

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
НАУЧНОЕ УЧРЕЖДЕНИЕ
«ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
ИНСТИТУТ ЦИТОЛОГИИ И ГЕНЕТИКИ СИБИРСКОГО ОТДЕЛЕНИЯ
РОССИЙСКОЙ АКАДЕМИИ НАУК»

на правах рукописи

Казанцев Федор Владимирович

ИНТЕГРИРОВАННАЯ ИНФОРМАЦИОННО-
КОМПЬЮТЕРНАЯ ПЛАТФОРМА ДЛЯ ИССЛЕДОВАНИЯ
МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКИХ СИСТЕМ

Математическая биология, биоинформатика
03.01.09

Диссертация на соискание учёной степени
кандидата биологических наук

Научный руководитель:

к.б.н. Лашин С.А.

д.б.н. Лихошвай В.А.

НОВОСИБИРСК - 2020

Оглавление

Перечень сокращений и условных обозначений.....	5
Введение	6
Научная новизна работы.....	9
Теоретическая и практическая значимость работы	10
Положения, выносимые на защиту.....	11
Вклад автора.....	11
Апробация работ.....	11
Публикации	12
Благодарности.....	15
Глава 1. Обзор литературы	16
1.1. Подход моделирования многокомпарментных МГС	18
1.2. Программные решения	21
1.2.1 Решения первого уровня	23
1.2.2 Решения второго уровня.....	25
1.2.2.1 BioNetGen	25
1.2.2.2 DBSolveOptimum	26
1.2.2.3 BioUml	27
1.2.2.4 CellDesigner	27
1.2.2.5 Genetic Network Analyzer (GNA)	28
1.2.2.6 JSim	28
1.2.2.7 CompuCell3D.....	29
1.2.2.8 VirtualLeaf	29
1.2.2.9 CellML	30
1.2.2.10 Copasi	30
1.2.2.11 E-cyanobacterium.org	31
1.2.2.12 BIGG	31
1.2.3 Решения третьего уровня	32
1.2.3.1 Cellzilla.....	32

1.2.3.2 SiBML	34
1.2.3.3 Whole-Cell Model <i>M. genitalium</i>	35
1.3. Заключение к главе 1.....	35
Глава 2. Разработка среды моделирования МГС	37
2.1. Язык спецификации моделей МГС SiBML.	37
2.1.1 Основные понятия SiBML.....	38
2.1.2 Основные лексемы SiBML	42
2.1.3 Определение отношений в SiBML	44
2.1.4 Определение модели в SiBML	47
2.1.5 Сборка конечной модели в SiBML.....	47
2.1.6 Объединение моделей в SiBML.....	49
2.2. Среда моделирования комплексных МГС MGSmodeller.....	50
2.2.1 Средства синтаксического и грамматического анализа текстов	51
2.2.2 Спецификация вычислительного эксперимента	52
2.2.3 Принцип представления результатов моделирования	53
2.3. Постобработка результатов моделирования.....	53
2.3.1 Возможности экспорта моделей	56
2.4. Модули высокопроизводительного анализа моделей МГС.....	57
2.4.1 Параллельные вычислительные эксперименты	58
2.4.2 Интеграция со сторонними программными средствами.....	58
2.4.3 Практика применения среды моделирования	60
2.5. Заключение к главе 2.....	61
Глава 3. Моделирование потока ауксина в корне <i>A. thaliana</i>	63
3.1.1 Модель «обратного фонтана»	65
3.1.2 Модель «отражённая волна»	67
3.1.3 Формальная постановка задачи моделирования	68
3.2. Модель меристемы корня «ММК»	70
3.2.1 Модели элементарных подсистем ММК	74
3.2.2 Математическое описание модели ММК	77
3.2.3 Методы численного исследования	80

3.2.4 Анализ моделей ММК	82
3.2.4.1 Конфигурация модели «ММК:4x20.отражённая волна»	83
3.2.4.2 Конфигурация модели «ММК:4x20.обратный фонтан».	84
3.2.4.3 Конфигурация модели «ММК:4x20.двойной_механизм»	85
3.2.4.4 Модель «ММК:4x20.двойной механизм» без потока в центр .	85
3.2.4.5 Модель «ММК:4x20.контроль»	87
3.2.5 Анализ модели «ММК:4x20.контроль»	89
3.2.6 Модели ММК мутантов по генам белков транспортёров	94
3.2.7. Заключение по модели ММК	97
3.3. Модель меристемы корня с учётом влияния салициловой кислоты	100
3.3.1 Постановка задачи.....	100
3.3.2 Модели элементарных подсистем МСК	103
3.3.3 Математическое описание модели МСК	106
3.3.4 Анализ модели МСК	109
3.3.4.1 Модель «МСК:5x25.контроль»	109
3.3.4.2 Модель «МСК:5x25.гетерогенный_эффект».....	112
3.3.5 Заключение по модели МСК.....	117
3.4. Заключение к главе 3.....	118
Глава 4. База моделей элементарных подсистем МАММОTh	120
4.1. База данных	123
4.2. Графический интерфейс базы данных МАММОTh	126
4.3. Программный интерфейс пользователя (API).....	128
4.4. Генерация шаблонов элементарных подсистем MGSgenerator..	129
4.5. Заключение к главе 4	131
Заключение	133
Выводы.....	135
Список используемой литературы.....	137
Приложение 1. Определение грамматики языка SiBML	158
Приложение 2. Пример использования МАММОTh API в PYTHON	162

Перечень сокращений и условных обозначений

МГС – молекулярно-генетические системы.

МЭП – модель элементарной подсистемы.

ММК – модель меристемы корня.

МСК – модель меристемы корня с учётом влияния салициловой кислоты.

НСК – ниша стволовых клеток. В контексте данной работы относится к меристеме корня растения.

ОДУ – обыкновенные дифференциальные уравнения.

Паттерн – в данной работе под этим словом подразумевается повторяющаяся раз за разом картина распределения веществ в структуре корня.

СУБД – система управления базами данных.

ЭСК – эмбриональные стволовые клетки.

мЭСК – эмбриональные стволовые клетки мыши.

Введение

Наблюдаемый прогресс в развитии экспериментальных технологий совершил революцию в биологии и позволил исследователям генерировать огромное количество многофакторных данных с высокой степенью детализации (Johnson et al. 2007; Robertson et al. 2007; Quail et al. 2012; Goodwin et al. 2016). Развитие и доступность высокопроизводительных вычислительных ресурсов и технологий хранения данных дало новый толчок в развитии математического моделирования биологических систем (Kitano 2002; Westerhoff and Palsson 2004; Karr et al. 2012; Najafi et al. 2014; Gilbert et al. 2017; Kim et al. 2017; Cowan et al. 2019). Сегодня подходы математического моделирования все чаще рассматриваются как основной инструмент для интеграции и анализа экспериментальных данных и системного исследования динамических биологических систем (Mori 2004; Demin and Goryanin 2008; Csikász-Nagy 2009; Mironova et al. 2010; Hübner et al. 2011; Akberdin et al. 2013; Danziger et al. 2014; Sanchez-Osorio et al. 2014; Mast et al. 2014; Kim et al. 2015; Hellerstein et al. 2019). Математические модели динамических биологических систем могут быть сформулированы в терминах систем обыкновенных дифференциальных уравнений (ОДУ), уравнений в частных производных или дифференциальных уравнений с запаздывающим аргументом, стохастических моделей, а также в терминах дискретных операций и, наконец, сочетая перечисленные выше техники, в терминах т.н. гибридных подходов (Chassagnole et al. 2002; Eungdamrong and Iyengar 2004; Demidenko and Likhoshvai 2005; Feist et al. 2007; Chen et al. 2010; Lashin et al. 2010; Karr et al. 2012; Ocone et al. 2013; Danziger et al. 2014; Shapiro and Mjolsness 2016). Выбор подходящего формализма обычно зависит от сложности рассматриваемой биологической системы в смысле точности и доступности экспериментальных данных.

На текущий момент доступно более 1000 различных опубликованных моделей комплексных био-молекулярных систем, которые разработаны для

нескольких организмов (Olivier and Snoep 2004; Funahashi et al. 2008; Lloyd et al. 2008; Chelliah et al. 2015; King et al. 2016; Kim et al. 2017; Malik-Sheriff et al. 2019). Количество таких моделей постоянно растёт, однако повторное использование, расширение или модификация таких готовых математических моделей довольно нетривиальная задача (Scharm et al. 2014; Henkel et al. 2015; Kirk et al. 2015; Stanford et al. 2015; Waltemath et al. 2016). Модель часто представлена в виде законченной, жёстко структурированной в некотором формализме системы (например, в виде системы уравнений или кода программы на языке высокого уровня). Для таких моделей отсутствуют средства, позволяющие модифицировать их структурное и/или функциональное содержание в автоматическом или полуавтоматическом режиме. Однако, декомпозиция моделей на отдельные «элементарные» подсистемы является альтернативным решением озвученной проблемы. Такие подсистемы могут рассматриваться как «строительные блоки», которые используются для описания/представления конкретных биомолекулярных функций и/или элементарных процессов биологических систем (Palsson 2000; Likhoshvai et al. 2001b; Guido et al. 2006; Bruggeman and Westerhoff 2007; Lopez et al. 2014; Beal et al. 2019; Cowan et al. 2019). *Моделью элементарной подсистемы (МЭП)* мы будем называть минимальную модель подсистемы, которой достаточно для осуществления заданной функции (Ратнер 2001; Kurata et al. 2013).

Подход представления моделей в виде комбинации *элементарных подсистем* является естественным, в силу самой природы организации биологических систем. Имеются естественные уровни иерархии в организации организмов – естественное разбиение на функциональные подсистемы. Именно такой подход является перспективным в создании комплексных многоуровневых моделей систем живой природы, поддерживающий методы размещения подсистем по компартаментам, организации модельных структур в органеллы и ткани. Такой подход в декомпозиции, каталогизации, интеграции, реконструкции и исследовании

моделей приводит к вовлечению в работу современных информационно-вычислительных технологий по хранению и доступу к данным (реляционным, объектным и NO-SQL СУБД), и технологий высокопроизводительного вычислительного анализа.

Развивающиеся подходы стандартного представления таких модульных моделей (Hucka et al. 2003; Miller et al. 2010; Galdzicki et al. 2014; Sorokin et al. 2015) и их поддержка программами вычислительного моделирования как общего назначения (Prlic et al. 2012; Lopez et al. 2014; Somogyi et al. 2015; Shapiro and Mjolsness 2016) так и узкоспециальными системами (Hoops et al. 2006; Funahashi et al. 2008; Harris et al. 2016), все ещё не покрывают многие вопросы в задачах реконструкции моделей, их численного анализа и визуализации результатов исследования. Разработанная в ИЦиГ СО РАН концепция/язык моделирования SiBML (Likhoshvai et al. 2001a; Likhoshvai and Ratushnyi 2007) закрывает большой пласт вопросов в реконструкции комплексных многокомпарментных моделей в полуавтоматическом режиме, что является перспективным заделом для дальнейшего развития. Данная работа посвящена развитию этого подхода с применением современных вычислительных возможностей.

Целью работы является 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.

В соответствии с поставленной целью были сформулированы следующие задачи:

1. Развитие пакета программ SiBML для реконструкции и анализа математических моделей МГС. Разработка средств анализа математических моделей МГС для использования на высокопроизводительных вычислительных машинах.
2. Разработка базы данных моделей элементарных подсистем МГС и пакета программ, обеспечивающих хранение, широкие возможности

поиска и экспорта моделей из базы, включая механизмы генерации комплексных моделей МГС.

3. Реконструкция и исследование математических моделей молекулярно-генетических систем бактерий (рода *Geobacillus*), растений (*Arabidopsis thaliana*) и животных (*Mus musculus*).

Научная новизна работы

С использованием современных методов и технологий программирования (модульная архитектура, мета-модель и объектно-ориентированный подход для представления биологических данных) 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 с применением разработанного инструментария позволило выявить минимальный и достаточный набор белков транспортёров ауксина для обеспечения нормальной работы 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 в университете города Фрайбург.

Теоретическая и практическая значимость работы

Разработанный инструментарий активно используется в работе отдела системной биологии ИЦиГ СО РАН. Он был использован 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 растений (Лихошвай и др. 2009; Mironova et al. 2010; Mironova et al. 2012; Novoselova et al. 2013; Novoselova et al. 2015; Hong et al. 2017).

Разработанная серия моделей распределения гормона ауксина 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 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.

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.

Положения, выносимые на защиту

1. Разработанные инструменты MGSgenerator, MGSmodeller и база данных MAMMOTH формируют информационно-компьютерную платформу для реконструкции, исследования и накопления математических моделей молекулярно-генетических систем.
2. Математическое моделирование выявило различный вклад белков-транспортёров PIN в распределение концентрации морфогена ауксина в нише столовых клеток корня *A. thaliana*.
3. Теоретически показано, что фитогормон салициловая кислота в концентрации 10-50 мкМоль/л вызывает накопление морфогена ауксина в нише столовых клеток корня *A. thaliana*.

Вклад автора

Основная часть работы выполнена автором самостоятельно. В работах по исследованию моделей эмбриональных столовых клеток животных и исследованию метаболизма бактерий *Geobacillus* автор принимал участие в создании моделей и обсуждении результатов моделирования. В исследовании растительных систем автор принимал непосредственное участие в обсуждении постановок задач, реконструкции моделей, проведении вычислительных экспериментов и анализе результатов.

Апробация работ

Результаты работы докладывались и обсуждались на следующих конференциях: Международная конференция “Беляевские Чтения”, посвящённая 100-летию со дня рождения академика АН СССР Д.К. Беляева (Новосибирск, Россия, 2017); Международная конференция “Bioinformatics of

Genome Regulation and Structure\Systems Biology” (Новосибирск. Россия. 2010, 2014, 2016 и 2020 года); Международная конференция "математика. компьютер. образование" (Дубна. Россия. 2016); Международная конференция “Математическое моделирование и высокопроизводительные вычисления в биоинформатике, биомедицине и биотехнологии” (Новосибирск, Россия. 2016); Международный симпозиум «Plant Organ Growth Symposium» (Bordeaux. France. 2019); Международная конференция «9th European Conference on Mathematical and Theoretical Biology» (Gothenburg. Sweden. 2014); Международная конференция «VI конгресс Вавиловского общества генетиков и селекционеров» (Ростов-на-Дону. Россия. 2014); Международный семинар «BREW 2013: Bioinformatics Research and Education Workshop» (Berlin. Germany. 2013); Конференция "Параллельные Вычислительные Технологии (ПаВТ) 2012" (Новосибирск. Россия. 2012); Международная конференция «European Conference on Mathematical and Theoretical Biology (ECMTB 2011)» (Krakow. Poland. 2011); VI московский международный конгресс “Биотехнология: состояние и перспективы развития” (Москва. Россия. 2011).

Публикации

По теме диссертации было опубликовано 11 печатных работ в изданиях из списка ВАК, а также 5 авторских свидетельств на программное обеспечение и базы данных.

1. Pasternak T, Groot E, **Kazantsev F**, Teale W, Omelyanchuk N, Kovrizhnykh V, Palme K, Mironova V. (2019) Salicylic Acid Affects Root Meristem Patterning via Auxin Distribution in a Concentration-Dependent Manner. *Plant Physiol* 180:1725–1739. doi: 10.1104/pp.19.00130.
2. **Kazantsev F**, Akberdin I, Lashin S, Ree N, Timonov V, Ratushny A, Khlebodarova T, Likhoshvai V. (2018) MAMMOTh: A new database for

- curated mathematical models of biomolecular systems. *J Bioinform Comput Biol* 16:1740010. doi: 10.1142/S0219720017400108.
3. Акбердин ИР, Иванисенко НВ, **Казанцев ФВ**, Ощепкова ЕА, Омелянчук НА, Матушкин ЮГ, Афонников ДА. (2014) Моделирование механизмов регуляции поддержания плюрипотентности эмбриональных стволовых клеток: кинетический и стохастический подходы. *Математическая биология и биоинформатика* 9(2):504-517.
 4. Novoselova ES, Mironova VV, Omelyanchuk NA, **Kazantsev FV**, Likhoshvai VA. (2013) Mathematical modeling of auxin transport in protoxylem and protophloem of *Arabidopsis thaliana* root tips. *Journal of bioinformatics and computational biology. J. Bioinform. Comput. Biol* 11(1):1340010. doi:10.1142/S0219720013400106.
 5. Нуриддинов МА, **Казанцев ФВ**, Розанов АС, Козлов КН, Пельтек СЕ, Колчанов НА, Акбердин ИР. (2013) Математическое моделирование синтеза биоэтанола и молочной кислоты термофильными бактериями рода *geobacillus*. *Вавиловский журнал генетики и селекции* 17(4/1):686-704.
 6. **Казанцев ФВ**, Акбердин ИР, Подколотный НЛ, Лихошвай ВА. (2012) Новые возможности системы MGSmodeller. *Вавиловский журнал генетики и селекции* 16(4/1):799-804.
 7. Mironova VV, Novoselova ES, Doroshkov AV, **Kazantsev FV**, Omelyanchuk NA, Kochetov AV, Mjolsness E, Likhoshvai VA. (2012) Combined *in silico/in vivo* analysis of mechanisms providing for root apical meristem self-organization and maintenance. *Annals of botany* 110(2): 349-360.
 8. Акбердин ИР, **Казанцев ФВ**, Лихошвай ВА, Фадеев СИ, Гайнова ИА, Королев ВК, Медведев АЕ. (2009) Компьютерная система интеграции модулей для автоматической генерации и численного анализа математических моделей молекулярно-генетических систем. *СибЭМЖ*. VI:440-456.

9. **Казанцев ФВ**, Акбердин ИР, Безматерных КД, Лихошвай ВА. (2009) Система автоматизированной генерации математических моделей генных сетей. Информационный вестник ВОГиС 13(1):163-169.
10. Акбердин ИР, **Казанцев ФВ**, Омелянчук НА, Лихошвай ВА. (2009) Математическое моделирование метаболизма ауксина в клетке меристемы побега растения. Информационный вестник ВОГиС 13(1):170-175.
11. Лихошвай ВА, Омелянчук НА, Миронова ВВ, **Казанцев ФВ**, Акбердин ИР, Королев ВК, Фадеев СИ, Колчанов НА. (2009) Моделирование регуляции ауксином инициации латеральных органов у *Arabidopsis thaliana* L. Информационный вестник ВОГиС 13(1):176-185.

Авторские свидетельства

1. Лихошвай В.А., **Казанцев Ф.В.**, Акбердин И.Р., Безматерных К.Д. Программа автоматической генерации математических моделей генных сетей (МГСгенератор) / Software tool for mathematical models autogeneration on basis of the gene networks structure (MGSgenerator), СВИДЕТЕЛЬСТВО №2008611941, 18 апреля 2008
2. Лихошвай В.А., **Казанцев Ф.В.**, Акбердин И.Р., Безматерных К.Д. Лашин С.А., Подколодная Н.Н., Ратушный А.В. Компьютерная система для конструирования, расчета и анализа моделей молекулярно-генетических систем (МГСмоделлер) / A computer system for reconstruction, calculation and analysis mathematical models of molecular genetic system (MGSmodeller), СВИДЕТЕЛЬСТВО №2008612820, 9 июня 2008.
3. Лихошвай В.А., Акбердин И.Р., **Казанцев Ф.В.**, Насонов В.В., Тимонов В.С. Программа визуализации и компиляции математических моделей генных сетей (МГСмодели) / Software tool for gene networks mathematical models view and compile (MGSmodelsDB). СВИДЕТЕЛЬСТВО № 2011616329, 12 августа 2011.

4. Лихошвай В.А., Акбердин И.Р., **Казанцев Ф.В.**, Ри М.Т., Ри Н.А., Хлебодарова Т.М., Лашин С.А., Ощепкова Е.А., Ратушный А.В. База данных «Элементарные подсистемы: метаболизм *E.coli*» (ЭлСи: *E.coli*) / Database «Elementary subsystems: *E.coli* metabolism» (ElSy: *E.coli*) СВИДЕТЕЛЬСТВО № 2012620064, 16 января 2012.
5. Акбердин И.Р., Нуриддинов М.А., **Казанцев Ф.В.**, Пельтек С.Е. Программа для исследования кинетической модели биосинтеза молочной кислоты термофильными бактериями рода *Geobacillus* (Элси: Geobac) СВИДЕТЕЛЬСТВО № 2014610722, 16 января 2014.

Структура и объём 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.

Благодарности

Автор выражает благодарность и глубокую признательность научному руководителю Лашину С.А. за поддержку, помощь, обсуждение результатов и терпение. Автор выражает благодарность д.б.н. Лихошваю В.А. и к.б.н. Ратушному А.В. как первым научным руководителям, которые направляли автора в ходе обучения в университете и аспирантуре. Отдельная благодарность к.б.н. Акбердину И.Р., к.б.н. Мироновой В.В., Омельянчук Н.А. и к.т.н. Деменкову П.С. за консультации и плодотворные научные дискуссии. Автор выражает благодарность академику Колчанову Н.А. и всем сотрудникам Отдела Системной Биологии ИЦиГ СО РАН. Без их поддержки эта работа не состоялась бы.

Отдельную благодарность хочу выразить своей супруге, Казанцевой Надежде Владимировне, за понимание и поддержку.

Глава 1. Обзор литературы

Математическое/компьютерное моделирование в биологии покрывает большой спектр задач: от исследования эффектов на уровне атомных взаимодействий (Medvedev et al. 2014; Chavent et al. 2016) (в случае исследования процессов молекулярной динамики), до агентного моделирования взаимодействия на уровне популяций и экосистем (Lashin and Matushkin 2011; Song et al. 2014; Klimentenko et al. 2019). В рамках данной работы мы ограничимся некоторым промежуточным уровнем – молекулярно-генетических систем органелл клетки, её жизнедеятельности и процессов организации работы клеточных структур (тканей и органов).

Процессы матричного синтеза молекул удобно исследовать в рамках «булевых моделей» (de Jong 2002; Kauffman 2004; Kwon and Cho 2007) или методик, рассматривающих стохастические события (Turner et al. 2004; Hoops et al. 2006; Xia et al. 2011; Rybiński et al. 2012). Для исследования путей передачи сигналов и белок-белковых взаимодействий применимы как «булевы модели» (Covert et al. 2008; Garg et al. 2008; Sanchez-Osorio et al. 2014), так непрерывные модели (Likhoshvai and Ratushnyi 2007; Chen et al. 2010; Osone et al. 2013), для которых существуют инструменты позволяющие учитывать природу объекта исследования, допускающую наличие нескольких активных состояний одной сущности (Faeder et al. 2009). Исследование метаболических путей требует учёта пространственных локализаций вовлечённых объектов (Feist et al. 2007; Cooling et al. 2008; McCloskey et al. 2013).

При всем многообразии доступных данных и технологий, на сегодняшний день разработано всего несколько полногеномных моделей бактериальных клеток с достаточной степенью проработки от процессов генетического синтеза молекул до уровня организации и взаимодействия органелл (Feist et al. 2007; Covert et al. 2008; Karr et al. 2012; Hastings et al. 2019). Причиной такого отставания является малое количество программных продуктов, решающих

задачи реконструкции и исследования моделей комплексных, гетерогенных молекулярно-генетических систем оперирующих на уровнях межклеточных взаимодействий, устройства тканей, органов и организма в целом.

Получение аналитических решений на моделях комплексных молекулярно-генетических систем практически невозможно. Поэтому все активнее применяются методы и подходы высокопроизводительного анализа включающие численные исследования с наглядной визуализацией результатов. Для получения более системного взгляда на функционирование модели, раз за разом повторяются вычислительные эксперименты с различными значениями параметров. Тестируется множество случайных наборов параметров в биологически-осмысленных границах для получения статистически значимых результатов. Тестируются несколько вариантов модели (наличие/отсутствие подсистем) на одной серии параметров в поисках изменений поведения моделей в одинаковых условиях внешней среды.

Таким образом, для успешной работы в области моделирования молекулярно-генетических систем необходимо иметь инструментарий, который:

1. Помогает в реконструкции математических моделей биологических систем, адекватно их особенностям строения и функционирования на разных уровнях организации.
2. Позволяет строить разнообразные «портреты» анализируемых систем, как в виде линейных диаграмм, поверхностей, так и в виде интерактивных инструментов, поддерживающих анимацию.
3. Позволяет выстраивать серию вычислительных экспериментов в конвейеры обработки данных, для выполнения высокопроизводительного анализа области параметров моделей.

Для пунктов 2 и 3 можно использовать готовые решения. В частности, встроенные в универсальные среды моделирования (Mathematica, Matlab, Scilab, Octave) инструменты отображения диаграмм или специальные системы, такие как ParaView (<https://www.paraview.org/>) или GnuPlot

(<http://www.gnuplot.info/>), хорошо подходят для визуализации данных. А готовая функциональность *bash* скриптов *nix систем (http://www.opennet.ru/docs/RUS/bash_scripting_guide/) или *.bat файлов команд для Windows (https://en.wikibooks.org/wiki/Windows_Batch_Scripting) является замечательным подспорьем в построении инструментальных конвейеров. Однако в пункте 1, в задаче реконструкции моделей, все ещё много ручной работы, требующей внимательности исследователя и знаний особенностей строения и функционирования объекта моделирования.

В процессе создания модели большой системы, на пути от наброска до конечной комплексной структуры, постоянно идёт процесс добавления/изменения/замены подсистем, под влиянием смены гипотез о механизмах или несоответствия модельной динамики экспериментальным данным. Рост ошибок при внесении правок в процессе «ручного» отслеживания растёт пропорционально размеру модели. Для цели минимизации вероятности внесения ошибок, вызванных человеческим фактором, и с целью автоматизации анализа структуры создаваемых моделей, необходимы современные подходы и решения реконструкции и структурной визуализации моделей.

1.1. Подход моделирования многокомпарментных МГС

Естественной, вытекающей из природы объекта исследования, является блочно-модульная концепция создания моделей МГС. В рамках этой концепции объект исследования рассматривается как набор взаимодействующих компонент – блоков/модулей, которые взаимодействуют друг с другом только через общих участников (Ратнер 1992; Ратнер 2001). В качестве такого блока могут выступать ферментативные реакции, подсистемы матричного синтеза, или даже целый метаболический путь. Удобством такого подхода служит возможность распределённой разработки модулей, и их

интеграция в серию конечных моделей с возможностью добавления/исключения/изменения ряда подсистем (Miller et al. 2011). Возможность получить варианты одной модели для разностороннего исследования явления, под влиянием разных условий среды, разных вариантов механизма протекания процесса, а также исследовать проявления мутаций.

Блочно-модульный подход позволяет строить модель сколь угодно сложной биомолекулярной системы итеративно, по уровням иерархии, от простого к сложному, объединяя модули в рамках единой *in silico* биологической системы (Milo et al. 2002; Фадеев и Лихошвай 2003; Alon 2003; Ратушный и др. 2005; Danino et al. 2010). При этом получается библиотека готовых блоков, которые можно повторно использовать в других моделях, имитируя тем самым естественный путь эволюции, где природа чаще использует готовые молекулярно-генетические конструкции, модифицируя и адаптируя их под новые условия окружающей среды, нежели, как инженер, создаёт новые функциональные подсистемы с нуля (Jacob 1977; Ратнер 1992). Именно опираясь на данную концепцию, авторы ресурса *Biomodels.net* (Le Novere 2006) увеличили содержимое базы на 140 000 моделей подсистем МГС в рамках проекта Path2Models (Büchel et al. 2013), проведя автоматическую генерацию моделей и взяв за основу структурные модели из базы метаболических путей KEGG (Kanehisa 2000).

В общем случае модуль-блок представлен абстракцией $B(S, P, V, F)$. Здесь: S – это набор участвующих биохимических субстанций; P – набор параметров; V – закон обработки информации в некотором формализме F . Серия блоков однозначно задаёт участников модели. Список задействованных в блоках формализмов (F) определяет выбор вычислительных алгоритмов и программных средств решения модели – или моно-формализм (булевы модели, непрерывные или в частных производных), или гибридный подход с понятными механизмами сообщения между модулями разных формализмов (Hogg et al. 2013; Osone et al. 2013; Le Novere 2015). Таким образом, получается небольшой набор программ-конвертеров в известные форматы систем

численного исследования, использование которых минимизирует время, требующееся на освоение новых форматов представления данных, подходов и правил реконструкции моделей и способов внесения изменений в них.

Преимущество блочной организации ярко проявляется при исследовании многокомпонентных многокомпарментных систем, характерных для многоклеточных организмов высших эукариот. Например, исследование формирования органов и тканей в растительных системах требуют построение математических моделей клеточных структур (Merks et al. 2011; Shapiro et al. 2013; Hirashima et al. 2017), где в каждой клетке работает своя генная сеть, обменивающаяся продуктами синтеза и переработки с соседями. Однако, тиражирование генных сетей, функционирующих в рамках одной клетки, на всю клеточную структуру без привлечения средств автоматизации – ресурсоёмкая задача. При использовании блочно-модульного подхода манипуляции с моделями происходят в рамках поддающихся тиражированию блоков.

Использование блочно-модульного подхода позволит не только создать модель одной тиражируемой клетки, но и позволяет централизованно вносить изменения в подсистемы. Изменения в структуре одного из блоков может/будет отражаться на всех его копиях в структуре модели, что затруднительно сделать в моделях представленных монолитной системой математических уравнений.

Существует множество программных продуктов, которые позволяют в той или иной степени создавать модели в рамках блочно-модульного подхода (Campagne et al. 2004b; Yang et al. 2005; Dräger et al. 2008; Faeder et al. 2009; Le Novere et al. 2009). Ниже будут подробно рассмотрены несколько из таких систем. Какой бы не был выбран инструментарий, нужно помнить, что результатами не только необходимо будет поделиться с мировым сообществом, но и постараться сделать так, чтобы «сообщество не прошло мимо» – одной лишь публикации может быть недостаточно. Результаты должны быть оформлены таким образом, чтобы требовали минимальных

усилий сторонних исследователей для повторения. Лучше, если результаты оформлены в формате, поддерживающем автоматический разбор набором программных инструментов (Hucka et al. 2003; Le Novere et al. 2009; Courtot et al. 2011; Galdzicki et al. 2014). Это увеличит шанс включения моделей в публичные банки данных (например, *Biomodels.net*). Централизованное хранилище модулей-блоков открывает новую парадигму работы с моделями. Реализуемые такими ресурсами программные интерфейсы доступа к данным позволяют:

- осуществлять поиск существующих блоков под созданную структурную модель исследуемых процессов
- осуществлять поиск модулей-блоков по количеству участников и их роли, в попытке предсказать поведение системы по известным для других организмов механизмам.
- строить модели искусственных метаболических систем, формируя цепочки блоков.
- осуществлять поиск фундаментальных и широко представленных в различных таксонах регуляторных мотивов (Milo et al. 2002; Alon 2007)
- реконструировать, исследовать и накапливать модели подсистем независимо (Kurata et al. 2013).

1.2. Программные решения

Приступая к задаче реконструкции и исследовании моделей молекулярно-генетических систем, необходимо определиться с инструментарием. Зачастую приоритет отдаётся тому инструментарию, с которым уже сложились отношения (преподавали в вузе, работа в лаборатории идёт на устоявшихся технологиях). Естественно, в области работы с последовательностями для языков программирования высокого уровня Java, Python и Perl появляются библиотеки работы с объектами предметной области: BioJava (Prlic et al. 2012;

Lafita et al. 2019), BioPython (Cock et al. 2009), BioPerl (Stajich 2007). В области системной биологии разрабатываются собственные инструменты и концепции (Hucka et al. 2002; Courtot et al. 2011; Lopez et al. 2014) ряд из которых не выдерживает проверки временем (Campagne et al. 2004a; Coskun et al. 2013; Najafi et al. 2014). Разработки, описанные в обзорах трёх-пяти летней давности, зачастую не поддерживаются авторами. Ориентация на список информационных систем, которые поддерживают SBML-стандарт представления моделей молекулярно-генетических систем (296 инструмента, на момент обращения http://sbml.org/SBML_Software_Guide), позволяет скорее проследить эволюцию и тренды в разработке инструментов системной биологии во времени.

С другой стороны, успешно проводятся исследования в универсальных средах моделирования (Matlab, Mathematica) или в собственных программах, написанных на языках высокого уровня (C++, Java). Проанализировав возможности существующих/поддерживаемых сейчас программных решений, эти решения разделили на три категории. Критерий отбора – принципы работы с данными (рис. 1.2.1).

Классификация осуществляется по объёму ручной работы при создании математических моделей. Решения из первой категории – много ручной работы. Решения из второй категории – наличие инструментов автоматизации рутинных процессов реконструкции моделей. К третьей категории относятся решения, в которых есть автоматизация в порождении целевых моделей подсистем/тканей.

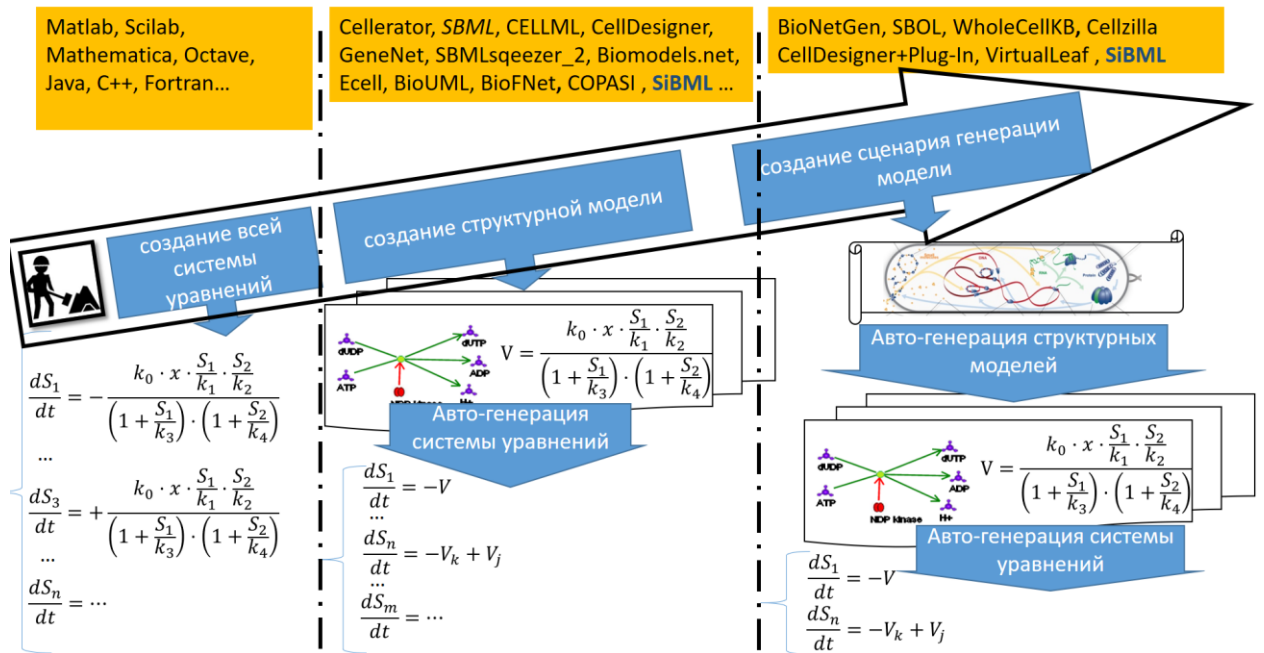


Рисунок 1.2.1 – Разделение программных решений на три категории по принципу работы с данными.

1.2.1 Решения первого уровня

На первом уровне работают универсальные среды математического и компьютерного моделирования (такие как Matlab, Mathematica, Scilab, Octave и др.) и языки программирования высокого уровня (C/C++, Java, Python, Perl, R и др.), в которых необходимо весь код модели писать самостоятельно. В этих условиях в основном проводятся исследования моделей, состоящих из десятков процессов. Большое количество процессов в такой методологии сложно контролировать в силу того, что вся ответственность за корректность кода модели лежит на авторе этого кода. Внесение изменений или смена базовых гипотез функционирования подсистем модели приводят к пересмотру всей структуры модели – ручному поиску мест, где должны быть отражены эти изменения.

Для моделей, представленных на этом уровне, характерны следующие свойства:

- Не к каждой опубликованной статье прилагается файл с оформленной в каком-либо формализме математической моделью (Приведены уравнения в тексте статьи, например). Динамика реконструированных по таким статьям моделей может не совпадать с результатами статьи. Для решения проблемы воспроизведения результатов был предложен стандарт “TEDDY” (Courtot et al. 2011), в котором прописываются все значения параметров модели и соответствующий им характер поведения кривых, что повышает точность и повторяемость результатов независимо от метода численного решения.
- Практика реконструкции математических моделей МГС как конечной системы, решающей поставленную задачу обработки вектора входных параметров/переменных в вектор параметров/переменных на выходе, где большая часть параметров модели обезличена. Т.е. узнать то, что означают переменные « x_1, \dots, x_n » в модели можно только из текста статьи, ассоциированной с моделью. Вносить изменения (расширять/дополнять/изменять уравнения) в такого рода модели, представленные системами ОДУ, например, затруднительно, выискивая вклад того или иного процесса по всей системе уравнений. Частично эту проблему решают подходы по аннотации моделей (Hucka et al. 2003; Le Novère et al. 2005; Courtot et al. 2011; Juty et al. 2012).
- Решение задачи расширения модели (добавление новых участников или изменение механизма формирующих модель подсистем) требует переработки и анализа всего кода модели.

При переходе на уровень моделей метаболических путей, путей передачи сигналов и далее уже необходимы другие подходы. Богатый выбор инструментов из универсальных инженерных сред моделирования не предусматривает учёт особенностей строения и функционирования моделей молекулярно-генетических систем как объектов живой природы. Эти особенности - компартментная организация, разделение мест синтеза и мест

влияния синтезируемых веществ, явления полиаллельности геномов (синтез молекул одной функциональной нагруженности с разных мест генома, под своими регуляторными конструкциями) и др. Учёт этих аспектов в процессе исследования приводит к разрастанию масштабов моделей и сложностям их дальнейшего сопровождения (модификация/анализ/интерпретация/интеграция).

1.2.2 Решения второго уровня

На втором уровне находятся программные решения, которые используют средства автоматизации рутинных процессов реконструкции моделей. Часто использующие собственный формализм описания конструкций предметной области, скрывая внутренности моделей за «фасадами». Ниже приведены несколько интересных инструментов.

1.2.2.1 BioNetGen

BioNetGen (Faeder et al. 2009; Smith et al. 2012; Chylek et al. 2015; Harris et al. 2016) Инструмент с открытым исходным кодом для построения моделей в терминах «основанных на правилах» концепции реконструкции (Rule-based modeling). Данный подход решает вопросы комбинаторики при росте модели в системах молекулярно-генетических взаимодействий: пути передачи сигналов в клетках, белок-белковые взаимодействия.

Процесс создания модели осуществляется в терминах собственного языка BNGL (the BioNetGen language). Задаётся набор молекул, их активные центры и набор правил взаимодействия молекул через активные центры. Далее алгоритм итеративно генерирует возможные комбинации молекул, объединяя их в процесс/цепочку реакций. В случае наличия свободных активных центров в новых сгенерированных комплексных молекулах, следующий шаг итерации

алгоритма строит новые связи. И так продолжается пока не закончатся все возможные комбинации, или алгоритм не достигнет точки останова.

Программный комплекс поддерживает вычислительные эксперименты с получившимися моделями в терминах систем обыкновенных дифференциальных уравнений и терминах стохастических вычислений. Поддерживается экспорт моделей в форматы SBML (Hucka et al. 2015) и Matlab. Есть режимы работы через консоль и с поддержкой графического интерфейса пользователя.

В рамках данного инструмента можно наглядно наблюдать что произойдёт если в организме начнётся синтезироваться новая молекула с существующим активным центром, т.е. молекула которая несёт ту же функцию что и существующая в системе. Алгоритм берет на себя все проблемы с комбинаторикой в построении всех комбинации реакций/молекул, увеличивая размер модели. Подходит для решения задач моделирования направленной генной инженерии.

1.2.2.2 DBSolveOptimum

DBSolveOptimum (Gizzatkulov et al. 2010) (<http://insysbio.ru/en/software/db-solve-optimum>, <http://sourceforge.net/projects/dbsolve>) – инструмент представляет собой некоторую графическую «обвязку» над серией рутинных операций, которые выполняет исследователь при решении задачи реконструкции моделей в терминах дифференциальных уравнений. Код модели все ещё пишется самостоятельно, включая сами уравнения. Реализован продвинутый анализ моделей: средства анализа вариабельности параметров, используемые в системной фармакологии; анализ особых точек; бифуркационный анализ; подбор (фитинг) параметров под экспериментальные точки. Реализован импорт/экспорт моделей в формализме SBML.

1.2.2.3 BioUml

BioUml (<http://wiki.biouml.org/index.php/Landing>) – Инструмент визуального моделирования и проведения вычислительных экспериментов. Поддерживает отображения SBML моделей в SBGN стандарте. Поддерживают вычислительные эксперименты с моделями следующих видов:

- системы обыкновенных дифференциальных уравнений
- стохастические модели
- агентное моделирование (Agent-based)
- модели гемодинамики
- потоковые модели

Поддерживает отображение моделей в формализме BioNetGen (представленного выше).

1.2.2.4 CellDesigner

Инструментарий графической реконструкции моделей молекулярно-генетических систем (Funahashi et al. 2008). Модель создаётся из набора блоков (субстанция, реакция и связи между ними) в рамках технологии *point-and-click*. Построение модели осуществляется в формализме SBML с дополнительно введёнными блоками аннотации. Модели реконструируются в рамках некоторого алфавита конструкций. Если придерживаться стандарта оформления, то получается самодокументированная модель (Типизация объекта, характер его связей и отношений с другими объектами). Реализована возможность подключения дополнительных инструментов анализа к самой системе в виде расширений (доступна библиотека расширений, Plug-in) и за счёт возможности экспорта/импорта моделей в стандартном формализме SBML. Однако при переносе моделей в другие инструменты через механизмы

экспорта возможна потеря информации (это касается информации в аннотациях, характерной для CellDesigner, но не являющейся частью SBML).

1.2.2.5 Genetic Network Analyzer (GNA)

Инструмент для построения моделей и проведения вычислительных экспериментов регуляторных генных сетей (Batt et al. 2012). Основное направление – помощь исследователю в построении модели генетической регуляции с использованием знаний о регуляторном взаимодействии в комбинации с данными о генетической экспрессии. Построение моделей идёт в виде **качественных** моделей, представленных кусочно-линейными дифференциальными уравнениями. Так как точные значения параметров сложно найти, пользователи системы вводят ограничения в виде неравенств. Этих данных достаточно для построения графа перехода состояний, который описывает качественное поведение регуляторной сети. Инструмент разработан как приложение на языке JAVA.

1.2.2.6 JSim

JSim (<https://www.physiome.org/jsim/>) – написанный на языке программирования JAVA инструмент для построения и анализа моделей в соответствии с экспериментальными данными. Ориентирован на решения задач физиологии и биомедицины. В процессе вычислительного эксперимента могут быть смешаны следующие подходы: дифференциальные уравнения, дифференциальные уравнения в частных производных, неявные уравнения, интегралы, сумматоры, дискретные события и процедурный код, при необходимости. Поддерживается импорт/экспорт в форматах SBML и CellML.

1.2.2.7 CompuCell3D

CompuCell3D (<http://www.compuCell3d.org/>) – инструментарий построения и проведения вычислительных экспериментов на моделях клеток, клеточных тканей (Swat et al. 2012). Инструмент для моделирования в терминах методологии «Клеточных горшочков» (Cellular Pots Models). Описание модели фиксируется в XML-подобном языке описания. Если необходимы расширенные возможности инструмента, поддерживается работа с использованием языка программирования Python.

1.2.2.8 VirtualLeaf

VirtualLeaf (Merks et al. 2011) – инструмент моделирования морфогенеза тканей растения в рамках методологии «Cellular Pots Models». Представляет из себя программу, написанную на языке C++. Это агентная модель, поддерживающая механизмы роста и деления клеток. Модель, в рамках данного инструмента, представляет собой скомпилированную C++ библиотеку. В рамках программного кода модели автор должен запрограммировать поведение каждой клетки: процессы жизнедеятельности клетки, процессы транспорта через/в клеточные мембраны, условия и характер деления клеток. Этот инструмент обладает хорошей визуализацией (Рисунок 1.2.2) с поддержкой анимации.

Однако возможности его использования ограничены заданной разработчиком функциональностью, расширение которой требует внесение изменений в код самой программы. Соответственно, перенос разработанной модели в другой инструментарий не возможен. В рамках конференции ЕСМТВ 2014 (<https://imc.zih.tu-dresden.de/wiki/multisysbio2014/doku.php>) разработчики похожих систем обсуждали решение проблемы совместимости данного класса моделей.

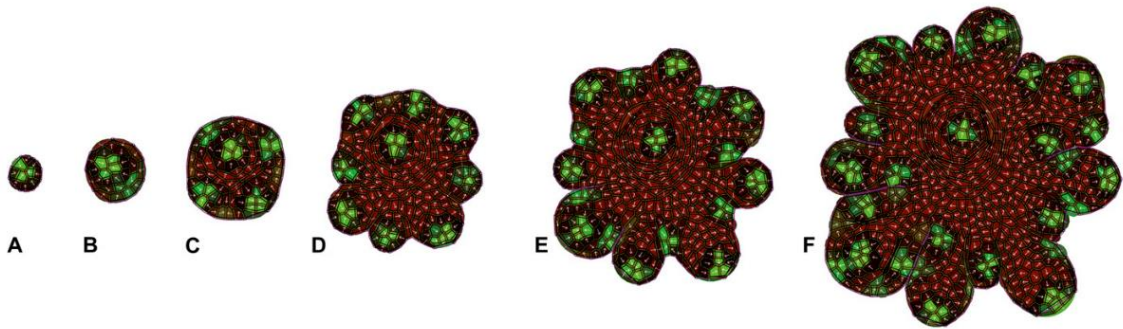


Рисунок 1.2.2 – Временные срезы вычислительного эксперимента по влиянию распределения гормона ауксин на рост меристемы побега растения (Merks et al. 2011).

Именно в невозможности отделить модель от инструмента – основной недостаток таких систем. То, что не заложено изначально в решении – нельзя получить. Только если модифицировать сам инструмент.

1.2.2.9 CellML

CellML (Miller et al. 2010) – язык описания и инструментарий проведения вычислительных экспериментов моделей. Язык описания является XML-подобным (как и SBML). Реконструкция моделей в терминах блоков-модулей, для решения задачи повторного использования блоков. Ориентация на представление моделей в терминах дифференциальных уравнений, терминах дифференциальных алгебраических уравнений и линейной алгебры. Имеется своя база готовых моделей (Lloyd et al. 2008). Поддерживается набор инструментов реконструкции, проведения вычислительных экспериментов и визуализации. Есть инструменты конвертации из SBML в CellML и обратно.

1.2.2.10 Copasi

Copasi (Hoops et al. 2006) – инструмент для построения моделей МГС в терминах модульного подхода. Модели описываются в привычном для

биологов биохимическом формализме. Прописывается сначала структурная модель – набор подсистем. После чего, для каждой из подсистем или выбирается один из предустановленных математических законов протекания процесса, или описывается свой собственный. Есть возможность разделения подсистем по компартментам, но нет инструментов генерации комплексных модельных структур. Благодаря блочно-модульной организации модели модификация математических законов подсистем осуществляется точно. Широкий набор готовых инструментов по анализу моделей (поиск стационарных решений, сводные характеристики модели). Есть возможность импорта/экспорта моделей в формате SBML. Имеется возможность выгрузки данных для комплексного анализа результатов в сторонних инструментах.

1.2.2.11 E-cyanobacterium.org

E-cyanobacterium.org (Trojáková et al. 2016) – это интернет-платформа для накопления и публичного обмена, аннотации, анализа и визуализации динамических моделей и экспериментальных данных относящихся к *цианобактериям*. Модели представлены в нескольких уровнях абстракции: как сеть биохимических процессов и как система обыкновенных дифференциальных уравнений. Ориентируются на сопоставление сущностей математических моделей формализованному биохимическому описанию, принятому консорциумом. Основная цель платформы – соединить биологические знания с теми преимуществами, что даёт математическое представление динамических процессов.

1.2.2.12 BIGG

BIGG (<http://bigg.ucsd.edu>) – интернет ресурс представляющий собой базу аннотированных потоковых моделей метаболических путей (King et al. 2016).

На ресурсе представлены 85 полногеномных (genome-scale) моделей для нескольких организмов. Элементы моделей аннотированы ссылками на публичные банки данных, такие как NCBI, KEGG, PubChem. Встроен инструмент для отображения и редактирования моделей метаболических путей Escher (King et al. 2015). Для выполнения вычислительных экспериментов используется инструментарий Cobra toolbox для инженерной среды моделирования Matlab (Becker et al. 2007). Есть поддержка экспорта в формат SBML.

1.2.3 Решения третьего уровня

На третьем уровне – решаются вопросы создания многокомpartmentных/многокомпонентных моделей, описывающих поведение систем масштабов тканей, органов или целых организмов. Применяются подходы «сценариев» конструирования систем. Принципиальная схема работы таких систем следующая: создание набора моделей подсистем как блоков-модулей; описание сценария взаимодействия этих блоков как по принадлежности к компартаментам, так и по межкомpartmentным взаимодействиям; помещение готовых подсистем в клеточный компартимент и прописывание процессов общения клеток со средой или другими клетками. И так далее, до искомого уровня организации модели.

1.2.3.1 Cellzilla

Cellzilla (Shapiro et al. 2013) – инструментарий для моделирования динамики тканей, с возможностью учёта роста и деления клеток. Использовалась для моделирования меристемы побега *A. thaliana*. Это библиотека методов для инженерной системы моделирования Mathematica. Является дополнением к пакету Cellerator (Shapiro et al. 2003),

представленному в решениях «второго уровня». Сначала модель описывается набором базовых подсистем процессов, протекающих в одном компартменте. Затем вводится дополнительный функционал: указываются, какие из элементов участвуют в процессах диффузии между клетками; указываются процессы направленного транспорта (с указанием законов протекания реакции); отдельно описывается структура моделируемой ткани. Есть серия предустановленных методов для генерации квадратных и гексагональных клеточных культур. Отдельно есть методы для ввода своего оформления ткани клеток. По заданной структуре ткани осуществляется автоматическая генерация моделей путём тиражирования подсистем для каждой клетки и прописывания процессов транспорта через мембраны соседствующих клеток.

Например, известная модель «репресселятора» (Elowitz and Leibler 2000) в структуре гексагональных клеток с процессами диффузии веществ между клетками в Cellzilla будет выглядеть так как показано на рисунке 1.2.3.

```
{mYx1 → ∅, kd1},
{mYx2 → ∅, kd2},
{mYx3 → ∅, kd3},
{mYx3 ⇨ mYx1, fdrRprsltr[kbas, 1, gam, 1, myFunction]},
{mYx1 ⇨ mYx2, fdrRprsltr[kbas, 1, gam, 1, myFunction]},
{mYx2 ⇨ mYx3, fdrRprsltr[kbas, 1, gam, 1, myFunction]},
{mYx1 → mYx1, Diffusion[DmYx1]},
{mYx2 → mYx2, Diffusion[DmYx2]},
{mYx3 → mYx3, Diffusion[DmYx3]},
```

Рисунок 1.2.3 – Набор процессов, в формате системы Cellerator, формирующих модель репресселятора.

Однако, данный набор инструментов ориентирован на решение задач морфогенеза меристем растений. Использование данного инструментария для задач из других областей может вызвать ряд трудностей в процессе адаптации конструкций формализма. Для представленных в главе 3 моделей использование пакета Cellzilla нецелесообразно, так как описание специфичных для клеток в слое транспортных потоков породит множество

неиспользуемых в модели конструкций, усложняя модель и повышая вычислительную нагрузку.

1.2.3.2 SiBML

В Институте Цитологии и Генетики СО РАН д.б.н. Лихошваем В.А. с коллегами была разработана методология реконструкции и исследованию моделей молекулярно-генетических систем, базирующаяся на блочно-модульной концепции моделирования, названная им Обобщённым Химико-Кинетическим методом Моделирования (ОХКММ) (Likhoshvai et al. 2001a; Likhoshvai and Ratushnyi 2007). Подход базируется на представлении моделей как совокупности независимых подсистем (ферментативные реакции, ассоциация/диссоциация белков, события транспорта и др.). Модель конечной системы собираются из блоков подсистем по заданному сценарию. Внося изменения в сценарий, можно получить вариации одной модели, например, для исследования воздействия разных факторов окружающей среды на динамику генной сети или проявления мутаций. Расширенные правила, сформулированные рекомендации и пожелания к работе по реконструкции моделей МГС в рамках ОХКММ в итоге получили название «*Метод Элементарных Подсистем*» (Лихошвай 2008). Следуя правилам и рекомендациям метода при реконструкции моделей целевого объекта, на выходе должны получаться математические модели искомым систем в удобном для сопровождения и анализа виде. Этот подход успешно применяется в ИЦиГ СО РАН при решении задач исследования комплексных МГС (Oshcherkova-Nedosekina and Likhoshvai 2007; Likhoshvai et al. 2010; Хлебодарова и др. 2013; Novoselova et al. 2015).

Заложенные в методологию концепции закрывают большую часть аспектов реконструкции моделей в рамках инструментария «третьего уровня».

Имеется свой формализм представления моделей и инструментарий проведения вычислительных экспериментов.

1.2.3.3 Whole-Cell Model *M. genitalium*

В 2012 году была представлена масштабная модель целого организма – *M. genitalium* (Karr et al. 2012). Модель представлена комбинацией из 28 функциональных процессов, представленных в наиболее подходящем формализме. Например, метаболизм представлен в терминах потоковой модели (flux-balance analysis), активность РНК полимеразы моделируют с применением стохастических подходов. Код модели жизненного цикла бактерии реализован в системе Matlab. Авторами был написан собственный вычислительный метод, который обеспечивал синхронизацию расчётов каждого из модулей модели. Собранный авторами набор параметров из доступных источников был оформлен как самостоятельная база данных (Karr et al. 2014).

1.3. Заключение к главе 1

В работах авторов, связанных с построением таких комплексных моделей (Macklin et al. 2014; Medley et al. 2016; Waltemath et al. 2016), отмечают, что для успешного построения масштабных моделей живых организмов недостаточно усилия одной группы – это должна быть целенаправленная работа всего сообщества по направлениям:

1. Экспериментальные вопросы – экспериментальные измерения в рамках одной клетки, пространственно-распределённые, в масштабах генома и наличие данных по динамике изменения.

2. Структурирование данных – организация в виде доступных аннотированных банков экспериментальных данных с унифицированным доступом.
3. Построение моделей и интеграция – расширение, унификация, внедрение новых методов анализа в системах моделирования.
4. Высокопроизводительные вычисления – Разработка параллельных вычислительных техник и частных аппаратных решений.
5. Анализ данных и визуализация – Применение алгоритмов машинного обучения и подходов визуальной интерпретации данных.
6. Валидация моделей – разработка метрик для оценки корректности моделей.
7. Разработка в рамках коллабораций и сообществ – организация конференций и научных соревнований.

Авторы отмечают нехватку методов и алгоритмов для работы на этом уровне, а также акцентируют внимание на необходимости решения задач автоматизации на всех этапах (создание, расширение, валидация моделей). Авторы призывают: «Только коллективный вклад сообщества в *разработку и наполнение информационных ресурсов* приведёт к качественному скачку в области биоинформатики/системной биологии».

Глава 2. Разработка среды моделирования МГС

2.1. Язык спецификации моделей МГС SiBML.

Существует две основных концепции к построению моделей молекулярно-генетических систем: от общего к частному (top-down), и от частного к общему (bottom-up). В первом случае рассматривается общая постановка задачи, которая делится на подзадачи меньшей сложности. Во втором случае, комбинируя готовые решения, наращивается сложность модели до уровня, необходимого для решения поставленной задачи. Концепция блочно-модульной реконструкции моделей молекулярно-генетических систем хорошо подходит в обоих случаях. Разработанная Лихошвай В.А. с коллегами методология реконструкции и исследования моделей базируется на этой концепции (Лихошвай 2008). В её рамках модель представляется как совокупность независимых подсистем (ферментативные реакции, ассоциация/диссоциация белков, события транспорта и др.). Конечная модель собирается из моделей-блоков подсистем по заданному сценарию. При следовании правилам и рекомендациям метода реконструированные математические модели целевого объекта получают в удобном для сопровождения и анализа виде.

В рамках диссертационной работы велось развитие данного инструментария. Были разработаны: современные инструменты синтаксического анализа текстов в заданном формализме, инструменты постановки и проведения вычислительных экспериментов. Это позволило получить полноценный *Язык Моделирования* – SiBML (Siberian Biology Modeling Language). Наличие собственного языка представления моделей, наиболее подходящего под конкретную задачу – довольно распространённый путь (Blinov et al. 2004; Cooling et al. 2008; Shapiro et al. 2013; Galdzicki et al. 2014).

Язык SiBML не является XML-подобным языком разметки, и для работы с ним была реализована серия инструментов на языке программирования JAVA: была формально описана грамматика (см. приложение 1) в виде контекстно-свободной $LL(1)$ грамматики без « e » правил (Рейуорд-Смит 1988), что позволило реализовать прямой алгоритм разбора текстов формализма, где по очередному символу, после уже выведенной части «слова», однозначно определяется какое правило из грамматики выбрать; средство синтаксического анализа; средство проверки полноты описания блока текстов и идентификации мест ошибок. Без таких инструментов анализа текстов судить о корректности (в смысле полноты присутствия процессов, которые планировалось ввести в модель) собранной комплексной модели можно было только по косвенным признакам.

2.1.1 Основные понятия SiBML

Основной рабочей единицей языка служит *модель элементарной подсистемы МГС*. *Элементарная подсистема МГС* – это молекулярный процесс или группа процессов, которые могут рассматриваться и моделироваться изолированно от других процессов (рисунок 2.1.1 А). Типичными представителями элементарных подсистем являются «Ферментативные реакции», процессы регуляции экспрессии генов, или даже метаболический путь. В зависимости от уровня детализации задачи, одна элементарная подсистема может быть представлена набором нескольких элементарных подсистем уровня ниже (Например, подсистема ферментативной реакции как набора комбинации этапов образования и распада комплекса из фермента, субстрата и продукта реакции) (рис. 2.1.1 Б). Формально, *элементарная подсистема (ЭП)* представлена тройкой (A, B, V) , где A - вектор входных переменных, B - вектор выходных переменных и V - внутренний закон функционирования подсистемы.

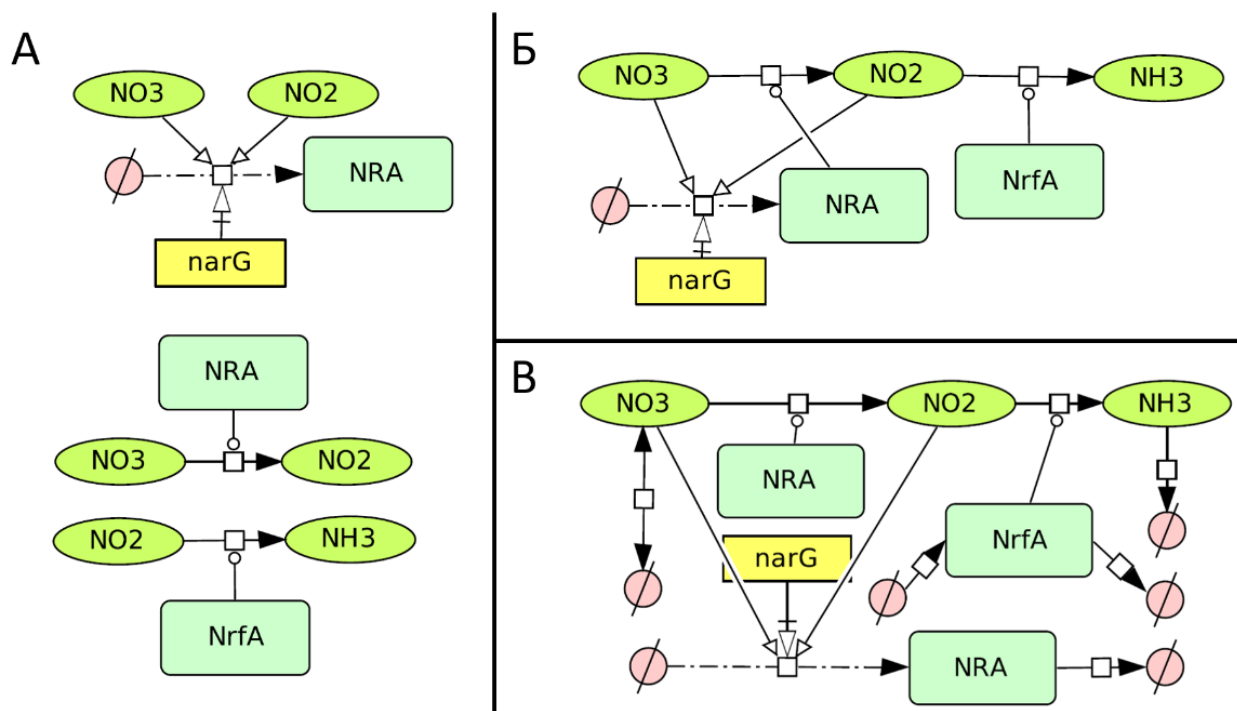


Рисунок 2.1.1 – Принципы отношений элементарных подсистем: А) Три независимые элементарных подсистемы; Б) Подсистема, являющаяся комбинацией из трёх; В) Модель, дополненная процессами притока и деградации веществ. (Kazantsev et al. 2018)

Важной особенностью *элементарных подсистем*, как биологических систем, служит то, что две и более подсистемы могут быть или *независимы*, или *связанными входами и/или выходами* (рис. 2.1.1 Б). Модель элементарной подсистемы МГС – это модель, в которой специфицированы участники взаимодействия и параметры математической модели, т.е. заданы конкретные биохимические элементы, вступающие во взаимодействие, и для закона протекания процесса заданы значения всех параметров (рис. 2.1.1 В).

Реконструкция математических моделей в терминах концепции элементарных подсистем при подходе «*top-down*» (от общего к частному) условно разделена на 4 основных этапа.

- 1) Определение области моделируемого явления. Построение структурной модели исследуемого процесса или явления, с целью определения задействованных объектов (субстанций, клеточных структур, органелл) и их отношений.

- 2) Декомпозиция на *элементарные подсистемы*. По имеющейся структурной модели осуществляется процесс декомпозиции (разбиения) на более мелкие подсистемы (органеллы, процессы транспорта, синтеза и другие функциональные элементы) до того уровня, на котором возможно описать математическую модель подсистемы заданного уровня подробности.
- 3) Разработка математических *моделей элементарных подсистем*. Для каждой из выделенных подсистем реконструируется самостоятельная математическая модель.
- 4) Объединение математических моделей. Объединение *математических моделей элементарных подсистем* в единую математическую модель. В итоге получая модель, которая с заданной степенью подробности описывает исследуемое явление.

После проведения такой декомпозиции в первый раз, часть *моделей элементарных подсистем* может быть повторно использована для исследования новых явлений или при других постановках задачи. Созданный набор *моделей элементарных подсистем* готов к решению новых задач уже в рамках подхода «*bottom-up*» (от частного к общему).

Возникающие конфликты при построении модели в рамках «*bottom-up*» подхода, связанные с возможными пересечениями по внутренней структуре (т.е. может нарушиться принцип элементарности) *моделей элементарных подсистем* при формировании комплексной структуры, решаются дополнительной декомпозицией «проблемных» подсистем до того уровня элементарности, пока не будет устранён конфликт.

Для корректной интеграции *моделей элементарных подсистем* в комплексные структуры необходимо максимально полно специфицировать моделируемые объекты (и формализм SiBML это позволяет). Необходимо отразить такие свойства моделируемых объектов как: название, принадлежность к классу объектов (ген, белок, РНК и др.), их локализация в структуре моделируемых компартментов (органеллы, клетки, принадлежность к ткани и др.), состояние (фосфорилирование, ацетилирование,

свободны/заняты активные центры молекулы), структура молекулярного комплекса и т.д. Это позволяет не только корректно идентифицировать объекты исследования, но и моделировать такие особенности организации МГС как:

- состояние молекулы ДНК в разные периоды клеточного цикла;
- существование обратимых состояний белковых молекул, меняющих их свойства;
- явления полиаллельности генома;
- анизотропия пространственных компартментов живых систем, разделяющая процессы биохимического синтеза, что добавляет регуляторные воздействия на систему процессами мембранного транспорта;
- эпигенетическая передача наследственной информации.

В итоге, в рамках формализма *SiBML* получается не только разрабатывать математические модели молекулярно-генетических систем в терминах, близких к естественной организации моделируемых явлений, но и создавать *базы математических моделей элементарных подсистем*. Такие базы будут содержать *модели элементарных подсистем* разной природы и функциональной направленности (транспорт молекул, ферментативный синтез, деградацию молекул и т.д.), модели органелл. Если рассматривать *модель элементарной подсистемы* как некоторое правило протекания биохимического процесса, то и базы самих *моделей элементарных подсистем*, и базы сценариев сборки в комплексные системы можно рассматривать как шаблоны (заготовки) для моделей явлений похожей природы. Например, создав *модель элементарной подсистемы* реакции ферментативного синтеза, можно повторно использовать эту модель для других ферментов, предполагая схожесть в механизмах протекания этих биохимических реакций.

2.1.2 Основные лексемы SiBML

В данном разделе описаны основные лексемы языка SiBML: *Идентификатор*, *Атрибут*, *Модель Элементарной Подсистемы МГС*, *Блок*, *Контейнер Блоков*, *Карта*, *Сценарий Компиляции*.

За определение моделируемых сущностей отвечает базовая единица языка – *ИДЕНТИФИКАТОР*. *Идентификатор* может служить как определение моделируемого объекта, определение параметра математической модели, определение служебного параметра и др. *Идентификатор* характеризуется набором *атрибутов*, которые определяют конкретные свойства определяемого *Идентификатором* объекта. Каждый *Атрибут* задаётся парой *наименование* и *значение*. Например, *Атрибут* «*type(protein)*» имеет *наименование* «*type*» и значение атрибута «*protein*», что интерпретируется как принадлежность определяемого объекта к типу биологических сущностей - «белки». *Атрибут* без заданного *наименования* не имеет смысла и считается ошибкой. *Атрибут* с не заданным *значением* (*пустой атрибут*) допустим, но средствами обработки игнорируется, интерпретируется как его отсутствие. Количество *атрибутов* в идентификаторе не ограничено. Для более тонкой настройки программ компиляции моделей, *наименование атрибутов* в *идентификаторах* объектов могут сопровождаться служебными символами (например, '*' или '#'. Интерпретацию и варианты использования см. в спецификации языка, приложение 1). В *идентификаторе* набор атрибутов задаётся перечислением через символ запятой «,».

Таблица 2.1 – Логика организации идентификаторов SiBML

идентификатор в языке SiBML	Семантика
<C(nucleus), I(CYP79B2), T(gene)>	Переменная, отвечающая за ген с названием <i>cyp79b2</i> находящийся в компартменте <i>nucleus</i>
<C(nucleus), I(CYP79B2), T(rna)>	Переменная, отвечающая за РНК с названием <i>CYP79B2</i> (синтезированная с гена <i>cyp79b2</i>) находящуюся в компартменте <i>nucleus</i>
<T(substance), I(tryptophane)>	Переменная, отвечающая за низкомолекулярное вещество с названием <i>tryptophane</i>
<T(protein), I(IAMT1p)>	Переменная, отвечающая за белок с названием <i>IAMT1p</i>
<C(c4.05), I(PIN3), T(protein)>	Переменная, отвечающая за белок с названием <i>PIN3</i> находящийся в клеточной структуре в клетке с индексом (4;5).
<C(c4.16), I(PIN3), T(protein)>	Переменная, отвечающая за белок с названием <i>PIN3</i> находящийся в клеточной структуре в клетке с индексом (4;16).

Смысл *идентификатора* зависит от контекста его расположения. *Атрибуты* используются как универсальные единицы фиксации в *идентификаторе* информации, используемой при реконструкции моделей (таблица 2.1). Например, при реконструкции многокомpartmentных моделей идентификаторы вовлечённых субстанций и параметры всех процессов, протекающих в конкретном компартменте, снабжаются соответствующим свойством выраженном атрибутом “С” (от слова «компартмент»).

2.1.3 Определение отношений в SiBML

Реконструкция математических моделей в терминах языка SiBML основывается на использовании *моделей элементарных подсистем* МГС как строительных блоков комплексной модели целевой МГС. Так как биологические процессы часто протекают по одинаковым механизмам, естественным подходом является организация библиотеки шаблонов *элементарных подсистем*. Используя такой шаблон, достаточно только специфицировать участников и получить элементарную подсистему, отвечающую за взаимодействие конкретных реагентов по заданному закону.

Отношения между объектами в терминах языка SiBML задаются в рамках лексемы *блок*, которая задаёт участников взаимодействия и правила их взаимодействия. *Блок* однозначно характеризуется своим типом (*список идентификаторов*), упорядоченным списком формальных динамических переменных \bar{X} (*список идентификаторов*), упорядоченным списком формальных параметров \bar{P} (*список идентификаторов*) и законом преобразования информации \bar{F} . На сегодня существуют 72 типа *блоков* (Лихошвай 2008), каждый из которых задаёт своё поведение (таблица 2.2).

В простейшем случае *модель элементарной подсистемы* МГС может быть задана одним *блоком* (рис. 2.1.2). В том случае, когда *элементарная подсистема* имеет сложный механизм, её описание может состоять из нескольких *блоков*. Для того, чтобы ассоциировать (связать) несколько *блоков* с одной *моделью элементарной подсистемы*, необходимо задать одинаковый *идентификатор принадлежности* для всех *блоков* (рис. 2.1.2).

Таблица 2.2 – Некоторые блоки SiBML (по Лихошвай В.А., 2008).

тип блока	Формальное представление блока	Описание блока
$\langle = \rangle$	$X = (x_1, x_2, x_3), P = (k_1, k_2),$ $\bar{F} : \frac{dx_1}{dt} = k_2 \cdot x_3 - k_1 \cdot x_1 \cdot x_2, \quad \frac{dx_1}{dt} = \frac{dx_2}{dt} = -\frac{dx_3}{dt}.$	<p>Обратимая бимолекулярная реакция</p> $x_1 + x_2 \xrightleftharpoons[k_2]{k_1} x_3$
1->N	$X = (x, y_1, y_2, \dots, y_n), P = (k),$ $\bar{F} : \frac{dx}{dt} = -\frac{dy_1}{dt} = -\frac{dy_2}{dt} = \dots = -\frac{dy_n}{dt} = -k \cdot x, n \geq 0.$	<p>Необратимая мономолекулярная реакция</p> $x \xrightarrow{-k} y_1 + y_2 + \dots + y_n$
1=>N	$X = (x_1, \dots, x_n), P = (k), F : \frac{dx_i}{dt} = k, i = 1, \dots, n, n \geq 1.$	<p>Конститутивный синтез</p> $\xrightarrow{-k} x_1 + \dots + x_n$

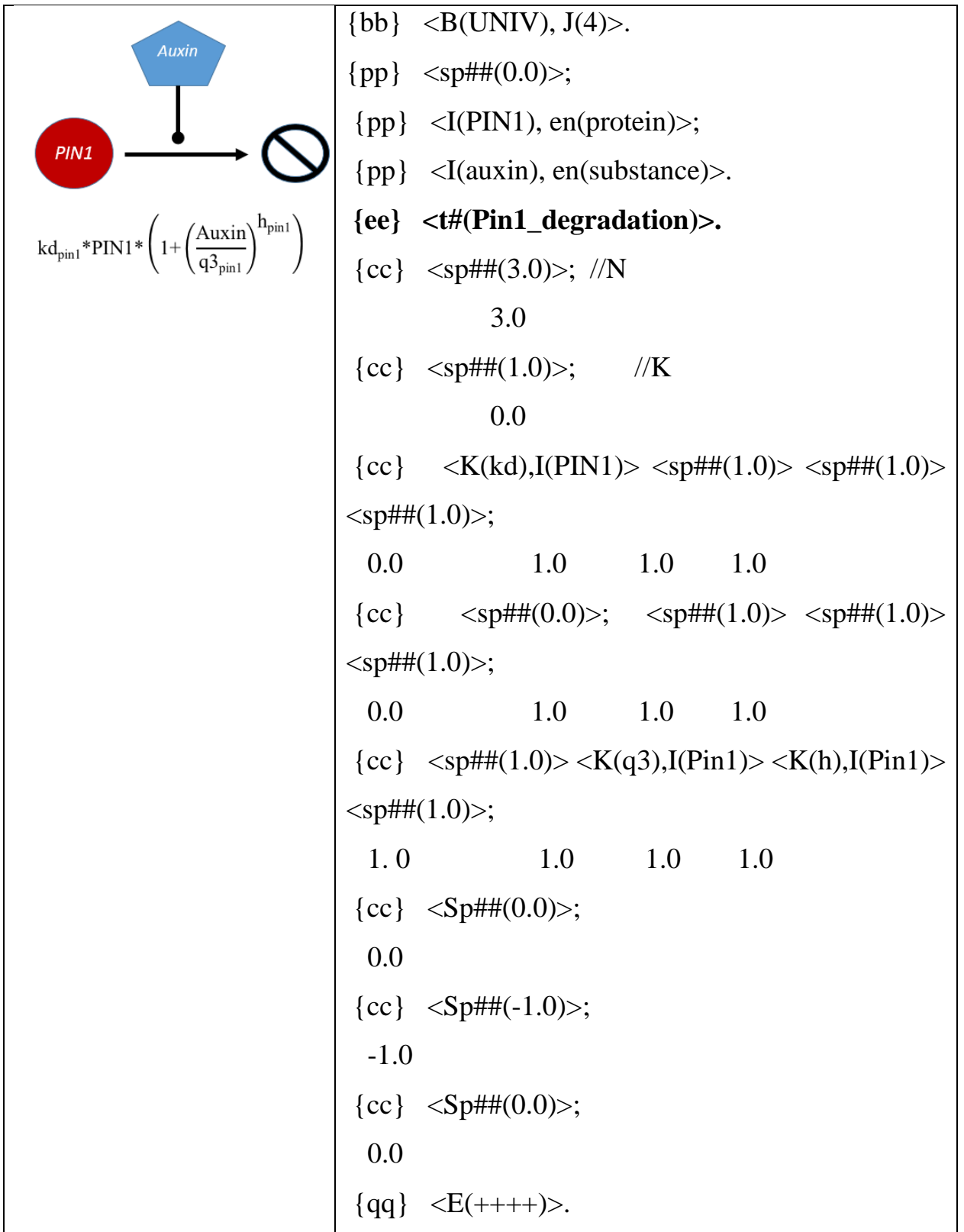


Рисунок 2.1.2 – Модель элементарной подсистемы деградации белка PIN1 при участии фитогормона ауксин. Возможное структурное представление и математическая формула (слева). Представление модели в терминах языка SiBML (справа). Жирным выделен идентификатор принадлежности блока к конкретной элементарной подсистеме.

2.1.4 Определение модели в SiBML

Математическая модель в рамках языка SiBML задаётся упорядоченным набором *блоков* (минимум 1 *блок*, максимум ограничен техническими возможностями компьютера). Это должен быть именно упорядоченный список, так как некоторые *элементарные подсистемы* могут быть представлены несколькими блоками со вспомогательными выражениями, и они должны быть рассчитаны до того, как будет вычислена основное выражение для *элементарной подсистемы*. Далее, в контексте реконструкции модели, под *элементарной подсистемой* будет пониматься *блок* или набор *блоков*, которые соответствуют *элементарной подсистеме*.

Блоки могут быть созданы самостоятельно, или же с использованием дополнительных инструментов, ряд которых уже встроен в инструментарий языка SiBML. Если не хватает возможностей существующего инструментария, всегда можно создать собственные подпрограммы на любом языке программирования высокого уровня (например, C++, Java, Python, R), которые, опираясь на принципы строения идентификаторов языка SiBML сгенерируют *блоки/модели* нужного процесса.

2.1.5 Сборка конечной модели в SiBML

Процесс реконструкции модели на языке SiBML строится на основе трёх сущностей: *контейнер блоков*, *карта*, *сценарий компиляции*.

Контейнер блоков – это файл (с расширением «.mod») с набором *моделей элементарных подсистем*. Модели в этом файле могут быть собраны по любым принципам, например, принадлежащие к определённому метаболическому пути, или представленные в одном компартменте, или формирующие модель целевого объекта исследования, или библиотека шаблонных подсистем.

Карта – это файл (с расширением «.map») со списком *идентификаторов* (заключённых в кавычки «'»), оканчивающийся служебным *идентификатором* « '*q#(++++)*'>». Набор *идентификаторов* в карте соответствует *идентификаторам принадлежности элементарных подсистем*. *Карта* служит правилом (шаблоном) для осуществления выборки элементарных подсистем из *контейнера блоков*. Так как в *контейнере блоков* могут быть *элементарные подсистемы* с различными *идентификаторами принадлежности* (библиотека *элементарных подсистем*), выбор конкретных подсистем регулируется *картой*.

Сценарий компиляции – это файл (с расширением «.сmp») с упорядоченным списком *идентификаторов* (заклучённых в одинарные кавычки «'»), оканчивающийся служебным *идентификатором* « '*q#(++++)*'>». Один из встроенных инструментов SiBML разбирает *сценарий*, осуществляя генерацию модели. Структура *идентификаторов* имеет служебные *атрибуты*, которые определяют процесс компиляции комплексной модели МГС. *Сценарий компиляции* фактически указывает какие *модели элементарных подсистем* из *контейнера блоков* должны быть извлечены в итоговую модель (*идентификаторы элементарных подсистем* указаны в файлах *карт*).

Такая организация работы алгоритма позволяет создавать модели несколькими путями:

- Описывать каждую из моделируемых *элементарных подсистем* МГС в рамках одного файла с одинаковыми *идентификаторами принадлежности*.
- Разносить функционально, логически или иным способом связанные *элементарные подсистемы* по отдельным *контейнерам блоков* и формировать конечную модель комбинируя для них *карты*.
- Комбинация первых двух подходов.

2.1.6 Объединение моделей в SiBML

Одним из преимуществ языка SiBML является возможность интегрировать несколько различных готовых моделей в одну комплексную. В случае, когда *элементарные подсистемы* каждой из объединяемых моделей были реконструированы в единых правилах формирования *атрибутов* в *идентификаторах* и не имеют двусмысленной интерпретации, даже пересечение *элементарных подсистем* в моделях (что характерно для моделей метаболических путей) будет согласовано.

Процесс создания математических моделей МГС – итеративный. Начинается все с небольших моделей подсистем, которые затем обрастают подробными механизмами и интегрируются с новыми подсистемами, образуя комплексную модель. Например, когда уже есть модель метаболического пути и необходимо добавить новую цепочку процессов, которая отвечает за альтернативный путь с вовлечением продуктов новых или внедрённых генов. Причём необходимо сравнить поведение этих двух систем при одинаковых условиях и проанализировать степень воздействия внедрённых конструкций. Для решения такой задачи с помощью SiBML надо: 1) Реализовать *модели недостающих элементарных подсистем*; 2) оформить для них *карту*; 3) Добавить блок правил в *сценарий компиляции* модели; 4) Запустить процесс сборки модели обоих *сценариев*; 5) провести вычислительные эксперименты и анализ.

Таким образом, модификация исходной модели в рамках языка SiBML может осуществляться на всех уровнях путём модификации, добавления, удаления или комбинацией этих действия как для *моделей элементарных подсистем*, так и для функционально связанных их наборов.

Результаты моделирования в SiBML представлены в следующих вариантах: динамика изменения всех переменных моделей, динамика изменения скоростей процессов, значения переменных в финальной точке

расчета, динамика и значения в финальной точке выбранных переменных модели. Структура *идентификаторов* переменных модели в SiBML позволяет написать сторонними средствами инструменты постобработки данных (используя язык программирования Python, например).

2.2. Среда моделирования комплексных МГС MGSmodeller.

Для процесса моделирования в терминах языка спецификации моделей МГС SiBML на современных аппаратных платформах была разработана среда MGSmodeller. MGSmodeller реализован на языке программирования JAVA по технологии Rich Client Platform Application (RCP) в среде ECLIPSE (http://wiki.eclipse.org/Rich_Client_Platform) и предоставляет набор программных инструментов, расширяющих возможности базовых инструментов SiBML (Казанцев и др. 2012). MGSmodeller включает ранее отсутствующие средства:

- Синтаксического анализа текстов SiBML. Подсказки о допущенных ошибках в коде моделей и указания этих мест с точностью до позиции в строке.
- Грамматического анализа текстов SiBML. Подсказки где и в каких блоках текста ошибки – некорректное количество элементов в описании блоков переменных и/или параметров.
- Постановки численных экспериментов на высокопроизводительных вычислительных комплексах.
- Модули экспорта моделей в серию форматов популярных сред моделирования (Matlab, Mathematica, SBML) для тех случаев, когда необходим анализ модели, не поддерживающийся в среде «MGSmodeller».

2.2.1 Средства синтаксического и грамматического анализа текстов

Основной принцип работы в рамках SiBML - работа в консоли и обычном текстовом редакторе. Разработанный инструментарий по синтаксическому анализу текстов языка моделирования позволил выявлять ошибки синтаксиса и грамматики до момента компиляции модели, показать место в тексте где они произошли. Разработанный графический интерфейс среды MGSmodeller позволил раскрасить текст модели и подсветить структуры языка (рис. 2.2.1). Это позволило понизить порог вхождения для новых пользователей и улучшить восприятие конструкций языка.

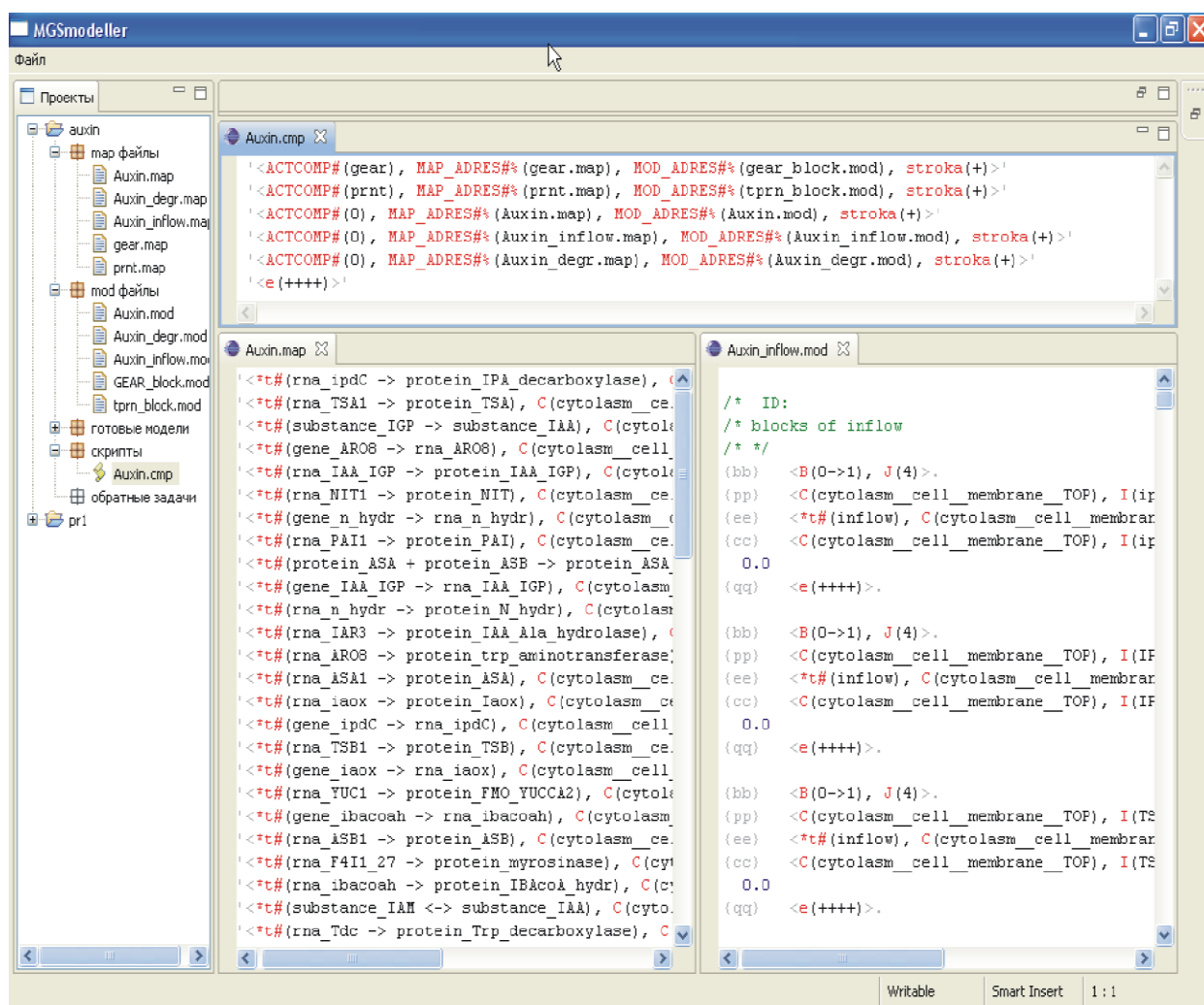


Рисунок 2.2.1 – Графический интерфейс системы MGSmodeller.

Инструментарий по анализу грамматики проверяет лексемы языка и их структуру (количество переменных и параметров в *блоке* и их функции в конечной модели). Автоматическое сопоставление описания *элементарной подсистемы* МГС с грамматически верным шаблоном контролирует количество переменных и параметров, минимизируя вероятность возникновения ошибок на этапах сборки модели.

2.2.2 Спецификация вычислительного эксперимента

В общем случае, модели МГС, реконструированные в рамках языка SiBML, являются гибридными. Т.е. модели содержат непрерывные, дискретные и стохастические части. Дискретные и стохастические события (в том случае, если они заданы) рассчитываются в определённые моменты времени, которые задаются как параметры модели. В промежутках между этими моментами времени рассчитывается непрерывная часть модели.

В рамках языка SiBML при постановке задачи численного эксперимента есть возможность управления параметрами модели:

- 1) Установка временного интервала для расчёта. Интересной особенностью средств численного эксперимента является то, что можно задавать неограниченное количество временных интервалов, в которых задан некоторый шаг. В итоге система будет рассчитывать модель от самой начальной временной точки, до конечной временной точки, параллельно выводя на печать результат во временных точках, которые заданы временным шагом в каждом из указанных интервалов.
- 2) Установка значений как параметров, так и переменных модели по шаблону. Такой механизм позволяет иметь именованные параметры и менять за один раз их значения во всех *элементарных подсистемах* модели, а также задавать начальные концентрации одних и тех же веществ в разных компартментах. Параметры вычислительного эксперимента задаются в

отдельном файле со всеми параметрами модели. Этот подход позволяет создавать несколько стартовых наборов параметров для одной и той же модели, симулируя исследования при разных условиях среды, например.

- 3) Благодаря особенностям архитектуры Eclipse RCP, реализована возможность интеграции дополнительных сценариев вычислительных экспериментов через подключаемые модули (Plug-In). По данному принципу были встроены инструменты интервального анализа (Kazantsev et al. 2008).

2.2.3 Принцип представления результатов моделирования

Результаты численного эксперимента на выходе представлены серией файлов со значением как всех переменных в заданных временных точках, так и только некоторого выбранного их подмножества в этих же точках. Такой встроенный фильтр помогает быстро отслеживать нужные переменные в серии вычислительных экспериментов, не выискивая их в матрице результатов. Новая сборка модели в общем случае может изменить порядок следования переменных в файлах. Т.е. новые матрицы результатов, полученные на тех же временных точках, могут быть не согласующимися с предыдущими сборками. Именно инструменты фильтрации, работающие со значениями *атрибутов* из *идентификаторов* переменных, помогают точно извлекать интересующие значения на всех модификациях анализируемых моделей.

2.3. Постобработка результатов моделирования

Язык моделирования SiBML позволяет не только строить серии моделей на основе одной базовой, добавляя/удаляя/модифицируя подсистемы, но и предоставляет большой простор для анализа результатов численного исследования в полуавтоматическом виде.

Результат численного эксперимента – это серия структурированных текстовых файлов, где каждая переменная модели представлена *идентификатором*. В *идентификаторе* заданы ряд *атрибутов* сущности: название, тип объекта, локализация в структуре компартментов и др. Используя стандартные методы и подходы обработки текстов (в рамках языков программирования или инженерных сред моделирования), можно разработать инструменты автоматической переработки таких данных, которые будут сортировать, упорядочивать и извлекать необходимые значения по заданным правилам.

Для модели, представленной в главе 3, на языке JAVA был написан инструмент по сортировке и упорядочению значений в соответствии с их локализацией в исследуемой структуре компартментов. Так как этот инструмент опирается только на *атрибуты* переменных, встраивание результатов в структурный шаблон происходит одинаково для всех версий модели, независимо от порядка переменных в выходном векторе данных. Внесение изменений в сами подсистемы модели и её дальнейшая сборка не влияют на инструменты постобработки, так как они опираются не на позиционную информацию переменной в векторе результатов, а на её семантику. Построенные упорядоченные и структурированные по шаблону данные результатов моделирования поступали на следующие этапы постобработки.

Эти возможности постобработки результатов моделирования в MGSmodeller были активно использованы в исследовании транспорта фитогормона ауксин в слоях протоксилемы и протофлоэмы корней растения *A. thaliana* (Novoselova et al. 2013). Модель представляла собой линейный массив клеток с компартментом апопласта между ними, где процессы синтеза, деградации и транспорта веществ были представлены всего из 11 элементарных подсистем.

В силу принципа конструирования математических моделей SiBML в среде MGSmodeller, итоговое распределение вектора переменных модели

может не совпадать для двух разных сборок. Это может случиться из-за внесённых изменений в структуру элементарных моделей или скриптов сборки. Для задачи извлечения серии необходимых концентраций переменных для нужных клеточных структур, был разработан инструментарий на языке JAVA, который сортировал, упорядочивал результаты вычислительных экспериментов и сохранял переменные в отдельных файлах. Эти файлы с уже упорядоченными данными легко визуализировались стандартными средствами (MS Excel, например).

Другой вариант постобработки результатов был сделан для распределения значений переменных по гексагональной сетке клеток корня (Novoselova et al. 2015). В этой работе исследовался транспорт гормона ауксин уже через поперечное сечение корня растения. Для более корректного соответствия модели процессам реального корня, структуру моделируемых клеток выбрали гексагональной (рис. 2.3.1).

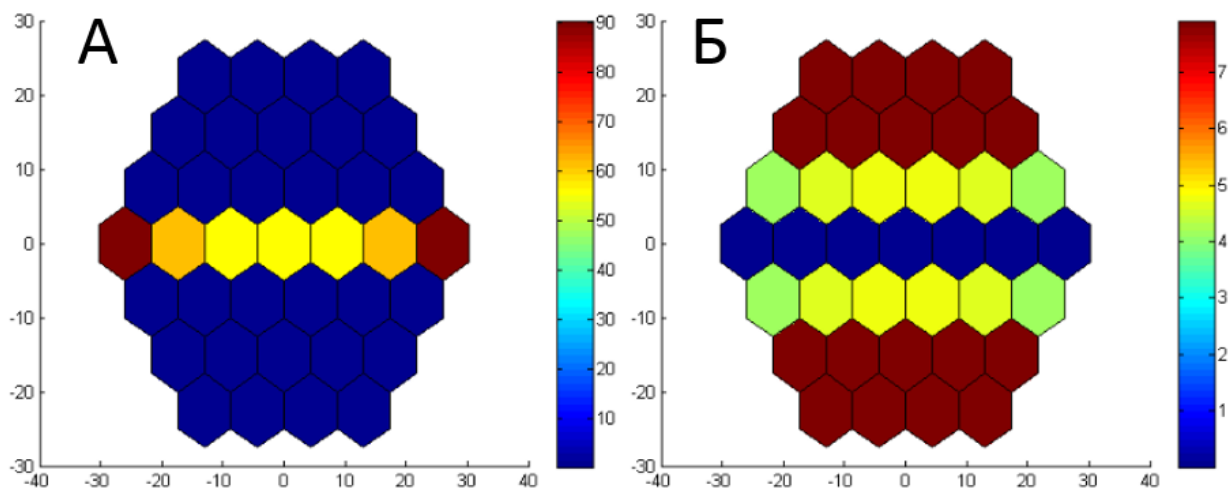


Рисунок 2.3.1 – Распределение переменных в стационаре на гексагональной сетке модели, полученные при заданной диархной симметрии поступающего потока ауксина из побега в корень. А) Распределение ауксина; Б) Распределение белка AUX1 (Novoselova et al. 2015).

Идентификаторы переменных модели в SiBML стандарте имели следующий вид: “<I(aux), str(c_i_j)>». Данный идентификатор переменной читается следующим образом: «I(aux)» - переменная, отвечающая за вещество ауксин; «str(c_i_j)» – переменная, отвечающая за вещество локализованное в клетке с индексом [i;j].

В этой модели была применена стандартная для гексагональных структур двумерная индексация ячеек (<https://www.redblobgames.com/grids/hexagons/>), которая была заложена в скрипты сборки модели. Разработанные инструментарии сортировки и упорядочивания переменных строили структурированные файлы, которые поступали скриптам построения гексагональных графиков в среде MATLAB.

2.3.1 Возможности экспорта моделей

На тот случай, когда средств языка SiBML недостаточно для проведения необходимого анализа модели, в среду моделирования MGSmodeller встроена серия программных модулей экспорта реконструированных математических моделей в следующие популярные среды моделирования: Matlab, Scilab, Octave, Mathematica (На момент написания диссертации доступен экспорт блоков SiBML задающих только непрерывные модели – системы ОДУ). Интеграция модулей была проведена по технологии «Plug-In» в ECLIPSE RCP. Использование богатого инструментария по анализу математических моделей инженерных сред моделирования позволило избежать временных затрат на повторную реализацию существующего функционала, и сосредоточится на решении содержательных задач. Интересной особенностью стал тот факт, что концепция именования переменных SiBML позволяет использовать уже реализованный функционал по сортировке, упорядочению и визуализации результатов численных экспериментов для данных полученных в сторонних средах моделирования.

Таким образом, концепции, заложенные в язык моделирования молекулярно-генетических систем SiBML и среду MGSmodeller, не ограничивают пользователя при решении задачи исследования моделей, но помогают системно подходить к вопросу реконструкции математической модели и к решению задач анализа результатов вычислительных экспериментов.

2.4. Модули высокопроизводительного анализа моделей МГС.

При реконструкции моделей комплексных МГС (модели масштабов в сотни штук компартментов) в рамках среды MGSmodeller, работа базовых программ и алгоритмов языка SiBML может составлять от нескольких минут до нескольких часов. Это относится как к процессам реконструкции моделей, так и к процессам их численного анализа. В частности, временные затраты на вычислительный эксперимент одной и той же модели могут сильно зависеть от исходных параметров. Или модель быстро выходит на стационарные значения (в SiBML используется метод Гира (Gear 1981) с адаптивным шагом интегрирования), или же долго обчисляет осциллирующую динамику модели. При втором режиме, в рамках одного персонального компьютера решение задачи оценки параметров модели может потребовать больших временных затрат.

Для разрешения этого вопроса комплекс программ среды MGSmodeller и инструментарий SiBML был перенесён и настроен на работу в рамках высокопроизводительного вычислительного кластера ЦКП «Биоинформатика» ССКЦ СО РАН (<http://pixie.bionet.nsc.ru/ratdna/ckp/newckp.php>). MGSmodeller был расширен средствами параллельной сборки математических моделей МГС, а также средствами постановки задач массового вычислительного эксперимента с варьированием параметров. Эти изменения позволили значительно (с

нескольких часов до десятка минут) сократить время сборки математической модели во время её реконструкции (Казанцев и др. 2012).

2.4.1 Параллельные вычислительные эксперименты

Решение задачи поиска значений параметров модели, приводящих решение модели к качественному соответствию экспериментально-наблюдаемым, может потребовать проведения нескольких десятков тысяч вычислительных экспериментов (и больше). В параллельном режиме на высокопроизводительном кластере на это потребуется всего несколько дней работы кластера по ночам. Использование вычислительного кластера позволяет сосредоточиться на сопоставлении и анализе данных моделирования, а не на их ожидании.

В частности, для решения задачи визуального анализа сотен результатов численного решения модели на кластере был создан ряд программных компонент постобработки данных по интеграции результатов вычислительного эксперимента, построению сводных таблиц, графиков и поверхностей решений. Была реализована визуализация полученных сводных данных с применением инструментов Matlab (mathworks.com) и Gnuplot (www.gnuplot.info).

2.4.2 Интеграция со сторонними программными средствами

В Политехническом Университете Санкт-Петербурга был разработан инструментарий по автоматическому подбору параметров модели – DEEP (Kozlov and Samsonov 2011). Данный инструментарий использует гибридный подход в задаче минимизации функционала – алгоритм градиентного спуска вместе с генетическими алгоритмами. Основой комплекса является модифицированный метод дифференциальной эволюции – стохастический

итеративный алгоритм многомерной математической оптимизации с применением идей генетических алгоритмов. В программном комплексе DEEP реализована многопоточная версия метода, обеспечивающая высокую скорость решения на суперкомпьютере.

Для работы этого инструментария необходимо специфицировать параметры, которые будут адаптироваться к искомому решению, и задать функционал для минимизации. Благодаря организации конструкций языка SiBML была реализована интеграция инструментария MGSmodeller и DEEP. Принципиальная схема решения представлена на рисунке 2.4.1.

На первом шаге метод DEEP генерирует фиксированное количество случайных векторов параметров как начальное условие в первом поколении «популяции». После этого проводятся вычислительные эксперименты модели с каждым из векторов параметров.

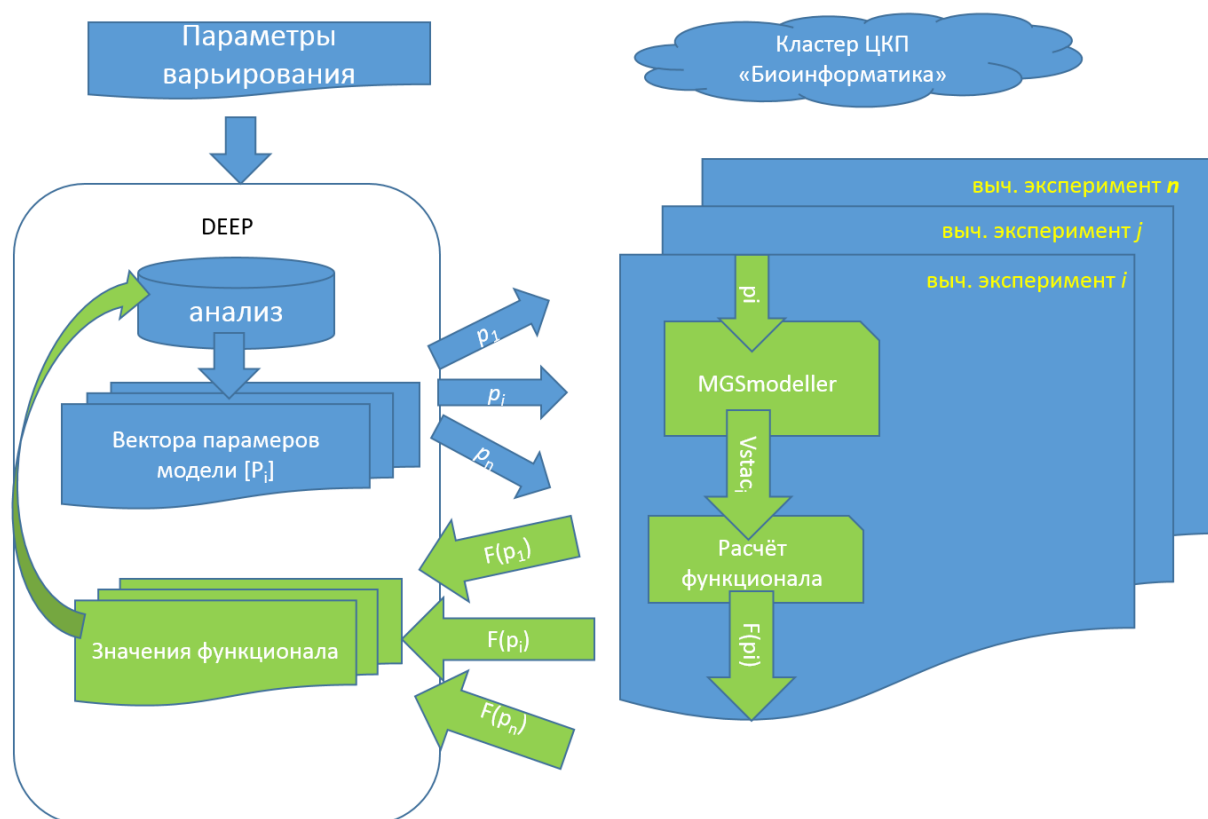


Рисунок. 2.4.1 – Схема интеграции инструментов DEEP и MGSmodeller. Стрелками указаны информационные потоки при решении задачи подбора набора параметров при которых минимизируется заданный пользователем функционал.

Для каждого результата расчётов вычисляется вес функционала этой «популяции». На каждом последующем шаге происходит «скрещивание» параметров векторов как аналог обмена генами между популяциями с получением новых векторов «популяций». Если новый вектор параметров имеет меньший вес функционала, чем родительский, то он заменяет родительский вектор в следующем поколении. Вычисления останавливаются при преодолении заданной границы функционала за фиксированное количество итераций.

2.4.3 Практика применения среды моделирования

Набор методов MGSmodeller, DEEP и суперкомпьютер был использован в реконструкции и исследовании интегрированной кинетической модели биосинтеза этанола, лактата и ацетата в клетке *Geobacillus spp.* (Нуриддинов и др. 2013). Модель включает 22 переменные, одна из которых характеризует рост биомассы, остальные уравнения описывают динамику изменения концентрации метаболитов. Модель содержит 127 параметров.

В результате параметрической адаптации математической модели методом DEEP к экспериментальным данным, полученные расчёты не только качественно отражают экспериментальную динамику изменения концентрации конечных продуктов (этанол, молочная кислота и ацетат), потребления глюкозы, но и количественно воспроизводят значения стационарных концентраций измеренных метаболитов. Разработанная математическая модель является мощным *in silico* инструментом для исследования режимов функционирования и метаболических потоков при «компьютерном создании» мутантных генотипов соответствующей бактерии.

Также системы MGSmodeller, DEEP и суперкомпьютер был применён в рамках исследования механизмов регуляции поддержания плюрипотентности эмбриональных стволовых клеток (ЭСК) (Акбердин и др. 2014b). Была

реконструирована генная сеть включающая транскрипционные факторы, которые контролируют основные молекулярно-генетические механизмы поддержания плюрипотентности ЭСК: Oct, Sox2 и Nanog. Эти составляют так называемую «ключевую генную сеть» плюрипотентности (Jaenisch and Young 2008). На модели исследовалась наблюдаемая в популяции ЭСК различной степени гетерогенность в экспрессии генов как плюрипотентности, так и дифференцировки. Структура модели была взята из работы Чикармане с соавторами (Chickarmane et al. 2006) и дополнена новыми данными. Также в модели были учтены процессы транскрипции и трансляции для «ключевых» транскрипционных факторов в мЭСК.

Модель содержит 12 переменных и 27 параметров. Константы деградации белков и их мРНК были взяты из литературных данных. Значения остальных параметров были подобраны при адаптации модели к экспериментальным данным на основе качественного анализа переходов из состояния плюрипотентности в дифференцированное и обратно. Численные эксперименты проводились на высокопроизводительном вычислительном кластере ЦКП «Биоинформатика» ССКЦ СО РАН. Анализ математической модели, построенной на этих данных, показал, что в динамике изменения концентрации как мРНК, так и белка Nanog потенциально может существовать осциллирующий режим функционирования, что может объяснять наблюдаемую в экспериментах гетерогенность экспрессии Nanog в популяции столовых клеток, поскольку в один и тот же момент времени в различных ЭСК в культуре может наблюдаться различная концентрация этого фактора.

2.5. Заключение к главе 2

Работы по совершенствованию инструментария моделирования SiBML в рамках MGSmodeller проходили по следующим пунктам:

- Формально сформулирована грамматика языка моделирования, опираясь на которую был реализован функционал синтаксического и грамматического анализа конструкций SiBML.
- Разработан программный комплекс MGSmodeller как среда моделирования в терминах языка SiBML. Среда моделирования включает графический интерфейс пользователя с подсветкой синтаксиса и подсказками по грамматике языка, а также средства постановки численного эксперимента и простые методы визуализации результатов.
- Для работы на суперкомпьютере ЦКП «Биоинформатика» ССКЦ СО РАН была проведена модернизация программного комплекса MGSmodeller и добавлены инструменты высокопроизводительного анализа.

Программный комплекс MGSmodeller/SiBML с расширенной функциональностью показал хорошие результаты при решении задач 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. 2014a); метаболизма гормона ауксина в клетке меристемы побега (Акбердин и др. 2009b) и транспортных потоков ауксина в ансамбле 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). Предложенная в представленных работах схема проведения вычислительных экспериментов может быть повторно использована для проведения исследований по моделированию в подобных задачах.

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

С использованием разработанных инструментариев SiBML и MGSmodeller было проведено исследование особенностей процесса распределения фитогормона ауксина в меристеме корня растения *A. thaliana*. Ауксин является одним из важнейших фитогормонов растений, отвечающий за регуляцию процессов деления, роста и дифференцировки клеток (Vanneste and Friml 2009). В апикальной меристеме корня *A. thaliana* ауксин синтезируется в небольших количествах. Основная часть ауксина синтезируется в молодых листьях и меристемах побегов и транспортируется в корень, способствуя его развитию. Поток из побега идёт по сосудистой ткани, но также имеет место диффузия в тканях. В меристеме корня ауксин распределяется, формируя максимум в зоне ниши стволовых клеток, а именно клетках покоящегося центра и инициалах. Ниша стволовых клеток (НСК) является источником клеток для всего корня (рис. 3.1.1). От сбалансированной активности НСК зависит будет ли нормально сформирован корень и как он будет расти.



Рисунок 3.1.1 – Анатомическая структура кончика корня *A. thaliana* и распределение ауксина в нем. а – клеточные типы в меристеме корня (Коврижных и др. 2014).; б – распределение ауксина, выявленное с помощью сенсора DR5::GUS (Pasternak et al. 2019). Шкала 30 мкм.

Известно, что распределение ауксина в корне с максимумом в покоем центре является необходимым условием поддержания и функционирования НСК (Коврижных и др. 2014). Важно, что это распределение формируется динамически, за счёт процессов перераспределения ауксина между клетками с помощью белков-транспортёров, расположенных на плазматической мембране. Основной вклад в транспорт ауксина в меристеме корня вносят белки семейства PIN (Bennett and Scheres 2010). В геноме *A. thaliana* закодировано 8 белков PIN семейства, из которых белки PIN1,2,3,4 и PIN7 располагаются на плазматической мембране (Feraru and Friml 2008). Так как PIN белки располагаются на определённой части мембраны клетки (полярно), и активно выводят ауксин из клетки, они играют существенную роль в полярном транспорте ауксина и формировании неравномерных распределений ауксина (Feraru et al. 2011). Результатом перераспределения ауксина с помощью PIN белков является наблюдаемый паттерн распределения концентрации ауксина с максимумом в районе НСК (рис. 3.1.2).

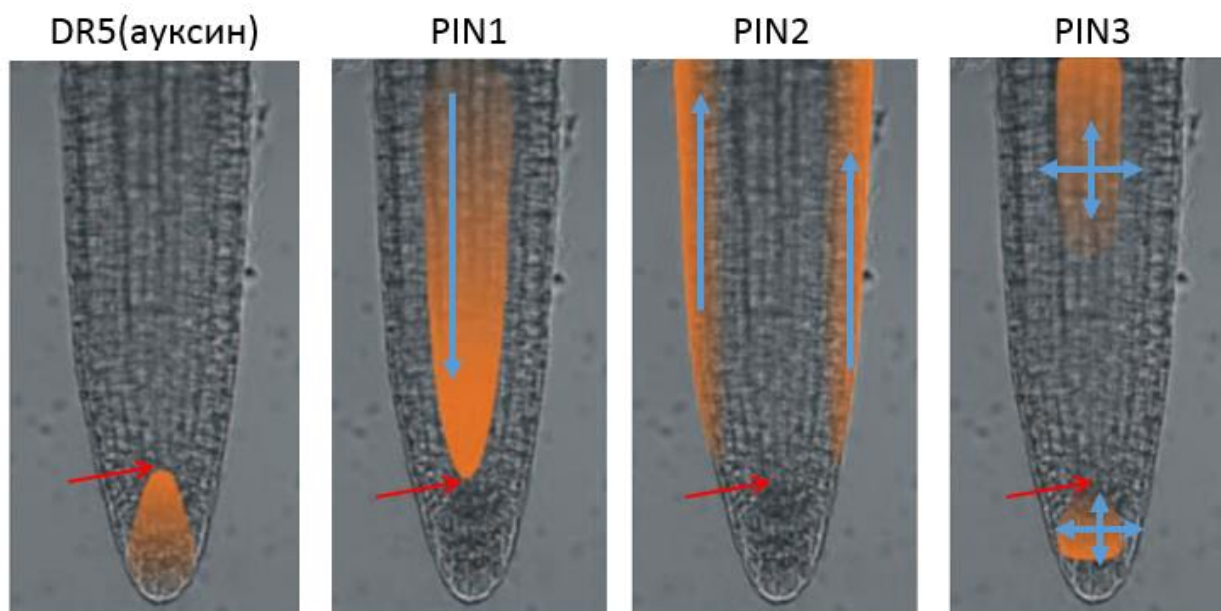


Рисунок 3.1.2 – Схематическое представление паттернов экспрессии белков транспортеров PIN и активности ауксин-чувствительного сенсора DR5 на микрофотографии корня (красной стрелкой отмечено место покоем центра, синими стрелками показано направление транспорта ауксина) (Mironova et al. 2012).

В то время как PIN белки контролируют внутриклеточную концентрацию ауксина, активно выводя его из клетки, ауксин тоже может контролировать концентрацию PIN белков на мембране. Ауксин регулирует экспрессию своих транспортёров на транскрипционном и посттранскрипционном уровнях (Vieten et al. 2005; Petrasek 2006; Teale et al. 2006; Tanaka et al. 2006). Отмечена концентрационно-зависимая регуляция ауксином PIN белков, например, при низкой концентрации ауксина активирует экспрессию PIN1, а при высокой ингибирует её (Omelyanchuk et al. 2016).

В рамках квалификационной работы представлен ответ на вопрос: «Способен ли ауксин сформировать характерное распределение белков транспортёров PIN1, PIN2, PIN3 в кончике корня, приводящее к самоорганизации своего паттерна распределения?».

Вопросами потоков ауксина в структурах корня занимаются несколько научных групп (Stoma et al. 2008; Lucas et al. 2011; Band et al. 2012; Band et al. 2014; Antonovici et al. 2015). Ниже рассмотрены наиболее соответствующие задаче модели.

3.1.1 Модель «обратного фонтана»

В работе Гринайсен и коллег (Grieneisen et al. 2007) был проведён *in silico* и *in vivo* анализ распределения ауксина в корне *A. thaliana*. Математическая модель корня представлена в терминах дифференциальных уравнений в частных производных. Рассмотрена статичная структура корня с фиксированной локализацией белков транспортёров, названная структурой «обратного фонтана». Авторами показано, что структура расположения PIN белков по типу «обратного фонтана» способна формировать максимум концентрации ауксина в области НСК. Максимум концентрации ауксина сам поддерживался длительное время при отсутствии внешнего потока ауксина (из побега в корень).

В модели были сделаны следующие допущения:

- Клетки имеют размеры сопоставимые с размерами клеток в соответствующих зонах корня.
- Активный транспорт ауксина через мембраны задан статичным распределением белков-транспортёров. Их локализация на мембранах клеток была фиксирована с учётом экспериментальных данных. Концентрация PIN белков не менялась в процессе расчёта модели.
- Специализация белков транспортёров отсутствует. Т.е. в модели присутствует некий обобщённый белок транспортёр, который осуществляет процесс транспорта ауксина только через ту мембрану, на которой закреплён.

В этой модели ауксин приходит в модель извне в сосудистые клетки корня. Затем, по этой сосудистой ткани, транспортируется к кончику корня (однонаправленно вниз, белки только на нижней мембране). В кончике корня ауксин распределяется по всем клеткам равномерно (во всех направлениях, белки равномерно распределены по всем мембранам клеток). А уже через клетки ткани эпидермиса и кортекса транспортируются обратно вверх, в сторону к побегу (белки на верхней мембране и мембране, ориентированной к центру корня).

Несмотря на то, что предложенная модель показывает формирование максимума концентрации ауксина в НСК, модель не объясняет процессов организации потоков ауксина в корне, и не даёт понять, как сформировались домены локализации каждого из белков транспортёров. Хотя в рамках своего исследования авторы попытались объяснить изменения в формировании пика максимума концентрации ауксина (смещение пика максимума к кончику корня и падение концентрации в нем) в тройном мутанте *pin3pin4pin7 A. thaliana*.

3.1.2 Модель «отражённая волна»

В ИЦиГ СО РАН вышла серия работ по исследованию процессов распределения ауксина по структуре клеток корня под руководством Лихошвая В.А. (Лихошвай и др. 2007; Фадеев и др. 2008; Mironova et al. 2010). В работах (Лихошвай и др. 2007; Фадеев и др. 2008) на одномерной модели был исследован механизм ауксин-зависимого транспорта ауксина по сосудистой ткани. В работах был введён механизм авторегуляции транспорта ауксина в зависимости от собственной концентрации. Предложенный механизм с положительной и отрицательной регуляциями воспроизводит феномен формирования максимума в нише ствольных клеток корня (НСК). Авторами был проведён подробный анализ параметров модели и найдены характерные стационарные решения: распределение ауксина с одним максимумом в зоне покоящегося центра, смещение точки максимума относительно НСК, множественные максимумы.

В работе (Mironova et al. 2010) этот механизм был исследован более подробно, с учётом конкретного белка транспортёра PIN1, биосинтез которого был задан ауксин-зависимым. В среде MGSmodeller авторы разработали серию математических моделей (в терминах обыкновенных дифференциальных уравнений) распределения ауксина в клетках корня при участии белка транспортёра PIN1 (Автор данной диссертации выступал в качестве технического консультанта). В отличие от модели «обратный фонтан», в модели введён ауксин-зависимый синтез белка PIN1 (осуществляет однонаправленный транспорт ауксина от побега к кончику корня). В рамках одномерной и двумерной моделей была показана самоорганизация паттернов экспрессии PIN1 и распределения ауксина, соответствующие экспериментальным данным. Было показано что даже один белок транспортёр PIN1, на биосинтез которого влияет ауксин (а в его отсутствии не синтезируется), способен сформировать в двумерной модели пик концентрации ауксина в зоне покоящегося центра корня. В работе Мироновой

и соавторов этот механизм распределения ауксина в корне получил название «отражённая волна». Параметры модели были подобраны методами градиентного спуска для соответствия стационара модели кривой, взятой из интенсивности свечения сигнальной конструкции DR5::GFP, чувствительной к концентрации ауксина, измеренной в эксперименте в кончике корня.

Данная модель дополняет результаты, полученные в работах (Лихошвай и др. 2007), знаниями о двух различных поведении модели в зависимости от эффективности концентрационно-зависимого ингибирования экспрессии PIN1, которые могут объяснять формирование стержневой и мочковатой корневых систем. Кроме того, в работе (Mironova et al. 2010) создали одномерную модель растущего корня, на которой показали самоорганизацию формирования максимума ауксина в зоне ниши стволовых клеток корня и воспроизвели феномен восстановления максимума концентрации ауксина (восстановление функции роста корня) в случае нарушения целостности его структуры.

3.1.3 Формальная постановка задачи моделирования

Приведённые модели решают поставленные перед ними задачи:

- Распределение белков транспортёров с чёткой локализацией на мембране формирует по механизму “обратного фонтана” паттерн распределения концентрации ауксина с максимумом в зоне покоящегося центра.
- Ауксин-зависимый биосинтез и деградация белка транспортёра PIN1 по механизму “отражённая волна” способны формировать распределение концентрации ауксина с максимумом в зоне покоящегося центра моделях с 1D, 2D структурами, и даже на 1D модели с учётом процессов роста и деления клеток.

В приведённых моделях остались неотвеченными следующие вопросы: Какова функция и вклад в общую картину каждого белка из семейства PIN (мембранные белки транспортёры: PIN1, PIN2, PIN3, PIN4 и PIN7)? Каким образом формируется наблюдаемое распределение этих белков транспортёров? Важны ли процессы биосинтеза ауксина в НСК? Какие процессы являются критичными при формировании максимума концентрации фитогормона ауксина в НСК?

В рамках данной диссертации с этой отправной точки продолжилось исследование, где методами математического и компьютерного моделирования в единой модели корня с двумерной структурой проведена интеграция описанных двух подходов (механизм «обратного фонтана» (Grieneisen et al. 2007), и механизм «отражённой волны» (Mironova et al. 2010)). В получившейся интегральной модели дополнительно учтены процессы ауксин-зависимого биосинтеза белков транспортёров PIN2 и PIN3 (в модели подразумевается интегральная активность белков PIN3, PIN4 и PIN7). Модель позволила проследить за динамикой изменения не только фитогормона в клетках, но и за экспрессией PIN белков в них.

Именно в этом исследовании раскрылись заложенные в SiBML и системе MGSmodeller возможности, методы и алгоритмы, включая поддержку расчётов на высокопроизводительном вычислительном кластере ЦКП «Биоинформатика» ССКЦ СО РАН.

В работе использовались экспериментальные данные, предоставленные коллегами из Университета Фрайбурга. В качестве экспериментальных объектов выступают корни проростков *A. thaliana* 3-5 дней после прорастания дикого типа и линии трансгенных растений DR5::GFP, PIN::PIN-GFP и другие.

3.2. Модель меристемы корня «ММК»

В модели рассмотрены три белка семейства PIN: PIN1, PIN2 и PIN3. Причём под белком PIN3 (см. рис. 3.1.2) будет подразумеваться интегральная активность белков PIN3, PIN4 и PIN7, так как домены их экспрессии пересекаются, и эти белки имеют схожую локализацию на мембране клеток (осуществляют не полярный транспорт). При решении задачи поиска параметров в качестве источника фактических данных по локализации и активности белков транспортёров использовались данные из работы (Vieten et al. 2005). В этой работе проведены серия экспериментов по обработке растений разными концентрациями ауксина и анализ экспрессии PIN транспортёров в них. В частности, в работе экспериментально подтверждено влияние ауксина на экспрессию генов, кодирующих белки PIN1, PIN2 и PIN3. В работе также показан наибольший вклад в регуляцию внутриклеточной концентрации ауксина процессов его транспорта, чем процессов его биосинтеза.

Если для белка PIN1 была использована готовая модель ауксин зависимой экспрессии (Mironova et al. 2010), то для белков PIN2 и PIN3 необходимо было сформулировать эти модели. При создании моделей подсистем биосинтеза и деградации PIN2 и PIN3 опирались на экспериментальные данные их локализации (см. рис. 3.1.2). Предполагается что разная концентрация ауксина в клетке активирует разных представителей PIN семейства. Так, PIN2 экспрессируется в клетках с низким содержанием ауксина; PIN1 в клетках со средними концентрациями ауксина; PIN3 в клетках с самым высоким содержанием ауксина. Приняв во внимание экспериментально подтверждённые данные о влиянии концентрации ауксина на процессы биосинтеза белков транспортёров была предложена гипотеза о характере экспрессии белков транспортёров (рис. 3.2.1 В). Т. е. в модели для каждого из PIN белков решено задать ауксин-зависимый биосинтез, параметры которого подбирались исходя из экспериментальных данных (Vieten et al. 2005).

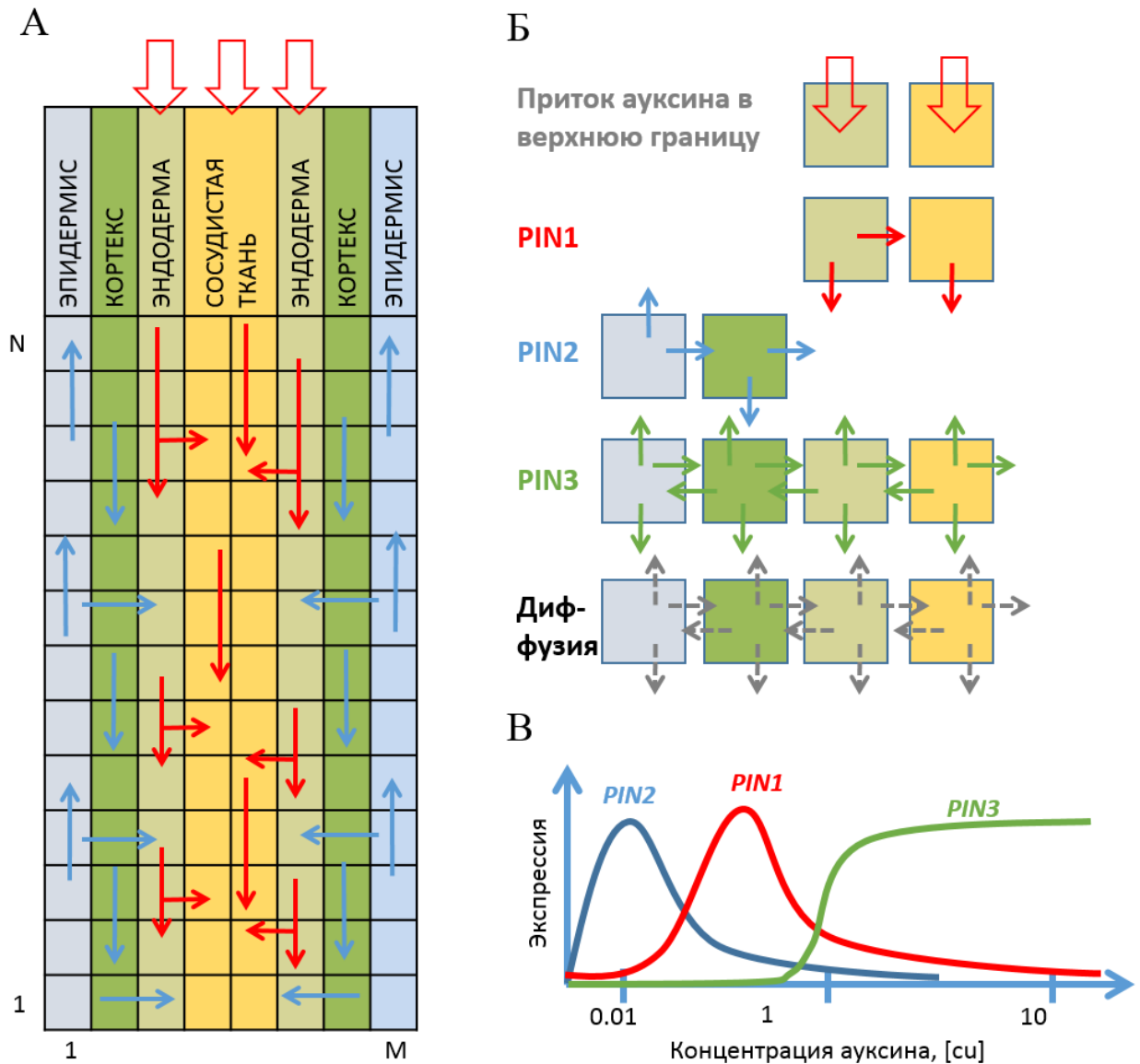


Рисунок 3.2.1 – Схема модели ММК. А) Структура моделируемого клеточного ансамбля с направлениями потоков ауксина. Б) схема мест синтеза белков транспортёров и определяющее ими направление транспорта ауксина. Цветом задана принадлежность к слоям клеток из пункта А. В) профиль экспрессии белков транспортёров в зависимости от концентрации ауксина в клетке.

В качестве математической модели для предложенного вида кривых были выбраны функции Хилла (Likhoshvai and Ratushnyi 2007), как наиболее подходящие. Причём деградация является ауксин-зависимой только для белков PIN1-2, а PIN3 деградирует без участия ауксина (рис. 3.2.1 В). На рисунке 3.2.1 показана структура модели с областью моделирования размером 8x20 клеток, которая соответствует продольному срезу корня – 8 слоёв тканей

корня в 20 рядов клеток. Математическая модель была построена на базе следующих установок:

- Переменными в модели являются концентрации ауксина и белков транспортёров (PIN1, PIN2 и PIN3) в каждой из клеток моделируемой области.
- Математическая модель, с точки зрения структуры, представляет собой организованный массив клеток с единственной точкой притока ауксина в систему – в верхние клетки сосудистой ткани.
- В каждой клетке происходят независимые от соседних клеток процессы синтеза белков транспортёров.
- Направление активного транспорта (локализация белков) между модельными клетками задаётся белками транспортёрами (см. ниже). Сам транспорт осуществляется при непосредственном участии белков транспортёров.
- В модели делается допущение, что все синтезируемые белки транспортёры моментально встраиваются в мембрану и начинают выполнять свою функцию.
- Помимо активного транспорта ауксина в модели присутствует пассивный транспорт ауксина, соответствующий процессам диффузии ауксина между клетками.
- Размер моделируемой области 8x20 клеток (ширина, длина). Моделируется зона меристемы корня, зона где клетки ещё не начали увеличивать свои линейные размеры.
- фитогормон ауксин поступает в модель извне в клетки сосудистой ткани (сосудистый слой и эндодерма) на границе (последний ряд 3 - 6 слоя клеток).
- Белок PIN2 экспрессируется только в клетках эпидермиса и кортекса (1,2,7 и 8 слой). Он транспортирует ауксин по направлению к побегу и к сосудистым клеткам (вверх и по направлению к центральной оси корня).

- Белок PIN1 экспрессируется только в сосудистых клетках и клетках эндодермы (3 - 6 слой). Он транспортирует ауксин по направлению к кончику корня (вниз).
- Белок PIN3 экспрессируется во всех клетках корня и осуществляет транспорт ауксина во всех направлениях одинаково.

В модели рассматриваются следующие элементарные процессы (в терминах формализма SiBML):

1. Процессы пассивного транспорта (диффузии) ауксина между клетками.
2. Процессы деградации ауксина в каждой клетке.
3. Процессы притока ауксина в систему (из побега в зону центрального цилиндра).
4. Процессы ауксин-зависимого биосинтеза белков PIN1 - PIN3.
5. Процессы ауксин-зависимой деградации белков PIN1 – PIN2.
6. Процессы ауксин-независимой деградации белков PIN3 (Vieten et al. 2005).
7. Процессы направленного активного мембранного транспорта ауксина посредством белков транспортёров семейства PIN.

В ходе реконструкции математической модели в виде системы ОДУ были применены средства автоматизации: язык моделирования SiBML и среда разработки MGSmodeller. Подходы, заложенные в данных решениях, позволили, описав 12 элементарных подсистем, формирующих процессы одной клетки, реконструировать математическую модель, соответствующую ансамблю клеток размером 8x20, и серии моделей размерностью 8x100, 8x50, 8x35, 8x15, 8x10 и 8x5 клеток, которые использовались для проверки ряда гипотез функционирования модели и для исследования её устойчивости.

3.2.1 Модели элементарных подсистем ММК

В рамках языка SiBML в среде MGSmodeller была описана серия элементарных подсистем (размерности далее: cu - концентрационные единицы, tu – временные единицы, dl – безразмерный параметр):

1. Ауксин-зависимый биосинтез белка PIN1 в клетке с индексом (i,j) со скоростью:

$$k_{pin1}^S * \frac{\left(\frac{Auxin_{ij}}{q_{pin1}^1}\right)^{S_{pin1}}}{\left(1 + \frac{Auxin_{ij}}{q_{pin1}^2}\right)^{S_{pin1}}},$$

где: $k_{pin1}^S (dl)$ – параметр скорости синтеза белка PIN1; $S_{pin1} (dl)$ – коэффициент Хилла, который определяет скорость ответа реакции синтеза белка PIN1 на изменения внутриклеточной концентрации ауксина; $q_{pin1}^1 (cu)$ – порог ауксин-зависимой активации синтеза; $q_{pin1}^2 (cu)$ – порог насыщения ауксин-зависимого синтеза.

2. Ауксин-зависимый биосинтез белка PIN2 в клетке с индексом (i,j) со скоростью:

$$k_{pin2}^S * \frac{\left(\frac{Auxin_{ij}}{q_{pin2}^1}\right)^{S_{pin2}}}{\left(1 + \frac{Auxin_{ij}}{q_{pin2}^2}\right)^{S_{pin2}}},$$

где: $k_{pin2}^S (dl)$ – параметр скорости синтеза белка PIN2; $S_{pin2} (dl)$ – коэффициент Хилла, который определяет скорость ответа реакции синтеза белка PIN2 на изменения внутриклеточной концентрации ауксина; $q_{pin2}^1 (cu)$ – порог ауксин-зависимой активации синтеза; $q_{pin2}^2 (cu)$ – порог насыщения ауксин-зависимого синтеза.

3. Ауксин-зависимый биосинтез белка PIN3 в клетке с индексом (i,j) со скоростью:

$$k_{pin3}^S * \frac{\left(\frac{Auxin_{ij}}{q_{pin3}^1}\right)^{S_{pin3}}}{\left(1 + \frac{Auxin_{ij}}{q_{pin3}^2}\right)^{S_{pin3}}},$$

где: k_{pin3}^S (dl) – параметр скорости синтеза белка PIN3; S_{pin3} (dl) – коэффициент Хилла, который определяет скорость ответа реакции синтеза белка PIN3 на изменения внутриклеточной концентрации ауксина; q_{pin3}^1 (cu) – порог ауксин-зависимой активации синтеза; q_{pin3}^2 (cu) – порог насыщения ауксин-зависимого синтеза.

4. Ауксин-зависимая деградация белка PIN1 в клетке с индексом (i,j) со скоростью:

$$k_{pin1}^d * Pin_1 * \left(1 + \frac{Auxin_{ij}}{q_{pin1}^3}\right)^{h_{pin1}},$$

где: k_{pin1}^d (l/tu) – параметр скорости деградации белка PIN1; h_{pin1} (dl) – коэффициент, который определяет эффективность ауксин-зависимой деградации PIN1; q_{pin1}^3 (cu) – порог ауксин-зависимой деградации PIN1.

5. Ауксин-зависимая деградация белка PIN2 в клетке с индексом (i,j) со скоростью:

$$k_{pin2}^d * Pin_2 * \left(1 + \frac{Auxin_{ij}}{q_{pin2}^3}\right)^{h_{pin2}},$$

где: k_{pin2}^d (l/tu) – параметр скорости деградации белка PIN2; h_{pin2} (dl) – коэффициент, который определяет эффективность ауксин-зависимой деградации PIN2; q_{pin2}^3 (cu) – порог ауксин-зависимой деградации PIN2.

6. Деградация белка Pin3 в клетке с индексом (i,j) со скоростью:

$$k_{pin3}^d * Pin_{zij},$$

где k_{pin3}^d (l/tu) – параметр скорости деградации белка PIN3.

7. Пассивный транспорт (диффузия) ауксина через конкретную мембрану, соединяющую две соседние клетки со скоростью:

$$D * Auxin_{ij} ,$$

где D (l/tu) – параметр диффузии через мембрану. Полная диффузия из клетки $4*D*Aux_{ij}$.

8. Деградация ауксина со скоростью:

$$k_{auxin}^d * Auxin_{ij} ,$$

где: k_{auxin}^d (l/tu) – параметр скорости деградации ауксина.

9. Приток ауксина в модель со скоростью:

$$k_{alpha} ,$$

где: k_{alpha} (cu/tu) – интенсивность потока ауксина в клетки сосудистой ткани и эндодермы на границе зоны меристемы корня (приток в модель).

10. Активный транспорт ауксина посредством PIN1 со скоростью:

$$k_{pin1}^0 * Pin1_{ij} * Auxin_{ij} ,$$

где: k_{pin1}^0 (l/tu) – параметр скорости активного транспорта посредством PIN1 по направлению к кончику корня (вниз).

11. Активный транспорт ауксина посредством PIN2 со скоростью:

$$(k_{pin2}^0 + k_{pin2}^{0,lat}) * Pin2_{ij} * Auxin_{ij} ,$$

где: k_{pin2}^0 (l/tu) – параметр скорости активного транспорта посредством PIN2 по направлению вверх в клетках эпидермиса и кортекса; $k_{pin2}^{0,lat}$ (l/tu) – параметр скорости активного транспорта посредством PIN2 по направлению к центральной оси корня. Причём $(k_{pin2}^0 + k_{pin2}^{0,lat}) = 1$. Эти два параметра контролируют характер деления потока ауксина.

12. Активный транспорт ауксина посредством PIN3 со скоростью:

$$k_{pin3}^0 * Auxin_{ij} * Pin3_{ij} ,$$

где k_{pin3}^0 (l/tu) – параметр скорости активного транспорта посредством PIN3 по всем направлениям.

Параметры экспрессии белков транспортёров были заданы на основе сопоставления экспериментально наблюдаемого распределения ауксина с паттернами экспрессии PIN белков в меристеме корня и с учётом влияния ауксина на экспрессию PIN (Vieten et al. 2005) (см. рис. 3.1.2).

3.2.2 Математическое описание модели ММК

На основе моделей элементарных подсистем средствами среды MGSmodeller была собрана математическая модель. Причём, с целью уменьшения затрат вычислительных ресурсов, моделировалась только половина корня (относительно центральной продольной оси). Так, основной размер модели стал не 8x20, а 4x20 клеток (далее, для ясности изложения, обозначение моделей будет задано через название и размеры клеточного ансамбля. Модель, построенная на правилах ММК с моделируемой областью 4x20 клеток, будет имеет обозначение – «ММК:4x20»). При этом подразумевается абсолютная симметрия процессов, с учётом граничных условий по линии симметрии. Итоговый размер математической модели «ММК:4x20» получился в 240 уравнений и 30 параметров (Mironova et al. 2012).

Для лучшего восприятия модели приведём ряд пояснений. Индексы $[i;j]$ обозначают принадлежность выбранной переменной клетке модели с таким же индексом (индексация начинается с левого нижнего края моделируемой области. См. рис. 3.2.1 А). Множество \mathbb{N}_{ij} представляет собой набор индексов существующих соседних клеток по отношению к клетке с индексом $[i;j]$. Соответственно параметр $k_{neighbor_{ij}}$ является *мощностью* (количество элементов множества) множества \mathbb{N}_{ij} . Таким образом, обобщённая модель «ММК» в виде системы дифференциальных уравнений имеет следующий вид:

$$\begin{cases} \frac{dPin_{1ij}}{dt} = Fs_{1ij}(Auxin_{ij}) - Fd_{1ij}(Auxin_{ij}) \\ \frac{dPin_{2ij}}{dt} = Fs_{2ij}(Auxin_{ij}) - Fd_{2ij}(Auxin_{ij}) \\ \frac{dPin_{3ij}}{dt} = Fs_{3ij}(Auxin_{ij}) - Fd_{3ij}(Auxin_{ij}) \\ \frac{dAuxin_{ij}}{dt} = -k_{Auxin}^d * Auxin_{ij} + J_{alpha}(i,j) + D * \sum_{g \in \mathbb{N}_{ij}} (Auxin_g - Auxin_{ij}) + J_{Pin1}(i,j) + J_{Pin2}(i,j) + J_{Pin3}(i,j) \end{cases}$$

Fs_{ij} отвечает за скорость синтеза, а Fd_{ij} за скорость деградации соответствующих белков транспортеров PIN семейства в клетках с индексами $[i;j]$. Формула скорости синтеза для белков PIN1-3 имеют одинаковую структуру:

$$Fs_{\#ij}(Auxin_{ij}) = k_{pin\#}^S * \frac{\left(\frac{Auxin_{ij}}{q_{pin\#}^1}\right)^{S_{pin\#}}}{\left(1 + \frac{Auxin_{ij}}{q_{pin\#}^2}\right)^{S_{pin\#}}}$$

(Здесь и далее, вместо символа # подставляется номер белка транспортера: для белка PIN1 это будет 1; для PIN2 – 2; а для PIN3 – 3).

Уравнения деградации белков транспортеров PIN1 и PIN2 также имеют похожую структуру

$$Fd_{\#ij}(Auxin_{ij}) = k_{pin\#}^d * Pin_{\#} * \left(1 + \frac{Auxin_{ij}}{q_{pin\#}^3}\right)^{h_{pin\#}}$$

Однако, деградация белка PIN3 не зависит от концентрации фитогормона ауксин:

$$Fd_{3ij}(Auxin_{ij}) = k_{pin3}^d * Pin_{3ij}$$

Поведение фитогормона ауксин имеет более сложное уравнение:

$$\begin{aligned} \frac{dAuxin_{ij}}{dt} = & k_{Auxin}^{syn} - k_{Auxin}^d * Auxin_{ij} + J_{alpha}(i,j) + \\ & D * \sum_{g \in \mathbb{N}_{ij}} (Auxin_g - Auxin_{ij}) + J_{Pin1}(i,j) + J_{Pin2}(i,j) + J_{Pin3}(i,j), \end{aligned}$$

где

$$J_{alpha}(i,j) = \begin{cases} k_{alpha}, & \text{если } i = 20 \text{ и } j = [3; 4] \\ 0, & \text{иначе} \end{cases}$$

$$J_{Pin1}(i,j) =$$

$$\begin{cases} k_{pin1}^0 * (Pin1_{i+1j} * Auxin_{i+1j} - Pin1_{ij} * Auxin_{ij}), \text{ если } 1 < i < 20 \\ -k_{pin1}^0 * Pin1_{ij} * Auxin_{ij}, \text{ если } i = 20 \\ k_{pin1}^0 * Pin1_{i+1j} * Auxin_{i+1j}, \text{ если } i = 1 \end{cases}$$

$$J_{Pin2}(i, j) = J_{Apical}^{Pin2}(i, j) + J_{Lateral}^{Pin2}(i, j)$$

$$J_{Apical}^{Pin2}(i, j) =$$

$$\begin{cases} -k_{pin2}^{0,apic} * k_{pin2}^0 * Pin2_{ij} * Auxin_{ij}, \text{ если } i = 1 \text{ и } (j = 1 \text{ or } j = 2) \\ k_{pin2}^{0,apic} * k_{pin2}^0 * \begin{pmatrix} Pin2_{i-1j} * Auxin_{i-1j} \\ Pin2_{ij} * Auxin_{ij} \end{pmatrix}, \text{ если } i > 1 \text{ и } (j = 1 \text{ или } j = 2) \end{cases}$$

$$J_{Lateral}^{Pin2}(i, j) =$$

$$\begin{cases} -k_{pin2}^{0,lat} * k_{pin2}^0 * Pin2_{ij} * Auxin_{ij}, \text{ if } j = 1 \\ k_{pin2}^{0,lat} * k_{pin2}^0 * (Pin2_{ij-1} * Auxin_{ij-1} - Pin2_{ij} * Auxin_{ij}), \text{ если } j = 2 \\ k_{pin2}^{0,lat} * k_{pin2}^0 * Pin2_{ij-1} * Auxin_{ij-1}, \text{ if } j = 3 \end{cases}$$

$$J_{Pin3}(i, j) =$$

$$k_{pin3}^0 \left(\begin{array}{c} \sum_{g \in \mathbb{N}_{ij}} \left(\frac{1}{k_{neighbor_g}} * Auxin_g * Pin3_g \right) - \\ \left\{ \begin{array}{ll} \left(\frac{2}{3} \right) * Auxin_{ij} * Pin3_{ij}, & \text{если } j = 4 \text{ и } (1 = i) \\ \left(\frac{3}{4} \right) * Auxin_{ij} * Pin3_{ij}, & \text{если } j = 4 \text{ и } (1 < i < 20) \\ Auxin_{ij} * Pin3_{ij}, & \text{в других случаях} \end{array} \right\} \end{array} \right)$$

$$k_{neighbor_{ij}} = \begin{cases} 4, & \text{если клетка } cell_{ij} \text{ имеет 4 соседних клетки} \\ 3, & \text{если клетка } cell_{ij} \text{ имеет 3 соседних клетки} \\ 2, & \text{если клетка } cell_{ij} \text{ имеет 2 соседних клетки} \end{cases}$$

В следствии граничных условий и при учёте потоков от второй половины корня, в уравнениях возникают коэффициенты 1/2, 1/3, 1/4, 2/3, 3/4.

Для удобства восприятия и сравнения параметров моделей ММК, они были оформлены в табличном виде. В таблице 3.2.1 представлена схема

размещения параметров. Читать эту таблицу следует следующим образом: в верхней части таблицы размещены параметры, относящиеся к ауксину ($kAlpha$ – приток ауксина в модель, D – диффузия ауксина, kd – скорость деградации ауксина) и название модели, ассоциированной с этими параметрами. Ниже приведены параметры, относящиеся к белкам транспортёрам. Серым выделены формирующие параметр блоки. Например, значение X в ячейке строки PIN2 и столбца $k0$ следует читать как: «значение параметра $k^0_{pin2}=X$ ».

Таблица 3.2.1 – Шаблон для чтения параметров модели ММК.

	$kAlpha$	D	kd	Модель ММК. Шаблон чтения параметров модели.					
Auxin	$kAlpha$	D	kd						
	$k0$	$k0_{lat}$	kS	S	kd	$q1$	$q2$	$q3$	h
PIN1	k^0_{pin1}	$k^0_{lat}_{pin1}$	k^S_{pin1}	S_{pin1}	kd_{pin1}	$q1_{pin1}$	$q2_{pin1}$	$q3_{pin1}$	h_{pin1}
PIN2	k^0_{pin2}	$k^0_{lat}_{pin2}$	k^S_{pin2}	S_{pin2}	kd_{pin2}	$q1_{pin2}$	$q2_{pin2}$	$q3_{pin2}$	h_{pin2}
PIN3	k^0_{pin3}		k^S_{pin3}	S_{pin3}	kd_{pin3}	$q1_{pin3}$	$q2_{pin3}$		

3.2.3 Методы численного исследования

Выполнение одного численного эксперимента на персональном компьютере занимало от нескольких минут, до нескольких часов. В таком режиме задача поиска значений параметров, которые бы выводили динамику изменения концентраций переменных модели на режимы согласующиеся с экспериментальными данными, на персональном компьютере является затратной (с точки зрения времени). Для решения этой проблемы был использован разработанный инструментарий системы MGSmodeller по запуску расчётов на высокопроизводительном вычислительном кластере ЦКП «Биоинформатика» ССКЦ СО РАН. На кластере за единицу времени удалось проверять сотни вариантов параметров одновременно, ограничиваясь лишь доступными вычислительными ресурсами суперкомпьютера.

Решив задачу запуска численных экспериментов в параллельном режиме, приходим к проблеме анализа получаемого массива данных. Встала задача

поиска подходящего метода сравнения результатов моделирования и данных эксперимента. Важно, модель была реконструирована в терминах языка моделирования SiBML, где каждой переменной модели сопоставлен идентификатор (в терминах языка моделирования SiBML, см. раздел 2.1.2). Это строка, по которой можно определить отношение переменной модели к конкретной моделируемой сущности (белок транспортёр или фитогормон) и её локализацию в клеточном ансамбле.

Такой идентификатор поддаётся относительно простому автоматическому разбору сторонними средствами. Для решения задачи сопоставления результатов моделирования был разработан инструментарий в рамках пакета MATLAB по фильтрации, реорганизации векторов результатов вычислительных экспериментов и генерации сводных изображений. Этот инструментарий позволил обрабатывать массив полученных на суперкомпьютере результатов, выдавая серию изображений с отображением концентрации переменных на структуру клеток модели. Результирующее изображение содержит отображение концентраций белков транспортёров PIN1-PIN3 и гормона ауксина, каждый в отдельной структуре (рисунок 3.2.2).

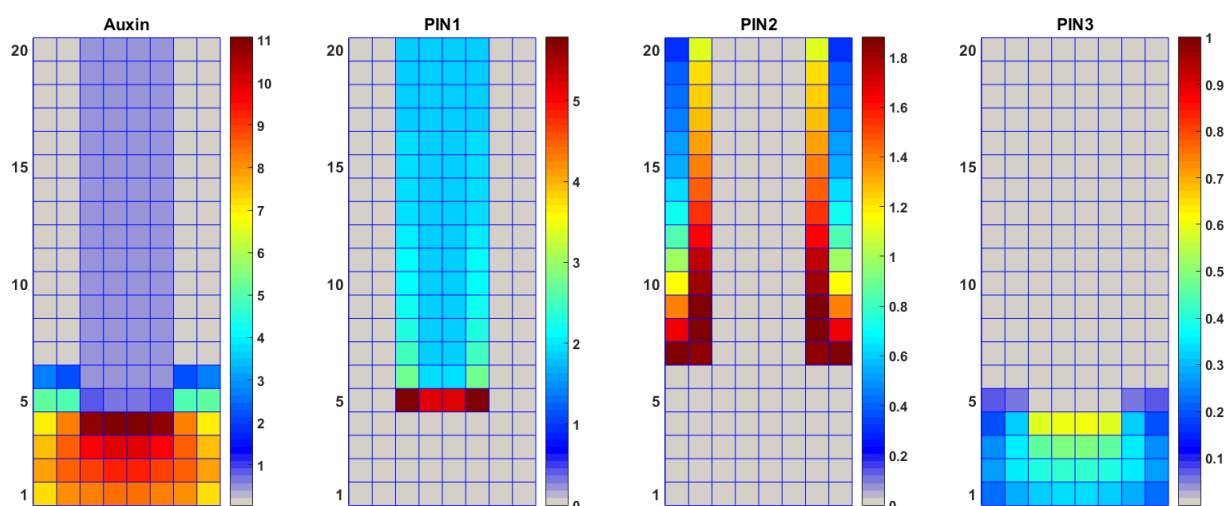


Рисунок 3.2.2 – Распределение концентраций ауксина и белков транспортеров по структуре корня в стационаре модели «ММК:4x20.контроль» (подразумевается позиция покоящегося центра в 4 ряду). Шкалы в условных концентрационных единицах.

Такие изображения поступали экспертам-биологам для оценки результатов и сопоставления с данными экспериментов. Для оценки корректности динамики поведения модели генерировалась анимация, по которой можно было отследить как достигается то или иное стационарное состояние.

В нашем случае необходимо было качественное совпадение результатов, так как получение количественных данных об интенсивности свечения флюоресцирующего сигнального белка сильно зависит от настроек экспериментальной установки и сильно варьируется от эксперимента к эксперименту. Отметим, что прямое сравнение вектора значений переменных и полученных с микроскопа изображений было затруднительно. Это делает проблемной количественную привязку по данным. Однако способ экспертного визуального анализа результатов моделирования показал себя с лучшей стороны.

3.2.4 Анализ моделей ММК

Напомним, что модель разрабатывалась как интеграция двух опубликованных моделей. В первой (Grieneisen et al. 2007) было статическое распределение белков транспортеров и динамика ауксина рассчитывалась в терминах дифференциальных уравнений в частных производных (условное название механизма «обратный фонтан»). Во второй (Mironova et al. 2010) рассматривался только один белок транспортер PIN1, но с ауксин-зависимым биосинтезом (условное название механизма «отраженная волна»). Обе этих модели демонстрировали формирование характерного, наблюдаемого в экспериментах паттерна распределения ауксина. Построенная интегральная модель демонстрирует поведение базовых моделей (с выбранными соответствующими параметрами).

3.2.4.1 Конфигурация модели «ММК:4x20.отражённая волна»

Данная версия модели имеет следующую конфигурацию: активный транспорт ауксина осуществляется только посредством белка PIN1 по сосудистым тканям; PIN1 имеет ауксин-зависимый биосинтез. Подбор параметров модели опирался на значения из модели «отражённая волна» добиваясь в стационаре формирования максимума концентрации ауксина в клетках покоящегося центра (4-ый ряд от кончика корня). Подбирался параметр k_{Alpha} – интенсивность поступления ауксина в сосудистые ткани из побега (рис. 3.2.3, таблица 3.2.2). Модель воспроизводит представленные в работе Мироновой и соавторов (Mironova et al. 2010) результаты.

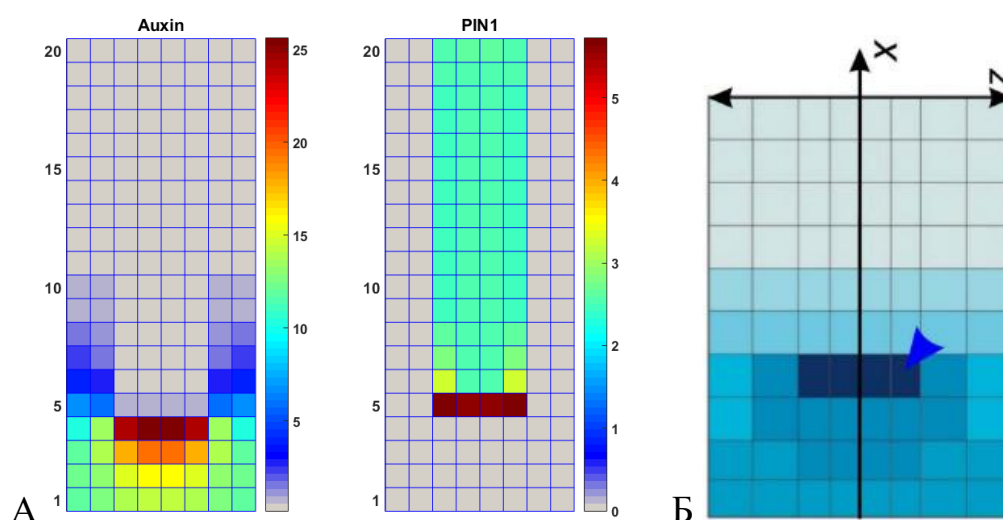


Рисунок 3.2.3 – А) распределение ауксина и белка PIN1 по структуре модели «ММК:4x20.отражённая волна» в состоянии стационара. Б) Распределение ауксина по структуре модели «отражённая волна» из статьи (Mironova et al. 2010). Шкалы в условных концентрационных единицах.

Таблица 3.2.2 – Значения параметров модели «ММК:4x20.отражённая волна»

	k_{Alpha}	D	kd	Значения параметров модели «ММК:4x20.отражённая волна»					
Auxin	0.8	0.08	0.005						
	$k0$	$k0_{lat}$	kS	S	kd	$q1$	$q2$	$q3$	h
PIN1	1.0	0.0	1000.0	2.0	1000	0.3	10	1.0	6.0
PIN2	0.0	0.0	1000.0	2.0	1000	0.1	0.8	0.2	4
PIN3	0.0		1000.0	4	1000	10	10		

3.2.4.2 Конфигурация модели «ММК:4x20.обратный фонтан».

Данная версия модели построена на установках из работы (Grieneisen et al. 2007) и имеет следующую конфигурацию: биосинтез и деградация белков транспортеров не учитывается; концентрация и локализация белков транспортеров заданы неизменными; PIN1 транспортирует ауксин по сосудистым тканям к кончику корня, PIN2 – по тканям эпидермиса и кортекса по направлению к побегу и в центральную часть (в отношении 70/30); PIN3 – способен транспортировать ауксин во всех тканях по всем направлениям. В вычислительных экспериментах осуществлялся поиск значений следующих параметров – k_{Alpha} (интенсивность притока ауксина в модель) и концентрации PIN белков. Добивались достижения в стационаре формирования максимума концентрации ауксина в клетках покоящегося центра (4-ый ряд от кончика корня). Наша модель с найденными параметрами (таблица 3.2.3) воспроизводит результаты представленные Гринайсен и соавторами (Grieneisen et al. 2007) (рис. 3.2.4).

Модель демонстрирует распределение ауксина по заданной структуре расположения белков транспортеров в тканях корня. Модель не способна дать ответа на вопрос: как белки транспортеры смогли прийти в это состояние?

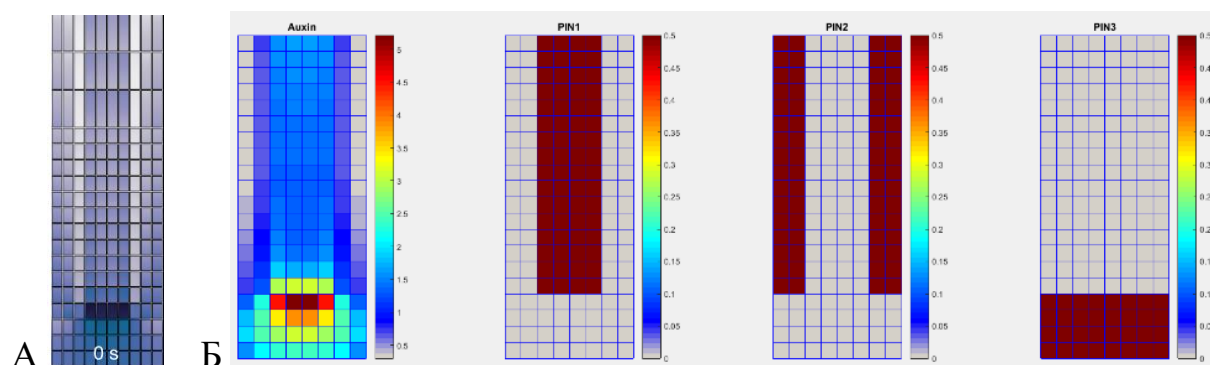


Рисунок 3.2.4 – А) Распределение ауксина по структуре модели «обратный фонтан» из статьи (Grieneisen et al. 2007). Б) Распределение ауксина и белков транспортеров PIN по структуре модели «ММК:4x20.обратный фонтан» в состоянии стационара. Шкалы в условных концентрационных единицах.

Таблица 3.2.3 – Значения параметров модели «ММК:4x20.обратный_фонтан»

	$kAlpha$	D	kd	Значения параметров модели «ММК:4x20.обратный_фонтан»					
Auxin	0.5	0.08	0.005	S	kd	$q1$	$q2$	$q3$	h
PIN1	1.0	0.0	1000.0	2.0	1000	0.3	10	1.0	6.0
PIN2	0.7	0.3	1000.0	2.0	1000	0.1	0.8	0.2	4
PIN3	1		1000.0	4	1000	10	10		

3.2.4.3 Конфигурация модели «ММК:4x20.двойной_механизм»

В объединении моделей к белку PIN1 добавлены белки транспортеры PIN2 и PIN3 с заданным ауксин-зависимым биосинтезом (профиль экспрессии см. рис. 3.2.1 В). В воспроизведении обеих версий механизмов («отраженная волна» и «обратный фонтан») была получена первичная оценка области значений параметров – отправная точка по подбору параметров биосинтеза белков транспортеров и их транспортной активности. Далее проводились вычислительные эксперименты с поиском соответствия динамики модели экспериментально наблюдаемым распределениям белков транспортеров (Vieten et al. 2005).

3.2.4.4 Модель «ММК:4x20.двойной механизм» без потока в центр

Модель включает ауксин-зависимый биосинтез всех трех белков транспортеров. Латеральный поток ауксина посредством белка PIN2 (по направлению к центральной оси корня) отключен. Несмотря на то, что латеральный поток через PIN2 присутствует в модели Гринайсен (Grieneisen et al. 2007), описание механизма «обратного фонтана» не подразумевает его наличие. Вычислительные эксперименты с отключённым латеральным потоком демонстрируют возможность установления как паттерна

концентрации ауксина с максимумом в 4-ом ряду клеток (в зоне «покоящегося центра» ниши стволовых клеток), так и корректного паттерна локализации белков транспортеров (рис. 3.2.5, таблица 3.2.4). Однако, наблюдаются пики концентрации ауксина во втором слое клеток (кортексе). Такие пики концентраций не наблюдаются в экспериментальных данных и поэтому были восприняты как артефакты модели. Эти пики возникали по причине неоднородного распределения концентрации ауксина, образующегося за счет пассивного транспорта (диффузии) из клеток центральных тканей.

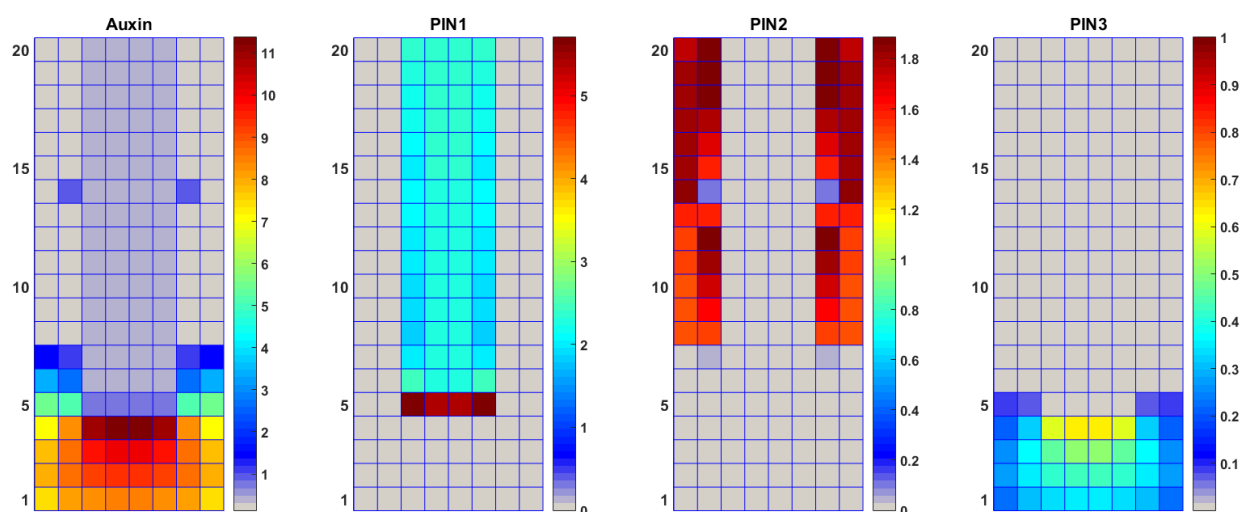


Рисунок 3.2.5. Распределение ауксина и белков транспортеров PIN1-3 по структуре модели «ММК:4x20.двойной_механизм» в состоянии стационара с отсутствующим латеральным потоком посредством белка PIN2. Шкалы в условных концентрационных единицах.

Таблица 3.2.4 – Параметры модели «ММК:4x20.двойной_механизм» без латерального потока

	k_{alpha}	D	kd	Значения параметров модели «ММК:4x20.двойной_механизм» с отключенным латеральным потоком через PIN2					
	k_0	k_{0_lar}	kS	S	kd	$q1$	$q2$	$q3$	h
Auxin	0.7	0.08	0.005						
PIN1	0.9	0.1	1000.0	2.0	1000	0.3	10	1.0	6.0
PIN2	1.0	0.0	1000.0	2.0	1000	0.1	0.8	0.2	4
PIN3	1.0		1000.0	4	1000	10	10		

Таким образом прямая интеграция двух механизмов не позволяет добиться согласования с данными эксперимента. Изучив подробнее экспериментальные данные, было обнаружено, что белок PIN2 локализуется не только на апикальной мембране клеток корня, но и на ориентированной в сторону центральной оси корня мембране. Данные с экспериментальных фотографий свидетельствуют в пользу того, что белок PIN2 осуществляет активный транспорт ауксина не только апикально (по направлению к побегу), но и частично латерально (по направлению к центральной оси симметрии корня). В следующей версии модели, «ММК:4x20.контроль», был учтен этот феномен.

С другой стороны, наблюдаемый артефакт в модели не стоит игнорировать как «небиологичный» пример. Это может свидетельствовать о возможности формирования пиков концентрации в слоях кортекса под воздействием потоков ауксина осуществляемых белком PIN2 в зонах корня где отсутствует латеральный поток (к центральной оси корня). Т.е. в тканях лежащих выше меристематической зоны корня. Этот эффект может служить причиной заложения латеральных корней и требует дополнительного изучения.

3.2.4.5 Модель «ММК:4x20.контроль»

Модель включает ауксин-зависимый биосинтез всех трех белков транспортеров. Активность PIN2 разделена на два потока: апикальный (вверх) и латеральный (к центральной оси корня). В проведенной серии вычислительных экспериментов не возникало артефактов, выявленных в предыдущей модели (рис. 3.2.6). Латеральный поток сглаживает возникающие «шумы». С применением высокопроизводительного вычислительного кластера ЦКП «Биоинформатика» была проведена серия вычислительных экспериментов с варьированием параметров модели, включающих параметры биосинтеза белков, параметров скорости активного транспорта и долей потока через PIN2. Отправной точкой в этом процессе служили параметры из моделей

«отраженной волны» и «обратного фонтана». В результате был найден набор параметров модели (таблица 3.2.5), который показывает качественное совпадение с экспериментальными данными как в паттерне ауксина (максимума точно в 4-ом ряду клеток), так и в локализации белков транспортеров (рис. 3.2.6). Этот набор параметров был принят как контроль.

Экспертная оценка результатов вычислительных экспериментов показала, что стационарное решение модели «ММК:4x20.контроль» согласуется с экспериментальными данными в следующих пунктах:

- Ауксин не только имеет характерный паттерн с максимумом в зоне покоящегося центра, но и присутствует в сосудистых тканях.
- Белок PIN1 экспрессируется в сосудистых клетках строго до зоны покоящегося центра.
- Белок PIN2 экспрессируется с разной интенсивностью во внешних слоях корня.
- Белок PIN3 экспрессируется в кончике корня

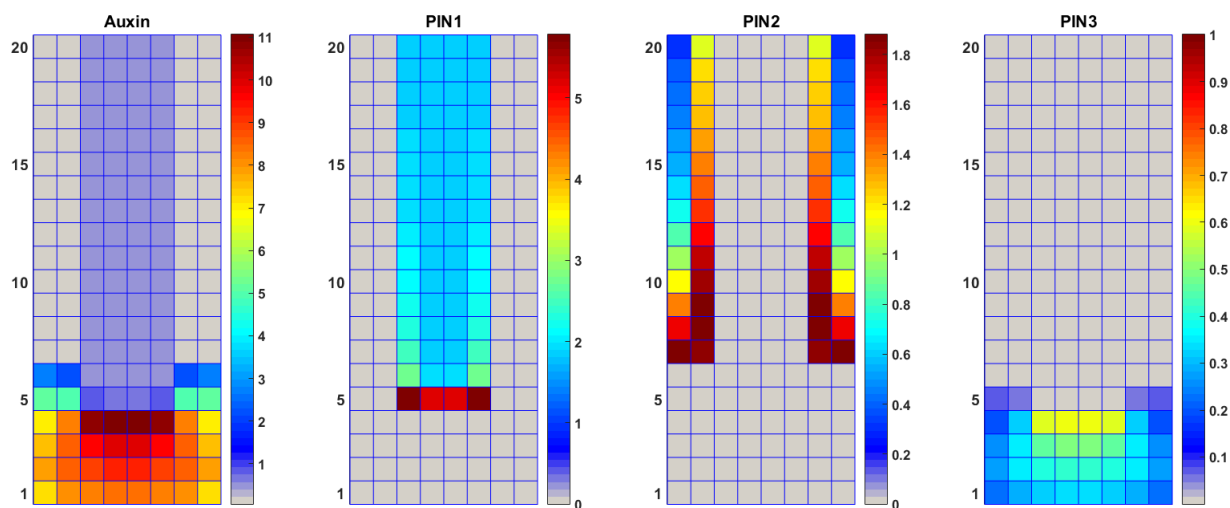


Рисунок 3.2.6 – Распределение ауксина и белков транспортеров PIN1-3 по структуре модели «ММК:4x20.контроль» в состоянии стационара. Шкалы в условных концентрационных единицах.

Таблица 3.2.5. Параметры модели 4x20, контроль.

	k_{alpha}	D	kd	Значения параметров модели «ММК:4x20.контроль»					
Auxin	0.5	0.08	0.005						
	k_0	k_{0_lat}	kS	S	kd	$q1$	$q2$	$q3$	h
PIN1	0.9	0.1	1000.0	2.0	1000	0.3	10	1.0	6.0
PIN2	0.7	0.3	1000.0	2.0	1000	0.1	0.8	0.2	4
PIN3	1.0		1000.0	4	1000	10	10		

3.2.5 Анализ модели «ММК:4x20.контроль»

Было проведено численное исследование модели по следующим пунктам:

- Устойчивость стационарного решения к изменению начальных значений переменных модели (распределения ауксина и белков транспортеров PIN1, PIN2 и PIN3).
- Исследование на устойчивость модели в формировании заданного паттерна распределения ауксина при варьировании набора параметров.

В рамках исследования устойчивости стационарного решения была проведена серия вычислительных экспериментов, в которых были найдены интервалы отклонения параметров и начальных значений переменных, в которых сохраняется стационарное решение модели с характерным паттерном концентрации ауксина с максимумом в 4-ом ряду клеток. Также проверялась серия случайных наборов параметров в этих интервалах, в поисках комбинаций приводящих к отклонению в организации паттерна.

Анализ устойчивости решения к начальным данным - исходному распределению ауксина и белков транспортеров - показал, что увеличение концентрации ауксина по всем клеткам свыше 10% приводит к смещению точки максимума. Повышение концентрации ауксина в случайных клетках модели «ММК:4x20.контроль» не сказывается на стационарном решении.

Была проведена серия вычислительных экспериментов, в которых начальное распределение как ауксина, так и белков PIN1-3 было задано случайными значениями. Причем максимальное значение случайного разброса значения варьировалось в пределах от $X*0.1$ до $X*10.0$ с шагом 0.2. Т.е. для каждого выбранного значения максимума было проведено 50 расчетов со случайными начальными концентрациями, не превышающими выбранного максимума (рис. 3.2.7).

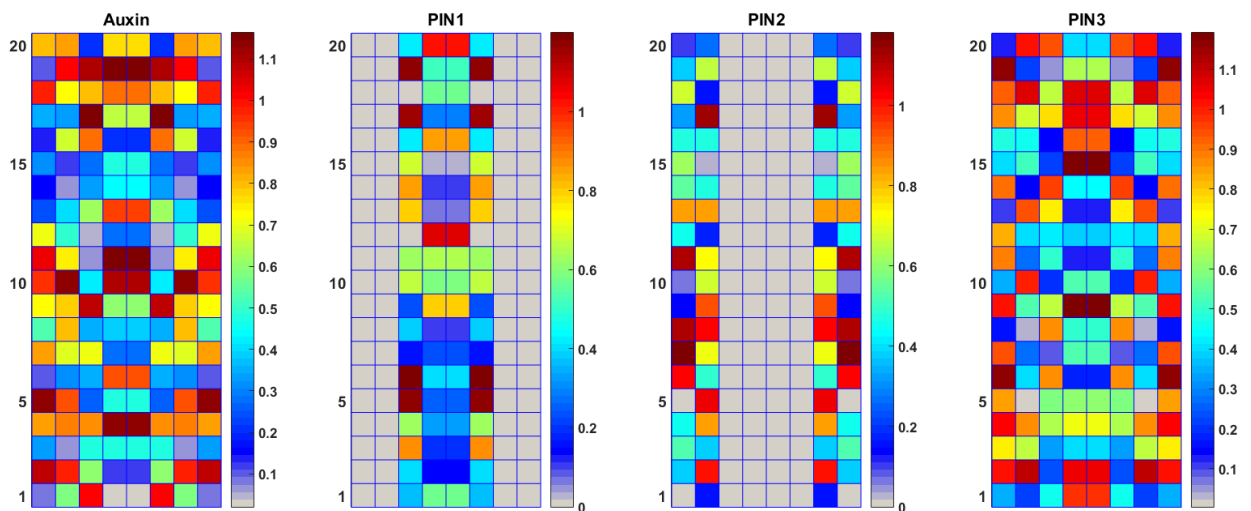


Рисунок 3.2.7 – Вариант исходного распределения начальных концентраций в модели «ММК:4x20.контроль». Шкалы в условных концентрационных единицах.

Модель «ММК:4x20.контроль» демонстрирует разные устойчивые решения в вычислительных экспериментах с варьированием начальных значений ауксина и белков транспортеров:

- Формирование максимума концентрации ауксина в зоне покоящегося центра. В серии вычислительных экспериментов, где в начальном распределении максимум концентрации ауксина в любой из клеток не превышал 140% от максимума концентрации ауксина модели «ММК:4x20.контроль» (рис. 3.2.8 А). Это состояние наблюдается и при превышении границы 140% от максимума контроля в случайном

распределении, но периодически начинают появляться остальные варианты.

- Формирование второго максимума как вблизи, так и на удалении от покоящегося центра (рис. 3.2.8 Б).
- Формирование нескольких максимумов концентрации ауксина (рис. 3.2.8 В,Г).
- Формирование характерного распределения ауксина в кончике корня со смещением максимума концентрации (рис. 3.2.8 Д).

Также была проведена серия вычислительных экспериментов с оценкой чувствительности модели на изменения в структуре клеточного ансамбля. Исследовалось влияние конфигурации модели на ее решения. Были реконструированы и исследованы модели «ММК:4x100.контроль», «ММК:4x50.контроль», «ММК:4x35.контроль», «ММК:4x15.контроль», «ММК:4x10.контроль» и «ММК:4x05.контроль». Характерный паттерн с максимумом концентрации ауксина в 4-ом ряду от кончика корня (в зоне клеток покоящегося центра ниши ствольных клеток) достигается на моделях всех озвученных размеров (рис. 3.2.9). При этом, очевидно, различаются только значение параметра k_{alpha} – интенсивность притока ауксина в систему (см. таблицу 3.2.6).

В вычислительных экспериментах показано, что процессы, на которых построена модель, способны удерживать максимум концентрации ауксина в широком диапазоне длин корня, осуществляя поддержку функции ниши ствольных клеток корня. Они достаточны для регулирования поддержания функции НСК в развивающемся корне. Этот феномен наблюдается на моделях меристемы разной длины (вплоть до 100 рядов клеток).

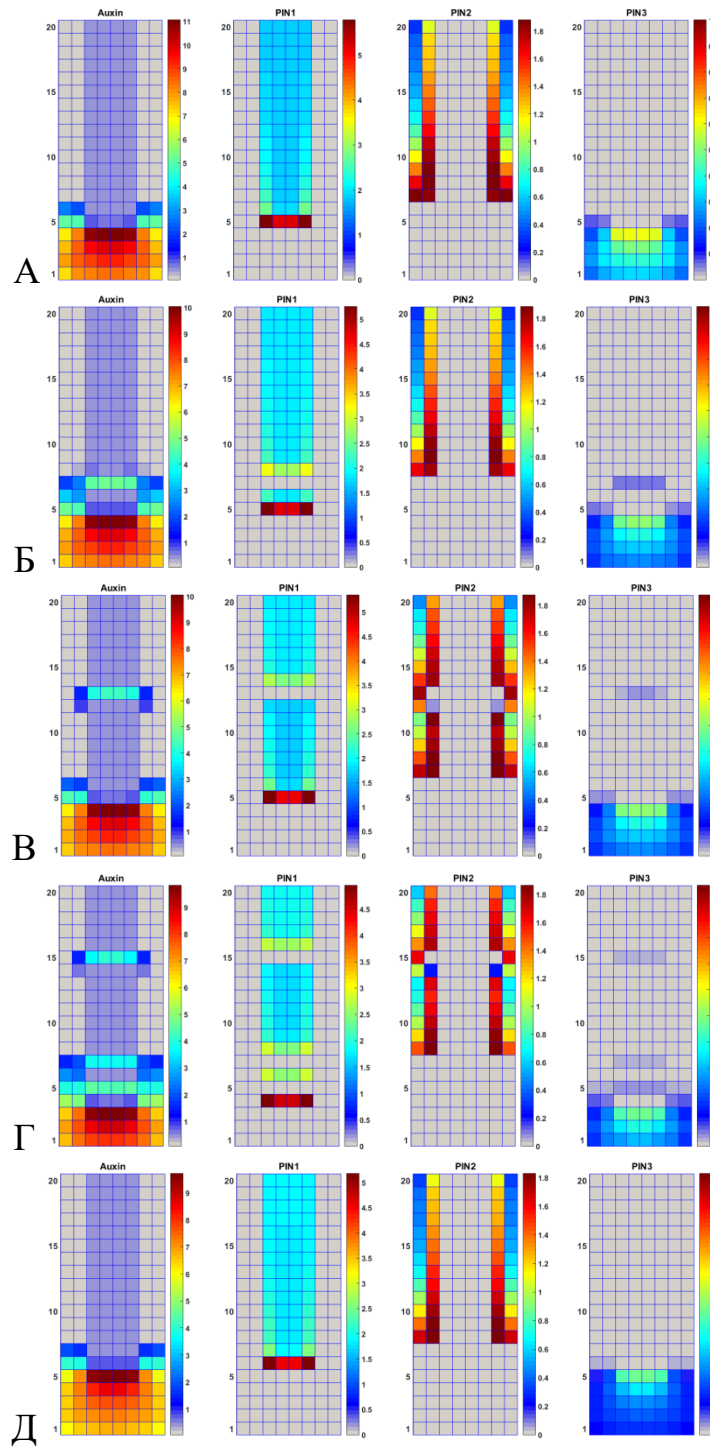


Рисунок 3.2.8 – Варианты распределения переменных модели в стационаре «ММК:4x20.контроль» по структуре клеток, полученные при различных начальных данных: А) Нормальный паттерн ауксина с максимумом в зоне покоящегося центра – 4 ряд от кончика корня; Б) Сформированный второй максимум ауксина вблизи клеток покоящегося центра; В) Сформированный второй максимум ауксина на удалении от клеток покоящегося центра; Г) Формирование больше двух максимумов; Д) Формирование паттерна ауксина со смещением точки максимума от клеток покоящегося центра. Шкалы в условных концентрационных единицах.

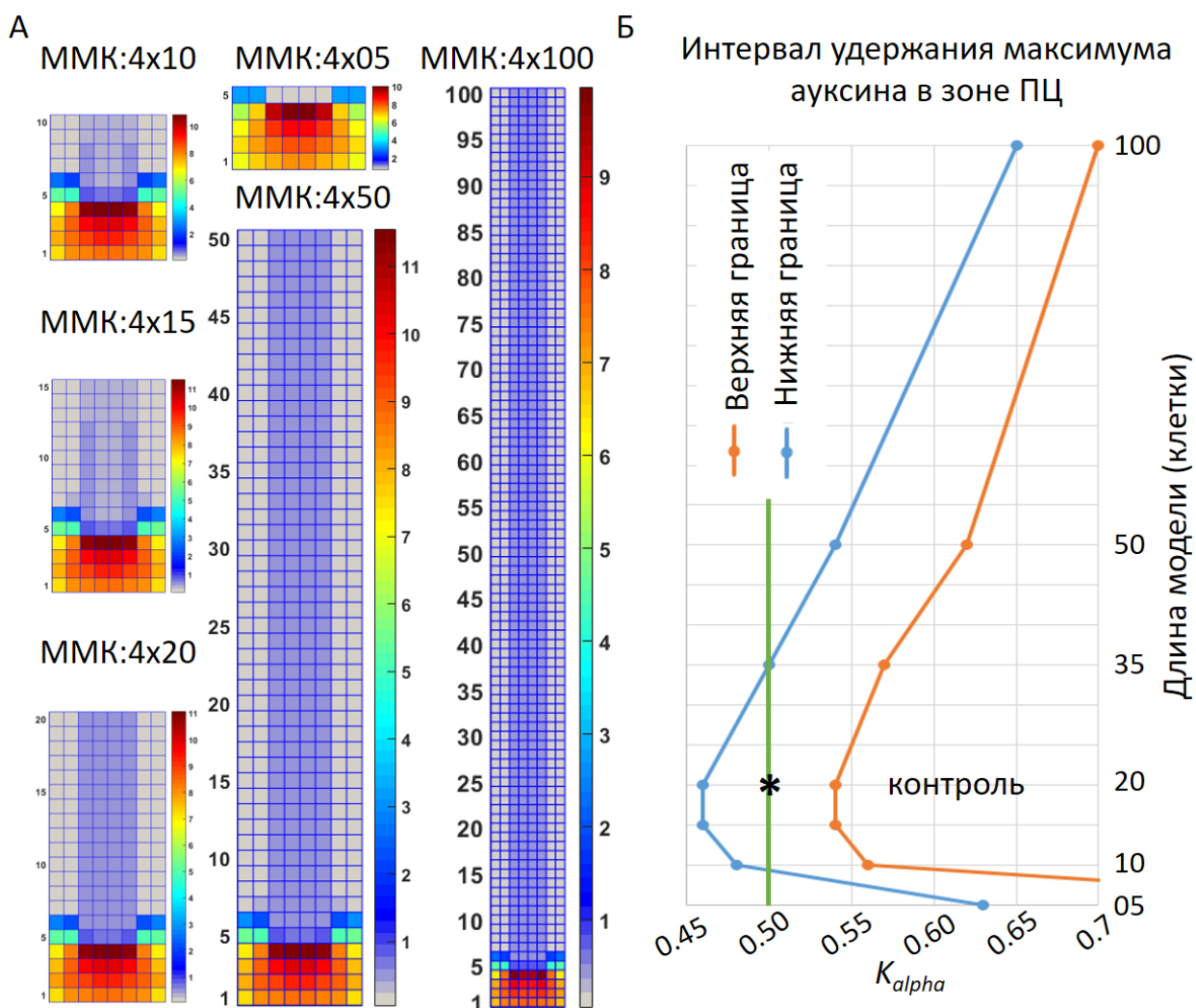


Рисунок 3.2.9 – А) Варианты распределения ауксина по структуре клеток моделируемого корня в стационаре на моделях «ММК:4х##.контроль». Б) Сводная диаграмма на которой отмечены интервалы значений параметра k_{alpha} в границах которого (Синяя – нижний предел, Оранжевая – верхний предел) удерживается максимум концентрации ауксина в зоне покоящегося центра (4 ряд клеток от кончика). По оси ОХ отложены значения интенсивности притока ауксина в модель (k_{alpha}). По оси ОУ отмечены размеры моделей на которых варьировались параметры. Звездочкой отмечены значения модели «ММК:4х20.контроль». Зеленым отмечен параметр k_{alpha} который был выбран как контроль. Шкалы в условных концентрационных единицах.

Таблица 3.2.6 – Интервалы параметра k_{alpha} в которых удерживается максимум в зоне покоящегося центра.

длина модельного корня (в клетках)	интервал k_{alpha}
100	[0.65; 0.70]
50	[0.54; 0.62]
35	[0.50; 0.57]
20	[0.46; 0.54]
15	[0.46; 0.54]
10	[0.48; 0.56]
5	[0.63; 0.93]

3.2.6 Модели ММК мутантов по генам белков транспортёров

Для *A. thaliana* существуют растения мутанты по генам белков транспортёров *PIN*, среди одиночных мутантов, только растения *pin2* имеют ярко выраженный фенотип в корне - нарушение гравитропического роста, у остальных мутантов происходит компенсация развития аномалий гомологами (Blilou et al. 2005). В рамках получившейся модели возможно исследовать эти проявления мутаций (рис. 3.2.10).

Модель нокаута *PIN2* «ММК:4x20.контроль.pin2_нокаут» способна сформировать паттерн ауксина с максимумом в зоне покоящегося центра. Причём поведение модели мало отличается от «ММК:4x20.контроль» и по интервалу параметра k_{alpha} в котором формируется паттерн ауксина (таблица. 3.2.8). Изменения проявляются в градиенте концентраций ауксина в слоях эпидермиса и кортекса, где в «ММК:4x20.контроль» имеется более чёткая граница из-за работы *PIN2* на отток ауксина. Есть экспериментальные работы где получали таких мутантов путём смены полярности клеток в слоях кортекса и эпидермиса (Wisniewska et al. 2006; Rahman et al. 2010).

Модель с нокаутом *PIN3* «ММК:4x20.контроль.pin3_нокаут» также способна формировать паттерн ауксина с максимумом в зоне покоящегося центра. Из-за отсутствия белка PIN3, который работает при высоких концентрациях ауксина и активно его выводит во все соседние клетки, в зоне покоящегося центра формируется узкое пятно с максимумом концентрации ауксина. И значение максимума значительно выше, относительно «ММК:4x20.контроль». Ауксина больше за счёт того, что интервал параметра k_{alpha} , в границах которого формируется нужный паттерн, выше (таблица. 3.2.8)

Модель с нокаутом *pin2pin3* «ММК:4x20.контроль.pin2pin3_нокаут» способна формировать паттерн ауксина с максимумом в зоне покоящегося центра. Причём этот феномен был воспроизведён на моделях разных линейных размеров вплоть до 4x100 клеток. Более того, такой нокаут уже был воспроизведён как модель «отражённой волны» на двумерной структуре клеток в работе (Mironova et al. 2010). Модель демонстрирует градиент ауксина в слоях эпидермиса и кортекса из-за отсутствия оттока через PIN2, и без PIN3 значение максимума значительно выше, относительно «ММК:4x20.контроль», как и значения границ интервала параметра k_{alpha} (таблицы 3.2.7 и 3.2.8).

Модель с нокаутом гена *PIN1* «ММК:4x20.контроль.PIN1_нокаут» в вычислительных экспериментах оказалась неспособна сформировать паттерн с максимумом ауксина в зоне покоящегося центра. Распределение ауксина идёт градиентом от места притока. Причём потока не хватает для формирования «отражённой волны» в кончике корня. Для растения в целом такая мутация является летальной, так как затрагивает фундаментальные процессы развития всего растения. Растение *pin1* или не формирует основной корень (Huang et al. 2010), или он формируется (Vieten et al. 2005) когда полярный поток ауксина в корень начинают регулировать гомологи из семейства PIN. Можно сделать вывод, что без направленного потока ауксина в сторону кончика корня НСК не сформируется и корень расти не будет.

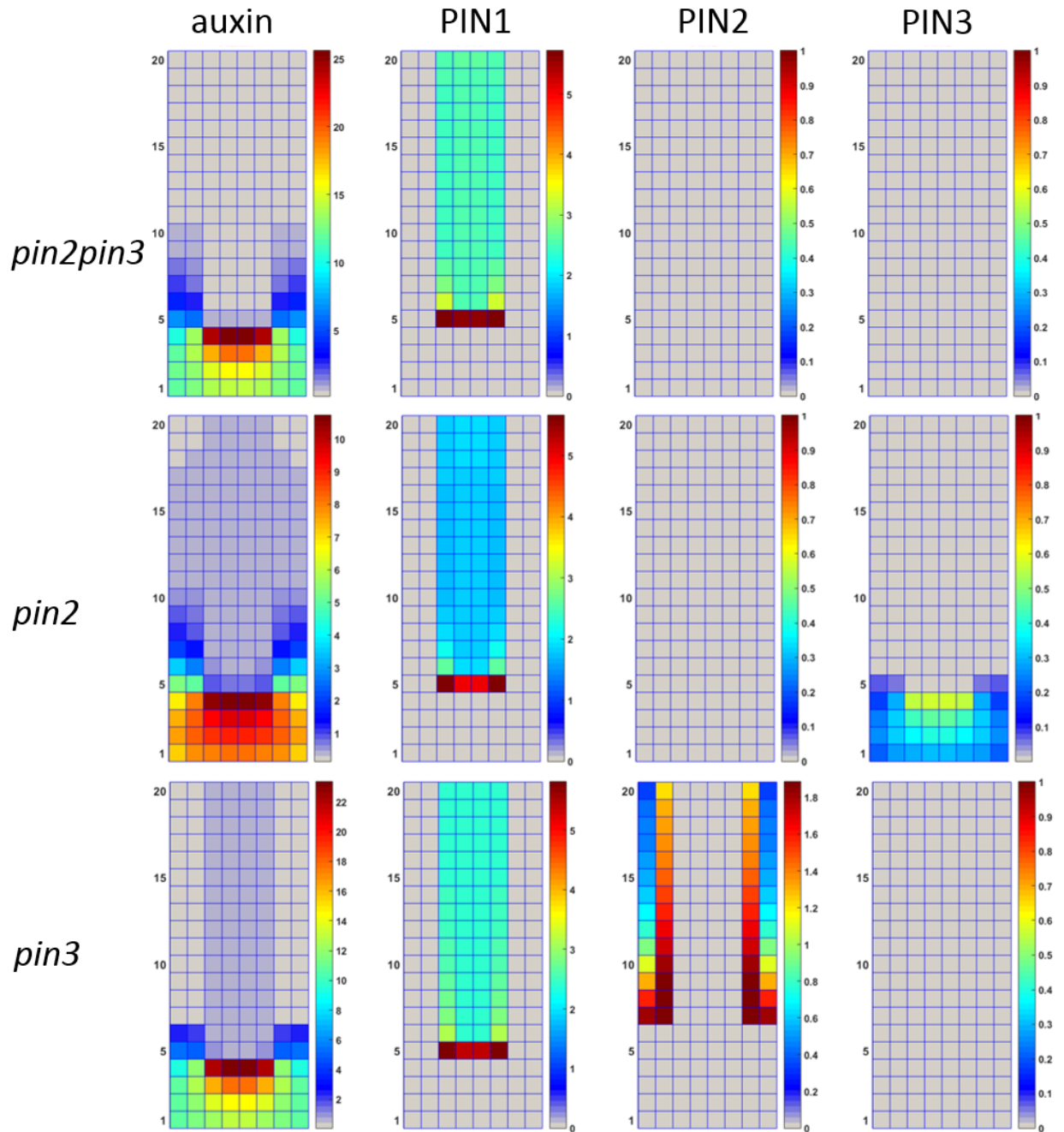


Рисунок 3.2.10 – Распределение ауксина и белков транспортёров по структуре корня растения в стационаре на моделях «ММК:4x20.#» мутантов нокауты. Показаны стационары при разных значениях параметра k_{alpha} : *pin2pin3* мутант – 0.8; *pin2* мутант – 0.5; *pin3* мутант – 0.72. Шкалы в условных концентрационных единицах.

Сводные характеристики по интервалам значения скорости притока ауксина в модель (k_{alpha}) на разных моделях мутантов при которых удерживается максимум ауксина в зоне ниши стволовых клеток корня представлен в таблице 3.2.8.

Таблица 3.2.7 – Интервалы значений параметра k_{alpha} в которых формируется и удерживается максимум ауксина в зоне покоящегося центра корня в моделях «ММК:##.контроль.PIN2PIN3_нокаут» размерами 4x20, 4x50, 4x100 клеток и в соответствующем размере моделей «ММК:##.контроль».

	ММК:PIN1 [k_{alpha}]	ММК:PIN1 k_{alpha}]	Контроль [k_{alpha}]	Контроль k_{alpha}]
4x20	0.75	0.80	0.45	0.55
4x30	0.79	0.88	0.5	0.57
4x50	0.84	0.95	0.54	0.62
4x100	1.09	1.15	0.65	0.7

Таблица 3.2.8 – Интервалы параметра k_{alpha} в рамках которых формируется максимум концентрации ауксина в зоне покоящегося центра на моделях «ММК:4x20###» мутантов нокаут.

Интервал \ ММК:4x20	контроль	мутант <i>pin2</i>	мутант <i>pin3</i>	мутант <i>pin2pin3</i>
нижняя граница	0.46	0.47	0.69	0.72
верхняя граница	0.54	0.54	0.78	0.83

3.2.7. Заключение по модели ММК

Разработанная математическая модель «ММК:4x##.контроль» распределения фитогормона ауксина по структуре клеток корня посредством белков транспортеров PIN семейства с учетом их ауксин-зависимого биосинтеза показала:

- 1) Представленных в модели процессов достаточно для формирования характерного паттерна концентрации ауксина с максимумом в 4-ом ряду клеток на моделях корней разной длины (показано на моделях ММК:4x05.контроль, ММК:4x10.контроль, ММК:4x20.контроль, ММК:4x30.контроль, ММК:4x45.контроль, ММК:4x50.контроль и

ММК:4x100.контроль). Причем эта конфигурация устойчива к изменению потока ауксина от побега.

- 2) Показано что формирование структуры корня (паттерна ауксина) сначала происходит за счет механизма «отраженной волны», а после формирования характерного распределения белков транспортеров поддержание структуры идет по механизму «обратного фонтана» (Mironova et al. 2012).
- 3) Подтверждены высказанные Виетеном с коллегами (Vieten et al. 2005) предположения:
 - а) возможность ауксина регулировать собственный поток в корне (в модели показан механизм и динамика самоорганизации);
 - б) Объяснено формирование распределения ауксина в нише стволовых клеток корня в мутантах по генам транспортерам ауксина (корень способен расти и развиваться, не леталь).

В рамках разработанных моделей воспроизводились результаты исходных моделей «обратного фонтана» и «отражённой волны». Единственное, не смогли найти набора параметров при котором проявились бы некоторые осциллирующие эффекты, которые наблюдались на одномерной модели в работах Лихошвая и Мироновой (Лихошвай и др. 2007; Mironova et al. 2010). Эти эффекты могли быть сглажены за счёт добавления белков PIN2 и PIN3 (Возможно, что тот же эффект что и в разделе 3.2.4.4). Однако, процесс достижения стационара с максимумом в 4-ой клетке идёт волнообразно, передвигая точку максимума от кончика корня постепенно (Анимированная динамика модели явно демонстрирует этот эффект).

Было показано, что паттерн ауксина достижим и при активности всего нескольких белков транспортёров (показано установление паттерна ауксина на моделях мутантов с нокаутами *PIN2*, *PIN3*, *PIN2 PIN3*).

В сравнении модели «ММК:4x20.контроль» и «ММК:4x20.контроль.*pin2pin3_нокаут*» (как модель с работающим одним

PIN1 белком транспортёром) показано, что увеличение количества белков транспортёров и разделение их по функциональной нагрузке позволяет при формировании максимума ауксина в зоне ствольных клеток корня использовать поток ауксина меньшей мощности (k_{alpha}) и с широким пределом варьирования этой мощности (см. таблицу 3.2.8).

Результаты данного исследования были опубликованы нами в статье (Mironova et al. 2012). Полученное знание о процессах формирования потоков в нише ствольных клеток позволит перенести его на процессы формирования латеральных корней и меристем побега, так как там работают похожие механизмы органов со ствольными клетками.

3.3. Модель меристемы корня с учётом влияния салициловой

КИСЛОТЫ

3.3.1 Постановка задачи

Результатами, полученными на моделях ММК, заинтересовались в лаборатории Клауса Пальме из университета г. Фрайбург. Ими были предложены результаты серии экспериментов по влиянию салициловой кислоты на меристему корня *A. thaliana* для дополнительного анализа. Салициловая кислота – это известный гормон стресса растений, который вырабатывается в ответ на воздействие патогенов (Rivas-San Vicente and Plasencia 2011). В предоставленных данных при обработке проростков *A. thaliana* экзогенной 30 μM салициловой кислотой проявляются следующие явления (Pasternak et al. 2019):

- Наблюдается характерный паттерн распределения ауксина с максимумом концентрации в зоне покоящегося центра ниши ствольных клеток корня.
- Наблюдается изменение (увеличение) интенсивности маркера ауксина в клетках меристемы корня, свидетельствующее о наличии базального синтеза ауксина во всех клетках.
- Наблюдается изменение интенсивности маркеров белков транспортёров PIN1, PIN2 и PIN3.

Эти аспекты было предложено исследовать методами математического и компьютерного моделирования в рамках модели ММК (см. пункт 3.2). Проанализировав предоставленные данные было решено развить математическую модель путём внесения ряда новых параметров, отвечающих за степень воздействия салициловой кислоты на процессы биосинтеза ауксина и белков транспортёров. Также, для более корректного соответствия тканей

модельного объекта реальному корню, область моделирования было решено расширить до размеров 10x25 клеток. Эта версия модели здесь и далее будет именоваться **МСК** – Модель меристемы корня с учётом влияния Салициловой Кислоты (рис. 3.3.1). Теперь в модели рассматриваются следующие слои тканей (выстроены в порядке от внешних слоёв к центральной оси корня): эпидермис, кортекс, эндодерма, перицикл (добавлен этот слой) и сосудистый слой.

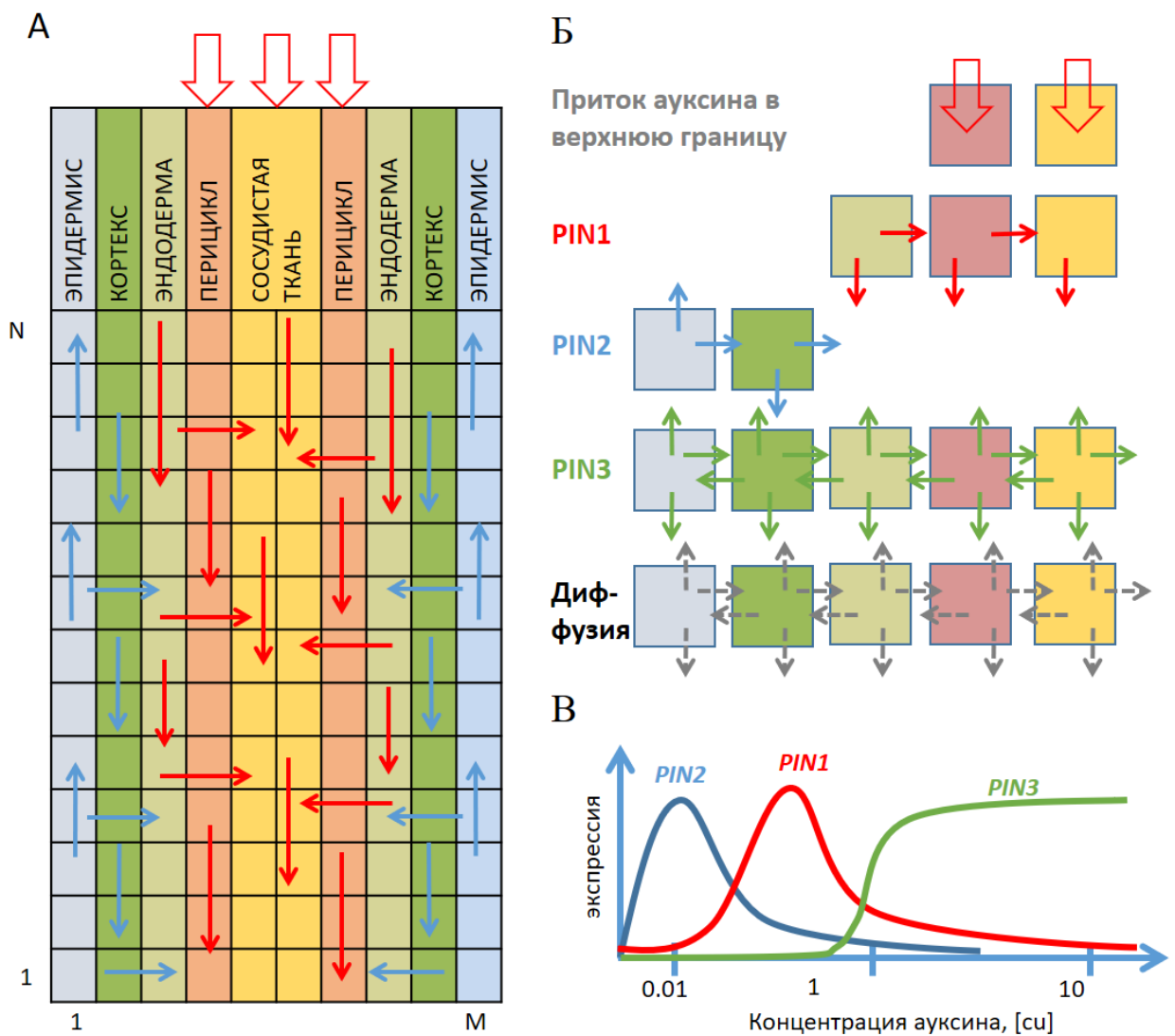


Рисунок 3.3.1 – Схема модели МСК. А) Структура моделируемого клеточного ансамбля с направлениями потоков ауксина. Б) схема мест синтеза белков транспортеров и определяющее ими направление транспорта ауксина. Цветом задана принадлежность к слоям клеток из пункта А. В) Схематично представленный профиль экспрессии белков транспортеров в зависимости от концентрации ауксина в клетке (Pasternak et al. 2019).

Таблица 3.3.1 – Распределение белков транспортёров в структуре модели МСК

Название слоя	Экспрессия белков в слое
эпидермис	PIN2, PIN3
кортекс	PIN2, PIN3
эндодерма	PIN1, PIN3
перицикл	PIN1, PIN3
сосудистый	PIN1, PIN3

На биосинтез белков транспортёров наложена серия ограничений (Таблица 3.3.1). В модели также полагается комплексная роль белка транспортёра PIN3 (включает в себя интегральную активность белков PIN3, 4 и 7, так как сложно разделить их активности в экспериментах).

Следующие положения были заложены в модели МСК:

- Модель, с точки зрения структуры, представляет собой организованный массив клеток с притоком ауксина из вне в клетки перицикла и сосудистой ткани.
- Переменными являются концентрации ауксина и белков транспортёров (PIN1, PIN2 и PIN3) в каждой из клеток моделируемой области.
- В каждой клетке происходят независимые от соседних клеток процессы биосинтеза белков транспортёров и биосинтеза ауксина.
- В модели делается допущение, что все синтезируемые белки транспортёры моментально встраиваются в мембрану и начинают выполнять свою функцию.
- Белок PIN1 транспортирует ауксин как по направлению к кончику корня (вниз), так и латерально (по направлению к центральной оси корня).
- Белок PIN2 в клетках эпидермиса транспортирует ауксин по направлению к побегу, а в клетках кортекса транспортирует ауксин по направлению к кончику корня. В каждом из слоев PIN2 осуществляет

латеральный активный транспорт ауксина (по направлению к центральной оси корня).

- Белок PIN3 экспрессируется во всех клетках корня и осуществляет транспорт ауксина во всех направлениях одинаково сильно.
- Помимо активного транспорта ауксина в модели присутствует пассивный транспорт ауксина, соответствующий процессам диффузии ауксина между клетками.

В модели МСК, относительно модели ММК, изменению подверглись следующие элементарные подсистемы (в терминах формализма SiBML):

1. Добавлена подсистема биосинтеза ауксина в каждой клетке модели.
2. В модели процессов ауксин-зависимого биосинтеза белков PIN1 - PIN3 был введён множитель, отвечающий за влияние салициловой кислоты.
3. Был модифицирован сценарий сборки модели добавлением дополнительного клеточного слоя.

Средствами автоматизации, язык моделирования SiBML и среда разработки MGSmodeller, сгенерирована математическая модель МСК в виде системы обыкновенных дифференциальных уравнений.

3.3.2 Модели элементарных подсистем МСК

Следующие модели элементарных подсистем формируют финальную модель (размерности далее: cu - концентрационные единицы, tu – временные единицы, dl – безразмерный параметр):

1. Биосинтез ауксина в клетках:

$$k_{Auxin}^{syn}$$

где k_{Auxin}^{syn} (cu/tu) – интенсивность биосинтеза ауксина в клетке.

2. Ауксин зависимый биосинтез белка PIN1 в клетке с индексом (i,j) :

$$k_{pin1}^S * k_{pin1}^{SA} * \frac{\left(\frac{Auxin_{ij}}{q_{pin1}^1}\right)^{S_{pin1}}}{\left(1 + \frac{Auxin_{ij}}{q_{pin1}^2}\right)^{S_{pin1}}},$$

где $k_{pin1}^S (l/tu)$ – Параметр скорости синтеза белка PIN1; $k_{pin1}^{SA} (dl)$ – степень влияния салициловой кислоты на эту подсистему (если равна 1, то нет влияния); $S_{pin1} (dl)$ – коэффициент Хилла, который определяет скорость ответа реакции синтеза белка PIN1 на изменения внутриклеточной концентрации ауксина; $q_{pin1}^1 (cu)$ – порог ауксин-зависимой активации синтеза; $q_{pin1}^2 (cu)$ – порог насыщения ауксин-зависимого синтеза.

3. Ауксин зависимый биосинтез белка PIN2 в клетке с индексом (i,j) :

$$k_{pin2}^S * k_{pin2}^{SA} * \frac{\left(\frac{Auxin_{ij}}{q_{pin2}^1}\right)^{S_{pin2}}}{\left(1 + \frac{Auxin_{ij}}{q_{pin2}^2}\right)^{S_{pin2}}},$$

где $k_{pin2}^S (dl)$ – параметр скорости синтеза белка PIN2; $k_{pin2}^{SA} (dl)$ – степень влияния салициловой кислоты на эту подсистему (если равна 1, то нет влияния); $S_{pin2} (dl)$ – коэффициент Хилла, который определяет скорость ответа реакции синтеза белка PIN2 на изменения внутриклеточной концентрации ауксина; $q_{pin2}^1 (cu)$ – порог ауксин-зависимой активации синтеза; $q_{pin2}^2 (cu)$ – порог насыщения ауксин-зависимого синтеза.

4. Ауксин зависимый биосинтез белка PIN3 в клетке с индексом (i,j) :

$$k_{pin3}^S * k_{pin3}^{SA} * \frac{\left(\frac{Auxin_{ij}}{q_{pin3}^1}\right)^{S_{pin3}}}{\left(1 + \frac{Auxin_{ij}}{q_{pin3}^2}\right)^{S_{pin3}}},$$

где $k_{pin3}^S (dl)$ – параметр скорости синтеза белка PIN3; $k_{pin3}^{SA} (dl)$ – степень влияния салициловой кислоты на эту подсистему (если равна 1, то нет влияния); $S_{pin3} (dl)$ – коэффициент Хилла, который определяет скорость ответа реакции синтеза белка PIN3 на изменения внутриклеточной концентрации

ауксина; q^1_{pin3} (cu) – порог ауксин-зависимой активации синтеза; q^2_{pin3} (cu) – порог насыщения ауксин-зависимого синтеза.

5. Ауксин зависимая деградация белка PIN1 в клетке с индексом (i,j):

$$k^d_{pin1} * Pin_1 * \left(1 + \frac{Auxin_{ij}}{q^3_{pin1}}\right)^{h_{pin1}},$$

где k^d_{pin1} (l/tu) – параметр скорости деградации белка PIN1; h_{pin1} (dl) – коэффициент, который определяет эффективность ауксин-зависимой деградации PIN1; q^3_{pin1} (cu) – порог ауксин-зависимой деградации PIN1.

6. Ауксин зависимая деградация белка PIN2 в клетке с индексом (i,j):

$$k^d_{pin2} * Pin_2 * \left(1 + \frac{Auxin_{ij}}{q^3_{pin2}}\right)^{h_{pin2}},$$

где k^d_{pin2} (l/tu) – параметр скорости деградации белка PIN2; h_{pin2} (dl) – коэффициент, который определяет эффективность ауксин-зависимой деградации PIN2; q^3_{pin2} (cu) – порог ауксин-зависимой деградации PIN2.

7. Деградация белка PIN3 в клетке:

$$k^d_{pin3} * Pin_3,$$

где k^d_{pin3} (l/tu) – параметр скорости деградации белка PIN3.

8. Пассивный транспорт (диффузия) ауксина через конкретную мембрану, соединяющую две соседние клетки:

$$D * Auxin_{ij},$$

где D (l/tu) – параметр диффузии через мембрану. Полная диффузия из клетки $4 * D * Auxin_{ij}$

9. Деградация ауксина:

$$k^d_{auxin} * Auxin,$$

где k^d_{auxin} (l/tu) – параметр скорости деградации ауксина.

10. Приток ауксина в модель:

$$k_{alpha},$$

где k_{alpha} (cu/tu) – интенсивность потока ауксина в клетки сосудистой ткани и перицикла на границе зоны меристемы корня (приток в модель).

11. Активный транспорт ауксина посредством PIN1:

$$(k_{pin1}^0 + k_{pin1}^{0,lat}) * Pin1 * Auxin,$$

где k_{pin1}^0 (l/tu) – параметр скорости активного транспорта посредством PIN1 по направлению к кончику корня (вниз); $k_{pin1}^{0,lat}$ (l/tu) – параметр скорости активного транспорта посредством PIN1 по направлению к центральной оси корня. Причём $(k_{pin1}^0 + k_{pin1}^{0,lat}) = 1$. Эти два параметра задают характер деления потока ауксина.

12. Активный транспорт ауксина посредством PIN2:

$$(k_{pin2}^0 + k_{pin2}^{0,lat}) * Pin2_{ij} * Auxin_{ij},$$

где k_{pin2}^0 (l/tu) – параметр скорости активного транспорта посредством PIN2 по направлению к кончику корня (вниз) в клетках кортекса и по направлению вверх в клетках эпидермиса; $k_{pin2}^{0,lat}$ (l/tu) – параметр скорости активного транспорта посредством PIN2 по направлению к центральной оси корня. Причём $(k_{pin2}^0 + k_{pin2}^{0,lat}) = 1$. Эти два параметра контролируют характер деления потока ауксина.

13. Активный транспорт ауксина посредством PIN3:

$$k_{pin3}^0 * Auxin_{ij} * Pin3_{ij},$$

где k_{pin3}^0 (l/tu) – параметр скорости активного транспорта посредством PIN3 по всем направлениям.

3.3.3 Математическое описание модели МСК

На основе 13 моделей элементарных подсистем средствами среды MGSmodeller была собрана математическая модель МСК. Так как в цели моделирования не входит исследование влияний неоднородного потока ауксина в корне и с целью уменьшения затрат вычислительных ресурсов моделировалась только половина корня 5x25 клеток (разделение относительно центральной продольной оси). При этом подразумевается абсолютная симметрия процессов, с учётом граничных условий по линии симметрии.

Модель «МСК:5x25» создана как система ОДУ в 456 уравнений и 35 параметров, сохранена в формате среды моделирования Matlab. Дальнейший анализ модели осуществлялся средствами инженерной среды моделирования Matlab.

В описании модели индексы $[i;j]$ обозначают принадлежность выбранной переменной клетке с таким же индексом (индексация начинается с левого нижнего края моделируемой области). Множество \mathbb{N}_{ij} представляет собой набор индексов существующих соседних клеток по отношению к клетке с индексом $[i;j]$. Соответственно параметр $k_{neighbor_{ij}}$ является мощностью множества \mathbb{N}_{ij} . Таким образом, обобщенная модель «МСК:5x25» в виде системы дифференциальных уравнений имеет следующий вид:

$$\begin{cases} \frac{dPin_{1ij}}{dt} = Fs_{1ij}(Auxin_{ij}) - Fd_{1ij}(Auxin_{ij}) \\ \frac{dPin_{2ij}}{dt} = Fs_{2ij}(Auxin_{ij}) - Fd_{2ij}(Auxin_{ij}) \\ \frac{dPin_{3ij}}{dt} = Fs_{3ij}(Auxin_{ij}) - Fd_{3ij}(Auxin_{ij}) \\ \frac{dAuxin_{ij}}{dt} = k_{Auxin}^{syn} - k_{Auxin}^d * Auxin_{ij} + J_{alpha}(i,j) + D * \sum_{g \in \mathbb{N}_{ij}} (Auxin_g - Auxin_{ij}) + J_{Pin1}(i,j) + J_{Pin2}(i,j) + J_{Pin3}(i,j) \end{cases}$$

Fs_{ij} отвечает за скорость синтеза, а Fd_{ij} за скорость деградации соответствующих белков транспортеров PIN семейства в клетках с индексами $[i;j]$. Формула скорости синтеза для белков PIN1-3 имеют одинаковую структуру:

$$Fs_{\#ij}(Auxin_{ij}) = k_{pin\#}^S * k_{pin\#}^{SA} * \frac{\left(\frac{Auxin_{ij}}{q_{pin\#}^1}\right)^{S_{pin\#}}}{\left(1 + \frac{Auxin_{ij}}{q_{pin\#}^2}\right)^{S_{pin\#}}}$$

Уравнения деградации белков транспортёров PIN1 и PIN2 также имеют похожую структуру

$$Fd_{\#ij}(Auxin_{ij}) = k_{pin\#}^d * Pin_{\#} * \left(1 + \frac{Auxin_{ij}}{q_{pin\#}^3}\right)^{h_{pin\#}}$$

Однако, деградация белка PIN3 не зависит от концентрации фитогормона ауксин:

$$Fd_{3ij}(Auxin_{ij}) = k_{pin3}^d * Pin_{3ij}$$

Поведение гормона ауксина имеет более сложное уравнение:

$$\frac{dAuxin_{ij}}{dt} = k_{Auxin}^{syn} - k_{Auxin}^d * Auxin_{ij} + J_{alpha}(i, j) + D * \sum_{g \in \mathbb{N}_{ij}} (Auxin_g - Auxin_{ij}) + J_{Pin1}(i, j) + J_{Pin2}(i, j) + J_{Pin3}(i, j)$$

$$J_{alpha}(i, j) = \begin{cases} k_{alpha}, & \text{если } i = 25 \text{ и } j = [4; 5] \\ 0, & \text{иначе} \end{cases}$$

$$J_{Pin1}(i, j) = J_{Basal}^{Pin1}(i, j) + J_{Lateral}^{Pin1}(i, j)$$

$$J_{Basal}^{Pin1}(i, j) = \begin{cases} k_{pin1}^{0,basal} * (Pin1_{i+1j} * Auxin_{i+1j} - Pin1_{ij} * Auxin_{ij}), & \text{если } 1 < i < 25 \\ -k_{pin1}^{0,basal} * Pin1_{ij} * Auxin_{ij}, & \text{если } i = 25 \\ k_{pin1}^{0,basal} * Pin1_{i+1j} * Auxin_{i+1j}, & \text{если } i = 1 \end{cases}$$

$$J_{Lateral}^{Pin1}(i, j) = \begin{cases} -k_{pin1}^{0,lat} * Pin1_{ij} * Auxin_{ij}, & \text{если } j = 3 \\ k_{pin1}^{0,lat} * (Pin1_{ij-1} * Auxin_{ij-1} - Pin1_{ij} * Auxin_{ij}), & \text{если } j = 4 \\ k_{pin1}^{0,lat} * Pin1_{ij-1} * Auxin_{ij-1}, & \text{if } j = 5 \quad (\text{компенсация от второй половины}) \end{cases}$$

$$J_{Pin2}(i, j) = J_{Apical}^{Pin2}(i, j) + J_{Basal}^{Pin2}(i, j) + J_{Lateral}^{Pin2}$$

$$J_{Apical}^{Pin2}(i, j) = \begin{cases} -k_{pin2}^{0,apic} * Pin2_{ij} * Auxin_{ij}, & \text{если } i = 1 \text{ и } j = 1 \\ k_{pin2}^{0,apic} * (Pin2_{i-1j} * Auxin_{i-1j} - Pin2_{ij} * Auxin_{ij}), & \text{если } i > 1 \text{ и } j = 1 \end{cases}$$

$$J_{Basal}^{Pin2}(i, j) = \begin{cases} k_{pin2}^{0,basal} * Pin2_{ij} * Auxin_{ij}, & \text{если } i = 25 \\ k_{pin2}^{0,basal} * (Pin2_{i+1j} * Auxin_{i+1j} - Pin2_{ij} * Auxin_{ij}), & \text{если } i < 25 \end{cases}$$

$$J_{Lateral}^{Pin2}(i, j) = \begin{cases} -k_{pin2}^{0,lat} * Pin2_{ij} * Auxin_{ij}, & \text{если } j = 1 \\ k_{pin2}^{0,lat} * (Pin2_{ij-1} * Auxin_{ij-1} - Pin2_{ij} * Auxin_{ij}), & \text{если } j = 2 \\ k_{pin2}^{0,lat} * Pin2_{ij-1} * Auxin_{ij-1}, & \text{если } j = 3 \end{cases}$$

$$J_{Pin3}(i, j) = k_{pin3}^0 \left(\sum_{g \in \mathbb{N}_{ij}} \left(\frac{1}{k_{neighbor_g}} * Auxin_g * Pin3_g \right) - \left. \begin{cases} \left(\frac{2}{3} \right) * Auxin_{ij} * Pin3_{ij}, & \text{если } j = 5 \text{ и } (1 = i) \\ \left(\frac{3}{4} \right) * Auxin_{ij} * Pin3_{ij}, & \text{если } j = 5 \text{ и } (1 < i < 25) \\ Auxin_{ij} * Pin3_{ij}, & \text{in other cases} \end{cases} \right) \right)$$

$$k_{neighbor_{ij}} = \begin{cases} 4, & \text{если клетка } cell_{ij} \text{ имеет 4 соседних клетки} \\ 3, & \text{если клетка } cell_{ij} \text{ имеет 3 соседних клетки} \\ 2, & \text{если клетка } cell_{ij} \text{ имеет 2 соседних клетки} \end{cases}$$

В следствии граничных условий и при учёте потоков от второй половины корня, в уравнениях возникают коэффициенты 1/2, 1/3, 1/4, 2/3, 3/4.

3.3.4 Анализ модели МСК

3.3.4.1 Модель «МСК:5x25.контроль»

Модель «МСК:5x25.контроль» была получена на параметрах, полностью взятых из модели «ММК:4x20.контроль» (см. раздел 3.2.4.5). Недостающий параметр базального синтеза ауксина в клетках k_{auxin}^{syn} был подобран по

критерию формирования точки максимума концентрации ауксина в покое центре корня (в четвертом ряду клеток). Параметры влияния салициловой кислоты на процессы биосинтеза PIN белков $k^{SA}_{pin1}=k^{SA}_{pin2}=k^{SA}_{pin3}=1.0$, т.е. салициловая кислота не оказывает эффекта.

Несмотря на изменившуюся структуру модели, поток ауксина в систему k_{alpha} – остался на прежнем уровне (см. рис. 3.3.2 и таб. 3.3.2). Для модели «МСК:5х25.контроль» была проведена серия вычислительных экспериментов по исследованию устойчивости модели к изменению параметров, соответствующих степени влияния салициловой кислоты. Сводная картина представлена на рис. 3.3.3.

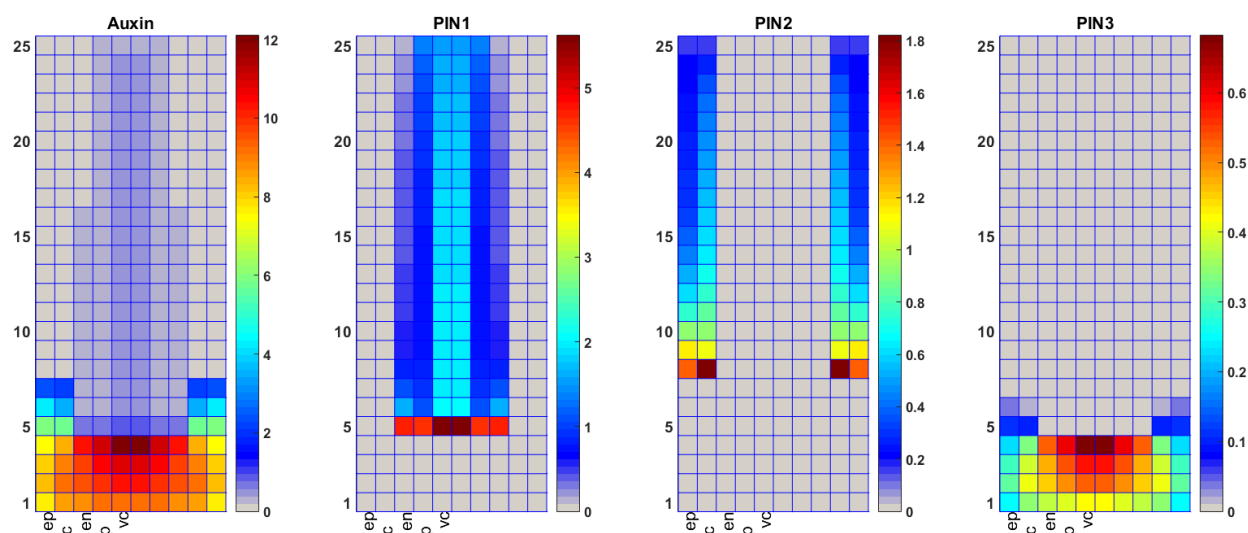


Рисунок 3.3.2 – Распределение ауксина и белков транспортёров PIN1-3 по структуре модели «ММК:5х25.контроль» в состоянии стационара. Параметры модели представлены в таблице 3.3.2. Шкалы в условных концентрационных единицах.

Таблица 3.3.2 – Параметры модели «МСК:5х25.контроль».

	k_{Alpha}	D	k^{syn}	kd	Значения параметров модели «МСК:5х25.контроль»						
					k^0	k^0_{lat}	k^{SA}	k^S	S	kd	$q1$
Auxin	0.5	0.08	0.002	0.005							
Pin1	0.9	0.1	1.0	1000.0	2.0	1000	0.3	10	1.0	6.0	
Pin2	0.7	0.3	1.0	1000.0	2.0	1000	0.1	0.8	0.2	4	
Pin3	1		1.0	1000.0	4	1000	10	10			

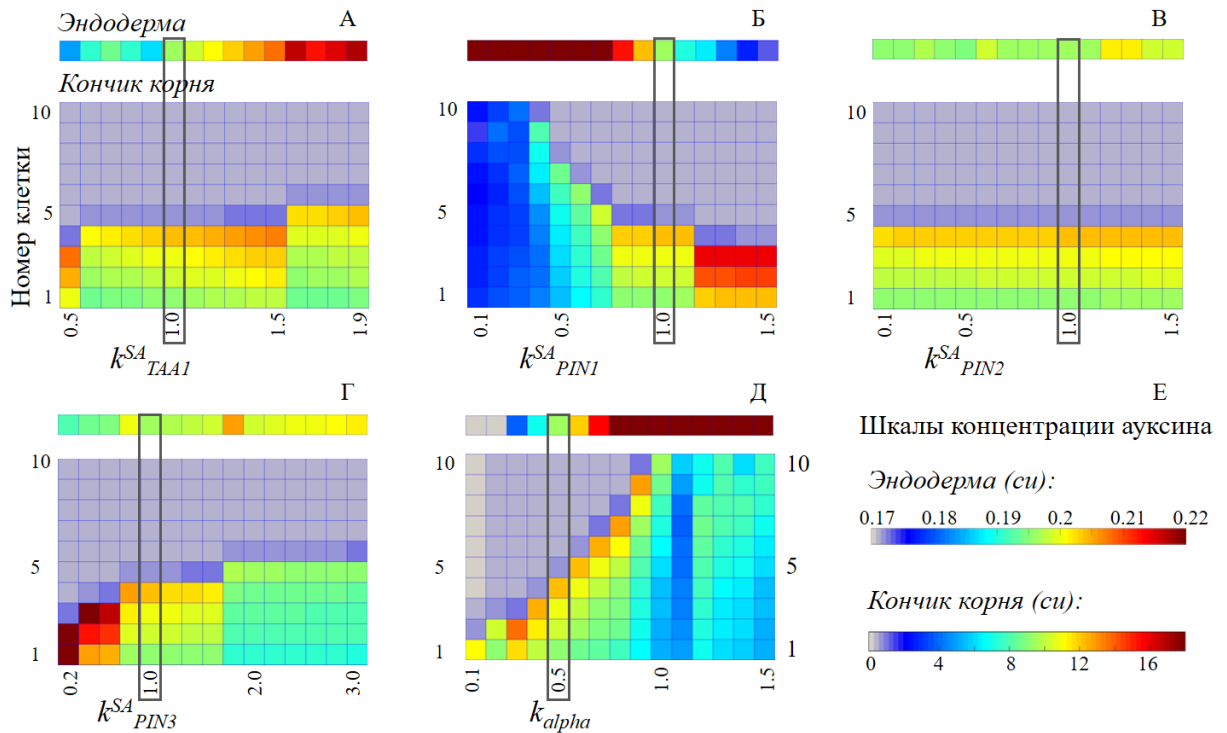


Рисунок 3.3.3 – Распределение концентрации ауксина в стационаре при варьировании параметров модели «МСК:5х25.контроль». По оси X отложены значения варьируемых параметров. По оси Y указаны индексы клеток в ряду $j=5$. В столбце цветом кодирована концентрация ауксина в соответствующей клетке сосудистой ткани. В 4 клетке $j=5$ находится «покоящийся центр» корня. Отдельной клеткой над столбцом обозначено среднее значение концентрации ауксина в клетках с 10 по 20 в слое эндодермы ($j=4$) при выбранном значении параметра. А) Варьирование параметра k_{TAA1}^{SA} , являющегося множителем параметра биосинтеза ауксина в каждой клетке. Обозначает во сколько раз меняется значение параметра k_{auxin}^{syn} . Б) Варьирование параметра k_{PIN1}^{SA} , являющегося характеристикой влияния салициловой кислоты на процесс биосинтеза PIN1. В) Варьирование параметра k_{PIN2}^{SA} , являющегося характеристикой влияния салициловой кислоты на процесс биосинтеза PIN2. Г) Варьирование параметра k_{PIN3}^{SA} , являющегося характеристикой влияния салициловой кислоты на процесс биосинтеза PIN3. Д) Варьирование параметра k_{alpha} , являющегося характеристикой притока ауксина в модель из вне. Е) Шкалы концентраций ауксина (ci) для эндодермы и кончика корня в ряду $j=5$. (Pasternak et al. 2019).

На рисунке 3.3.3 видны границы варьирования параметров, отвечающих за степень влияния салициловой кислоты, в которых максимум концентрации уходит из зоны покоящегося центра (4 ряд клеток). Столбец отражает распределение концентрации ауксина в сосудистом слое с 1 по 10 ряд клеток. Верхняя ячейка отражает среднее значение концентрации ауксина с 10 по 20

клетках в слое эндодермы. Модель «МСК:5х25.контроль» устойчива к изменению интенсивности биосинтеза ауксина в пределах $\pm 50\%$ (рис. 3.3.3.А). Менее устойчива модель к изменению интенсивности биосинтеза белков PIN1 и PIN3 (рис. 3.3.3.Б,Д). Характер активности белка PIN2 не является критическим для изменения положения максимума концентрации ауксина (рис. 3.3.3.В). Наиболее чувствительна модель к изменению параметра k_{alpha} – величине потока ауксина из побега (рис. 3.3.3.Д). Это может быть следствием добавления процессов биосинтеза ауксина в каждую клетку.

3.3.4.2 Модель «МСК:5х25.гетерогенный_эффект».

В предоставленных лабораторией Клауса Пальме экспериментальных данных по воздействию салициловой кислоты наблюдаются следующие эффекты (рис. 3.3.4):

1. Повышение активности маркера ТАА1, отвечающего за биосинтез ауксина, в 3 раза.
2. Повышение сигнала на мембране для белка PIN1 до 130%.
3. Понижение сигнала на мембране для белка PIN2 до 40%.
4. Понижение сигнала на мембране для белка PIN3 до 60%.

Напомню, эти эффекты наблюдаются при сохранении характерного паттерна распределения ауксина с максимумом концентрации в зоне покоящегося центра ниши ствольных клеток корня.

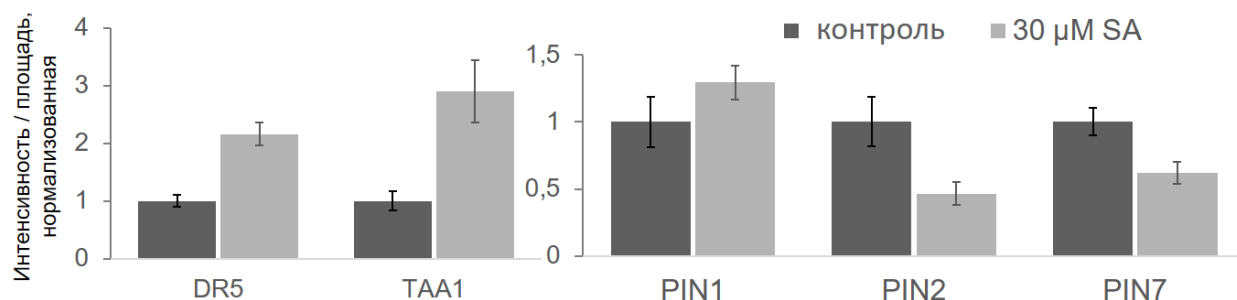


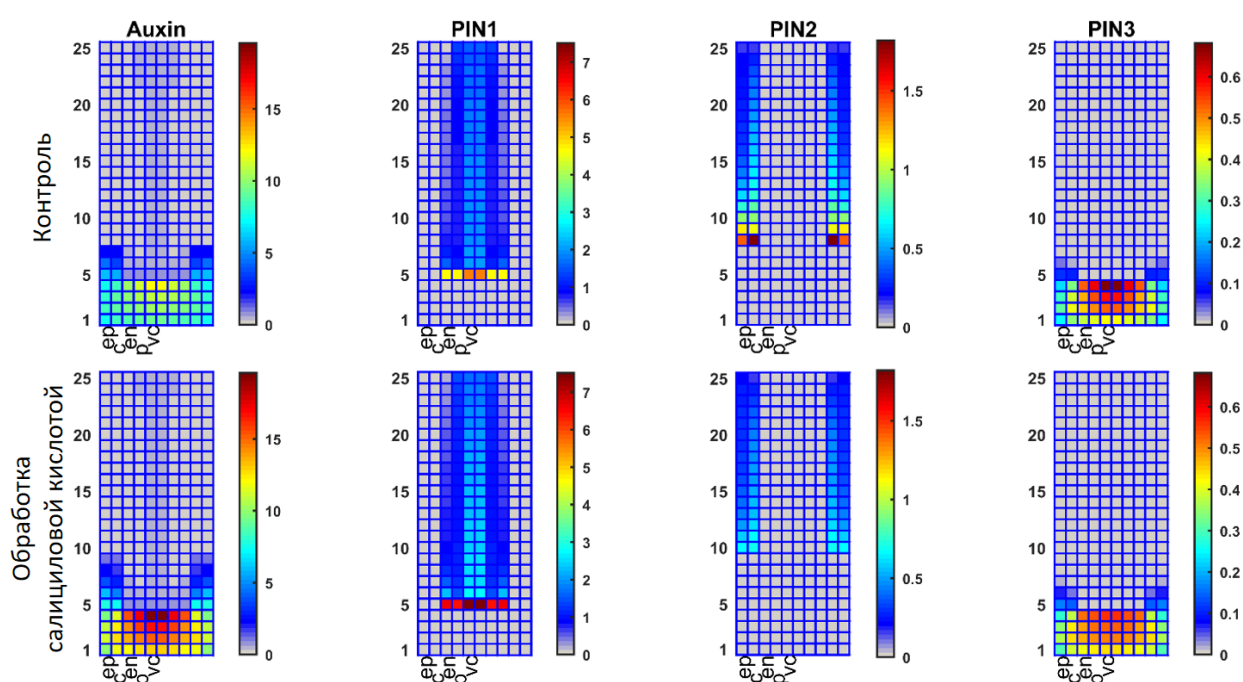
Рисунок 3.3.4 – Количественное измерение интенсивности свечения репортёрной конструкции в кончиках корня после 24ч обработки 30 μM салициловой кислотой (светло серый). Относительное значение флуоресценции было получено с зоны корня отвечающей нише стволовых клеток. DR5 – активность маркера соответствующего концентрации ауксина в клетках. TAA1 – активность маркера биосинтеза ауксина в клетках. PIN1 – уровень экспрессии белка PIN1 в клетках. PIN2 – уровень экспрессии белка PIN2 в клетках. PIN3 – уровень экспрессии белка PIN3 в клетках (Pasternak et al. 2019).

Из диаграммы с варьированием параметров модели «МСК:5x25.контроль» (рис. 3.3.3) видно, что только изменением одного из параметров $K^{SA}_{PIN\#}$ или K^{syn}_{auxin} до измеренных в эксперименте значений невозможно получить паттерн ауксина с максимумом концентрации в зоне покоящегося центра (кроме случая с k^{SA}_{PIN2} , но этот параметр практически не влияет на поведение модели). Изменение значений выбранных параметров в модели «МСК:5x25.контроль» по отдельности будет приводить к смещениям точки максимума концентрации ауксина из зоны «покоящегося центра» корня, что соответствует нарушению функции ниши стволовых клеток корня.

Далее, измеренные в эксперименте изменения характеристик биосинтеза ауксина и активности белков транспортёров (рис. 3.3.4) были введены в модель (параметры k^{syn}_{auxin} и $k^{SA}_{pin\#}$). Получившаяся модель была названа «МСК:5x25.гетерогенный_эффект». Параметры модели представлены в таб. 3.3.3.

Таблица 3.3.3 – Параметры модели «МСК:5х25.гетерогенный эффект».

	k_{alpha}	D	k_{syn}	kd	Параметры модели «МСК:5х25.гетерогенный эффект»					
Auxin	0.4	0.08	0.006	0.005						
	k_0	k^{0-lat}	k^{SA}	k^S	S	kd	$q1$	$q2$	$q3$	h
PIN1	0.9	0.1	1.3	1000.0	2.0	1000	0.3	10	1.0	6.0
PIN2	0.7	0.3	0.4	1000.0	2.0	1000	0.1	0.8	0.2	4
PIN3	1		0.6	1000.0	4	1000	10	10		

**Рисунок 3.3.5** – Распределение ауксина и белков транспортёров в стационаре для моделей «МСК:5х25.контроль» (контроль) и «МСК:5х25.гетерогенный эффект» (результат воздействия салициловой кислоты). Шкалы в условных концентрационных единицах.

Результаты вычислительных экспериментов (рис. 3.3.5) показали совпадение с измеренным в эксперименте поведением ауксина после обработки салициловой кислотой (рис. 3.3.4 DR5).

Также была проверена гипотеза, что наблюдаемое в эксперименте при воздействии салициловой кислоты увеличение концентрации ауксина в клетках корня вызвано не повышением скорости его биосинтеза в каждой

клетке, а только в клетках покоящегося центра. Версия модели, с биосинтезом ауксина только в клетках покоящегося центра, показала, что локальное увеличение биосинтеза ауксина в клетках покоящегося центра не приводит к каким-либо качественным изменениям. Таким образом именно гетерогенный эффект влияния салициловой кислоты на биосинтез ауксина и активности белков транспортёров позволяет получить характерный паттерн ауксина с максимумом в зоне покоящегося центра.

Также был проведён сравнительный анализ контрольных соотношений уровней ауксина и белков PIN модели «МСК:5х25.гетерогенный_эффект» и *in vivo* данных измеренными для соответствующих репортёрных линий после 36 часов воздействия салициловой кислотой (рис 3.3.6). Анализ дал количественное совпадение результатов вычислительных экспериментов экспериментальным данным.

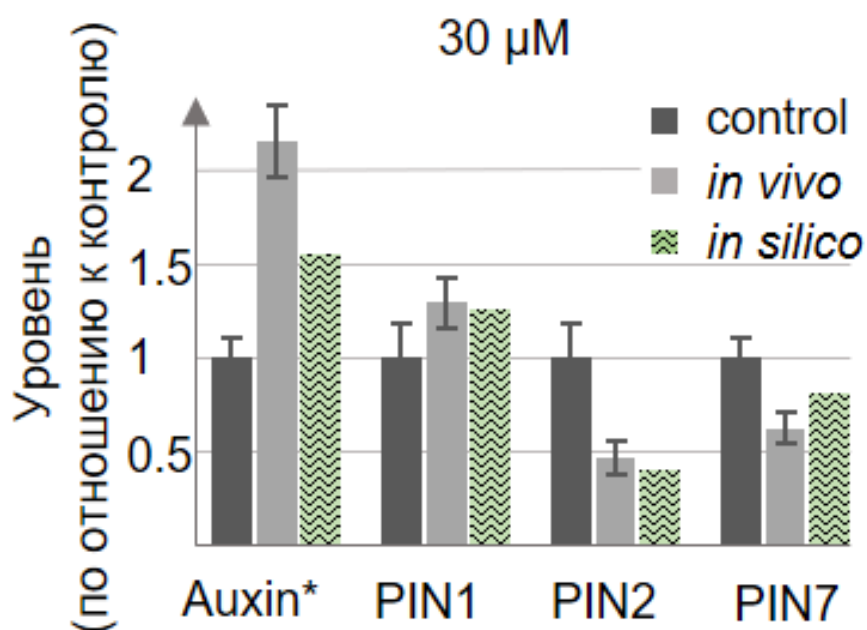


Рисунок 3.3.6 – Диаграмма сравнения стационарного решения модели «МСК:5х25.гетерогенный_эффект» (зелёный) с экспериментальными данными (светло серый) по флюоресценции белка GFP конструкций DR5::GFP (Auxin*), PIN1::PIN1-GFP, PIN2::PIN1-GFP, PIN7::PIN7-GFP в кончиках корней после обработки в течении 36 часов Салициловой Кислотой 30 мкМоль/литр. *in silico* и *in vivo* данные были нормализованы на их контроль (темно серая колонка). Интервалы на серых столбцах означают Стандартное Отклонение. (Pasternak et al. 2019)

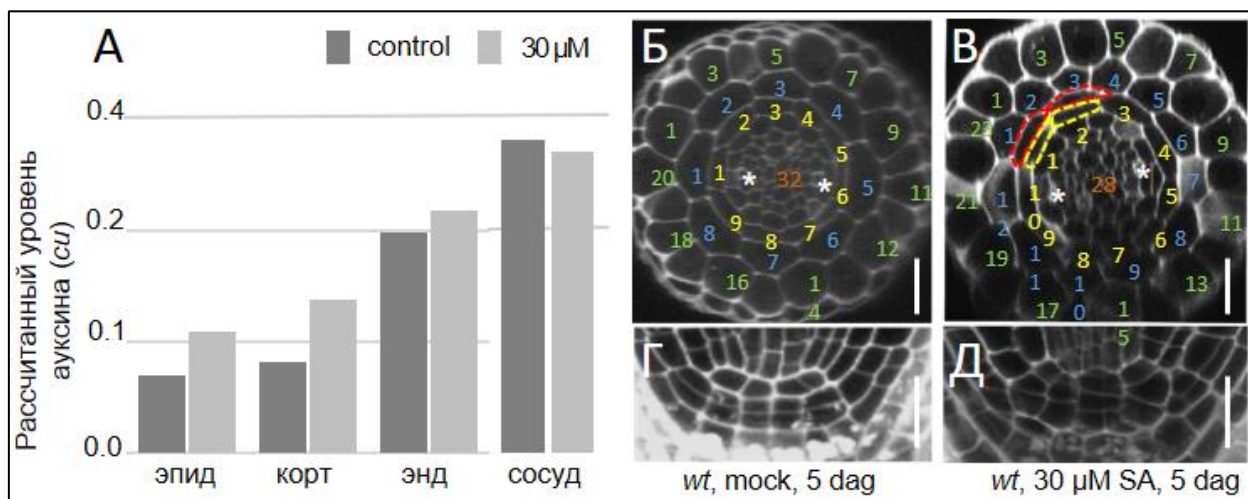


Рисунок 3.3.7 – А) Модель предсказала накопление ауксина во внешних слоях корня (эпидермис, кортекс и эндодерма) и истощение в сосудистой ткани в корнях, обработанных 30 μM салициловой кислоты. Гистограмма показывает среднюю концентрацию ауксина (*ci* – концентрационные единицы) в клетках с 10 по 20 в соответствующем слое клеток. Анатомическая структура корня дикого типа возраста 5 дней после прорастания (Б, Г), в сравнении с корнями того же возраста, но прошедшие обработку 30 μM салициловой кислоты (В, Д). Белыми звёздочками отмечены клетки протофлоэмы. В) Увеличение количества клеточных слоёв эпидермиса, кортекса и эндодермы и уменьшение клеточных слоёв сосудистой ткани. Переклиналильные деление клеток в эндодерме порождают клетки среднего кортекса (красный) и эндодермы (жёлтый). (Pasternak et al. 2019)

Анализ результатов вычислительных экспериментов показал повышение концентрации ауксина в слоях эпидермиса, кортекса и эндодермы, по отношению к контролю, и понижение в сосудистых слоях (рис 3.3.7.А).

Ауксин – морфоген. Если в клетках повышается концентрация морфогена, то можно предположить, что это должно отразиться на развитии корня. Опираясь на данное предположение мы более подробно изучили экспериментальные данные (корни после обработки 30 μM салициловой кислоты) в системе iRoCS Toolbox (Schmidt et al. 2014).

Действительно, при обработке салициловой кислотой на 5 день после прорастания наблюдаются изменения в структуре ниши стволовых клеток корня, а также наблюдаются периклиналильные деления клеток (рис. 3.3.7 Б-Д). В норме, такие деления наблюдаются в тканях корня только в поздних

проростках, на 10 день после прорастания. Также в корнях обработанных салициловой кислотой были выявлены множественные случаи дополнительных делений клеток во внешних слоях корня. В эпидермисе и кортексе дополнительные деления были радиально ориентированы, формируя новые клеточные слои этих тканей (рис. 3.3.7.Б-Д). В дополнение, воздействие 30 μM салициловой кислоты нарушает структурирование клеток инициалей эндодермы-кортекса и приводит к преждевременному формированию дополнительного слоя кортекса (так называемого, middle cortex) (рис. 3.3.7.В). Эктопические периклинальные деления в клетках эндодермы детектируются уже на 36 часу воздействия 30 μM салициловой кислотой в 5 дневных проростках, и их количество постоянно растёт. Радиальное количество клеток эндодермы также увеличивается. Напротив, происходит уменьшение количества слоёв сосудистой системы. Все эти изменения подтверждают предсказания модели.

3.3.5 Заключение по модели МСК

В вычислительных экспериментах на модели «МСК:5x25.гетерогенный_эффект» было показано:

- салициловая кислота влияет на биосинтез ауксина в каждой клетке.
- только увеличение синтеза ауксина в клетках покоящегося центра не дают эффекта, наблюдаемого в эксперименте.
- для достижения наблюдаемых в эксперименте эффектов, недостаточно изменение активности только одного из белков транспортёров, необходимо гетерогенное влияние на активность всех белков транспортёров.
- Предсказано увеличению концентрации ауксина в слоях эпидермиса, кортекса и эндодермы при воздействии низкими концентрациями экзогенной салициловой кислоты. Это увеличение концентрации

морфогена предсказало увеличение интенсивности делений в этих тканях, что было подтверждено экспериментально группой Клауса Пальме в университете города Фрайбург.

Более подробные результаты этого исследования можно найти в опубликованной нами работе (Pasternak et al. 2019).

3.4. Заключение к главе 3

Разработанные инструментарии SiBML и MGSmodeller показали себя с лучшей стороны в исследовании моделей распределения ауксина по структурам тканей корня. Именно благодаря возможности по автоматизации построения моделей получилось оперативно создать серии моделей разных линейных размеров на базе небольшого набора подсистем. Возможности по экспорту математических моделей в формат инструментария Matlab позволили использовать богатую библиотеку методов инженерной среды моделирования в анализе результатов и построения сводных диаграмм. С привлечением суперкомпьютера удалось не только оперативно провести тысячи вычислительных экспериментов, но и выполнить экспертный визуальный анализ результатов моделирования.

В исследовании математической модели распределения фитогормона ауксина по структуре клеток корня посредством белков транспортёров PIN семейства было показано, что ауксин-зависимый биосинтез PIN белков транспортёров обеспечивает самоорганизацию собственных потоков в корне приводя к его накоплению в нише ствольных клеток корня. Объяснено сохранение нормального развития корня в растениях мутантах по генам транспортёрам ауксина (корень способен расти и развиваться) (Mironova et al. 2012).

На расширенной модели в исследовании влияния салициловой кислоты на процессы развития корня показали её гетерогенный эффект на процессы

биосинтеза ауксина и процессы активности всех белков транспортёров ауксина. Выявленное в исследовании модели увеличение концентрации ауксина в слоях эпидермиса, кортекса и эндодермы при воздействии экзогенной салициловой кислоты предсказало увеличение интенсивности делений в этих тканях, что было подтверждено экспериментально группой Клауса Пальме в университете города Фрайбург (Pasternak et al. 2019). Впервые было показано, что даже низкие концентрации (10-50 мкМоль/литр) экзогенной салициловой кислоты влияют на анатомию корня.

Также, расширенная модель – «МСК:5х25» может служить заготовкой для исследования различных аспектов влияния факторов внешней среды на меристему корня. Эта модель уже была использована в исследовании влияния холодового стресса (Hong et al. 2017).

Глава 4. База моделей элементарных подсистем MAMMOTH

В процессе работы с моделями в рамках концепции элементарных подсистем естественным желанием является сохранять их в некотором структурированном хранилище с открытым доступом для коллег. В ИЦиГ СО РАН уже накопилась коллекция из 300 адаптированных к экспериментальным данным моделей элементарных подсистем молекулярно-генетических систем, ассоциированных с несколькими метаболическими путями бактерии *E. Coli* (авторское свидетельство №2012620064 от 16 января 2012). Расширение этой коллекции и повторное использование подсистем в новых моделях позволит постепенно прийти к описанию полной модели всего организма.

В рамках диссертационной работы решалась задача повторного использования разрабатываемых/разработанных моделей (сделанных по концепции «элементарных подсистем») как ресурса в открытом доступе. Следующими критерии были выбраны при разработке:

- Информационный ресурс, включающий коллекцию моделей в виде базы данных.
- Простой формат модели, доступный к пониманию без дополнительных инструментов, или же формат широко используемой среды моделирования (желательно со свободной лицензией).
- Помимо самих моделей, описание содержит некоторую аннотацию, позволяющую получить информацию о формирующих модель элементах, не погружаясь в код, или ассоциированную с моделью публикацию.
- Графический интерфейс пользователя для поиска моделей по их структурной организации и задействованным в них компонентам.
- Модульность – возможность расширения форматов экспорта данных.
- Экспертная валидация данных

- Построение комплексных моделей по набору выбранных элементарных подсистем из базы.
- Программный доступ напрямую к данным, минуя графический интерфейс пользователя.

Опираясь на эти установки, получаем настраиваемые механизмы и критерии поиска по централизованному хранилищу, имея расширенные возможности программного доступа. Основным недостатком существующих ресурсов с коллекциями моделей (Miller et al. 2010; Chelliah et al. 2015) является отсутствие полного контроля над операциями с данными. В результате – ограничения в инструментах поиска подсистем тем, что сделали разработчики.

В качестве системы управления базой данных была выбрана свободно-распространяемая СУБД PostgreSQL (<https://postgresql.org/>), так как она поддерживает работу с информацией в формате XML. В формате XML удобно представить структурную модель подсистемы. В качестве средства разработки графического интерфейса пользователя была выбрана технология Vaadin (vaadin.com). Сервер приложений – Apache Tomcat (tomcat.apache.org). Все выбранные технологии обладают свободной лицензией. Такая конфигурация позволяет и решить поставленную задачу и отдать всю базу третьей стороне для более плодотворного использования данных в их системе.

Разработанный ресурс получил название MAMMOTH – mathematical models of biomolecular systems (<http://mammoth.biomodelsgroup.ru>) (рис. 4.1.1). Он представляет собой инструмент хранения и обработки моделей молекулярно-генетических систем в рамках блочно-модульного подхода.



MAMMOTh

Mathematical Models of biomolecular systems

Selected models About

(a) Structural models search

Name to search

gene

scheme_name

- > NRA
- > Nap
- > NirB
- > NrfA
- > CarA

(б) Structural model description: [GN0000014]

-> NRA

Model consists of:

Alias	Short name	Function	SID
p1	NRA	product	SS0000018
r1	narG		SS0000019
r2	NO3	activator	SS0000016
r3	NO2	activator	SS0000017

Information (press to expand)

(в) Mathematical models

MM0000259

$$\frac{ks \left(k_0 + \left(\frac{r_2}{k_{1no3}} \right)^{h_{1no3}} w_{1no3} + \left(\frac{r_2}{k_{2no3}} \right)^{h_{2no3}} w_{2no3} + \left(\frac{r_3}{k_{no2}} \right)^{h_{no2}} w_{no2} \right)}{1 + \left(\frac{r_2}{k_{1no3}} \right)^{h_{1no3}} + \left(\frac{r_2}{k_{2no3}} \right)^{h_{2no3}} + \left(\frac{r_3}{k_{no2}} \right)^{h_{no2}}}$$

Data:

Name	Value	Units	Description
ks	1	mM/s	rate of NRA complex synthesis
k0	0.01		level of NRA basal synthesis
wno2	0.228		level of nar operon expression in response to nitrite
kno2	3.5	mM	concentration of nitrite defining increase of nar operon expression
hno2	2.8		hill coefficient of nitrite influence on nar operon expression
w1no3	0.5		level of nar operon expression in response to nitrate

Information (press to expand)

(г) Select model

Parameters set

PS0000249

Рисунок. 4.1.1 – Внешний вид интернет приложения MAMMOTh. а) форма поиска. б) раздел биохимического представления подсистем с информацией об участниках. в) Раздел представления математической модели со списком параметров. На верхней кромке размещены кнопки экспорта математической формулы в форматы MathML и LaTeX. г) Раздел выбора набора значений параметров модели. Кнопка «Select model» позволяет отобрать выбранную модель для экспорта в составе комбинированной модели. Каждый из разделов содержит блок с сопроводительной информацией.

Каждая модель в базе имеет следующие три части: (1) Структурное представление (рис. 4.1.1 б), которое содержит сущности, формирующие

процесс и их стехиометрические параметры, а также ссылки на публичные источники (EcoCyc, KEGG, PubChem); (2) математическое представление (рис. 4.1.1 в) с законом протекания процесса и заданными параметрами модели; (3) список конкретных значений параметров модели (рис. 4.1.1 г).

Каждая математическая модель в базе имеет ссылки на сопутствующие публикации, которые послужили основой при разработке модели или описывали экспериментальные данные, к которым адаптировалось динамическое поведение модели. Все математические модели, внесённые в базу, были верифицированы экспертами – специалистами в области молекулярной биологии и математического моделирования динамики генных сетей (Kazantsev et al. 2018).

Большая часть математических моделей, представленных в базе, была реконструирована с помощью методологии Обобщённых функций Хилла (Likhoshvai and Ratushnyi 2007). Данная методология позволяет адаптировать математические модели к существующим кинетическим данным даже при нехватке экспериментальных данных о механизмах конкретных биохимических реакций (Oshchepkova-Nedosekina and Likhoshvai 2007).

Именно факт адаптации к экспериментальным данным даёт преимущество в использовании элементарных подсистем – при конструировании комплексной модели часть параметров будут уже определены.

4.1. База данных

База данных является реляционной (см. рис. 4.1.2), где ряд полей в таблицах представлены в виде XML-структур. Язык XML используется для описания структуры элементарных подсистем, параметров математических моделей и блоков информации.

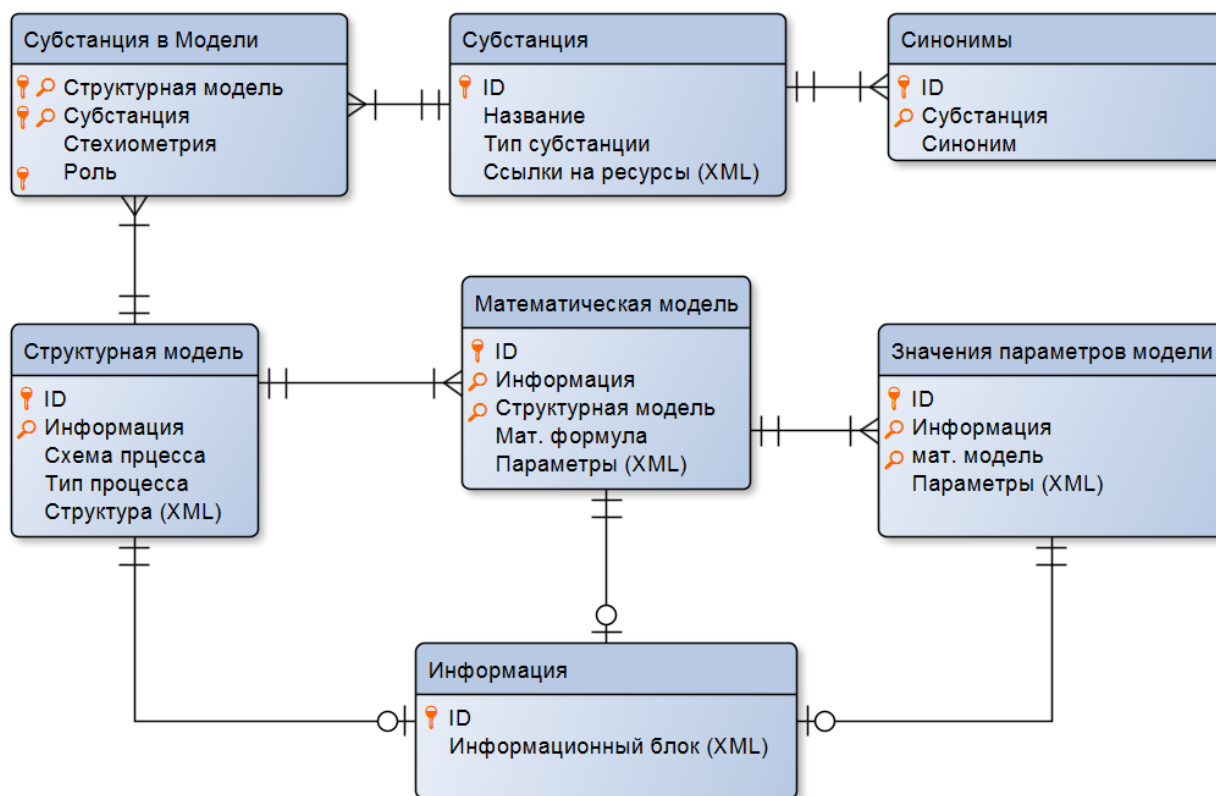


Рисунок. 4.1.2 – Схема отношений основных объектов базы данных MAMMOTh в нотации «вороньи лапки» (Everest 1976). Поля в формате XML содержат основное структурированное описание разделов модели.

Такая организация таблиц позволяет в одном месте хранить данные о связи между участниками структурной модели и ролью этих веществ в математическом представлении, включая ссылки на сторонние банки данных. Структурная модель в данном случае – это базовое описание элементарной подсистемы где описаны участники процесса и их роли: 0 или более субстратов; 0 или более продуктов; 0 или более регуляторов. Причём не допускается ситуация где нет ни одного участника во всех трёх группах. Каждый участник – субстанция, специфицированная именем, набором синонимов и принадлежностью к типу вещества (ген, рнк, белок или низкомолекулярное соединение). Если описывается регулятор, то добавляется параметр – тип регуляции (ген, ингибитор, активатор). В данном контексте ген как регулятор выступает в смысле необходимого условия для протекания процесса и в реакции не расходуется (реакция транскрипции). Каждый

участник подсистемы имеет псевдоним, который используется в математической модели (см. рис. 4.1.1 б).

Математическая модель – это математическое представление закона протекания процесса. Для одной структурной модели допускается наличие нескольких математических (возможны разные скорости и принципы протекания реакций, в зависимости от компартмента локализации процесса). Характеризуется математическим законом (формулой) где прописаны псевдонимы элементов из структурной модели и параметры модели.

Набор параметров – это список параметров модели, специфицированных конкретными значениями. Для одной математической модели может быть несколько наборов параметров (для разных условий среды или для моделирования мутантных подсистем, например).

Элементарная подсистема в данном контексте однозначно определяется тройкой: структурная модель, математическая модель и конкретный набор параметров. В силу того, что в каждом из трёх разделов элементарной подсистемы может быть различное количество сущностей, было решено полное описание представлять в виде структурированного блока текста в формате XML (см. рис. 4.1.3). Этот формат имеет богатый набор стандартных инструментов по обработке/поиску/манипуляции данными и поддерживается средствами поиска СУБД PostgreSQL.

Именно из-за наличия богатого инструментария в качестве формата хранения моделей, был выбран XML, а не SiBML. Задача помещения данных в базу и экспорта из неё разрешена дополнительными инструментами, которые извлекают из XML информацию, генерируя модель в нужном формате (SiBML и Mathematica, например), или наоборот, заносят информацию из Интернет-форм приложения. Использование базы данных даёт дополнительные преимущества в манипуляции структурами. Так как модели построены на основе единого словаря сущностей – реально сравнивать/комбинировать модели осуществляя сравнения только по идентификаторам базы.

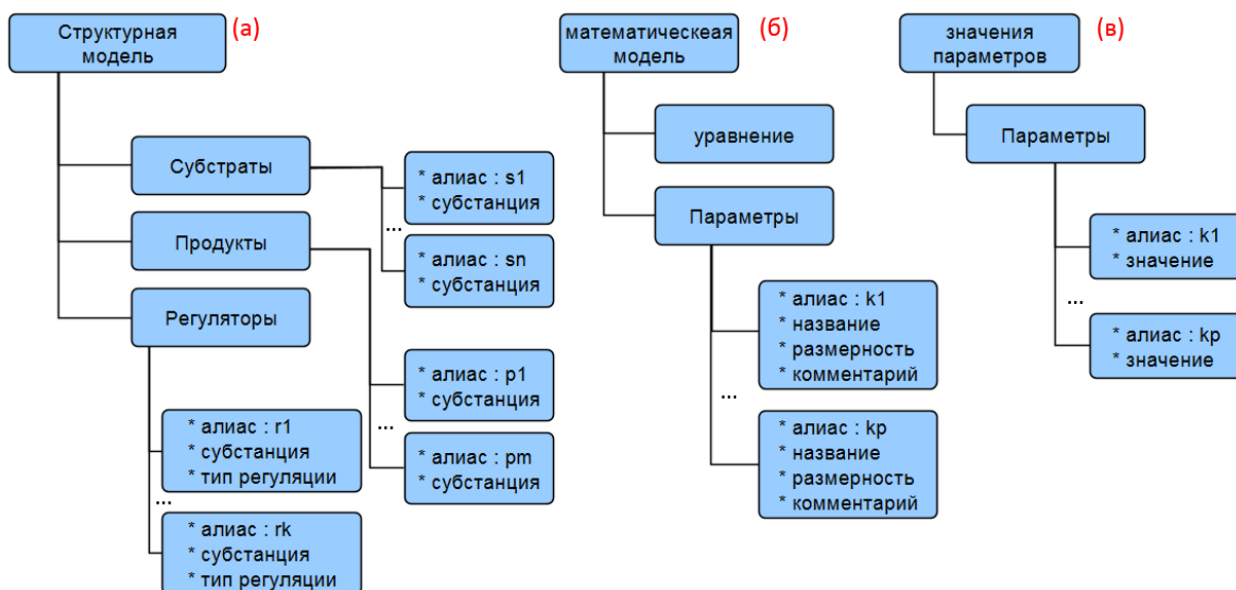


Рисунок. 4.1.3 – Структура данных лежащая в XML полях таблиц. а) Описание структурной модели элементарной подсистемы. б) Описание математической модели элементарной подсистемы. в) Значения параметров математической модели элементарной подсистемы.

В частности, был реализован механизм экспорта комплексной модели, являющейся композицией из серии выбранных моделей «элементарных подсистем» (опция «выбрать модель», см. рис. 4.1.1 г). Экспорт итоговых математических моделей доступен в форматах инженерных сред моделирования Matlab и Mathematica, SiBML и в формализме SBML (Hucka et al. 2015) – международном признанном стандарте представления моделей МГС (sbml.org).

4.2. Графический интерфейс базы данных MAMMOTH

Графический интерфейс пользователя реализован как Интернет приложение, расположенное по адресу *mammoth.biomodelsgroup.ru*. На стартовой странице присутствует форма поиска моделей (см. рис. 4.1.1 а). Поиск осуществляется по принципу совпадения названия молекул, входящих в модели, и их принадлежности к типу биологических субстанций (ген, белок, РНК или низкомолекулярное соединение). При выборе найденной подсистемы

автоматически отображается описание структурной (см. рис. 4.1.1 б), математической (см. рис. 4.1.1 в) частей модели и набора параметров (см. рис. 4.1.1 г). Каждый блок модели содержит вкладку с информацией и ссылками на публикации, основываясь на которых строилась модель.

Ключевая особенность ресурса – возможность выбрать серию моделей и экспортировать их как единую математическую модель в одном из формализмов: Mathematica, MATLAB, SiBML или SBML. Для того что бы поместить выбранную модель в список кандидатов на экспорт необходима выбрать опцию «select model» в разделе параметров (см. рис. 4.1.1 г). Полный список выбранных моделей доступен по ссылке «selected models» вверху главной страницы. При переходе по ссылке появляется форма экспорта моделей, в которой указывается название для модели и формат экспорта. Экспорт модели осуществляется в виде архива, в котором хранятся файлы, формирующие модель в выбранном формализме.

Для пополнения базы моделей MAMMOTh новыми моделями элементарных подсистем МГС не обязательно знать язык SiBML. Реализован инструментарий, который позволяет поэтапно заводить сущности в базе в рамках простых форм:

- Первым этапом служит создание структурной модели, со спецификацией всех участников процесса и всеми ссылками на сторонние источники.
- Вторым этапом создаётся математическая модель – для выбранной структурной модели специфицируется математический закон протекания процесса, с использованием коротких синонимов структурной модели и добавлением параметров, и специфицируется блок информации характеризующий математическую модель, с привязкой к литературе. В автоматическом режиме проверяется корректность введённых данных – информирование о незадействованных реагентах в математическом выражении, а также

выводится список получившихся параметров, которым должны быть заданы размерности.

- Когда специфицирована структурная модель и соответствующая ей математическая, появляется возможность добавить набор конкретных параметров. На этом этапе можно только задать значения параметров конкретной математической модели, указав источник информации, на основе которого были выбраны параметры.

Если же возникает желание отредактировать математическую модель выбранной структурной модели, то это осуществляется через создание новой мат. модели, с удалением предыдущей (и всех ассоциированных с ней объектов наборов параметров), если необходимо. Вся эта процедура выполняется экспертами в специально созданном приложении, в котором указывается кто вносит данные. После дополнительной проверки администратором данные попадают в публичный доступ.

4.3. Программный интерфейс пользователя (API)

Программный интерфейс пользователя реализован как интернет приложение, выполненное по технологии REST API. Благодаря чему возможно обращаться к данным базы из кода программ написанных на Языках Высокого Уровня (C++, Java, Python и др.) или инженерных сред моделирования (Matlab, Mathematica, Scilab, Octave и др.).

REST API представляет собой серию обычных URL адресов. Обращению по такому адресу в ответ получает данные в структурированном виде. Центральная точка входа в интерфейс: *mammoth.biomodelsgroup.ru/api*. При обращении по данному адресу выводится инструкция по использованию API.

API поддерживает следующие сценарии поиска элементарных подсистем:

- По названию молекул, входящих в подсистему

- По названию молекул и её принадлежности к типу (белок, низкомолекулярное соединение, ген ...)
- По роли молекулы в подсистеме (субстрат, продукт или регулятор)
- По количеству участников в подсистеме в одной из ролей. Например, модели только с одним субстратом, или те в которых три продукта.

Такой стиль доступа к данным базы даёт обширные возможности не только в идентификации подсистем, но и в использовании заданных в них математических законов обработки информации для похожих, в смысле структуры, подсистем. Пример использования API в среде программирования Python приведён в приложении 2. Таким образом, существует возможность реализовать собственный программный модуль, который будет обрабатывать данные из базы MAMMOTh по нужному пользователю сценарию.

4.4. Генерация шаблонов элементарных подсистем MGSgenerator

Пополнение базы готовых, адаптированных к экспериментальным данным математических моделей – продолжительный, трудоёмкий процесс, требующий усилия множества экспертов. С другой стороны, накоплены большие коллекции моделей структурной организации молекулярно-генетических систем в таких ресурсах как KEGG (Kanehisa 2000), GeneNet (Ananko et al. 2004), ANDSsystem (Demenkov et al. 2012; Ivanisenko et al. 2015) и многих других (информация о более 700 таких ресурсах для разных организмов собрана на сайте *pathguide.org*). Опираясь на накопленный опыт моделирования (Likhoshvai et al. 2001a; Лихошвай и др. 2003; Фадеев и Лихошвай 2003) был предложен механизм генерации шаблонных математических моделей под известную организацию элементарных подсистем МГС. Разработанный инструмент получил название MGSgenerator (Казанцев и др. 2009). Этот инструмент перерабатывает поданный на вход граф структурной модели МГС и, пробегаая по

формирующим его элементарным подсистемам, дописывает к структуре математический закон, наиболее соответствующий участникам подсистемы. На выходе получается объединённая информация – структурная модель и соответствующая ей математическая. Эта интегральная информация экспортируется или в формате SBML (Hucka et al. 2015), или в формате инструментов проведения вычислительных экспериментов. На сегодня поддерживаются форматы систем Matlab, Mathematica, SiBML/MGSmodeller и STEP+ (Акбердин и др. 2009а). Похожий механизм автогенерации математических моделей был сделан в рамках проекта Path2models (Büchel et al. 2013), в рамках которого ресурс *biomodels.net* (Le Novère 2006) пополнился 140000 математических моделей элементарных подсистем МГС, взятых из БД KEGG.

Инструменты MGSgenerator и MGSmodeller были использованы в исследовании метаболизма ауксина в клетке меристемы побега растения (Акбердин и др. 2009b). В этой работе структурная модель внутриклеточной регуляции метаболизма ауксина сначала была реконструирована в системе GeneNet (Ananko et al. 2004). Модель описывает основной триптофан-зависимый путь биосинтеза ауксина и триптофан-независимый путь его биосинтеза из индола. Модель также включает основные пути конъюгации и деградации ауксина в клетках меристемы побега и корня. Структурная модель, сформированная на основе 245 молекулярно-генетических подсистем (представленных в 8 компартментах), была обработана в инструменте MGSgenerator. MGSgenerator сгенерировал набор математических моделей элементарных молекулярно-генетических подсистем в формате SiBML. Дальнейшее построение комплексной модели и её анализ проводились в инструментари MGSmodeller. Реконструированная математическая модель содержит порядка двухсот параметров, значения большинства которых взяты из литературных источников для *A. thaliana*, кукурузы и риса. Другая часть значений параметров была подобрана численно. В результате было получено соответствие поведения модели экспериментальным данным по содержанию

различных форм ауксина в клетке. Полученная математическая модель позволила исследовать мутантные фенотипы по генам *trp3-1* и *trp2-1* в *in silico* экспериментах (Акбердин и др. 2009b).

4.5. Заключение к главе 4

К концепции «блочно-модульной» реконструкции и исследования моделей молекулярно-генетических систем приходят все большее количество исследовательских групп. Все чаще задумываются о гибридных подходах в моделировании таких систем, где на каждом уровне иерархии описания и функционирования процессов работают алгоритмы с более узкой специализацией. Именно здесь разделение на блоки/модули подсистем позволит распределять процессы по своим алгоритмам, сшивая по определённым временным, событийным, пространственным меткам вещества между этапами алгоритмов.

Поэтому представление данных в информационном ресурсе MAMMOTH (mammoth.biomodelsgroup.ru) в рамках концепции «блочно-модульной» реконструкции и исследования моделей является перспективной точкой роста. Ресурс содержит коллекцию математических моделей элементарных подсистем, адаптированных к экспериментальным данным, и позволяет не только поделиться с сообществом имеющимися моделями, но и служит прекрасным справочным подспорьем в освоении навыков моделирования биологических систем. Используя программный доступ к коллекции (REST API), можно интегрировать базу в цепочки программ по подбору готовых математических моделей (в качестве первого приближения) для новых молекулярно-генетических подсистем по критериям структурной схожести или совпадения в типе взаимодействующих сущностей (рис. 4.5.1).

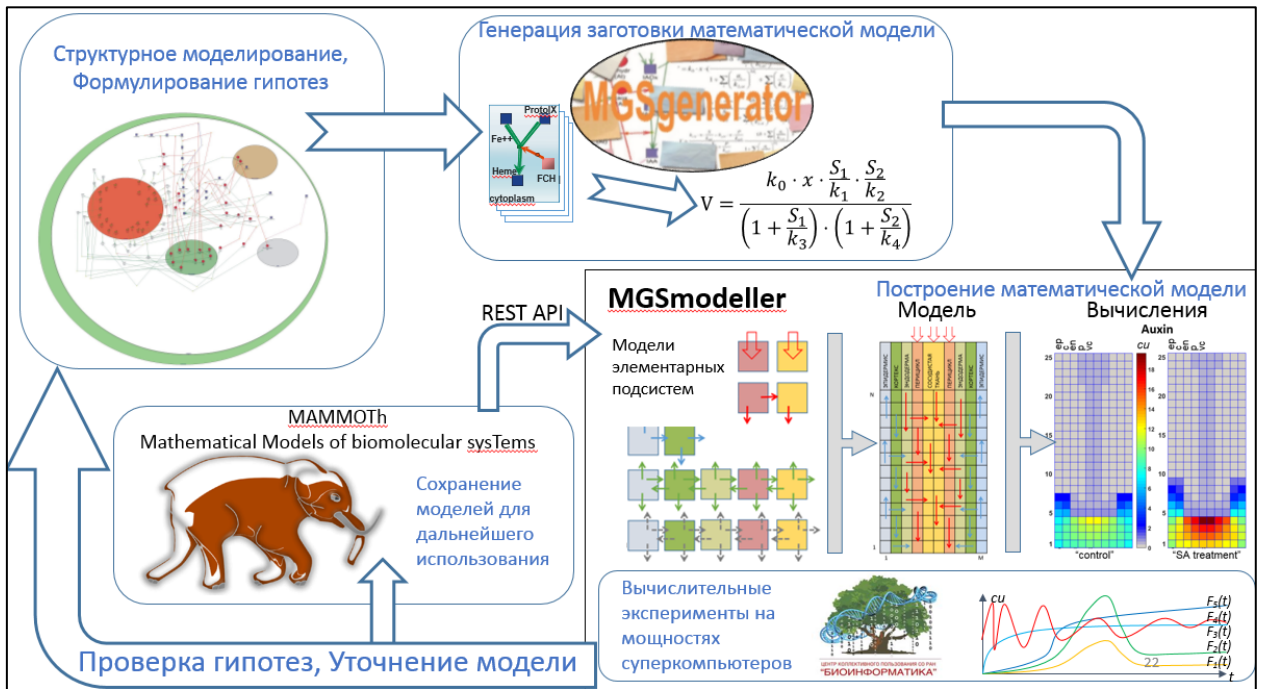


Рисунок 4.5.1 – Схема информационных потоков в интегральном подходе моделирования молекулярно-генетических систем, построенном на наборе разработанных инструментов.

Развивая базу математических моделей и ранжируя данные по принадлежности к организмам, тканям и органеллам клетки во взаимодействии с параллельно развивающимися базами абстрактных моделей «функциональных сетей» (Shen-Orr et al. 2002; Milo et al. 2002; Pieper et al. 2006; Kurata et al. 2013; Fraser et al. 2013; Lim et al. 2013) и подходами автогенерации, придём к интеграции этих подходов. Абстрактные модели, с одной стороны, и согласующиеся с экспериментальными данными модели, с другой, позволят проводить полномасштабное моделирование различных типов клеток, целых организмов *in silico*, исследуя на всех уровнях их иерархической организации влияние внесённых генетических модификации и факторов внешней среды на формирование фенотипических признаков.

Заключение

В рамках диссертационной работы реализован интегральный комплекс из программ MGSmodeller, MGSgenerator и базы данных MAMMOTH, обеспечивающих технологию реконструкции и исследования иерархических математических моделей комплексных многокомпонентных, многокомпаратментных молекулярно-генетических систем. Технология, лежащая в основе интегрального комплекса, базируется на концепции модульной организации объектов живой природы.

Техника итеративного наращивания сложности модели с постепенным включением функциональных блоков, соответствующих новым механизмам протекания процессов, метаболическим путям или компартментам обеспечивается в языке SiBML и среде моделирования MGSmodeller. Реализация поддержки вычислительных экспериментов на современном высокопроизводительном кластере позволила оперативно закрыть запросы 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).

Развитие моделей распределения гормона ауксина по тканям меристемы корня *A. thaliana* с поддержкой больших двумерных ансамблей клеток, схожих по структуре и устройству формирующих потоков с реальной организацией биологического объекта, показали достаточность процессов ауксин-зависимого биосинтеза ключевых белков транспортёров ауксина для самоорганизации потоков гормона в формировании наблюдаемых в эксперименте структур распределений веществ (Mironova et al. 2012). В этом исследовании показано что для поддержания ниши стволовых клеток

меристемы корня достаточно одного транспортёра ауксина, дополнительные транспортёры обеспечивают устойчивость ниши стволовых клеток к изменяющимся условиям среды. Организация моделей в рамках блочно-модульной концепции моделирования позволила расширить модель на большее количество вовлечённых структур корня в исследовании влияния экзогенной салициловой кислоты на процессы биосинтеза ауксина и его ключевых белков транспортёров. В этом исследовании был предсказан и экспериментально подтверждён морфогенетический эффект от низких концентрации салициловой кислоты (Pasternak et al. 2019).

Разработанный ресурс МАММОTh (Kazantsev et al. 2018) с адаптированными к экспериментальным данным математическими моделями обеспечивает доступ к наработанной базе знаний, и может служить опорной точкой для развития новых технологий массовой обработки данных в концепции автоматической аннотации вновь секвенируемых геномов.

Дальнейшее развитие блочно-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.

Выводы

1. Разработан программный комплекс MGSmodeller, ориентированный на поддержку вычислительных экспериментов по реконструкции и исследованию математических моделей молекулярно-генетических систем в терминах формализма SiBML. В рамках данного комплекса решена задача реконструкции и анализа моделей комплексных многокомпарментных биологических систем с применением высокопроизводительных вычислительных техник на суперкомпьютере. Показана возможность интеграции MGSmodeller со сторонними программами для расширения возможностей по анализу данных.
2. Разработан интернет-ресурс MAMMOTh, предоставляющий сервис по накоплению моделей молекулярно-генетических систем, созданных в рамках концепции «элементарных подсистем». Ресурс предоставляет возможность поиска моделей в коллекции с последующим их экспортом в разных форматах. Дополняет его программный комплекс MGSgenerator, генерирующий шаблонные математические модели для подсистем, которые ещё не появились в базе.
3. MGSmodeller, MGSgenerator и MAMMOTh, как интегральная информационно-компьютерная платформа, позволили решить задачи реконструкции и анализа моделей в серии исследований: исследование синтеза биоэтанола и молочной кислоты термофильными бактериями рода *Geobacillus*; исследование регуляторных механизмов поддержания плюрипотентности и дифференцировки эмбриональных стволовых клеток животных (*Mus musculus*); исследование транспортных потоков гормона ауксина в структурах клеток корня растения (*A. thaliana*)
4. С использованием информационно-компьютерной платформы в исследовании моделей меристемы корня *A. thaliana*:
 - Показано что для поддержания ниши стволовых клеток меристемы корня достаточно одного транспортёра ауксина, дополнительные

транспортёры обеспечивают устойчивость ниши стволовых клеток к изменяющимся условиям среды.

- Впервые показана морфогенетическая роль низких концентраций салициловой кислоты (10-50 мкМоль/л) в развитии корня *A.thaliana*.

Список используемой литературы

1. Akberdin IR, Kazantsev F V., Ermak T V., Timonov VS, Khlebodarova TM, Likhoshvai VA (2013) In Silico Cell: Challenges and Perspectives. *Math Biol Bioinforma* 8: . <https://doi.org/10.17537/2013.8.295>
2. Alon U (2003) Biological networks: the tinkerer as an engineer. *Science* 301:1866–1867 . <https://doi.org/10.1126/science.1089072>
3. Alon U (2007) Network motifs: theory and experimental approaches. *Nat Rev Genet* 8:450–461 . <https://doi.org/10.1038/nrg2102>
4. Ananko EA, Podkolodny NL, Stepanenko IL, Podkolodnaya OA, Rasskazov DA, Miginsky DS, Likhoshvai VA, Ratushny A V., Podkolodnaya NN, Kolchanov NA (2004) GeneNet in 2005. *Nucleic Acids Res* 33:D425–D427 . <https://doi.org/10.1093/nar/gki077>
5. Antonovici CC, Boas SEM, Rens EG, Tahir H, Merks RMH (2015) *Multiscale Analysis of Morphogenesis*. Elsevier Ltd.
6. Band LR, Fozard JA, Godin C, Jensen OE, Pridmore T, Bennett MJ, King JR (2012) Multiscale Systems Analysis of Root Growth and Development: Modeling Beyond the Network and Cellular Scales. *Plant Cell* 24:3892–3906 . <https://doi.org/10.1105/tpc.112.101550>
7. Band LR, Wells DM, Fozard JA, Ghetiu T, French AP, Pound MP, Wilson MH, Yu L, Li W, Hijazi HI, Oh J, Pearce SP, Perez-Amador MA, Yun J, Kramer E, Alonso JM, Godin C, Vernoux T, Hodgman TC, Pridmore TP, Swarup R, King JR, Bennett MJ (2014) Systems Analysis of Auxin Transport in the Arabidopsis Root Apex. *Plant Cell* 26:862–875 . <https://doi.org/10.1105/tpc.113.119495>
8. Batt G, Besson B, Ciron P-E, de Jong H, Dumas E, Geiselmann J, Monte R, Monteiro PT, Page M, Rechenmann F, Ropers D (2012) Genetic Network Analyzer: A Tool for the Qualitative Modeling and Simulation of Bacterial Regulatory Networks. pp 439–462
9. Beal J, Nguyen T, Goroehowski TE, Goñi-Moreno A, Scott-Brown J, McLaughlin JA, Madsen C, Aleritsch B, Bartley B, Bhakta S, Bissell M, Castillo

- Hair S, Clancy K, Luna A, Le Novère N, Palchick Z, Pocock M, Sauro H, Sexton JT, Tabor JJ, Voigt CA, Zundel Z, Myers C, Wipat A (2019) Communicating Structure and Function in Synthetic Biology Diagrams. *ACS Synth Biol* 8:1818–1825 . <https://doi.org/10.1021/acssynbio.9b00139>
10. Becker SA, Feist AM, Mo ML, Hannum G, Palsson BØ, Herrgard MJ (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat Protoc* 2:727–738 . <https://doi.org/10.1038/nprot.2007.99>
11. Bennett T, Scheres B (2010) Root development—two meristems for the price of one? *Curr Top Dev Biol* 91:67–102 . [https://doi.org/10.1016/S0070-2153\(10\)91003-X](https://doi.org/10.1016/S0070-2153(10)91003-X)
12. Blilou I, Xu J, Wildwater M, Willemsen V, Paponov I, Friml J, Heidstra R, Aida M, Palme K, Scheres B (2005) The PIN auxin efflux facilitator network controls growth and patterning in *Arabidopsis* roots. *Nature* 433:39–44 . <https://doi.org/10.1038/nature03184>
13. Blinov ML, Faeder JR, Goldstein B, Hlavacek WS (2004) BioNetGen: Software for rule-based modeling of signal transduction based on the interactions of molecular domains. *Bioinformatics* 20:3289–3291 . <https://doi.org/10.1093/bioinformatics/bth378>
14. Bruggeman FJ, Westerhoff H V. (2007) The nature of systems biology. *Trends Microbiol* 15:45–50 . <https://doi.org/10.1016/j.tim.2006.11.003>
15. Büchel F, Rodriguez N, Swainston N, Wrzodek C, Czauderna T, Keller R, Mittag F, Schubert M, Glont M, Golebiewski M, van Iersel M, Keating S, Rall M, Wybrow M, Hermjakob H, Hucka M, Kell DB, Müller W, Mendes P, Zell A, Chaouiya C, Saez-Rodriguez J, Schreiber F, Laibe C, Dräger A, Le Novère N (2013) Path2Models: Large-scale generation of computational models from biochemical pathway maps. *BMC Syst Biol* 7: . <https://doi.org/10.1186/1752-0509-7-116>
16. Campagne F, Neves S, Chang C -w., Skrabanek L, Ram PT, Iyengar R, Weinstein H (2004a) Quantitative Information Management for the Biochemical Computation of Cellular Networks. *Sci Signal* 2004:pl11–pl11 .

- <https://doi.org/10.1126/stke.2482004pl11>
17. Campagne F, Neves S, Chang C, Skrabanek L, Ram PT, Iyengar R, Weinstein H (2004b) Quantitative information management for the biochemical computation of cellular networks. *Sci STKE* 2004:pl11 . <https://doi.org/10.1126/stke.2482004pl11>
 18. Chassagnole C, Noisommit-Rizzi N, Schmid JW, Mauch K, Reuss M (2002) Dynamic modeling of the central carbon metabolism of *Escherichia coli*. *Biotechnol Bioeng* 79:53–73 . <https://doi.org/10.1002/bit.10288>
 19. Chavent M, Duncan AL, Sansom MSP (2016) Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale. *Curr Opin Struct Biol* 40:8–16 . <https://doi.org/10.1016/j.sbi.2016.06.007>
 20. Chelliah V, Juty N, Ajmera I, Ali R, Dumousseau M, Glont M, Hucka M, Jalowicki G, Keating S, Knight-Schrijver V, Lloret-Villas A, Natarajan KN, Pettit JB, Rodriguez N, Schubert M, Wimalaratne SM, Zhao Y, Hermjakob H, Le Novère N, Laibe C (2015) BioModels: Ten-year anniversary. *Nucleic Acids Res* 43:D542–D548 . <https://doi.org/10.1093/nar/gku1181>
 21. Chen WW, Niepel M, Sorger PK (2010) Classic and contemporary approaches to modeling biochemical reactions. *Genes Dev* 24:1861–1875 . <https://doi.org/10.1101/gad.1945410>
 22. Chickarmane V, Troein C, Nuber UA, Sauro HM, Peterson C (2006) Transcriptional Dynamics of the Embryonic Stem Cell Switch. *PLoS Comput Biol* 2:e123 . <https://doi.org/10.1371/journal.pcbi.0020123>
 23. Chylek LA, Harris LA, Faeder JR, Hlavacek WS (2015) Modeling for (physical) biologists: An introduction to the rule-based approach. *Phys Biol* 12:45007 . <https://doi.org/10.1088/1478-3975/12/4/045007>
 24. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, de Hoon MJL (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25:1422–1423 . <https://doi.org/10.1093/bioinformatics/btp163>
 25. Cooling MT, Hunter P, Crampin EJ (2008) Modelling biological modularity with

- CellML. *IET Syst Biol* 2:73–9 . <https://doi.org/10.1049/iet-syb:20070020>
26. Coskun SA, Cicek AE, Lai N, Dash RK, Ozsoyoglu ZM, Ozsoyoglu G (2013) An online model composition tool for system biology models. *BMC Syst Biol* 7: . <https://doi.org/10.1186/1752-0509-7-88>
27. Courtot M, Juty N, Knüpfer C, Waltemath D, Zhukova A, Dräger A, Dumontier M, Finney A, Golebiewski M, Hastings J, Hoops S, Keating S, Kell DB, Kerrien S, Lawson J, Lister A, Lu J, Machne R, Mendes P, Pocock M, Rodriguez N, Villegier A, Wilkinson DJ, Wimalaratne S, Laibe C, Hucka M, Le Novère N (2011) Controlled vocabularies and semantics in systems biology. *Mol Syst Biol* 7:543 . <https://doi.org/10.1038/msb.2011.77>
28. Covert MW, Xiao N, Chen TJ, Karr JR (2008) Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics* 24:2044–2050 . <https://doi.org/10.1093/bioinformatics/btn352>
29. Cowan AE, Mendes P, Blinov ML (2019) ModelBricks—modules for reproducible modeling improving model annotation and provenance. *npj Syst Biol Appl* 5:37 . <https://doi.org/10.1038/s41540-019-0114-3>
30. Csikász-Nagy A (2009) Computational systems biology of the cell cycle. *Brief Bioinform* 10:424–434 . <https://doi.org/10.1093/bib/bbp005>
31. Danino T, Mondragón-Palomino O, Tsimring L, Hasty J (2010) A synchronized quorum of genetic clocks. *Nature* 463:326–330 . <https://doi.org/10.1038/nature08753>
32. Danziger SA, Ratushny A V., Smith JJ, Saleem RA, Wan Y, Arens CE, Armstrong AM, Sitko K, Chen W, Chiang J, Reiss DJ, Baliga NS, Aitchison JD (2014) Molecular mechanisms of system responses to novel stimuli are predictable from public data. *Nucleic Acids Res* 42:1442–1460 . <https://doi.org/10.1093/nar/gkt938>
33. de Jong H (2002) Modeling and Simulation of Genetic Regulatory Systems: A Literature Review. *J Comput Biol* 9:67–103 . <https://doi.org/10.1089/10665270252833208>
34. Demenkov PS, Ivanisenko TV, Kolchanov NA, Ivanisenko VA (2012)

- ANDVisio: A new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem. In *Silico Biol* 11:149–161 .
<https://doi.org/10.3233/ISB-2012-0449>
35. Demidenko G V., Likhoshvai VA (2005) On differential equations with retarded argument. *Sib Math J* 46:417–430 . <https://doi.org/10.1007/s11202-005-0045-7>
36. Demin O, Goryanin I (2008) Kinetic Modelling in Systems Biology
37. Dräger A, Hassis N, Supper J, Schröder A, Zell A (2008) SBMLsqueezer: A CellDesigner plug-in to generate kinetic rate equations for biochemical networks. *BMC Syst Biol* 2: . <https://doi.org/10.1186/1752-0509-2-39>
38. Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–338 . <https://doi.org/10.1038/35002125>
39. Eungdamrong NJ, Iyengar R (2004) Computational approaches for modeling regulatory cellular networks. *Trends Cell Biol* 14:661–669 .
<https://doi.org/10.1016/j.tcb.2004.10.007>
40. Everest G (1976) Basic Data Structure Models Explained with a Common Example. *Proc Fifth Texas Conf Comput Syst* 39–45
41. Faeder JR, Blinov ML, Hlavacek WS (2009) Rule-based modeling of biochemical systems with BioNetGen. *Methods Mol Biol* 500:113–167 .
https://doi.org/10.1007/978-1-59745-525-1_5
42. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson B (2007) A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 3:1–18 .
<https://doi.org/10.1038/msb4100155>
43. Feraru E, Feraru MI, Kleine-Vehn J, Martinière A, Mouille G, Vanneste S, Vernhettes S, Runions J, Friml J (2011) PIN Polarity Maintenance by the Cell Wall in Arabidopsis. *Curr Biol* 21:338–343 .
<https://doi.org/10.1016/j.cub.2011.01.036>
44. Feraru E, Friml J (2008) PIN Polar Targeting. *Plant Physiol* 147:1553–1559 .
<https://doi.org/10.1104/pp.108.121756>

45. Fraser JS, Gross JD, Krogan NJ (2013) From systems to structure: bridging networks and mechanism. *Mol Cell* 49:222–31 . <https://doi.org/10.1016/j.molcel.2013.01.003>
46. Funahashi A, Matsuoka Y, Jouraku A, Morohashi M, Kikuchi N, Kitano H (2008) CellDesigner 3.5: A versatile modeling tool for biochemical networks. *Proc IEEE* 96:1254–1265 . <https://doi.org/10.1109/JPROC.2008.925458>
47. Galdzicki M, Clancy KP, Oberortner E, Pocock M, Quinn JY, Rodriguez CA, Roehner N, Wilson ML, Adam L, Anderson JC, Bartley BA, Beal J, Chandran D, Chen J, Densmore D, Endy D, Grünberg R, Hallinan J, Hillson NJ, Johnson JD, Kuchinsky A, Lux M, Misirli G, Peccoud J, Plahar HA, Sirin E, Stan G-B, Villalobos A, Wipat A, Gennari JH, Myers CJ, Sauro HM (2014) The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. *Nat Biotechnol* 32:545–550 . <https://doi.org/10.1038/nbt.2891>
48. Garg A, Di Cara A, Xenarios I, Mendoza L, De Micheli G (2008) Synchronous versus asynchronous modeling of gene regulatory networks. *Bioinformatics* 24:1917–1925 . <https://doi.org/10.1093/bioinformatics/btn336>
49. Gear CW (1981) Numerical solution of ODE: is there anything left to do? *SIAM Rev* 23:10–24
50. Gilbert D, Heiner M, Jayaweera Y, Rohr C (2017) Towards dynamic genome-scale models. *Brief Bioinform.* <https://doi.org/10.1093/bib/bbx096>
51. Gizzatkulov NM, Goryanin II, Metelkin EA, Mogilevskaya EA, Peskov K V., Demin O V. (2010) DBSolve Optimum: A software package for kinetic modeling which allows dynamic visualization of simulation results. *BMC Syst Biol* 4: . <https://doi.org/10.1186/1752-0509-4-109>
52. Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* 17:333–351 . <https://doi.org/10.1038/nrg.2016.49>
53. Grieneisen VA, Xu J, Marée AFM, Hogeweg P, Scheres B (2007) Auxin transport is sufficient to generate a maximum and gradient guiding root growth.

- Nature 449:1008–1013 . <https://doi.org/10.1038/nature06215>
54. Guido NJ, Wang X, Adalsteinsson D, McMillen D, Hasty J, Cantor CR, Elston TC, Collins JJ (2006) A bottom-up approach to gene regulation. *Nature* 439:856–860 . <https://doi.org/10.1038/nature04473>
55. Harris LA, Hogg JS, Tapia JJ, Sekar JAP, Gupta S, Korsunsky I, Arora A, Barua D, Sheehan RP, Faeder JR (2016) BioNetGen 2.2: Advances in rule-based modeling. *Bioinformatics* 32:3366–3368 . <https://doi.org/10.1093/bioinformatics/btw469>
56. Hastings J, Mains A, Virk B, Rodriguez N, Murdoch S, Pearce J, Bergmann S, Le Novère N, Casanueva O (2019) Multi-Omics and Genome-Scale Modeling Reveal a Metabolic Shift During *C. elegans* Aging. *Front Mol Biosci* 6:2 . <https://doi.org/10.3389/fmolb.2019.00002>
57. Hellerstein JL, Gu S, Choi K, Sauro HM (2019) Recent advances in biomedical simulations: a manifesto for model engineering. *F1000Research* 8:261 . <https://doi.org/10.12688/f1000research.15997.1>
58. Henkel R, Wolkenhauer O, Waltemath D (2015) Combining computational models, semantic annotations and simulation experiments in a graph database. *Database* 2015:1–16 . <https://doi.org/10.1093/database/bau130>
59. Hirashima T, Rens EG, Merks RMH (2017) Cellular Potts modeling of complex multicellular behaviors in tissue morphogenesis. *Dev Growth Differ* 59:329–339 . <https://doi.org/10.1111/dgd.12358>
60. Hogg JS, Harris LA, Stover LJ, Nair NS, Faeder JR (2013) Exact hybrid particle/population simulation of rule-based models of biochemical systems. 1–15 . <https://doi.org/10.1371/journal.pcbi.1003544>
61. Hong JH, Savina M, Du J, Devendran A, Kannivadi Ramakanth K, Tian X, Sim WS, Mironova V V., Xu J (2017) A Sacrifice-for-Survival Mechanism Protects Root Stem Cell Niche from Chilling Stress. *Cell* 170:102-113.e14 . <https://doi.org/10.1016/j.cell.2017.06.002>
62. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, Singhal M, Xu L, Mendes P, Kummer U (2006) COPASI--a COMplex PATHway SIMulator. *Bioinformatics*

- 22:3067–3074 . <https://doi.org/10.1093/bioinformatics/btl485>
63. Huang F, Kemel Zago M, Abas L, van Marion A, Galván-Ampudia CS, Offringa R (2010) Phosphorylation of Conserved PIN Motifs Directs Arabidopsis PIN1 Polarity and Auxin Transport. *Plant Cell* 22:1129–1142 . <https://doi.org/10.1105/tpc.109.072678>
64. Hübner K, Sahle S, Kummer U (2011) Applications and trends in systems biology in biochemistry. *FEBS J* 278:2767–2857 . <https://doi.org/10.1111/j.1742-4658.2011.08217.x>
65. Hucka M, Bergmann FT, Hoops S, Keating SM, Sahle S, Schaff JC, Smith LP, Wilkinson DJ (2015) The Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 1 Core. *J Integr Bioinforma* 12:382–549 . <https://doi.org/10.2390/biecoll-jib-2015-266>
66. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle J, Kitano H (2002) The ERATO Systems Biology Workbench: enabling interaction and exchange between software tools for computational biology. *Pac Symp Biocomput* 461:450–461
67. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novère N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J (2003) The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics* 19:524–531 . <https://doi.org/10.1093/bioinformatics/btg015>
68. Ivanisenko VA, Saik O V., Ivanisenko N V., Tiys ES, Ivanisenko T V., Demenkov PS, Kolchanov NA (2015) ANDSys: an Associative Network Discovery System for automated literature mining in the field of biology. *BMC Syst Biol* 9:S2 . <https://doi.org/10.1186/1752-0509-9-S2-S2>

69. Jacob F (1977) Evolution and tinkering. *Science* (80-) 196:1161–1166
70. Jaenisch R, Young R (2008) Stem Cells, the Molecular Circuitry of Pluripotency and Nuclear Reprogramming. *Cell* 132:567–582 .
<https://doi.org/10.1016/j.cell.2008.01.015>
71. Johnson DS, Mortazavi A, Myers RM, Wold B (2007) Genome-Wide Mapping of in Vivo Protein-DNA Interactions. *Science* (80-) 316:1497–1502 .
<https://doi.org/10.1126/science.1141319>
72. Juty N, Le Novère N, Laibe C (2012) Identifiers.org and MIRIAM Registry: Community resources to provide persistent identification. *Nucleic Acids Res* 40:580–586 . <https://doi.org/10.1093/nar/gkr1097>
73. Kanehisa M (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28:27–30 . <https://doi.org/10.1093/nar/28.1.27>
74. Karr JR, Phillips NC, Covert MW (2014) WholeCellSimDB: a hybrid relational/HDF database for whole-cell model predictions. *Database* (Oxford) 2014:1–8 . <https://doi.org/10.1093/database/bau095>
75. Karr JR, Sanghvi JC, MacKlin DN, Gutschow M V., Jacobs JM, Bolival B, Assad-Garcia N, Glass JI, Covert MW (2012) A whole-cell computational model predicts phenotype from genotype. *Cell* 150:389–401 .
<https://doi.org/10.1016/j.cell.2012.05.044>
76. Kauffman S (2004) A proposal for using the ensemble approach to understand genetic regulatory networks. *J Theor Biol* 230:581–590 .
<https://doi.org/10.1016/j.jtbi.2003.12.017>
77. Kazantsev F V., Akberdin IR, Bezmaternykh KD, Lashin SA, Podkolodnaya NN, Likhoshvai VA (2008) MGSmodeller—a computer system for reconstruction, calculation and analysis mathematical models of molecular genetic system. In: THE SIXTH INTERNATIONAL CONFERENCE ON BIOINFORMATICS OF GENOME REGULATION AND STRUCTURE. p 113
78. Kazantsev F V., Akberdin IR, Lashin SA, Ree NA, Timonov VS, Ratushnyi A V., Khlebodarova TM, Likhoshvai VA (2018) MAMMOTH: A new database for curated mathematical models of biomolecular systems. *J Bioinform Comput Biol*

- 16:1740010 . <https://doi.org/10.1142/S0219720017400108>
79. Kim B, Kim WJ, Kim DI, Lee SY (2015) Applications of genome-scale metabolic network model in metabolic engineering. *J Ind Microbiol Biotechnol* 42:339–348 . <https://doi.org/10.1007/s10295-014-1554-9>
80. Kim WJ, Kim HU, Lee SY (2017) Current state and applications of microbial genome-scale metabolic models. *Curr Opin Syst Biol* 2:10–18 . <https://doi.org/10.1016/j.coisb.2017.03.001>
81. King ZA, Dräger A, Ebrahim A, Sonnenschein N, Lewis NE, Palsson BO (2015) Escher: A Web Application for Building, Sharing, and Embedding Data-Rich Visualizations of Biological Pathways. *PLoS Comput Biol* 11:1–13 . <https://doi.org/10.1371/journal.pcbi.1004321>
82. King ZA, Lu J, Dräger A, Miller P, Federowicz S, Lerman JA, Ebrahim A, Palsson BO, Lewis NE (2016) BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res* 44:D515–D522 . <https://doi.org/10.1093/nar/gkv1049>
83. Kirk PDW, Babbie AC, Stumpf MPH (2015) Systems biology (un)certainties. *Science* (80-) 350:386–388 . <https://doi.org/10.1126/science.aac9505>
84. Kitano H (2002) Systems biology: a brief overview. *Science* 295:1662–4 . <https://doi.org/10.1126/science.1069492>
85. Klimenko AI, Matushkin YG, Kolchanov NA, Lashin SA (2019) Spatial heterogeneity promotes antagonistic evolutionary scenarios in microbial community explained by ecological stratification: a simulation study. *Ecol Modell* 399:66–76 . <https://doi.org/10.1016/j.ecolmodel.2019.02.007>
86. Kozlov K, Samsonov A (2011) DEEP-differential evolution entirely parallel method for gene regulatory networks. *J Supercomput* 57:172–178 . <https://doi.org/10.1007/s11227-010-0390-6>
87. Kurata H, Maeda K, Onaka T, Takata T (2013) BioFNet: Biological functional network database for analysis and synthesis of biological systems. *Brief Bioinform* 15:699–709 . <https://doi.org/10.1093/bib/bbt048>
88. Kwon Y-K, Cho K-H (2007) Analysis of feedback loops and robustness in

- network evolution based on Boolean models. *BMC Bioinformatics* 8:430 .
<https://doi.org/10.1186/1471-2105-8-430>
- 89.Lafita A, Bliven S, Prlić A, Guzenko D, Rose PW, Bradley A, Pavan P, Myers-Turnbull D, Valasatava Y, Heuer M, Larson M, Burley SK, Duarte JM (2019) BioJava 5: A community driven open-source bioinformatics library. *PLOS Comput Biol* 15:e1006791 . <https://doi.org/10.1371/journal.pcbi.1006791>
- 90.Lashin SA, Matushkin YG (2011) Haploid evolutionary constructor: New features and further challenges. In *Silico Biol* 11:125–135 .
<https://doi.org/10.3233/ISB-2012-0447>
- 91.Lashin SA, Suslov V V., Matushkin YG (2010) Comparative modeling of coevolution in communities of unicellular organisms: adaptability and biodiversity. *J Bioinform Comput Biol* 08:627–643 .
<https://doi.org/10.1142/S0219720010004653>
- 92.Le Novère N (2006) BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res* 34:D689–D691 . <https://doi.org/10.1093/nar/gkj092>
- 93.Le Novère N (2015) Quantitative and logic modelling of molecular and gene networks. *Nat Rev Genet* 16:146–158 . <https://doi.org/10.1038/nrg3885>
- 94.Le Novère N, Finney A, Hucka M, Bhalla US, Campagne F, Collado-Vides J, Crampin EJ, Halstead M, Klipp E, Mendes P, Nielsen P, Sauro H, Shapiro B, Snoep JL, Spence HD, Wanner BL (2005) Minimum information requested in the annotation of biochemical models (MIRIAM). *Nat Biotechnol* 23:1509–1515 .
<https://doi.org/10.1038/nbt1156>
- 95.Le Novère N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, Demir E, Wegner K, Aladjem MI, Wimalaratne SM, Bergman FT, Gauges R, Ghazal P, Kawaji H, Li L, Matsuoka Y, Villéger A, Boyd SE, Calzone L, Courtot M, Dogrusoz U, Freeman TC, Funahashi A, Ghosh S, Jouraku A, Kim S, Kolpakov F, Luna A, Sahle S, Schmidt E, Watterson S, Wu G, Goryanin I, Kell DB, Sander C, Sauro H, Snoep JL, Kohn K, Kitano H (2009) The Systems Biology Graphical Notation. *Nat Biotechnol* 27:735–741 . <https://doi.org/10.1038/nbt.1558>

96. Likhoshvai VA, Khlebodarova TM, Ree MT, Kolchanov NA (2010) Metabolic engineering in silico. *Appl Biochem Microbiol* 46:671–687 .
<https://doi.org/10.1134/S0003683810070021>
97. Likhoshvai VA, Matushkin YG, Ratushnyi A V., Anako EA, Ignatieva E V., Podkolodnaya OA (2001a) Generalized chemokinetic method for gene network simulation. *Mol Biol* 35:919–925 . <https://doi.org/10.1023/A:1013254822486>
98. Likhoshvai VA, Matushkin YG, Ratushnyi A V., Anako EA, Ignat'eva E V., Podkolodnaia OA (2001b) A generalized chemical-kinetic method for modeling gene networks. *Mol Biol (Mosk)* 35:1072–1079
99. Likhoshvai VA, Ratushnyi A V. (2007) Generalized Hill Function Method for Modeling Molecular Processes. *J Bioinform Comput Biol* 05:521–531 .
<https://doi.org/10.1142/S0219720007002837>
100. Lim WA, Lee CM, Tang C (2013) Design principles of regulatory networks: searching for the molecular algorithms of the cell. *Mol Cell* 49:202–212 .
<https://doi.org/10.1016/j.molcel.2012.12.020>
101. Lloyd CM, Lawson JR, Hunter PJ, Nielsen PF (2008) The CellML Model Repository. *Bioinformatics* 24:2122–2123 .
<https://doi.org/10.1093/bioinformatics/btn390>
102. Lopez CF, Muhlich JL, Bachman JA, Sorger PK (2014) Programming biological models in Python using PySB. *Mol Syst Biol* 9:646–646 .
<https://doi.org/10.1038/msb.2013.1>
103. Lucas M, Laplace L, Bennett MJ (2011) Plant systems biology: Network matters. *Plant, Cell Environ* 34:535–553 . <https://doi.org/10.1111/j.1365-3040.2010.02273.x>
104. Macklin DN, Ruggero NA, Covert MW (2014) The future of whole-cell modeling. *Curr Opin Biotechnol* 28:111–115 .
<https://doi.org/10.1016/j.copbio.2014.01.012>
105. Malik-Sheriff `Rahuman S., Glont M, Nguyen TVN, Tiwari K, Roberts MG, Xavier A, Vu MT, Men J, Maire M, Kananathan S, Fairbanks EL, Meyer JP, Arankalle C, Varusai TM, Knight-Schrijver V, Li L, Dueñas-Roca C, Dass G,

- Keating SM, Park YM, Buso N, Rodriguez N, Hucka M, Hermjakob H (2019) BioModels—15 years of sharing computational models in life science. *Nucleic Acids Res* 48:D407–D415 . <https://doi.org/10.1093/nar/gkz1055>
106. Mast FD, Ratushny a. V., Aitchison JD (2014) Systems cell biology. *J Cell Biol* 206:695–706 . <https://doi.org/10.1083/jcb.201405027>
107. McCloskey D, Palsson B, Feist AM (2013) Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. *Mol Syst Biol* 9:1–15 . <https://doi.org/10.1038/msb.2013.18>
108. Medley JK, Goldberg AP, Karr JR (2016) Guidelines for Reproducibly Building and Simulating Systems Biology Models. *IEEE Trans Biomed Eng* 63:2015–2020 . <https://doi.org/10.1109/TBME.2016.2591960>
109. Medvedev KE, Alemasov NA, Vorobjev YN, Boldyreva E V., Kolchanov NA, Afonnikov DA (2014) Molecular dynamics simulations of the Nip7 proteins from the marine deep- and shallow-water *Pyrococcus* species. *BMC Struct Biol* 14:23 . <https://doi.org/10.1186/s12900-014-0023-z>
110. Merks RMH, Guravage M, Inze D, Beemster GTS (2011) VirtualLeaf: An Open-Source Framework for Cell-Based Modeling of Plant Tissue Growth and Development. *Plant Physiol* 155:656–666 . <https://doi.org/10.1104/pp.110.167619>
111. Miller AK, Marsh J, Reeve A, Garny A, Britten R, Halstead M, Cooper J, Nickerson DP, Nielsen PF (2010) An overview of the CellML API and its implementation. *BMC Bioinformatics* 11: . <https://doi.org/10.1186/1471-2105-11-178>
112. Miller AK, Yu T, Britten R, Cooling MT, Lawson J, Cowan D, Garny A, Halstead MDB, Hunter PJ, Nickerson DP, Nunns G, Wimalaratne SM, Nielsen PMF (2011) Revision history aware repositories of computational models of biological systems. *BMC Bioinformatics* 12:22 . <https://doi.org/10.1186/1471-2105-12-22>
113. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U (2002) Network motifs: simple building blocks of complex networks. *Science* 298:824–

- 827 . <https://doi.org/10.1126/science.298.5594.824>
114. Mironova V V., Omelyanchuk NA, Novoselova ES, Doroshkov A V., Kazantsev F V., Kochetov A V., Kolchanov NA, Mjolsness E, Likhoshvai VA (2012) Combined in silico/in vivo analysis of mechanisms providing for root apical meristem self-organization and maintenance. *Ann Bot* 110:349–360 . <https://doi.org/10.1093/aob/mcs069>
115. Mironova V V., Omelyanchuk NA, Yosiphon G, Fadeev SI, Kolchanov NA, Mjolsness ED, Likhoshvai VA (2010) A plausible mechanism for auxin patterning along the developing root. *BMC Syst Biol* 4: . <https://doi.org/10.1186/1752-0509-4-98>
116. Mori H (2004) From the Sequence to Cell Modeling: Comprehensive Functional Genomics in Escherichia coli. *J Biochem Mol Biol* 37:83–92 . <https://doi.org/10.5483/BMBRep.2004.37.1.083>
117. Najafi A, Bidkhorji G, Bozorgmehr J, Koch I, Masoudi-Nejad A (2014) Genome Scale Modeling in Systems Biology: Algorithms and Resources. *Curr Genomics* 15:130–159 . <https://doi.org/10.2174/1389202915666140319002221>
118. Novoselova ES, Mironova V V., Khlebodarova TM, Likhoshvai VA (2015) On the distribution of auxin concentrations in root horizontal layer cells. *Russ J Genet Appl Res* 5:293–299 . <https://doi.org/10.1134/S2079059715030120>
119. Novoselova ES, Mironova V V., Omelyanchuk NA, Kazantsev F V., Likhoshvai VA (2013) Mathematical modeling of auxin transport in protoxylem and protophloem of arabidopsis thaliana root tips. *J Bioinform Comput Biol* 11: . <https://doi.org/10.1142/S0219720013400106>
120. Ocone A, Millar AJ, Sanguinetti G (2013) Hybrid regulatory models: a statistically tractable approach to model regulatory network dynamics. *Bioinformatics* 29:910–916 . <https://doi.org/10.1093/bioinformatics/btt069>
121. Olivier BG, Snoep JL (2004) Web-based kinetic modelling using JWS Online. *Bioinformatics* 20:2143–2144 . <https://doi.org/10.1093/bioinformatics/bth200>
122. Omelyanchuk NA, Kovrizhnykh VV, Oshchepkova EA, Pasternak T, Palme K, Mironova VV (2016) A detailed expression map of the PIN1 auxin transporter

- in *Arabidopsis thaliana* root. *BMC Plant Biol* 16:5 .
<https://doi.org/10.1186/s12870-015-0685-0>
123. Oshchepkova-Nedosekina EA, Likhoshvai VA (2007) A mathematical model for the adenylosuccinate synthetase reaction involved in purine biosynthesis. *Theor Biol Med Model* 4:11 . <https://doi.org/10.1186/1742-4682-4-11>
124. Palsson B (2000) The challenges of in silico biology Moving from a reductionist paradigm to one that views cells as systems will necessitate. *18:1147–1150*
125. Pasternak T, Groot EP, Kazantsev F V., Teale W, Omelyanchuk NA, Kovrizhnykh V V., Palme K, Mironova V V. (2019) Salicylic Acid Affects Root Meristem Patterning via Auxin Distribution in a Concentration-Dependent Manner. *Plant Physiol* 180:1725–1739 . <https://doi.org/10.1104/pp.19.00130>
126. Petrasek J (2006) PIN Proteins Perform a Rate-Limiting Function in Cellular Auxin Efflux. *Science* (80-) 312:914–918 .
<https://doi.org/10.1126/science.1123542>
127. Pieper U, Eswar N, Davis FP, Braberg H, Madhusudhan MS, Rossi A, Marti-Renom M, Karchin R, Webb BM, Eramian D, Shen M-Y, Kelly L, Melo F, Sali A (2006) MODBASE: a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res* 34:D291–D295 .
<https://doi.org/10.1093/nar/gkj059>
128. Prlic A, Yates A, Bliven SE, Rose PW, Jacobsen J, Troshin P V., Chapman M, Gao J, Koh CH, Foisy S, Holland R, Rimsa G, Heuer ML, Brandstatter-Muller H, Bourne PE, Willis S (2012) BioJava: an open-source framework for bioinformatics in 2012. *Bioinformatics* 28:2693–2695 .
<https://doi.org/10.1093/bioinformatics/bts494>
129. Quail M, Smith ME, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y (2012) A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics* 13:341 . <https://doi.org/10.1186/1471-2164-13-341>
130. Rahman A, Takahashi M, Shibasaki K, Wu S, Inaba T, Tsurumi S, Baskin TI

- (2010) Gravitropism of *Arabidopsis thaliana* Roots Requires the Polarization of PIN2 toward the Root Tip in Meristematic Cortical Cells. *Plant Cell* 22:1762–1776 . <https://doi.org/10.1105/tpc.110.075317>
131. Rivas-San Vicente M, Plasencia J (2011) Salicylic acid beyond defence: Its role in plant growth and development. *J Exp Bot* 62:3321–3338 . <https://doi.org/10.1093/jxb/err031>
132. Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, Euskirchen G, Bernier B, Varhol R, Delaney A, Thiessen N, Griffith OL, He A, Marra M, Snyder M, Jones S (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* 4:651–657 . <https://doi.org/10.1038/nmeth1068>
133. Rybiński M, Lula M, Banasik P, Lasota S, Gambin A (2012) Tav4SB: integrating tools for analysis of kinetic models of biological systems. *BMC Syst Biol* 6: . <https://doi.org/10.1186/1752-0509-6-25>
134. Sanchez-Osorio I, Ramos F, Mayorga P, Dantan E (2014) Foundations for modeling the dynamics of gene regulatory networks: a multilevel-perspective review. *J Bioinform Comput Biol* 12:1330003 . <https://doi.org/10.1142/S0219720013300037>
135. Scharm M, Wolkenhauer O, Waltemath D (2014) An algorithm to detect and communicate the differences in computational models describing biological systems. <https://doi.org/10.7287/peerj.preprints.640v1>
136. Schmidt T, Pasternak T, Liu K, Blein T, Aubry-Hivet D, Dovzhenko A, Duerr J, Teale W, Ditengou FA, Burkhardt H, Ronneberger O, Palme K (2014) The iRoCS Toolbox - 3D analysis of the plant root apical meristem at cellular resolution. *Plant J* 77:806–814 . <https://doi.org/10.1111/tpj.12429>
137. Shapiro BE, Levchenko A, Meyerowitz EM, Wold BJ, Mjolsness ED (2003) Cellerator: extending a computer algebra system to include biochemical arrows for signal transduction simulations. *Bioinformatics* 19:677–678 . <https://doi.org/10.1093/bioinformatics/btg042>
138. Shapiro BE, Meyerowitz EM, Mjolsness E (2013) Using cellzilla for plant

- growth simulations at the cellular level. *Front Plant Sci* 4:1–9 .
<https://doi.org/10.3389/fpls.2013.00408>
139. Shapiro BE, Mjolsness E (2016) Pycellator: an arrow-based reaction-like modelling language for biological simulations. *Bioinformatics* 32:629–631 .
<https://doi.org/10.1093/bioinformatics/btv596>
140. Shen-Orr SS, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat Genet* 31:64–68 .
<https://doi.org/10.1038/ng881>
141. Smith AM, Xu W, Sun Y, Faeder JR, Marai GE (2012) RuleBender: integrated modeling, simulation and visualization for rule-based intracellular biochemistry. *BMC Bioinformatics* 13 Suppl 8:S3 .
<https://doi.org/10.1186/1471-2105-13-S8-S3>
142. Somogyi ET, Bouteiller JM, Glazier JA, König M, Medley JK, Swat MH, Sauro HM (2015) LibRoadRunner: A high performance SBML simulation and analysis library. *Bioinformatics* 31:3315–3321 .
<https://doi.org/10.1093/bioinformatics/btv363>
143. Song H-S, Cannon W, Beliaev A, Konopka A (2014) Mathematical Modeling of Microbial Community Dynamics: A Methodological Review. *Processes* 2:711–752 . <https://doi.org/10.3390/pr2040711>
144. Sorokin A, Le Novere N, Luna A, Czauderna T, Demir E, Haw R, Mi H, Moodie S, Schreiber F, Villeger A (2015) Systems Biology Graphical Notation: Entity Relationship language Level 1 Version 2. *J Integr Bioinform* 12:264 .
<https://doi.org/10.1038/npre.2011.5902.1>
145. Stajich JE (2007) An Introduction to BioPerl. In: *Plant Bioinformatics*. Humana Press, Totowa, NJ, pp 535–548
146. Stanford NJ, Wolstencroft K, Golebiewski M, Kania R, Juty N, Tomlinson C, Owen S, Butcher S, Hermjakob H, Le Novere N, Mueller W, Snoep J, Goble C (2015) The evolution of standards and data management practices in systems biology. *Mol Syst Biol* 11:851–851 . <https://doi.org/10.15252/msb.20156053>
147. Stoma S, Lucas M, Chopard J, Schaedel M, Traas J, Godin C (2008) Flux-

- based transport enhancement as a plausible unifying mechanism for auxin transport in meristem development. *PLoS Comput Biol* 4: . <https://doi.org/10.1371/journal.pcbi.1000207>
148. Swat MH, Thomas GL, Belmonte JM, Shirinifard A, Hmeljak D, Glazier JA (2012) Multi-Scale Modeling of Tissues Using CompuCell3D. In: *Methods in Cell Biology*. Elsevier Inc., pp 325–366
149. Tanaka H, Dhonukshe P, Brewer PB, Friml J (2006) Spatiotemporal asymmetric auxin distribution: a means to coordinate plant development. *Cell Mol Life Sci* 63:2738–2754 . <https://doi.org/10.1007/s00018-006-6116-5>
150. Teale WD, Paponov IA, Palme K (2006) Auxin in action: signalling, transport and the control of plant growth and development. *Nat Rev Mol Cell Biol* 7:847–859 . <https://doi.org/10.1038/nrm2020>
151. Troják M, Šafránek D, Hrabec J, Šalagovič J, Romanovská F, Červený J (2016) E-Cyanobacterium.org: A Web-Based Platform for Systems Biology of Cyanobacteria. pp 316–322
152. Turner TE, Schnell S, Burrage K (2004) Stochastic approaches for modelling in vivo reactions. *Comput Biol Chem* 28:165–178 . <https://doi.org/10.1016/j.compbiolchem.2004.05.001>
153. Vanneste S, Friml J (2009) Auxin: A Trigger for Change in Plant Development. *Cell* 136:1005–1016 . <https://doi.org/10.1016/j.cell.2009.03.001>
154. Vieten A, Vanneste S, Wisniewska J, Benkova E, Benjamins R, Beeckman T, Luschnig C, Friml J (2005) Functional redundancy of PIN proteins is accompanied by auxin-dependent cross-regulation of PIN expression. *Development* 132:4521–4531 . <https://doi.org/10.1242/dev.02027>
155. Waltemath D, Karr JR, Bergmann FT, Chelliah V, Hucka M, Krantz M, Liebermeister W, Mendes P, Myers CJ, Pir P, Alaybeyoglu B, Aranganathan NK, Baghalian K, Bittig AT, Burke PEP, Cantarelli M, Chew YH, Costa RS, Cursons J, Czauderna T, Goldberg AP, Gomez HF, Hahn J, Hameri T, Gardiol DFH, Kazakiewicz D, Kiselev I, Knight-Schrijver V, Knupfer C, König M, Lee D, Lloret-Villas A, Mandrik N, Medley JK, Moreau B, Naderi-Meshkin H,

- Palaniappan SK, Priego-Espinosa D, Scharm M, Sharma M, Smallbone K, Stanford NJ, Song JH, Theile T, Tokic M, Tomar N, Toure V, Uhlenhof J, Varusai TM, Watanabe LH, Wendland F, Wolfien M, Yurkovich JT, Zhu Y, Zardilis A, Zhukova A, Schreiber F (2016) Toward Community Standards and Software for Whole-Cell Modeling. *IEEE Trans Biomed Eng* 63:2007–2014 . <https://doi.org/10.1109/TBME.2016.2560762>
156. Westerhoff H V., Palsson BO (2004) The evolution of molecular biology into systems biology. *Nat Biotechnol* 22:1249–1252 . <https://doi.org/10.1038/nbt1020>
157. Wisniewska J, Xu J, Seifertová D, Brewer PB, Ruzicka K, Blilou I, Rouquié D, Benková E, Scheres B, Friml J (2006) Polar PIN localization directs auxin flow in plants. *Science* 312:883 . <https://doi.org/10.1126/science.1121356>
158. Xia T, Van Hemert J, Dickerson JA (2011) CytoModeler: A tool for bridging large-scale network analysis and dynamic quantitative modeling. *Bioinformatics* 27:1578–1580 . <https://doi.org/10.1093/bioinformatics/btr150>
159. Yang CR, Shapiro BE, Mjolsness ED, Hatfield GW (2005) An enzyme mechanism language for the mathematical modeling of metabolic pathways. *Bioinformatics* 21:774–780 . <https://doi.org/10.1093/bioinformatics/bti068>
160. Акбердин ИР, Иванисенко НВ, Казанцев ФВ, Ощепкова ЕА, Омелянчук НА, Матушкин ЮГ, Афонников ДА (2014а) Modeling of Regulatory Mechanisms for mESC Self-Renewal: Kinetic and Stochastic Approaches. *Математическая биология и биоинформатика* 9:504–517 . <https://doi.org/10.17537/2014.9.504>
161. Акбердин ИР, Иванисенко НВ, Казанцев ФВ, Ощепкова ЕА, Омелянчук НА, Матушкин ЮГ, Афонников ДА (2014б) Моделирование механизмов регуляции поддержания плюрипотентности эмбриональных стволовых клеток: кинетический и стохастический подходы. *Математическая биология и биоинформатика*
162. Акбердин ИР, Казанцев ФВ, Лихошвай ВА, Фадеев СИ, Гайнова ИА, Королев ВК, Медведев АЕ (2009а) Компьютерная Система Интеграции

- Модулей Для Автоматической Генерации И Численного Анализа Математических Моделей Молекулярно-Генетических Систем. Сибэми 6:440–456
163. Акбердин ИР, Казанцев ФВ, Омелянчук НА, Лихошвай ВА (2009b) Математическое моделирование метаболизма ауксина в клетке меристемы побега растения. Вестник ВОГиС 13:170–175
164. Казанцев ФВ, Акбердин ИР, Безматерных КД, Лихошвай ВА (2009) Система автоматизированной генерации математических моделей генных сетей. Вавиловский Журнал Генетики И Селекции 13:163–169
165. Казанцев ФВ, Акбердин ИР, Подколодный НЛ, Лихошвай ВА (2012) Новые возможности системы MGSmodeller. Вестник Вогис 16:799–804
166. Коврижных ВВ, Омелянчук НА, Пастернак ТП, Миронова ВВ (2014) Ключевая роль PIN-белков в транспорте ауксина в корне *Arabidopsis thaliana* L. Вавиловский Журнал Генетики И Селекции 18:797–806
167. Лихошвай ВА (2008) Математическое моделирование и компьютерный анализ генных сетей. Дис. ... докт. биолог. наук. Новосибирск. 2008. - 378 с
168. Лихошвай ВА, Матушкин ЮГ, Фадеев СИ (2003) Задачи Теории Функционирования Генных Сетей. Сибжим 6:64–80
169. Лихошвай ВА, Омелянчук НА, Миронова ВВ, Казанцев ФВ, Акбердин ИР, Королев ВК, Фадеев СИ, Колчанов НА (2009) Моделирование регуляции ауксином инициации латеральных органов у *Arabidopsis thaliana* L. Вестник Вогис 13:176–185
170. Лихошвай ВА, Омелянчук НА, Миронова ВВ, Фадеев СИ, Мелснесс ЭД, Колчанов НА (2007) Математическая Модель Распределения Ауксина В Корне Растения. Онтогенез 38:446–456
171. Нуриддинов МА, Казанцев ФВ, Розанов АС, Козлов КН, Пельтек СЕ, Акбердин ИР (2013) Математическое моделирование синтеза биоэтанола и молочной кислоты термофильными бактериями рода *Geobacillus*. Вавиловский Журнал Генетики И Селекции 17:686–704
172. Ратнер ВА (2001) Молекулярно-Генетическая Система Управления.

Природа 3:16–22

173. Ратнер ВА (1992) Блочно-модульный принцип организации и эволюции молекулярно-генетических систем управления (МГСУ). Генетика 28:5–13
174. Ратушный АВ, Лихошвай ВА, Ананько ЕА, Владимиров НВ, Гунбин КВ, Лашин СА, Недосекина ЕА, Николаев СВ, Омелянчук ЛВ, Матушкин ЮГ, Колчанов НА (2005) Новосибирская Школа Системной Компьютерной Биологии : Исторический Экскурс И Перспективы Развития. 232–261
175. Рейурд-Смит ВД (1988) Теория формальных языков. Вводный курс. Радио и связь
176. Фадеев СИ, Лихошвай ВА (2003) О Гипотетических Генных Сетях. Сибжим 6:134–153
177. Фадеев СИ, Лихошвай ВА, Когай ВВ, Омелянчук НА (2008) О Математическом Моделировании Паттерна Распределения Ауксина В Корне Растений. Сибжим 48:25–41
178. Хлебодарова ТМ, Когай ВВ, Акбердин ИР, Ри НА, Фадеев СИ, Лихошвай ВА (2013) Моделирование утилизации нитрита клетками *Escherichia coli*: анализ потоков. Математическая Биология И Биоинформатика 8:276–294

Приложение 1. Определение грамматики языка SiBML

Грамматика языка относится к классу контекстно свободных грамматик $LL(1)$, без “е” правил. В описании грамматики введены следующие обозначения: ‘...’ – символ; $\langle \dots \rangle$ – сущность языка (нетерминал). В таблице п1 описаны сущности, представленные в языке SiBML.

Таблица п1. Описание основных сущностей, представленных в грамматике языка SiBML.

Название	описание
\langle допустимое имя \rangle	имя атрибута
\langle допустимое значение \rangle	значение атрибута
\langle список значений \rangle	значение атрибута как список “(a,b,c)”
\langle атрибут \rangle	атрибут языка
\langle лист атрибутов \rangle	список атрибутов, перечисление через пробел и символ запятой.
\langle под лист атрибутов \rangle	вспомогательная сущность для списка атрибутов
\langle элемент \rangle	имя любого элемента SiBML (реагент, константа ...)
\langle список элементов \rangle	список имён
\langle под список \rangle	вспомогательная сущность для списка
\langle именованный список \rangle	именованный список (типа “{pp} ”, “{cc} ” и т.д.)
\langle pp \rangle	список динамических переменных блока
\langle cc \rangle	список параметров блока
\langle ee \rangle	список расширения блока
\langle bb \rangle	список шапки блока
\langle блок \rangle	блок языка SiBML
\langle лист блоков \rangle	список блоков языка SiBML
\langle под лист блоков \rangle	вспомогательная сущность для списка
\langle модель \rangle	
\langle space \rangle	пробельный символ (1 штука) (или символ пробела -‘\s’, или символ табуляции ‘\t’)

Далее представлено формальное описание правил грамматики для каждой сущности из таблицы п1. Выражение вида « \langle сущность \rangle : A/B/C»

означает, что <сущность> может быть представлена выражением или *A*, или *B*, или *C*. Для более наглядного описания перечисление правил может быть со смещением на следующую строку. Выражение ‘ничего’ означает отсутствие символа. Выражение вида «// текст» означает, что *текст* является комментарием к выбранному блоку информации:

```
<допустимый флаг атрибута>:
  [#-+@|$%!^~] <допустимый флаг атрибута>
  | ‘(’
```

```
< допустимое имя атрибута >:
  ‘ничего’
  | [A..Z] <допустимый символ>
  | [a..z] <допустимый символ>
  | ‘_’ <допустимый символ>
  // (пробелы не допустимы)
```

```
<допустимый символ>:
  ‘ничего’
  | [A..Z] <допустимый символ>
  | [a..z] <допустимое значение>
  | [0..9] <допустимое значение>
  | ‘_’ <допустимый символ>
  | <допустимый флаг атрибута>
  | ‘(’
```

```
<допустимое значение>:
  любая последовательность символов не включающая
  символов [‘, ’ ’)’, ‘(’] // пробелы допустимы
```

```
<лист значений>:
  <допустимое значение> <список значений>
  | ‘)’ ‘ ’ // пустой список
```

```
<список значений>:
  ‘,’ <допустимое значение> <список значений>
  | ‘)’ ‘ ’ // окончание списка
```

```
<атрибут>:
  <допустимое имя атрибута> <лист значений>
```

```
<элемент>:
  ‘<’ <лист атрибутов>
```

```
<лист атрибутов>:
  <атрибут> <под лист атрибутов>
  | ‘>’ // пустой элемент
```

```
<под лист атрибутов>:
  ‘,’ <атрибут> <под лист атрибутов>
  | <sp> <атрибут> <под лист атрибутов>
  | ‘>’ // окончание списка
```

```
<именованный список>:
  ‘p’<pp> | ‘c’ <cc> | ‘e’<ee> | b<bb>
```

```

<список элементов>:
  <элемент> <под список>

<под список>:
  <space> <список элементов >
  | ‘;’ ‘\n’ ‘{’ <именованный список >
  | ‘ничего’ // переход на новую строку
  | ‘.’ // окончание списка

<pp>:
  “p} ” <список элементов>
<cc>:
  “c} ” <список элементов>
<ee>:
  “e} ” <список элементов>
<qq>:
  “q} ” <список элементов>
<bb>:
  “b} ” <список элементов>

<блок>:
  <bb> // название блока
  ‘{p’<pp> // список реагентов
  <расширения>

<расширения>:
  ‘{e’<ee> // список расширений блока
  ‘{c’<cc> // список параметров блока
  <расширения>
  | ‘{q’<qq>

<лист блоков>:
  ‘{’<под лист блоков>

<под лист блоков>:
  ‘b’ <блок> <под лист блоков>
  | ‘q’<qq> // признак окончания модели

<модель>:
  <лист блоков>

```

Представленные правила для конструкции <список элементов> порождают цепочки вида:

```

<atr1(znach1), atr2(znach2)>;
<C(cytoplasm), I(TSA1), T(rna), k(k_inflow)> ;
<C(cytoplasm), I(IAA_IGP), T(protein)> ;
<C(nucleus), I(UGT84B1), T(gene)> ;
<t*#(inflow), C(nucleus), I(UGT84B1), T(gene)>.

```

Такие правила определения объектов позволяют не только однозначно восстановить семантику, заложенную в атрибутах, но и надстраивать

автоматические средства анализа и репрезентации данных. Последнее качество является большим преимуществом, так как моделирование динамики объектов разной природы, формы строения и структурных связей требует применение индивидуальных подходов к представлению данных для их комфортного анализа.

Приложение 2. Пример использования МАММОТh API в PYTHON

```

import requests
import json
# URL path to the API
pathToAPI='http://mammoth.biomodelsgroup.ru/api'
searchSyn='atp' # the entity to search
# request a list of available entities in the DB
# by their synonyms
url = pathToAPI+'/substances-syn/'+searchSyn
resp=requests.get(url) # run a request.
synonyms= resp.json() # get the response as JSON object
# The response is an array of JSON structures,
# which comprises the information about synonyms.
# request models where the selected entity is involved.
# substanceId property is the main identifier
for entry in synonyms: print(entry['substanceId']+" is "+entry['name']) # print
the result
### 1) request models for the entity and its role.
# parameter to search. SS0000065 – is an ATP
parameters={'substance':'SS0000065'}
### 2) request models where entity has the substrate role
parameters={'substance': 'SS0000065', 'substrate':1}
### 3) request models where entity has the product role
parameters={'substance': 'SS0000065', 'product':1}
### 4) request models where entity has the modifier role
parameters={'substance': 'SS0000065', 'modifier':1}
### 5) request models where entity has modifier role in # reversible reactions
parameters={'substance': 'SS0000065', 'modifier':1, 'reversible ':1}
### run a request.
requestURL = pathToAPI+'/models-short/entityRole'

# run a request.
resp=requests.get(requestURL, parameters)
entries = resp.json()
# each of the returned entries contains structural model
# identifier and note of the substance role
# print example
for entry in entries:
    print('modelID: '+entry['structuralModelID']+
' .Here ' + entry['substance']
+ ' is ['
+ (' substrate' if entry['substrate'] else '')
+ (' product' if entry['product'] else '')
+ (' modifier' if entry['modifier'] else '')

```

```

    + ']'
  )
  ### requesting model description by structural
  # model identifier
  requestURL = pathToAPI+'/models/GN0000068'
  resp=requests.get(requestURL)          # run a request.
  theModels= resp.json()
  # we may have several mathematical models for single
  # scheme
  for theModel in theModels:
    # print the model ID
    print([''+theModel['gnid']+':'+theModel['mmid']+
          ':'+theModel['psid']+']')
  # print the model scheme
    print(theModel['scheme'])
  # print the model equation
    print(theModel['theMathModel'])

  ### requesting models by their structure. In example,
  # requesting models with one substrate, two products and
  # one modifier.
  parameters={'substrates':2, 'products':1,'modifiers':1}
  requestURL = pathToAPI+'/models-short/entitiesAmount'
  # run a request.
  resp=requests.get(requestURL,parameters)
  shortModels= resp.json()
  # we may have several mathematical models for single
  # scheme
  for theModel in shortModels:
    print('Model: '+theModel['structuralModelID']
          + ' is '
          + ('reversible' if theModel['reversible'] else 'irreversible')
          + ' and has ['
              +str(theModel['substrates'])
          + ':'+str(theModel['products'])
          + ':'+str(theModel['modifiers'])
          + ']'
          + ' [substrates:products:modifiers]'
          )

```