

ОТЗЫВ официального оппонента

на диссертационную работу КАЗАНЦЕВА ФЕДОРА ВЛАДИМИРОВИЧА на тему «ИНТЕГРИРОВАННАЯ ИНФОРМАЦИОННО-КОМПЬЮТЕРНАЯ ПЛАТФОРМА ДЛЯ ИССЛЕДОВАНИЯ МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКИХ СИСТЕМ», представленную к защите на соискание ученой степени кандидата биологических наук по специальности 03.01.09 – математическая биология, биоинформатика

29 сентября 2020 г.

Актуальность Подходы математического моделирования в последнее время увеличивают свое значение как основной инструмент интеграции и анализа экспериментальных данных а также системного исследования динамических биологических систем. Независимо от того, в каком формализме будет разработана конечная математическая модель, процесс их реконструкции лучше осуществлять путём декомпозиции на отдельные «элементарные» подсистемы. Такие подсистемы могут рассматриваться как «строительные блоки» или модули, которые используются для описания/представления конкретных биомолекулярных функций и/или элементарных процессов биологических подсистем. Подход представления моделей в виде комбинации элементарных подсистем является естественным, в силу самой природы биологических систем. Существуют естественные уровни иерархии организации организмов, которое определяет естественное разбиение на функциональные подсистемы. Развивающиеся подходы стандартного представления таких модульных моделей и их поддержка программами вычислительного моделирования как общего назначения, так и узкоспециальными системами, к сожалению, пока не позволяют решать все многообразие имеющихся задач в биологии.

Диссертация Казанцева Федора Владимировича посвящена разработке и развитию методов, алгоритмов и программ для построения и анализа иерархических математических моделей молекулярно-генетических систем и решению с их помощью задач системной биологии. С применением разработанных инструментов проведено исследование моделей распределения гормона ауксина в тканях меристемы корня растения *A. thaliana*.

Научная новизна С использованием современных методов и технологий программирования разработан оригинальный компьютерный инструмент MGSmodeller для реконструкции моделей молекулярно-генетических систем широкого профиля, основанный на принципах блочно-модульной концепции моделирования. Инструментарий позволяет проводить комплексный анализ моделей с применением высокопроизводительных вычислительных систем. Исследование серии моделей особенностей распределения фитогормона ауксина в структуре клеток корня с применением разработанного инструментария позволило выявить минимальный и достаточный набор белков транспортёров ауксина для обеспечения нормального развития меристемы корня растения *Arabidopsis thaliana*. Впервые методами математического моделирования был исследован эффект влияния салициловой кислоты на активность белков транспортёров фитогормона ауксина в клетках корня *A. thaliana*. В вычислительных экспериментах показано увеличение концентрации морфогена в клетках эпидермиса, кортекса и эндодермы при воздействии низкими концентрациями экзогенной салициловой кислоты. Увеличение концентрации морфогена предсказало увеличение интенсивности делений в этих тканях, что было подтверждено экспериментально группой Клауса Пальме университета города Фрайбург.

Разработанная интегральная информационно-компьютерная платформа активно используется в работе отдела системной биологии ИЦиГ СО РАН. Она была использована

в исследованиях: синтеза биоэтанола и молочной кислоты термофильными бактериями рода *Geobacillus* (Нуриддинов и др., 2013); регуляторных механизмов поддержания плюрипотентности и дифференцировки эмбриональных стволовых клеток животных (Акбердин и др., 2014); транспортных потоков гормона ауксина по структурам клеток корня растений (Лихошвай и др., 2009; Mironova et al., 2010, 2012; Novoselova et al., 2013, 2015; Hong et al., 2017). Полученные в диссертации результаты, и в частности, созданная интернет доступная база моделей MAMMOTH с коллекцией адаптированных элементарных подсистем дыхания *E. coli*, используются в чтении лекционного курса на кафедре информационной биологии НГУ.

Структура диссертации

Характеристика диссертационной работы Диссертация состоит из введения, четырех глав, заключения, выводов, списка литературы из 178 наименований и 2-х приложений. Содержание основного текста диссертации изложено на 157 страницах, содержит 31 иллюстрацию и 13 таблиц.

Первая глава является обзором литературы, в рамках которого описан модульный подход моделирования молекулярно-генетических систем (МГС), представлены преимущества в создании, сопровождении и повторном использовании моделей из таких модулей. Далее описаны существующие программные инструменты для реконструкции и исследования моделей МГС, разделенные по степени автоматизации работы в них в процессе создания математических моделей.

Во второй главе описаны оригинальные инструменты моделирования, разработанные автором. Описан язык спецификации моделей МГС – SiBML, представлена среда моделирования MGSmodeller в терминах этого языка. Описаны основные понятия, лексемы отношения структур, принципы описания математических моделей и правила интеграции нескольких моделей в более комплексные структуры моделей целевых молекулярно-генетических систем. Описаны возможности по проведению и анализу результатов вычислительных экспериментов в полуавтоматическом виде на высокопроизводительных вычислительных платформах. В этом же разделе представлены примеры применения возможностей высокопроизводительного анализа моделей в исследовании биосинтеза этанола, лактата и ацетата в бактериальной клетке *Geobacillus* spp. и в исследовании механизмов регуляции поддержания состояния плюрипотентности эмбриональных стволовых клеток животных (*Mus musculus*).

В третьей главе представлено исследование аспектов распределения гормона ауксина в корне растения *A. thaliana* методами математического и компьютерного моделирования. В разделе 3.1 описаны базовые процессы и модели, которые послужили основой при построении моделей.

В разделе 3.2 сформулирован класс моделей меристемы корня - «ММК». В этих моделях описываются потоки гормона ауксина по клеткам меристемы корня с участием мембранных белков транспортеров PIN1, PIN2 и PIN3. Подходы, заложенные в инструментари MGSmodeller, позволили, описав 12 элементарных подсистем, формирующих процессы для одной клетки, реконструировать математические модели, соответствующие ансамблям клеток размерами 8x100, 8x50, 8x35, 8x20, 8x15, 8x10 и 8x5 клеток. Эти 12 подсистем представляют собой следующие процессы: ауксин-зависимый биосинтез белка PIN1 в клетке; ауксин-зависимый биосинтез белка PIN2 в клетке; ауксин-зависимый биосинтез белка PIN3 в клетке; ауксин-зависимая деградация белка PIN1 в клетке; ауксин-зависимая деградация белка PIN2 в клетке; деградация белка Pin3 в клетке; пассивный транспорт (диффузия) ауксина через конкретную мембрану, соединяющую две соседние клетки корня; деградация ауксина; приток ауксина в модель. В вычислительных экспериментах показано, что представленные процессы способны сформировать поток ауксина так, что его распределение в меристеме корня сформирует максимум концентрации аук-

сина в районе ниши ствольных клеток корня на каждом из представленных вариантов моделей длин корня. Формирование такого максимума в распределении ауксина является необходимым условием для поддержки функции ниши ствольных клеток корня (НСК). Таким образом показано, что рассмотренных процессов достаточно для регулирования поддержания функции НСК в развивающемся корне. В рамках получившейся модели автор также исследовал проявления мутаций по генам белков транспортеров ауксина. Было показано, что паттерн распределения ауксина с максимумом в зоне НСК достижим и при активности всего нескольких белков транспортеров (показано установление паттерна ауксина на моделях мутантов с нокаутами PIN2, PIN3, PIN2 PIN3).

В разделе 3.3 рассмотрена расширенная версия модели ММК на которой проводилось исследование влияния салициловой кислоты на процессы распределения ауксина в меристеме корня *A. thaliana* – «МСК». Экспериментальные данные для анализа модели были предоставлены лабораторией Клауса Пальме из университета г. Фрайбург (Германия). Экспериментальные данные по обработке проростков *A. thaliana* экзогенной 30 μ M салициловой кислотой. Были внесены изменения в модели подсистем биосинтеза белков транспортеров (PIN1, PIN2 и PIN3) и добавлена подсистема биосинтеза ауксина в каждой клетке. В вычислительных экспериментах показано, что именно гетерогенный эффект (одновременное изменение параметров влияния) влияния салициловой кислоты на биосинтез ауксина и активности белков транспортеров позволяет получить характерный паттерн ауксина с максимумом в зоне покоящегося центра. Кроме того, анализ результатов вычислительных экспериментов показал повышение концентрации ауксина в слоях эпидермиса, кортекса и эндодермы, по отношению к контролю, и понижение в сосудистых слоях. Повышение концентрации морфогена (ауксина) должно было отразиться на развитии корня. Опираясь на данное предположение, были более подробно изучены экспериментальные данные (корни после обработки 30 μ M салициловой кислоты) в системе iRoCS Toolbox. И действительно, на фотографиях меристемы корней со сканирующего микроскопа при обработке салициловой кислотой на 5 день после прорастания были найдены изменения в структуре ниши ствольных клеток корня, и наблюдались периклинальные деления клеток.

В четвертой главе представлен разработанный интернет ресурс MAMMOTH, являющийся базой данных с коллекцией моделей подсистем, ассоциированных с несколькими метаболическими путями бактерии *E. coli*. Данный ресурс позволяет не только просмотреть содержимое базы данных, но извлечь модели как в виде одиночных подсистем, так и в виде моделей, являющихся комбинацией нескольких выбранных подсистем. Преимуществом этой базы данных является тот факт, что представленные модели являются адаптированными экспертами к конкретным экспериментальным данным из опубликованных источников. Доступ к данным базы также можно получить напрямую, используя программный инструментарий доступа по технологии REST API. Дополняет эту базу моделей разработанный инструментарий MGSgenerator, который позволяет получить рамочную (по шаблону) математическую модель для поданной на вход программе структурной организации подсистем.

В Заключении описаны варианты использования разработанных инструментариев как платформы моделирования моделей МГС. Дальнейшее развитие модульной концепции в разделе моделирования в биоинформатике приведёт к созданию «элементной базы» и инструментов реконструкции конечных модельных объектов (по тому же принципу, что и в области электроники, где уже имеется унифицированная база электронных компонент), что, по-видимому, и определит развитие системной, а также синтетической биологии в ближайшие десятилетия.

Теоретическая и практическая значимость работы

Результаты диссертации и выводы будут интересны для ученых, сотрудников научно-исследовательских организаций, а также предприятий, использующих математическое моделирование молекулярно-генетических систем. Исследователи, изучающие развитие корней у растений, могут воспользоваться результатами математического моделирования распределения фитогормона ауксина по структуре клеток корня посредством белков-транспортёров семейства PIN. Результаты работы являются интересными и для исследователей, изучающих эмбриональные стволовые клетки животных и метаболизм бактерий *Geobacillus*. Результаты могут быть рекомендованы для включения в курсы биоинформатики, математической биологии и биотехнологии.

В развитии моделирования растительных систем, накапливая модели подсистем корня, а затем и остальных органов, поэтапно приблизимся к модели виртуального растения, на которой можно будет исследовать и прогнозировать воздействие факторов окружающей среды на промышленно-важные свойства растений.

Автореферат

Текст автореферата соответствует содержанию диссертационной работы. Содержание работы детально и качественно отражено в публикациях. По теме работы опубликованы 11 статей в изданиях из списка ВАК. Разные этапы и результаты работы были доложены на профильных российских и международных конференциях.

Выводы, сделанные по результатам работы, обоснованы и полностью соответствуют задачам, поставленным в исследовании

Замечания

1. В тексте диссертации встречаются досадные опечатки и пропущенные запятые. Местами встречаются пробелы в изложении. Например на стр. 32. В части при описании инструментария Cellzilla упомянуто, что он является расширением модуля Cellerator, описаного раньше. Но этого описания раньше не встречалось. А рисунок 1.2.3, на котором представлена модель в Cellerator, не содержит описание использованных символы и обозначений.
2. Предложенные автором лексемы языка SiBML сложны, не очевидны и затрудняют его использование не разработчиком даже при наличии у последнего опыта программирования (например см. стр. 46 диссертации)

Данные замечания не снижают значимости полученных результатов и не влияют на общую положительную оценку диссертации Казанцева Ф.В.

Заключение Диссертационная работа Казанцева Федора Владимировича «Интегрированная информационно-компьютерная платформа для исследования молекулярно-генетических систем», представленная на соискание научной степени кандидата биологических наук по специальности 03.01.09 – математическая биология, биоинформатика, является завершённой научно-исследовательской работой в области математического моделирования молекулярно-генетических систем, создания инструментов автоматизации реконструкции моделей.

В рамках работы получил развитие язык моделирования SiBML, берущий начало с работ Лихошвая В.А. по обобщенному химико-кинетическому методу моделирования. Созданный инструмент MGSmodeller позволил перенести вычислительные эксперименты и анализ моделей на современные высокопроизводительные вычислительные машины. Разработанный инструментарий MGSgenerator и база данных MAMMOTH позволили

получать в полуавтоматическом виде заготовки для новых моделей с частично адаптированными к экспериментальным данным параметрами. Полученные возможности высокопроизводительного анализа позволили оперативно решить задачи моделирования в исследованиях: синтеза биоэтанола и молочной кислоты термофильными бактериями рода *Geobacillus* (Нуриддинов и др. 2013); регуляторных механизмов поддержания плюрипотентности и дифференцировки эмбриональных стволовых клеток животных (Акбердин и др. 2014а); транспортных потоков гормона ауксин по структурам клеток корня растения (Лихошвай и др. 2009; Mironova et al. 2010; Mironova et al. 2012; Novoselova et al. 2013; Novoselova et al. 2015; Hong et al. 2017; Pasternak et al. 2019).

Разработанные в рамках диссертации модели меристемы корня продолжают использоваться в новых исследованиях по влиянию факторов внешней среды на функционирование корня.

Тема диссертации, публикации по работе и положения, выносимые на защиту, полностью соответствуют специальности 03.01.09 – математическая биология, биоинформатика.

Представленная работа по своей актуальности, научной значимости и новизне полученных результатов, соответствует всем требованиям пп. 9-14 «Положения о присуждении научных степеней», утвержденного постановлением Правительства Российской Федерации № 842 от 24 сентября 2013 г. (в редакции с изменениями, утвержденными Постановлением Правительства РФ от 21 апреля 2016 г. №335), предъявляемым к диссертациям на соискание ученой степени кандидата биологических наук, а ее автор Казанцев Федор Владимирович заслуживает присуждения ему искомой степени кандидата биологических наук по специальности 03.01.09 – математическая биология, биоинформатика.

Официальный оппонент
заведующий лабораторией геносистематики
Федерального государственного бюджетного
учреждения науки Лимнологического института
Сибирского отделения Российской академии наук (ЛИН СО РАН)

доктор биологических наук

Щербаков Дмитрий Юрьевич

640033, г. Иркутск, ул. Улан-Баторская, 3
тел. +79248330122
E-mail: sherb@lin.irk.ru

Подпись заведующего лабораторией, д.б.н. Д.Ю. Щербакова
Ученый секретарь ЛИН СО РАН

ЗАВЕРЯЮ.
к.б.н. Максимова Н.В.

