



МИНОБРАЗОВАНИЯ РОССИИ
федеральное государственное автономное
образовательное учреждение высшего образования
«Санкт-Петербургский политехнический
университет Петра Великого»
(ФГАОУ ВО «СПбПУ»)

ИНН 7804040077, ОГРН 1027802505279,
ОКПО 02068574

Политехническая ул., 29, Санкт-Петербург, 195251
тел.: +7(812)297 2095, факс: +7(812)552 6080
office@spbstu.ru

_____	№
на № _____	от _____

УТВЕРЖДАЮ:

Проректор по научной работе
Федерального государственного
автономного образовательного учреждения
высшего образования «Санкт-
Петербургский политехнический
университет Петра Великого»

член-корр. РАН, д.т.н. Сергеев В.В.

(подпись)

«25» августа 2020 г.



ОТЗЫВ

ведущей организации 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

Актуальность темы диссертационной работы

Математическое моделирование рассматривается как основной инструмент для интеграции и анализа экспериментальных данных и системного исследования динамических биологических систем. Независимо от того, в каком формализме будет разработана конечная математическая модель, процесс реконструкции удобно осуществлять путём декомпозиции на отдельные «элементарные» подсистемы, которые могут рассматриваться как «строительные блоки», описывающие конкретные биомолекулярные функции. Такой подход является естественным в силу самой природы организации биологических систем.

Развивающиеся подходы стандартного представления таких модульных моделей (Huska et al., 2003; Miller et al., 2010; Sorokin et al., 2015; Galdzicki et al., 2014) и их поддержка программами вычислительного моделирования как общего назначения (Somogyi et al., 2015;

Lopez et al., 2014; Shapiro and Mjolsness, 2016; Prlic et al., 2012) так и узкоспециальными системами (Funahashi et al., 2008; Hoops et al., 2006; Harris et al., 2016), все еще не покрывают многие вопросы в задачах реконструкции моделей, их численного анализа и визуализации результатов исследования. Разработанная в ИЦиГ СО РАН система моделирования SiBML (Likhoshvai, Matushkin et al., 2001; Likhoshvai and Ratushnyi, 2007) является перспективным заделом для дальнейшего развития. Данная работа посвящена развитию этого подхода с применением современных вычислительных возможностей.

Научная новизна исследования

С использованием современных методов и технологий программирования разработан оригинальный компьютерный инструментарий 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 позволило выявить минимальный и достаточный набор белков транспортёров ауксина для обеспечения нормальной работы корня растения *A. 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 в университете города Фрайбург.

Теоретическая и практическая значимость исследования

Разработанная 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).

Разработанная серия моделей распределения гормона ауксин в тканях корня растения: (а) позволила проследить в динамике самоорганизацию его потоков с помощью белков транспортёров PIN; (б) показала механизмы поддерживающие нормальное развитие корня в

растениях нокаутах по генам белков транспортёров PIN; (в) позволила выявить изменения анатомии кончика корня в ответ на обработку салициловой кислотой (10-50 мкМоль/л). В перспективе, поэтапно разрабатывая модели подсистем корня, а затем других органов, можно приблизиться к модели виртуального растения, на которой можно будет исследовать и прогнозировать воздействие факторов окружающей среды на промышленно-важные свойства растений.

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 и общая характеристика диссертационной работы

Диссертация имеет стандартную структуру, состоит из разделов: 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.

Введение

Во Введении четко и понятно изложены актуальность работы, в том числе описаны математические модели динамических биологических систем, декомпозиция моделей на модели элементарных подсистем, а также реконструкция комплексных моделей с помощью языка моделирования SiBML (стр. 6-8). Далее (стр.8-9) корректно сформулированы цель работы: разработка и развитие методов, алгоритмов и программ для построения и анализа иерархических математических моделей молекулярно-генетических систем и решение с их помощью содержательных задач системной биологии, а так же задачи работы.

Далее описаны Научная новизна работы (стр.9), Теоретическая и практическая значимость работы (стр.10), Положения, выносимые на защиту (стр.11), Вклад автора (стр.11), Апробация работ (стр.11-12), Публикации (стр.12-15), а также Благодарности (стр.15). Положения, выносимые на защиту, базируются на основных результатах диссертационной работы.

Глава 1 (Обзор литературы)

В Обзоре литературы описаны различные подходы к математическому моделированию в биологии, при этом наибольшее внимание уделено моделированию на уровне молекулярно-генетических систем органелл клетки, ее жизнедеятельности и процессов организации работы клеточных структур. Сделан вывод о необходимости создания новых инструментов для реконструкции и анализа структуры комплексных моделей (стр.16-18).

Далее (раздел 1.1.) приводится исчерпывающее описание блочно-модульного подхода к созданию моделей МГС. Дано обобщенное описание программных инструментов для использования такого подхода (стр.18-21). Далее (раздел 1.2.) определяются 3 категории программ: первая категория – много ручной работы, вторая – наличие инструментов автоматизации рутинных процессов реконструкции моделей, третья — решения, в которых есть автоматизация в порождении целевых моделей подсистем/тканей (стр. 21-23). В последующих разделах подробно разобраны примеры программ каждого уровня (стр.23-35).

Таким образом, в Обзоре литературы раскрыто состояние изучаемой проблемы в мировой науке, обоснованы актуальность целей и задач. Выбранные методы и подходы к проведению исследования являются современными и адекватными для решения поставленных задач.

Глава 2. Разработка среды моделирования МГС

Разработка начинается с языка спецификации моделей МГС SiBML (раздел 2.1). Формальное описание контекстно-свободной грамматики LL(1) без «е» правил приводится в приложении 1. В разделе 2.1.1. подробно описаны основные понятия SiBML (стр.38-41). В разделе 2.1.2 детально разъясняются основные лексемы языка SiBML: Идентификатор, Атрибут, Модель Элементарной Подсистемы МГС, Блок, Контейнер Блоков, Карта, Сценарий Компиляции (стр.42-43). Далее приведены исчерпывающие описания отношений, определения модели, сборки конечной модели и объединения моделей (стр.44-50, разделы 2.1.3-2.1.6). В целом, описание языка спецификации дано на высоком уровне. Продолжает изложение раздел, посвященный разработке среды моделирования MGSmodeller, которое позволяет сделать вывод о важности и широком наборе возможностей данной программы (стр.51-53). Описание возможностей языка дано в разделе 2.3 (стр.53-57). В разделе 2.4 раскрываются возможности разработанной платформы для уменьшения временных затрат на реконструкцию комплексных моделей и поиск неизвестных параметров уравнений путем адаптации программы к использованию на высокопроизводительном кластере и интеграции с сторонней программой многомерной оптимизации (стр.58-60). Изложение проиллюстрировано четкой схемой интеграции и примерами применения среды моделирования для решения биологических задач (стр.60-61). Таким образом, разработка среды моделирования понятно описана и хорошо проиллюстрирована. Среда MGSmodeller и язык SiBML соответствуют поставленным задачам, полученные с помощью их применения результаты убедительны, а сделанное заключение обосновано.

Глава 3. Моделирование потока ауксина в корне *A. thaliana*

Ауксин является одним из важнейших фитогормонов, отвечающий за регуляцию процессов деления, роста и дифференцировки клеток у растений. Целью моделирования являлось исследование способности ауксина формировать характерное распределение белков-транспортеров семейства PIN в кончике корня, ведущее к самоорганизации его паттерна распределения.

На основе литературных данных, а именно модели «Обратного фонтана» Гринайсен с соавторами и одномерной модели, разработанной в ИЦиГ СО РАН под руководством Лихошвая В.А., в среде MGSmodeller было создано семейство моделей в терминах ОДУ. В отличие от предыдущих работ, в модели введён ауксин-зависимый синтез белка PIN1, который осуществляет однонаправленный транспорт ауксина от побега к кончику корня. Показана самоорганизация паттернов экспрессии PIN1 и распределения ауксина, соответствующие экспериментальным данным. Этот механизм распределения ауксина в корне получил название «отражённая волна». Параметры модели были подобраны методами градиентного спуска для соответствия стационара модели экспериментальной кривой (стр.63-68, разделы 3.1.1-3.1.2).

В рамках представленной диссертации проводилось дальнейшее развитие модели. Была поставлена задача (раздел 3.1.3. стр.68-70) выяснения механизмов формирования распределения мембранных белков-транспортеров PIN1, PIN2, PIN3, PIN4 и PIN7, их вкладов в общую картину, а также влияния различных процессов, в частности синтеза, на формирование максимума концентрации ауксина в НСК. Решение задачи стало возможным благодаря языку SiBML и методам и алгоритмам системы MGSmodeller, включая поддержку расчётов на высокопроизводительном вычислительном кластере.

С помощью разработанной среды была построена модель меристемы корня «ММК» (раздел 3.2). Процессы в одной клетке формируются 12 элементарными подсистемами, описанными в рамках языка SiBML. Далее средствами MGSmodeller реконструируется серия моделей разных размерностей (раздел 3.2.1-3.2.2, стр.70-80). Расчёты по определению параметров модели проводились на высокопроизводительном кластере. Достигалось качественное совпадение результатов (раздел 3.2.3, стр.80-82). Анализ экспериментальных данных показал, что белок PIN2 может осуществлять активный транспорт ауксина не только апикально, но и по направлению к центральной оси симметрии корня (раздел 3.2.4, стр.82-87). Этот факт был учтен в модели «ММК:4x20.контроль». Определен набор параметров, при котором максимум ауксина располагается в 4-м ряду клеток, а локализация белков-транспортеров так же качественно совпадает с экспериментом (стр.87-89). Для всестороннего анализа модели было проведено несколько серий численных экспериментов, результаты которых подробно и наглядно проиллюстрированы в работе (стр.89-97).

Модель была дополнена рядом параметров, отвечающих за степень воздействия салициловой кислоты на процессы биосинтеза на основе данных лаборатории Клауса Пальме (г. Франкфурт, Германия) (раздел 3.3). В работе приводится подробное описание 13 элементарных подсистем модели меристемы корня с учётом влияния салициловой кислоты (МСК), а также результирующих уравнений. Проведенные расчеты подробно описаны и проиллюстрированы рисунками распределения концентрации ауксина и белков-транспортеров в стационаре. Значения параметров приводятся в таблицах. Модель предсказала увеличение концентрации ауксина в слоях эпидермиса, кортекса и эндодермы при воздействии низкими концентрациями экзогенной салициловой кислоты, что было подтверждено экспериментально группой Клауса Пальме в университете города Фрайбург. (стр.97-118).

Заключение Главы 3 обосновано, в нем резюмируется, что создание описанных моделей во многом стало возможно благодаря удобному инструментарию, разработанному в работе, и приводятся основные результаты анализа численных экспериментов.

Глава 4. База моделей элементарных подсистем МАММОTh

Для решения поставленной задачи повторного использования разработанных с использованием концепции элементарных подсистем моделей в диссертации создана база данных МАММОTh. В диссертации приведена информация структуре базы данных, которая является реляционной, с использованием XML-структур. Организация таблиц проиллюстрирована схемой (рис.4.1.2). Была использована СУБД PostgreSQL, средствами поиска которого поддерживается формат XML, который используется по этой причине вместо SiBML (стр.120-125). Далее описан графический интерфейс к базе данных (стр.126-128), в том числе приводится его внешний вид (рис.4.1.1). На основе технологии REST к базе данных выполнен программный интерфейс, позволяющий обращаться к ресурсы программам на различных языках (раздел 4.3, стр.128-129). Опираясь на накопленный опыт моделирования в работе предложен инструмент MGSGenerator для генерации шаблонных математических моделей, который был использован в исследовании метаболизма ауксина в клетке меристемы побега растения (раздел 4.4, стр.129-131).

В Заключение к главе 4 на основе изложенной ранее информации дается описание перспектив интеграции базы в цепочки программ по подбору готовых математических моделей для новых молекулярно-генетических подсистем по критериям структурной схожести или совпадения в типе взаимодействующих сущностей, что наглядно иллюстрируется приведенной схемой информационных потоков в интегральном подходе моделирования (рис. 4.5.1, стр. 131-132).

Заключение.

Приведенное Заключение по работе основывается на достигнутых в ходе работы результатах, тщательно описанных в предыдущих главах. Также делается прогноз о дальнейшем развитии блочно-модульной концепции в моделировании в биоинформатике, которое приведёт к созданию «элементной базы» и инструментов реконструкции конечных модельных объектов, по принципу, аналогичному электронике, где уже имеется унифицированная база электронных компонент.

Выводы.

По проделанной работе автором сделаны выводы о разработке языка моделирования SiBML и составляющих платформы: MGSmodeller, MGSgenerator и MAMMOTh, а также о решенных с их помощью задачах моделирования.

Сделанные выводы целиком следуют из выполненной работы, твердо аргументированы и полностью соответствуют полученным результатам, целям и задачам поставленным в работе.

Всего по теме работы опубликованы 11 статей в изданиях из списка ВАК. Из них четыре статьи опубликованы на английском языке, в том числе 2 статьи за последние 2 года. Журналы Plant Physiology и Journal of Bioinformatics and Computational Biology имеют высокий импакт-фактор. Таким образом, статьи по теме диссертации находятся на высоком мировом уровне.

Результаты работы были представлены на многих российских и международных конференциях.

В качестве общего заключения о диссертационной работе необходимо отметить ее высокое качество, научную новизну, теоретическую и практическую ценность. Все разделы информативны, изложены понятно и содержат всю необходимую информацию. Методы и разработки описаны детально, численные эксперименты были проведены и проанализированы тщательно. Результаты исследования высоко достоверны и не вызывают сомнений.

Автореферат.

Текст автореферата полно и точно соответствует содержанию диссертационной работы. Содержание работы качественно и детально отражено в публикациях. Всего по теме работы опубликованы 11 статей в изданиях из списка ВАК, в том числе четыре статьи на английском языке. Таким образом, статьи по теме диссертации находятся на высоком мировом уровне.

Результаты и выводы диссертации 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 ствольные клетки животных и метаболизм бактерий *Geobacillus*. Результаты могут быть рекомендованы для включения в курсы биоинформатики, математической биологии и биотехнологии.

Замечания.

На стр. 59 автор впервые использует термин функционал, однако его следовало подробно раскрыть, т. к. в контексте работы он до этого момента не применялся и не обозревался, и может быть неясен читателю-биологу, не знакомому с математическим аппаратом.

На стр. 115, в описании рис. 3.3.6 утверждается, что «анализ дал количественное совпадение результатов вычислительных экспериментов экспериментальным данным», однако, совпадение визуально не очень хорошее для ауксина и PIN7 (которого кстати в модели нет, должен быть PIN3, наверно). Стоило посчитать среднюю ошибку численно и применить статистический критерий для проверки гипотезы о совпадении.

На стр. 116, рис 3.3.7 А, на гистограмме не хватает «усов», поэтому утверждение у накоплении нельзя считать доказанным, его требуется проверять с помощью статистического критерия.

На стр. 116 утверждается, что «если в клетках повышается концентрация морфогена, то можно предположить, что это должно отразится на развитии корня», однако, это может быть не так если существует компенсационный регуляторный механизм.

Также в тексте присутствуют опечатки и пропущенные запятые, например:

стр 9. для обеспечение нормальной работы – обеспечения

стр 10. механизмы поддерживающие нормальное развитие корня – запятая

стр 16. существует инструменты позволяющие-существуют инструменты, позволяющие

стр 36 - системной биологии». - биологии

и др.

Замечено использование слэнга, как то:

Стр 26 — фитинг (название сантехнической арматуры)

Стр 98 корень способен расти и развиваться, не леталь

Небольшие логические провалы:

стр 65 — п. 3.1.1, но пункта 3.1 не было, не тот уровень нумерации

стр 93 рис 3.2.9. - осей ОУ и ОХ нет на рисунке.

Стр 32. В части про CellZilla упомянуто, что про Cellerator написано раньше. Но в решениях 2 уровня его нет. Рис 1.2.3 ничего не проясняет, т. к. содержит нигде не описанные символы и обозначения.

Стр 72 - организованный массив клеток с единственной точкой притока ауксина в систему – в верхние клетки сосудистой ткани. - это не точка, а граница восьми клеток, как кажется из этого пункта, но потом оказывается, что только для 4 клеток с 3 по 6.

стр 74 и 75 почему параметры скорости синтеза белков ПИН 1-3 безразмерные, а скорости деградации — неизвестная величина l (обозначение не описано) в единицу времени? Получается, что в уравнениях (плохо, что нумерации нет) на стр 78 F_s и F_d имеют разную размерность.

Следует заметить, что эти замечания не носят принципиальный характер и не умаляют научно-практической значимости диссертационной работы, они не касаются основных результатов диссертационной работы, определяющих ее новизну и научную значимость, и не снижают общей высокой положительной оценки диссертационной работы.

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 математическое моделирование молекулярно-генетических систем, создания инструментов автоматизации реконструкции моделей. В результате выполнения работы был развит язык моделирования SiBML и созданы инструменты MGSmodeller, MGSgenerator и база данных MAMMOTH. С их помощью были построены и проанализированы модели распределения фитогормона ауксина по структуре клеток корня посредством белков транспортёров семейства PIN.

Проведенная научная экспертиза диссертации, автореферата и списка опубликованных работ даёт основание положительно оценить выполненное исследование с точки зрения актуальности, степени обоснованности научных положений и выводов, степени их достоверности и новизны. Основные идеи исследования изложены ясно, выводы соответствуют полученным результатам. Все представленные в диссертационной работе результаты опубликованы в рецензируемых профильных ведущих отечественных и зарубежных журналах, рекомендованных ВАК, и широко обсуждены на научных конференциях. 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», утверждённого 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.

Отзыв на диссертацию и автореферат обсужден на заседании научного семинара научно-исследовательской лаборатории «Математической биологии и биоинформатики» ИПММ от «03» июля 2020 г.


Данные об организации

Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский политехнический университет Петра Великого», адрес: 195251, Санкт-Петербург, ул. Политехническая 29.; телефон: 8 (812) 297-20-95; электронный адрес: office@spbstu.ru; сайт: www.sbstu.ru.

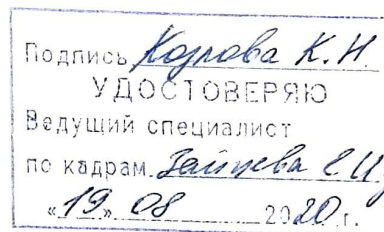
Отзыв составлен:

ведущим научным сотрудником научно-исследовательской лаборатории «Математической биологии и биоинформатики» ИПММ, кандидатом биологических наук по специальности 03.01.09 – математическая биология, биоинформатика,

Козлов К. Н.



Подпись к.б.н. К. Н. Козлова удостоверяю



Сведения о составителе отзыва:

Козлов Константин Николаевич, кандидат биологических наук по специальности 03.01.09 – математическая биология, биоинформатика, ведущий научный сотрудник научно-исследовательской лаборатории «Математической биологии и биоинформатики» ИПММ.

Адрес: 195251, Санкт-Петербург, ул. Политехническая 29.; телефон: 8 (812) 290-96-42; электронный адрес: kozlov_kn@spbstu.ru;