

УДК 575.2:577.29:004.942

**МАТЕМАТИЧЕСКАЯ МОДЕЛЬ РЕГУЛЯТОРНОГО КОНТУРА
Rob, MarR и MarA ГЕННОЙ СЕТИ *Escherichia coli***

М.Т. Ри, С.С. Хайрулин, О.В. Сайк

Аннотация

В работе приводятся результаты исследования математической модели, описывающей регуляторный контур генной сети *Escherichia coli*, состоящей из трех белков Rob, MarR и MarA. Выбор контура связан с анализом механизмов прямой и обратной связи с учетом запаздывающего аргумента (время задержки на регуляцию). Результаты расчета модели демонстрируют, что система может иметь стационарное решение и незатухающие автоколебания. Оба режима являются устойчивыми в отношении внутреннего шума и периодического внешнего воздействия. Численный анализ модели показывает, что период предельного цикла связан с запаздывающим аргументом линейно, а амплитуда – нелинейно. Выявлено также, что дупликация *marR* гена приводит к решениям сложной природы. На данном этапе разработаны две версии модели: детерминистская и стохастическая. При расчетах были использованы параллельные вычисления, позволившие уменьшить время расчета модели без запаздывания в 12 раз и модели с запаздыванием в 640 раз.

Ключевые слова: генная сеть, математическое моделирование, стохастические системы, колебания, запаздывающий аргумент, *Escherichia coli*, OpenCL, параллельные вычисления.

Введение

Одной из актуальных задач системной биологии является исследование динамических свойств генных сетей в зависимости от структурно-функциональной организации с использованием математического моделирования и высоко воспроизводительных вычислений [1, 2]. Концепция исследования генной сети *E. coli* состоит в разработке математических моделей отдельных регуляторных контуров, их анализа и последующего объединения подмоделей. К настоящему моменту разработан ряд математических моделей, описывающих генетическую регуляцию некоторых процессов *E. coli* [3–11]. Структура моделей учитывает механизмы генетической регуляции репарационного процесса [7], экспрессии основных компонентов генной сети, контролирующих дыхание в клетке *E. coli* [5, 9], метаболизма нуклеотидов [6, 11], триптофанового и лактозного оперонов [3, 4]. Приведенные примеры – это лишь небольшая часть существующих генетических процессов, однако перечисленные механизмы являются фундаментальными.

Выбор генной сети *E. coli* связан с тем, что элементы сети и процессы наиболее хорошо исследованы и представлены как в научной литературе, так и в базах данных, таких как RegulonDB и EcoCyc. Особенностью *E. coli* является способность адаптироваться к быстро меняющимся условиям окружающей среды посредством переключения метаболизма на новый режим функционирования.

Данное переключение осуществляется за счет работы регуляторных белков, элементов генной сети, которые либо непосредственно реагируют на внешний сигнал, либо участвуют в системе распространения сигнала внутри клетки. Передача сигнала обеспечивается за счет положительных и отрицательных связей между регуляторными белками. Ранее в работе [11] была представлена генная сеть¹ *E. coli*. Основываясь на данной сети, нами рассмотрена взаимная регуляция трех регуляторных белков – Rob, MarR и MarA.

Rob является мономером, относится к AraC/XylS белковому семейству, которые регулируют гены, вовлеченные в ответ на действия антибиотиков, органических солей и тяжелых металлов [12, 13]. В частности, к таким генам относятся гены белков MarR и MarA, входящие в состав оперона marRAB (рис. 1, а). Белок Rob положительно влияет на регуляцию оперона marRAB и повышает экспрессию в 1.5-2 раза [14]. В свою очередь, регуляция экспрессии гена белка Rob находится под негативным контролем со стороны мономерного белка MarA и собственного белка Rob [13, 15]. Сайты связывания MarA и Rob перекрываются с сайтом посадки РНК полимеразы (рис. 1, б) и между собой, в результате чего экспрессия оперона rob снижается в 2–4 раза благодаря MarA и в 3 раза при действии Rob. [15]. Регуляторный белок MarA повышает экспрессию оперона marRAB в 1.5-3 раза [14, 16]. Белок MarR является гомодимером [17] и связывается с двумя сайтами связывания в регуляторной области оперона marRAB, один из которых перекрывается с сайтом посадки РНК полимеразы (рис. 1, а), что приводит к снижению экспрессии в 19 раз [18]. Помимо Rob, MarR и MarA в регуляции оперона marRAB задействованы регуляторные белки – SoxS, Fis, Crp и Cra [14, 19, 20], также SoxS негативно регулирует оперон rob [15].

Сложность взаимной регуляции между генами и их продуктами является основной особенностью природных генных сетей, которая помогает клетке выживать в самых разнообразных условиях. Представленные белки Rob, MarR и MarA взаимно регулируются пятью регуляторными связями (см. рис. 1, в). Таким образом, на основе имеющихся данных целью настоящей работы было продемонстрировать динамические свойства данного контура с помощью математического моделирования и проанализировать механизмы прямой и обратной связей. Поскольку большинство динамических процессов со сложной регуляцией демонстрирует стохастическое поведение и процессы запаздывания в регуляции, в настоящей статье сначала рассматриваются детерминистическая и стохастическая модели без учета запаздывающего аргумента, а затем в обе модели вводится запаздывающий аргумент. Важно отметить, что ранее данный контур не моделировался и не представлен в научной литературе.

1. Материалы и методы

Математическая модель (М) строилась как система обыкновенных дифференциальных уравнений, описывающая глобальные скорости изменения концентраций белков. Глобальные скорости вычислялись на основании закона суммирования локальных скоростей, описывающих процесс синтеза или деградации

¹ Генная сеть – ориентированный граф, вершины которого являются регуляторными белками, а связи – тип регуляции.

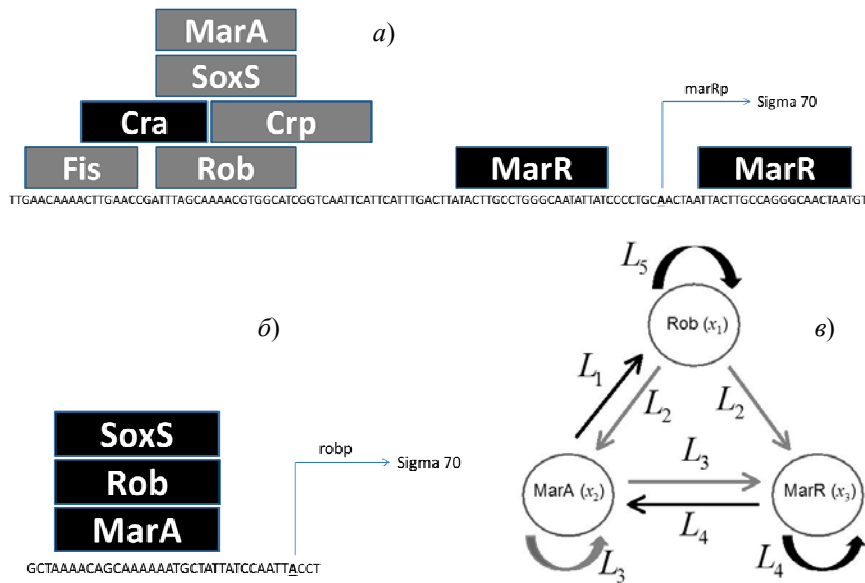


Рис. 1. (а) Регуляторная область оперона *marRAB*. (б) Регуляторная область оперона *gob*. (в) Общая схема регуляторного контура, состоящего из трех белков. Обозначения для (а) и (б): прямоугольником черного цвета обозначены регуляторные белки, выступающие в роли ингибиторов, серого цвета – активаторы. *marRp* и *gobp* – название промоторов оперона *marRAB* и *gob* соответственно. *Sigma 70* – сигма-фактор в составе РНК полимеразы. Обозначения для (в): овалами белого цвета обозначены белки. Черные стрелки – ингибирование экспрессии гена соответствующего белка, Серые – активация. x_1 , x_2 и x_3 – переменные математической модели (М). L_1 , L_2 , L_3 , L_4 , L_5 – обозначения регуляторных контуров. Поскольку гены белков *MarR* и *MarA* входят в состав одного оперона *marRAB*, L_2 , L_3 и L_4 повторяются дважды

белка. Верификация модели (М) осуществлялась на основании работ [14–19, 21–26]. Численные расчеты и анализ модели (М) проводились средствами программы Mathematica 7.0. При разработке стохастической модели без запаздывания использовался алгоритм Гиллеспи [27], а стохастическая модель с запаздыванием основана на обобщенном алгоритме Гиллеспи, для случая немарковских систем [28]. Для реализации распараллеленного алгоритма была выбрана платформа OpenCL. Алгоритм был реализован на языках Python и OpenCL C, и протестирован на процессоре Intel i7 (4 ядра, 8 потоков) и GPU Tesla C1060 (240 ядер, 30 процессоров). Математическая модель, код программы и результаты анализа данных доступны по адресу http://dl.dropbox.com/u/52461630/programs_model.zip.

2. Результаты исследования

2.1. Детерминистическая модель.

2.1.1. Математическая модель регуляторного контура *Rob*, *MarA* и *MarR*.

Переменные математической модели (М) – *Rob* (x_1), *MarA* (x_2), *MarR* (x_3). f_1 и f_2 – функции, описывающие регуляцию экспрессии оперонов *gob* и *marRAB*, соответственно. Структура функций f_1 и f_2 записана с учетом структурно-функцио-

нальной организации регуляторной области оперонов (рис. 1, а, б). Параметры модели представлены в табл. 1.

$$\begin{cases} \frac{dx_1}{dt} = k_{10}f_1 - k_8x_1[t], \\ \frac{dx_2}{dt} = k_{11}f_2 - k_9x_2[t], \\ \frac{dx_3}{dt} = k_{12}f_2 - k_{13}x_3[t], \end{cases} \quad (M)$$

где

$$f_1 = \frac{1}{1 + \frac{x_1[t]}{k_0} + \frac{x_2[t]}{k_1}}, \quad f_2 = \frac{1 + \frac{x_1[t]}{k_2} + \frac{x_2[t]}{k_4}}{1 + \frac{x_1[t]}{k_3} + \frac{x_2[t]}{k_5}} \frac{1}{1 + \left(\frac{x_3[t]}{k_6}\right)^{k_7}}.$$

2.1.2. Верификация параметров математической модели (M). Подбор параметров осуществлялся следующим образом.

1. Константа ингибирования (k_0) экспрессии оперона *rob* белком Rob (x_1) вычисляется на основании функции f_1 , а также информации о том, что количество белка Rob приблизительно равно 7000 молекул/клетку [21, 23, 24] и $\text{MarR} = 0$. Поскольку Rob снижает уровень экспрессии в 3 раза [15], то $f_1 = 0.33$. Константа ингибирования (k_1) экспрессии оперона *rob* белком MarA (x_2) вычисляется на основании функции f_1 , а также информации о том, что количество белка MarA ≈ 20000 молекул/клетку [19] и $\text{Rob} = 0$. Поскольку MarA снижает уровень экспрессии в 2–4 раза [15], то $f_1 \in [0.25, 0.5]$ при условии, что $f_1 = 1$ при $\text{MarA} = 0$. В модели (M) $f_1 = 0.25$.

2. Используя функцию f_2 , мы оцениваем константы k_2 и k_3 , характеризующие активацию экспрессии оперонов *marRAB* белком Rob (x_1) при $f_2 = 2$ [14], количество белка Rob ≈ 7000 молекул/клетку [21, 23, 24] и $\text{MarR} = \text{MarA} = 0$.

Аналогично вычисляются константы k_4 и k_5 при условии, что $f_2 = 3$ [16], количество белка MarA ≈ 20000 молекул/клетку [19] и $\text{MarR} = \text{Rob} = 0$.

3. Константа ингибирования (k_6) экспрессии оперона *marRAB* белком MarR вычисляется на основании функции $f_2 = 1/19$ и $\text{MarA} = \text{Rob} = 0$ при двух предположениях. Во-первых, константа k_7 , характеризующая степень нелинейности влияния MarR на регуляцию экспрессии оперона *marRAB*, равна 4, поскольку MarR является гомодимером [17] и взаимодействует с двумя неперекрывающимися сайтами связывания [18]. Во-вторых, количество белка MarR ≈ 6000 молекул/клетку, что соответствует теоретической оценке на основании количественных данных по мРНК *marR* [25].

4. Время полужизни белка лежит в интервале от десятков секунд до десятков минут [22, 26]. Предполагая, что время полураспада белка равно 5 мин, вычисляем константы деградации для белков Rob (k_8) и MarR (k_9) из формулы $X(t) = X_0 e^{-kt}$, где X – концентрация белка, k – константа деградации белка, t – время полураспада белка при $X(t) = 0.5X_0$. Следовательно, $k_8 = k_9 = (\ln 2)/t$.

Табл. 1

Параметры математической модели (М)

Параметр	Описание
$k_0 = 3448$ [молекул]	Константа ингибирования экспрессии оперона <i>gob</i> белком Rob
$k_1 = 6667$ [молекул]	Константа ингибирования экспрессии оперона <i>gob</i> белком MarA
$k_1 = 1000$ [молекул]	Константы, характеризующие активацию экспрессии оперона <i>marRAB</i> белком Rob
$k_3 = 2333$ [молекул]	
$k_4 = 1000$ [молекул]	Константы, характеризующие активацию экспрессии оперона <i>marRAB</i> белком MarA
$k_5 = 3333$ [молекул]	
$k_6 = 2913$ [молекул]	Константа ингибирования экспрессии оперона <i>marRAB</i> белком MarR
$k_7 = 4$	Константа, характеризующая степень нелинейности влияния MarR на регуляцию экспрессии оперона <i>marRAB</i>
$k_8 = 0.0023$ [с ⁻¹]	Константа деградации белка Rob
$k_9 = 0.0023$ [с ⁻¹]	Константа деградации белка MarA
$k_{10} = 97$ [молекул/с]	Обобщенная константа синтеза белка Rob
$k_{11} = 312$ [молекул/с]	Обобщенная константа синтеза белка MarA
$k_{12} = 312$ [молекул/с]	Обобщенная константа синтеза белка MarR
$k_{13} = 0.0077$ [с ⁻¹]	Константа деградации белка MarR

5. Исходя из предположения, что система находится в квазиравновесном состоянии, вычисляем обобщенные константы синтеза для белков Rob (k_{10}) и MarR (k_{11}) из уравнения $dx_1/dt = dx_2/dt = 0$. Поскольку гены белков MarR и MarA входят в состав одного оперона, то предполагаем, что $k_{11} = k_{12}$. На основании этого равенства и уравнения $dx_3/dt = 0$ вычисляем константу деградации белка MarA (k_{13}).

2.1.3. Результаты численного расчета математической модели (М). Результаты модели (М) демонстрируют, что система имеет стационарное решение. Полученные стационарные значения концентраций регуляторных белков соответствуют литературным данным для Rob [21, 23, 24] и MarA [19, 29]. Значение концентрации MarR соответствует теоретической оценке на основании количественных данных по мРНК *marR* [25]. Решение является асимптотически устойчивым, поскольку собственные числа матрицы Якоби являются отрицательными при начальных данных трех переменных Rob (x_1), MarA (x_2), MarR (x_3), соответствующих значениям из литературы.

2.1.4. Параметрический анализ модели (М). На первом этапе проведен параметрический анализ поведения модели (М). Численный анализ показал, что варьирование по отдельности каждого параметра k_0, k_1, \dots, k_{13} обеспечивает наличие разных наборов устойчивых стационарных значений трех переменных Rob (x_1), MarA (x_2), MarR (x_3). На рис. 2 представлен пример изменения значения MarR (x_3) в зависимости от варьирования каждого параметра на один порядок в сторону как уменьшения, так и увеличения. Для данного интервала параметры $k_0 - k_5$ и $k_8 - k_{11}$ изменяют стационарное значение MarR (x_3) не более чем на 56%. Оставшиеся два параметра k_6 и k_7 являются более чувствительными. Стоит отметить, что параметры $k_2 - k_5$, характеризующие активацию экспрессии перона *marRAB* белком Rob и белком MarA, представляют наибольший интерес для

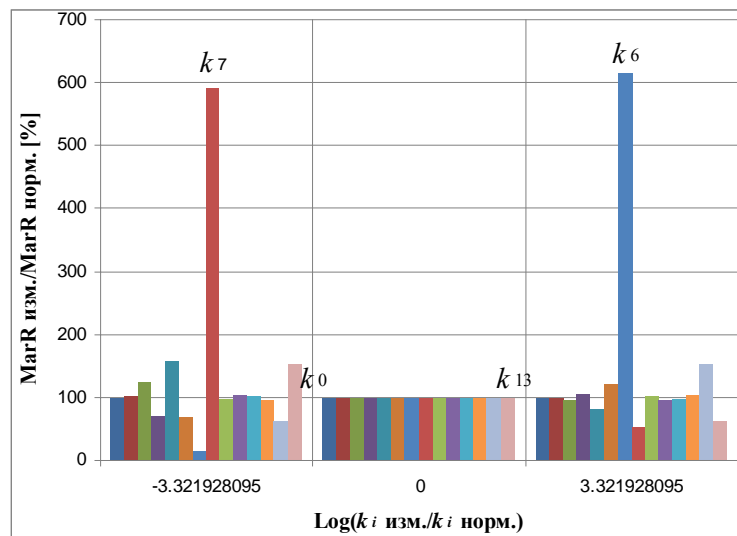


Рис. 2. Изменение стационарного значения MarR в зависимости от изменения параметров $k_0 - k_{13}$. По оси абсцисс отложен логарифм отношения измененного параметра (обозначено изм.) к неизмененному параметру (обозначено норм.). По оси ординат отношение измененного значения MarR к неизмененному значению

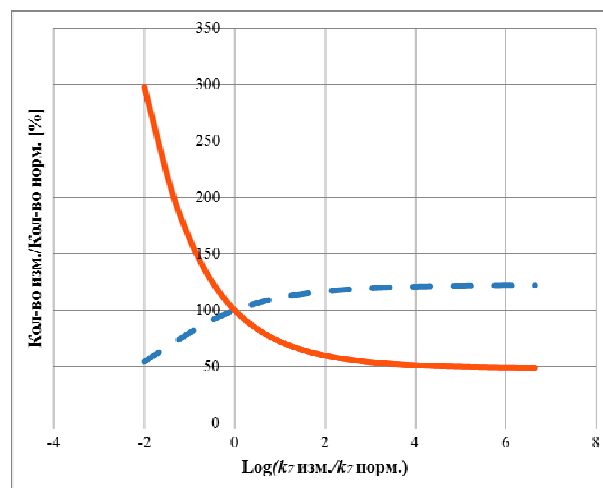


Рис. 3. Изменение стационарного значения Rob (штриховая линия) и MarR (сплошная линия) в зависимости от изменения параметра k_7 . По оси абсцисс отложен логарифм отношения измененного параметра (обозначено изм.) к неизмененному параметру (обозначено норм., $k_{7 \text{ норм}} = 4$). По оси ординат отношение измененного значения к неизмененному значению. Кривая по MarA совпадает с кривой MarR

изучения, поскольку входят в функцию f_2 , как и параметры k_6 и k_7 . Результаты расчета для переменных Rob (x_1) и MarA (x_2) доступны по адресу http://dl.dropbox.com/u/52461630/programs_model.zip. Более детально рассмотрим влияние варьирования параметра k_7 на динамику изменения стационарных значений трех переменных. Напомним, параметр k_7 характеризует степень нелинейности влияния MarR на регуляцию экспрессии оперона marRAB . Значение параметра

равно 4, поскольку MarR является гомодимером [17] и взаимодействует с двумя неперекрывающимися сайтами связывания [18]. При увеличении параметра k_7 на два порядка значения трех переменных изменяются не более чем в 2 раза, а при уменьшении до 1 количество Rob снижается в 1.8 раза, а MarA и MarR увеличивается в 3 раза (рис. 3). Соответствующие расчеты показывают, что параметр k_7 является чувствительным для всех трех переменных и может быть выбран в качестве дальнейшего изучения в экспериментах *in silico*.

2.1.5. Запаздывающий аргумент. На следующем этапе модель (M) рассматривается с учетом запаздывающего аргумента T . В результате мы переписем функции f_1 и f_2 следующим образом:

$$f_1 = \frac{1}{1 + \frac{x_1[t-T]}{k_0} \frac{x_2[t-T]}{k_1}}, \quad f_2 = \frac{1 + \frac{x_1[t-T]}{k_2} + \frac{x_2[t-T]}{k_4}}{1 + \frac{x_1[t-T]}{k_3} + \frac{x_2[t-T]}{k_5}} \frac{1}{1 + \left(\frac{x_3[t-T]}{k_6}\right)^{k_7}}. \quad (M1)$$

Система дифференциальных уравнений остается той же самой. Стоит отметить, что для генетических процессов в клетки *E. coli* время задержки лежит в интервале от десятков секунд до нескольких минут [30–33]. Численные расчеты демонстрируют, что в зависимости от численного значения запаздывающего аргумента возможны два предельных случая поведения системы: стационарное решение (СР) и периодические автоколебания (ПА). На рис. 4 продемонстрированы примеры ПА. Интересно, что режим функционирования ПА реализуется при значениях запаздывающего аргумента 67 с и более. Возможность существования предельного цикла в виде незатухающих автоколебаний порождает ряд интересных вопросов. Во-первых, какие регуляторные петли генерируют автоколебания? Во-вторых, как варьирование величины параметра k_7 , характеризующего степень нелинейности влияния MarR на регуляцию экспрессии оперона *marRAB*, меняет динамику системы? И, в-третьих, являются ли колебания внутренним шумом в зависимости от величины периода и амплитуды?

2.1.6. Влияние регуляторных петель на динамику модели (M1). На рис. 1 представлены пять регуляторных связей, которые обозначены через $L_1–L_5$. Две регуляторных петли L_1 и L_2 регулируют по механизму прямой связи. Остальные три L_3 , L_4 и L_5 контролируют по механизму обратной связи. L_1 , L_4 и L_5 – отрицательные связи, а L_2 и L_3 – положительные связи. Введем данные обозначения в функции f_1 и f_2 модели (M1):

$$f_1 = \frac{1}{1 + (1-L_5) \frac{x_1[t-T]}{k_0} + (1-L_1) \frac{x_2[t-T]}{k_1}},$$

$$f_2 = \frac{1 + (1-L_2) \frac{x_1[t-T]}{k_2} + (1-L_3) \frac{x_2[t-T]}{k_4}}{1 + (1-L_2) \frac{x_1[t-T]}{k_3} + (1-L_3) \frac{x_2[t-T]}{k_5}} \frac{1}{1 + (1-L_4) \left(\frac{x_3[t-T]}{k_6}\right)^{k_7}}. \quad (M2)$$

Положим, что при $L_i=1$, $i=1,2,3,4,5$, регуляторная связь отсутствует, а если $L_i=0$, то петля присутствует. Таким образом, численный анализ 32 версий (32 –

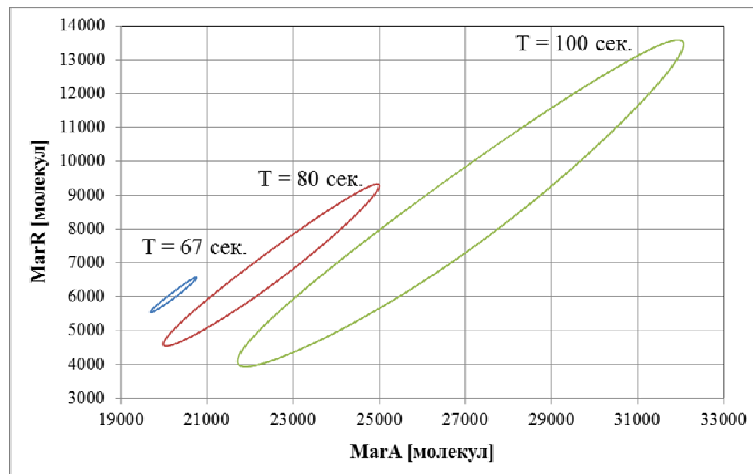


Рис. 4. Фазовый портрет двух переменных MarA (x_2) и MarR (x_3). Предельные циклы в виде незатухающих автоколебаний представлены при разных значениях времени запаздывания

комбинаторика пяти связей) модели (M2) показал, что автоколебания возникают в 16 версиях модели (M2), в которых учитывается влияние регуляторной петли L_1 и L_4 (рис. 1). Если убрать из рассмотрения контур L_1 , то только белки MarR и MarA изменяются периодически. Таким образом, можно заключить, что регуляторная петля L_4 по механизму обратной отрицательной связи генерирует автоколебания, а через механизмы прямой отрицательной связи запускает автоколебания белка Rob. При этом регуляторная петля L_5 по механизму обратной отрицательной связи не генерирует автоколебания.

2.1.7. Влияние параметра k_7 на динамику системы. Для любого $T \geq 67$ уменьшение k_7 приводит к стационарному решению, а увеличение – к предельному циклу. Важно отметить, что параметр k_7 характеризует нелинейный процесс, осуществляемый через регуляторную петлю L_4 по механизму обратной отрицательной связи, наличие которой обеспечивает колебания в системе.

2.1.8. Взаимосвязь периода и амплитуды с запаздывающим аргументом. Исходя из того, что для генетических процессов в клетки *E. coli* время задержки лежит в интервале от десятков секунд до нескольких минут [30–33] нами выбран интервал запаздывающего аргумента от 67 с – величина, с которой реализуются колебания до 5 мин. Численные расчеты модели (M1) продемонстрировали, что для данного интервала значение периода колебаний концентраций белков Rob, MarA и MarR растет линейно в интервале от 229 с (≈ 4 мин) до 985 с (≈ 16 мин) (рис. 5, а). При этом амплитуда изменяется нелинейно (рис. 5, б). Численные значения амплитуды для переменной Rob составляют несколько десятков молекул, а для MarA и MarR – тысячи молекул. Таким образом, колебания, как внутреннее свойство системы, не является внутренним шумом.

2.2. Стохастическая модель. Стохастические процессы в экспрессии генов играют важную роль в жизнедеятельности клетки, поскольку позволяют ей справиться с изменениями в окружающей среде через фенотипическое разнообразие в популяции [34, 35]. Ранее в нескольких работах экспериментально продемонстрирована стохастичность транскрипции и трансляции на индивидуальных про- и эукариотических клетках [36, 40].

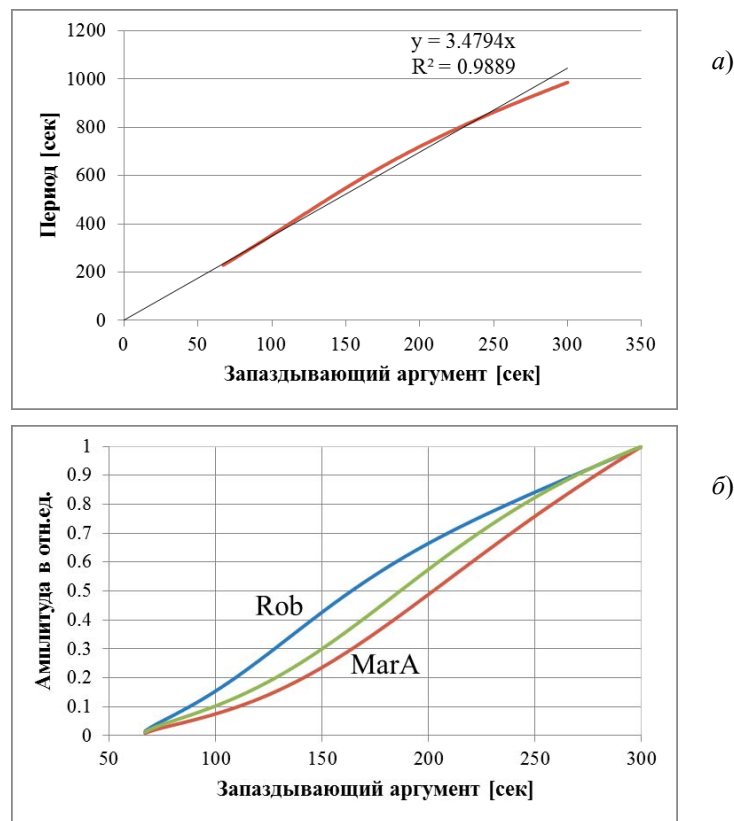


Рис. 5. Зависимость периода (а) и амплитуды (б) от запаздывающего аргумента для трех переменных Rob, MarA и MarR. Для (б) величина оси ординат вычисляется как значение амплитуды в каждой точке, поделенное на значение амплитуды при запаздывающем аргументе 300 с

2.2.1. Стохастическая модель без запаздывания. Нами разработана стохастическая модель, основанная на алгоритме Гиллеспи [27], для того чтобы исследовать роль флуктуаций в генетически измененной системе, где количество белков невелико. Например, изменим величину константы деградации (k_8) белка Rob с $0.0023 \text{ [с}^{-1}\text{]}$ на $0.69 \text{ [с}^{-1}\text{]}$. В этом случае соответствующая система монотонным образом перейдет на новый стационарный уровень белка Rob (рис. 6, а).

2.2.2. Стохастическая модель с запаздыванием. Впервые в работе Д.А. Брацуна [28] для генетических процессов транскрипции/трансляции генов были рассмотрены вопросы взаимодействия запаздывания и стохастических флуктуаций. В частности, был предложен обобщенный алгоритм Гиллеспи на случай немарковских систем. Полное описание идеи и численных расчетов модификации алгоритма представлено в работе [41]. Используя данный алгоритм, нами разработана стохастическая версия модели (M1) с запаздывающим аргументом для генетически измененной системы ($k_6 = 29.13 \text{ [молекул]}$, $k_8 = k_9 = k_{13} = 0.23 \text{ [с}^{-1}\text{]}$, $k_{10} = k_{11} = k_{12} = 17 \text{ [молекул/с]}$; остальные параметры в табл. 1). На рис. 6, б приведены результаты расчета. Хорошо видно, что стохастическая модель качественно описывает результаты детерминистической модели, а именно присутствуют флуктуации и колебания. Для увеличения производительности программы

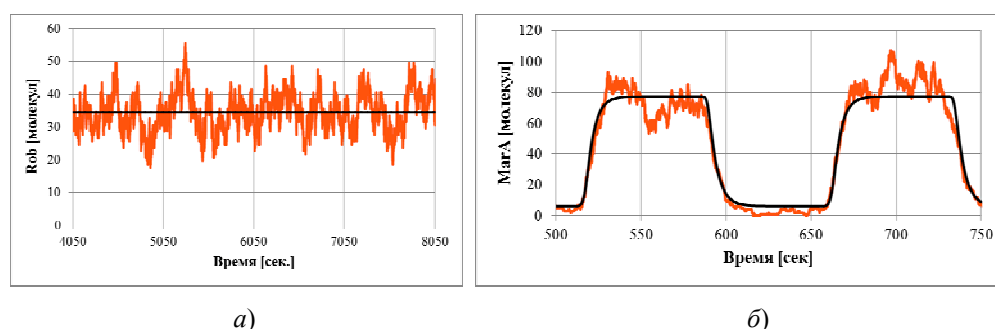


Рис. 6. Временная эволюция количества белка Rob (x_1) при внезапном увеличении деградации белка Rob, полученная при численном расчете стохастической модели без запаздывания, (а) и белка MagA (x_2), полученная при численном расчете стохастической модели с запаздыванием, $T = 70$ с (б). Численные расчеты стохастической модели получены для одной клетки. Сплошной линией показано численное решение детерминистической модели

Табл. 2

Время расчета стохастической модели на процессорах Intel i7 core 8 и Tesla C1060 core 30 (приведено в секундах)

Количество клеток	Intel i7 core 8				Tesla C1060 core 30	
	C1	C2	C3	C4	C3	C4
100	88	9037.36	9.94	14.05	7.01	242.11
300*	—	—	10.89	26.71	7.84	296.43
600**	—	—	11.28	49.88	8.33	312.92

Расчет осуществлен без учета вывода в файл. Тесты проводились на операционной системе Ubuntu 12.04 x64. Обозначения: стохастическая модель без запаздывания (C1) и с запаздыванием (C2), однопоточная реализация; стохастическая модель без запаздывания (C3) и с запаздыванием (C4), многопоточная реализация. Для всех версий модели было положено 10^5 итераций. Для C1 и C3 параметры представлены в табл. 1, а для C2 и C4 – в разд. 2.2.2.

* Количество клеток в экспоненциальную фазу роста.

** Количество клеток в стационарную фазу роста. Количество клеток рассчитано на $\times 10^{-10}$ л минимальной среды.

Оценка для * и ** приведена в http://dl.dropbox.com/u/52461630/programs_model.zip. Для «→» расчеты не проводились.

мы распараллелили существующие алгоритмы для однопоточной реализации, что привело к уменьшению времени расчета более чем в 12 раз (табл. 2).

2.3. In silico эксперимент.

2.3.1. Влияние внешнего шума. Под внешним воздействием будем считать влияние регуляторных белков на экспрессию генов белка Rob, MagA и MarR. На наш взгляд, одним из интересных кандидатов выступает мультифункциональный фактор DnaA. Он задействован в процессе инициации репликации ДНК [42], и его количество циклически изменяется. Экспериментально показано, что минимум приходится на момент клеточного деления [42]. Однако в литературе не представлены сайты связывания белка DnaA в регуляторной области оперона *gob* и *magRAB*. Для выяснения этого вопроса мы использовали 12 известных сайтов связывания для белка DnaA из базы данных RegulonDB (см. в http://dl.dropbox.com/u/52461630/programs_model.zip), которые выровняли с регуляторной областью оперона *gob* и *magRAB* с шагом в один нуклеотид,

начиная с -100 -позиции по $+38$ -позицию с учетом одноименных цепей. Результаты представлены на рис. 7. Найдены сайты связывания, совпадающие на 77.8% и 88.9% с известными сайтами (7 и 8 из 9 нуклеотидов совпали). Часть сайтов перекрывается с уже известными сайтами (рис. 1, *a*, *b*) и сайтом посадки РНК полимеразы. На основании этого можно предположить, что DnaA выступает в роли ингибитора. Однако для того чтобы выяснить, действительно ли существует связывание и ингибирующий эффект, необходимо проводить эксперименты, и это является отдельной задачей. На данном этапе полученную информацию мы использовали для расширения модели (M1). Важно отметить, что активная форма фактора DnaA связана с АТФ, при этом DnaA также способна связываться с молекулой АДФ. Поскольку значения констант диссоциации равны 0.03 и 0.1 мкМ соответственно и константы значительно ниже внутриклеточной концентрации 3 мМ АТФ и 0.25 мМ АДФ [43], то все образовавшиеся молекулы DnaA свяжутся с АТФ и АДФ, при этом равновесная концентрация DnaA равна 0.0017 мМ [23]. Запишем формулу активной формы белка следующим образом:

$$\text{DnaA}_a = f \cdot g(\text{DnaA}, \text{ATP}, \text{ADP}), \quad \text{где } f = 0.56 + 0.44 \cos\left(\frac{t}{390} + \pi\right).$$

При этом величина функции g численно равна концентрации DnaA. Отметим, что при $f_{\min} = 0.12$ концентрация активных форм белка будет составлять $\sim 10\%$ от общего количества, а при $f_{\max} = 1$ концентрация активных форм равна $\sim 100\%$. 80% достаточно, чтобы активировать инициацию репликации.

Обновим функцию f_1 и f_2 в модели (M1):

$$\begin{aligned} f_1 &= \frac{1}{1 + \frac{x_1[t-T]}{k_0} + \frac{x_2[t-T]}{k_1} + \frac{\text{DnaA}_a}{k_{\text{DnaA}_a}}}, \\ f_2 &= \frac{1 + \frac{x_1[t-T]}{k_2} + \frac{x_2[t-T]}{k_4}}{1 + \frac{x_1[t-T]}{k_3} + \frac{x_2[t-T]}{k_5} + \frac{\text{DnaA}_a}{k_{\text{DnaA}_a}}} \frac{1}{1 + \left(\frac{x_3[t-T]}{k_6}\right)^{k_7}}. \end{aligned} \quad (\text{M3})$$

Здесь k_{DnaA_a} – константа ингибирования экспрессии оперонов *gov* и *magRAB* активной формой DnaA. Для простоты положим k_{DnaA_a} одинаковой для сайта связывания в регуляторной области оперона *gov* и *magRAB*. Пренебрегая нелинейными эффектами, примем коэффициент Хилла равным единице. Исследуем, как будет меняться динамика при значениях параметра $k_{\text{DnaA}_a} = 20$ нМ и 2 мкМ. На рис. 8 представлены результаты численного расчета модели (M3) для детерминистической и стохастической версий. Если запаздывание отсутствует и $k_{\text{DnaA}_a} = 20$ нМ, то возникают автоколебания, которые не являются внутренним свойством системы, а генерируются за счет периодического внешнего шума (рис. 8, *a*). При $k_{\text{DnaA}_a} = 2$ мкМ также возникают автоколебания, однако амплитуда составляет менее 3% от стационарного значения. Данные колебания можно

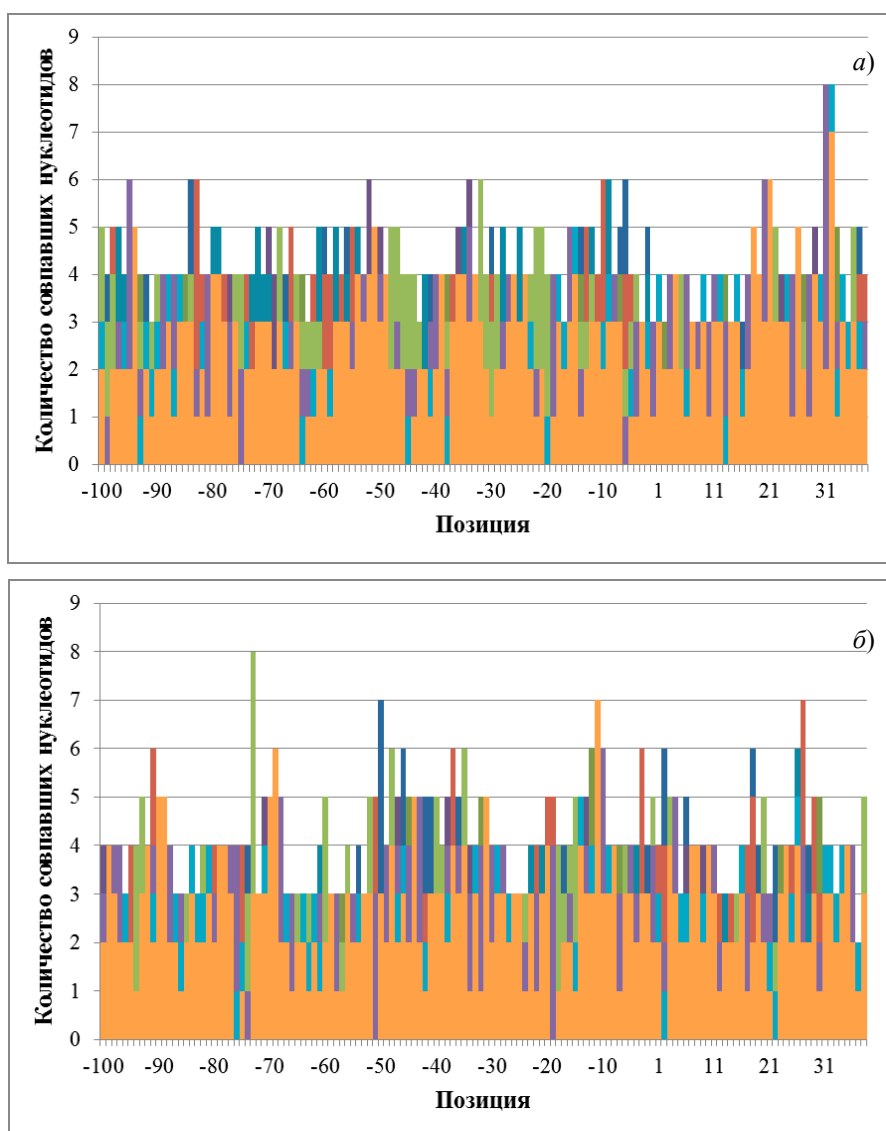


Рис. 7. Диаграмма, показывающая количество совпавших нуклеотидов известного сайта связывания (на других регуляторных областях) белка DnaA с регуляторной областью оперона *gov* (а) и *marRAB* (б). Каждый цвет показывает один из 12 известных сайтов (диаграмма построена с учетом перекрытия). Длина известных сайтов – 9 нуклеотидов. Выравнивание каждого сайта проводилось для одноименных цепей, начиная с –100-позиции от старта транскрипции с шагом в один нуклеотид и заканчивая +38-позицией. Каждый столбик показывает правую границу 9-нуклеотидного сайта и количество совпавших нуклеотидов

рассматривать как шум, что позволяет утверждать, что система устойчива к внешнему периодическому шуму. Введение запаздывающего аргумента приводит к периодическим режимам функционирования системы. При определенном наборе параметров можно получить периодическое решение сложной природы, например с параметрами $k_{\text{DnaA}_v} = 20$ нМ (рис. 8, б) и 2 мкМ (рис. 8, в).

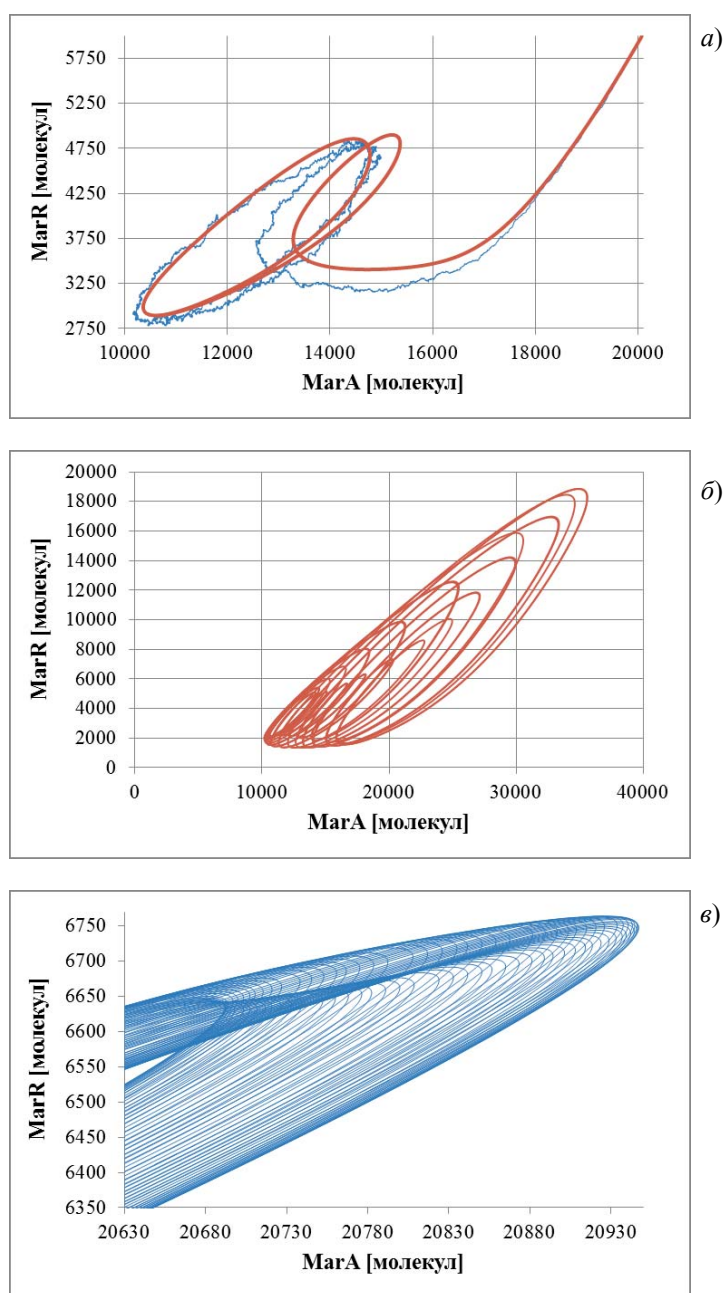


Рис. 8. Фазовый портрет двух переменных MagA (x_2) и MarR (x_3). (а) Эволюция фазовой траектории при расчетах стохастической модели. Численные расчеты стохастической модели получены для одной клетки. Сплошная линия показывает численное решение задачи Коши для модели (M4). Время запаздывания $T = 0$ с; константа $k_{\text{DnaA}_s} = 20$ нМ. (б) Время запаздывания $T = 200$ с; константа $k_{\text{DnaA}_s} = 20$ нМ. (в) Фрагмент аттрактора. Время запаздывания $T = 68$ с; константа $k_{\text{DnaA}_s} = 2$ мкМ. Отметим, что для (а) и (б) это не эволюция выхода на предельный цикл

2.3.2. Дупликация гена *marR*. Как было показано выше, регуляторная петля L_4 по механизму обратной отрицательной связи генерирует автоколебания. Напомним, что под этим подразумевается ингибирование экспрессии оперона *marRAB* белком *MarR*. Проведем *in silico* эксперимент, а именно учтем в нашей модели наличие плазмидной конструкции, состоящей из гена *marR* и регуляторной области, включающей в себя только сайты посадки белка *MarR* (рис. 1, *a*). Предположим, что скорость синтеза белка *MarR* и эффективность ингибирования в конструкции и в геноме одинаковы. При этом времена запаздывания будем считать различными. Запишем формально данные дополнения в M1:

$$\frac{dx_3}{dt} = k_{12}f_2 + k_{14}f_3 - k_{13}x_3[t], \quad (M4)$$

$$f_2 = \frac{1 + \frac{x_1[t-T]}{k_2} + \frac{x_2[t-T]}{k_4}}{1 + \frac{x_1[t-T]}{k_3} + \frac{x_2[t-T]}{k_5}} \frac{1}{1 + \left(\frac{x_3[t-T] + x_3[t-T1]}{k_6} \right)^{k_7}},$$

$$f_3 = \frac{1}{1 + \left(\frac{x_3[t-T] + x_3[t-T1]}{k_6} \right)^{k_7}}.$$

Правые части переменных x_1 и x_2 , а также функция f_1 останутся без изменения. Как и в разд. 3.1, исследуем динамические свойства системы.

Численный анализ модели M4 показал, что существуют три режима функционирования системы при варьировании запаздывающего аргумента $T1$ – стационарное решение, периодическое решение и решение сложной природы. Важно отметить, что время запаздывания лежит как в секундном, так и минутном диапазоне для первых двух режимов, что хорошо согласуется с экспериментальными данными. Для решений со сложным динамическим поведением время запаздывания может составлять несколько десятков минут.

3. Обсуждение результатов

Используя ранее опубликованную генную сеть, мы выделили небольшую подсеть из трех белков *Rob*, *MarA* и *MarR*, которые составляют ~1.5% от общего количества регуляторных белков, и разработали математическую модель. На основе экспериментальных данных модель демонстрирует, что система может иметь два режима функционирования: стационарное решение и незатухающие периодические автоколебания.

Существует ли автоколебания для данного контура *in vivo*? Вопрос, на который можно ответить, только поставив эксперимент. Численные же расчеты математической модели регуляторного контура генной сети *E. coli* демонстрируют, что возможно существование двух режимов функционирования, которое обеспечивается наличием отрицательных связей, нелинейными эффектами, запаздыванием и определенным набором параметров. Важно, что данные автоколебания

являются внутренним свойством системы. Теоретический анализ ранее опубликованных моделей, которые учитывают негативные обратные связи, демонстрируют осциллирующие режимы функционирования [44–46]. Динамические характеристики режима зависят от разного набора параметров, в том числе от количества негативных петель [47].

Нелинейность является необходимым, но недостаточным условием формирования колебаний в рассматриваемой системе. Достаточность определяется количеством субъединиц белка *MagR* и сайтов связывания, с которыми взаимодействует *MagR*, поскольку *MagR* ингибирует экспрессию собственного гена.

Запаздывание в несколько минут является неотъемлемой частью генетических процессов и формируется от начала инициации транскрипции до поиска сайта связывания активной формой регуляторного белка. Таким образом, используя запаздывающий аргумент в математической модели, мы отказываемся от детального описания процессов синтеза мРНК, диффузии и других процессов. Возникает вопрос, а правомерно ли использовать запаздывающий аргумент наравне с детальным описанием генетических процессов. Ответ является утвердительным, поскольку ранее в работах [48–50] были доказаны теоремы и проведены численные расчеты, показывающие, что возможен переход от систем большой размерности к системам более низких размерностей, используя уравнения с запаздывающим аргументом. При этом изменение величины запаздывающего аргумента приводит к изменению поведения динамики системы. Переключение происходит при $T \geq 67$ с. Интересно отметить, что взаимосвязь между периодом и временем задержки является линейной, а между запаздывающим аргументом и амплитудой – нелинейной в интервале от 67 с до 5 мин. Возможно, это свидетельствует в пользу того, что нелинейные эффекты отдельных процессов не проявляются на уровне регуляции всего метаболизма клетки.

Важной особенностью биологических систем является способность приобретать мутации. Анализ мутаций в математических моделях осуществляется через варьирование параметров, которые характеризуют те или иные свойства регуляторных белков – деградацию, ассоциацию с сайтом связывания, проявление нелинейных эффектов и др. Таким образом, параметрический анализ модели (М) показал возможность предсказывать последствия мутаций через изменения значения параметров и подбирать оптимальные параметры. Данный анализ является важным с практической точки зрения, поскольку выявление ключевых параметров модели, при которых концентрация всех белков минимальна, является состоянием клетки, когда она не устойчива к антибиотикам. Напомним, что данные белки вовлечены в ответ на действия антибиотиков, органических солей и тяжелых металлов [12, 13]. Отметим также, что именно *MagA* и *Rob* являются вирулентными факторами при пиелонефрите. Делеция по генам *rob* и *marA* способствуют уменьшению проявления пиелонефрита [51]. Для исследования генетически измененной системы, когда концентрация белков невелика, было использовано стохастическое моделирование. Результаты продемонстрировали качественное соответствие с детерминистической моделью как без запаздывания, так и с запаздыванием. Развитие моделей такого типа имеют огромное значение для системной биологии, геной инженерии и медицины, так как позволяют детально изучать динамические свойства регуляторных контуров, конкретных генных сетей и поддерживать эксперимент.

Численные расчеты моделей, описывающих динамические системы со сложными регуляторными петлями, как правило, позволяют получить набор решений, характеризующихся различными типами функционирования – стационарные решения и периодические колебания. Возможны также колебания сложной природы, обусловленные через связь с другими регуляторными петлями или запаздывающим аргументом. Возможность существования сложного поведения теоретически была показана на генетической системе циркадных ритмов в дрозофиле [52], на системе гемопозза человека [53]. Теоретические исследования хаотического поведения в генных сетях представлены в ряде работ [54, 55]. Важно отметить, что данная динамика может и не проявляться *in vivo* либо быть скрыта за другими физиологическими процессами.

Кроме всего прочего, численные расчеты были сделаны с использованием высоко воспроизводительных вычислений, что позволяет в дальнейшем делать численные расчеты на системах большой размерности. Напомним, что концепция исследования генной сети *E. coli* состоит в разработке математических моделей отдельных регуляторных контуров, их анализа и последующего объединения подмоделей. Поэтому в последующих работах предполагается дополнение представленной модели подмоделями, описывающими транспорт и окислительный стресс, а также проведение экспериментов *in silico* с обновленной моделью для выяснения действия лекарственных средств на динамику системы при лечении пиелонефрита и подагры.

Авторы выражают признательность Ри Наталье Александровне за обсуждение научных результатов.

Работа выполнена при финансовой поддержке стипендиальной программы для молодых ученых Forschungsstipendien für Doktoranden und Nachwuchswissenschaftler (DAAD).

Литература

1. Лихошвай В.А., Голубятников В.П., Демиденко Г.В., Евдокимов А.А., Матвеева И.И., Фадеев С.И. Теория генных сетей // Системная компьютерная биология / Отв. ред. Н.А. Колчанов, С.С. Гончаров. В.А. Лихошвай, В.А. Иванисенко. – Новосибирск: Изд-во СО РАН, 2008. – Вып. 14. – С. 397–480.
2. Казанцев Ф.В., Акбердин И.Р., Безматерных К.Д., Лихошвай В.А. Система автоматизированной генерации математических моделей генных сетей // Вестник ВОГиС. – 2009. – Т. 13, № 1. – С. 163–169.
3. Wong P., Gladney S., Keasling J.D. Mathematical model of the lac operon: inducer exclusion, catabolite repression, and diauxic growth on glucose and lactose // Biotechnol. Prog. – 1997. – V. 13, No 2. – P. 132–143.
4. Santillan M., Mackey M.C. Dynamic behavior in mathematical models of the tryptophan operon // Chaos. – 2001. – V. 11, No 1. – P. 261–268.
5. Ратушный А.В. Исследование динамики функционирования генных сетей методами математического моделирования: Дис. ... канд. биол. наук. – Новосибирск, 2006. – 246 с.
6. Ri M.T., Khlebodarova T.M., Likhoshvai V.A. The mathematical model of genetic regulation of pyrimidine biosynthesis in *Escherichia Coli*. – 2008. – URL: <https://dl.dropbox.com/u/52461630/THE-MATHEMATICAL-MODEL-OF-GENETIC-REGULATION-OF-PYRIMIDINE-BIOSYNTHESIS-IN-ESCHERICHIA-COLI.pdf>, свободный.

7. Белов О.В., Красавин Е.А., Пархоменко А.Ю. Математическая модель индуцированного мутационного процесса в бактериальных клетках *Escherichia coli* при ультрафиолетовом облучении // Радиационная биология. Радиоэкология. – 2009. – Т. 49, № 5. – С. 617–628.
8. Martínez-Antonio A., Lomnitz J.G., Sandoval S., Aldana M., Savageau M.A. Regulatory design governing progression of population growth phases in bacteria // PLoS One. – 2012. – V. 7, No 2. – P. e30654-1–e30654-12. – doi: 10.1371/journal.pone.0030654.
9. Пу Н.А., Хлебодарова Т.М., Козай В.В., Фадеев С.И., Лихошвай В.А. Бистабильность системы утилизации нитрита *Escherichia coli*: анализ математической модели // Сиб. журн. индустр. матем. – 2012. – Т. 15, № 4. – С. 110–117.
10. Tanouchi Y., Pai A., Buchler N.E., You L. Programming stress-induced altruistic death in engineered bacteria // Mol. Syst. Biol. – 2012. – V. 8. – P. 626-1–626-11. – doi: 10.1038/msb.2012.57.
11. Пу М.Т. Анализ и математическое моделирование генной сети *Escherichia coli* // Учен. зап. Казан. ун-та. Сер. Естеств. науки. – 2012. – Т. 154, кн. 2. – С. 247–262.
12. Kwon H.J., Bennik M.H., Demple B., Ellenberger T. Crystal structure of the *Escherichia coli* Rob transcription factor in complex with DNA // Nat. Struct. Biol. – 2000. – V. 7, No 5. – P. 424–430.
13. Domínguez-Cuevas P., Ramos J.-L., Marqués S. Sequential XylS-CTD binding to the Pm promoter induces DNA bending prior to activation // J. Bacteriol. – 2010. – V. 192, No 11. – P. 2682–2690. – doi: 10.1128/JB.00165-10.
14. Martin R.G., Rosner J.L. Fis, an accessory factor for transcriptional activation of the mar (multiple antibiotic resistance) promoter of *Escherichia coli* in the presence of the activator MarA, SoxS, or Rob // J. Bacteriol. – 1997. – V. 179, No 23. – P. 7410–7409.
15. Schneiders T., Levy S.B. MarA-mediated transcriptional repression of the rob promoter // J. Biol. Chem. – 2006. – V. 281, No 15. – P. 10049–10055.
16. Martin R.G., Rosner J.L. Promoter discrimination at class I MarA regulon promoters mediated by glutamic acid 89 of the MarA transcriptional activator of *Escherichia coli* // J. Bacteriol. – 2011. – V. 193, No 2. – P. 506–515. – doi: 10.1128/JB.00360-10.
17. Wilkinson S.P., Grove A. Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins // Curr. Issues Mol. Biol. – 2006. – V. 8, No 1. – P. 51–62.
18. Martin R.G., Rosner J.L. Transcriptional and translational regulation of the marRAB multiple antibiotic resistance operon in *Escherichia coli* // Mol. Microbiol. – 2004. – V. 53, No 1. – P. 183–191.
19. Wall M.E., Markowitz D.A., Rosner J.L., Martin R.G. Model of transcriptional activation by MarA in *Escherichia coli* // PLoS Comput. Biol. – 2009. – V. 5, No 12. – P. e1000614-1–e1000614-11. – doi: 10.1371/journal.pcbi.1000614.
20. Shimada T., Yamamoto K., Ishihama A. Novel members of the Cra regulon involved in carbon metabolism in *Escherichia coli* // J. Bacteriol. – 2011. – V. 193, No 3. – P. 649–659. – doi: 10.1128/JB.01214-10.
21. Skarstad K., Thöny B., Hwang D.S., Kornberg A. A novel binding protein of the origin of the *Escherichia coli* chromosome // J. Biol. Chem. – 1993. – V. 268, No 8. – P. 5365–5370.
22. Varshavsky A. The N-end rule pathway and regulation by proteolysis // Protein Sci. – 2011. – V. 20, No 8. – P. 1298–1345. – doi: 10.1002/pro.666.
23. Ali Azam T., Iwata A., Nishimura A., Ueda S., Ishihama A. Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid // J. Bacteriol. – 1999. – V. 181, No 20. – P. 6361–6370.

24. Rosner J.L., Dangi B., Gronenborn A.M., Martin R.G. Posttranscriptional activation of the transcriptional activator Rob by dipyrindyl in *Escherichia coli* // J. Bacteriol. – 2002. – V. 184, No 5. – P. 1407–1416.
25. Lu P., Vogel C., Wang R., Yao X., Marcotte E.M. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation // Nat. Biotechnol. – 2007. – V. 25, No 1. – P. 117–124.
26. Carstea A.S., Grecu A.T. Time delay effects on oscillatory signals in genetic repressilators with two promoters // Rom. Rep. Phys. – 2010. – V. 62, No 1. – P. 169–178.
27. Gillespie D.T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions // J. Comput. Phys. – 1976. – V. 22, No 4. – P. 403–434.
28. Bratsun D., Volfson D., Tsimring L.S., Hasty J. Delay-induced stochastic oscillations in gene regulation // Proc. Natl. Acad. Sci. USA. – 2005. – V. 102, No 41. – P. 14593–14598.
29. Martin R.G., Gillette W.K., Martin N.I., Rosner J.L. Complex formation between activator and RNA polymerase as the basis for transcriptional activation by MarA and SoxS in *Escherichia coli* // Mol. Microbiol. – 2002. – V. 43, No 2. – P. 355–370.
30. Rosenfeld N., Elowitz M.B., Alon U. Negative autoregulation speeds the response times of transcription networks // J. Mol. Biol. – 2002. – V. 323, No 5. – P. 785–793.
31. Friedman N., Vardi S., Ronen M., Alon U., Stavans J. Precise temporal modulation in the response of the SOS DNA repair network in individual bacteria // PLoS Biol. – 2005. – V. 3, No 7. – P. 1261–1268. – doi: 10.1371/journal.pbio.0030238.
32. Wunderlich Z., Mirny L.A. Spatial effects on the speed and reliability of protein–DNA search // Nucleic Acids Res. – 2008. – V. 36, No 11. – P. 3570–3578. – doi: 10.1093/nar/gkn173.
33. Hilbert L., Albrecht D., Mackey M.C. Small delay, big waves: a minimal delayed negative feedback model captures *Escherichia coli* single cell SOS kinetics // Mol. Biosyst. – 2011. – V. 7, No 9. – P. 2599–2607. – doi: 10.1039/c1mb05122a.
34. Samoilo M.S., Price G., Arkin A.P. From fluctuations to phenotypes: the physiology of noise // Sci. STKE. – 2006. – V. 2006, No 366. – doi: 10.1126/stke.3662006re17.
35. Acar M., Mettetal J.T., van Oudenaarden A. Stochastic switching as a survival strategy in fluctuating environments // Nat. Genet. – 2008. – V. 40, No 4. – P. 471–475. – doi: 10.1038/ng.110.
36. Elowitz M.B., Levine A.J., Siggia E.D., Swain P.S. Stochastic gene expression in a single cell // Science. – 2002. – V. 297. – P. 1183–1186.
37. Golding I., Paulsson J., Zawilski S.M., Cox E.C. Real-time kinetics of gene activity in individual bacteria // Cell. – 2005. – V. 123, No 6. – P. 1025–1036. – doi: 10.1016/j.cell.2005.09.031.
38. Kaern M., Elston T.C., Blake W.J., Collins J.J. Stochasticity in gene expression: from theories to phenotypes // Nat. Rev. Genet. – 2005. – V. 6, No 6. – P. 451–464.
39. Cai L., Friedman N., Xie X.S. Stochastic protein expression in individual cells at the single molecule level // Nature. – 2006. – V. 440. – P. 358–362.
40. Yu J., Xiao J., Ren X., Lao K., Xie X.S. Probing gene expression in live cells, one protein molecule at a time // Science. – 2006. – V. 311. – P. 1600–1603.
41. Браун Д.А., Захаров А.П. Моделирование пространственно-временной динамики циркадных ритмов *Neurospora crassa* // Компьютерные исследования и моделирование. – 2011. – Т. 3, № 2. – С. 191–213.
42. Kurokawa K., Nishida S., Emoto A., Sekimizu K., Katayama T. Replication cycle-coordinated change of the adenine nucleotide-bound forms of DnaA protein in *Escherichia coli* // EMBO J. – 1999. – V. 18, No 23. – P. 6642–6652.

43. *Kaguni J.M.* DnaA: controlling the initiation of bacterial DNA replication and more // *Annu. Rev. Microbiol.* – 2006. – V. 60. – P. 351–375.
44. *Griffith J.S.* Mathematics of cellular control processes. I. Negative feedback to one gene // *J. Theor. Biol.* – 1968. – V. 20, No 2. – P. 202–208.
45. *Tyson J.J.* On the existence of oscillatory solutions in negative feedback cellular control processes // *J. Math. Biol.* – 1974. – V. 1, No 4. – P. 311–315.
46. *Xiao M., Cao J.* Genetic oscillation deduced from Hopf bifurcation in a genetic regulatory network with delays // *Math. Biosci.* – 2008. – V. 215, No 1. – P. 55–63.
47. *Nguyen L.K., Kulasiri D.* On the functional diversity of dynamical behaviour in genetic and metabolic feedback systems // *BMC Syst. Biol.* – 2009. – V. 3. – P. 51-1–51-30. – doi: 10.1186/1752-0509-3-51.
48. *Лихошвай В.А., Матушкин Ю.Г., Фадеев С.И.* Задачи функционирования генных сетей // *Сиб. журн. индустр. матем.* – 2003. – Т. 2, № 14. – С. 64–80.
49. *Штокало Д.Н.* О предельном переходе к уравнению с запаздывающим аргументом в модели синтеза вещества с учетом обратимости и стоков // *Сиб. журн. индустр. матем.* – 2009. – Т. 2, № 38. – С. 143–156.
50. *Фадеев С.И., Лихошвай В.А., Штокало Д.Н., Королев В.К.* Об исследовании математических моделей матричного синтеза нерегулярных полимеров ДНК, РНК и белков // *Сиб. электрон. матем. изв.* – 2010. – Т. 7. – С. 467–475.
51. *Casaz P., Garrity-Ryan L.K., McKenney D., Jackson C., Levy S.B., Tanaka S.K., Aleksun M.N.* MarA, SoxS and Rob function as virulence factors in an *Escherichia coli* murine model of ascending pyelonephritis // *Microbiology.* – 2006. – V. 152, Pt. 12. – P. 3643–3650.
52. *Gonze D., Halloy J., Leloup J.C., Goldbeter A.* Stochastic models for circadian rhythms: effect of molecular noise on periodic and chaotic behaviour // *C. R. Biol.* – 2003. – V. 326, No 2. – P. 189–203.
53. *Mackey M.C., Glass L.* Oscillation and chaos in physiological control systems // *Science.* – 1977. – V. 197. – P. 287–289.
54. *Goldbeter A., Gonze D., Houart G., Leloup J.C., Halloy J., Dupont G.* From simple to complex oscillatory behavior in metabolic and genetic control networks // *Chaos.* – 2001. – V. 11, No 1. – P. 247–260.
55. *Likhoshvai V.A., Fadeev S.I., Kogai V.V., Khlebodarova T.M.* On the chaos in gene networks // *J. Bioinform. Comput. Biol.* – 2013. – V. 11, No 1. – P. 1340009-1–1340009-25. – doi: 10.1142/S021972001340009X.

Поступила в редакцию
25.02.13

Ри Максим Тексенович – программист, ООО «АкадемДжин», г. Новосибирск, Россия.

E-mail: maxim.ree@gmail.com

Хайрулин Сергей Сергеевич – младший научный сотрудник, Институт систем информатики им. А.П. Ершова СО РАН, г. Новосибирск, Россия.

E-mail: skhayrulin@iis.nsk.su

Сайк Ольга Владимировна – младший научный сотрудник, Институт цитологии и генетики СО РАН, г. Новосибирск, Россия

E-mail: sayk.olga@gmail.com

* * *

**MATHEMATICAL MODEL OF Rob, MarR, MarA
REGULATORY CIRCUIT
OF *Escherichia coli* GENE NETWORK**

M.T. Ri, S.S. Khairulin, O.V. Saik

Abstract

This paper presents the results of the analysis of a mathematical model describing the regulatory circuit of the *Escherichia coli* gene network, which consists of three proteins Rob, MarR, and MarA. The choice of the circuit is due to the analysis of the mechanisms of forward and backward links taking into account the delay argument (regulation delay). The results of the model calculation demonstrate that the system may have a steady-state solution and undamped oscillations. Both modes are stable with respect to intrinsic noise and periodic extrinsic force. The numerical analysis of the model shows that the period of the limit cycle depends linearly on the delay argument, while the amplitude – nonlinearly. It is also revealed that the duplication of the *marR* gene leads to the solutions of a complex nature. At this stage two versions of the model are developed: deterministic and stochastic. Parallel computations allowed a 12-fold reduction of the calculation time for the model without a delay and a 640-fold reduction for the model with a delay.

Keywords: gene network, mathematical modeling, stochastic systems, oscillations, delay argument, *Escherichia coli*, OpenCL, parallel computations.

References

1. Likhoshvai V.A., Golubyatnikov V.P., Demidenko G.V., Evdokimov A.A., Matveeva I.I., Fadeev S.I. Theory of Gene Networks. N.A. Kolchanov, S.S. Goncharov, V.A. Likhoshvai, V.A. Ivanisenko (Eds.) *Sistemnaya kompyuternaya biologiya* [Systems Computer Biology]. Novosibirsk: Izd. SO RAN, 2008, no. 14, pp. 397–480. (In Russian)
2. Kazantsev F.V., Akberdin I.R., Bezmaternykh K.D., Likhoshvai V.A. System of automated generation of mathematical models of gene networks. *Vestnik VOGiS*, 2009, vol. 13, no. 1, pp. 163–169. (In Russian)
3. Wong P., Gladney S., Keasling J.D. Mathematical model of the lac operon: inducer exclusion, catabolite repression, and diauxic growth on glucose and lactose. *Biotechnol. Prog.*, 1997, vol. 13, no. 2, pp. 132–143.
4. Santillan M., Mackey M.C. Dynamic behavior in mathematical models of the tryptophan operon. *Chaos*, 2001, vol. 11, no. 1, pp. 261–268.
5. Ratushnyi A.V. Research on the dynamics of gene network functioning by the methods of mathematical modeling: Cand. Biol. Sci. Diss. Novosibirsk, 2006. 246 p. (In Russian)
6. Ri M.T., Khlebodarova T.M., Likhoshvai V.A. The mathematical model of genetic regulation of pyrimidine biosynthesis in *Escherichia Coli*. 2008. Available at: <https://dl.dropbox.com/u/52461630/THE-MATHEMATICAL-MODEL-OF-GENETIC-REGULATION-OF-PYRIMIDINE-BIOSYNTHESIS-IN-ESCHERICHIA-COLI.pdf>.
7. Belov O.V., Krasavin E.A., Parkhomenko A.Yu. Mathematical model of induced mutagenesis in the bacterial cells of *Escherichia coli* under ultraviolet irradiation. *Radiatsionnaya biologiya. Radioekologiya*, 2009, vol. 49, no. 5, pp. 617–628. (In Russian)
8. Martinez-Antonio A., Lomnitz J.G., Sandoval S., Aldana M., Savageau M.A. Regulatory design governing progression of population growth phases in bacteria. *PLoS One*, 2012, vol. 7, no. 2, pp. e30654-1–e30654-12. doi: 10.1371/journal.pone.0030654.
9. Ri N.A., Khlebodarova T.M., Kogai V.V., Fadeev S.I., Likhoshvai V.A. Bistability of the utilization system of nitrite in *Escherichia coli*: Analysis of the mathematical model. *Sibirskii Zh. Industr. Mat.* 2012, vol. 15, no. 4, pp. 110–117. (In Russian)
10. Tanouchi Y., Pai A., Buchler N.E., You L. Programming stress-induced altruistic death in engineered bacteria. *Mol. Syst. Biol.*, 2012, vol. 8, pp. 626-1–626-11. doi: 10.1038/msb.2012.57.

11. Ri M.T. Analysis and Mathematical Modeling of *Escherichia coli* Gene Network. *Uchenye Zapiski Kazanskogo Universiteta. Seriya Estestvennye Nauki*, 2012, vol. 154, no. 2, pp. 247–262. (In Russian)
12. Kwon H.J., Bennik M.H., Demple B., Ellenberger T. Crystal structure of the *Escherichia coli* Rob transcription factor in complex with DNA. *Nat. Struct. Biol.*, 2000, vol. 7, no. 5, pp. 424–430.
13. Domínguez-Cuevas P., Ramos J.-L., Marqués S. Sequential XylS-CTD binding to the Pm promoter induces DNA bending prior to activation. *J. Bacteriol.*, 2010, vol. 192, no. 11, pp. 2682–2690. doi: 10.1128/JB.00165-10.
14. Martin R.G., Rosner J.L. Fis, an accessorial factor for transcriptional activation of the mar (multiple antibiotic resistance) promoter of *Escherichia coli* in the presence of the activator MarA, SoxS, or Rob. *J. Bacteriol.*, 1997, vol. 179, no. 23, pp. 7410–7409.
15. Schneiders T., Levy S.B. MarA-mediated transcriptional repression of the rob promoter. *J. Biol. Chem.*, 2006, vol. 281, no. 15, pp. 10049–10055.
16. Martin R.G., Rosner J.L. Promoter discrimination at class I MarA regulon promoters mediated by glutamic acid 89 of the MarA transcriptional activator of *Escherichia coli*. *J. Bacteriol.*, 2011, vol. 193, no. 2, pp. 506–515. doi: 10.1128/JB.00360-10.
17. Wilkinson S.P., Grove A. Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins. *Curr. Issues Mol. Biol.*, 2006, vol. 8, no. 1, pp. 51–62.
18. Martin R.G., Rosner J.L. Transcriptional and translational regulation of the marRAB multiple antibiotic resistance operon in *Escherichia coli*. *Mol. Microbiol.*, 2004, vol. 53, no. 1, pp. 183–191.
19. Wall M.E., Markowitz D.A., Rosner J.L., Martin R.G. Model of transcriptional activation by MarA in *Escherichia coli*. *PLoS Comput. Biol.*, 2009, vol. 5, no. 12, pp. e1000614-1–e1000614-11. doi: 10.1371/journal.pcbi.1000614.
20. Shimada T., Yamamoto K., Ishihama A. Novel members of the Cra regulon involved in carbon metabolism in *Escherichia coli*. *J. Bacteriol.*, 2011, vol. 193, no. 3, pp. 649–659. doi: 10.1128/JB.01214-10.
21. Skarstad K., Thöny B., Hwang D.S., Kornberg A. A novel binding protein of the origin of the *Escherichia coli* chromosome. *J. Biol. Chem.*, 1993, vol. 268, no. 8, pp. 5365–5370.
22. Varshavsky A. The N-end rule pathway and regulation by proteolysis. *Protein Sci.*, 2011, vol. 20, no. 8, pp. 1298–1345. doi: 10.1002/pro.666.
23. Ali Azam T., Iwata A., Nishimura A., Ueda S., Ishihama A. Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J. Bacteriol.*, 1999, vol. 181, no. 20, pp. 6361–6370.
24. Rosner J.L., Dangi B., Gronenborn A.M., Martin R.G. Posttranscriptional activation of the transcriptional activator Rob by dipyriddyI in *Escherichia coli*. *J. Bacteriol.*, 2002, vol. 184, no. 5, pp. 1407–1416.
25. Lu P., Vogel C., Wang R., Yao X., Marcotte E.M. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.*, 2007, vol. 25, no. 1, pp. 117–124.
26. Carstea A.S., Grecu A.T. Time delay effects on oscillatory signals in genetic repressilators with two promoters. *Rom. Rep. Phys.*, 2010, vol. 62, no. 1, pp. 169–178.
27. Gillespie D.T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 1976, vol. 22, no. 4, pp. 403–434.
28. Bratsun D., Volfson D., Tsimring L.S., Hasty J. Delay-induced stochastic oscillations in gene regulation. *Proc. Natl. Acad. Sci. USA*, 2005, vol. 102, no. 41, pp. 14593–14598.
29. Martin R.G., Gillette W.K., Martin N.I., Rosner J.L. Complex formation between activator and RNA polymerase as the basis for transcriptional activation by MarA and SoxS in *Escherichia coli*. *Mol. Microbiol.*, 2002, vol. 43, no. 2, pp. 355–370.
30. Rosenfeld N., Elowitz M.B., Alon U. Negative autoregulation speeds the response times of transcription networks. *J. Mol. Biol.*, 2002, vol. 323, no. 5, pp. 785–793.
31. Friedman N., Vardi S., Ronen M., Alon U., Stavans J. Precise temporal modulation in the response of the SOS DNA repair network in individual bacteria. *PLoS Biol.*, 2005, vol. 3, no. 7, pp. 1261–1268. doi: 10.1371/journal.pbio.0030238.

32. Wunderlich Z., Mirny L.A. Spatial effects on the speed and reliability of protein–DNA search. *Nucleic Acids Res.*, 2008, vol. 36, no. 11, pp. 3570–3578. doi: 10.1093/nar/gkn173.
33. Hilbert L., Albrecht D., Mackey M.C. Small delay, big waves: a minimal delayed negative feedback model captures *Escherichia coli* single cell SOS kinetics. *Mol. Biosyst.*, 2011, vol. 7, no. 9, pp. 2599–2607. doi: 10.1039/c1mb05122a.
34. Samoilov M.S., Price G., Arkin A.P. From fluctuations to phenotypes: the physiology of noise. *Sci. STKE*, 2006, vol. 2006, no. 366. doi: 10.1126/stke.3662006re17.
35. Acar M., Mettetal J.T., van Oudenaarden A. Stochastic switching as a survival strategy in fluctuating environments. *Nat. Genet.*, 2008, vol. 40, no. 4, pp. 471–475. doi: 10.1038/ng.110.
36. Elowitz M.B., Levine A.J., Siggia E.D., Swain P.S. Stochastic gene expression in a single cell. *Science*, 2002, vol. 297, pp. 1183–1186.
37. Golding I., Paulsson J., Zawilski S.M., Cox E.C. Real-time kinetics of gene activity in individual bacteria. *Cell*, 2005, vol. 123, no. 6, pp. 1025–1036. doi: 10.1016/j.cell.2005.09.031.
38. Kaern M., Elston T.C., Blake W.J., Collins J.J. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.*, 2005, vol. 6, no. 6, pp. 451–464.
39. Cai L., Friedman N., Xie X.S. Stochastic protein expression in individual cells at the single molecule level. *Nature*, 2006, vol. 440, pp. 358–362.
40. Yu J., Xiao J., Ren X., Lao K., Xie X.S. Probing gene expression in live cells, one protein molecule at a time. *Science*, 2006, vol. 311, pp. 1600–1603.
41. Bratsun D.A., Zakharov A.P. Modeling of the spatiotemporal dynamics of circadian rhythms in *Neurospora crassa*. *Kompyuternye issledovaniya i modelirovanie*, 2011, vol. 3, no. 2, pp. 191–213. (In Russian)
42. Kurokawa K., Nishida S., Emoto A., Sekimizu K., Katayama T. Replication cycle-coordinated change of the adenine nucleotide-bound forms of DnaA protein in *Escherichia coli*. *EMBO J.*, 1999, vol. 18, no. 23, pp. 6642–6652.
43. Kaguni J.M. DnaA: controlling the initiation of bacterial DNA replication and more. *Annu. Rev. Microbiol.*, 2006, vol. 60, pp. 351–375.
44. Griffith J.S. Mathematics of cellular control processes. I. Negative feedback to one gene. *J. Theor. Biol.*, 1968, vol. 20, no. 2, pp. 202–208.
45. Tyson J.J. On the existence of oscillatory solutions in negative feedback cellular control processes. *J. Math. Biol.*, 1974, vol. 1, no. 4, pp. 311–315.
46. Xiao M., Cao J. Genetic oscillation deduced from Hopf bifurcation in a genetic regulatory network with delays. *Math. Biosci.*, 2008, vol. 215, no. 1, pp. 55–63.
47. Nguyen L.K., Kulasiri D. On the functional diversity of dynamical behaviour in genetic and metabolic feedback systems. *BMC Syst. Biol.*, 2009, vol. 3, pp. 51–1–51–30. doi: 10.1186/1752-0509-3-51.
48. Likhoshvai V.A., Matushkin Yu.G., Fadeev S.I. The problems of functioning of gene networks. *Sibirskii Zh. Industr. Matem.*, 2003, vol. 2, no. 14, pp. 64–80. (In Russian)
49. Shtokalo D.N. On the limiting passage to an equation with lagging argument in the model of substance synthesis taking into account reversibility and drainage. *Sibirskii Zh. Industr. Matem.*, 2009, vol. 2, no. 38, pp. 143–156. (In Russian)
50. Fadeev S.I., Likhoshvai V.A., Shtokalo D.N., Korolev V.K. On the investigation of mathematical models of the matrix synthesis of DNA and RNA irregular polymers and proteins. *Sib. Elektron. Matem. Izv.*, 2010, vol. 7, pp. 467–475. (In Russian)
51. Casaz P., Garrity-Ryan L.K., McKenney D., Jackson C., Levy S.B., Tanaka S.K., Alekshun M.N. MarA, SoxS and Rob function as virulence factors in an *Escherichia coli* murine model of ascending pyelonephritis. *Microbiology*, 2006, vol. 152, Pt. 12, pp. 3643–3650.
52. Gonze D., Halloy J., Leloup J.C., Goldbeter A. Stochastic models for circadian rhythms: effect of molecular noise on periodic and chaotic behaviour. *C. R. Biol.*, 2003, vol. 326, no. 2, pp. 189–203.
53. Mackey M.C., Glass L. Oscillation and chaos in physiological control systems. *Science*, 1977, vol. 197, pp. 287–289.
54. Goldbeter A., Gonze D., Houart G., Leloup J.C., Halloy J., Dupont G. From simple to complex oscillatory behavior in metabolic and genetic control networks. *Chaos*, 2001, vol. 11, no. 1, pp. 247–260.

-
55. Likhoshvai V.A., Fadeev S.I., Kogai V.V., Khlebodarova T.M. On the chaos in gene networks. *J. Bioinform. Comput. Biol.*, 2013, vol. 11, no. 1, pp. 1340009-1–1340009-25. doi: 10.1142/S021972001340009X.

Received
February 25, 2013

Ri Maksim Teksenovich – Programmer, AkademDzhin Company Ltd., Novosibirsk, Russia.
E-mail: *maxim.ree@googlemail.com*

Khairulin Sergei Sergeevich – Junior Research Fellow, A.P. Ershov Institute of Informatics Systems, Russian Academy of Sciences, Siberian Branch, Novosibirsk, Russia.
E-mail: *skhayrulin@iis.nsk.su*

Saik Olga Vladimirovna – Junior Research Fellow, Institute of Cytology and Genetics, Russian Academy of Sciences, Siberian Branch, Novosibirsk, Russia.
E-mail: *sayk.olga@gmail.com*