

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ

Siberian Electronic Mathematical Reports

<http://semr.math.nsc.ru>

Том 6, стр. 440–456 (2009)

УДК 519.61; 577.21

MSC 37M05

КОМПЬЮТЕРНАЯ СИСТЕМА ИНТЕГРАЦИИ МОДУЛЕЙ
ДЛЯ АВТОМАТИЧЕСКОЙ ГЕНЕРАЦИИ И ЧИСЛЕННОГО
АНАЛИЗА МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ
МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКИХ СИСТЕМИ. Р. АКБЕРДИН, Ф. В. КАЗАНЦЕВ, В. А. ЛИХОШВАЙ, С. И. ФАДЕЕВ,
И. А. ГАЙНОВА, В. К. КОРОЛЕВ, А. Е. МЕДВЕДЕВ

ABSTRACT. At the turn of the 20th and 21st centuries there was such meaningful scientific progress in molecular biology as mapping of the human genome. This event is a start of the new period in development of the life science often called as post genome era. At the same time there were also qualitative changes in computer technology that multiply increase possibilities of computer's technology using for processing and storage of the huge information data. The intensive increased volume of experimental data generating by methods of the modern system biology stimulates a development of computer data bases for information processing and storage from literature and also stimulates a development of software for analysis and studying of complex models from these data base. The existing variety of computer tools solving the system biology tasks and modeling of the molecular genetic systems (MGS) in particular allows us to use the different approaches for solving of the assigned tasks. At the same time there is necessary to face the challenges under selection of the definite software. There are specificity of the model formats in this software and doubling of his functionality in other computer systems. In the judgment of specialists the decision of the compatibility problems of

AKBERDIN, I.R., KAZANTSEV, F.V., LIKHOSHVAI, V.A., FADEEV, S.I., GAINOVA, I.A., KOROLEV, V.K., MEDVEDEV, A.E., COMPUTER SYSTEM OF MODULES INTEGRATION FOR AUTOMATIC CONSTRUCTION AND NUMERICAL ANALYSIS OF MOLECULAR GENETIC SYSTEMS.

© 2009 АКБЕРДИН И.Р., КАЗАНЦЕВ Ф.В., ЛИХОШВАЙ В.А., ФАДЕЕВ С.И., ГАЙНОВА И.А., КОРОЛЕВ В.К., МЕДВЕДЕВ А.Е.

Работа частично поддержана грантом НШ-2447.2008.4, междисциплинарными интеграционными проектами СО РАН N 107, 119, программами РАН 21.26, 22.8, 23.29, грантом РФФИ N 08-01-91204-ЯФа.

Поступила 29 июня 2009 г., опубликована 16 ноября 2009 г.

different specialized systems and development of free and open formats on data presentation are key in the achievement of purpose to develop the integrative systems supporting technologies of complex model analysis.

We have represented the integrative computer system that oriented on construction and numerical analysis of models describing dynamics of the MGS functioning in pro- and eukaryotes. The system consists from the next program modules: module MGSgenerator and module STEP+. Module MGSgenerator is an intermediate unit in generation process of the mathematical models on basis of the gene networks reconstructed in GeneNet. Module STEP+ allows us to realize the numerical analysis of the model representing by autonomous system of ordinary differential equations. The mathematical model converts from our specialized format SBML in input STEP+ format by MGSgenerator module. We have tested the integrative computer system on basis of the MGS model describing intracellular auxin metabolism in a plant cell.

Keywords: integrative computer system, gene network, molecular genetic systems, intracellular auxin metabolism in a plant cell.

1. ВВЕДЕНИЕ

Конец 20-го - начало 21-го столетия ознаменован значительным научным достижением в области молекулярной биологии — расшифровкой генома человека (всей генетической информации, содержащейся в каждой клетке организма). Это событие считается началом нового периода в развитии наук о жизни, часто называемым постгеномной эрой. Одновременно с этим в области компьютерной техники также произошли качественные изменения, которые многократно повысили возможности использования компьютерных технологий для обработки и хранения огромных информационных массивов. Это привело к резкому ускорению темпов накопления информации во всех областях знания, в том числе и в области наук о жизни. В частности, в молекулярной биологии и генетике разработка скоростных методов секвенирования геномов привела к быстрому росту количества организмов с расшифрованными геномами. Развитие высокоэффективных экспериментальных ДНК-чиповых, ПЦР технологий позволила изучать динамику экспрессии (активности) тысяч генов одновременно, строить карты белок-белковых, ДНК-белковых и иных классов взаимодействий. Появилось новое поколение методов высокоразрешающей масс-спектрометрии, которые позволяют наблюдать за динамикой изменения РНК, белков, изучать потоки низкомолекулярных соединений и т.п. в режиме реального времени.

Разработка новых экспериментальных методик привела к стремительному росту объемов данных, которые нельзя характеризовать иначе как информационный «взрыв». Все возрастающий объем накапливаемой информации стимулировал разработку компьютерных технологий ее обработки, хранения и анализа. Одним из эффективных подходов к анализу биологических данных является метод математического моделирования.

Математические модели позволяют естественным образом объединять в рамках единой концептуальной схемы экспериментальные данные, касающиеся закономерностей строения и функционирования биологических систем на всех уровнях их организации, начиная с молекулярного, генетического, клеточного и заканчивая уровнями органов, тканей, организмов, популяций и экосистем.

При этом возникает возможность сопоставления и анализа разномасштабных экспериментальных данных, анализа причинно-следственных связей между молекулярной структурой, динамикой и фенотипическими характеристиками живых систем, разработки компьютерных систем для проведения численных экспериментов и т. д.

В настоящее время арсенал методов, который может быть использован для моделирования живых систем, достаточно широк и включает дискретные, непрерывные, стохастические и комбинированные подходы [1, 2, 3, 4, 6]. В силу новых вызовов, сформировавшихся в постгеномную эру, получили развитие компьютерные системы, ориентированные на математическое моделирование биологических систем. Укажем лишь некоторые из них, в основном используемые в научном сообществе: E-Cell (www.e-cell.org/software/e-cell-system), Dbsolve (<http://biosim.genebee.msu.su/en/>), Virtual Cell (<http://www.vcell.org/>), Cellerator (www.aig.jpl.nasa.gov/public/mls/cellerator), Pathway lab (<http://innetics.com/>), Netbuilder (<http://strc.herts.ac.uk/bio/maria/NetBuilder/>). Авторы хорошо известного пакета Mathematica (<http://www.wolfram.com>) — универсального средства для решения математических задач, также не оставили без внимания биологические проблемы и приступили к разработке предметно-ориентированных ресурсов Cellerator и MathSBML. Для совмещения различных средств моделирования биологических процессов разработчиками компьютерных систем предложен открытый компьютерный формат SBML (<http://sbml.org/index.psp>). Формат SBML является стандартом представления моделей биохимических реакций, в том числе — связанных с метаболизмом, передачей сигнала и управлением в биологических системах. Осуществляет построение моделей реакций в виде блок-схем с указанием внутренних связей.

К настоящему времени общее количество разрабатываемых компьютерных систем моделирования уже превысило сотню. Каждая система имеет собственный набор специализированных функциональностей, предназначенных для решения задач моделирования молекулярно-генетических систем (МГС). В то же время можно отметить, что даже в совокупности разработанные системы не могут предложить средств конструирования моделей с учетом всех специфических особенностей строения МГС, а именно, учитывая взаимное расположение и ориентацию генов в составе хромосом и геномов, полиаллельность (аллели — различные формы одного гена, отличающиеся нуклеотидным составом) и полиморфизм генов; поликомпарментность и иерархическую организацию генных сетей, полифункциональность молекул, динамическую перестройку организации биологических систем в процессе их роста и развития. Таким образом, проблема разработки универсальных средств моделирования биологических систем до конца не решена. В тоже время наличие большого количества специализированных компьютерных систем и широкий спектр функциональностей позволяет развивать на их основе более универсальные системы моделирования.

Ранее нами были разработаны программные модули MGSgenerator и STEP+. Программный модуль MGSgenerator [7] предназначен для автоматизированного конструирования математических моделей на основе

структурно-функциональной организации генных сетей (ГС), аннотированных в системе GeneNet [8]. Программный модуль STEP+ [9] позволяет проводить численное исследование математических моделей, представленных автономной системой обыкновенных дифференциальных уравнений.

В работе мы представляем компьютерную систему, разработанную на основе модулей MGSgenerator и STEP+. Представляемая система позволяет автоматизировать генерацию модели по структуре генной сети и проводить ее численный анализ. Апробация взаимодействия указанных программных модулей была проведена на модели МГС внутриклеточного метаболизма ауксина в побеге растения [10, 11].

2. ОПИСАНИЕ ПРОГРАММНОГО МОДУЛЯ MGSGENERATOR

Программный модуль MGSgenerator является ключевым звеном в процессе генерации математических моделей. Его цель — построение математической модели по имеющейся структурно-функциональной организации генной сети и ряду дополнительных комментариев. Реконструкция генных сетей осуществляется при помощи системы GeneNet, входящей в интегрированную компьютерную систему. MGSgenerator имеет собственный стандарт организации данных и программный интерфейс (API), используя который, можно разрабатывать модули генерации математических моделей для различных источников, описывающих ГС. Принцип работы модуля MGSgenerator состоит в импорте структурной модели ГС из системы GeneNet, построении математической модели и сохранении построенной модели МГС в базе данных.

Построение математической модели основано на обобщенном химико-кинетическом подходе [12, 13], который позволяет описывать молекулярно-генетические процессы, связывающие компоненты генной сети, системой дифференциальных уравнений. Модуль MGSgenerator ставит в соответствие каждому биологическому процессу некоторую шаблонную функцию и, исходя из этой информации, генерирует математическую модель.

Суть обобщенного химико-кинетического подхода состоит в следующем. Для описания биологических процессов, таких, например, как трансляция, транскрипция и т.д., применяются шаблонные модели трех различных уровней сложности. Ниже приведены шаблонные модели для всех трех уровней. Так, процессы репликации, транскрипции и трансляции без учёта регуляторных воздействий описываются в виде простого уравнения:

$$(1) \quad V = xk_0$$

где V — скорость данного процесса, x — количество (концентрация) гена, мРНК в клетке и т.д., k_0 — константа скорости реакции.

При наличии молекулярно-генетической регуляции данных процессов используется другая шаблонная модель описания:

$$(2) \quad V = xk_0 \left[\frac{\delta + \sum_i \left(\frac{a_i}{k_{1,ai}} \right)^{h_i}}{1 + \sum_i \left(\frac{a_i}{k_{2,ai}} \right)^{hd_i} + \sum_i \left(\frac{r_i}{k_{ri}} \right)^{g_i}} \right],$$

где x – выделенный элемент (ген/РНК/фермент), a_i – концентрация активатора, r_i – концентрация ингибитора, k_0 – константа скорости реакции, $k_{1,ai}$, $k_{2,ai}$ – константы активатора, k_{ri} – константа ингибитора, δ – базальная активность (т.е. константа скорости данного процесса без учета каких-либо механизмов регуляции – активации, ингибирования), h_i , g_i , hd_i – коэффициенты Хилла. По умолчанию полагается конкурентное взаимодействие регуляторов между собой и неконкурентное взаимодействие с выделенным элементом (x).



Для описания процессов ферментативного синтеза с учётом молекулярно-генетической регуляции и обратимости реакции используется еще более сложный способ математического описания:

$$(3) \quad V = x \frac{k_{s,0} \prod_i \frac{S_i}{k_{mS,i}} - k_{p,0} \prod_i \frac{P_i}{k_{mP,i}}}{\prod_i \left(1 + \frac{S_i}{k_{mS,i}} + \frac{P_i}{k_{mP,i}} \right)} \times \frac{\delta + \sum_i \left(\frac{a_i}{k_{1,ai}} \right)^{h_i}}{1 + \sum_i \left(\frac{a_i}{k_{2,ai}} \right)^{hd_i} + \sum_i \left(\frac{r_i}{k_{ri}} \right)^{g_i}},$$

где x – фермент (регулятор), a_i , r_i – соответственно: активатор, ингибитор, S - субстрат (вход), P - продукт (выход), k_0 – константа оборота, $k_{S,0}$ – константа прямой реакции, $k_{P,0}$ – константа обратной реакции, k_{mS} – константа Михаэлиса для субстрата, k_{mP} – константа Михаэлиса для продукта, k_{ai} , k_{ri} – константы активатора и ингибитора, соответственно, δ – базальная активность, h_i , g_i , hd_i – коэффициенты Хилла.

Из таких шаблонных блоков формируются функции, задающие правые части системы обыкновенных дифференциальных уравнений, т. е. математическую модель. На выходе модуль MGSgenerator генерирует три файла во входном формате программного модуля STEP+:

- (1) файл с количеством уравнений модели;
- (2) файл с количеством параметров модели;
- (3) файл, содержащий запись уравнений исследуемой системы.

На рисунке 1 представлена элементарная ферментативная реакция превращения одного низкомолекулярного вещества в другое (реакция на рисунке обозначена зеленой стрелкой, вещества – ). Данная реакция осуществляется под действием фермента – белка  (реакция показана на рисунке красной стрелкой). Связывание третьего низкомолекулярного вещества с белком (синяя стрелка) приводит к изменению конформационной (структурной) организации фермента. В результате белок функционирует хуже (ингибируется) и скорость ферментативной реакции уменьшается.

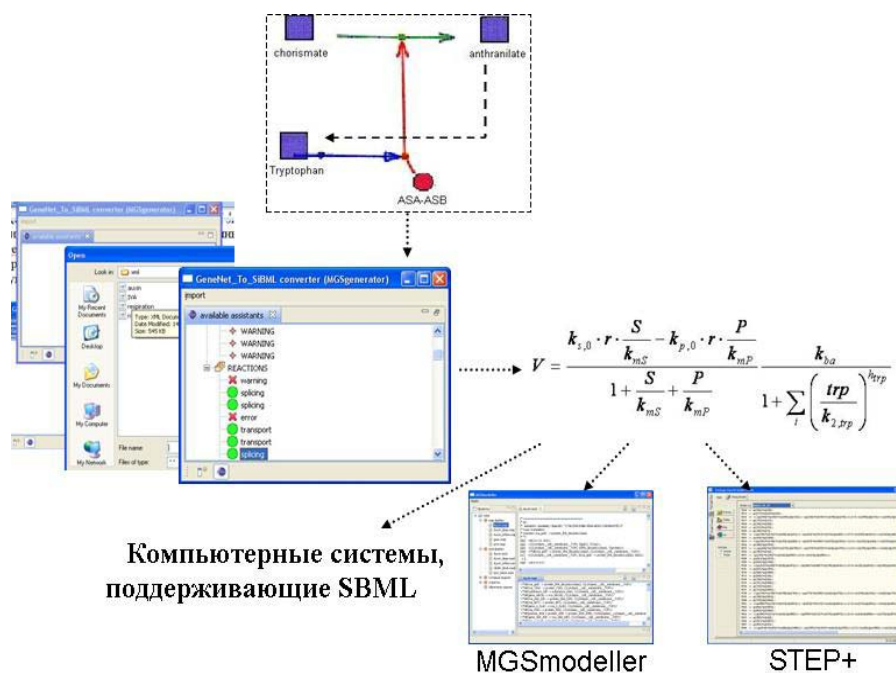


Рис. 1. Блок-схема автоматической генерации и экспорта математической модели в различные компьютерные системы в соответствующих входных форматах.

Прерывистая линия указывает на то, что через ряд промежуточных стадий второе вещество превращается в третье.

Структурно-функциональная организация ферментативной реакции превращения веществ, представленная в формате XML (формат программного модуля MGSmodeller [14], разработанного авторами), является входными данными для программного модуля MGSgenerator. Для данной реакции модуль MGSgenerator сопоставляет соответствующую шаблонную модель в терминах обобщенного химико-кинетического подхода и, в итоге, генерирует математическую модель реакции. При этом MGSgenerator предоставляет возможность экспорта математической модели в формате различных компьютерных систем: 1) поддерживающих стандартный формат SBML; 2-3) форматы программных модулей MGSmodeller и STEP+, разработанные авторами статьи.

Программный модуль MGSgenerator реализован по технологии "Plug-in" и имеет модульную архитектуру. Такая структура генератора позволяет подключать новые источники данных о структурно-функциональной организации МГС, экспортировать модели из других баз данных, а также дает возможность минимальными усилиями изменять и дополнять имеющиеся программные модули, интегрировать данный модуль в другие компьютерные системы.

3. ОПИСАНИЕ ПРОГРАММНОГО МОДУЛЯ STEP+

Программный модуль STEP+ дает возможность проводить комплексное изучение математической модели, которая генерируется модулем MGSgenerator в виде автономной системы из n дифференциальных уравнений с вектором параметров p :

$$(4) \quad \frac{dx}{dt} = f(x, p), \quad t > 0,$$

$$x = x^0, \quad t = 0.$$

Здесь $f(x, p)$ — достаточно гладкая вектор-функция векторных аргументов $x \in R^n$ и $p \in R^m$ в области ее определения. Если установить соответствие между шаблонной функцией, описанной, например, уравнением (3), и правой частью уравнения (4), вектор-функцией $f(x, p)$, то вектор переменных x будет составлен из концентраций веществ, а вектор параметров p — из констант биохимических реакций и коэффициентов Хилла шаблонной функции (3):

$$x = (x, S_i, P_i, a_i, r_i); \quad p = (k_0, k_{S,0}, k_{P,0}, k_{mS}, k_{mP}, k_{ai}, k_{ri}, \delta, h_i, g_i, hd_i).$$

Модуль STEP+ содержит алгоритмы численного исследования решения автономной системы в зависимости от параметров:

- методы интегрирования — многошаговый метод Гира переменного порядка точности для жестких систем, метод Адамса для интегрирования нежестких систем [15];
- метод продолжения решения по параметру для построения диаграммы стационарных решений автономной системы в зависимости от скалярного параметра $\alpha \in p$ [16, 17];
- числовой критерий Годунова-Булгакова определения гарантированной асимптотической устойчивости стационарных решений в зависимости от параметра α [18].

Метод продолжения решения по параметру применяется также для изучения зависимости решения системы нелинейных уравнений (не связанной с автономной системой)

$$(5) \quad f(x, \alpha) = 0,$$

от скалярного параметра α , $\alpha \in p \in R^m$. Графики одной или нескольких компонент вектор-функции $x = x(\alpha)$, являющейся решением (5), будем называть диаграммой стационарных решений. Метод продолжения по параметру опирается на известную в анализе теорему о неявных функциях. Согласно этой теореме графиком решения (5) в $(n + 1)$ -мерном пространстве будет гладкая пространственная кривая, если в окрестности пространственной кривой ранг матрицы A производных правых частей: $A = [f_x, f_\alpha]$ независимо от α всегда равен n . При этом гладкая пространственная кривая может несколько раз пересекать гиперплоскость $\alpha = \alpha^*$, $\alpha^* \in [\alpha_0, \alpha_1]$, что означает множество решений (5) при $\alpha = \alpha^*$.

Идея метода продолжения по параметру для построения гладкой пространственной кривой состоит в следующем. Пусть при некотором значении α известно решение системы (5), и A — матрица $[f_x, f_\alpha]$, построенная на этом решении. Поскольку ранг A равен n , то существует невырожденная

квадратная матрица $B = [f_z]$, получаемая из A вычеркиванием некоторого k -го столбца. Обозначим через z вектор, компоненты которого совпадают с компонентами вектора (x, α) после вычеркивания компоненты с индексом k , которую обозначим через μ . Определив вектор производных z_μ как решение системы с матрицей B и правой частью f_μ , найдем максимальную по модулю компоненту вектора $(z_\mu, 1)$ с индексом j . Вектор, состоящий из компонент вектора (x, α) после вычеркивания компоненты с номером j , обозначим через u , а вычеркиваемую компоненту – через λ . Вектор u можно рассматривать как решение системы (5) при том же значении α , но с параметром λ . При этом вектор производных решения по параметру λ находится нормировкой компонент вектора $(z_\mu, 1)$ с использованием для нормировки максимальной по модулю компоненты вектора $(z_\mu, 1)$. В результате максимальная по модулю компонента вектора производных решения по параметру λ будет равна 1. Это означает, что λ можно рассматривать как параметр системы (5) в окрестности рассматриваемого решения.

Процедуру определения λ будем называть параметризацией, а сам параметр λ – текущим параметром системы (5), поскольку он определяется регулярно на один шаг $\Delta\lambda$ в методе продолжения решения. В роли текущего параметра могут равноправно оказаться компоненты вектора (x, α) . Продолжение решения по параметру в сочетании с параметризацией и адаптацией текущего шага позволяет построить гладкую пространственную кривую, определяемую системой (5), которая может содержать точки поворота, и, следовательно, в некоторых областях изменения параметра α иметь несколько решений.

В программный модуль STEP+ включён числовой критерий гарантированной асимптотической устойчивости стационарного решения, разработанного под руководством академика С.К. Годунова. Метод, не требующий вычисления собственных чисел матрицы Якоби $J = [f_x]$ системы (4), основан на оценке нормы решения H матричного уравнения Ляпунова:

$$J^*H + HJ = -I,$$

где I – единичная матрица, J^* – матрица, сопряженная с J . Матрица J – гурвицева (т. е. исследуемое стационарное решение асимптотически устойчиво), если H – эрмитова положительно определенная матрица. В этом случае H имеет интегральное представление, в котором используются матричные экспоненты матриц J и J^* :

$$H = \int_0^\infty \exp(tJ^*) \exp(tJ) dt .$$

Числовая характеристика $\kappa(J)$ асимптотической устойчивости, если J – гурвицева, имеет вид:

$$\kappa(J) = 2\|J\| \sup_{v(0) \neq 0} \frac{\int_0^\infty \|v(t)\|^2 dt}{\|v(0)\|^2} = 2\|J\|\|H\| ,$$

где $v(t) = [\exp(tJ)]v(0)$.

В итерационном процессе вычисления $\kappa(J)$ на каждой итерации предусмотрена проверка выполнения неравенства: $\bar{\kappa}(J) < k_0$, где $\bar{\kappa}(J)$ – приближение $\kappa(J)$ на итерации, а k_0 – константа, значение которой зависит только от точности представления числа в ЭВМ. Асимптотическая

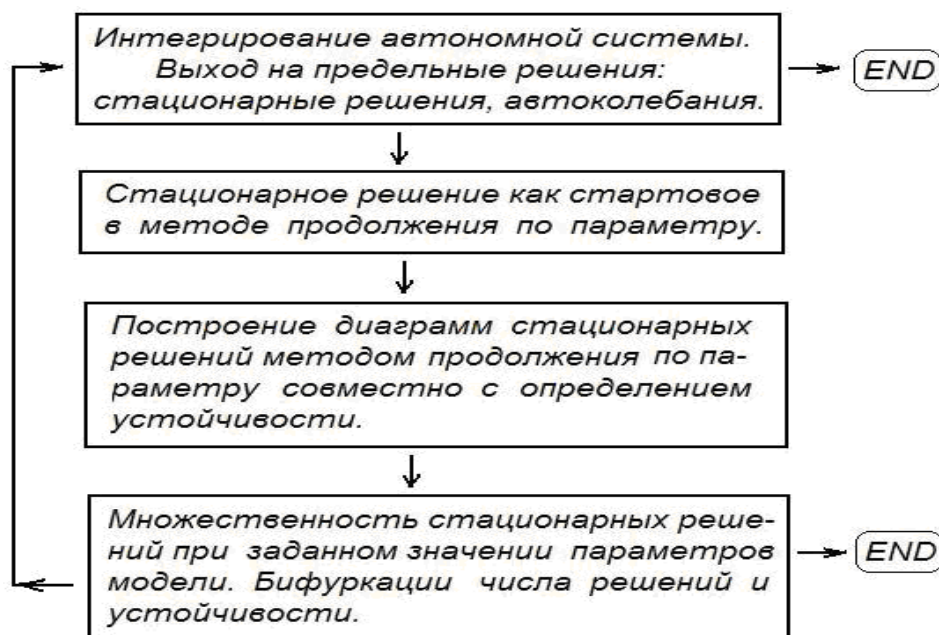


Рис. 2. Блок-схема численного исследования математической модели.

устойчивость гарантируется, если и предельное значение $\kappa(J)$ удовлетворяет этому неравенству. Если же на очередной итерации неравенство нарушается, то сообщается гарантированное заключение о «практической» неустойчивости исследуемого стационарного решения.

Пользовательский интерфейс модуля STEP+ формирует матрицу Якоби $J = [f_x]$ и матрицу частных производных от правых частей системы по параметрам $\Phi = [f_p]$. Аналитическое представление элементов матриц дает возможность пользователю проводить численное исследование зависимости решения от любого параметра модели по указанной на рисунке 2 схеме. Также в автоматическом режиме осуществляются трансляция, генерация и запуск расчетных программ.

В случае выхода решения задачи Коши на автоколебания, имеется возможность графического представления траекторий в физической и фазовой плоскостях, а также в 3D-мерном фазовом пространстве, с возможностью вращения трехмерных объектов. Анализ диаграмм стационарных решений с помощью таблиц позволяет выявить бифуркационные значения параметра системы (точки поворота и точки бифуркации Андронова-Хопфа), области изменения параметра, характеризующиеся самовозбуждением автоколебаний.

4. АПРОБАЦИЯ КОМПЬЮТЕРНОЙ СИСТЕМЫ НА МАТЕМАТИЧЕСКОЙ МОДЕЛИ ГС ВНУТРИКЛЕТОЧНОЙ РЕГУЛЯЦИИ МЕТАБОЛИЗМА АУКСИНА

Тестирование функциональных возможностей каждого программного модуля системы проводилось для математической модели генной сети метаболизма ауксина (основной гормон растений, регулирующий их рост

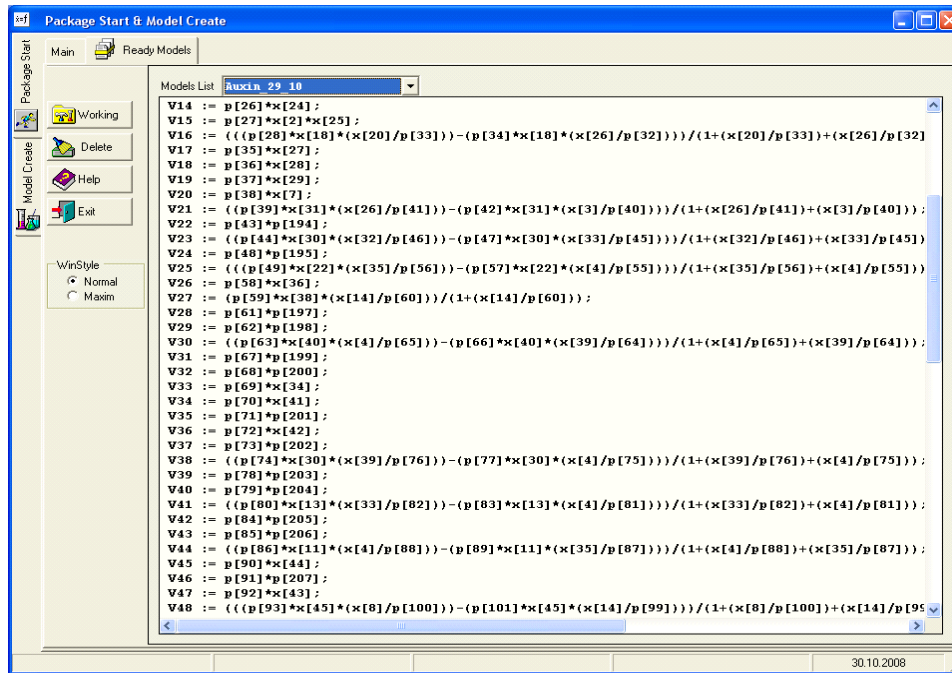


Рис. 3. Фрагмент записи шаблонных функций модели метаболизма ауксина с учётом генетической регуляции в формате STEP+.

и развитие [19]) в клетке растения. Реконструкция строения генной сети внутриклеточной регуляции метаболизма ауксина в клетке растения осуществлена на основе литературных данных (107 научных публикаций) и представлена в компьютерной системе GeneNet. Так как процесс реконструкции генных сетей по сути является структурным моделированием молекулярно-генетических систем, то генные сети, представленные в компьютерной системе GeneNet, в дальнейшем будем также называть структурными моделями генных сетей.

Структурная модель генной сети метаболизма ауксина (ГСМА) с учётом генетической регуляции процессов (активации и ингибирования) содержит 62 гена, 56 мРНК и 44 белка, включает в себя 245 молекулярно-генетических процессов, часть из которых аналогична процессу, описанному на рисунке 1.

Затем был проведен сравнительный анализ структурно-функциональной организации метаболизма ауксина у высших и низших растений [20]. На основе ГСМА построены упрощенные (редуцированные) структурные модели генной сети метаболизма ауксина внутри клетки высших (покрытосеменные, голосеменные) (ГСМАв) и низших (папоротник, мхи) видов растений (ГСМАН). Структурные модели редуцированных генных сетей были конвертированы в две математические модели с помощью программного модуля MGSgenerator. На рисунке 3 представлен фрагмент записи в формате пакета STEP+ математической модели ГСМАв. Эта модель содержит 45 переменных и 207 параметров, описывает 48 элементарных подсистем

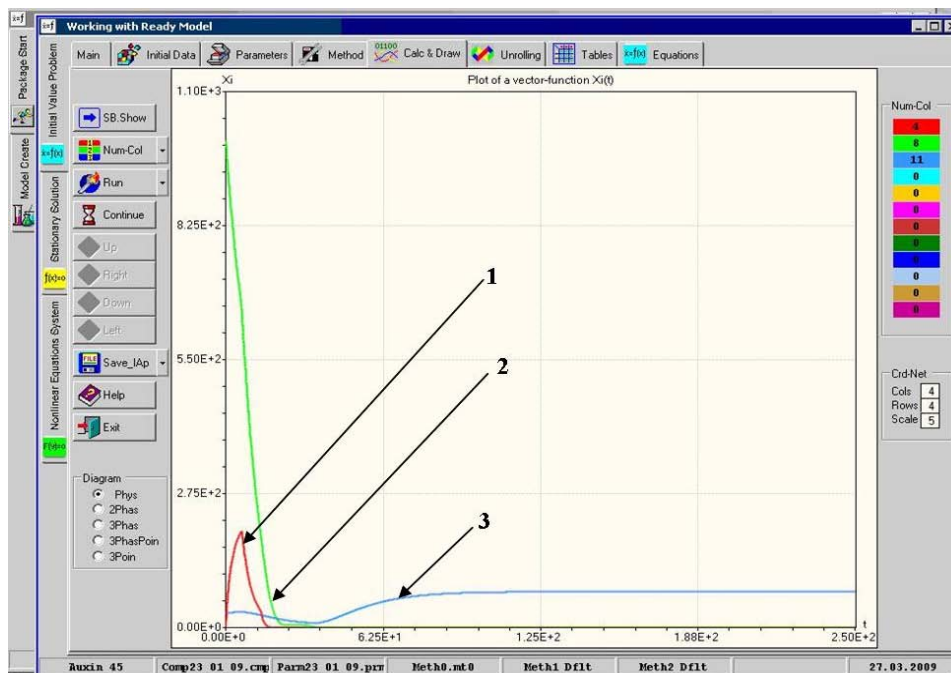


Рис. 4. Динамика изменения концентрации в модели ГСМАв: 1) ауксина; 2) триптофана; 3) белка GH3. По оси y – концентрация (mM), по оси x – время (сек).

метаболизма ауксина с учётом генетической регуляции. Математическая модель ГСМАн содержит 18 переменных и 81 параметр, описывает 17 элементарных подсистем.

Начальные значения большинства переменных и параметров моделей взяты из литературных источников [21, 22, 23, 24, 25, 26]. Остальные значения были оценены на основе косвенных экспериментальных данных и биологических представлениях о природе и механизмах регуляции молекулярно-генетической системы метаболизма ауксина в клетке растения.

5. ЧИСЛЕННЫЙ АНАЛИЗ В ПРОГРАММНОМ МОДУЛЕ STEP+

Для расчётов динамики изменения концентраций ауксина, триптофана (низкомолекулярное вещество, из которого синтезируется ауксин) и одного из важных ферментов системы (белка GH3) решалась задача Коши на интервале по $t \in [0, 250]$ (сек). Как видно из рисунков 4 и 5, с течением времени решение задачи Коши выходит на устойчивое стационарное решение. При численном исследовании моделей, включая исследование стационарного решения методом продолжения по параметру в положительной области значений переменных, другие аттракторы не выявляются. Устойчивость стационарных решений подтверждалась методом Годунова-Булгакова. Численные результаты согласуются с биологической природой моделируемых генных сетей, назначение которых состоит в поддержании гомеостаза целевого продукта (ауксина).

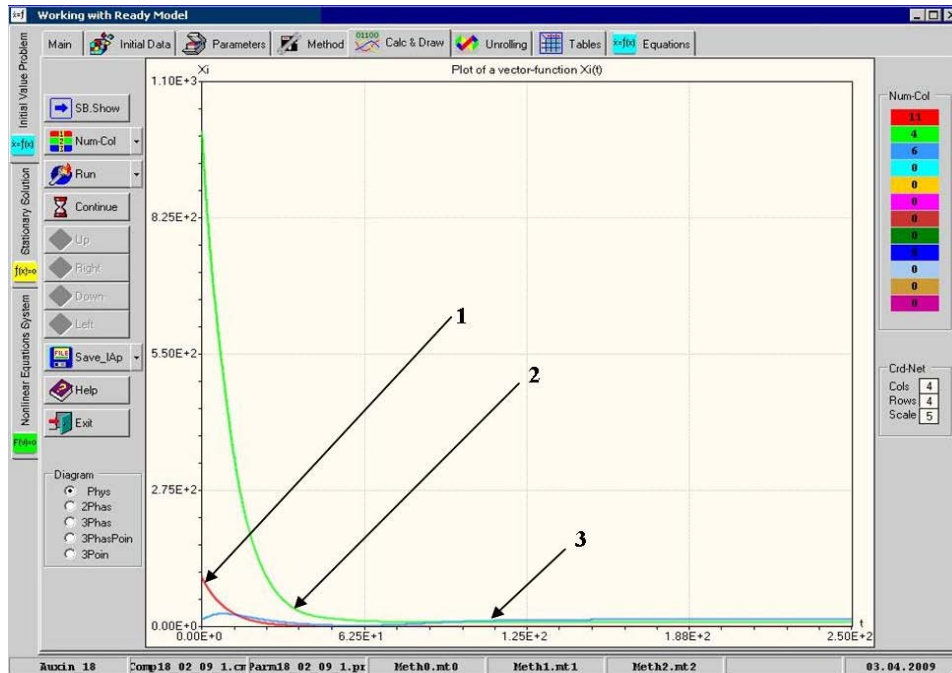


Рис. 5. Динамика изменения концентрации в модели ГСМАН: 1) ауксина; 2) триптофана; 3) белка ГНЗ. По оси y – концентрация (mM), по оси x – время (сек).

Молекулярно-генетические системы являются динамическими нелинейными системами, функционирование которых существенно зависит от природы молекулярно-генетических механизмов и значения констант, определяющих скорости синтеза и деградации веществ. В моделях генных сетей механизмы трансформируются в конкретные формулы, а константы трансформируются в параметры, которые входят в формулы. Одна из задач, которая стандартно решается при моделировании МГС, заключается в определении влияния различных типов мутаций, которые могут происходить в целевой генетической системе в процессе ее функционирования. В живых системах изменения чаще всего происходят на генетическом уровне. Изменения могут быть разных типов: точечные мутации (замена одного нуклеотида на другой), делеции (удаление некоторой нуклеотидной последовательности), вставки (добавление в состав генома некоторой нуклеотидной последовательности) и т. д. На уровне математического описания имитация генетических мутаций проводится либо изменением структуры модели (это требуется в том случае, если мутация меняет некоторые механизмы в генной сети), либо изменением значений некоторых параметров (если мутация не меняет механизмы генной сети, но изменяет константы, входящие в них). Необходимость в рассмотрении того или иного структурно-модифицированного варианта модели, по сути, неотъемлемо диктуется спецификой моделируемой системы. Напротив, исследование параметрически модифицированных вариантов

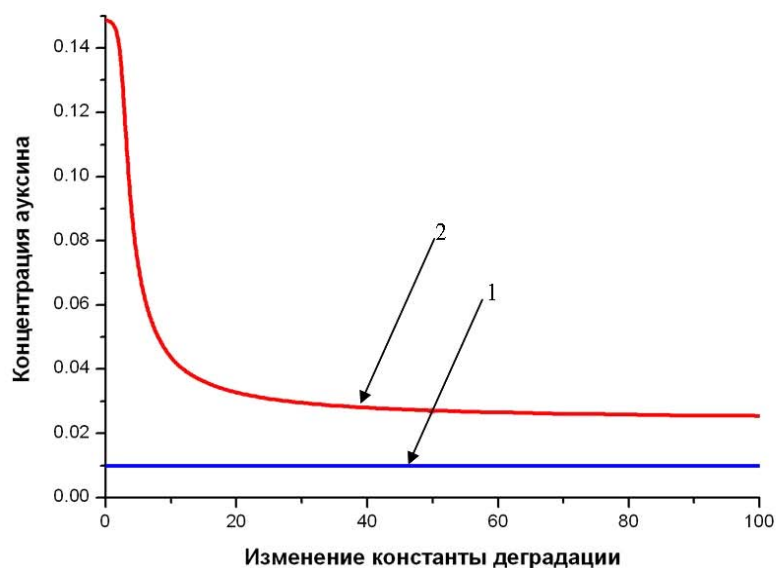


Рис. 6. Изменение концентрации ауксина в зависимости от константы деградации Рнк гена *gh3*: 1) кривая для модели ГСМАв; 2) кривая для модели ГСМАН. По оси y — концентрация ауксина (мМ), по оси x — константа деградации (1/сек).

модели представляет стандартную вычислительную проблему исследования чувствительности модели от параметров.

В настоящей работе мы исследовали параметрическую чувствительность стационарных состояний, в которые приходят обе модели на бесконечном интервале времени. Для анализа мы использовали метод продолжения по параметру, реализованный в модуле STER+ для решения данного класса задач. Варьирование значений параметров проводили в заданном диапазоне: в интервале $[10^{-2}\alpha_0, 10^{+2}\alpha_0]$, где α_0 — исходное значение варьируемого параметра. Графики изменения концентрации ауксина в моделях для ГСМАв и ГСМАН в зависимости от варьируемого значения параметра представлены на рис. 6, 7 (рис. 6: в зависимости от константы деградации Рнк гена *gh3* ($k_degrad_Рнк_gh3$), рис. 7: в зависимости от коэффициента Хилла реакции превращения хоризмата в антранилат, которая описана на рисунке 1, ($kg1_хоризмат \Leftrightarrow антранилат$)). Из рисунков видно, что варьирование одного и того же параметра в двух моделях приводит к качественно различным кривым изменения его стационарной концентрации от параметра. Представленные кривые демонстрируют влияние структурно-функциональной организации на характеристики поведения метаболической системы ауксина в клетке растения и нелинейный характер влияния параметров системы на концентрацию ауксина в модели ГСМАН.

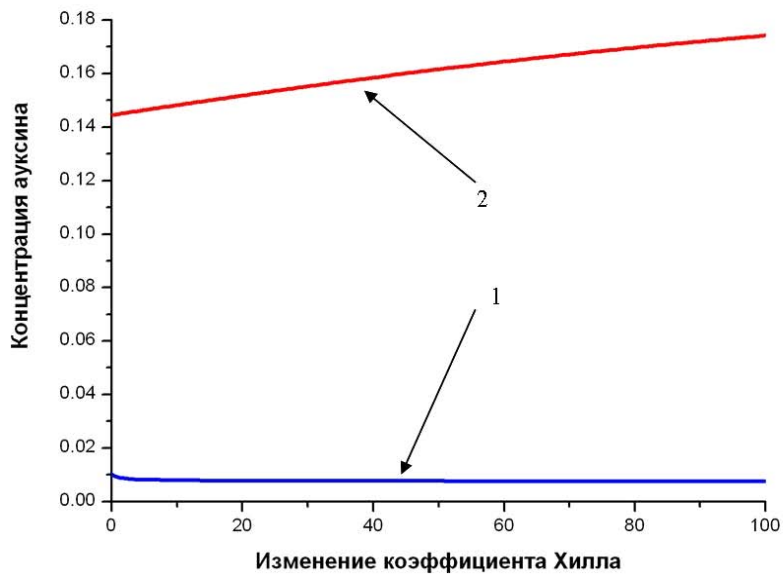


Рис. 7. Изменение концентрации ауксина в зависимости от коэффициента Хилла реакции превращения хоризмата в антранилат: 1) кривая для модели ГСМАв; 2) кривая для модели ГСМАН. По оси y – концентрация ауксина (мМ), по оси x – коэффициент Хилла.

6. ЗАКЛЮЧЕНИЕ

В публикации представлена новая компьютерная система генерации и анализа математических моделей молекулярно-генетических систем на основе двух программных модулей — MGSgenerator и STEP+, которые входят в интегрированную компьютерную систему, разрабатываемую в ИЦиГ СО РАН. Актуальность разработки данной системы диктуется необходимостью решения ряда компьютерных проблем, стоящих на современном этапе системной биологии: совместимость различных специализированных компьютерных программ и систем, разработка открытых форматов представления данных, построение и численное исследование математических моделей МГС, создание базы данных математических моделей указанного типа. Возможности компьютерной системы демонстрируются на примере моделирования генной сети метаболизма ауксина в клетке растений.

Реконструкция ГСМА была проведена на основе более ста литературных источников в компьютерной системе GeneNet. На основе структурно-функциональной организации и модуля MGSgenerator были автоматически сгенерированы две математические модели метаболизма ауксина в формате программного средства STEP+ для высших и низших видов растений, соответственно. Численный анализ математических моделей показал, что исследуемая система имеет одно устойчивое стационарное решение.

Полученный результат согласуется с биологическими представлениями о природе моделируемых генных сетей: поддержание гомеостаза целевого продукта (ауксина). Сравнительный анализ зависимости стационарного решения от изменения значения параметров был проведен для двух математических моделей метаболизма ауксина, описывающих исследуемую молекулярно-генетическую систему у высших и низших видов растений.

Автоматическая генерация математических моделей метаболизма ауксина в клетке растения с помощью программного модуля MGSgenerator и анализ моделей в программном модуле STEP+ показывают не только состоятельность разрабатываемой комплексной компьютерной системы, но и позволяют получить интересные биологические результаты о динамике функционирования и эволюционных этапах преобразования молекулярно-генетических систем и метаболизма ауксина, в частности. Разработанная компьютерная система может быть полезна при разработке и анализе математических моделей широкого круга МГС. Функциональные возможности каждого из программных модулей разработанной системы позволяют, в отличие от других аналогичных систем, представленных во введении:

а) моделировать молекулярно-генетические системы с учетом всех специфических особенностей их строения на основе обобщенного химико-кинетического метода и метода обобщенных функций Хилла;

б) проводить полномасштабный анализ чувствительности математической модели на основе подходов, реализованных в программном модуле STEP+.

Отметим, что в настоящее время не существует аналогичных компьютерных систем, которые бы реализовали конвейерную технологию автоматической генерации математической модели на основе структурно-функциональной организации МГС с последующим анализом результатов расчета и чувствительности модели к варьированию параметров. Несомненно, разработанная система не является универсальным средством, позволяющим решать все проблемы в области моделирования живых систем, именно поэтому мы планируем проводить дальнейшее развитие данной системы в направлении повышения степени автоматизации процесса генерации математических моделей, а также развивать методы анализа моделей.

СПИСОК ЛИТЕРАТУРЫ

- [1] S.A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford Univ. Press, N.Y., 1993.
- [2] R. Thomas, D. Thieffry, M. Kaufman, *Dynamical behaviour of biological regulatory networks-I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state*, Bull. Math. Biol., **57**: 2 (1995), 247–276.
- [3] P. Smolen, D.A. Baxter, J.H. Byrne, *Modeling transcriptional control in gene networks - methods, recent results, and future directions*, Bull. Math. Biol., **62**: 2 (2000), 247–292.
- [4] В.А. Лихошвай, С.И. Фадеев, Г.В. Демиденко, Ю.Г. Матушкин, *Моделирование многостадийного синтеза вещества без ветвления уравнением с запаздывающим аргументом*, Сибирский журнал индустриальной математики, **7**: 1 (2004), 73–94.
- [5] V.A. Likhoshvai, S.I. Fadeev, Yu.G. Matushkin, *The Global Operation Modes Of Gene Networks Determined By The Structure Of Negative Feedbacks*, "Bioinformatics of genome regulation and structure". Ed. by N. Kolchanov and R. Hofestaedt, Kluwer Academic Publishers, Boston/Dordrecht/London, (2004), 319–330.
- [6] T.E. Turner, S. Schnell, K. Burrage, *Stochastic approaches for modelling in vivo reactions*, Comput. Biol. Chem., **28**: 3 (2004), 165–178.

- [7] Ф.В. Казанцев, И.Р. Акбердин, К.Д. Безматерных, В.А. Лихошвай, *Система автоматизированной генерации математических моделей генных сетей*, Информационный Вестник ВОГиС, **13**: 1 (2009), 163–170.
- [8] E.A. Ananko, N.L. Podkolodny, I.L. Stepanenko, O.A. Podkolodnaya, D.A. Rasskazov, D.S. Miginsky, V.A. Likhoshvai, A.V. Ratushny, N.N. Podkolodnaya and N.A. Kolchanov, *GeneNet in 2005*, Nucleic Acids Res., **33** (2005), 425–427.
- [9] S.I. Fadeev, V.K. Korolev, I.A. Gainova, A.E. Medvedev, *The package Step+ for numerical study of autonomous systems arising when modeling dynamics of genetic-molecular systems*, Proc. of the 6th Intern. Conf. on Bioinformatics of Genome Regulation and Structure, **2** (2006), 118–120.
- [10] I.R. Akberdin, N.A. Omelyanchuk, S.I. Fadeev, V.M. Efimov, I.A. Gainova, V.A. Likhoshvai, *Mathematical model of auxin metabolism in shoots of arabidopsis thaliana L*, Proc. of the 6th Intern. Conf. on Bioinformatics of Genome Regulation and Structure, **2** (2008), 23.
- [11] И.Р. Акбердин, Ф.В. Казанцев, Н.А. Омелянчук, В.А. Лихошвай, *Математическое моделирование метаболизма ауксина в клетке меристемы побега растения*, Информационный Вестник ВОГиС, bf13: 1 (2009), 170–176.
- [12] В.А. Лихошвай, Ю.Г. Матушкин, А.В. Ратушный, Е.А. Ананько, Е.В. Игнатьева, О.В. Подколдная, *Обобщенный химико-кинетический метод моделирования генных сетей*, Молекулярная биология, **35**: 6 (2001), 1072–1079.
- [13] V. Likhoshvai and A. Ratushny, *Generalized Hill function method for modeling molecular processes*, Journal of Bioinformatics and Comp. Biology, **5**: 2 (2007) 521–531.
- [14] V.A. Likhoshvai, F.V. Kazantsev, I.R. Akberdin, et al. *A computer system for reconstruction, calculation and analysis of mathematical models of molecular genetic system (MGSmodeller)*, The certificate N 2008612820, (2008).
- [15] G.W.Gear, *The automatic integration of ordinary differential equations*, Comm. ACM.V., **14**: 3, (1971), 176–179.
- [16] С.И.Фадеев, С.А.Покровская, А.Ю.Березин, И.А.Гайнова, *Пакет программ STEP для численного исследования систем нелинейных уравнений и автономных систем общего вида. Описание работы пакета STEP на примерах задач из учебного курса «Инженерная химия каталитических процессов»*, Изд-во НГУ, Новосибирск, 1998.
- [17] S.I. Fadeev, V.V. Kogai, *Using parameter continuation based on the multiple shooting method for numerical research of nonlinear boundary value problems*, International Journal of Pure and Applied Mathematics, **14**: 4 (2004), 467–498.
- [18] С.К. Годунов, *Обыкновенные дифференциальные уравнения с постоянными коэффициентами*, Изд-во НГУ, Новосибирск, 1994.
- [19] A.W. Woodward, B. Bartel, *Auxin: regulation, action, and interaction*, Ann Bot (Lond), **95**: 5 (2005), 707–735.
- [20] E. Sztein, J.D. Cohen and T.J. Cooke, *Evolutionary patterns in the auxin metabolism of green plants*, Int. J. Plant Sci., **161**: 6 (2000), 849–859.
- [21] R.S. Bandurski, A. Schulze, *Concentration of Indole-3-acetic Acid and Its Derivatives in Plants*, Plant Physiol., **60**: 2 (1977), 211–213.
- [22] Y.S. Momonoki, A. Schulze, R.S. Bandurski, *Effect of Deseeding on the Indole-3-acetic Acid Content of Shoots and Roots of Zea mays Seedlings*, Plant Physiol., **72**: 2 (1983), 526–529.
- [23] R.T. Davies, D.H. Goetz, J. Lasswell, M.N. Anderson, B. Bartel, *IAR3 encodes an auxin conjugate hydrolase from Arabidopsis*, Plant Cell, **11**: 3 (1999), 365–376.
- [24] J. Normanly, B. Bartel, *Redundancy as a way of life - IAA metabolism*, Curr Opin Plant Biol., **2**: 3 (1999), 207–213.
- [25] M.D. Mikkelsen, C.H. Hansen, U. Wittstock, B.A. Halkier, *Cytochrome P450 CYP79B2 from Arabidopsis catalyzes the conversion of tryptophan to indole-3-acetaldoxime, a precursor of indole glucosinolates and indole-3-acetic acid*, J. Biol. Chem., **275**: 43 (2000), 33712–33717.
- [26] S. Park, A. Walz, Y.S. Momonoki, J. Ludwig-Muller, J.P. Slovin, J.D. Cohen, *Partial characterization of major IAA conjugates in Arabidopsis*, Proc. of the American Society of Plant Biologists Conference, (2003), Honolulu, HI. <http://abstracts.aspb.org/pb2003/public/P46/0961.html>

Илья Ринатович Акбердин
Институт цитологии и генетики СО РАН,
пр. академика Лаврентьева 10,
630090, Новосибирск, Россия
E-mail address: akberdin@bionet.nsc.ru

Федор Владимирович Казанцев
Институт цитологии и генетики СО РАН,
пр. академика Лаврентьева 10,
630090, Новосибирск, Россия
E-mail address: kazfdr@bionet.nsc.ru

Виталий Александрович Лихошвай
Институт цитологии и генетики СО РАН,
пр. академика Лаврентьева 10,
630090, Новосибирск, Россия
E-mail address: likho@bionet.nsc.ru

Станислав Иванович Фадеев
Институт математики им. С.Л. Соболева СО РАН,
пр. академика Коптюга 4,
630090, Новосибирск, Россия
E-mail address: fadeev@math.nsc.ru

Ирина Алексеевна Гайнова
Институт математики им. С.Л. Соболева СО РАН,
пр. академика Коптюга 4,
630090, Новосибирск, Россия
E-mail address: gajnova@math.nsc.ru

Виктор Кузьмич Королев
Институт математики им. С.Л. Соболева СО РАН,
пр. академика Коптюга 4,
630090, Новосибирск, Россия
E-mail address: korolev@math.nsc.ru

Алексей Елизарович Медведев
Институт теоретической и прикладной механики
им. С.А. Христиановича СО РАН,
ул. Институтская 4/1,
630090, Новосибирск, Россия
E-mail address: medvedev@itam.nsc.ru