

РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ –
МОСКОВСКАЯ СЕЛЬСКОХОЗЯЙСТВЕННАЯ АКАДЕМИЯ
имени К.А. ТИМИРЯЗЕВА

А.В. Смиряев, Л.К. Панкина

МОДЕЛИРОВАНИЕ ГЕННЫХ СЕТЕЙ

Издательство РГАУ-МСХА
Москва
2013 г.

УДК 577.21.001.573

ББК 28.04в6

C50

Учебное пособие одобрено и рекомендовано методической комиссией агрономического факультета РГАУ-МСХА им. К.А. Тимирязева

Протокол № 23 от 17.03.2008

Рецензент – **А.А. Миронов** доктор биологических наук, профессор (Московский государственный университет им. М.В. Ломоносова)

Смиряев А.В., Панкина Л.К.

C50 Моделирование генных сетей. Учебное пособие. Издание 2-е исправленное – М., ФГОУ ВПО РГАУ – МСХА им. К.А. Тимирязева, 2013. - 52 с.

ISBN 978-5-9675-02224-8

Изложены принципы функционирования и методические основы моделирования генных сетей, регулирующих формирование сложных признаков и процессов прокариот и эукариот. Приведен краткий обзор баз данных, необходимых для построения моделей функционирования генных сетей. Рассмотрены графические, дискретные и непрерывные математические модели. Изложены результаты и перспективы применения этого современного направления моделирования в генетике, медицине и теории эволюции.

Пособие предназначено для подготовки магистров сельского хозяйства по направлению «Агрономия», а также для курсов повышения квалификации специалистов в области генетики, селекции и биотехнологии.

Введение

В последние десятилетия накоплен огромный объем экспериментальных данных о молекулярно-генетических системах. Расшифрованы геномы отдельных видов. Биоинформатика позволила выявлять гены и их функциональные участки. Революция в технологиях (генные, протеомные чипы, и т.д.) дала возможность изучать совместную динамику функционирования многих генов, продуктов их экспрессии, взаимодействий с низкомолекулярными соединениями. Осмысление подобной информации становится актуальной задачей.

В частности, по современным представлениям многие адаптивные признаки определяются целыми блоками коадаптированных генов. Поэтому необходимо исследовать не только функционирование отдельных генов блока, но также их взаимодействие в онтогенезе, регуляцию в стрессовых ситуациях, учитывать влияние мутаций и других факторов эволюции. Но какие именно гены входят в блок, влияющий на признак, каковы механизмы и динамика коадаптации генов при формировании признака в изменчивых условиях среды? Естественным является представление о *генной сети* координировано экспрессирующихся генов, выполняющих общую функцию – формирование признака или реализацию процесса. Однако, строго говоря, все гены экспрессирующиеся до и в процессе формирования сложного адаптивного признака в той или иной степени влияют на этот признак. Необходимо также учитывать все изменения условий среды. Ввиду сложности таких биологических процессов попытка учесть все генетические и средовые факторы при формировании признаков не конструктивна.

Подобные задачи невозможно решать без упрощенных математических моделей и методов анализа данных с привлечением современных информационных технологий. Многие десятилетия в селекционно-генетических исследованиях используют генетико-математические популяционные модели изменчивости количественных признаков. В них учитывается влияние лишь тех генов, которые вызывают изменчивость признака в популяции. Причем взаимовлияние таких генов рассматривается

упрощенно, формально. Остальные же гены, влияющие на формирование количественного признака в онтогенезе, рассматриваются как генетический «фон». С учетом современного представления о сложной генетической и средовой детерминации важных признаков высших организмов, этих моделей уже не достаточно.

Перспективной оказалась концепция *модели генной сети* (далее используется сокращение ГС), т.е. модели функционирования группы координировано экспрессирующихся генов, наиболее важных при выполнении общей функции – формировании признака или процесса (биохимического, физиологического, морфологического, поведенческого и т.п.).

Изучению изложенной в пособии концепции предшествует освоение курсов моделирования в биологии, общей, популяционной и биометрической генетики, а также теории эволюции.

1. Принципы функционирования и модели генных сетей

Можно выделить 5 основных составляющих элементов ГС:

1. *Ядро генной сети* – группа координировано работающих генов.
2. РНК и белки, кодируемые этими генами.
3. Низкомолекулярные компоненты (гормоны и другие сигнальные молекулы), переключающие функции генных сетей в ответ на внешние воздействия, энергетические компоненты, различные метаболиты и т.д.
4. Пути передачи сигналов, обеспечивающие активацию или подавление экспрессии генов.
5. *Отрицательные и положительные обратные связи*. Первые из них стабилизируют количественные характеристики ГС на определенном уровне. Вторые, напротив, отклоняют их от исходных значений, обеспечивая переход к новому функциональному состоянию.

Важнейшее свойство ГС – способность к адекватному изменению динамики функционирования в ответ на изменение условий внешней и внутренней среды. Это достигается за счет сложных взаимодействий

определенных групп веществ-регуляторов с некоторыми сайтами-мишенями в ДНК, РНК и белках. Генные сети регулируются на уровне транскрипции, сплайсинга, трансляции, фосфорилирования, посттрансляционной деградации белков, мембранного транспорта и др. Продукты генной сети одновременно могут являться ее регуляторами.

Так, транскрипция регулируется при воздействии белков на специальные сайты связывания – регуляторные участки ДНК, что приводит к инициализации, подавлению или изменению скорости процесса транскрипции. Поэтому одним из важных вопросов перед построением модели генной сети является нахождение таких регуляторных участков ДНК и изучение механизма их взаимодействия с регуляторами. Последние могут также осуществлять взаимодействие между разными сетями.

Особенностью протекания биохимических реакций в организме является их разделение во времени и пространстве – компоненты ГС могут быть разнесены по различным органам, тканям, клеткам и клеточным компартментам.

В качестве элементов ГС могут выступать не только гены и продукты их экспрессии, но также различные низкомолекулярные соединения, метаболиты, вовлеченные в регуляцию и т.д. Регуляторные воздействия (связи) в зависимости от типа влияния делятся на 4 класса: включение, выключение, усиление или ингибирование процессов.

В настоящее время наиболее разработаны упрощенные модели ГС таких процессов в клетке, как регуляция противовирусного ответа и синтеза холестерина, биосинтез гемоглобина в эритроидной клетке, выработка белков теплового шока, механизм апоптоза и т.п. Подобные модели позволят понять переходы в иерархической схеме организации клетки: ДНК - РНК - белки - ферменты - сети химических реакций - физиология клетки. В перспективе, как полагают специалисты, концепция ГС позволит выстраивать иерархическую модель: локальные ГС более низкого уровня, решая свои частные задачи, обмениваются друг с другом сигналами,

объединяются в сети более высокого уровня и т.д. до модели целого организма.

Пример анализа природной ГС. У растений наиболее глубоко изучена генная сеть развития цветка *Arabidopsis thaliana*, характеризующаяся блочностью и иерархичностью.

Развитие цветка определяется сложной генной сетью (рис. 1), контролирующей формирование меристемы соцветия, образование на ней меристемы цветка, дифференцировку зачатков органов цветка и их спецификацию в органы. Формирование органов цветка детерминируется концентрическими кругами паттернов экспрессии по т.н. ABC-модели (рис. 2). В процессе участвуют группы MADS-генов, активных (или неактивных) в меристеме цветка: А (семейство AP1-подобных белков), В (семейство AP3/PI-подобных белков) и С (семейство AG-подобных белков).

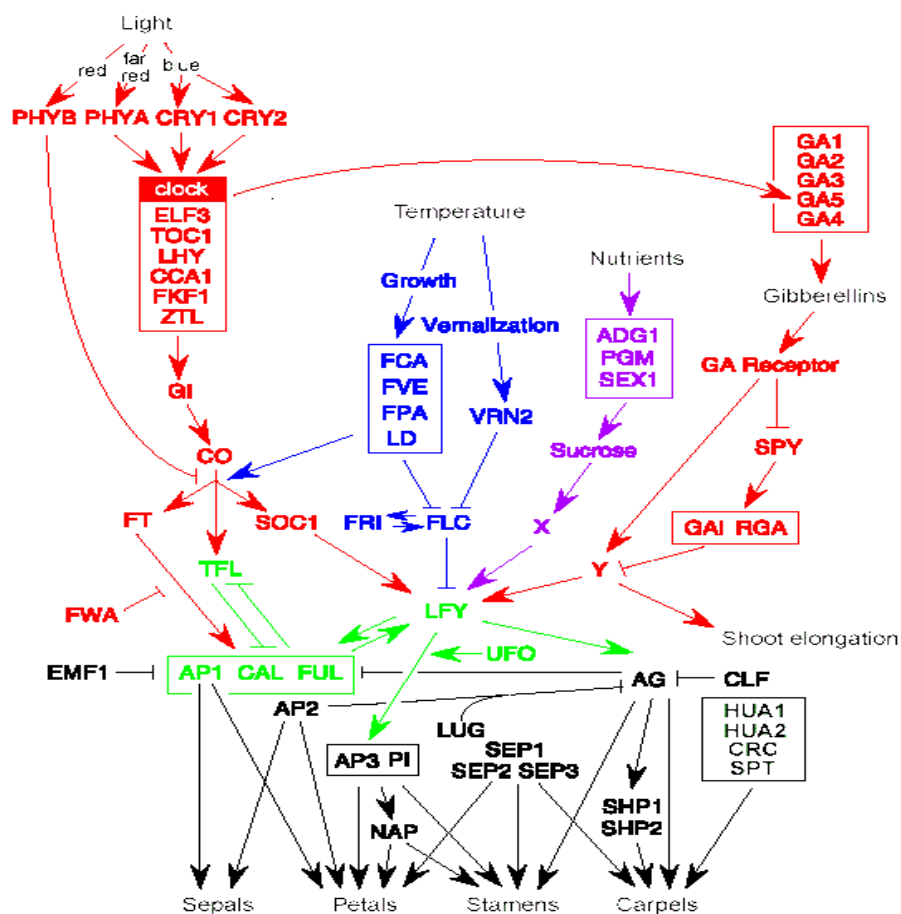


Рис. 1. Генная сеть развития органов цветка (по: Blázquez, 2000).

Обозначения: $\longrightarrow$ – положительные обратные связи, $\longrightarrow|$ - отрицательные

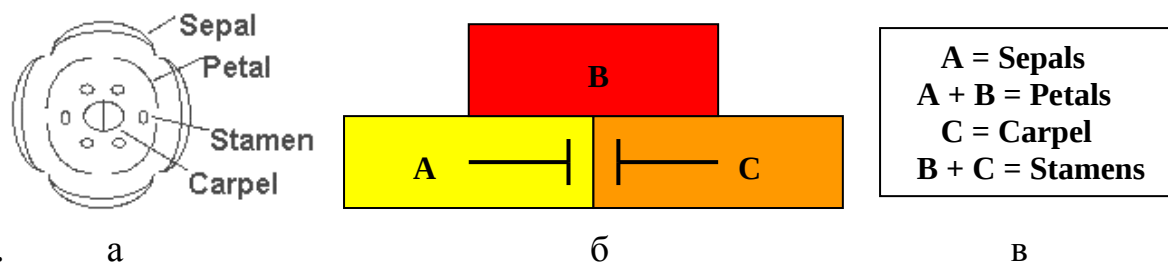


Рис. 2. Закономерности формирования органов цветка в зависимости от активности групп MADS-генов:

а) схема цветка: Sepal – чашелистик, Petal – лепесток, Carpel – пестик, Stamen – тычинка.

б, в) ABC-модель (по: Bowman, Meyerowitz, 1991).

Важнейшими генами ГС, представленной на рис. 1, являются TFL и AP, т.к. группа генов А контролирует переключение от вегетативного развития меристемы побега к генеративному. А именно, активность гена AP1 необходима для индукции цветения; продукт гена TFL подавляет активность гена AP1, что обеспечивает устойчивое вегетативное состояние меристемы побега.

Интегратором всех сигналов, индуцирующих цветение (удлинением светового дня, температура, питательные вещества, гормоны и т.п.), является ген LFY. Этот ген образует контур положительной обратной связи с геном AP1. В результате уровень экспрессии LFY и AP1 быстро растет, ингибируя активность гена TFL, подавляющего активность LFY и AP1 в вегетативной меристеме. Таким образом, подавление активности гена TFL обеспечивает необратимый переход к цветению.

Вопросы:

1. Каковы основные элементы в генных сетях?
2. Поясните смысл ABC-модели формирования цветка?
3. Почему ген LFY можно считать интегратором в генной сети формирования цветка *Arabidopsis thaliana*?

2. Генные сети про- и эукариот

Хотя прямое сравнение эффективности и надежности механизмов регуляции генов эукариот и прокариот не совсем корректно (они используют разные эволюционные стратегии), существуют некоторые особенности их ГС. У прокариот сложность организации в целом коррелирует с размерами

геномов и числом генов, хотя есть исключения. Тенденция роста размера геномов и числа генов прослеживается лишь при переходе от прокариот к эукариотам. У эукариот же отсутствует связь между биологической сложностью, размерами геномов и числом генов (табл. 1). Например, человек и рыба фугу имеют примерно равное количество генов. В чем причина?

Таблица 1. Сравнительная характеристика геномов про- и эукариот (по: Колчанов и др., 2001 с изменениями)

Таксон	Вид	Гаплоидный геном млн. п.н.	Число генов в геноме
Прокариоты			
Микоплазмы	<i>Mycoplasma genitalium</i>	0,58	470
	<i>Mycoplasma pneumoniae</i>	0,82	670
Рикетсии	<i>Rickettsia prowazekii</i>	1,1	834
Археобактерии	<i>Archaeoglobus fulgidus</i>	2,18	2436
	<i>Methanopyrus kandleri</i>	1,69	1738
Цианобактерии	<i>Synechocystis sp.</i>	3,57	3168
Эубактерии	<i>Escherichia coli</i>	4,4-5,5	4288
	<i>Campylobacter jejuni</i>	1,64	1654
	<i>Aquifex aeolicus</i>	1,55	1512
	<i>Neisseria meningitidis</i>	2,27	2121
	<i>Bacillus subtilis</i>	4,2	4100
Низшие эукариоты			
Грибы	<i>Saccharomyces cerevisiae</i>	11,4	6241
	<i>Schizosaccharomyces pombe</i>	13,8	4824
	<i>Aspergillus nidulans</i>	31	
Протисты	<i>Amoeba dubia</i>	670000	
	<i>Entamoeba histolytica</i>	20	
	<i>Dictyostelium discoideum</i>	32	11000
Высшие эукариоты			
Высшие растения	<i>Lilium longiflorum</i>	90000	
	<i>Arabidopsis thaliana</i>	115,7	27540
	<i>Oryza sativa</i>	466	46022-55615
Первичноротые	<i>Caenorhabditis elegans</i>	97	19049
	<i>Drosophila melanogaster</i>	120	13600
Вторичноротые	<i>Protopterus nethiopicus</i>	139000	
	<i>Fugu rubricens</i>	365-400	30000-40000
	<i>Homo sapiens</i>	3500	25000
	<i>Mus musculus</i>	2500	30000

Есть предположение, что усложнение генетических самовоспроизводящихся систем невозможно без роста надежности хранения и копирования генетической информации (Eigen, 1971). В ходе эволюции усложнение прокариот сопровождалось увеличением этой информации.

Прокариоты, как полагают, вплотную подошли к границе т.н. мутационной катастрофы ошибок, что стало препятствием роста их гаплоидных геномов. Поэтому бактерии пошли по пути минимизации некодирующей ДНК. Преодоление границы при дальнейшем усложнении эукариот стало возможным на пути усложнения способов кодирования генетической информации.

Прокариоты чаще используют оперонную структуру генома: основой функционального единства является структурное сцепление генов. Оперон – группа генов с общими регуляторными районами, транскрибирующаяся как один цистрон. Пример оперона из 5-ти генов представлен на рис. 3.

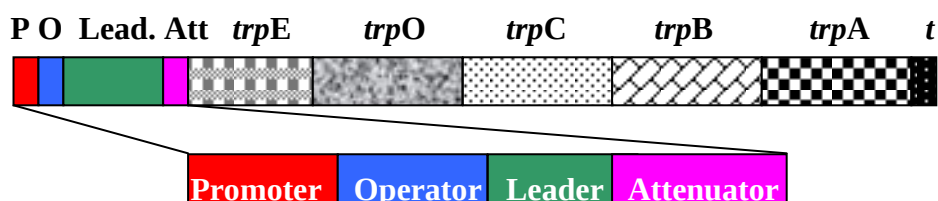


Рис. 3. Триптофановый оперон *Escherichia coli*. Пояснения в тексте (<http://wwwmgs.bionet.nsc.ru/systems/mgl/genenet/>)

Преимуществом оперона является быстрота и четкость регуляции экспрессии. Другое преимущество, вытекающее из его структуры, – экономия на размере ДНК и на числе молекул-регуляторов, т.к. регуляторный район (нижняя часть рис. 3) один для всех генов оперона. Но структура оперона таит и недостатки – затруднения в тонкой регуляции каждого гена и чувствительность к повреждениям: мутация в общем регуляторном районе или мутация ниже старта транскрипции может вывести из строя весь оперон. Конечно, подобные упрощенные оценки нельзя признать исчерпывающими, картина значительно сложнее.

Эукариоты используют оперонную структуру для оптимизации регулирования генных сетей значительно реже, чем прокариоты. Порядок эволюционно стабилен лишь у наиболее жестко корегулируемых генов.

В отличие от генов прокариот регуляторные районы, концентрирующие транскрипцию генов эукариот часто имеют большую

длину и содержат большое количество регуляторных элементов (рис. 4). Так, регуляторный район гена тирозин-аминотрансферазы имеет размер 10000 пар нуклеотидов и содержит более 40 сайтов.

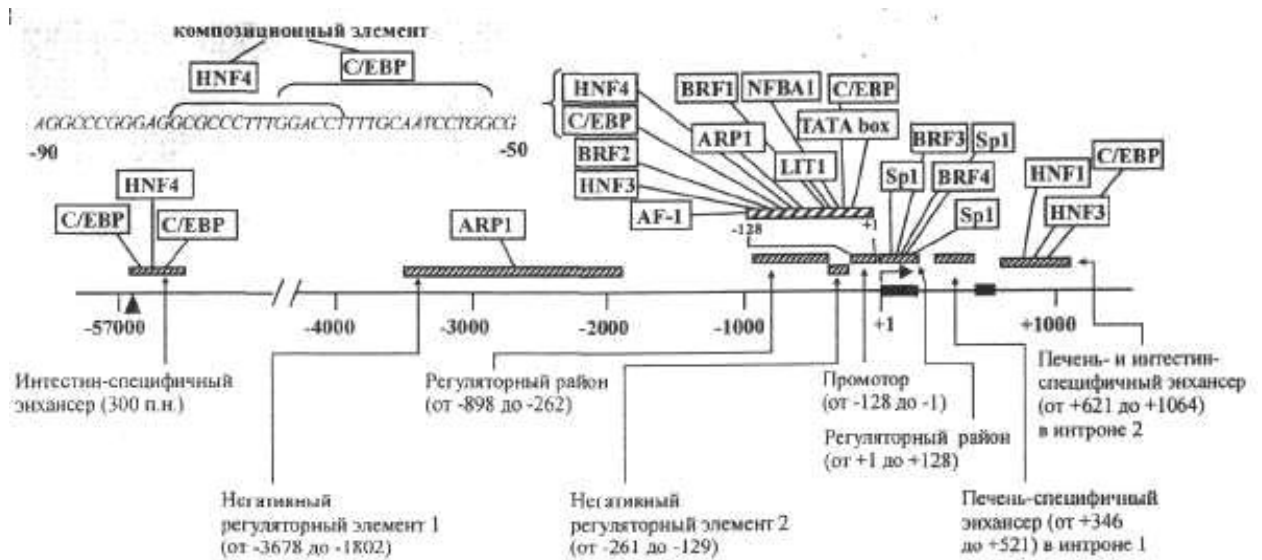


Рис. 4. Фрагмент иерархически организованного регуляторного района гена аполипопротеина В человека (по: Колчанов и др., 2004).

Представим, что регуляторный район гена содержит N сайтов связывания регуляторных белков. Каждый сайт может быть свободным или связанным с регуляторным белком. Тогда максимальное количество состояний регуляторного района равно 2^N .

Другим способом комбинаторного кодирования является альтернативный сплайсинг у эукариот. Комбинаторика экзонов и интронов обеспечивает огромную емкость кодирования генетической информации. Убедительный пример – ген DSCAM дрозофилы, белок которого участвует в формировании нервной системы. На основе альтернативного сплайсинга один этот ген кодирует десятки тысяч вариантов белка. Такой комбинаторный принцип кодирования генетической информации у эукариот – эволюционное приобретение широкого профиля. Он позволяет фактически безгранично наращивать сложность генетических программ без существенного увеличения размеров геномов.

Генные сети про- и эукариот организованы по общим принципам: ядро сети представляют функциональные объединения регулируемых генов. Но

реализация этих принципов различна: прокариоты действуют крайне экономно – функциональное объединение их генов в значительной мере обеспечивается их структурным объединением.

Можно сказать, что перед нами две стратегии развития ГС – экономичность и, как следствие, простота регуляции у прокариот и избыточная, но эволюционно прогрессивная сложность у эукариот. В общем же, генная сеть – это функциональное объединение генов. Структурное их объединение, если оно есть, лишь следствие участия этих генов в общей функции.

На основе компьютерной аннотации экспериментальных данных более чем из 10000 публикаций в Институте цитологии и генетики (ИЦиГ) СО РАН (Новосибирск) создана база данных TRRD (Transcription Regulatory Regions Database). Она содержит описание регуляторных районов генов, контролирующих транскрипцию.

Вопросы:

1. В чем сходство и отличие регуляции в генных сетях про- и эукариот.
2. Что называется опероном у прокариот?
3. Перечислите механизмы комбинаторного кодирования белков у эукариот.

3. Цели и задачи моделирования генных сетей

Представленный ранее пример ГС развития цветка ограничен качественным, логическим анализом. Современный уровень изучения формирования признаков требует также моделей количественной динамики переменных – характеристик ГС. Обобщенно, концентрации продуктов и регуляторов ГС выступают в качестве таких переменных. При моделировании ГС требуется установить качественную и количественную динамику этих процессов во времени и пространстве – в отдельных клетках, участках (блоках, компартментах).

Исследование ГС преследует следующие цели:

- выяснение механизмов работы биологических систем на разных уровнях, начиная с клеточного;

- проверка методов воздействия на биосистему с использованием моделей ГС;

- конструирование искусственных ГС, в частности, с использованием трансгенных технологий;

- изучение влияния различных факторов эволюции на ГС.

Моделирование ГС удешевляет и ускоряет достижение названных целей.

В связи с этим возникают три основных задачи:

- создание компьютерных систем, обеспечивающих удобное для анализа ведение информации о ГС вместе с методами визуализации ее структуры;

- построение адекватных моделей ГС;

- анализ этих моделей (т.е. структуры и динамики важнейших переменных).

Выделяют также прямую и обратную (более сложную) задачи, связанные с моделированием ГС:

- прямая задача: дана структура ГС - предсказать динамику ее переменных во времени;

- обратная задача: из экспериментов известна динамика переменных ГС – построить модель ГС.

Применяются не только математические, компьютерные, но и аналоговые (биологические) модели ГС. Например, выявлена высокая степень консервативности некоторых ГС, выполняющих общие функции у млекопитающих: в процессе эволюции сохраняются не только удачные гены, но и структуры их ГС. Поэтому в качестве модели таких ГС человека иногда используют аналогичные сети мыши.

В качестве примеров можно указать некоторые области применения моделей ГС в медицине и ветеринарии.

1. Изучение патологических процессов, протекающих на клеточном и организменном уровне (это могут быть различные мутации, врожденные либо приобретенные заболевания, синдромы и т.д.). Так, например,

математическая модель системы регуляции внутриклеточного гомеостаза холестерина может быть потенциально применена для изучения таких заболеваний как атеросклероз, гиперхолестеринемия и т.д. Модель ГС регуляции дифференцировки эритроидной клетки пригодна для изучения различных видов анемий и гипоксий.

2. Идентификации генетических и биохимических дефектов и анализ их воздействий на функцию ГС.

3. Развитие оптимальных методов воздействия на исследуемые биологические системы для нормализации их функции, например, при разработке новых терапевтических, либо фармакологических средств.

Следует учесть, что фармакологическое воздействие, обеспечивающее нормализацию функции ГС по ее критической переменной, может вызывать существенные отклонения от нормы по другим переменным, то есть обладать нежелательными побочными эффектами. Коррекцию нарушения функционирования ГС следует признать допустимой, если в результате не только восстановится основная функция ГС, но и побочные эффекты окажутся минимальными. Следовательно, задача фармакологической коррекции функционирования ГС является многокритериальной задачей оптимального управления.

Часто ГС должна оставаться в равновесии по приемлемым значениям нескольких переменных – в *стационарном состоянии*. В этом случае необходимо с помощью модели выяснить существует ли вообще в «мутантной» системе (патологическом варианте ГС) стационарное состояние, соответствующее норме, или близкое к ней. Затем найти допустимое управление (фармакологическое воздействие), переводящее основные переменные поврежденной ГС в это состояние.

В прикладном смысле необходимо с помощью моделей ГС научиться определять, какие воздействия (генно-инженерные, физиологические и пр.) и их комбинации приведут к желаемому признаку – фенотипу.

Вопросы:

1. Назовите цели и задачи моделирования генных сетей.

2. В чем причина консервативности генных сетей у разных видов?
3. Укажите возможные области практического применения моделей генных сетей.

4. Классификация генных сетей и их составляющие

Качественный анализ работы генных сетей позволяет выделить несколько их классов.

1) ГС, обеспечивающие гомеостаз организма, обычно содержат контуры отрицательной обратной связи (ОС), обеспечивающие возвращение контролируемых (критических) переменных к норме. Пример: ГС регуляции уровня холестерина в клетке. Если уровень слишком высок, то активация его предшественника (preSREBP) замедляется самими избыточными молекулами холестерина. Уровень внутриклеточного холестерина постепенно возвращается к норме.

2) ГС, обеспечивающие адаптивный ответ организма на резкие изменения внешних условий. Подобные ГС часто имеют двухэтапную схему работы: сначала происходит быстрая активация процессов за счет положительной ОС, затем включается отрицательная ОС, которая подавляет ответ. В результате контролируемые переменные ГС (концентрации каких-либо метаболитов) поддерживаются на определенном уровне. Примеры: ГС противовирусного ответа и ГС теплового шока.

Так, при действии теплового шока усиливается экспрессия многих генов по принципу положительной ОС. Ключевой транскрипционный фактор ответа на тепловой шок *HSF1* (heat shock factor 1) в неактивном состоянии находится в цитоплазме в комплексе с молекулярными шаперонами, например, *HSP90* (heat shock protein 90). При действии теплового шока появляются частично денатурированные, неправильно свернутые белки. Комплекс распадается. *HSP90* связывается с белками и способствует их правильному свертыванию. Фактор *HSF1* активируется, затем транспортируется в ядро, где связывается с *HSE*-сайтами и активирует транскрипцию генов теплового шока, являющимися в большинстве молекулярными шаперонами. При избытке шаперонов, что регулируется

концентрацией в клетке неправильно свернутых белков, *HSP70* связывается с *HSF1*, который транспортируется в цитоплазму. То есть срабатывает отрицательная ОС и ответ на тепловой шок завершается.

3) ГС, обеспечивающие циклические процессы. Пример: клеточный цикл. В подобных ГС наблюдается четкий баланс между положительными и отрицательными обратными связями.

4) ГС необратимых процессов, контролирующие рост и развитие организмов, дифференцировку клеток, морфогенез органов и отмирание тканей. Они содержат контуры положительной ОС, обеспечивающие эффективное отклонение основной переменной от исходного значения. Примеры: ГС дифференцировки эритроцитов, ГС, регулирующая прорастание семян и ГС апоптоза. Последний пример будет разобран ниже для иллюстрации *каскадного принципа* усиления сигнала.

5) Генные сети, обеспечивающие передачу наследственной информации и воспроизведение организма.

Характер изменения критических переменных для первых 4-х классов ГС представлен на рис. 5.

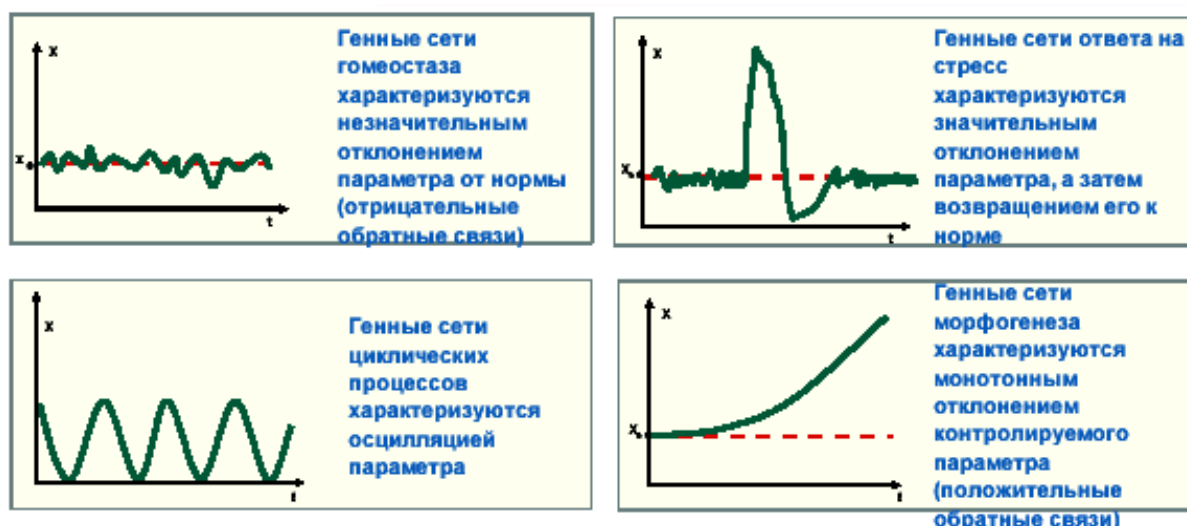
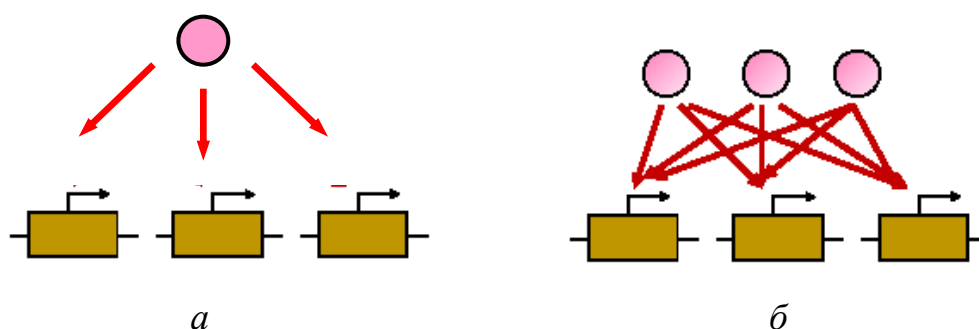


Рис. 5. Динамика основных переменных (x) для 4-х классов генных сетей (<http://wwwmgs.bionet.nsc.ru/systems/mgl/genenet/>)

Логический анализ позволяет выявлять *ключевые элементы, регуляторные контуры и мотивы* ГС.

В каждой локальной генной сети есть такие ключевые элементы, как *центральные гены* и центральные регуляторы. Первые обеспечивают координацию функций остальных генов этой сети. Пример – ген LFY в ГС развития цветка, о которой говорилось выше.

Центральный регулятор – это транскрипционный фактор (на рисунке *а* обозначен кружком), который активируют или репрессирует (стрелки) одновременно множество генов (прямоугольники) – *генную кассету*.



Мутации центральных регуляторов могут менять функции сразу многих регулируемых генов, приводя к выраженным изменениям.

Другая типичная ситуация (рисунок *б*)– несколько транскрипционных факторов регулируют кассету генов (не обязательно все стрелки-регуляции должны присутствовать).

Совокупность механизмов, посредством которых регулируются ГС, называют *регуляторным контуром*. В большинстве случаев адаптивные ГС содержат регуляторные контуры, как с положительными, так и с отрицательными ОС (пример на рис. *б*), что обеспечивает тонкую количественную авторегуляцию ГС.

Мотивы ГС – это регуляторные контуры, встречающиеся в реальных сетях достоверно чаще остальных. Прежде всего, это контуры положительных и отрицательных обратных связей.

ГС развития цветка арабидопсиса (рис. 1) имеет несколько контуров положительных обратных связей. Например, ген AP1 в цветочных почках активирует транскрипцию гена LFY, тогда как белок LFY позитивно регулирует экспрессию AP1 (рис. 146).

Типичным мотивом является также *биологический триггер* – конфигурация взаимодействующих генов, которая обеспечивает длительное пребывание части или всей ГС в одном из двух (реже многих) состояний устойчивого равновесия, но позволяет скачкообразно переключаться из одного состояния в другое под действием внешнего сигнала.

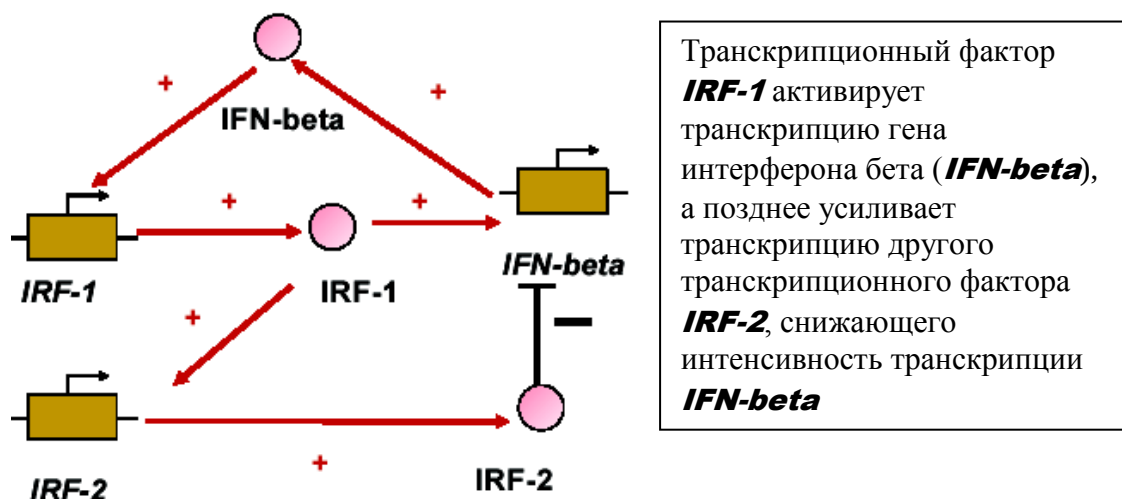


Рис. 6. Фрагмент ГС антивирусного ответа (по: Степаненко, 2003, с изменениями).

Пример – ген TFL в ГС развития цветка (рис. 1). Пока он активен – продолжается вегетативное развитие растения. Его подавление обеспечивает необратимость перехода к цветению. Более общий пример – каскады фосфорилирования у высших эукариот, контролирующие активацию или репрессию процессов клеточного цикла.

Конкретный вариант конфигурации из двух генов-регуляторов, обеспечивающей триггерный эффект, представлен на рис. 7. Это генетический триггер Жакоба и Моно – схема взаимной регуляции двух систем синтеза ферментов.

Ген-регулятор каждой системы (**Re**) синтезирует неактивный репрессор (**r**). Соединяясь с ферментом **P** – продуктом структурного гена противоположной системы, он образует активный комплекс. Последний, обратимо реагируя с сайтом связывания своего структурного гена, блокирует синтез его мРНК. Таким образом, **P1** – продукт первой системы является репрессором второй системы, а **P2** – продукт второй системы – репрессором

первой. Очевидно, что при таком характере взаимодействий интенсивная работа первой системы заблокирует вторую и наоборот. В результате будет активен, например, только ген **P1** пока какой-либо достаточно мощный внешний сигнал не вызовет переключения триггера – активным станет только **P2**.

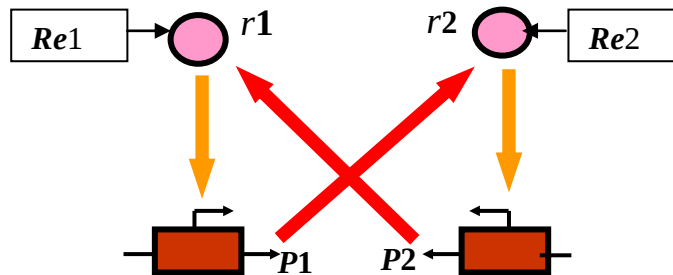


Рис. 7. Триггерная схема синтеза одного из двух ферментов – **P1** или **P2**.

В ГС часто выявляется *каскадный принцип* усиления сигнала, действующий по принципу положительной обратной связи. Такой принцип лежит в основе организации ГС апоптоза.

Митохондриальный путь запуска апоптоза инициируется белком *FasL*. Он активируют каспазы (**cysteine aspartate proteases**), которые представляют собой протеолитические ферменты, расщепляющие свои специфические субстраты по остаткам аспарагиновой кислоты. Степень усиления сигнала в контуре положительной обратной связи ГС апоптоза весьма высока. Активизация происходит по каскадному принципу усиления: инициирующие каспазы активируют исполняющие. Последние содержат «домены смерти» (*death domains*) и активируются в результате увеличения их локальной концентрации, а также белок-белковых взаимодействий с адапторными белками. Наступает фаза деструкции, в ходе которой хроматин клетки конденсируется и ДНК подвергается деградации, погибшие клетки опознаются и фагоцитируются другими клетками.

В любой физиологической системе все элементы и механизмы, названные выше, действуют совместно. Так при функционировании ГС эмбрионального развития дрозофилы на клеточной основе важнейшую роль

играют: молекулярно-генетические триггеры (обеспечивают 2 типа специфической дифференцировки клеток), петли положительных обратных связей (механизм поддержания дифференцированного состояния клеток), петли отрицательных обратных связей (механизм ограничения времени экспрессии генов в развитии).

В генной сети, обычно, можно выделить *исполняющую и регуляторную компоненты*. Например, в ГС цикла трикарбоновых кислот *Escherichia coli* в исполняющую (метаболическую) компоненту входят гены, кодирующие ферменты, непосредственно участвующие в этом метаболическом цикле, а в регуляторную – гены, их регулирующие. Регуляторная компонента этого цикла существенно сложнее метаболической: в среднем на 1 метаболический процесс приходится 13 регуляторных.

Вопросы:

1. Назовите основные классы генных сетей.
2. Объясните на примерах действие положительных и отрицательных обратных связей в генных сетях разных классов.
3. Дайте определения ключевым элементам, регуляторным контурам и мотивам генных сетей.
4. Приведите схему функционирования биологического триггера Жакоба и Моно.
5. Что называется регуляторной и исполняющей компонентой генной сети?

5. Построение моделей генных сетей

В чем состоят проблемы моделирования ГС адаптивного признака? Реальная сеть обычно обладает очень сложной, иерархической структурно-функциональной организацией, нередко содержащей сотни компонентов. Можно выделить несколько уровней организации ГС, которые учитываются при моделировании этих объектов.

Самым нижним является биохимический уровень. Здесь рассматривается огромная сеть биохимических процессов, происходящих в клетке или организме. Следующим уровнем является генетический. Здесь уже учитывается расположение и ориентация генов, регуляторных элементов, полиаллельность и т.д. Затем следует выделить клеточный

уровень, где учитываются процессы активного и пассивного переноса, выделить зоны (компартменты) локализации и действия ГС и т.д. Также выделяют более высокие уровни организации – тканевый, органный, организменный, популяционный и т.д. В большинстве случаев, объекты моделирования обладают сложным нелинейным поведением в силу наличия в них отрицательных и положительных ОС.

Даже если ограничиться рассмотрением ГС на уровне отдельной клетки, можно выделить несколько зон, например, межклеточное пространство, клеточная мембрана, цитоплазма, ядро. Кроме общих, в каждой зоне идут свои процессы, продукты которых воздействуют на ГС других зон. Далее необходимо описать ГС, отдельные части которых распределены по более крупным компартментам организма, таким, как группы клеток, органы и ткани. Для полного описания генных сетей желательно рассмотрение протекающих в них взаимосвязанных процессов на уровне целого организма.

При этом возникают трудности, связанные как с моделированием систем очень большой размерности, так и проведением экспериментов с реальными ГС. Поэтому пока исследования концентрируются на относительно небольших ГС, допускающих прямое сравнение с экспериментальными данными.

5.1. Базы данных

Для построения модели любой ГС используют информацию об ее элементах и связях между ними. Это означает, что необходимо знать какие гены участвуют в формировании сети, их регуляторные элементы (промоторы и др.), синтезируемые ими белки, пути передачи сигналов в клетке. Привлечение этой информации идет путем обобщения литературных данных и получения новых данных в ходе натурных экспериментов с конкретными ГС. Далее, составляют уравнения, характеризующие динамику ГС, и подбирают значения *параметров* – констант, входящих в эти уравнения.

Первые теоретические работы, посвященные анализу организации молекулярно-генетических систем управления у прокариот появились еще в 60-х годах прошлого века. До середины 80-х годов развитие теории генных сетей ограничивалось недостаточным объемом экспериментальных данных. В последние десятилетия появились эффективные методы изучения молекулярно-биологических механизмов регуляции экспрессии генов. В результате началось стремительное накопление экспериментальных данных о функционировании ГС (подробнее на сайте http://www.iti.cs.uni-magdeburg.de/iti_bm/ibss/courses/). Однако данные слабо структурированы и разбросаны по многим публикациям, в каждой из которых находятся результаты исследований лишь отдельных аспектов функционирования разных ГС.

Основной проблемой стала не слабость методов моделирования динамики ГС, а отсутствие эффективных компьютерных технологий для ведения информации об их структурно-функциональной организации, прежде всего для эукариот. С целью накопления и обобщения информации создан целый ряд баз данных (см. приложение 1). В них содержится качественная, описательная информация (о структурно-функциональной организации генной сети, механизмы различных процессов, организация регуляторных участков генов и т.д.), а также количественная информация (параметры ферментативных реакций, массы белков, и т.д.).

В качестве примера отметим базу данных GeneNet, созданную в ИЦиГ и доступную через интернет (<http://wwwmgs.bionet.nsc.ru/mgs/gnw/genenet/>). Она является частью системы GeneExpress (<http://wwwmgs.bionet.nsc.ru/mgs/systems/geneexpress/>) - электронной библиотеки по структуре и функционированию ДНК и белков. Классификация компонентов ГС в БД GeneNet представлена на рис. 8.



Рис. 8. Классификация компонентов генной сети в системе GeneNet (по: Ананько и др., 2001)

Для моделирования ГС в GeneNet разработан специальный язык и удобный графический интерфейс. Можно, в частности, указать элемент модели ГС, изменить его характеристики и проследить на графике ответные изменения в поведении различных компонентов ГС.

Технология GeneNet позволяет эксперту реконструировать ГС на основе экспериментальных данных, опубликованных в научных статьях. В базе данных описано несколько сотен генов, белков и небелковых веществ 43 организмов по данным из сотен публикаций. В результате удалось построить модели некоторых ГС: липидного метаболизма, кроветворения, противовирусного ответа, функционирования эндокринной системы, регуляции развития семени, азотфиксации, теплового шока, регуляции холестерина и др. Сведения об этих разработках в виде научных и учебных материалов можно найти по адресу <http://wwwmgs.bionet.nsc.ru/>.

Компьютерная технология GeneNet, развиваемая в лаборатории теоретической генетики ИЦиГ, является одним из первых примеров успешного решения задачи, которая открывает путь к быстрому накоплению и анализу данных экспериментальной информации о структурно-функциональной организации генных сетей.

5.2. Графические схемы функционирования генных сетей

Генные сети в упрощенных вариантах моделей описывают на языке *генетических элементов* (ГЭ). Каждый из них содержит ген,

соответствующие ему мРНК и белки. Функционирование генетического элемента в простейшем случае – процесс синтеза белка. В генетический элемент неявно включаются все процессы, обеспечивающие синтез мРНК, белка, формирование мультимера белка (мономерная форма белка есть частный случай мультимерной). Белки обеспечивают регулирующие взаимодействия между ГЭ.

Регуляторный механизм ГС – элементарный процесс, посредством которого некоторая совокупность генетических элементов действует на конкретный генетический элемент, уменьшая или увеличивая скорость синтеза его продукта. В реальной ГС кроме белков генетические элементы, возможно, содержат низкомолекулярные соединения, которые также регулируют состояния других ГЭ.

Схему функционирования ГС, включающую ГЭ и их воздействия, можно задать *ориентированным графом*. Он состоит из конечного множества вершин (g_i) и набора упорядоченных их пар, изображаемых в виде стрелок. Любая стрелка называется дугой, или ориентированным ребром. Она отражает направление и характер воздействия (отрицательное или положительное регуляторное воздействие) одного ГЭ на другой. Пример схемы функционирования простой гипотетической ГС – на рис. 9. Отметим, что ген 1 не регулирует экспрессию гена 2 непосредственно, но может повлиять на нее через ген 3. В некоторых случаях отрицательные регуляторные связи на схемах функционирования ГС обозначают пунктирными, а положительные – сплошными стрелками.

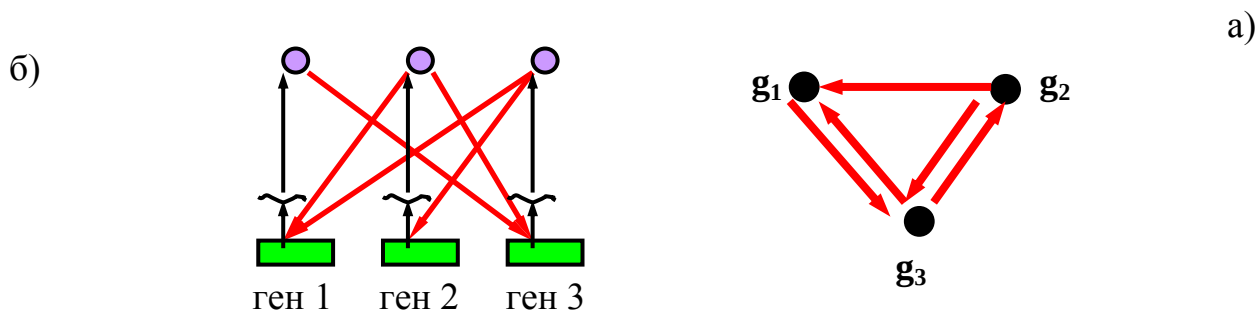


Рис.9. а) Исходная ГС из трех генов (внизу), их белков (наверху) и регуляторных воздействий (стрелки) на другие гены; б) соответствующий ей ориентированный граф – структурная схема функционирования ГС.

На рис. 10 в этих терминах, но в других обозначениях, показан фрагмент схемы реальной ГС прорастания семени растения. Элементы гиббереллины (gibberelins) и абсцизовая кислота (ABA) соответствуют низкомолекулярным соединениям, остальные ГЭ – гены и их белки. Стрелки, помеченные кружками с плюсом – положительные (усиливающие) регуляторные воздействия, с минусом – отрицательные (ингибирующие). Более подробно эта ГС рассмотрена в [4].

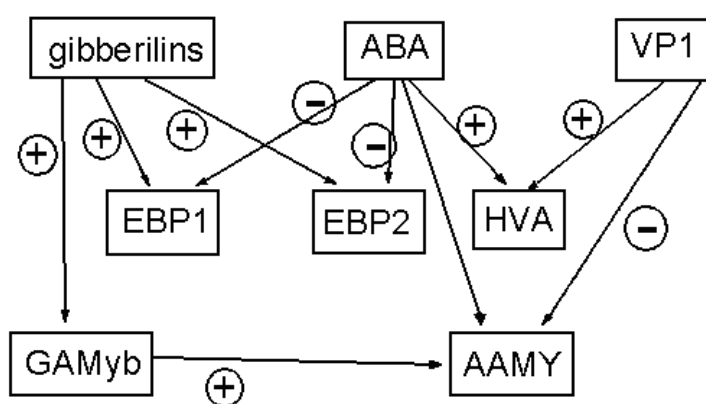


Рис. 10. Граф фрагмента внутриклеточной ГС прорастания семени (http://www.mathcell.ru/ru/obzors/obzor_Elkin.shtml)

Иногда ГЭ может находиться не только в двух (активен – не активен), а в нескольких состояниях. Например, когда регуляторный белок, продуцируемый геном должен пройти несколько состояний – стадий формирования. В подобных случаях на схеме каждому ГЭ соответствуют все его состояния – вершины графа. Считается, что известна схема перехода состояний друг в друга, по-прежнему отраженная дугами – переходами между ними. Каждый переход осуществляется с определенной скоростью, которая зависит от концентраций ряда белков-регуляторов. Это выражается в конкретной математической зависимости скорости от концентраций белков-регуляторов и учитывается в более сложных, непрерывных моделях ГС (см. рис. 13). Они будут рассмотрены в разделе 5.4.

В базе данных GeneNet в соответствии с выделением элементарных структур и событий (рис. 8) для графического представления используется

также *двудольный граф ГС*. В нем обозначаются кружками ГЭ с их переменными, квадратами – процессы в ГС, стрелками – участие ГЭ в процессах. Пример подобного графа представлен на рис. 11

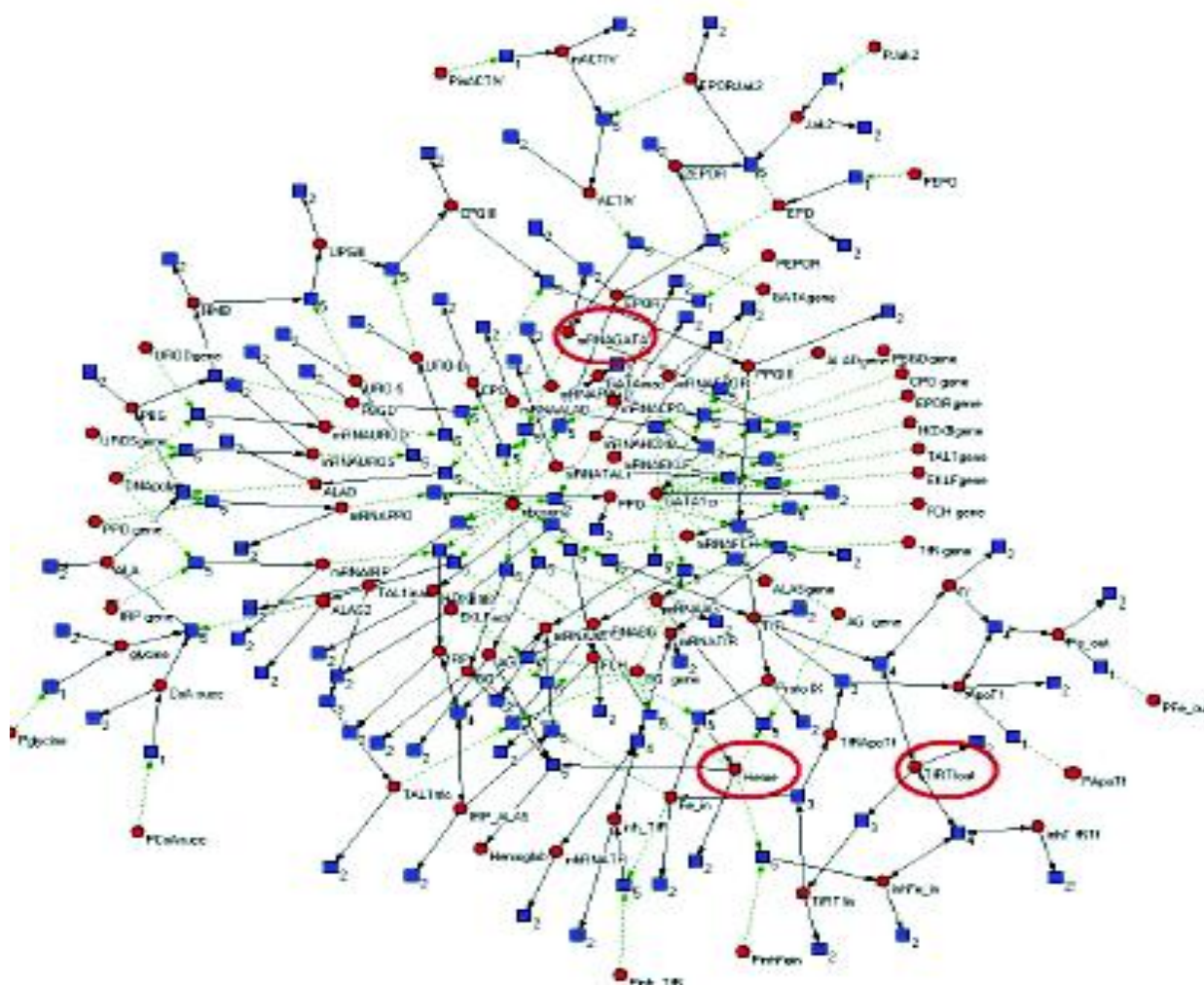


Рис. 11. Двудольный граф ГС регуляции дифференцировки эритроидной клетки. (65 динамических переменных, 165 процессов). (<http://wwwmgs.bionet.nsc.ru/>)

Графическое моделирование – простейший способ описания и первичного анализа структурно-функциональной организации ГС. В некоторых случаях непосредственно по структурному графу можно выявить важные регуляторные контуры, мотивы, а также предсказывать свойства ГС, связанные с *предельными режимами* ее функционирования. Например, в некоторых случаях удастся оценить число *точек покоя* или *предельных циклов*. Смысл этих терминов следующий.

Иногда при достаточно длительном функционировании ГС (в пределе) из всех ее возможных (n) продуктов синтеза остается только один из m ($m \leq n$).

Такие состояния ГС называют точками покоя, т.к. они соответствуют положениям равновесия. Какой именно из m продуктов останется, по графу определить не всегда возможно – необходим количественный анализ. Для этого следует дополнять графы ГС, как минимум, их дискретными моделями, рассмотренными в следующем разделе.

В других случаях непосредственно по графу ГС удастся выявить сочетание из двух или большего числа продуктов синтеза, которые даже при длительном функционировании будут присутствовать в ГС. Такие сочетания называются предельными циклами, т.к. их концентрации продолжают циклически меняться во времени. Однако по графу ГС невозможно оценить закономерности изменения концентраций во времени, для этого их необходимо дополнить непрерывными моделями.

Таким образом, графическое моделирование используется совместно с более сложными математическими моделями.

5.3. Дискретный способ моделирования генных сетей

Существует два основных количественных детерминистических подхода к описанию ГС – дискретный и непрерывный. В некоторых случаях эти математические модели объединяют, а также пытаются учесть случайную составляющую процессов в ГС с помощью стохастических моделей. Последние применяют, в частности, для таких генных сетей, в которых значительные события обеспечиваются очень малым количеством регуляторного агента (например, транскрипционного фактора).

В приложении 2 приведен список названий разнообразных вариантов количественных математических моделей ГС с указанием некоторых фамилий авторов соответствующих статей.

В простейших вариантах *дискретных моделей* ГС предполагается, что каждый ген в ГЭ может находиться лишь в одном из двух состояний: 1-активен (экспрессируется) и 0 - неактивен. Для простоты предполагается, что процесс функционирования ГС состоит из элементарных этапов (шагов). На очередном шаге состояние каждого гена сохраняется или меняется на

противоположное. Это зависит от активизирующих и ингибирующих воздействий на него со стороны других генов, активных перед очередным шагом. Схема и знаки воздействий отражаются в структуре графа изучаемой ГС (см., например, рис. 10).

Исчерпывающей характеристикой ГС в рамках подобной модели является *вектор состояний*. Это набор нулей и единиц, упорядоченный по номерам ГЭ, представленных на структурном графе ГС. Например, для ГС на рис. 9 вектор состояний содержит три компонента, каждый из которых может принимать значение 0 или 1. Некоторые значения меняются по шагам.

Алгоритм определения состояния каждого ГЭ по шагам формализован в компьютерной программе. Для этого используют специальные функции, оперирующие только с 0 и 1 (*алгебра логики – булева алгебра*), или *пороговые функции*.

Для примера рассмотрим алгоритм переключения состояний в модели, включающий пороговую функцию. Предполагается, что состояние (активен или неактивен) для i -го гена ГС в результате перехода к следующему $(t+1)$ шагу будет зависеть только от средневзвешенной активности $a_j(t)$ – состояний генов на предыдущем шаге (t) . Тогда для оценки $a_i(t+1)$ – нового состояния j -го ГЭ можно использовать пороговую функцию η :

$$a_i(t+1) = \eta \left[\sum_{j=1}^N m_{ij} a_j \right]$$

Здесь N – общее число ГЭ в сети, m_{ij} – вес, т.е. степень влияния j -го гена на i -ый. В простейшем случае m_{ij} может принимать значения 0 (не влияет), 1 (активирует), -1 или $-N$ (ингибирует частично или полностью).

Если сумма, стоящая в квадратных скобках, больше или равна «высоте» *порога* – заранее заданного натурального числа, то значение пороговой функции η становится равным единице. Это значит, что i -ый ген на шаге $t+1$ активен [$a_i(t+1) = 1$]. Если сумма в скобках меньше порога, то – ноль [$a_i(t+1) = 0$ – ген неактивен]. Более сложные варианты дискретных моделей позволяют учесть не два, а несколько $(1, \dots, p)$ состояний каждого ГЭ.

Результат моделирования ГС, т.е. изменение ее вектора состояний по шагам, иногда существенно зависит от начального условия – исходного вектора состояний перед первым шагом.

Главным преимуществом дискретных моделей ГС является их простота. Такие модели позволяют проследивать динамику ГС (изменение вектора состояний) в течение длительного времени, прогнозировать предельные режимы функционирования регуляторных контуров генных сетей. Но, безусловно, подобные модели являются довольно грубым упрощением.

5.4. Непрерывные модели

Наиболее сложная задача – изучение динамики количественных переменных ГС. Сюда относятся концентрации продуктов: «входных» - действующих на ГС извне, «внутренних» - регуляторов процессов и «выходных» - выделяемых сетью во вне. Для количественного описания динамики этих непрерывных количественных переменных во времени необходимы сложные модели. Наиболее полное описание обеспечивают дифференциальные уравнения и их системы, хотя этот подход имеет ряд недостатков.

С помощью непрерывных моделей пытаются ответить на некоторые вопросы теории ГС. Среди них:

- выявление условий существования в ГС множества состояний устойчивых во времени, т.е. состояний, равновесных по значениям основных переменных, а также выявление условий переключения между такими состояниями.

- изучение «степени устойчивости» положений равновесия и влияния на нее различных факторов (мутаций, гетерозиготности, стрессов и пр.). Понятно, например, что наличие отрицательной ОС увеличивает, а положительной - снижает устойчивость равновесия.

- нахождение условий возникновения колебательных, циклических режимов функционирования ГС. В частности, оценка минимального уровня

сложности сети и необходимости присутствия в ней положительных и отрицательных ОС.

Кроме теоретических исследований ведутся работы по математическому моделированию реальных ГС и сравнению результатов с данными экспериментов.

Для построения непрерывных моделей требуется формализовать, например, с помощью дифференциальных уравнений «элементарные события» ГС: функционирование регуляторных механизмов, связанных с управлением процессами транскрипции, трансляции, биохимические реакции, включая формирование белков, передачу внешних и внутренних сигналов, работу контуров управления и пр. Частные примеры подобных элементарных событий и уравнений представлены на рис. 12-14.

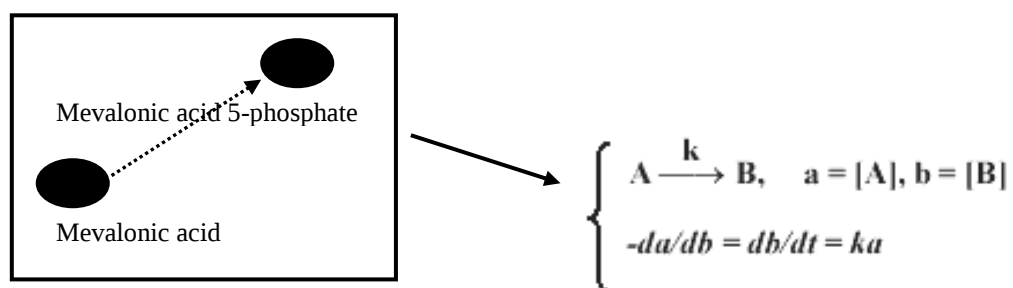


Рис. 12. Схема и дифференциальные уравнения – описание химико-кинетической закономерности простейшей мономолекулярной биохимической реакции фосфорилирования. Здесь a и b – концентрации исходного и результирующего продуктов реакции, k – параметр, характеризующий скорость прохождения реакции (<http://wwwmgs.bionet.nsc.ru/>)

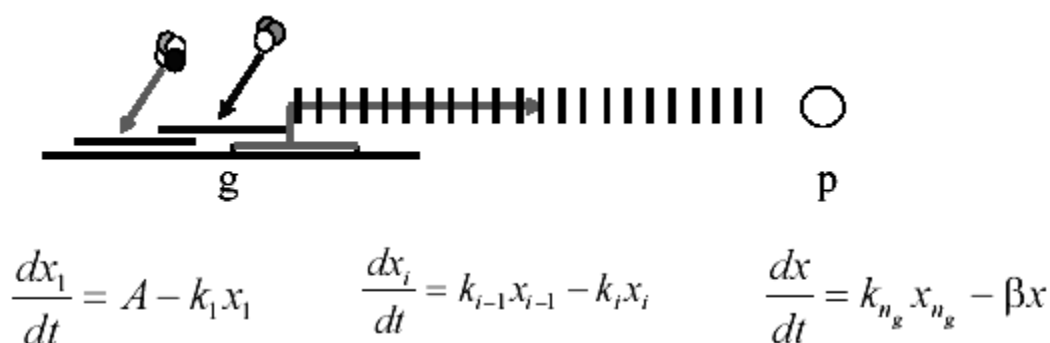


Рис. 13. Схема и дифференциальные уравнения, описывающие функционирование генетического элемента сети при многоэтапном формировании регуляторного белка p (<http://wwwmgs.bionet.nsc.ru/>).

Здесь: n_g – количество промежуточных стадий синтеза белка, x_i – концентрации промежуточных продуктов синтеза белка, x – концентрация белка, A – параметр – скорость инициации конститутивного синтеза, k_i – параметры – скорость перехода из i -го промежуточного состояния в $(i+1)$ -ое промежуточное состояние, β – параметр – скорость деградации белка p .

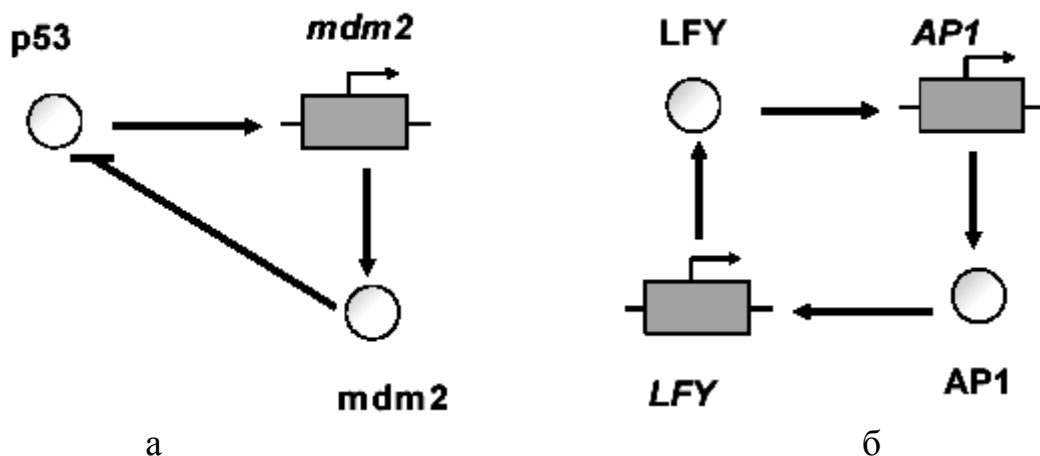


Рис. 14. Схема некоторых контуров ГС с обратной связью (<http://wwwmgs.bionet.nsc.ru/>)

а) Отрицательная ОС: транскрипционный фактор гена p53 активирует транскрипцию его гена-ингибитора mdm2. Уравнение для переменной x – концентрации транскрипционного фактора гена имеет вид: $dx/dt = \alpha / (1 + K^c x^c)$. α , K , c – параметры реакций. Скорость синтеза белка (dx/dt) убывает с ростом концентрации x .

б) Положительная ОС: взаимное усиление транскрипции двух генов. Уравнение для переменной x – концентрации транскрипционного фактора гена AP1 имеет вид: $dx/dt = \alpha / (1 + K_1^c x^c) / (1 + K_2^d x^d)$. Скорость синтеза белка (dx/dt) возрастает с ростом концентрации x .

В рамках непрерывных и дискретных моделей возможен *гибридный подход* к моделированию динамики ГС. В нем разделяют сеть на 2 подсистемы. Первая – регуляторная, включает совокупность регуляторных генов и их макромолекулы. Ее состояния моделируются на языке дискретных моделей (например, с использованием пороговых функций). Вторая – исполняющая, содержит гены, обеспечивающие выполнение различных биохимических процессов. Она моделируется системами дифференциальных уравнений. В итоге общая модель формулируется на языке кусочно-непрерывных систем уравнений.

Подобные принципы построения моделей функционирования ГС заложены в обобщенном химико-кинетическом методе моделирования (ОХКММ), разработанном в ИЦиГ. Вначале проводится иерархическая декомпозиция исходной ГС до элементарных структур и процессов. Затем процессы формализуются в терминах стандартных блоков – систем уравнений. Из них проводят сборку имитационной компьютерной модели всей ГС.

Динамика отдельных молекулярно-генетических систем в ОХКММ формально описывается моделями общего вида:

$$\frac{dx_i}{dt} = F_i(X, K) - x_i G_i(X, K)$$

Здесь $X = (x_1, \dots, x_n)$ – вектор динамических переменных модели ГС, т.е. концентраций регуляторных продуктов. $K = (k_1, \dots, k_m)$ – вектор параметров, входящих в уравнения, $F_i(X, K)$ и $G_i(X, K)$ – рациональные функции.

В ОХКММ учитываются временные задержки и локализация процессов в однородных областях, закономерности пассивного и активного переноса веществ и энергии между областями, возможная полиплоидность и гетерозиготность генов, случайные составляющие процессов и пр. В результате изучение динамики функционирования ГС можно рассматривать как математическую проблему анализа решений сложных систем уравнений.

К настоящему времени разработан специальный язык моделирования ОХКММ и программное обеспечение, позволяющее строить имитационные модели сложных многокомпарментных биологических систем, вплоть до организменного уровня.

5.5. Проблема адекватности моделей

Одной из наиболее важных и актуальных проблем при построении математических моделей ГС является ее настройка, т.е. подбор значений многочисленных констант-параметров, входящих в уравнения. Простейший пример – параметр k в уравнении фосфорилирования на рис. 12.

Как правило, экспериментально измеренные значения известны только для ограниченного числа параметров. Существует также проблема вариабельности набора экспериментальных данных из литературы, по которым оценивают параметры и строят модель. Их получали в конкретных условиях, при разных экспериментальных воздействиях в различные моменты времени и т.д. Недостающие и уточненные данные приходится получать в ходе дополнительных экспериментов с реальными ГС.

Получение значений недостающих параметров назовем *настройкой модели*. Кроме того, далеко не всегда детально известна структура организации некоторых частей ГС (пути передачи сигналов, механизмы протекания процессов и т.п.). Проверку правильности гипотез о структуре ГС и проверку «устойчивости» ранее полученных оценок параметров при использовании новых экспериментальных данных можно определить как *верификацию модели ГС*.

Обе проверки можно рассматривать как единую оптимизационную задачу, решение которой проводится на основе специальных численных экспериментов с моделью. Целевой функцией, подлежащей минимизации, выступает общее расхождение реальной экспериментальной динамики от динамики переменных ГС, рассчитанной по ее модели при заданной структуре и параметрах. В качестве динамики могут выступать как зависимости переменных ГС от времени, так и друг от друга.

Рассмотрим, например, метод оптимизации, используемый в разных областях науки и называемый *генетическим алгоритмом*.

Идея заключается в имитации последовательной «эволюции» популяции, состоящей из так называемых «особей» – моделей генных сетей. Каждая из «особей» обладает «генотипом» – конкретным набором значений параметров модели, несколько отличным от значений у других особей. Допускается также включение в расчет одновременно нескольких моделей сети с несколько отличной структурой. Генотип однозначно определяет «фенотип» – расчетную динамику переменных этой модели генной сети. Особи образуют исходную «популяцию». Она является объектом действия

«эволюции», включающей «мутации», «скрещивание родителей» с «рекомбинацией», «размножением потомков», а также «отбор» среди них наиболее приспособленны. Эволюция вызывает изменение состава популяции, которое возникает на каждом эволюционном такте алгоритма. Эволюционный такт, в свою очередь, складывается из двух подтактов.

Первый подтакт состоит в «скрещивании и размножении» особей популяции. Каждая пара моделей-родителей может порождать некоторое количество «потомков». Перед скрещиванием возможен «мутационный процесс», после скрещивания – свободная «рекомбинация» параметров-аллелей. Существует несколько вариантов порождения мутаций. Например, в одном из вариантов от каждого «родителя» возникает несколько «мутантных особей», отличающихся по величине не более чем одного параметра. Небольшие величины отличий задаются методом Монте-Карло случайно для каждой особи. После скрещивания с рекомбинацией возникает новая популяция из родителей и потомков.

Второй подтакт заключается в выделении из нее «наилучшей» субпопуляции. А именно, для каждой особи рассчитывается фенотип (динамика функционирования модели ГС), который сравнивается с едиными и неизменными экспериментальными данными об этой динамике, т.е. вычисляется значение функции приспособленности. Функция приспособленности, подлежащая минимизации, может представлять, например, взвешенное среднеквадратическое расхождение расчетной динамики каждой особи (модели ГС) и «идеала» – экспериментальной динамики. Для последующего размножения отбирается фиксированное количество особей, имеющих наилучшие характеристики индивидуальной «приспособленности». В результате возникает новая популяция, в которой содержатся наилучшие родительские особи и потомки.

После того как новая популяция порождена, такт размножения с мутациями, рекомбинацией и отбором снова повторяется. Постепенно лучшие модели ГС приближаются к реальной экспериментальной динамике.

Только проведя процедуру оптимизации и убедившись в адекватности модели, имеет смысл проводить анализ функционирования ГС на основании расчетов по ее модели.

В качестве примера на рис. 15 представлен расчет динамики одной из переменных по адекватной модели ГС биосинтеза холестерина, полученной в результате применения подобного генетического алгоритма.

Вопросы:

1. В чем состоят проблемы моделирования генных сетей?
2. Какие базы данных используют при разработке и проверке моделей генных сетей?
3. Назовите типы графов генных сетей.
4. Что такое вектор состояний в дискретных моделях генных сетей?
5. Приведите структуру пороговой функции, используемой для описания состояния гена.
6. Какие «элементарные события» требуется формализовать при построении непрерывной модели генной сети?
7. Для чего используют генетический алгоритм?

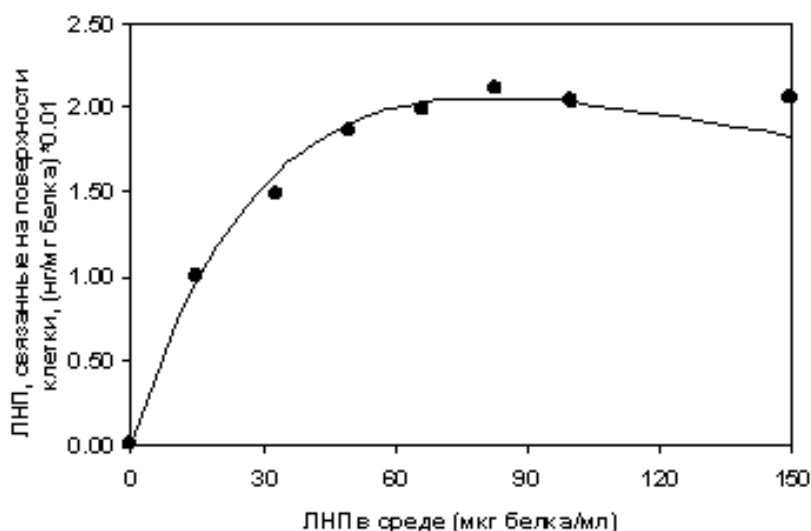


Рис. 15. Сравнение расчетных данных (гладкая линия) лучшей модели ГС биосинтеза холестерина и экспериментальных данных (точки) об эффективности связывания липопroteинов низкой плотности (ЛНП) их рецепторами (по: Ратушный, 2006).

6. Методика и результаты анализа моделей генных сетей

Адекватные математические модели ГС и их компьютерные реализации являются незаменимым инструментом для анализа

функционирования сложных биологических систем. С их помощью возможно:

- предсказывать динамику тех или иных процессов в ГС;
- исследовать механизмы их протекания;
- выявлять ключевые звенья в генных сетях;
- анализировать влияние различных воздействий на живые системы;
- подбирать оптимальные воздействия;
- изучать эволюцию ГС.

6.1. Анализ функционирования генных сетей

В качестве примера приведем результаты моделирования ГС морфогенеза цветка арабидопсиса. На рис. 16 представлен прогноз динамики белковых продуктов нескольких генов в режиме R_0^1 – перехода к формированию чашелистиков.

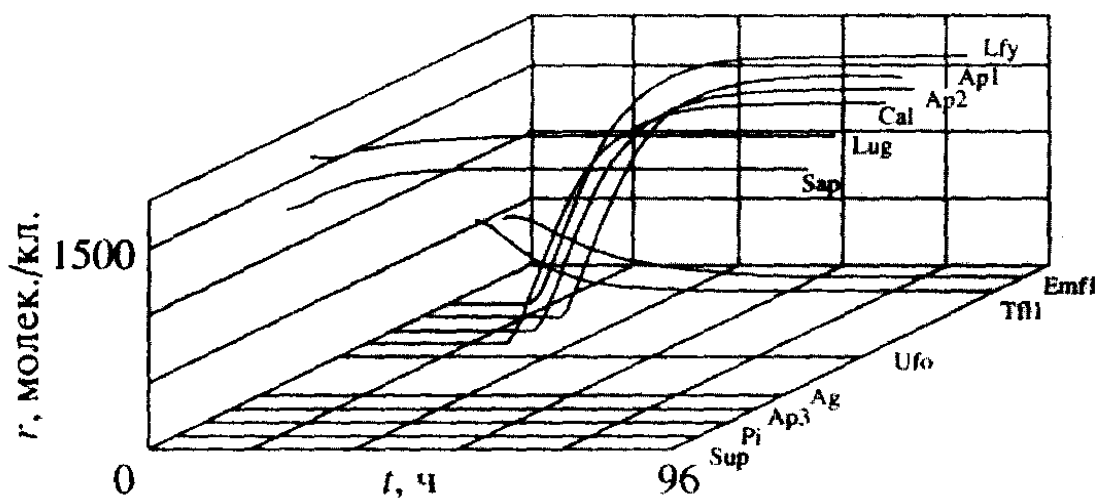


Рис. 16. Расчетные кинетические кривые концентраций белков (в молекулах на клетку) при переходе к стационарному режиму R_0^1 (по: Чураев, Галимзянов, 2001).

При анализе подобных расчетных кривых для клеточных ГС морфогенеза арабидопсиса и дрозофилы были выявлены возможности функционирования одной ГС в разных режимах, наличие нескольких стационарных состояний с разными режимами переключения. Динамика существенно зависит от начальных условий (начальных концентраций

белков-морфогенов) и параметров ГС (пороговых концентраций регуляторных веществ, временных задержек протекания элементарных процессов и т.п.).

Изучение ГС различных организмов, признаков и процессов, описанных в базе данных GeneNet, позволило сформулировать общую концепцию интерференции. *Интерференция генных сетей* – это прохождение возбуждения по определенному компартменту глобальной генной сети организма, приводящее к последовательному возбуждению ряда его генных сетей. Интерференция может приводить к координированной экспрессии ГС, либо к подавлению функционирования одной ГС сигналом другой. Интерференция генных сетей в глобальной ГС организма реализуется с участием генетических *элементов-интеграторов*. Они могут рассматриваться как ключевые элементы, обеспечивающие передачу сигнала в глобальной ГС всего организма.

С помощью моделирования динамики гипотетических ГС получен ряд теоретических результатов. В частности, установлено, при каких условиях ГС могут иметь устойчивые предельные режимы, т.е. предельные во времени особенности ее функционирования: стационарные, циклические, а также динамические состояния переменных более сложной природы. В качестве примера на рис. 17 представлена динамика перехода к предельным траекториям переменных $(p_1, \dots, p_4)$ гипотетической ГС, состоящей из 4-х элементов с отрицательными регуляторными воздействиями. Динамика существенно зависит от начальных условий – концентраций их продуктов в начальный момент ее функционирования.

Установлено, что предельные режимы динамики ГС определяются характером структурно-функциональной регуляции генетических элементов. Это выразилось в формулировках ряда математических критериев, которые связывают предельные режимы со свойствами соответствующих структурных графов ГС. Применительно к реальным ГС это означает, что необходимым условием существования в них некоторого количества стационарных и/или циклических режимов является наличие определенной

минимальной сложности (нелинейности) процессов регуляции ГЭ сети. Достаточный уровень сложности может достигаться, в частности, мультимеризацией белков-репрессоров и/или наличием минимального числа промежуточных стадий, ведущих от гена к белку-регулятору.

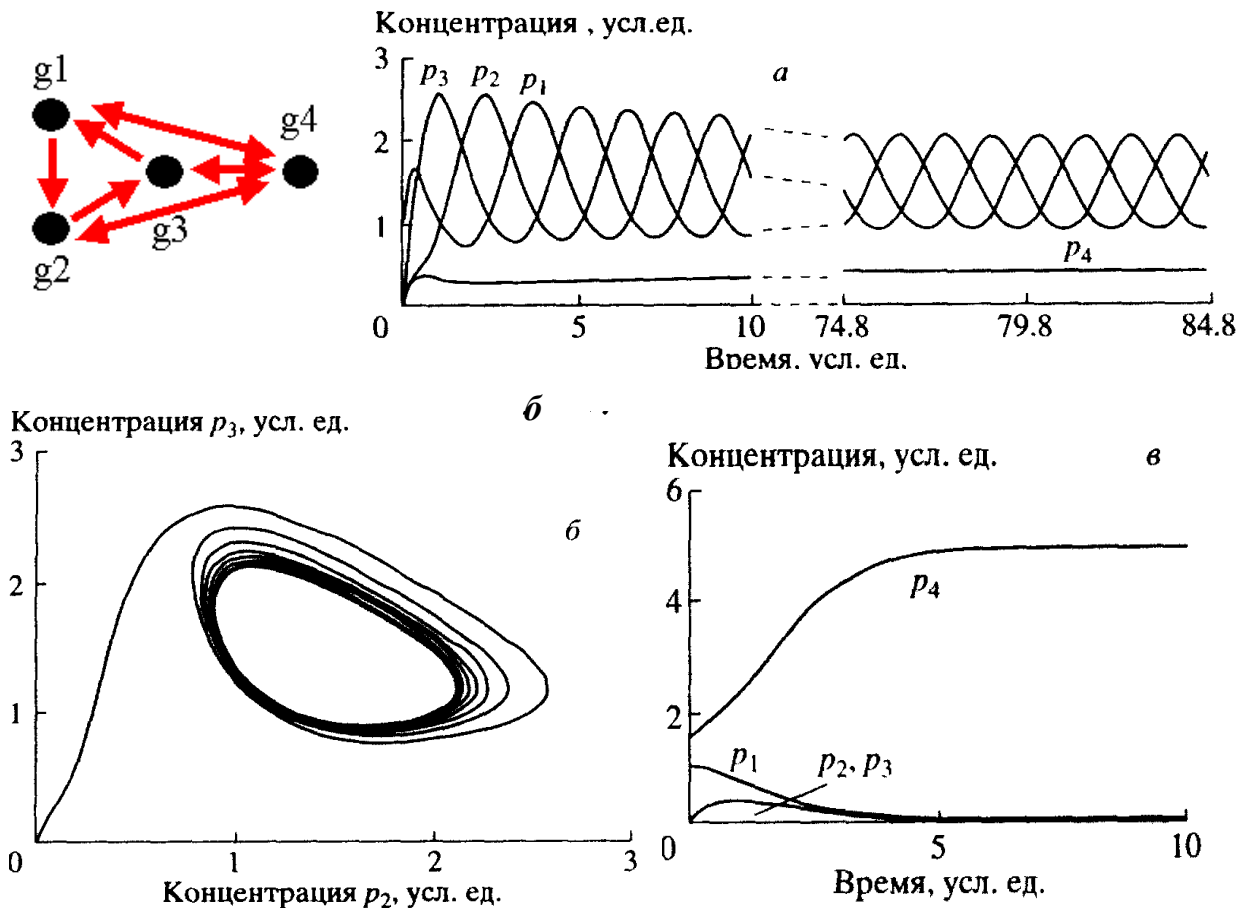


Рис. 17. Ориентированный граф гипотетической ГС и траектории изменений ее переменных $p_1, \dots, p_4$ при разных начальных условиях $p_1^0, \dots, p_4^0$ функционирования (по: Лихошвай и др., 2001):

а, б) $p_1^0=1, p_2^0=p_3^0=p_4^0=0$ – в пределе ГС приходит к осциллирующему режиму (незатухающие периодические колебания p_1, p_2, p_3 при $p_4=0,3$);

в) $p_1^0=1, p_2^0=p_3^0=0, p_4^0=1,5$ – ГС приходит к стационарному режиму $p_1=p_2=p_3=0, p_4=5$

6.2. Анализ влияния мутаций

Адекватные модели ГС позволяют проводить компьютерные эксперименты, имитирующие влияние мутаций, мобильных элементов, трансгеноза и внешних воздействий. Методика моделирования состоит в определении узких мест и их смены в процессе функционирования ГС.

Учитывается, что мутации генов ГС, гетерозиготность, полиплоидия и другие особенности вызывают изменения ее параметров. Для

количественного анализа последствий подобных воздействий в компьютерной модели ГС задают отклонения параметров. При сопоставлении результатов расчетов по моделям ГС с отклонениями получают *матрицу чувствительности*. Она состоит из коэффициентов, характеризующих относительное влияние варьирования отдельных параметров на предельные режимы основных переменных ГС.

При анализе этой матрицы выявляют, прежде всего, *узкие места* – лимитирующие (ключевые) параметры, отклонения которых ведут к резкому изменению функционирования ГС. На этой основе делают предварительные выводы о возможностях управления ГС за счет мутаций отдельных генов сети, а также внешних воздействий. В частности, величины параметров-констант реакций ГС в рамках современной фармакологии могут быть изменены путем введения высокоспецифичных активаторов или ингибиторов.

Например, на рис. 18 представлены результаты сравнения динамики ГС синтеза гемоглобина при нормальном и мутантных типах системы дифференцировки эритроидной клетки (расчет по модели).

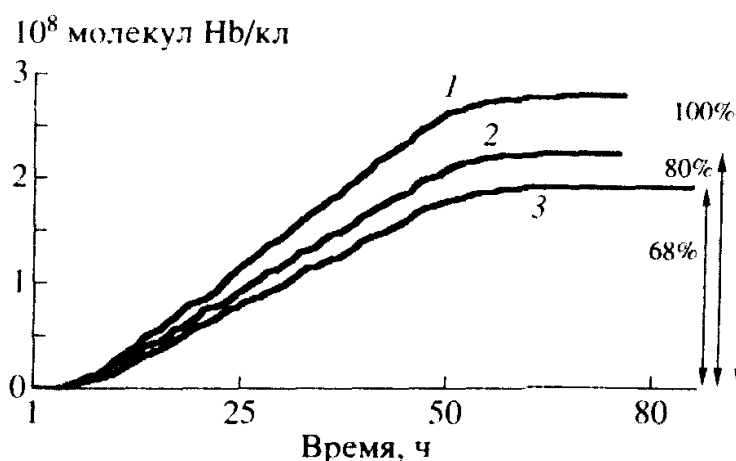


Рис. 18. Сравнение динамики синтеза гемоглобина при нормальном (1) и мутантных (2,3) типах системы дифференцировки эритроидной клетки (по: Лихошвай и др., 2001)

Здесь 2 – влияние мутации гена *EKLF* (прекращение синтеза транскрипционного фактора *EKLF*); 3 – влияние мутации гена *GATA-1*

(снижение активности транскрипционного фактора *GATA-1* на 30%). По оси абсцисс – время в часах, по оси ординат – число молекул гемоглобина (Hb) на клетку.

Сравнительный анализ мутаций и приведенных выше кривых показывает, что роль гена *GATA-1* в биосинтезе гемоглобина в эритроидной клетке важнее роли гена *EKLF*. Экспрессия гена *GATA-1* в данном случае является лимитирующим процессом.

Сложные вычислительные эксперименты с моделью ГС регуляции биосинтеза холестерина в клетке позволили сделать более общие выводы. Выяснилось, что стационарная концентрация холестерина, несвязанного с мембранами, весьма устойчива по отношению к мутационным изменениям большого количества параметров модели. Около 40% параметров модели ГС при изменении в заданных пределах практически не влияли на стационарную концентрацию холестерина. При изменении другой части параметров (~45%) она колебалась незначительно – в пределах $\pm 35\%$. Однако, изменения некоторых параметров (~15%) весьма сильно сказывались на равновесной концентрации холестерина. Дело в том, что их изменение влечет за собой нарушение регуляторных связей внутри ГС.

6.3. Эволюция генных сетей

В основе эволюции ГС лежат известные факторы, основными из которых являются мутации и отбор. Однако их влияние имеет особенности, некоторые из них удалось установить в результате изучения ГС с использованием разных методов.

Так, роль мутаций, изменяющих динамический режим функционирования генной сети можно проиллюстрировать примером из предыдущего раздела. На рис. 17 было представлено сложное переключение предельных режимов гипотетической сети из 4-х генов. Если в результате мутации оборвется лишь одна из многочисленных связей (от первого гена к четвертому), то ГС утратит возможность переключения режимов. Независимо от начальных условий ($p_1^0, \dots, p_4^0$) у ГС сохранится единственное

устойчивое равновесие (рис. 17с), когда в клетке остается только белок p_4 , а остальные белки отсутствуют.

Специфическое значение для ГС, имеют мутации центральных регуляторов. Даже их небольшие мутационные отклонения могут повлечь резкую перестройку фенотипа. Экспериментальным подтверждением служат перестройки морфогенеза *A. thaliana* вследствие трансгенеза MADS-генов – центральных регуляторов генной сети развития цветка. В разных экспериментальных вариантах трансгенез приводил к превращению стеблевых листочков в лепестки, появлению цветка из двух рядов лепестков и тычинок и т.д. Подобные резкие изменения фенотипа ставят вопрос о механизмах защиты от подобных мутаций в процессе эволюции.

По-видимому, один из таких механизмов – комплексность центральных регуляторов, хорошо изученная у растений. Так, отличительной чертой групп генов, участвующих в формировании меристемы цветка и его органов (рис. 1), является то, что регуляторами генов-интеграторов этих групп являются не отдельные белки, а совместно действующие ансамбли белков и/или белковые комплексы. Поэтому, например, мутации в гене AP1 не изменяют экспрессию гена-регулятора LFY, приводя только к частичным потерям функции меристемы цветка.

Другим механизмом защиты таксона от мутаций центральных регуляторов является сильная интеграция генных сетей, вследствие чего мутации обычно имеют резко выраженный плейотропный эффект. Поэтому центральные регуляторы постоянно находятся под жестким давлением стабилизирующего отбора. В результате их мутантные варианты с большой вероятностью элиминируют из популяции.

Ослабление интеграции генных сетей или выделение относительно самостоятельных и слабо перекрывающихся подсетей формируют горячие точки эволюции. Например, у растений лепестки и тычинки являются самыми лабильными органами цветка. Этот факт коррелирует с тем, что участвующие в их формировании MADS-гены групп В и С задействованы только в сети формирования органов цветка. Наоборот, гены группы А,

кроме того, участвуют в формировании меристемы цветка. Соответственно эволюция генов, входящих в подсети В и С, по-видимому, началась раньше. Гены же семейства AP1, судя по наличию высокомолекулярных паралогов, переживают эволюционные изменения в настоящий период, причем механизмом ослабления интеграции генных сетей в этом случае является дупликация гена, что типично для растений.

Другим механизмом ослабления интеграции генных сетей может служить усложнение регуляции гена, позволяющее по-разному регулировать один и тот же ген в разных группах клеток многоклеточного организма. Функционально это равносильно дупликации гена. Например, одни и те же *Hox*-гены, *Ubx* и *abdominal-A*, по-разному экспрессируются в различных брюшных сегментах гусеницы.

Роль мутаций в контурах положительных и отрицательных обратных связей несходна в разных ситуациях отбора. Возникновение новых отрицательных ОС минимизирует фенотипическое проявление мутаций, охваченных этим контуром, отчасти уводя их из-под действия отбора. Такая ситуация выгодна при стабилизирующем естественном отборе.

Напротив, новая положительная ОС способна вызвать критические изменения функционирования всего организма. Так, роль уникальных мутаций в промоторе гена *tb1* у дикого предка кукурузы (теосинте) можно объяснить лишь возникновением контуров с каскадным усилением в генных сетях онтогенеза по механизму положительной ОС. В результате этих мутаций проявилось большое разнообразие ранее скрытых мутантных форм, что обеспечило успех селекции кукурузы. Таким образом, мутации, вызывающие новые положительные ОС, могут повысить эффективность движущего (направленного) отбора.

Таксоны, потерявшие отрицательную ОС, также будут взрывообразно демонстрировать весь спектр ранее накопленных латентных мутаций. Широкий спектр новой мутационной изменчивости потенциально способен вызвать всплеск видообразования.

Эти теоретические рассуждения получили подтверждение при моделировании эволюции ОС в генных сетях при разных типах отбора (рис. 19). Стабилизирующий отбор в популяции благоприятствует особям с отрицательной ОС, преимущественно фиксируя такие связи высокого уровня иерархии, что ведет к росту иерархии регуляторных систем (рис. 19а). При этом на нижних уровнях иерархии накапливаются мутации, эволюционирующие в нейтральном режиме.

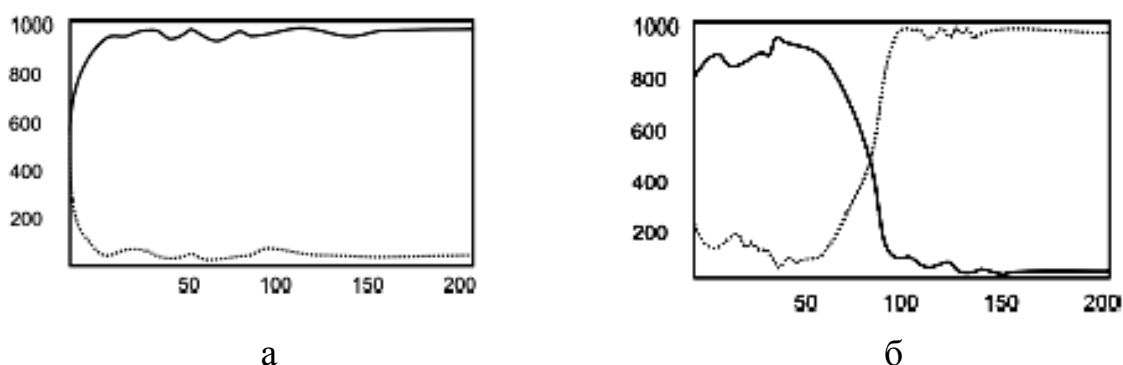


Рис. 19. Результаты моделирования конкуренции особей с отрицательной ОС (сплошная линия) и без нее (пунктир) в ходе эволюции популяции под действием:

- а) стабилизирующего отбора,
- б) движущего отбора.

По оси абсцисс отложено количество эволюционных «шагов», по оси ординат – число особей (численность «популяции ГС» постоянна – 1000 особей). В начальный момент 50% особей имеют контур с отрицательной ОС и 50 % особей не имеют такого контура (по: Колчанов и др., 1991).

При движущем отборе ситуация противоположна: преимущество получают особи без отрицательной ОС или ее потерявшие, в то время как особи с отрицательной ОС элиминируют из популяции (рис. 19б). Эффект дестабилизирующего и дестабилизирующего отбора аналогичен движущему.

Таким образом, стабилизирующий и движущий отборы противоположным образом влияют на регуляторные системы организмов, что приводит к *эволюционным качелям*. Поочередно при стабилизирующем отборе возникают и усиливаются отрицательной ОС, а при сменяющем его в ходе эволюции движущем отборе – ослабляются или полностью разрушаются некоторых отрицательных ОС (рис. 20).



ДВИЖУЩИЙ И ДЕСТАБИЛИЗИРУЮЩИЙ ОТБОР

Рис. 20. Эволюционные качели: эффект чередования стабилизирующего и движущего/дестабилизирующего отбора (<http://wwwmgs.bionet.nsc.ru/>).

Анализ функционирования ГС и структуру их графов позволил выделить следующие тенденции, связанные с их усложнением в процессе макроэволюции:

- 1) увеличение числа элементов ГС,
- 2) увеличение числа связей между элементами,
- 3) рост числа уровней иерархии,
- 4) рост числа элементов и связей ГС, работающих в единицу времени и/или в единице пространства;
- 5) рост разнообразия режимов динамического поведения ГС.

Вопросы:

1. На какие теоретические и прикладные вопросы можно ответить с помощью моделей генных сетей?
2. От чего зависят предельные режимы функционирования генных сетей?
3. Что содержится в матрице чувствительности генной сети?
4. Какие выводы можно сделать при анализе влияния мутаций на динамику переменных генной сети?
5. Чем различается влияние стабилизирующего и движущего отборов на эволюцию регуляторных контуров генных сетей?
6. Какие тенденции проявляет структура генных сетей в процессе макроэволюции?

Заключение

Одной из попыток взглянуть на жизнь, как на системный феномен, является концепция генных сетей. Создание математических моделей функционирования живых клеток, детально и с исчерпывающей полнотой описывающих процессы внутриклеточного метаболизма и его генетической регуляции – одна из ключевых проблем системной компьютерной биологии. Ее называют также проблемой электронной клетки.

В настоящее время исследователи далеки от возможности построения полных математических моделей не только целого организма, но даже эукариотических клеток. Что же касается клеток прокариотических, создание их полных математических моделей – дело самого ближайшего будущего.

Пока из-за высокого уровня сложности и комплексности решение подобных задач под силу лишь экспертам-специалистам в молекулярной генетике, биохимии, физиологии, в математических, компьютерных методах моделирования. Что же тормозит прогресс в этой области? Какие проблемы возникают при моделировании динамики ГС?

Их разделяют на два класса. Первые – методологические, возникающие на стадии построения, проверки и работы с моделями. Сюда относятся: анализ огромных объемов переменных экспериментальных данных, отражающих сложные процессы функционирования молекулярно-генетических систем, исследование структурной организации, молекулярных механизмов функционирования, оценка влияния мутаций на функцию ГС, их реконструкция на основе экспериментальных данных и т.д. Сюда же следует отнести задачи исследования устойчивости функционирования ГС, управления ими, исследование эволюции моделируемых биологических систем и т.д.

Второй класс проблем – математические, которые возникают на стадии построения, расчета и адаптации моделей ГС. К ним относятся такие проблемы как решение больших систем дифференциальных уравнений. Ситуация особенно усложняется если модели состоят из системы

дифференциальных уравнений и дискретной части. Не существует пока универсального способа решения обратной задачи (верификация структуры генной сети и определение параметров модели ГС по экспериментальным данным).

Все это типичные проблемы имитационного моделирования сложных биологических систем.

В настоящее время разработки в области ГС получают полезные применения, прежде всего, в медицине и фармакологии. В будущем более совершенные подходы и методы моделирования ГС, по-видимому, станут полезны и в других областях, в т.ч. для научной селекции.

Так, генетический алгоритм, описанный в разделе 5.5, в принципе может быть использован не только для настройки и верификации моделей ГС. Если вместо экспериментальной динамики переменных использовать динамику «идеального генотипа», к которой следует стремиться в процессе селекции, то этот метод позволит подбирать наиболее подходящие структуры и параметры ГС. Далее, используя современные селекционно-генетические технологии, можно реализовать полученные рекомендации (ГС с новыми, «идеальными» параметрами и структурами) и добиться создания организмов с фенотипом, близким к желаемому.

Конечно, нет гарантии, что подобный организм вообще может быть создан с учетом биологических ограничений селектируемого вида, которые должны быть учтены в модели. Но, во-первых, есть надежда, что на это укажет отсутствие положительных результатов применения алгоритма: динамика даже лучших подобранных ГС окажется слишком далекой от идеала. Во-вторых, сам анализ рекомендаций, полученных в результате успешной работы алгоритма, возможно, приведет эксперта к выводу об их нереализуемости и пересмотру постановки задачи.

Приложение 1. Некоторые базы данных по параметрам ферментативных реакций, регуляторным путям и метаболическим генным сетям (<http://www.mgs.bionet.nsc.ru/>).

BRENDA: (<http://www.brenda.bc.uni-koeln.de/>)

KEGG: Regulatory Pathways (<http://www.genome.ad.jp/kegg/regulation.html>)

SPAD: Signal Transduction (<http://www.grt.kyushu-u.ac.jp/spad/>)

CSNDB: Cell Signaling Networks (<http://www.geo.nihs.go.jp/csndb/>)

MIPS: Yeast Pathways

(<http://www.mips.biochem.mpg.de/proj/yeast/pathways/index.html>)

Interactive Fly: Drosophila Genes

(<http://www.sdb.bio.purdue.edu/fly/aimain/1aahome.htm>)

Wnt Signaling Pathway (<http://www.stanford.edu/~rnusse/wntwindow.html>)

TRANSPATH: Gene Regulatory Pathways (<http://www.transpath.gbf.de/>) или (<http://www.gene-regulation.com/pub/databases.html>)

KEGG: Metabolic Pathways (<http://www.genome.ad.jp/kegg/metabolism.html>)

EMP: Enzymes and Metabolic Pathways (<http://www.emp.mcs.anl.gov/>)

WIT: Metabolic Reconstruction (<http://www.wit.mcs.anl.gov/WIT2/>) или (<http://www.unix.mcs.anl.gov/compbio/>)

UM-BBD: Microbial Biocatalysis/Biodegradation

(<http://www.umbbd.ahc.umn.edu/>)

BioCyc: Procarotic Genes and Metabolism (<http://www.biocyc.org/>)

MetaCyc: Metabolism from 150 species (<http://www.metacyc.org/>)

SoyBase: Soybean Metabolism (<http://www.cgsc.biology.yale.edu/metab.htm>)

Boehringer Mannheim: Biochemical Pathways (<http://www.expasy.ch/cgi-bin/search-biochem-index>)

IUBMB-Nicholson Minimaps (<http://www.tcd.ie/Biochemistry/IUBMB-Nicholson/>)

TRRD: Transcription Regulatory Regions Database

(http://www.humbio.ru/Humbio/02022001/erit_gen/x0002aa8.htm).

Приложение 2. Классификация методов математического моделирования динамики генных сетей, основанная на типах формализации процессов (<http://www.wmgs.bionet.nsc.ru/>)

- Булевы сети [Weisbuch, 1986; Kauffman, 1993; Somogyi and Sniegowski, 1996; Huang, 1999; Shmulevich & Kauffman, 2004] и вероятностные Булевы сети [Shmulevich et al., 2002a; Shmulevich et al., 2002b];
- обобщенные логические сети [Thomas & d'Ari, 1990; Thieffry, Thomas, 1995, Thomas et al., 1995; Sanchez et al., 1997; Mendoza et al., 1999];
- Байесовские сети [Friedman et al., 2000; Sachs et al., 2002] и динамические Байесовские сети [Ong et al., 2002; Kim et al., 2004];
- сети Петри [Reddy et al., 1993; Hofstadt & Meineke, 1995, Peleg et al., 2002; Moore & Hahn, 2003], стохастические сети Петри [Goss, Pessoud, 1998], гибридные сети Петри [Matsuno et al., 2000], функциональные гибридные сети Петри [Doi et al., 2004];
- нейронные сети и их обобщения [Taylor, Beiu, 1996; Hanai & Honda H, 2004];
- обобщенный пороговый метод моделирования [Tchuraev, 1991, Prokudina et al., 1991; Tchuraev & Galimzyanov, 2001; Tchuraev & Galimzyanov, 2003];
- подходы, основанные на применении обыкновенных нелинейных дифференциальных уравнений [Bliss et al., 1982; Goldbeter, 1995; Mestl et al., 1995; Endy et al., 1997; Koh et al., 1998; Carrier & Keasling, 1999; Smolen et al., 2000; Chen et al, 2000, 2004; Novak et al, 2001; de Jong, 2002];
- подходы, основанные на применении обыкновенных нелинейных дифференциальных уравнений с запаздывающим аргументом [Landahl, 1969; Mahaffy, 1984; MacDonald, 1989; Smolen et al., 2000, Лихошвай и соавт., 2004];
- обобщенный химико-кинетический метод моделирования [Likhoshvai et al., 2000; Ratushny et al., 2000a; Ratushny et al., 2000b; Лихошвай и соавт., 2001; Nedosekina et al., 2002; Ратушный и соавт., 2003; Ratushny et al., 2003];
- подходы, основанные на применении дифференциальных уравнений в частных производных (модели пространственно распределенных систем) [Turing, 1951; Meinhardt, 1977, 1982, 2000, 2004; Bunow, 1980; Gierer, 1981; Britton, 1986; Kauffman, 1993; Reinitz et al, 1995; Reinitz & Sharp, 1995; Sanchez et al, 1997];
- стохастическое моделирование [Gillespie, 1977; McAdams & Arkin, 1997; Gibson & Bruck, 2000; Gillespie, 2001; Kepler & Elston, 2001; Rao & Arkin, 2003; Rathinam et al., 2003, Turner et al., 2004; Vladimirov & Likhoshvai, 2004];
- логика, основанная на теории нечетких множеств [Pan Du et al., 2005].

Словарь основных терминов и понятий

Алгебра логики (булева алгебра) – специальные алгебраические функции, оперирующие только с 0 и 1.

Биологический триггер – регуляторный контур ГС (см.), который обеспечивает переключение группы генов в одно из двух устойчивых состояний.

Вектор состояний ГС – Исчерпывающей характеристикой ГС в рамках дискретной модели (см.): набор нулей и единиц, упорядоченный по номерам генетических элементов (см.) и обозначающих: 1 – элемент активен, 0 – не активен.

Верификация модели ГС – проверка структуры ГС и «устойчивости» ранее полученных оценок параметров при использовании новых экспериментальных данных.

Генетический алгоритм – математический метод оптимизации, основанный на аналогии с эволюцией популяции и использующийся, в частности, для уточнения модели ГС.

Генетический элемент (ГЭ) – элементарная функциональная единица модели ГС, включающая ген, соответствующие варианты мРНК и процессы синтеза белков.

Генная сеть (ГС) – группа координировано экспрессирующихся генов, выполняющих общую функцию – формирование признака или реализацию процесса.

Гибридный подход в моделировании ГС – разделение сети на две подсистемы. Первая – регуляторная – включает совокупность регуляторных генов и их продуктов. Ее состояния моделируются в рамках дискретных моделей (см.). Вторая – исполняющая – содержит гены, обеспечивающие выполнение различных биохимических процессов, и их продукты. Ее динамику описывают в рамках непрерывных моделей (см.).

Граф ГС двудольный – графическая схема функционирования ГС. В нем кружками обозначаются генетические элементы (см.), квадратами – процессы в ГС, стрелками – участие генетических элементов в процессах.

Граф ГС ориентированный – графическая схема функционирования ГС. Изображается набором вершин, которые обозначают генетические элементы (см.). Некоторые вершины соединены стрелками, которые обозначают взаимодействия генетических элементов.

Дискретные модели ГС – упрощенные модели, где каждый ген в ГЭ может находиться лишь в одном из двух состояний: 1- активен (экспрессируется) и 0 - неактивен. Предполагается, что процесс функционирования ГС состоит из элементарных этапов (шагов). На очередном шаге состояние каждого гена сохраняется или меняется на противоположное.

Интегратор ГС – ключевой элемент (см.), обеспечивающий сигнал-возбуждение при интерференции генных сетей (см.).

Интерференция генных сетей – прохождение возбуждения по определенному компартменту глобальной ГС организма, приводящее к последовательному возбуждению ряда его ГС.

Каскадный принцип – усиление сигнала в ГС по принципу положительной обратной связи (см.).

Кассета генов – группы генов ГС, экспрессией которых управляет один транскрипционный фактор.

Ключевые элементы ГС – центральные гены, центральные регуляторы и интеграторы (см.).

Матрица чувствительности ГС – состоит из чисел, характеризующих относительное влияния вариации отдельных параметров ГС (см.) на предельные режимы (см.) основных ее переменных.

Модель генной сети (ГС) – упрощенная схема функционирования группы наиболее важных генов генной сети (см.).

Мотивы ГС – регуляторные контуры (см), встречающиеся в реальных сетях достоверно чаще остальных.

Настройка модели ГС – получение или уточнение значений параметров в уравнениях, описывающих функционирование ГС.

Непрерывная модель ГС – описывает динамику функционирования ГС уравнениями (например, дифференциальными), оперирующими непрерывными величинами.

Отрицательная обратная связь (ОС) в ГС – регуляторная связь генов, обеспечивающая стабилизацию количественных характеристик ГС на определенном уровне.

Параметры ГС – константы, которые входят в уравнения, моделирующие динамику ГС.

Положительная обратная связь (ОС) в ГС – регуляторная связь генов, обеспечивающая отклонение количественных характеристик ГС от исходных значений и переход к новому функциональному состоянию.

Порог – заранее заданное натуральное число, определяющее границу переключения пороговой функции (см.).

Пороговая функция – принимает значение 0, если аргумент меньше порога (см.), или 1 – если больше или равен. Применяется в дискретных моделях ГС.

Предельные режимы ГС – состояния динамических переменных ГС при длительном функционировании в постоянных внешних условиях. Возможны, в частности, стационарные состояния (см.) и/или предельные циклы – когда некоторые переменные продолжают циклически меняться во времени при постоянных внешних условиях.

Регуляторная и исполняющая компоненты ГС – регуляторные гены, кодирующие ферменты управления исполняющими генами, которые непосредственно синтезируют метаболиты.

Регуляторный контур ГС – группа генов и совокупность регуляторных механизмов (см.), посредством которых эти гены регулируют активность экспрессии генов ГС.

Регуляторный механизм ГС – элементарный процесс, посредством которого некоторая совокупность генетических элементов (см.) действуют на конкретный генетический элемент, уменьшая или увеличивая скорость синтеза его продукта.

Стационарное состояние – равновесие (точке покоя) по значениям одного или нескольких переменных ГС.

Узкие места ГС – лимитирующие (ключевые) параметры ГС (см.), отклонения которых влечет резкое изменение функционирования ГС.

Центральный ген ГС – ген, обеспечивающий координацию функций остальных генов сети.

Центральный регулятор ГС – транскрипционный фактор, который активируют или репрессирует одновременно несколько генов – кассету генов (см.).

Эволюционные качели – возникновение и усиление отрицательных обратных связей (см.) в регуляторных контурах ГС (см.) при стабилизирующем отборе, а при сменяющем его в ходе эволюции движущем отборе – ослабление или разрушение части отрицательных обратных связей.

Ядро ГС – группа основных координировано работающих генов ГС.

Литература

1. Ананько Е.А., Лихошвай В.А., Колпаков Ф.А. и др. Электронная библиотека GeneNet: описание и моделирование генных сетей животных и растений// Вторая Всероссийская научная конференция «Электронные библиотеки: перспективные методы и технологии, электронные коллекции», 2000, Протвино, стр. 66-72
2. Елькин Ю.Е. Генные сети.
http://www.mathcell.ru/ru/obzors/obzor_Elkin.shtml
3. Колчанов Н.А., Шиндялов И.Н. Теоретическое исследование эволюции регуляторных контуров при различных типах отбора// Проблемы генетики и теории эволюции. Ред. В.К. Шумный, А.О. Рувинский. Новосибирск: Наука, 1991. стр. 268-279.
4. Колчанов Н.А., Ананько Е.А., Колпаков Ф.А. и др. Генные сети// Мол. Биология, 2000, т. 34, стр. 533 - 544.
5. Колчанов Н.А., Суслов В.В., Гунбин К.В.. Моделирование биологической эволюции: регуляторные генетические системы и кодирование сложности биологической организации// Вестник ВОГиС, 2004, т. 8., №2. стр. 86-99.
6. Колчанов Н.А. Генные сети липидного метаболизма// Бюллетень СО РАМН, 2006, №2 (120), стр. 29-42
7. Лихошвай В.А., Матушкин Ю.Г., Ратушный А.В., Ананько Е.А., Игнатьева Е.В., Подколотная О.В. 2001. Обобщенный химико-кинетический метод моделирования генных сетей. // Молекул. биология, 2001, т. 35, стр. 1072 - 1079.
8. Лихошвай В.А., Матушкин Ю.Г., Фадеев СИ. О связи графа генной сети с качественными режимами ее функционирования// Мол. биология. 2001, т. 35, № 6, стр. 1080-1087.
9. Ратушный А.В. Исследование динамики и функций ГС методами математического моделирования// Автореф. канд. биол. наук., Новосибирск, 2006. 20 с.
10. Чураев Р.Н., Галимзянов А.В. Моделирование реальных эукариотических управляющих генных подсетей на основе метода обобщенных пороговых моделей// Мол. Биология, 2001, т. 35, стр. 1088-1094.

Содержание

	Введение.....	3
1.	Принципы функционирования и модели генных сетей.....	4
2.	Генные сети про- и эукариот.....	7
3.	Цели и задачи моделирования генных сетей.....	11
4.	Классификация генных сетей и их составляющие	14
5.	Построение моделей генных сетей.....	19
5.1.	Базы данных.....	20
5.2.	Графические схемы функционирования генных сетей.....	22
5.3.	Дискретный способ моделирования генных сетей.....	26
5.4.	Непрерывные модели	28
5.5.	Проблема адекватности моделей.....	31
6.	Методика и результаты анализа моделей генных сетей.....	34
6.1.	Анализ функционирования генных сетей.....	35
6.2.	Анализ влияния мутаций.....	37
6.3.	Эволюция генных сетей.....	39
	Заключение.....	44
	Приложение 1.....	46
	Приложение 2.....	47
	Словарь основных терминов и понятий.....	48
	Литература.....	51