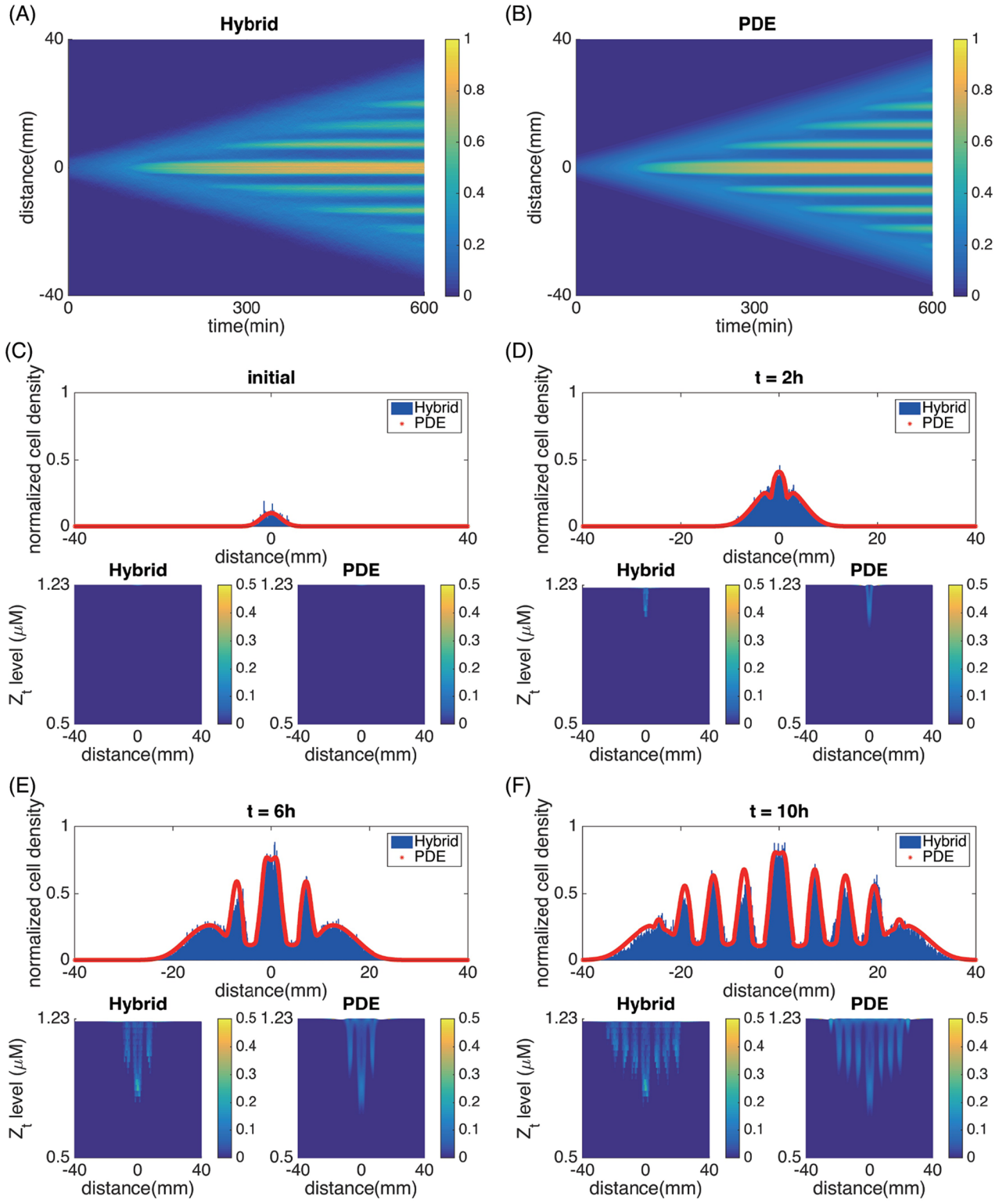

Fig 5. Sequential stripe formation in 1D for engineered mutants. (A) spatial-temporal evolution of the cell density recovered from averages of 6 realizations of the hybrid model. (B) the evolution of the cell density from the PDE model. (C)—(F): spatial plots of the cell density and Z_t distribution at $t = 0$h, $t = 2$h, $t = 6$h and $t = 10$h, respectively. Top: the normalized cell density $\rho(x, t)$. Blue bars are obtained from histograms using the hybrid model, averaged over 6 realizations. Red Lines are calculated from the PDE model. Bottom left: distribution of $\rho^z(x, z, t)$ in space and time from the hybrid model. Bottom right: $\rho^z(x, z, t)$ from the PDE model. The parameters used here are the same as in Tables 1–3.

https://doi.org/10.1371/journal.pcbi.1006178.g005

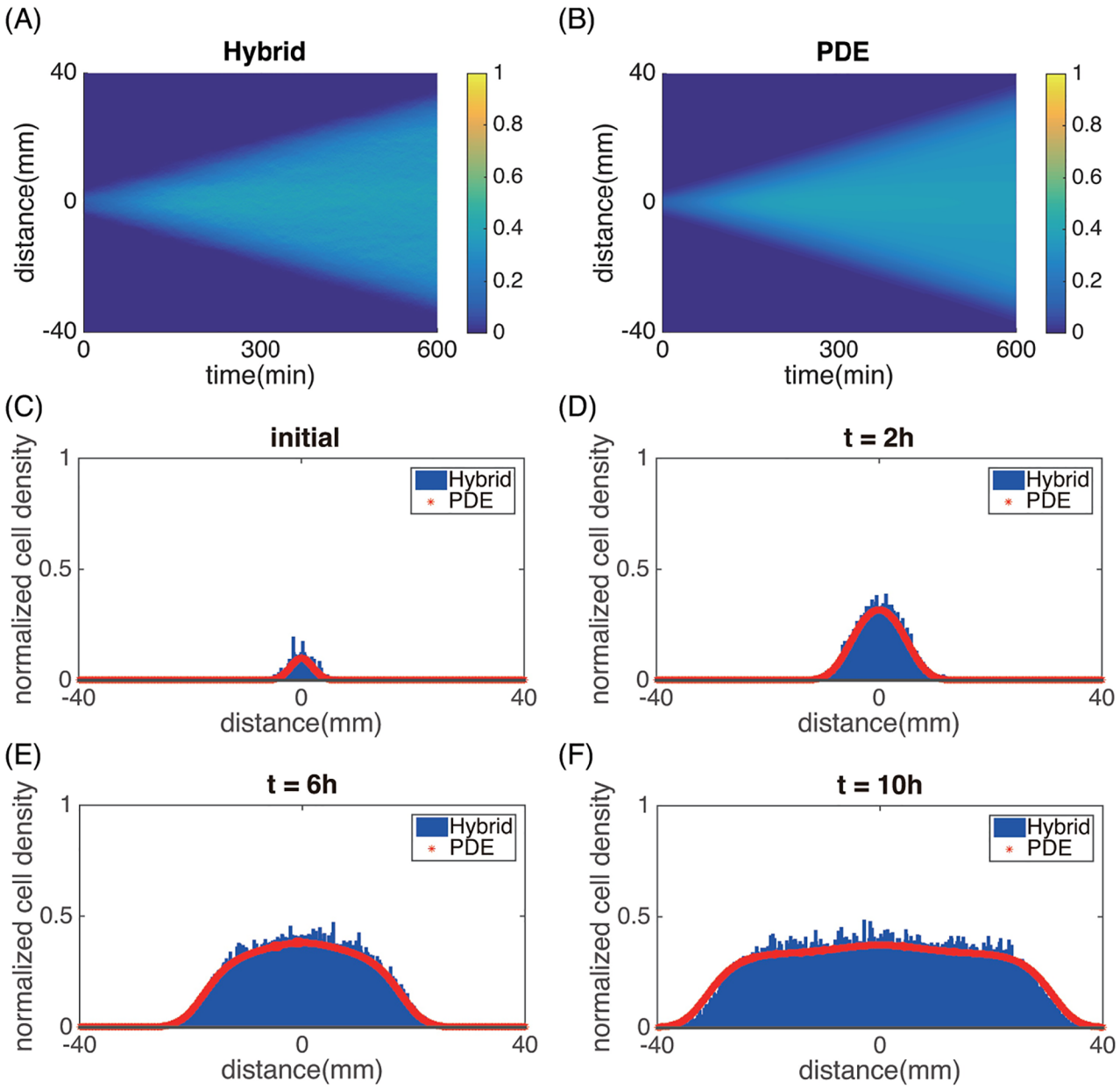

Fig 6. Uniform colony expansion for wild-type cells. (A) spatial-temporal evolution of the cell density recovered from averages of 6 realizations of the hybrid model. (B) The evolution of the cell density from the PDE model. (C)—(F) Spatial plots of the cell density at $t = 0$h, $t = 2$h, $t = 6$h and $t = 10$h, respectively. Blue bars are obtained from histograms using the hybrid model averaged over 6 realizations. Red Lines are calculated from the PDE model. In these simulations we used $\alpha_d = \alpha = 0$. Other parameters are the same as in Tables 1–3.

https://doi.org/10.1371/journal.pcbi.1006178.g006

Figs 5 and 6 suggest quantitative agreement between the hybrid model and the PDE model. This justifies using the PDE model for further investigations to save computational cost. We also note that the colony front expansion speed for both the engineered mutant and wild-type are identical. This is because the front speed is primarily determined by the growth and motility of cells at the colony front, where AHL does not reach the threshold h_0 required for quorum-sensing. Hence, cells therein have the wild-type phenotype for both cases.

Concentric stripe formation in 2D predicted by the PDE model

Stripe patterns on a plate. Using the PDE model, we next investigated how the concentric stripe pattern in Fig 1 forms from a single inoculant in the center of the petri dish. Assuming radial symmetry, the PDE model takes the following form

$$\begin{aligned}\partial_t \rho^z &= \frac{1}{\xi}\partial_\xi(D(z)\xi\partial_\xi\rho^z) - \partial_z(g(z,h(\xi,t))\rho^z) + rn\rho^z,\\ \partial_t h(\xi,t) &= \frac{D_h}{\xi}\partial_\xi(\xi\partial_\xi h(\xi,t)) + \alpha\rho(\xi,t) - \beta h(\xi,t),\\ \partial_t n(\xi,t) &= \frac{D_n}{\xi}\partial_\xi(\xi\partial_\xi n(\xi,t)) - \gamma\rho(\xi,t)n(\xi,t),\end{aligned} \tag{20}$$

where $\xi \in (0, R]$ is the polar coordinate. We imposed the following boundary conditions

$$\begin{aligned}&\partial_\xi\rho^z(0,z,t) = \partial_\xi\rho^z(R,z,t) = 0, \qquad \rho^z(\xi,0,t) = \rho^z(\xi,Z_{\max},t) = 0,\\ &\partial_\xi h(0,t) = \partial_\xi h(R,t) = 0, \qquad \partial_\xi n(0,t) = \partial_\xi n(R,t) = 0\end{aligned} \tag{21}$$

The initial conditions mimics the initial conditions in the experiments [12]

$$\rho^z(\xi,z,0) = \frac{1}{4\sqrt{2\pi}}\exp\left(-\frac{\xi^2}{8}\right)\delta(z-Z_w), \quad h(\xi,0) = 0, \quad n(\xi,0) = 1. \tag{22}$$

According to the experiments in [12], the initial cell colony expanded for several hours before its density reaches the threshold to form a stripe. Each stripe is composed of a high density part and a low density part with the average wavelength approximately 0.5cm. The average time to form one stripe is around 200mins (see Fig 1). Simulations of our PDE model with baseline parameters predict spatiotemporal dynamics that agrees with experiments semi-quantitatively (Fig 7).

Effect of CheZ inhibitor. In [12], the expression of cheZ gene was varied using an aTc-inducible module. As aTc level increases, the *CheZ* mRNA level decreases gradually. At high aTc level, the engineered *E. coli* colonies do not form the spatial stripe patterns as they grow. Motivated by these experiments, we investigated how the population pattern formation depends on the synthesis rate of CheZ protein. Specifically, we varied the parameter Z_w in Eq (1), which is the steady state of CheZ concentration in the absence of quorum-sensing effect.

Fig 8 shows that as Z_w decreases, the colony expansion rate decreases and the stripe patterns disappear, consistent with experimental data. In our simulations, small changes in Z_w induce large variances in the colony expansion rate. This is because the effective diffusion coefficient of the cells is very sensitive to the total CheZ level (Fig 4C). The experiments in [12] measured the correlation between *CheZ* mRNA level and the pattern formation. We note that the magnitudes of changes in CheZ protein and *CheZ* mRNA can be different in these experiments.

Effect of CheR CheB mutation. In [12], a secondary mutation was introduced to the engineered cell line by knocking out the *CheR* and *CheB* genes. The purpose was to investigate whether canonical chemotaxis induced by external ligands is a critical factor in the stripe pattern formation. Cells can still swim and tumble, but cannot adapt to chemotactic signals because the methylation level of the receptors are not modified. Experiments showed that these mutants can still form sequential stripe patterns, but there was little discussion on how the pattern differs from those formed by non-mutants (Fig S2(J) in [12]). The conclusion was that canonical chemotaxis is not necessary for the spatial pattern formation.

We investigated this aspect numerically using the PDE model (Fig 9). Knocking out *CheR*, *CheB* indicates that R, B_t and B_p are all zeros in (2)–(6). Thus according to our model, m

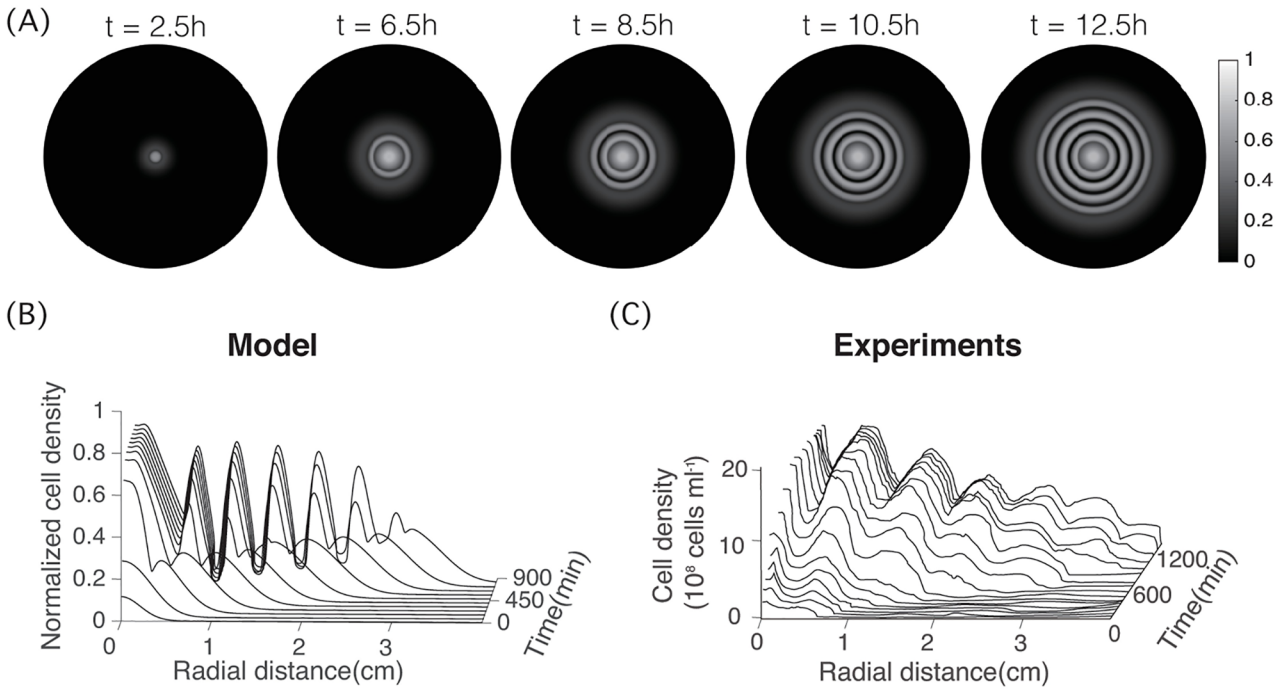

Fig 7. Concentric stripe patterns in 2D predicted by the PDE model. (A) spatial density plots at $t = 2.5$h, $t = 6.5$h, $t = 8.5$h, $t = 10.5$h and $t = 12.5$h. (B) Dynamics of the cell density in the radial direction predicted by the PDE model. (C) Dynamics of the cell density in the radial direction measured in experiments. (C) Reproduced from Fig. S3 in SOM of Liu et al, Science, Vol 334, 238–241, 2011 [12]. Parameters used here are either directly taken from Tables 1–3 or calculated using the conversion formulas.

https://doi.org/10.1371/journal.pcbi.1006178.g007

remains a constant which depends on the initial methylation level of the cells. Due to mutation, the cell turning rates and the effective cell diffusion coefficients are modified (A, B, C). The effective cell diffusion coefficient is very sensitive to the methylation level m (C). This is because the dependence of Y_p on the methylation level m has a sharp transition, similar to Fig 3A. Simulations of the corresponding PDE model showed that the colony expansion rate and the interior spatial pattern are both very sensitive to the initial methylation level of the cells:

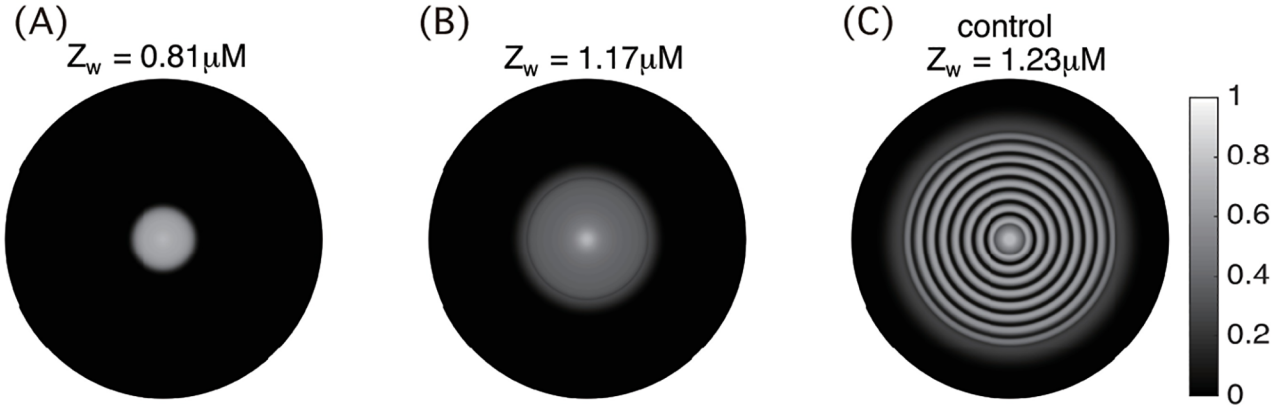

Fig 8. CheZ inhibitor disrupts the stripe pattern formation. The effect of CheZ inhibitor aTc is modeled by a reduction of total CheZ protein level Z_w. (A)–(C): cell density predicted by the PDE model at $t = 20$h with $Z_w = 0.81\mu M$, $Z_w = 1.17\mu M$ and $Z_w = 1.23\mu M$ (no inhibitor), respectively. Other parameters are the same as in Fig 7.

https://doi.org/10.1371/journal.pcbi.1006178.g008

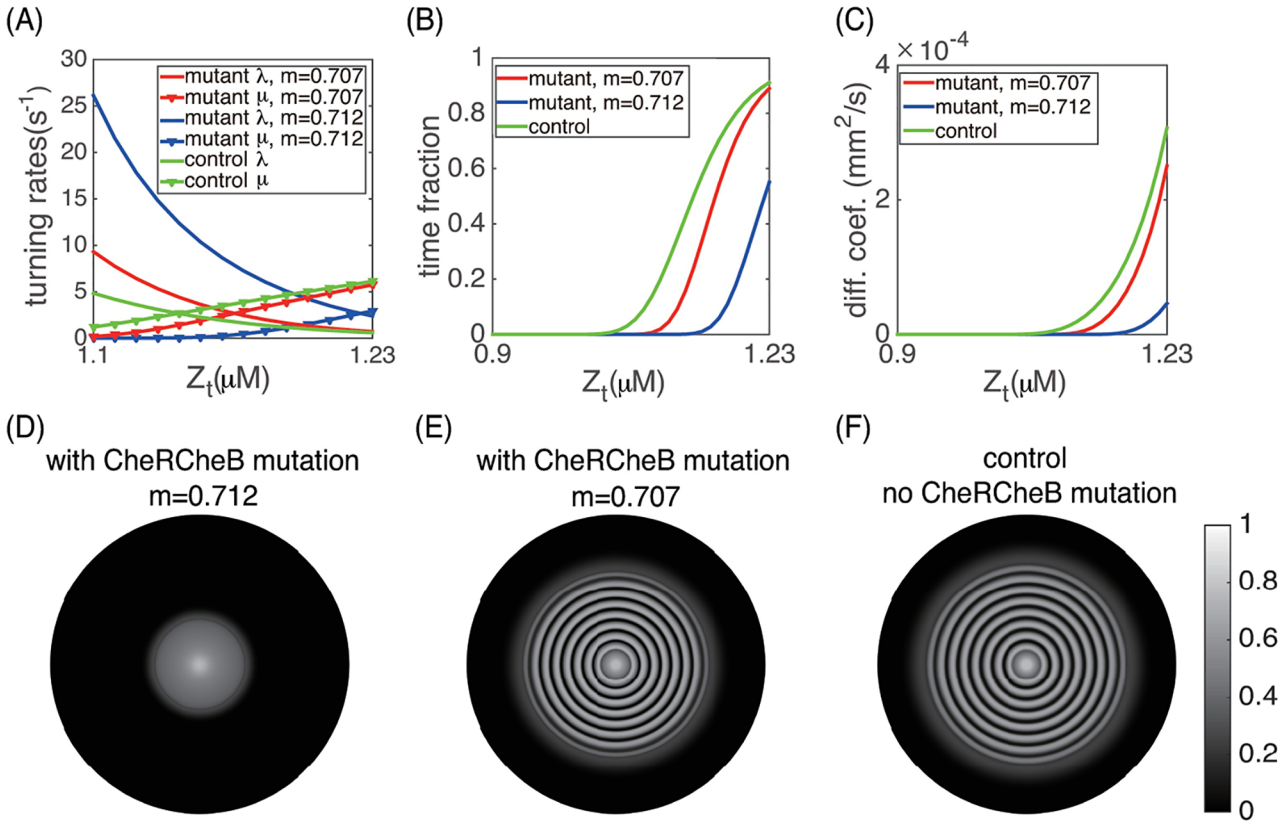

Fig 9. The effect of *CheR*$^{-1}$ *CheB*$^{-1}$ double mutation. (A) Dependence of cell turning rates on Z_t. (B) Mean time fraction that cell spends running as a function of Z_t, denoted as $\tilde{\mu}$. (C) Effective diffusion coefficients (Eq (18) for different Z_t. The methylation level m is kept constant for the mutant cells. (D) and (E): cell density plots at t = 20h for mutant cells with m = 0.712 and m = 0.707 respectively. (F) cell density plot for the engineered cells without secondary mutation as a comparison. Other parameters are the same as in Fig 7.

https://doi.org/10.1371/journal.pcbi.1006178.g009

small changes in m can lead to big differences in the population dynamics (D, E). In experiments, the methylation level of each bacterium also fluctuates due to internal noise. These considerations suggest that it is not justified to use this cell line to draw conclusions on whether canonical chemotaxis is involved in the pattern formation process.

How does the stripe formation depend on cell-level parameters

We use three important features to characterize the spatial-temporal pattern: the colony front propagation speed, the wavelength of the spatial stripes and the internal structure within a spatial period. We investigated how these features depend on intracellular dynamics, cell movement and cell growth.

We calculated the front speed as the average speed between t = 10 hr and 20 hr and the wavelength as the average distance between the maximum densities of two successive high density stripes (Fig 10A). To characterize the internal structure of the stripes, we defined the height ratio and the density ratio of the stripes (Fig 9B and 9C). The height ratio is the minimum density (h_2) divided by the maximum density (h_1) within a stripe. The density ratio is the volume of the shaded region over the region defined by the rectangle ABCD, factoring in the radially symmetric profile of the solution, i.e., $\int_B^C \xi\rho(\xi,t)/h_1 \mathrm{d}\xi$. The height ratio measures the

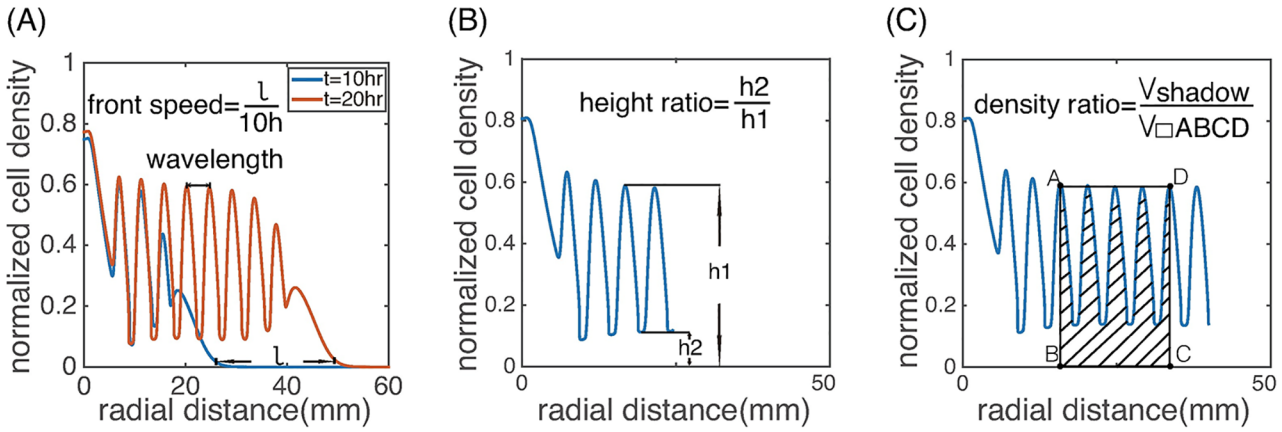

Fig 10. Schematic illustration of the front speed, wavelength, height ratio and density ratio.

https://doi.org/10.1371/journal.pcbi.1006178.g010

fluctuations of the cell density in the spatial pattern; while the density ratio quantifies the area fraction of the high-density regions when h_2 is small.

Dependence on cell speed and cell doubling time. We investigated how the stripe pattern depends on the individual cell speed (s_0) and the cell doubling time (log2/r) by varying the parameters s_0 and r in (18) and (20). Simulations of the PDE model show that as the cell speed increases, both the colony front speed and the pattern wavelength increase linearly (Fig 11A and 11B). If individual cells move faster near the colony front, the effective diffusion rate of the population becomes larger and thus the colony spreads out faster. Indeed, our simulation shows that the colony front speed is roughly proportional to the cell speed. As the cell doubling time increases, the front speed decreases, but the pattern wavelength increases, reflecting the slower growth of the total population size (Fig 11D and 11E). Our model predicts a linear dependence of the front speed and the pattern wavelength as one varies the cell speed (Fig 11B) or the cell doubling time (Fig 11E). The height ratio and the density ratio only decrease slightly with the cell speed and doubling time (Fig 11C and 11F).

Dependence on the CheZ turnover rate. In our models, we associated the turnover rate of CheZ protein with the cell growth rate k_V, see Eq (1). We next investigated the effect of a slower or faster CheZ turnover rate on the spatial patterning. To do that, we introduced a non-dimensional parameter κ preceding the z-flux term in Eq (17), i.e.,

$$\partial_t \rho^z = \frac{1}{\xi}\partial_\xi\left(D(z)\xi\partial_\xi\rho^z\right) - \kappa\partial_z\left(g(z,h(\xi,t))\rho^z\right) + r\rho^z n. \tag{23}$$

Mathematically speaking, κ parameterizes the convective speed in z, which corresponds to the response speed of intracellular CheZ to the external signal AHL: $\kappa = 1$ represents the baseline model, $\kappa > 1$ represents faster response, and $\kappa < 1$ represents slower response.

Fig 12A plots the cell density profiles at $t = 20$h with $\kappa = 0.1$, 0.6, 1, 3 and 10. For $\kappa = 0.1$, CheZ concentration in cells does not change much over the whole computational time, and as a result cells remain highly mobile and form no stripes. As κ increases, the spatial stripes appear and become increasingly more prominent. Interestingly, both the colony front speed and the wavelength of the spatial pattern do not change much given different κ (Fig 12B). In contrast, the height ratio decreases to 0 as κ increases, due to the increase of the maximum cell density in each stripe and the decrease of the minimum cell density (Fig 12C); while the density ratio shows a biphasic dependence on κ: it first decreases significantly as the height ratio decreases and then increases as the height ratio becomes close to 0 and κ increases further. The

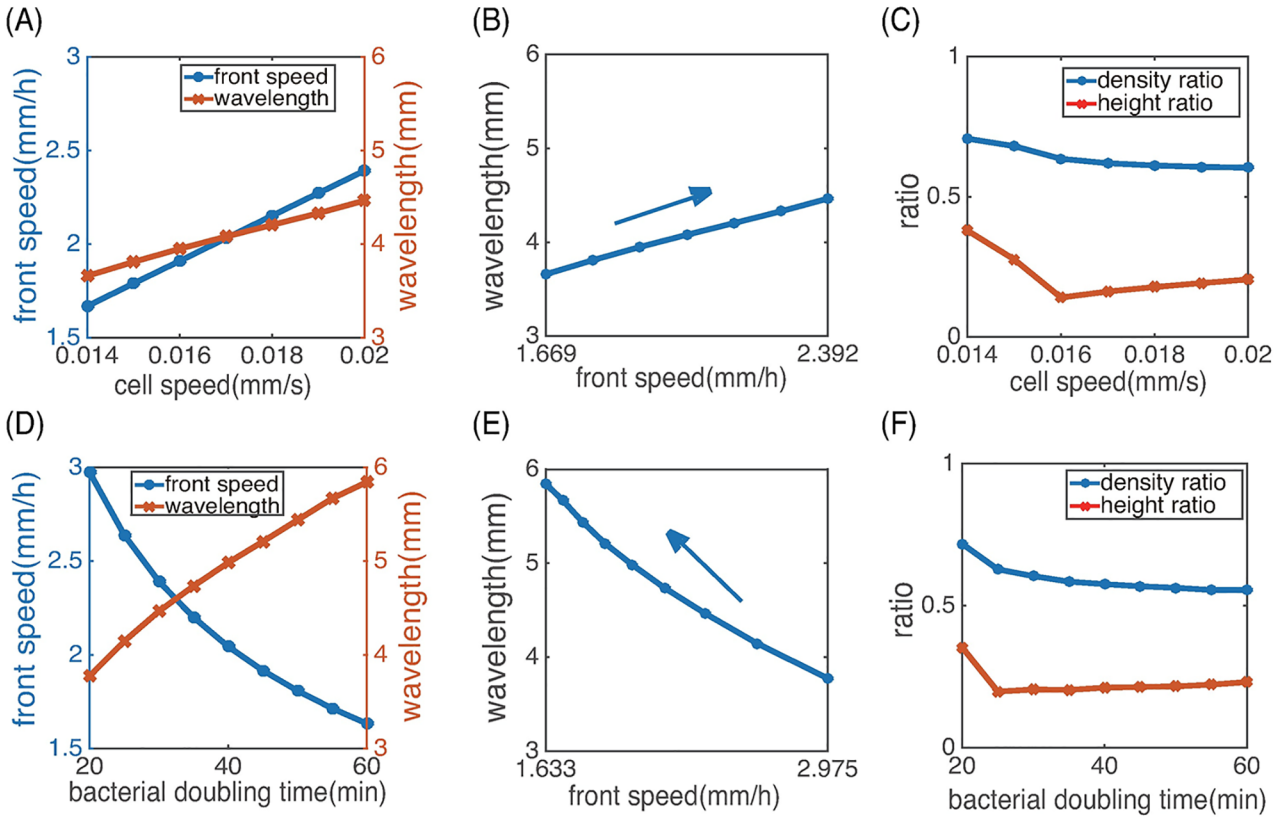

Fig 11. Dependence on individual cell speed and cell doubling time. (A), (D) Dependence of the front speed and the pattern wavelength on the cell speed and doubling time. (B), (E) The linear dependence between the pattern wavelength and the front speed. The cell speed or doubling time increases along the directions of the arrows. (C), (F) The height ratio and the density ratio as functions of the cell speed and doubling time. All other parameters are the same as in Fig 7.

https://doi.org/10.1371/journal.pcbi.1006178.g011

rebound of the density ratio is primarily due to the widening of the high-density region in a stripe. Finally, we note that when κ is small, the peak in each stripe has a more or less symmetric shape (Fig 12D, blue); while as κ becomes large, the peak shows a higher density at locations with smaller radii (Fig 12D, red).

Discussion

Synthetic biology has been used to design relatively simple systems to help understand how regularly-spaced structures form in nature. In [12], *E. coli* was engineered to couple chemotaxis and quorum sensing and these cells establish sequential stripe patterns when grown in semi-solid agar. In this paper, we developed multiscale models to help explain how these population patterns arise and predict their dependence on cell-level parameters.

We first developed a hybrid model that takes into account great details of intracellular signaling and movement of each individual cell. This model provides a method to connect cell-level dynamics and population-level behavior in a quantitative manner, but simulating it is very time-consuming as the cell number becomes large. To overcome this challenge, we mathematically derived a PDE model from our hybrid model. All the parameters of the PDE model can be calculated from measurable cell-level parameters used in the hybrid model. The PDE model matches the hybrid model quantitatively and is much more efficient in terms of computation. Our benchmark comparisons showed that the computation of the PDE model was over

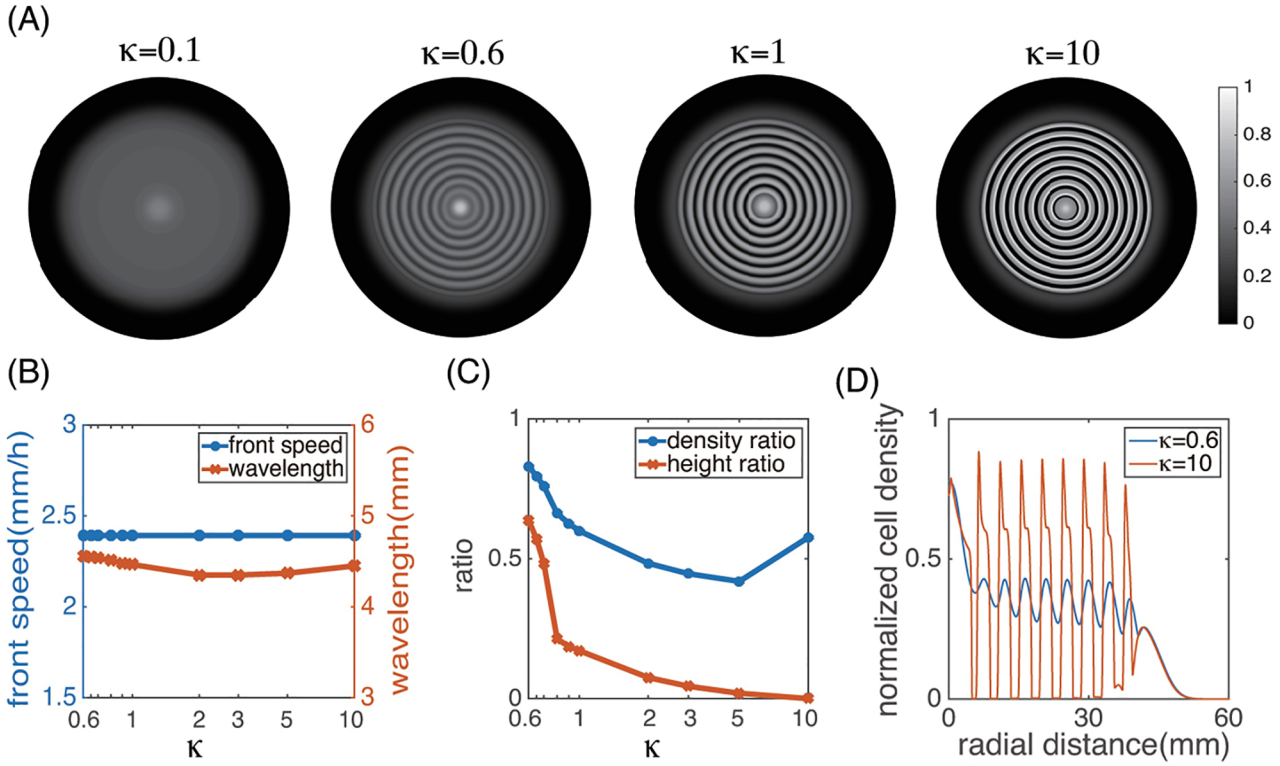

Fig 12. The cell density profile for different CheZ turnover rate κ. (A) Spatial density plots at t = 20h with κ = 0.1, 0.6, 1, 10, respectively. (B) Dependence of the front speed and wavelength on κ. (C) Dependence of the height ratio and density ratio on κ. (D) Plots of $\rho(\xi, 20)$ for κ = 0.6 and 10. All other parameters are the same as in Fig 7.

https://doi.org/10.1371/journal.pcbi.1006178.g012

100 times faster than that of the hybrid model. This justifies using the PDE model as a quantitative and predictive tool to explore the relation between population patterning and individual behavior.

Simulations of our models showed that the stripes arise sequentially due to suppression of CheZ in cells near the front of the expanding colony. At first, the self-secreted AHL reaches the threshold concentration for quorum sensing at these regions. This turns off the production of CheZ proteins in cells locally. The gradual drop of total CheZ inside these cells causes them to tumble excessively. As more and more cells move into these regions and get trapped, a high-density stripe develops. In the meantime, the colony grows and expands outward, and after some time, another high-density stripe establishes at a larger radius for the same reason. The self-trapping is due to the density-dependent suppression of motility, which has been studied before in [13, 25]. The model in [13] eliminates CheZ level by enslaving it to the AHL level, while the model in [25] directly links motility to the cell density, however both models are qualitative. The main contribution of our model is that it can not only reproduce the pattern, but also predict how the patterns varies when the individual cell signaling or movement changes. The spatial-temporal dynamics predicted by our simulations match experimental data *semi-quantitatively*.

We also made a number of predictions on the relation between the population patterns and cell level dynamics. Our simulations showed that the individual cell speed and the cell doubling time primarily affect the colony front speed and the wavelength of the stripe pattern (Fig 11). As the cell speed increases, the front speed and the pattern wavelength increases linearly. As the cell doubling time increases, the front speed decreases while the pattern wavelength

increases. Moreover, the turnover rate of CheZ protein does not alter the colony front speed and pattern wavelength, but changes the spatial structure of each stripe characterized by the height ratio and density ratio (Fig 12). These predictions can be tested by further experiments.

Our PDE model gives a detailed characterization of the anisotropic movement of the whole cell population in response to AHL. Cells with different intracellular CheZ concentration z have different mobility coefficient, given by $D(z)$ (Eq (18)). As a cell moves around, its internal state evolves with the extracellular environment. The change of z in each cell leads to the average mobility change of the whole population. We note that if z can be approximated by its steady state, which equals Z_w if $h < h_0$ and 0 otherwise, then Eq (17) can be "formally" reduced to the anisotropic diffusion model used in [12]

$$\partial_t \rho = \nabla_{\mathbf{x}} \cdot (\nabla_{\mathbf{x}} (\bar{D}(h)\rho)) + rn\rho. \tag{24}$$

where $\bar{D}$ is a step function of h. Specifically, we have $\rho^z(\mathbf{x}, t, z) = \rho(\mathbf{x}, t)Q(\mathbf{x}, t, z)$ with

$$Q(\mathbf{x}, t, z) = \begin{cases} \delta(z - Z_w) & h(\mathbf{x}, t) < h_0, \\ \delta(z) & h(\mathbf{x}, t) \geq h_0. \end{cases}$$

Integrating (17) with respect to z, one obtains Eq (24) with

$$\bar{D}(h) = \int D(z)Q(\mathbf{x}, t, z)\, \mathrm{d}z = \begin{cases} D(Z_w) & h < h_0 \\ D(0) & h \geq h_0. \end{cases}$$

However, during the stripe formation, CheZ turnover correlates with cell growth, which is much slower than single cell movement and intracellular signal adaptation. As a result, CheZ has a broad distribution among all cells and so it is far from its steady state (Fig 5). This suggest that it is important for models to take into account the internal state of cells individually rather than averaging it out.

We note that in this paper we used a multiscale modeling approach: start with a detailed, individual-based model for cell dynamics, then derive a PDE model and justify it using numerical simulations, and finally use the PDE model to make predictions on relations of phenomena at different scales. This multiscale approach allowed the macroscopic model to go beyond qualitative and can be used as a predictive tool. This type of multiscale modeling approach has also been used for classical bacterial chemotaxis [17, 21].

Supporting information

S1 Text. Parameter estimation for the turning rates.
(PDF)

S2 Text. Derivation of the PDE model.
(PDF)

S3 Text. Details of the numerical methods.
(PDF)

Acknowledgments

MT would like to thank Prof. Leihan Tang for many insightful discussions during her visits to HKBU. CX and MT would like to thank the Mathematical Bioscience Institute, Ohio State University for support as long-term visitors.

Author Contributions

Conceptualization: Chuan Xue, Min Tang.

Formal analysis: Xiaoru Xue, Min Tang.

Investigation: Chuan Xue, Min Tang.

Methodology: Xiaoru Xue, Chuan Xue, Min Tang.

Software: Xiaoru Xue.

Supervision: Chuan Xue, Min Tang.

Validation: Xiaoru Xue, Chuan Xue, Min Tang.

Visualization: Xiaoru Xue, Chuan Xue, Min Tang.

Writing – original draft: Xiaoru Xue, Chuan Xue, Min Tang.

Writing – review & editing: Chuan Xue, Min Tang.

References

1. Murray JD. Mathematical biology. vol. 2. Springer; 2002.
2. Turing AM. The Chemical Basis of Morphogenesis. Philosophical Transactions of the Royal Society B. 1952; 237(641):37–72. https://doi.org/10.1098/rstb.1952.0012
3. Baker RE, Schnell S, Maini PK. Waves and patterning in developmental biology: vertebrate segmentation and feather bud formation as case studies. The International journal of developmental biology. 2009; 53:783. https://doi.org/10.1387/ijdb.072493rb PMID: 19557684
4. Wang Q, Oh JW, Lee HL, Dhar A, Peng T, Ramos R, et al. A multi-scale model for hair follicles reveals heterogeneous domains driving rapid spatiotemporal hair growth patterning. eLife. 2017; 6. https://doi.org/10.7554/eLife.22772
5. Painter KJ, Maini PK, Othmer HG. Stripe formation in juvenile *Pomacanthus* explained by a generalized turing mechanism with chemotaxis. Proc Natl Acad Sci U S A. 1999; 96(10):5549–5554. https://doi.org/10.1073/pnas.96.10.5549 PMID: 10318921
6. Volkening A, Sandstede B. Modelling stripe formation in zebrafish: an agent-based approach. Journal of the Royal Society Interface. 2015; 12(112):20150812. https://doi.org/10.1098/rsif.2015.0812
7. Marrocco A, Henry H, Holland IB, Plapp M, Seror SJ, Perthame B. Models of Self-Organizing Bacterial Communities and Comparisons with Experimental Observations. Mathematical Modelling of Natural Phenomena. 2010; 5(1):148–162. https://doi.org/10.1051/mmnp/20105107
8. Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R. A synthetic multicellular system for programmed pattern formation. Nature. 2005; 434(7037):1130. https://doi.org/10.1038/nature03461 PMID: 15858574
9. Mukherji S, Van Oudenaarden A. Synthetic biology: understanding biological design from synthetic circuits. Nature reviews Genetics. 2009; 10(12):859. https://doi.org/10.1038/nrg2697 PMID: 19898500
10. Elowitz M, Lim WA. Build life to understand it. Nature. 2010; 468(7326):889–890. https://doi.org/10.1038/468889a PMID: 21164460
11. Khalil AS, Collins JJ. Synthetic biology: applications come of age. Nature reviews Genetics. 2010; 11 (5):367. https://doi.org/10.1038/nrg2775 PMID: 20395970
12. Liu C, Fu X, Liu L, Ren X, Chau CKL, Li S, et al. Sequential Establishment of Stripe Patterns in an Expanding Cell Population. Science. 2011; 334(6053):238–241. https://doi.org/10.1126/science.1209042 PMID: 21998392
13. Fu X, Tang L, Liu C, Huang J, Hwa T, Lenz P. Stripe formation in bacterial systems with density-suppressed motility. Physical Review Letters. 2012; 108(19):198102. https://doi.org/10.1103/PhysRevLett.108.198102 PMID: 23003092
14. Dallon JC., Othmer HG. A discrete cell model with adaptive signaling for aggregation of Dictyostelium discoideum, Philosophical Transactions of the Royal Society B. 1997; 352:391–418. https://doi.org/10.1098/rstb.1997.0029
15. Xue C. Macroscopic equations for bacterial chemotaxis: integration of detailed biochemistry of cell signaling. J Math Biol. 2015; 70(1-2):1–44. https://doi.org/10.1007/s00285-013-0748-5 PMID: 24366373

16. Xin X, Othmer HG. A "trimer of dimers"-based model for the chemotactic signal transduction network in bacterial chemotaxis. Bull Math Biol. 2012; 74(10):2339–2382. https://doi.org/10.1007/s11538-012-9756-7 PMID: 22864951

17. Si G, Tang M, Yang X. A pathway-based mean-field model for E. coli chemotaxis: Mathematical derivation and Keller-Segel limit. Multiscale Modeling and Simulations. 2014; 12(2):907–926. https://doi.org/10.1137/130944199

18. Xue C, Othmer HG. Multiscale models of taxis-driven patterning in bacterial populations. SIAM J Appl Math. 2009; 70(1):133–167. https://doi.org/10.1137/070711505 PMID: 19784399

19. Berg HC, Brown D. Chemotaxis in *Escherichia Coli* analyzed by three-dimensional tracking. Nature. 1972; 239:502–507. https://doi.org/10.1038/239500a0

20. Macnab RM. Sensing the Environment: Bacterial Chemotaxis. In: Goldberg R, editor. Biological Regulation and Development. New York: Plenum Press; 1980. p. 377–412.

21. Xue C. Bacterial chemotaxis: a classic example of multiscale modeling in biology. Springer;.

22. Cluzel P, Surette M, Leibler S. An Ultrasensitive Bacterial Motor Revealed by Monitoring Signaling Proteins in Single Cells. Science. 2000; 287:1652–1655. https://doi.org/10.1126/science.287.5458.1652 PMID: 10698740

23. Franz B, Xue C, Painter K, Erban R. Travelling waves in hybrid chemotaxis models. Bull Math Biol. 2014;. https://doi.org/10.1007/s11538-013-9924-4 PMID: 24347253

24. Chalub FACC, Markowich PA, Perthame B, Schmeiser C. Kinetic models for chemotaxis and their drift-diffusion limits. Monatshefte für Mathematik. 2004; 142(1):123–141. https://doi.org/10.1007/s00605-004-0234-7

25. Cates ME, Marenduzzo D, Pagonabarraga I, Tailleur J, Arrested phase separation in reproducing bacteria creates a generic route to pattern formation. PNAS, 2010; 107: 11715. https://doi.org/10.1073/pnas.1001994107 PMID: 20498087