

УДК 57.011:577.121:57.087.2:57.05:57.052:575.85:519.716

ГЕННЫЕ СЕТИ

© 2013 г. **Н.А. Колчанов^{1,2,3}, Е.В. Игнатьева^{1,2}, О.А. Подколотная¹,
В.А. Лихошвай^{1,2}, Ю.Г. Матушкин^{1,2}**

¹ Федеральное государственное бюджетное учреждение науки Институт цитологии и генетики
Сибирского отделения Российской академии наук, Новосибирск, Россия;
e-mail: kol@bionet.nsc.ru

² Новосибирский национальный исследовательский государственный университет,
Новосибирск, Россия;

³ НИЦ «Курчатовский институт», Москва, Россия

Поступила в редакцию 5 сентября 2013 г. Принята к публикации 1 ноября 2013 г.

ВВЕДЕНИЕ

Исследования последнего десятилетия свидетельствуют о том, что подавляющее большинство фенотипических признаков человека, животных, растений и микроорганизмов (молекулярных, биохимических, клеточных, физиологических, морфологических, поведенческих и т. д.) контролируются очень сложным образом и что в основе их формирования лежат генные сети, т. е. группы координированно функционирующих генов, взаимодействующих друг с другом как через свои первичные продукты (РНК и белки), так и через разнообразные метаболиты и другие вторичные продукты функционирования генных сетей.

БАЗОВЫЕ ОСОБЕННОСТИ ОРГАНИЗАЦИИ ГЕННЫХ СЕТЕЙ

В каждой генной сети можно выделить несколько обязательных типов структурных и функциональных компонентов: 1) группа координированно экспрессирующихся генов, составляющая ядро сети; 2) белки, кодируемые этими генами; 3) низкомолекулярные компоненты (гормоны и другие сигнальные молекулы, энергетические компоненты, метаболиты); 4) отрицательные и положительные обратные связи, стабилизирующие параметры генной сети на определенном уровне или, напротив, отклоняющие их от исходного значения, обес-

печивая переход к новому функциональному состоянию. Генные сети могут включать такие крупные функциональные модули, как: а) метаболические пути; б) пути передачи сигналов от клеточных мембран в ядра клеток, обеспечивающие активацию или подавление транскрипции генов в ответ на внешние по отношению к клетке стимулы (рис. 1, а).

Генные сети могут функционировать как на уровне отдельных клеток (тогда данный тип клеток будет включать все структурные компоненты генной сети), так и на уровне ткани, органа либо целого многоклеточного организма. В этом случае в реализации обратных связей генной сети будут участвовать гены и белки, экспрессирующиеся в различных типах клеток.

Рассмотрение структурно-функциональной организации генных сетей (Колчанов и др., 2000) позволяет выделить следующие базовые принципы их организации.

1. Существование большого разнообразия молекулярных механизмов, обеспечивающих функционирование обратных связей, в том числе за счет изменения интенсивности про-

Генные сети – группы координированно функционирующих генов, взаимодействующих друг с другом как через свои первичные продукты (РНК и белки), так и через разнообразные метаболиты и другие вторичные продукты функционирования генных сетей.

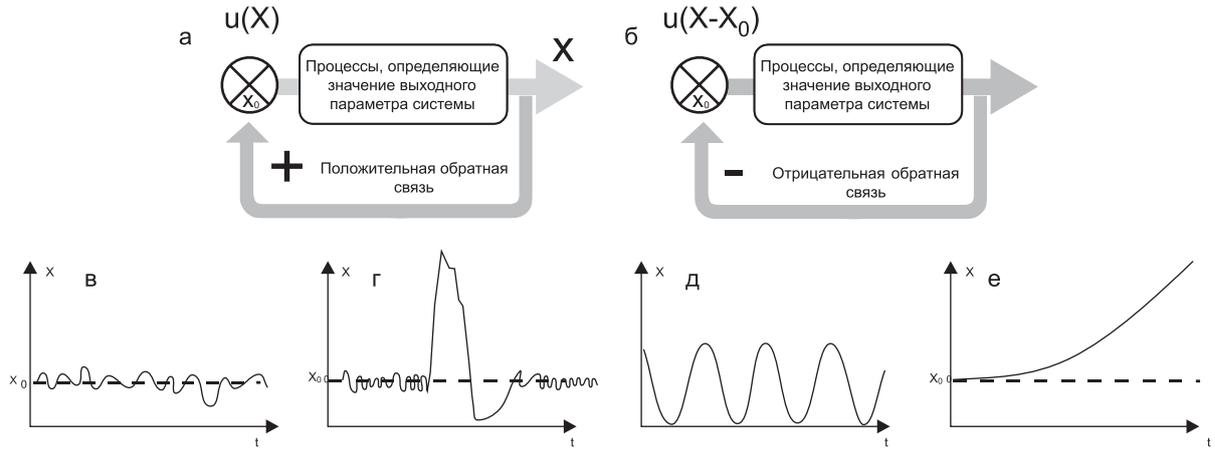


Рис. 2. Регуляторные контуры с обратными связями и типы динамики критических параметров, контролируемых генными сетями.

а – регуляторный контур с отрицательной обратной связью; б – регуляторный контур с положительной обратной связью; в – постоянство контролируемых параметров для генных сетей гомеостаза; г – выраженное отклонение контролируемых параметров с последующим возвращением к норме для генных сетей стрессового ответа; д – осцилляция контролируемых параметров для генных сетей циклических процессов; е – монотонное отклонение параметров от текущего состояния для некоторых генных сетей морфогенеза.

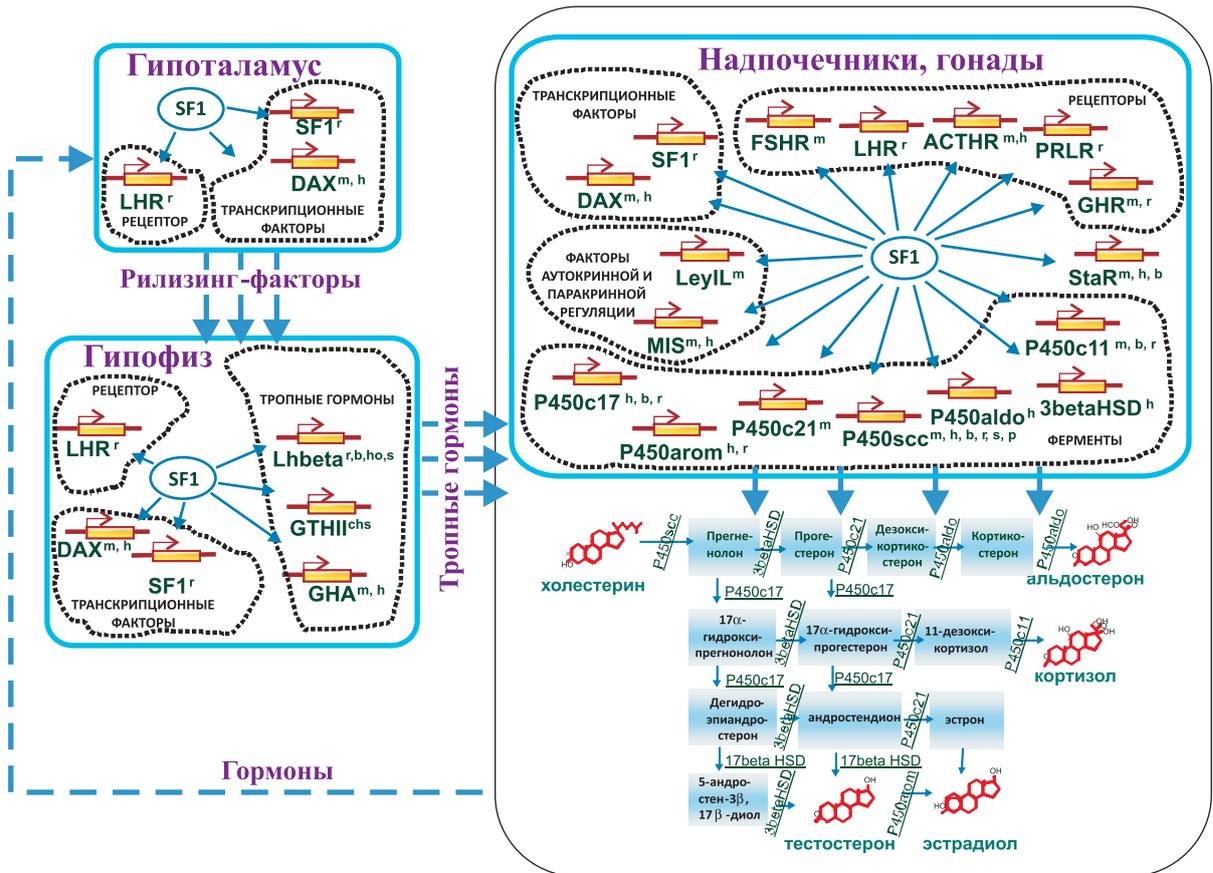


Рис. 3. Генные сети регуляции стероидогенеза у млекопитающих.

Функционирование генных сетей трех иерархических уровней регуляции контролируется транскрипционным фактором SF1. Реконструировано по: (Бусыгина и др., 2003. С. 364–382). В обозначениях генов h – человек, m – мышь, r – крыса, b – корова, chs – лосось, ho – лошадь, s – овца, p – свинья.

ляторных процессов. Например, цикл трикарбоновых кислот у *E. coli K-12*, снабжающий клетку энергией, включает 1882 регуляторных и только 139 метаболических процессов. То есть на один рабочий процесс приходится более 10 контролирующих процессов. Наличие мощной регуляторной компоненты проявляется и на уровне экспрессии генов: регуляторные районы генов эукариот содержат большое количество сайтов связывания транскрипционных факторов, что обеспечивает огромное разнообразие вариантов регуляции экспрессии одного и того же гена. Таким образом, характерное свойство генных сетей, функционирующих в живых системах, состоит в том, что в них регуляторная компонента на порядки больше, чем рабочая. Это касается как метаболической, так и генетической компоненты генных сетей, и проявляется на каждом иерархическом уровне их организации.

МЕТОДЫ РЕКОНСТРУКЦИИ ГЕННЫХ СЕТЕЙ

Среди множества подходов к реконструкции генных сетей можно выделить следующие основные группы.

1. Ручная реконструкция генных сетей экспертами с использованием специальных программных средств – редакторов генных сетей.

2. Извлечение знаний о молекулярно-генетических взаимодействиях и семантических ассоциациях различного типа из текстов научных публикаций компьютерными методами (методы text-mining).

3. Методы автоматической интеграции гетерогенных данных о молекулярно-генетических взаимодействиях (ДНК-белковых и белок-белковых и т. д.), метаболических и сигнальных путях, генных сетях и других из различных источников данных, основанные на использовании онтологических подходов.

Заметим, что современные системы для реконструкции и анализа генных сетей, как правило, построены на сочетании нескольких подходов и используют, кроме того, различные подходы, основанные на многомерном анализе данных и компьютерных методов предсказания (Taylor, Singhal, 2009; Haibe-Kains *et al.*, 2012; Wang *et al.*, 2013).

Онтология – согласование понятий предметной области, их определений и атрибутов, отношений между ними, способов их описания и использования, а также связанных с ними аксиом и правил вывода.

Первый исторически возникший способ реконструкции генных сетей был основан на ручном экстрагировании информации из научных публикаций и других источников экспериментальных данных. При этом ввод информации о структуре генных сетей и функции их элементов осуществлялся с помощью специальных графических и текстовых редакторов. Этот подход является трудоемким, но по-прежнему остается актуальным при необходимости реконструкции максимально точной и подробной генной сети. Наиболее известными из систем, использующих данный подход, являются Reactome, KEGG, MetaCyc, WikiPathways и др. (Kanehisa *et al.*, 2006; Croft *et al.*, 2011; Caspi *et al.*, 2012; Kelder *et al.*, 2012). К этому же классу относится разработанная в ИЦиГ СО РАН в 1998 г. система GeneNet, в которой все компоненты генных сетей делятся на два больших класса: 1) элементарные структуры (гены, РНК, белки, метаболиты) и 2) элементарные события (реакции и регуляторные события), позволяющая экспертам вводить в базу данных GeneNet информацию о генных сетях как с использованием графических образов элементарных структур и событий, так и в текстовой форме (Колчанов *et al.*, 2000).

Стремительный рост объемов экспериментальных данных и, как следствие, аккумулирующих их баз данных, научных публикаций, патентов явился предпосылкой для создания методов реконструкции генных сетей на основе автоматического анализа тестов научных публикаций и баз данных (Rzhetsky *et al.*, 2004; Cooper *et al.*, 2005; Harmston *et al.*, 2010). Одна из систем построения таких сетей на основе автоматического анализа текстов, названная AND System (Associative Network Discovery System), была создана в ИЦиГ СО РАН (Demchenkova *et al.*, 2011). Система AND System может быть настроена на конкретную область биологических знаний. Настройка осуществляется путем построения

специфических для нее словарей объектов, включая гены, белки, малые молекулы, типы клеток тканей и органов, метаболические пути и пр. Непрерывный рост как объемов, так и типов экспериментальных данных потребовал разработки методов интеграции гетерогенной информации, полученной в результате экспрессионного и ChIP-seq анализа, исследования ДНК-белковых, белок-белковых взаимодействий, а также исследования протеомов, метаболомов, проведенного на уровне клеток, тканей, организмов в различных экспериментальных условиях (Mungall *et al.*, 2007; Mi, Thomas, 2009; Zamboni *et al.*, 2012). Одним из основных этапов семантической интеграции такого типа данных является построение онтологий. Примерами широко используемых онтологий служат онтологии BioPAX и GO (Gene Ontology Consortium ..., 2006; Demir *et al.*, 2010).

Еще одно направление в реконструкции генных сетей основано на использовании данных о динамике экспрессии больших массивов генов, полученных с помощью ДНК-чипов или метода RNA-seq, который позволяет секвенировать транскрибируемые участки генома. На основе серии таких экспериментов в различных временных точках можно с помощью методов биоинформатики реконструировать сеть межгенных взаимодействий, отвечающих за наблюдаемую динамику всего массива изучаемых генов. Качественно идею этого подхода можно представить на простых примерах (рис. 4). Если два гена (A и B) имеют очень похожие профили экспрессии, как это показано для первого случая (рис. 4, а), можно предположить наличие общего регулятора их экспрессии. Отметим, что это только одна из возможных гипотез. Если профиль экспрессии гена B повторяет профиль

экспрессии гена A с неким временным запаздыванием τ (рис. 4, б), это может означать, что ген B регулируется геном A. Если профиль экспрессии гена A возрастает, а профиль экспрессии гена B падает (рис. 4, в), это может означать, что они находятся во взаимно антагонистических регуляторных отношениях. Профиль экспрессии генов на рис. 4, г может быть следствием целого ряда разнообразных процессов, позволяющих построить множество гипотез.

Анализ такого рода экспрессионных данных осуществляется с привлечением различных методов многомерной статистики, результатом чего является построение моделей межгенных взаимодействий, отвечающих за динамику всего массива изучаемых генов. Даже из этих простых примеров становится ясно, что анализ динамических механизмов регуляции экспрессии генов в клетке является одной из сложнейших задач системной биологии. Конечной целью этого направления исследований является построение цельной картины биологических процессов, протекающих на всех иерархических уровнях организма, что значительно сложнее, чем просто сумма отдельных частей. Решение этой задачи может быть осуществлено только с привлечением самых современных методов молекулярной биологии и генетики, а также и мощного аппарата биоинформатики и компьютерной техники.

МАЛЫЕ ГЕННЫЕ СЕТИ И ТЕОРИЯ ИХ ДИНАМИКИ

Для понимания фундаментальных особенностей динамики регуляторных генных сетей (ГС) очень ценную информацию дают компьютерный анализ и моделирование генных сетей малой размерности. Теория динамики таких

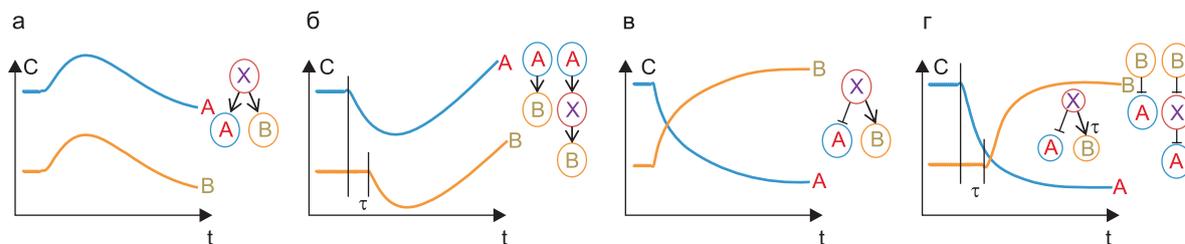


Рис. 4. Примеры, описывающие потенциально возможные особенности межгенных взаимодействий, отвечающих за динамику экспрессии генов.

На оси OY представлен уровень экспрессии гена (в условных единицах), на оси OX – время (в условных единицах).

генных сетей разрабатывается В.А. Лихошваем с коллегами в ИЦиГ СО РАН и ИМ СО РАН (Колчанов и др., 2001; Лихошвай и др., 2001; Лихошвай и др., 2003; Демиденко и др., 2004, 2009; Лихошвай и др., 2004; Фадеев и др., 2004; Голубовский и др., 2010; Demidenko *et al.*, 2010; Лихошвай, 2013).

В данной теории элементарными единицами ГС являются генетические элементы G . Генетический элемент представляет собой пару $G = (g, p)$, состоящую из гена g и целевого (конечного) продукта p , синтезируемого с него (рис. 5, а). Активность генетического элемента G характеризуется скоростью синтеза V целевого продукта p и скоростью его деградации/диссипации D . Целевым продуктом может быть как молекула РНК, так и белок. Если изменение концентрации продукта p_1 снижает/увеличивает скорость изменения концентрации продукта p_2 , то полагаем, что между генетическими элементами G_1 и G_2 существует регуляторная (ориентированная) связь $\sigma = \sigma(G_1, G_2)$. Механизмы осуществления регуляторных связей могут быть самыми различными. На рис. 5, б показан пример реализации регуляторной связи посредством влияния гомомультимера белка p_1 на скорость синтеза продукта p_2 .

Скорость синтеза целевого продукта p описывается уравнением общего вида:

$$\frac{dp}{dt} = V(\bar{p} \in B) - D(\bar{p} \in C), \quad (1)$$

где V и D – функции регуляции активности синтеза целевого продукта p и скорости его деградации/диссипации; B и C – множества целевых продуктов, участвующих в регуляции процессов синтеза и деградации/диссипации соответственно. Данные уравнения также могут содержать запаздывающие аргументы, учитывающие время, затрачиваемое на синтез целевых продуктов, модификацию и доставку в компартмент функционирования.

Основной задачей теории генных сетей является выявление причинно-следственных связей между структурно-функциональной организацией генных сетей и их динамическими свойствами. При этом под структурно-функциональной организацией ГС понимается совокупность всех регуляторных связей и молекулярно-генетических механизмов, их осуществляющих, а под динамическими свой-

Динамический портрет генной сети – совокупность траекторий динамики изменения концентраций синтезируемых веществ, реализуемых в некоторой заданной параметрической области из некоторой заданной области начальных данных.

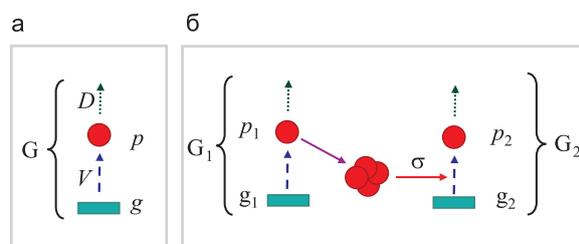


Рис. 5. Генетические элементы и регуляторные связи ГС.

G, G_1, G_2 – генетические элементы, σ – регуляторная связь, V – синтез целевого продукта, D – деградация/диссипация целевого продукта.

ствами понимаются характеристики траекторий концентраций продуктов, которые могут реализовываться в данной системе при определенных значениях параметров и начальных данных. Часто изменение параметров или начальных данных ведет к изменению качественных характеристик динамического поведения системы. Например, при одних условиях концентрация всех продуктов со временем стремится к некоторому постоянному значению – стационару, а при других – наблюдаются незатухающие колебания. Поэтому важно знать динамические свойства ГС не просто для отдельных наборов значений параметров и начальных данных, но и для некоторых областей значений параметров и начальных значений. В этом случае уместно говорить о динамическом портрете ГС.

В общем случае системы уравнений, описывающие ГС, имеют настолько общий вид, что их полное изучение вряд ли возможно. Поэтому в теории генных сетей выделяются более узкие классы, которые, с одной стороны, допускают более полное изучение, а с другой – позволяют выявить новые биологически значимые закономерности.

В частности, изучаются генные сети с отрицательными механизмами регуляции, описываемые системами уравнений:

$$\frac{dp_i(t)}{dt} = \frac{\alpha_i}{1 + \sum_{j \in B_i} \left[\frac{p_j(t - \tau_i)}{K_{i,j}} \right]^{\gamma_{i,j}}} - \beta_i p_i, \quad i = 1, \dots, n, \quad (2)$$

здесь p_i – концентрация белка, синтезируемого i -м генетическим элементом; α_i – базисная скорость синтеза i -го белка в отсутствие регуляторов; β_i – константа скорости деградации i -го белка; $K_{i,j}$ – константа ингибирования скорости синтеза белка p_i белком p_j ; γ_{ij} – константа, характеризующая степень нелинейности механизма репрессии синтеза белка p_i белком p_j ; n – число генетических элементов G_i ; $B_i \subseteq \{1, \dots, n\}$ – множество номеров белков, являющихся регуляторами активности синтеза белка p_j ; t – текущее время; τ_i – суммарное время, необходимое для синтеза, модификации и транспорта конечного продукта p_i в компартмент его функционирования.

В них (а) регулируемыми являются только стадии инициации синтеза конечных продуктов

(для простоты – белков), стадии деградации/диссипации не регулируются, но каждый продукт имеет конечное время функциональной жизни в пространственном компартменте функционирования генной сети; (б) функциональными молекулами (белками-регуляторами) являются гомомультимеры; (в) белки-регуляторы конкурируют между собой в процессе регуляции активности экспрессии каждого конкретного генетического элемента.

На рис. 6, а представлена генная сеть, содержащая четыре генетических элемента. Активность каждого генетического элемента регулируется двумя белками-репрессорами (красные стрелки). Для этой сети модель без учета стадий промежуточного синтеза, созревания и транспорта целевых продуктов ($\tau_i = 0$) имеет вид:

$$\begin{cases} \frac{dp_1}{dt} = \alpha_1 \left/ \left[1 + \left(\frac{p_3}{K_{1,3}} \right)^{\gamma_{1,3}} + \left(\frac{p_4}{K_{1,4}} \right)^{\gamma_{1,4}} \right] \right. - \beta_1 p_1, & \frac{dp_2}{dt} = \alpha_2 \left/ \left[1 + \left(\frac{p_1}{K_{2,1}} \right)^{\gamma_{2,3}} + \left(\frac{p_4}{K_{2,4}} \right)^{\gamma_{2,4}} \right] \right. - \beta_2 p_2, \\ \frac{dp_3}{dt} = \alpha_3 \left/ \left[1 + \left(\frac{p_2}{K_{3,2}} \right)^{\gamma_{3,3}} + \left(\frac{p_4}{K_{3,4}} \right)^{\gamma_{3,4}} \right] \right. - \beta_3 p_3, & \frac{dp_4}{dt} = \alpha_4 \left/ \left[1 + \left(\frac{p_2}{K_{4,2}} \right)^{\gamma_{4,2}} + \left(\frac{p_3}{K_{4,3}} \right)^{\gamma_{4,3}} \right] \right. - \beta_4 p_4. \end{cases} \quad (3)$$

Для нее при $K_{i,j} = 1$, $\beta_i = 1$, $\gamma_{i,j} = 4$, $\alpha_1 = \alpha_2 = \alpha_3 = 5$, $\alpha_4 = 10$ численно выявляется единственный стационарный режим, к которому со временем стремятся концентрации продуктов p_i (рис. 6, а). С биологической точки зрения это означает, что соответствующая генетическая система имеет только один устойчивый режим

функционирования: четвертый генетический элемент активно синтезирует белок-репрессор, который подавляет синтез других белков.

На рис. 6, б показана вышерассмотренная генная сеть с дополнительной регуляторной связью (синяя стрелка). В этом случае система уравнений принимает вид:

$$\begin{cases} \frac{dp_1}{dt} = \alpha_1 \left/ \left[1 + \left(\frac{p_3}{K_{1,3}} \right)^{\gamma_{1,3}} + \left(\frac{p_4}{K_{1,4}} \right)^{\gamma_{1,4}} \right] \right. - \beta_1 p_1, & \frac{dp_2}{dt} = \alpha_2 \left/ \left[1 + \left(\frac{p_1}{K_{2,1}} \right)^{\gamma_{2,3}} + \left(\frac{p_4}{K_{2,4}} \right)^{\gamma_{2,4}} \right] \right. - \beta_2 p_2, \\ \frac{dp_3}{dt} = \alpha_3 \left/ \left[1 + \left(\frac{p_2}{K_{3,2}} \right)^{\gamma_{3,3}} + \left(\frac{p_4}{K_{3,4}} \right)^{\gamma_{3,4}} \right] \right. - \beta_3 p_3, & \frac{dp_4}{dt} = \alpha_4 \left/ \left[1 + \left(\frac{p_1}{K_{4,1}} \right)^{\gamma_{4,3}} + \left(\frac{p_2}{K_{4,2}} \right)^{\gamma_{4,2}} + \left(\frac{p_3}{K_{4,3}} \right)^{\gamma_{4,3}} \right] \right. - \beta_4 p_4. \end{cases} \quad (4)$$

Здесь член уравнения выделен и соответствует добавленной регуляторной связи. Ее добавление и задание значений $K_{4,1} = 1$, $\gamma_{4,1} = 4$ приводят к радикальному изменению динамики поведения системы. В этом случае для соответствующей генной сети прогнозируются два качественно различ-

ных устойчивых состояния: стационарный режим, который подобен стационару, описанному выше (рис. 6, б (2)); циклический режим, при котором концентрации белков p_1 , p_2 и p_3 претерпевают значительные по амплитуде периодические изменения, а концентрация белка p_4 колеблется с малой амплитудой, всегда

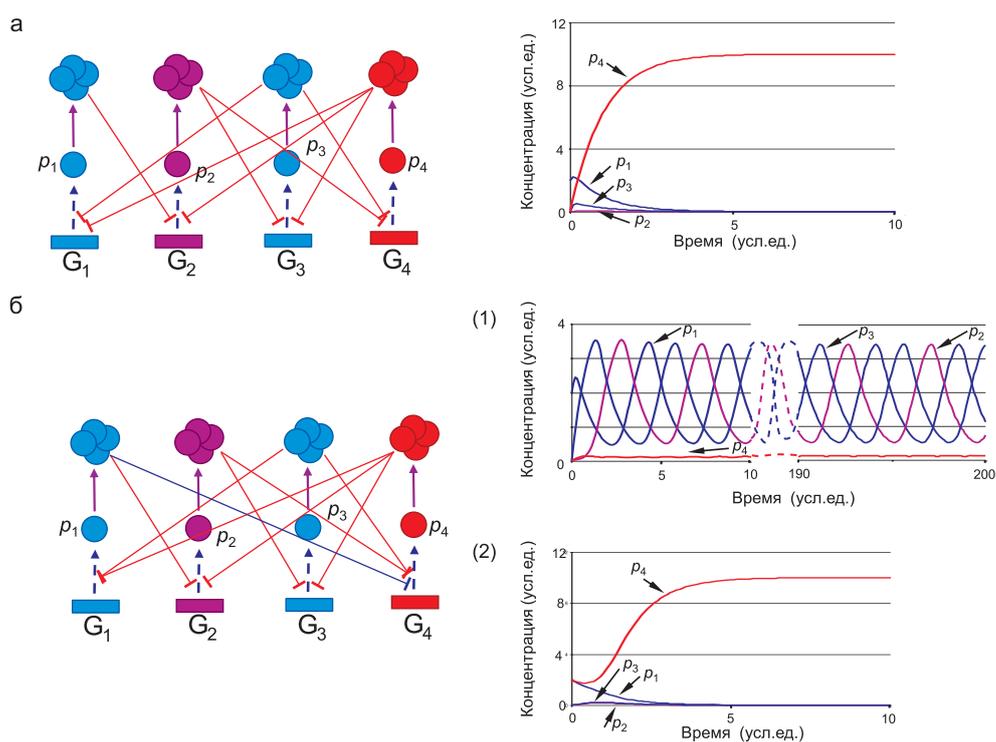


Рис. 6. Пример качественного изменения динамики поведения малой генной сети в результате добавления одной регуляторной связи.

На панелях а и б слева представлена схема взаимодействия объектов в генной сети, справа – динамическое поведение сети во времени.

оставаясь много ниже значений, необходимых для его функционирования в качестве репрессора (рис. 6, б (1)). Здесь следует подчеркнуть, что рассмотренная ситуация весьма реалистична, так как исчезновение существующей или появление новой регуляторной связи может быть результатом одиночной аминокислотной замены в регуляторном белке или нуклеотидной замены в ДНК-сайте его связывания.

Данный пример интересен также тем, что из него прогнозируется, что генная сеть (рис. 6, б) обладает эпигенетической памятью – зависимостью ее динамики от начальных концентраций белков в клетке, которые могут, например, задаваться ее предыдущим делением.

Анализ показывает, что даже простейшие генные сети могут иметь очень большое разнообразие режимов функционирования. В качестве примера на рис. 7, а показан граф малой генной сети, содержащей 7 генетических элементов и 12 регуляторных взаимодействий между ними. Вид соответствующей системы не

приводим, так как ее легко выписать из системы (2) и структурного графа, представленного на рис. 7, а. На рис. 7, б показан фрагмент динамического портрета генной сети, прогнозируемого на основе численного анализа модели.

В начальный момент семимерная траектория концентраций продуктов генной сети (кривая f_1) стартует в гиперпространстве $p_1 = p_5, p_2 = p_6, p_3 = p_7$ из точки $p_1 = p_5 = 1, p_2 = p_3 = p_4 = p_6 = p_7 = 0$ и притягивается к предельному циклу f_2 , лежащему в данном многообразии. В точке А на генную сеть оказывается слабое внешнее воздействие, выводящее ее за пределы данного гиперпространства, после чего концентрационная траектория притягивается к новому стабильному циклу (кривая f_3). Для данной генной сети численно выявляются три таких стабильных цикла и одно устойчивое стационарное состояние. Имеются также пять дополнительных циклов, лежащих в различных многообразиях. Возможно, имеются и другие замкнутые траектории. Из рассмотрения этой модели можно сделать несколько выводов.

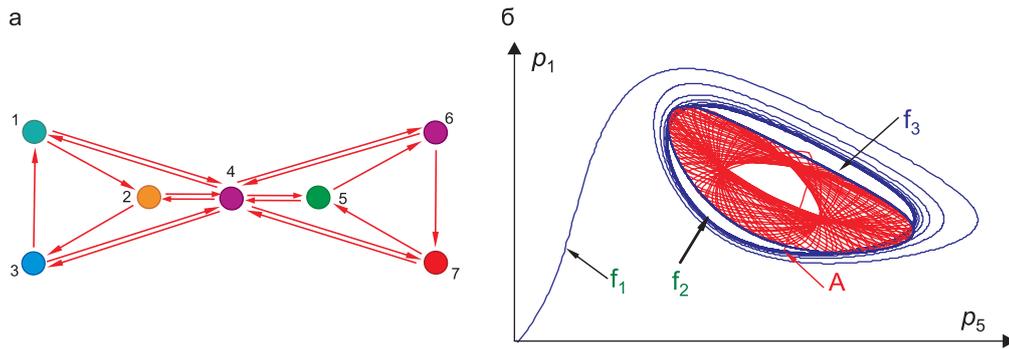


Рис. 7. Сложное динамическое поведение малой генной сети.

а – структурный граф малой ГС: вершины – генетические элементы, стрелки – регуляторные связи отрицательного типа действия; б – фазовый портрет эволюции малой ГС: в нулевой момент времени траектория стартует в гиперпространстве $p_1 = p_5, p_2 = p_6, p_3 = p_7$ (кривая f_1) и притягивается к визуальному циклу, лежащему на данном аттракторе (кривая f_2); в точке А на систему оказывается слабое внешнее воздействие, выводящее ее за пределы данного гиперпространства (красная кривая); после стимула А система притягивается к новому визуальному циклическому аттрактору (кривая f_3). Расчеты проведены при следующих значениях параметров: $K_{i,j} = \beta_i = 1, \gamma_{i,j} = 4, \alpha_i = 5$.

1. Здесь имеет место эффект эпигенетической памяти системы – зависимость ее динамики от начальных концентраций белков в клетке. Впервые, еще в начале 70-х годов прошлого века (т. е. еще до начала эпохи эпигенетики), это показал методами математического моделирования российский ученый Р.Н. Чураев, который работал в те годы в ИЦиГ СО РАН (Чураев, 2010).

2. Эти результаты показывают, что биологическая сложность кодируется, прежде всего, не на уровне геномов, а на уровне генных сетей, которые даже при малом количестве генетических элементов могут обеспечивать огромное разнообразие режимов функционирования.

3. Следует иметь в виду, что на протяжении большей части времени генная сеть пребывает в том или ином устойчивом состоянии или близко к нему, которых хотя и много, но все же конечное количество. А суть функционирования любой генной сети заключается в переходах между этими устойчивыми состояниями, из которых наибольшее значение имеют два типа – стабильные стационарные состояния и стабильные циклы. Именно эти состояния оцениваются естественным отбором.

На рис. 8 показан авторепрессиллятор – сверхмалая генная сеть, состоящая всего из одного генетического элемента, обеспечивающего по механизму отрицательной обратной связи регуляцию экспрессии гена g , кодирующего белок-авторепрессор p . Модели таких систем

изучаются уже давно и различными математическими методами. Модель (2), примененная к авторепрессиллятору, преобразуется к следующему виду:

$$\frac{dp(t)}{dt} = \frac{\alpha}{1 + \left[\frac{p(t-\tau)}{K} \right]^\gamma} - \beta p. \quad (5)$$

Если предполагать, что процессы синтеза белков протекают быстро, то τ мало и с течением времени в системе устанавливается постоянный уровень репрессора (рис. 8, а). Если же последовательность процессов транскрипции, трансляции, транспорта и модификации молекул занимает достаточно времени, то в системе устанавливается устойчивая циклическая динамика изменения уровня белка авторепрессора p (рис. 8, б, в). В этом случае период и амплитуда цикла зависят от параметров модели (рис. 8, б, в).

В качестве примера природной генной сети малого размера рассмотрим генную сеть регуляции циркадного ритма (рис. 9). Такие генные сети функционируют практически в каждой клетке живого организма млекопитающих и синхронизируются центральным циркадным осциллятором, локализованным у млекопитающих в супрахиазматических ядрах гипоталамуса.

Минимальная (коровая) генная сеть циркадного осциллятора млекопитающих основана на регуляции транскрипции небольшой группы генов по механизму отрицательной обратной

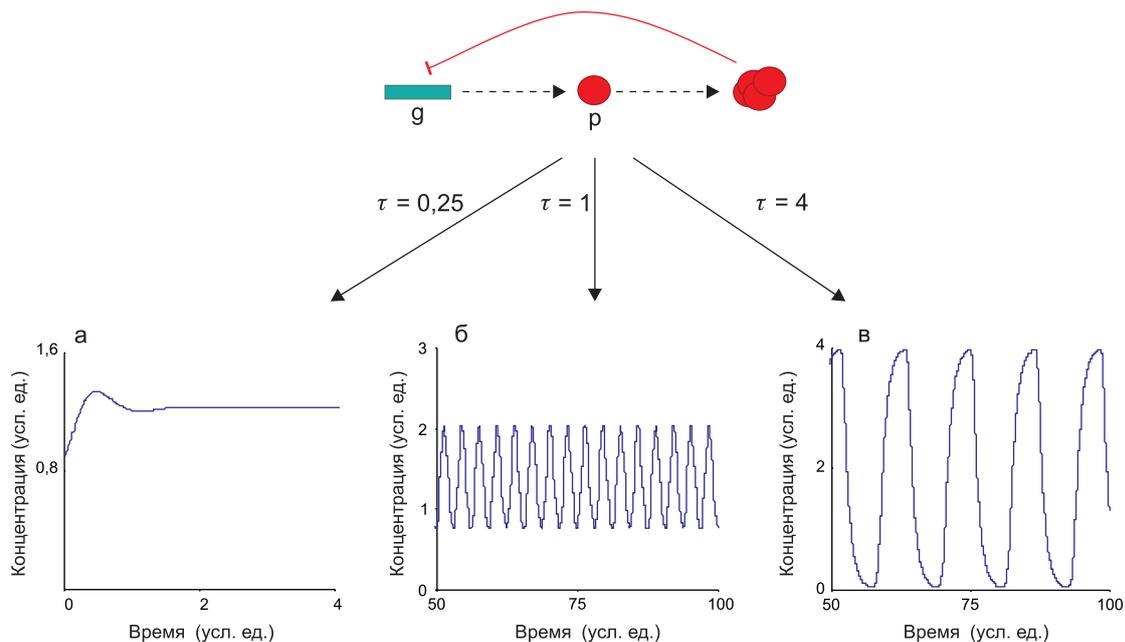


Рис. 8. Различные режимы функционирования авторепрессилатора.

а – стационар; б – цикл с малыми периодом и амплитудой; в – цикл с большими периодом и амплитудой. Расчеты проведены по модели (5), во всех расчетах полагали $\alpha = 4, \gamma = 4, \beta = 1$, а в качестве начальной функции использовали постоянную функцию $p(t) = 0,9, -\tau \leq t \leq 0$.

Авторепрессилатор – генетический элемент, обеспечивающий по механизму отрицательной обратной связи регуляцию экспрессии гена, кодирующего белок-авторепрессор.

связи. Гетеродимерный транскрипционный фактор CLOCK/BMAL1 активирует транскрипцию генов *Per* и *Cry*, а димер *Per/Cry* за счет белок-белковых взаимодействий подавляет активность CLOCK/BMAL1. Это в свою очередь приводит к замыканию регуляторного цикла за счет снижения уровня транскрипции генов *Per*

и *Cry*, обеспечивая формирование циркадного ритма. Стабильность этого цикла поддерживается двумя дополнительными регуляторными контурами. В положительном регуляторном контуре CLOCK/BMAL1 активирует транскрипцию гена *Ror α*, а фактор *Ror α* активирует транскрипцию гена *Bmal1*. В отрицательном регуляторном контуре CLOCK/BMAL1 активирует транскрипцию гена *Rev-erba*, кодирующего негативный регулятор *Rev-erba*, который подавляет транскрипцию гена *Bmal1*. При этом белки *Ror α* и *Rev-erba* конкурируют за связывание с сайтом в регуляторных районах гена *Bmal1*. Эти две дополнительные петли обеспечивают

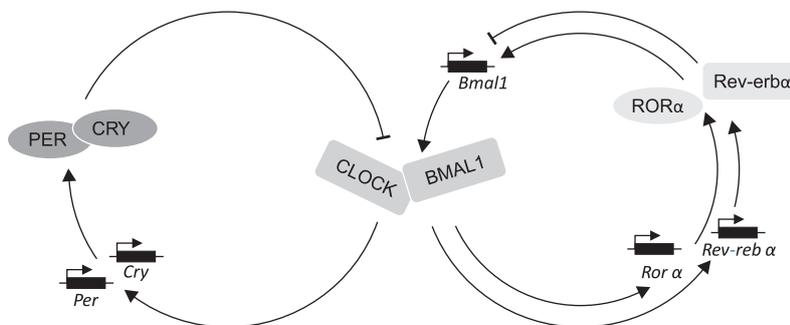


Рис. 9. Минимальная генная сеть циркадного осциллятора млекопитающих.

устойчивое функционирование циркадного осциллятора с периодом, близким к 24 ч.

РЕАЛЬНАЯ СЛОЖНОСТЬ ГЕННЫХ СЕТЕЙ

Генная сеть регуляции уровня холестерина

Рассмотрим более детально механизмы функционирования генных сетей и их реальную сложность на примере генной сети регуляции внутриклеточного уровня холестерина в клетках человека и животных.

Холестерин – незаменимая структурная компонента клеточных мембран и наружного слоя липопротеинов плазмы крови, предшественник целого ряда других стероидов, а именно: кортикостероидов, половых гормонов, желчных кислот и витамина D. Отклонения от нормального уровня холестерина приводят к тяжелым патологиям, включая атеросклероз сосудов головного мозга, сердечной мышцы и других органов.

У человека холестерин поступает в организм с пищей, а также синтезируется в клетках многих тканей *de novo* из предшественника ацетил-СоА. Генная сеть регуляции внутриклеточного уровня холестерина включает несколько регуляторных контуров, функционирующих при участии транскрипционных факторов семейства SREBP (Sterol Regulatory Element-Binding Proteins). У человека белки семейства SREBP (SREBP1a, SREBP1c, SREBP2) кодируются двумя генами, *SREBF1* и *SREBF2*. Активность транскрипционных факторов семейства SREBP регулируется в обратной зависимости от уровня холестерина в клетке: чем выше уровень холестерина, тем ниже активность факторов SREBP.

У человека активность гена *SREBF2* контролируется регуляторным **контуром с положительной обратной связью** (рис. 10, а): транскрипция гена *SREBF2* активируется фактором SREBP2 (авторегуляция) (Fernández-Hernando, Moore, 2011).

Генная сеть включает также **контур** с отрицательной обратной связью, регулирующие уровень холестерина.

Центральными элементами первых двух контуров с отрицательной обратной связью являются транскрипционные факторы семейства

SREBP (рис. 10, б), регулирующие активность целой каскады генов, включая гены ферментов мевалонатного пути биосинтеза холестерина (*HMGCoAS*, *HMGCoAR*, *FDFS*, *SS*, *DHCR7* и др.), а также рецептора липопротеинов низкой плотности (*LDLR*). В пределах **первого контура** по механизму обратной связи регулируется активность генов мевалонатного пути биосинтеза холестерина: чем выше концентрация холестерина, тем ниже активность транскрипционных факторов SREBP и тем ниже уровень транскрипции ферментов *HMGCoAS*, *HMGCoAR*, *FDFS*, *SS*, *DHCR7*. В пределах **второго контура** по аналогичному механизму обратной связи регулируется транскрипция гена *LDLR*, кодирующего рецептор LDL, обеспечивающий транспорт липопротеинов низкой плотности в клетку, вследствие чего интенсивность поступления холестерина в цитоплазму снижается (Игнатьева и др., 1997; Меркулова и др., 2013).

Третий контур с отрицательной обратной связью, регулирующий уровень холестерина в клетке, был выявлен недавно, когда было обнаружено, что район 16-го интрона гена *SREBF2* кодирует микроРНК (Fernández-Hernando, Moore, 2011). Эта микроРНК (miR-33a) экспрессируется в составе первичного транскрипта гена *SREBF2* и подавляет трансляцию мРНК, кодирующей белок ABCA1. Белок ABCA1 является мембранным белком, осуществляющим перенос холестерина из цитоплазмы во внеклеточное пространство, где этот метаболит захватывается частицами липопротеинов высокой плотности. Регуляция уровня холестерина в пределах третьего регуляторного контура с отрицательной обратной связью осуществляется следующим образом. При низком уровне холестерина активность транскрипционных факторов семейства SREBP повышается. Далее белки SREBP активируют экспрессию гена *SREBF2*. МикроРНК (miR-33a), экспрессируемая в составе первичного транскрипта гена *SREBF2*, ингибирует трансляцию белка ABCA1, осуществляющего экспорт холестерина из клетки. Замедление экспорта холестерина приводит к нормализации внутриклеточного уровня холестерина.

Помимо контуров с обратной связью, функционирующих на уровне клетки, существуют дополнительные механизмы активации транс-

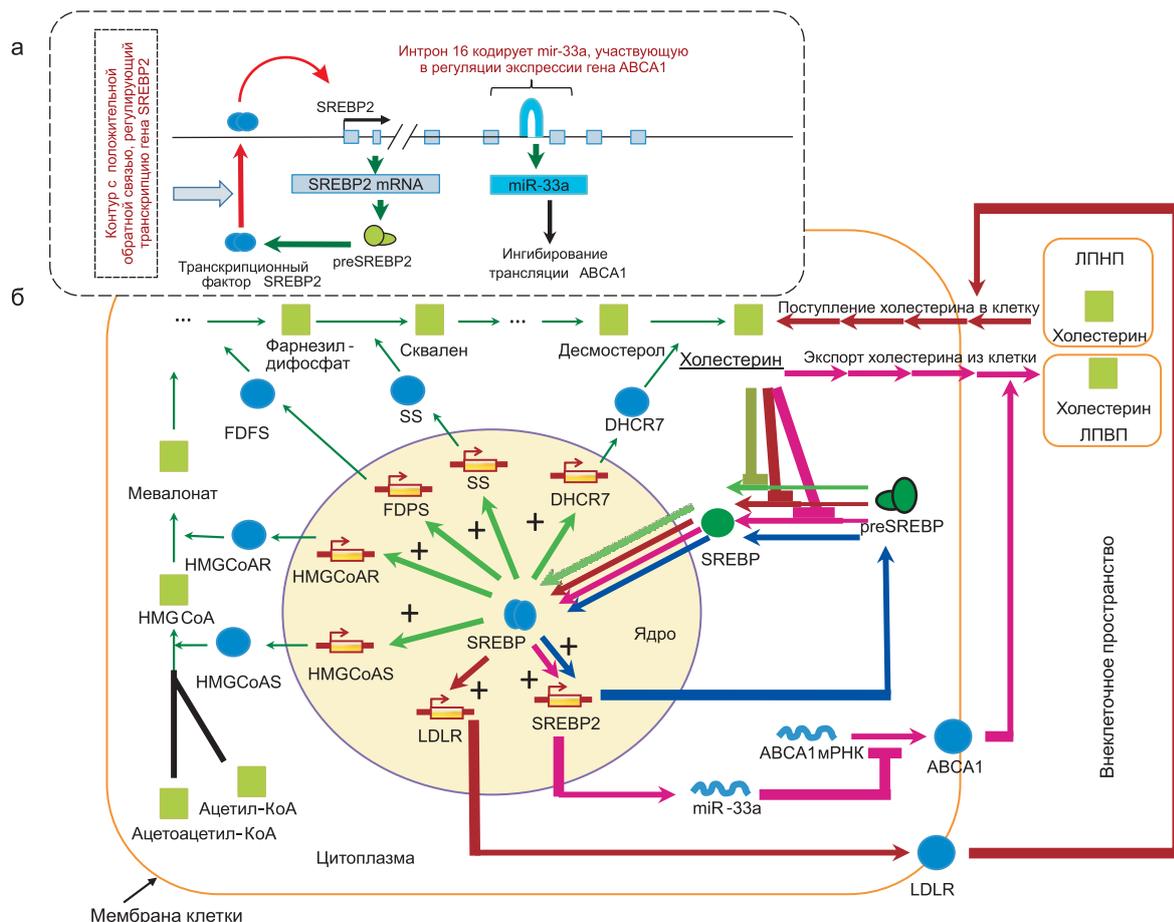


Рис. 10. Регуляторные контуры с обратной связью, функционирующие в генной сети, контролирующей уровень холестерина в клетках человека.

а – экспрессия микроРНК в составе первичного транскрипта гена *SREBF2* и контур с положительной обратной связью, регулирующий экспрессию *SREBF2*; б – четыре регуляторных контура (обозначены стрелками различных цветов), функционирующих с участием транскрипционных факторов семейства SREBP.

крипционных факторов семейства SREBP, обеспечивающие настройку активности белков SREBP в соответствии с потребностями многоклеточного организма (рис. 11).

Первый механизм активации факторов семейства SREBP – это протеолитический процессинг неактивного предшественника preSREBP в активную форму SREBP. Как уже упоминалось выше, этот процесс запускается при низком уровне холестерина за счет изменения конформации якорного белка INSIG, удерживающего комплекс белков SCAP/preSREBP на мембране эндоплазматического ретикулума. В результате связь комплекса белков SCAP/preSREBP с мембраной нарушается и preSREBP вместе с белком SCAP перемещается в аппарат Гольджи, где под действием протеаз S1P и S2P происходит

расщепление preSREBP и образуются активные факторы SREBP (Gimpl *et al.*, 2002).

Второй механизм активации факторов SREBP запускается под действием ростовых факторов и диеты, богатой аминокислотами и глюкозой. Как ростовые факторы, так и полноценная диета активируют киназу mTOR, которая фосфорилирует белок lipin1. Фосфорилированная форма белка lipin1 не может проникать в ядро, поэтому концентрация белка lipin1 в ядре снижается. В ядре клетки белок lipin1 подавляет активность факторов SREBP за счет того, что lipin1 способствует транспорту транскрипционных факторов SREBP из активных центров ядра к ядерной мембране, где SREBP инактивируется, связываясь с белком lamin A. Таким образом, второй механизм активации

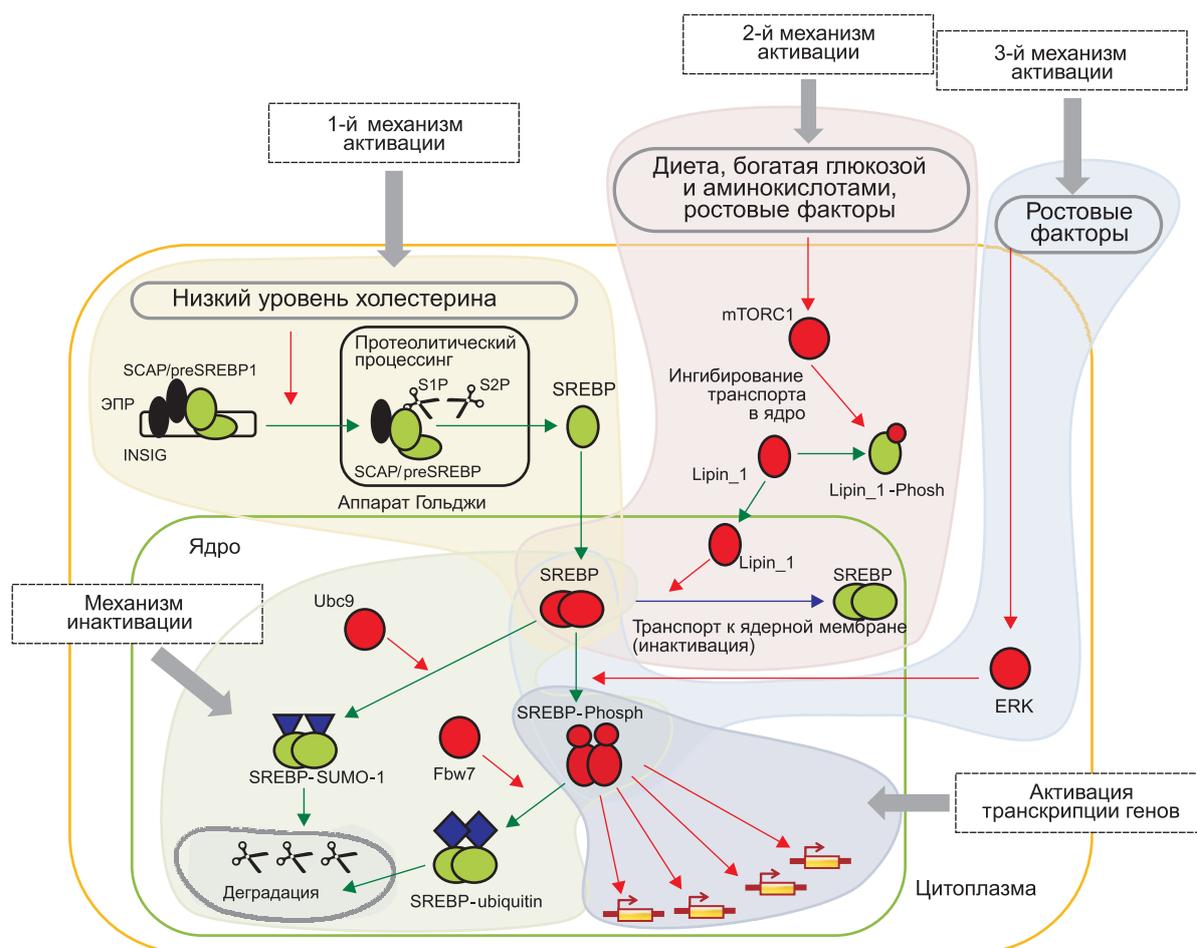


Рис. 11. Пути регуляции активности транскрипционных факторов семейства SREBP человека и животных внеклеточными и внутриклеточными сигналами.

факторов SREBP реализуется через ингибирование транспорта в ядро инактиваторного белка lipin 1 (Peterson *et al.*, 2011).

Третий механизм активации факторов SREBP опосредуется ERK-киназами. Этот механизм запускается под действием ростовых факторов (например insulin-like growth factor-1), активирующих ERK-киназы (ERK1 и ERK2). ERK-киназы осуществляют фосфорилирование факторов SREBP, что повышает их ДНК-связывающую и транскрипционную активность (Arito *et al.*, 2008).

Наряду с путями активации факторов SREBP в клетке постоянно осуществляются и **процессы инактивации**, запускаемые посттрансляционными модификациями. Известно, что сумоилирование может осуществляться под действием SUMO-1-конъюгирующего фермента Ubc9 (Arito *et al.*, 2008), а убиквитинирование

запускается убиквитин-лигазой Fbw7 (Sundqvist *et al.*, 2005).

Таким образом, итоговый уровень активности SREBP контролируется сложно организованной генной сетью, интегрирующей все описанные выше процессы и включающей многие десятки генов и белков. Очевидно, что данное описание генной сети, составленное на основании доступных к настоящему времени экспериментальных данных, еще далеко от полного представления о реальной сложности процесса регуляции уровня холестерина в клетках и тканях организма, в контроле которого могут быть задействованы продукты многих сотен генов.

Взаимодействие генных сетей

Нормальное функционирование живого организма предполагает координацию и син-

хронизацию всех процессов на всех уровнях его иерархии. Одним из ярких примеров такой координации является настройка активности генных сетей метаболизма генными сетями регуляции циркадного ритма. Генная сеть регуляции циркадного ритма управляет временной разверткой всех функций организма (молекулярных, клеточных, физиологических и др.), контролируя на уровне транскрипции значительную часть генома через транскрипционный фактор CLOCK/Bmal1. Следует отметить, что среди мишеней CLOCK/Bmal1 представлены различные функциональные группы генов. Так, например, в печени мыши наиболее обогащены сайтами связывания CLOCK/Bmal1 группы генов липидного и углеводного метаболизма, а также регуляторы транскрипции, в частности ядерные рецепторы (Rey *et al.*, 2011). Циркадные часы могут контролировать периферические системы как прямо (через CLOCK/Bmal1), так и опосредованно, через активируемые им регуляторы транскрипции. На рис. 12 в интегрированном виде представлены результаты метаболомного анализа печени мыши, демонстрирующие наличие циркадной ритмики метаболитов, участников путей обмена аминокислот, нуклеотидов, углеводов, липидов и ксенобиотиков. Эта ритмика является следствием согласованной работы транскриптома и протеома и определяет состояние метаболического гомеостаза (Eckel-Mahan *et al.*, 2012).

Можно отметить, что максимум метаболитов путей обмена ксенобиотиков и аминокислот приходится на вторую половину суток, в то время как нуклеотидов, углеводов и липидов – на первую.

Поскольку существует контроль метаболических процессов со стороны циркадных часов, естественно предположить, что функциональные

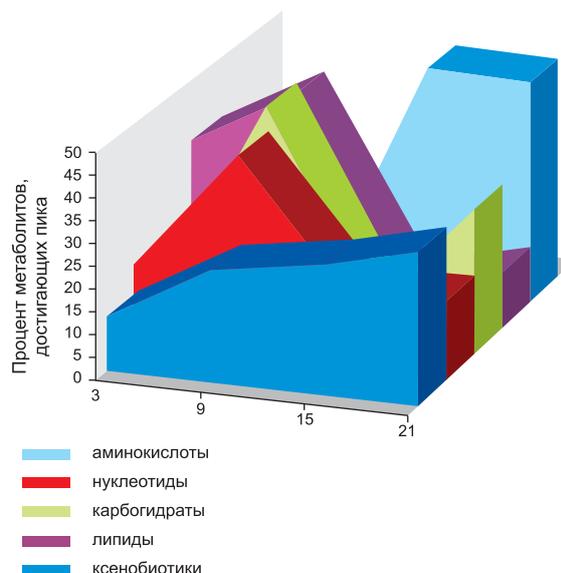


Рис. 12. Циркадная ритмика метаболитов, принадлежащих различным путям обмена веществ в печени мыши.

По оси X – время суток в часах (модифицировано по: (Eckel-Mahan *et al.*, 2012. P. 5541–5546)).

и структурные нарушения элементов циркадного осциллятора могут приводить к нарушениям метаболизма и развитию различных патологических состояний. Это предположение нашло частичное подтверждение в ряде эпидемиологических исследований. Некоторые примеры однонуклеотидных полиморфизмов (ОНП) гена *CLOCK* человека и ассоциированных с ними фенотипов или патологий приведены в табл. 1.

ЭВОЛЮЦИЯ ГЕННЫХ СЕТЕЙ

В ходе микроэволюции популяций человека отбор неоднократно приводил к возникновению комплексов адаптивных полиморфизмов,

Таблица 1

Ассоциации некоторых ОНП гена *CLOCK* человека с метаболическими нарушениями

ОНП	Патология или фенотип
rs1801260; T > C	Ожирение (Garaulet <i>et al.</i> , 2010) Устойчивость к воздействиям, направленным на потерю веса (Garaulet <i>et al.</i> , 2010) Предрасположенность к высокому индексу массы тела (Monteleone <i>et al.</i> , 2008) Изменение пищевого поведения со сдвигом на вечернее время (Garaulet <i>et al.</i> , 2011) Высокое потребление жиров (Garaulet <i>et al.</i> , 2011)
rs11932595; A > G	Стеатоз (Sookoian <i>et al.</i> , 2007)
rs6850524; C > G	Высокий индекс массы тела или ожирение (Sookoian <i>et al.</i> , 2008)

оптимизировавших функцию генных сетей под конкретные условия жизнеобитания. Оптимизация затрагивала как особенности экспрессии отдельных генов, так и другие аспекты организации и функции генных сетей. Такого рода адаптивные комплексы являлись относительно стабильными единицами наследственной памяти, устойчиво передавались в ряду поколений и обеспечивали формирование оптимальных для данной среды фенотипических характеристик организмов. Простейший вариант адаптивного комплекса – гаплотип.

Проведенный нами ранее анализ «мутационного портрета» генной сети, контролирующей гомеостаз холестерина в клетках животных, выявил интересную особенность генной сети, контролируемой контурами с отрицательной обратной связью (Ананько и др., 2008). Сеть гомеостатирует любые флуктуации: как изменения параметров, вызванные физиологическими воздействиями на клетку, так и их отклонения, вызванные мутациями генов. Исследование влияния мутаций в математических моделях показало, что мутационные повреждения многих генов либо практически не влияют на уровень холестерина в клетке, либо изменяют этот показатель в незначительной степени. Зато есть группа генов, повреждения которых вызывают значительные отклонения уровня холестерина (от 0 до более чем 200 % относительно нормы). Доля таких генов невелика (около 15 %), однако все они выполняют очень важные регуляторные функции.

Сходная картина наблюдается и в популяциях человека: как показывает рассмотренный далее пример, наиболее существенными оказываются последствия в регуляторных звеньях генных сетей, контролирующих те или иные фенотипические признаки.

Комплексы полиморфизмов, адаптивные для популяций человека в историческом прошлом, в настоящее время могут быть источником системных заболеваний. Выявлено, что японские популяции отличаются от европейских по частотам полиморфизмов в 8 генах, регулирующих потребление, запасание и расход энергии (табл. 2). Предполагается, что этот комплекс аллельных вариантов генов фиксировался в японских популяциях в результате отбора голодом, например, за период между 501 и 1947 годами японцы пе-

Гаплотип – устойчивый набор полиморфных вариантов (аллелей) в пределах одного или нескольких генетических локусов, расположенных на одной хромосоме.

режили 226 периодов сильнейшего неурожая и голода (Kagawa *et al.*, 2002). Однако в настоящее время, когда потребление основных продуктов (жиров, углеводов и мяса) возросло в Японии в 10–20 раз (по сравнению с 1950 г.), этот комплекс полиморфизмов создает предпосылки к нарушению метаболизма углеводов и липидов и возникновению сопутствующих этому патологий. В частности, заболеваемость диабетом второго типа у японцев превышает таковую у жителей США в 1,5 раза.

Примечательно, что половина из 8 генов (*AGT*, *PPAR γ* , *ADRB3* и *LEPR*) кодирует белки с регуляторными функциями и большинство генов (*AGT*, *PPAR γ* , *ADRB3*, *UCP2*, *CAPN10*) экспрессируется в адипоцитах, клетках, в которых и происходит запасание энергии в виде триглицеридов.

Ген *AGT*, кодирующий ангиотензиноген, экспрессируется как в печени, так и в адипоцитах и является предшественником регуляторного белка ангиотензина 2. Примечательно, что в адипоцитах человека и животных экспрессируются и ренин – фермент, под действием которого из ангиотензиногена образуется ангиотензин 1 и ангиотензин-превращающий белок (АПФ), который расщепляет ангиотензин 1 с образованием ангиотензина 2. Ангиотензин 2 воздействует на мышечные клетки кровеносных сосудов, что приводит к повышению кровяного давления (Goossens *et al.*, 2007). У индивидов с нормальной массой тела основная масса ангиотензиногена образуется в печени, ренин синтезируется в почках, а АПФ в основном нарабатывается в легких. В результате секреции этих белковых продуктов в кровь нарабатывается регулятор кровяного давления – ангиотензин 2. Однако, поскольку при накоплении избыточной массы тела возникает проблема сердечно-сосудистой системы, необходимо было фиксировать дополнительные механизмы, которые позволяли бы более эффективно и динамично осуществлять подрегуляцию артериального давления и сер-

Таблица 2

Гены энергетического метаболизма, для которых обнаружены различные частоты полиморфных аллелей у монголоидов и европеоидов (Kagawa *et al.*, 2002)

Ген	Продукт гена	Функция	Орган (ткань), где преимущественно наблюдается экспрессия
<i>AGT</i>	Ангиотензиноген, предшественник ангиотензина 2	Регуляция кровяного давления	Печень, адипоциты
<i>PPARγ</i>	Транскрипционный фактор PPAR γ	Регуляция дифференцировки адипоцитов и поддержание их зрелого фенотипа	Адипоцит
<i>ADRB3</i>	Бета 3 адренергический рецептор	Регуляция липолиза и термогенеза	Адипоциты
<i>LEPR</i>	Рецептор гормона лептина	Регуляция пищевого поведения	Гипоталамус
<i>UCP2</i>	Белки митохондрий UCP2 и UCP3, разобщающие окислительное фосфорилирование и синтез АТФ	Активация теплопродукции	Адипоцит
<i>UCP3</i>			Мышцы
<i>CAPN10</i>	Кальций-зависимая протеаза	Разрушение внутриклеточных белков	Экспрессируется повсеместно, включая адипоцит
<i>FABP2</i>	Транспортный белок	Внутриклеточный транспорт жирных кислот	Кишечник

дечно-сосудистой системы. И действительно, у японцев частота аллеля 235T (*AGT/M235T* полиморфизм), ассоциированного с повышенным давлением, оказалась в два раза выше, чем у европеоидов (Katsuya *et al.*, 2003).

Ген PPAR γ кодирует транскрипционный фактор, который регулирует транскрипцию большого количества генов, экспрессирующихся в адипоците, и обеспечивает дифференцировку адипоцитов и поддержание их зрелого фенотипа (Кузнецова и др., 2008).

Третий ген с регуляторными функциями – *ADRB3* (adrenoceptor beta 3) – экспрессируется преимущественно в жировых клетках и кодирует катехоламинный рецептор (бета 3 адренергический рецептор). Рецептор *ADRB3* воспринимает сигналы, идущие от нервных окончаний, и вовлечен в регуляцию липолиза и термогенеза.

Четвертый ген с регуляторными функциями из числа генов, в которых фиксировались адаптивные замены, кодирует **рецептор лептина (*LEPR*)**. Рецептор лептина экспрессируется в нейронах аркуатных ядер гипоталамуса, отвечающих за восприятие чувства голода либо насыщения. *LEPR* взаимодействует с гормоном лептином, который нарабатывается жировыми

клетками, в результате чего пищевое поведение подавляется (Колчанов и др., 2005).

Рассмотренные примеры показывают, что когда возникла жесткая необходимость приспособиться к суровым условиям среды, большинство адаптивных вариантов аллелей зафиксировались не в генах метаболизма, а в небольшом количестве регуляторных звеньев (иерархически высоких точек), которые обеспечили настройку на оптимальный уровень функционирования генной сети.

БЛАГОДАРНОСТИ

Работа выполнена при частичной поддержке Совета по грантам Президента Российской Федерации (НШ-5278.2012.4), Президиумов СО РАН (проект VI.61.1.2. из направления VI.61) и РАН (программа «Молекулярная и клеточная биология» – Интеграция РАН 6.6), Междисциплинарного интеграционного проекта № 136 программы Президиума СО РАН, РФФИ (проект № 11-04-01748-а), а также проекта № 14.В25.31.0033 от 28.07.2013 в рамках постановления Правительства России N 220 от 09.04.2010.

ЛИТЕРАТУРА

- Ананько Е.А., Гунбин К.В., Суслов В.В. и др. Генные сети: описание в базах данных и моделирование // Системная Компьютерная биология / Ред. Н.А. Колчанов, С.С. Гончаров, В.А. Лихошвай, В.А. Иванисенко. Новосибирск: Изд-во СО РАН, 2008. С. 313–393.
- Бусыгина Т.В., Игнатъева Е.В., Осадчук А.В. Регуляция транскрипции генов, контролирующей биосинтез стероидных гормонов: описание в базе данных ES-TRRD // Усп. соврем. биологии. 2003. Т. 123. С. 364–382.
- Голубятников В.П., Голубятников И.В., Лихошвай В.А. О существовании и устойчивости циклов в пятимерных моделях генных сетей // Сиб. журн. вычисл. математики. 2010. Т. 13. № 4. С. 403–412.
- Демиденко Г.В., Колчанов Н.А., Лихошвай В.А. и др. Математическое моделирование регуляторных контуров генных сетей // Журн. вычисл. матем. и матем. физики. 2004. Т. 44. С. 2276–2295.
- Демиденко Г.В., Лихошвай В.А., Мудров А.В. О связи между решениями дифференциальных уравнений с запаздывающим аргументом и бесконечномерных систем дифференциальных уравнений // Дифференциальные уравнения. 2009. Т. 45. С. 34–46.
- Игнатъева Е.В., Меркулова Т.И., Вишневский О.В., Кель А.Э. Регуляция транскрипции генов липидного метаболизма: описание в базе данных TRRD // Молекуляр. биология. 1997. Т. 31. С. 684–700.
- Колчанов Н.А., Лихошвай В.А., Матушкин Ю.Г. Регуляторные контуры генетических систем: принципы организации и эволюции // Информ. вестн. ВОГиС. 2001. № 16. С. 5–10.
- Колчанов Н.А., Ананько Е.А., Колпаков Ф.А. и др. Генные сети // Молекуляр. биология. 2000. Т. 34. С. 533–544.
- Колчанов Н.А., Подколдная О.А., Игнатъева Е.В. и др. Интеграция генных сетей, контролирующей физиологические функции организма // Информ. вестн. ВОГиС. 2005. Т. 9. № 2. С. 179–198.
- Кузнецова Т.Н., Игнатъева Е.В., Мордвинов В.А. и др. Анализ структуры инсулин-зависимых регуляторных контуров зрелых адипоцитов // Усп. физиол. наук. 2008. Т. 39. № 1. С. 3–22.
- Лихошвай В.А., Матушкин Ю.Г., Фадеев С.И. Задачи теории функционирования генных сетей // Сиб. журн. индустриальной математики. 2003. Т. 6. С. 64–80.
- Лихошвай В.А., Матушкин Ю.Г., Фадеев С.И. О связи графа генной сети с качественными режимами ее функционирования // Молекуляр. биология. 2001. Т. 35. С. 1080–1088.
- Лихошвай В.А., Фадеев С.И. О сдвиге регуляторного сигнала в моделях матричного синтеза // Сиб. журн. индустриальной математики (СИБЖИМ). 2013. Т. XVI. № 1(53). С. 66–74.
- Лихошвай В.А., Фадеев С.И., Демиденко Г.В., Матушкин Ю.Г. Моделирование многостадийного синтеза вещества без ветвления уравнением с запаздывающим аргументом // Сиб. журн. индустриальной математики. 2004. Т. 7. № 1. С. 73–94.
- Меркулова Т.И., Ананько Е.А., Игнатъева Е.В., Колчанов Н.А. Регуляторные коды транскрипции геномов эукариот // Генетика. 2013. Т. 49. № 1. С. 37–54.
- Фадеев С.И., Гайнова И.А., Березин А.Ю. и др. Исследование стационарных решений в моделях генных сетей методом гомотопии // Сиб. электрон. матем. известия (SEMR). 2004. Т. 1. С. 64–75.
- Чураев Р.Н. Эпигены – наследственные единицы надгенного уровня // Экол. генетика. 2010. Т. 8. С. 17–24.
- Arito M., Horiba T., Hachimura S. *et al.* Growth factor-induced phosphorylation of sterol regulatory element-binding proteins inhibits sumoylation, thereby stimulating the expression of their target genes, low density lipoprotein uptake, and lipid synthesis // *J. Biol. Chem.* 2008. V. 283. No. 22. P. 15224–15231.
- Caspi R., Altman T., Dreher K. *et al.* The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases // *Nucl. Acids Res.* 2012. V. 40. P. D742–D753.
- Cooper J.W., Kershenbaum A. Discovery of protein-protein interactions using a combination of linguistic, statistical and graphical information // *BMC Bioinformatics.* 2005. V. 6. P. 143.
- Croft D., O’Kelly G., Wu G. *et al.* Reactome: a database of reactions, pathways and biological processes // *Nucl. Acids Res.* 2011. V. 39. P. D691–D697.
- Demenev P.S., Ivanisenko T.V., Kolchanov N.A., Ivanisenko V.A. ANDVisio: A new tool for graphic T. V. visualization and analysis of literature mined associative gene networks in the ANDSystem // *In Silico Biol.* V. 11. P. 149–161.
- Demidenko G.V., Likhoshvai V.A., Melnik I.A. On properties of solutions to equations of multistage substance synthesis // *J. Anal. Appl.* 2010. V. 8. P. 47–61.
- Demir E., Cary M.P., Paley S. *et al.* The BioPAX community standard for pathway data sharing // *Nat. Biotechnol.* 2010. V. 28. P. 935–942.
- Eckel-Mahan K.L., Patel V.R., Mohney R.P. *et al.* Coordination of the transcriptome and metabolome by the circadian clock // *Proc. Natl Acad. Sci. USA.* 2012. V. 109. P. 5541–5546.
- Fernández-Hernando C., Moore K.J. MicroRNA modulation of cholesterol homeostasis // *Arterioscler. Thromb. Vasc. Biol.* 2011. V. 31. No. 11. P. 2378–2382.
- Garulet M., Corbalán M.D., Madrid J.A. *et al.* CLOCK gene is implicated in weight reduction in obese patients participating in a dietary programme based on the Mediterranean diet // *Int. J. Obes. (Lond).* 2010. V. 34. P. 516–523.
- Garulet M., Sánchez-Moreno C., Smith C.E. *et al.* Ghrelin, sleep reduction and evening preference: relationships to CLOCK 3111 T/C SNP and weight loss // *PLoS One.* 2011. V. 6. P. e17435.
- Gene Ontology Consortium. The Gene Ontology (GO) project in 2006 // *Nucl. Acids Res.* 2006. V. 34. P. D322–D326.
- Gimpl G., Burger K., Fahrenholz F. A closer look at the cholesterol sensor // *Trends Biochem. Sci.* 2002. V. 27. P. 596–599.
- Goossens G.H., Jocken J.W., Blaak E.E. *et al.* Endocrine role of the renin-angiotensin system in human adipose tissue and muscle: effect of beta-adrenergic stimulation // *Hypertension.* 2007. V. 49. No. 3. P. 542–547.
- Haibe-Kains B., Olsen C., Djebbari A. *et al.* Predictive

- networks: a flexible, open source, web application for integration and analysis of human gene networks // *Nucl. Acids Res.* 2012. V. 40. P. D866–875.
- Harmston N., Filsell W., Stumpf M.P. What the papers say: Text mining for genomics and systems biology // *Hum. Genomics.* 2010. V. 5. P. 17–29.
- Kagawa Y., Yanagisawa Y., Hasegawa K. *et al.* Single nucleotide polymorphisms of thrifty genes for energy metabolism: evolutionary origins and prospects for intervention to prevent obesity-related diseases // *Biochem. Biophys. Res. Commun.* 2002. V. 295. No. 2. P. 207–222.
- Kanehisa M., Goto S., Hattori M. *et al.* From genomics to chemical genomics: new developments in KEGG // *Nucl. Acids Res.* 2006. V. 34. P. D354–D357.
- Katsuya T., Ishikawa K., Sugimoto K. *et al.* Salt sensitivity of Japanese from the viewpoint of gene polymorphism // *Hypertens. Res.* 2003. V. 26. No. 7. P. 521–525.
- Kelder T., van Iersel M.P., Hanspers K. *et al.* WikiPathways: building research communities on biological pathways // *Nucl. Acids Res.* 2012. V. 40. P. D1301–D1307.
- Meléndez-Hevia E., Sicilia J., Ramos M.T. *et al.* Molecular bureaucracy: who controls the delays? Transient times in branched pathways and their control // *Theor. Biol.* 1996. 182. No. 3. P. 333–339.
- Mi H., Thomas P. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools // *Methods Mol. Biol.* 2009. V. 563. P. 123–140.
- Monteleone P., Tortorella A., Docimo L. *et al.* Investigation of 3111T/C polymorphism of the CLOCK gene in obese individuals with or without binge eating disorder: association with higher body mass index // *Neurosci. Lett.* 2008. V. 435. P. 30–33.
- Mungall C.J., Emmert D.B. FlyBase Consortium. A Chado case study: an ontology-based modular schema for representing genome-associated biological information // *Bioinformatics.* 2007. V. 23. P. i337–i346.
- Peterson T.R., Sengupta S.S., Harris T.E. *et al.* mTOR complex 1 regulates lipin 1 localization to control the SREBP pathway // *Cell.* 2011. V. 146. No. 3. P. 408–420.
- Rey G., Cesbron F., Rougemont J. *et al.* Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver // *PLoS Biol.* 2011. V. 9. P. e1000595.
- Rzhetsky A., Iossifov I., Koike T. *et al.* GeneWays: a system for extracting, analyzing, visualizing, and integrating molecular pathway data // *J. Biomed. Inform.* 2004. V. 37. P. 43–53.
- Sookoian S., Castaco G., Gemma C. *et al.* Common genetic variations in CLOCK transcription factor are associated with nonalcoholic fatty liver disease // *World J. Gastroenterol.* 2007. V. 13. P. 4242–4248.
- Sookoian S., Gemma C., Gianotti T.F. *et al.* Genetic variants of Clock transcription factor are associated with individual susceptibility to obesity // *Am. J. Clin. Nutr.* 2008. V. 87. P. 1606–1615.
- Sundqvist A., Bengoechea-Alonso M.T., Ye X. *et al.* Control of lipid metabolism by phosphorylation-dependent degradation of the SREBP family of transcription factors by SCF(Fbw7) // *Cell. Metab.* 2005. V. 1. No. 6. P. 379–391.
- Taylor R., Singhal M. Biological Network Inference and analysis using SEBINI and CABIN // *Methods Mol. Biol.* 2009. V. 541. P. 551–576.
- Wang M., Verdier J., Benedito V.A. *et al.* LegumeGRN: a gene regulatory network prediction server for functional and comparative studies // *PLoS ONE.* 2013. V. 8. P. e67434.
- Zambon A.C., Gaj S., Ho I. *et al.* GO-Elite: a flexible solution for pathway and ontology over-representation // *Bioinformatics.* 2012. V. 28. P. 2209–2210.