

М. В. Казанцев

*Алтайский государственный технический университет
пр. Ленина, 46, Барнаул, 656038, Россия*

markynaz.astu@gmail.com

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ МОДЕЛИРОВАНИЯ РАННИХ СТАДИЙ НЕКОТОРЫХ БИОЛОГИЧЕСКИХ ПРОЦЕССОВ *

Описаны программные продукты, разработанные нами для моделирования различных биологических процессов, в том числе циклических генных сетей и регуляторных контуров механорецепторов плодовой мушки. Описаны математические модели исследуемых процессов и их реализация в виде программ.

Ключевые слова: генные сети, динамические системы, моделирование, программное обеспечение, системы дифференциальных уравнений с запаздыванием.

Введение

Математическое моделирование биологических процессов широко применяется при решении проблем биоинформатики, см., например, [1]. Часто в качестве моделей таких систем системы дифференциальных уравнений (обыкновенных или с запаздыванием). Аналитическое решение подобных систем, ввиду их сложности и размера, часто бывает невозможным. Поэтому для изучения свойств таких динамических систем широко используются численные методы решения уравнений и компьютерное моделирование. К настоящему времени разработано множество программных пакетов, предоставляющих исследователям широкий набор различных численных методов для исследования указанных динамических систем.

В данной работе мы описываем ряд программных продуктов, которые были нами разработаны для моделирования и визуализации моделей некоторых биологических процессов. Примерами таких моделей являются нелинейные циклические динамические системы с положительными, отрицательными и переменными обратными связями, а также центральный регуляторный контур механорецепторов плодовой мушки *D.melanogaster*.

Генные сети – это группы координированно функционирующих генов, взаимодействующих друг с другом как через свои первичные продукты (РНК и белки), так и через разнообразные метаболиты и другие вторичные продукты функционирования генных сетей [1; 2]. Одной из простейших их разновидностей являются циклические генные сети, которые рассматривались в работах [2–4]. Это такие гипотетические сети, в которых скорость экспрессии каждого агента зависит только от концентраций этого агента и другого, предшествующего ему. В работе [5] подобные сети названы «кольцевыми». В данной работе мы рассматриваем программные продукты, направленные на моделирование таких сетей с положительными, отрицательными или переменными обратными связями. Особый интерес для исследования

* Автор выражает искреннюю благодарность Т. А. Бухариной и Д. П. Фурман (ИЦиГ СО РАН) за полезное обсуждение.

представляют устойчивые состояния и циклические режимы функционирования генных сетей. Они соответствуют стационарным точкам и периодическим траекториям, или циклам, в соответствующих динамических системах. В рамках данной работы была разработана серия программ, строящих и визуализирующих такие точки и траектории.

Одной из конкретных биологических задач, для решения которых могут быть использованы генные сети, является задача о моделировании ранних этапов развития макрохетов плодовой мушки. Макрохеты (большие щетинки) – внешние сенсорные органы с функцией механорецепторов. Их развитие начинается с выделения в крыловом имагинальном диске пронеуральных кластеров – групп из 20–30 клеток, содержащих белки *Achaete* и *Scute* (AS-C). Далее из каждой такой группы выделяется одна родительская клетка, дающая начало механорецептору. Ранее в работах [6–8] была построена генная сеть, описывающие этапы этого процесса. В данной работе мы рассмотрим систему дифференциальных уравнений, описывающую эту сеть, а также ее модифицированную версию, учитывающую эффекты запаздывания во времени между некоторыми из происходящих в биологической системе явлений.

На основе полученных дифференциальных уравнений было разработано программное обеспечение, моделирующее данную систему. Для разработки был выбран язык программирования R (см. официальный сайт ¹). Он и ранее применялся для моделирования биологических систем (см. [9-12]). Пакеты данного языка содержат множество различных решателей систем дифференциальных уравнений, в том числе и уравнений с запаздыванием. Также его пакеты предлагают широкий набор инструментов визуализации полученных результатов (в частности, в виде графиков и проекций траекторий многомерных динамических систем на плоскости размерности 2 и 3). Разработанная программа построена как Shiny-приложение ² языка R и выложена в открытом доступе ³.

Математические модели

В рамках данной работы мы рассматриваем модели биологических процессов, основанные на динамических системах, описываемых системами нелинейных дифференциальных уравнений (обыкновенных или с запаздыванием). В данной системе строится решение задачи Коши, начальным условием при этом служит точка, соответствующая начальному состоянию биологического процесса.

Циклическая генная сеть моделируется системой обыкновенных нелинейных дифференциальных уравнений вида (1).

$$\frac{dx_i}{dt} = f_i(x_{i-1}) - x_i, \quad x \in 1..n, \text{ при } i = 1 \quad x_{i-1} = x_n. \quad (1)$$

Здесь f_i – функции, моделирующие положительные, отрицательные или переменные обратные связи. Они являются соответственно монотонно возрастающими, монотонно убывающими или унимодальными. Мы также рассматривали модели со ступенчатыми (частично-постоянными) функциями f_i . Для удобства в дальнейшем будем обозначать их L-, Г- и П-функциями. Некоторые системы такого класса рассматривались и ранее (см. [2, 3, 13–17]).

В работах [18–20] установлены необходимые условия существования циклов в различных частях фазового пространства систем с L- и Г-функциями, а также определены области, в которых эти циклы могут существовать. В работе [21] установлено, что для систем с П-функциями существуют некоторые аналогии, позволяющие частично сводить поиск областей существования циклов в них к системам с L- и Г-функциями.

Регуляторный контур механорецепторов. Развитие механорецепторов плодовой мушки занимает около 14–16 часов и разбито на два этапа. Первый этап характеризуется отсутствием в клетке белка Phyllopod. Он занимает первые 12 часов развития. На втором этапе начинается экспрессия соответствующего ему гена. Переменные данных систем соответствуют концентрациям белков, концентрации которых изменяется: AS-C (x), Hairy (y), Senseless (z), Scratch (u), Charlatan (w) и Phyllopod (p). Концентрации белков Daughterless (DA/D),

¹ <https://www.r-project.org/>

² <http://shiny.rstudio.com/>

³ <https://maxim-kazantsev.shinyapps.io/AS-C/>

Extramacrochaete (EMC/E), Groucho (Gro/G), Seven in one (SINA/S) и Ubiquitin (Ub/U) остаются неизменными. В простейшей модели развитие механорецепторов описывается динамической системой (2):

$$\left\{ \begin{array}{l} \frac{dx}{dt} = \frac{\sigma_1(Dx) + \sigma_3(z) + \sigma_5(w)}{(1 + Gy)(1 + Ex)} - (1 + pUS)x, \\ \frac{dy}{dt} = \frac{C}{1 + u} - y, \\ \frac{dz}{dt} = s_3(Dx) - z, \\ \frac{du}{dt} = s_4(Dx) - u, \\ \frac{dw}{dt} = s_5(Dx) - w, \\ \frac{dp}{dt} = \frac{s_6(Dx) h(t - 12)(t - 12)^2}{1 + (t + 12)^2} - p. \end{array} \right. \quad (2)$$

Данная модель, в частности, рассматривается в работе [6]. Здесь $h(v)$ – функция Хевисайда, σ_i и s_i – сигмоидные функции, имеющие вид

$$\sigma_i(v) = \frac{d_i v^2}{1 + v^2}, \quad s_i(v) = \frac{a_i \exp\left(\frac{x - b_i}{c_i}\right)}{1 + \exp\left(\frac{x - b_i}{c_i}\right)}.$$

Здесь a_i, b_i, c_i, d_i – некоторые константы. Таким образом, на первом этапе $p = 0$.

Мы также рассматриваем более сложную модель, описываемую системой дифференциальных уравнений с запаздыванием. Она учитывает то обстоятельство, что в действительности между моментом, когда начинает синтезироваться белок Phyllopod, и моментом, когда он начинает разлагать AS-C, существует заметный временной промежуток. Чтобы учесть этот эффект, мы заменяем первое уравнение системы (2) на

$$\frac{dx}{dt} = \frac{\sigma_1(Dx) + \sigma_3(z) + \sigma_5(w)}{(1 + Gy)(1 + Ex)} - (1 + p(t - \Delta t)US)x,$$

здесь Δt – время запаздывания, $p(t - \Delta t)$ – концентрация Phyllopod в момент времени $(t - \Delta t)$.

Также интерес в биологических исследованиях представляют мутанты – особи *D. melanogaster*, у которых некоторые из белков не синтезируются вообще, либо их концентрация заметно ниже, чем у прочих особей. Чтобы моделировать генные сети мутантов, для каждого из участвующих в системе генов был введен коэффициент мутации (соответственно k_x, k_y и т. д.), который находится в интервале $[0; 1]$ (значение коэффициента 0 соответствует полному отсутствию соответствующего гена белка, 1 – его нормальной концентрации). Таким образом, итоговая система приобретает вид (3).

$$\left\{ \begin{array}{l} \frac{dx}{dt} = k_x \frac{\sigma_1(Dx) + \sigma_3(z) + \sigma_5(w)}{(1 + Gy)(1 + Ex)} - (1 + p(t - \Delta t)US)x, \\ \frac{dy}{dt} = k_y \frac{C}{1 + u} - y, \\ \frac{dz}{dt} = k_z s_3(Dx) - z, \\ \frac{du}{dt} = k_u s_4(Dx) - u, \\ \frac{dw}{dt} = k_w s_5(Dx) - w, \\ \frac{dp}{dt} = k_p \frac{s_6(Dx) h(t - 12)(t - 12)^2}{1 + (t + 12)^2} - p. \end{array} \right. \quad (3)$$

В начальный момент времени у немутировавшей особи концентрации белков AS-C, Hairy и Senseless выше нуля и равны соответственно некоторым значениям x_0, y_0, z_0 . Концентрации остальных белков равны нулю. Таким образом, начальное состояние динамической системы $s_0 = (k_x x_0, k_y y_0, k_z z_0, 0, 0, 0)$. Динамика концентрации каждого из рассматриваемых генных комплексов описывается решением задачи Коши для системы (3) с начальным условием $s(t=0) = s_0$.

Описание программных продуктов

Циклические генные сети. Для моделирования поведения циклических генных сетей с L-, Г- и П-функциями нами был разработан набор программных продуктов на языке R, которые позволяют задавать динамические системы вида (1) и по заданным начальным условиям строить в них траектории, которые впоследствии могут быть визуализированы. Для решения задачи Коши системы (1) используется численный метод *lsoda* пакета *deSolve* языка R (см. [22–24]). Данный пакет был разработан в Ливерморской Национальной Лабораторией США и предназначается для решения систем жестких и нежестких систем дифференциальных уравнений. Содержащиеся в нем численные методы имеют низкоуровневую имплементацию на языке Fortran, доступ к ним осуществляется через программный интерфейс приложения (API) языка R.

Для визуализации проекций построенных траекторий на трехмерное пространство в виде интерактивных векторных графиков использовался пакет *RGL*⁴ языка R. Функция *plot3d* данного пакета позволяет создавать интерактивные трехмерные модели, позволяющие взглянуть на траекторию с разных сторон и при различном масштабе. На рис. 1 представлены примеры траекторий в трехмерных циклических динамических системах с кусочно-линейными правыми частями вида ЛГП и ЛГГ.

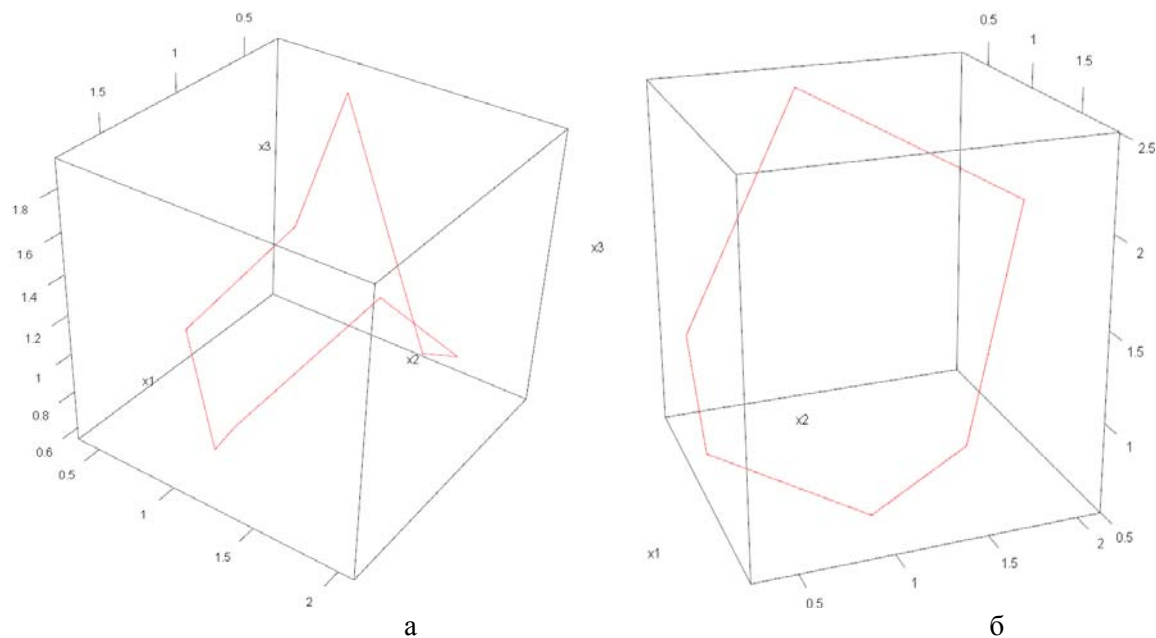


Рис. 1. Примеры циклов некоторых трехмерных систем вида (1):
а – ЛГП; б – ЛГГ

⁴ <https://cran.r-project.org/web/packages/rgl/index.html>

AS-C dynamics

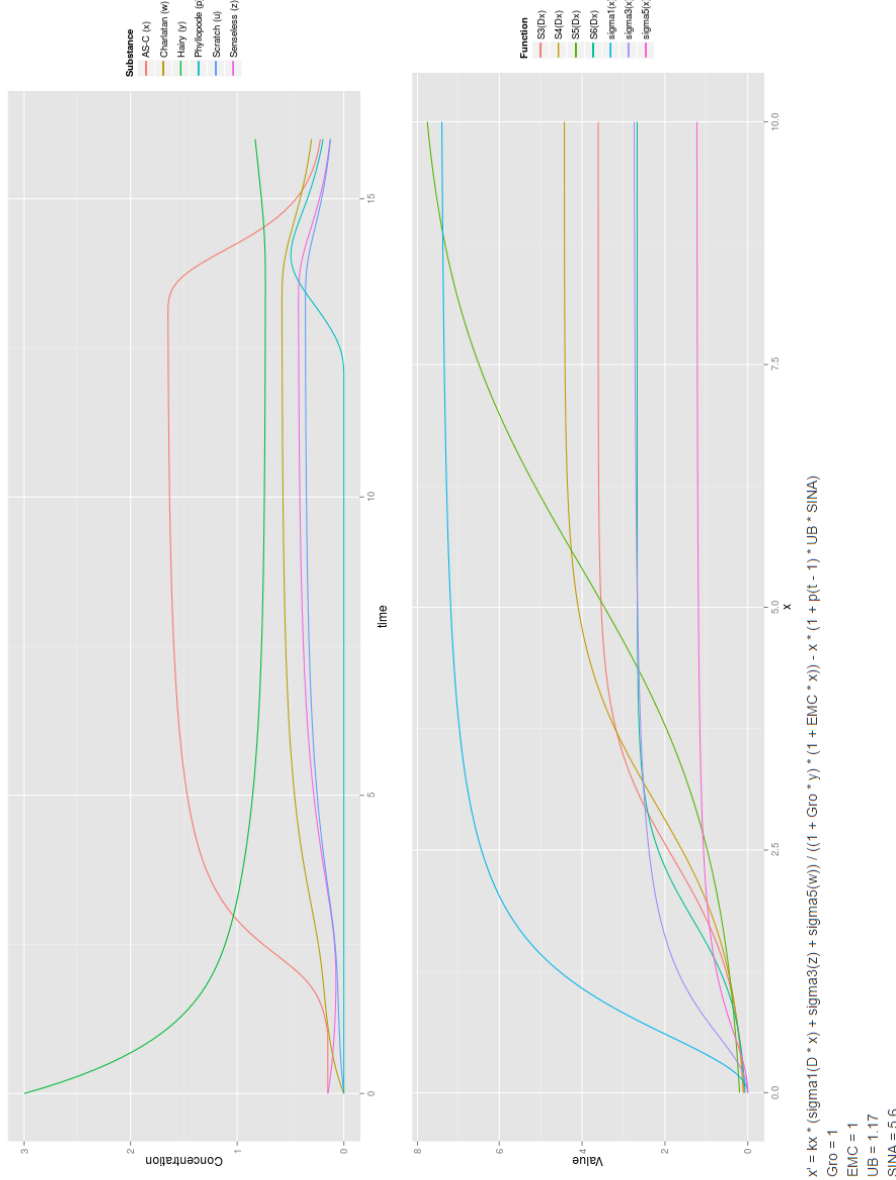
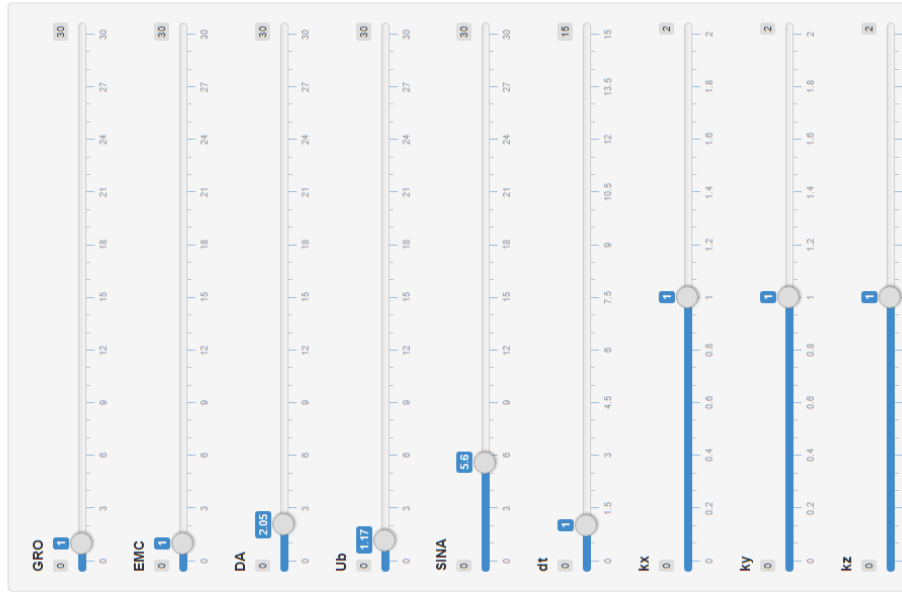


Рис. 2. Графический интерфейс приложения

Рендеринг траекторий осуществляется с использованием решений, основанных на открытой графической библиотеке *OpenGL*. Существует также возможность сохранения отдельных графиков и траекторий в форматах *png* и *jpeg*.

Разработанные скрипты также поддерживают ряд дополнительных возможностей:

- 1) поиск точек пересечений траекториями с плоскостями, параллельными координатным и проходящим через стационарные точки системы вида (1);
- 2) трассировка траекторий при переходе между этими плоскостями;
- 3) автоматическое определение заикливания траектории или попадания ее в стационарную точку;
- 4) возможность построения нескольких траекторий в рамках одной модели.

В дальнейшем планируется работа по расширению функциональности и разработка единого графического интерфейса для всех созданных скриптов с целью объединения их в целостный программный продукт. Разработка ведется на языке R с использованием бесплатных программных пакетов с открытым исходным кодом.

Регуляторный контур механорецепторов. Программное обеспечение для моделирования регуляторного контура механорецепторов представляет собой Shiny-приложение, написанное на языке R. Приложение разделено на клиентскую и серверную часть. Клиентская часть представляет собой Web-страницу с графическим интерфейсом программы. Элементы интерфейса, ответственные за ввод, сгруппированы в левой части страницы (рис. 2) и представляют собой набор ползунков для ввода значений:

- 1) концентраций белков DA, EMC, Gro, SINA и Ub.
- 2) времени запаздывания δt ;
- 3) коэффициентов мутации k_i ;
- 4) параметров сигмоидных функций a_i , b_i , c_i и d_i .

В правой части страницы находятся элементы вывода, а именно:

- 1) графики концентраций белков $x(t)$, $y(t)$, $z(t)$, $u(t)$, $w(t)$, $p(t)$;
- 2) графики вспомогательных функций $s_i(t)$ и $\sigma_i(t)$;
- 3) формулы системы (3) с подставленными в них числовыми коэффициентами.

Серверная часть приложения представляет собой скрипт на языке R, который на основе выставленных пользователем значений параметров моделирует поведение регуляторного контура. Для решения краевой задачи Коши используется численный метод *lsoda*⁵ пакета *deSolve* языка R с переменным шагом интегрирования. Данный метод реализован на языке Fortran и обладает высокой производительностью.

Архитектура Shiny-приложения позволяет автоматически реагировать на изменение параметров системы пользователем и обновлять выходные данные в соответствии с ними в режиме реального времени. Программный продукт находится в свободном доступе в сети Интернет по адресу <https://maxim-kazantsev.shinyapps.io/AS-C/>.

Список литературы

1. Колчанов Н. А., Игнатьева Е. В., Подколотная О. А. и др. Генные сети // Вавиловский журнал генетики и селекции. 2013. Т. 17, № 4/2. С. 833–850.
2. Лихошвай В. А., Голубятников В. П., Демиденко Г. В. и др. Теория генных сетей // Системная компьютерная биология. Новосибирск: Изд-во СО РАН, 2008. С. 395–480.
3. Glass L., Pasternack J. S. Stable Oscillations in Mathematical Models of Biological Control Systems // J. Math. Biology. 1978. No. 6. P. 207–223.
4. Glass L. Combinatorial aspects of dynamics in biological systems // Statistical mechanics and statistical methods in theory and application. 1977. P. 585–611.
5. Колесов А. Ю., Розов Н. Х., Садовничий В. А. Периодические решения типа бегущих волн в кольцевых генных сетях // Изв. РАН. Серия математическая. 2016. Т. 80, № 3. С. 67–94.
6. Furman D. P., Bukharina T. A., Golubyatnikov V. P. A model study of the morphogenesis of *Drosophila melanogaster* mechanoreceptors: The central regulatory circuit // Journal of Bioinformatics and Computational Biology. 2015. Vol. 13. No. 1. 1540006 (15 pages).

⁵<http://www.inside-r.org/packages/cran/deSolve/docs/lsoda>

7. Furman D. P., Bukharina T. A. How *Drosophila melanogaster* forms its mechanoreceptors // *Curr Genomics*. 2008. No. 8. P. 312—323.
8. Furman D. P., Bukharina T. A. Genetic control of macrochaetae development in *Drosophila melanogaster* // *Russ. Journ. of Dev. Biol.* 2008. No. 39. P. 195–206.
9. Акинъшин А. А. Пример использования языка R для решения задач биоинформатики // Журнал «Горизонты образования». 10-я Всероссийская научно-техническая конференция студентов, аспирантов и молодых ученых «Наука и молодежь». Барнаул, 2013. С. 6–8.
10. Казанцев М. В. Сравнение функций моделирования генных сетей // XXII Конференция серии «Математика. Компьютер. Образование». Тезисы. Пущино, 2015. С. 132.
11. Казанцев М. В., Акинъшин А. А., Голубятников В. П. Сравнительный анализ некоторых численных методов моделирования генных сетей с использованием языка R // Тр. междунар. конф. «Ломоносовские чтения на Алтае. Фундаментальные проблемы науки и образования». Барнаул, 2014. С. 548–554.
12. Казанцев М. В. Анализ некоторых методов моделирования генных сетей с использованием языка R // Тр. междунар. конф. «Математическое и компьютерное моделирование». Барнаул, . 2014. С. 8.
13. Gedeon T. Cyclic feedback systems // *Memoirs of AMS*. 1998. Vol. 134. No. 637. P. 72.
14. Golubyatnikov V. P., Golubyatnikov I. V. On periodic trajectories in odd-dimensional gene network models // *Russian Journal of Numerical Analysis and Mathematical Modelling*. 2011. Vol. 26. No. 4. P. 397–412.
15. Акинъшин А. А., Голубятников В. П., Голубятников И. В. О некоторых многомерных моделях функционирования генных сетей // *Сибирский журнал индустриальной математики*. 2013. Т. XVI, № 1 (53). С. 3–9.
16. Акинъшин А. А., Голубятников В. П. Циклы в симметричных динамических системах // *Вестн. Новосиб. гос. ун-та. Серия: Математика, механика, информатика*. 2012. Т. 12, № 2. С. 3–12.
17. Казанцев М. В., Голубятников В. П. Дискретные модели некоторых нелинейных динамических систем // XII Всерос. науч.-техн. конф. студентов, аспирантов и молодых ученых «Наука и молодежь 2015». Барнаул, 2015. С. 9–10.
18. Казанцев М. В. О некоторых свойствах графов доменов динамических систем // *Сибирский журнал индустриальной математики*. 2015. Т. 18, № 4 (64). С. 42–48.
19. Голубятников В. П., Казанцев М. В., Каленых А. Э. О дискретизации фазовых портретов некоторых динамических систем // *Математика и ее приложения: фундаментальные проблемы науки и техники: Сб. тр. всерос. конф.* Барнаул, 2015. С. 28–31.
20. Казанцев М. В. Некоторые свойства дискретных моделей циклических генных сетей // *Ломоносовские чтения на Алтае: фундаментальные проблемы науки и образования: Сб. науч. ст. междунар. конф.* Барнаул, 2015. С. 671.
21. Golubyatnikov V., Kazantsev M. On control of cycles in some gene networks models // *International conference “Metric Structures and Control Systems”*. Abstracts. Novosibirsk, 2015. P. 17–19.
22. Soetaert K., Petzoldt Th., Setzer R. Woodrow. Solving Differential Equations in R: Package deSolve // *Journal of Statistical Software*. 2010. Vol. 33. No. 9. P. 1–25. URL: <http://www.jstatsoft.org/v33/i09>.
23. Hindmarsh C. A. ODEPACK, A Systematized Collection of ODE Solvers // *Scientific Computing*. 1983. P. 55–64.
24. Petzold R. L. Automatic Selection of Methods for Solving Stiff and Nonstiff Systems of Ordinary Differential Equations // *SIAM Journal on Scientific and Statistical Computing*. 1983. No. 4. P. 136–148.

M. V. Kazantsev

*Altai State Technical University
46 Lenin Ave., Barnaul, 656038, Russian Federation*

markynaz.astu@gmail.com

**SOFTWARE FOR MODELING
OF EARLY STAGES OF SOME BIOLOGICAL PROCESSES**

This article describes software products for modeling of some biological objects, such as cyclic gene networks and regulatory circuits of mechanoreceptors of fruit fly. We describe mathematical models of these processes and their software implementation.

Keywords: gene networks, dynamical systems, modeling, software, delayed differential equations.

References

1. Kolchanov N.A., Ignatieva E.V, Podkolodnaya O.A. et al. Gene networks // Vavilov journal of genetics and selection. 2013. T. 17, № 4/2. P. 833–850.
2. Likhoshvai V.A., Golubyatnikov V.P., Demidenko G.V. et al. Theory of gene networks // System computer biology. Novosibirsk: publ. SB RAS, 2008. P. 395–480.
3. Glass Leon, Pasternack Joel S. Stable Oscillations in Mathematical Models of Biological Control Systems // J. Math. Biology. 1978. no. 6. P. 207–223.
4. Glass Leon. Combinatorial aspects of dynamics in biological systems // Statistical mechanics and statistical methods in theory and application. 1977. P. 585–611.
5. A. Yu. Kolesov, N. Kh. Rozov, V. A. Sadovnichii. Periodic solutions of travelling-wave type in circular gene networks // Izvestiya: Mathematics, 2016, 80:3, 523–548.
6. Furman D.P., Bukharina T.A., Golubyatnikov V.P. A model study of the morphogenesis of *Drosophila melanogaster* mechanoreceptors: The central regulatory circuit // Journal of Bioinformatics and Computational Biology. 2015. vol. 13 no. 1. 1540006 (15 pages).
7. Furman D.P., Bukharina T.A. How *Drosophila melanogaster* forms its mechanoreceptors // Curr Genomics. 2008. no. 8. P. 312–323.
8. Furman D.P., Bukharina T.A. Genetic control of macrochaetae development in *Drosophila melanogaster* // Russ J Dev Biol. 2008. no. 39. P. 195–206.
9. Akinshin A.A. Example of using R language for bioinformatics problem solving // Education horizons: Information technologies. 2013. P. 6–8.
10. Kazantsev M.V. Comparison of gene networks modeling functions // XXII conference «Mathematics. Computer. Education». Abstracts. 2015. P. 132.
11. Kazantsev M.V., Akinshin A.A., Golubyatnikov V.P. Comparison of some numeric methods for gene networks modeling using R // «Lomonosov readings on Altai». Abstracts. Barnaul, AltSU. 2014. P. 548–554.
12. Kazantsev M.V. Analysis of some methods for gene networks modeling using R // Abstracts of international conference «Mathematical and computer modeling». 2014. P. 8.
13. Gedeon T. Cyclic feedback systems // Memoirs of AMS. 1998. Vol. 134, no. 637. P. 72.
14. [14] Golubyatnikov V.P., Golubyatnikov I.V. On periodic trajectories in odd-dimensional gene network models // Russian Journal of Numerical Analysis and Mathematical Modelling. 2011. Vol. 26, no. 4. P. 397–412.
15. Akinshin A.A., Golubyatnikov V.P., Golubyatnikov I.V. On some multi-dimensional models of gene networks functioning // Sib. Journ. Ind. Math. 2013. T. XVI, № 1 (53). P. 3–9.
16. Akinshin A.A., Golubyatnikov V.P. Loops in symmetrical dynamical systems // NSU Herald: Math-mech-inf. 2012. T. 12, № 2. P. 3–12.

17. Kazantsev M.V., Golubyatnikov V.P. Discrete models of some non-linear dynamical systems // XII conference «Science and Youth 2015»: Information technologies. 2015. P. 9–10.
18. Kazantsev M.V. On some properties of domain graphs of dynamical systems // Sib. Journ. Ind. Math. 2015. T. XVIII, № 4 (64). P. 42–48.
19. Golubyatnikov V.P., Kazantsev M.V., Kalenykh A.E. On discretion of phase portraits of some dynamical systems // Mathematics and application: fundamental problems of science of technics. 2015. P. 28–31.
20. Kazantsev M.V. On some properties of discrete models of cyclic gene networks // Lomonosov readings on Altai: abstracts. 2015. P. 671.
21. Golubyatnikov V., Kazantsev M. On control of cycles in some gene networks models // International conference “Metric Structures and Control Systems”. Abstracts. 2015. P. 17–19.
22. Soetaert Karline, Petzoldt Thomas, Setzer R. Woodrow. Solving Differential Equations in R: Package deSolve // Journal of Statistical Software. 2010. Vol. 33, no. 9. P. 1–25. <http://www.jstatsoft.org/v33/i09>.
23. Hindmarsh C. Alan. ODEPACK, A Systematized Collection of ODE Solvers // Scientific Computing. 1983. P. 55–64.
24. Petzold R. Linda. Automatic Selection of Methods for Solving Stiff and Nonstiff Systems of Ordinary Differential Equations // SIAM Journal on Scientific and Statistical Computing. 1983. no. 4. P. 136–148.