

УДК 61:575; 658.011.56

ЦИРКАДНЫЕ ЧАСЫ МЛЕКОПИТАЮЩИХ: ГЕННАЯ СЕТЬ И КОМПЬЮТЕРНЫЙ АНАЛИЗ

© 2014 г. О.А. Подколотная¹, Н.Н. Подколотная^{1,2}, Н.Л. Подколотный^{1,3}

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

² Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия;

³ Федеральное государственное бюджетное учреждение науки Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук, Новосибирск, Россия

Поступила в редакцию 25 сентября 2014 г. Принята к публикации 21 октября 2014 г.

В работе представлены результаты реконструкции и анализа генной сети циркадных часов млекопитающих. Применение методов теории графов позволило провести анализ структуры генной сети и выделить центральную компоненту регуляции циркадного ритма, которая включает базовые регуляторные контуры, проходящие через ключевой элемент циркадных часов – белок Clock/Bmal1. Использование кластерного анализа позволило выявить подсистемы, имеющие четкую биологическую интерпретацию и участвующие в функционировании циркадных часов путем взаимодействия с центральной компонентой. Такая структурная модель, включающая центральную компоненту и взаимодействующие с ней функциональные подсистемы, может быть основой для построения математической модели динамики генной сети регуляции циркадного ритма.

Ключевые слова: циркадные часы, генные сети, методы анализа графов.

ВВЕДЕНИЕ

Циркадные часы являются универсальным адаптивным механизмом эукариот. Они обеспечивают согласованное протекание процессов в живых организмах на всех уровнях от молекулярно-генетических до поведенческих. Общепринято мнение, что основой функционирования этого механизма служат молекулярно-генетические осцилляторы, присутствующие практически в каждой клетке организма (Albrecht, 2012). У млекопитающих главный водитель циркадного ритма, сформированный нейронами супрахиазматических ядер гипоталамуса (СХЯ), синхронизует ритм сети периферических циркадных осцилляторов, расположенных как в мозге (вне СХЯ), так и в периферических тканях (Там же). Принципиальная структура клеточных циркадных осцилляторов одинакова в нейронах СХЯ и клетках перифе-

рических тканей (Yagita *et al.*, 2001). Универсальным регуляторным модулем циркадных часов млекопитающих является отрицательная обратная связь, обеспечивающая ритмическую регуляцию транскрипции генов, кодирующих белки *Per1-3* и *Cry1-2* (Reppert, Weaver, 2001). Коротко она может быть описана следующим образом: гетеродимерный транскрипционный фактор (ТФ) Clock/Bmal1 активирует транскрипцию генов *Per* и *Cry*, а белковый гетеродимерный комплекс *Per/Cry* за счет белок-белковых взаимодействий подавляет активность ТФ Clock/Bmal1. Это, в свою очередь, приводит к замыканию регуляторного контура за счет снижения уровня транскрипции генов *Per* и *Cry* и соответствующему восстановлению активности Clock/Bmal1. В ходе этого цикла ритмически изменяется уровень гетеродимеров *Per/Cry* и активность ТФ Clock/Bmal1. Кроме данной основной петли в циркадных часах млекопи-

тающих описаны и другие транскрипционно-трансляционные петли как с отрицательными, так и с положительными обратными связями, которые обеспечивают стабильность функционирования циркадных часов (Albrecht, 2012).

Помимо этого, функционирование циркадных часов регулируется различными механизмами, включая дополнительные обратные связи с регуляторами транскрипции, пост-трансляционную модификацию белков (циклы фосфорилирования/дефосфорилирования и ацетилирования/деацетилирования, полиADP-рибозилирование, протеосомная деградация белков), стабилизацию, деградацию, транспорт RNA и белков и др. (Virshup *et al.*, 2007; Asher, Schible, 2011; Albrecht, 2012; Morf *et al.*, 2012; Подколотная, 2014). Можно отметить характерную особенность молекулярно-генетического механизма циркадных часов – наличие обратных связей их компонентов различных уровней с центральными компонентами часов Clock/Bmal1 и Per(1-3)/Cry(1-2) и между собой и избыточность компонент часового механизма.

Оба фактора существенно затрудняют исследование механизмов и построение более детальных математических моделей, учитывающих тканеспецифические особенности регуляции циркадных часов (Ripperger, Brown, 2010). Первым и одним из важнейших этапов построения моделей молекулярно-генетических систем является реконструкция и структурный анализ геновой сети, описывающей основные молекулярно-генетические события и их взаимосвязи в рассматриваемой системе. Цель работы – реконструкция геновой сети циркадного осциллятора млекопитающих и анализ ее структуры с помощью методов теории графов.

МАТЕРИАЛЫ И МЕТОДЫ

Методы реконструкции геновых сетей

На первом этапе работы геновая сеть Circadian Rhythm была реконструирована с использованием системы GeneNet (Ananko *et al.*, 2005). Система позволяет накапливать и систематизировать данные о генах, РНК, белках, малых молекулах, реакциях и регуляторных событиях, экстрагированных из научных публикаций и баз данных, используя специализированный редак-

тор геновых сетей. Экстрагированная на основе ручной аннотации информация представляется в текстовом и графическом виде. Для дальнейшей работы по реконструкции геновой сети на основе первичных знаний, накопленных в базе данных GeneNet, о молекулярно-генетических взаимодействиях при регуляции циркадного ритма млекопитающих разработаны специальные программные средства, позволяющие:

1. Экстрагировать информацию из XML-представления геновой сети в системе GeneNet в стандартные форматы gml, graphml и sbml.
2. Осуществлять реконструкцию геновой сети для определенного вида организма на основе использования информации по близким видам организмов из базы данных GeneNet.
3. Объединять различные геновые сети, представленные в базе данных GeneNet.
4. Осуществлять проверку целостности геновой сети, восстановление недостающей информации из других баз данных и удаление дублирующей информации, включая описание одних и тех же молекулярных событий с различной степенью детальности.
5. Осуществлять проверку связности графа геновой сети и выделять связанные компоненты графа.

Методы анализа графов геновых сетей

Для анализа структуры геновой сети использованы библиотека программ Networkx 1.9 (<http://networkx.github.io/>), системы Cytoscape v.3.1.1 (<http://www.cytoscape.org/>) и Gephi v.0.8.2-beta (<http://gephi.github.io/>), а также собственные программы, написанные на языке Python 2.7, в рамках которых осуществлялись загрузка и преобразование графа геновой сети, подключение внешних библиотек, вызов методов анализа и интеграция результатов обработки.

Для выявления центральных или наиболее важных вершин в геновой сети применяли методы расчета различных показателей центральности вершин (индексы структурной важности вершин графа) (Koschützki, Schreiber, 2008), реализованные в библиотеке программ Networkx 1.9, включая:

- центральность по степени вершин (degree centrality) – наиболее распространенный и простой показатель, который соответствует

степени вершины, т. е. числу дуг, связанных с вершиной, нормированной на число вершин графа. Для ориентированного графа генной сети вводятся отдельные показатели для входных и выходных степеней вершины. Вершины с максимальным количеством дуг, или хабы, обычно соответствуют белкам – регуляторным молекулам;

- центральность вершины графа по близости (*closeness centrality*) – обратно пропорциональна сумме кратчайших расстояний от этой вершины до других вершин в графе. Этот показатель определяет важность конкретной вершины графа для быстрой передачи информации в графе;
- центральность вершины по посредничеству (*betweenness centrality*) – взвешенная сумма всех кратчайших путей между вершинами в графе, проходящих через данную вершину. Этот показатель определяет важность вершины графа с точки зрения ее влияния на пути передачи информации между вершинами графа;
- эксцентриситет вершин – максимальное расстояние от вершины до других вершин графа. Вершины с минимальным эксцентриситетом находятся в центре графа.

Важным методом анализа ориентированных графов является поиск сильносвязных компонент и регуляторных контуров в генной сети, который осуществлен нами с использованием алгоритма R. Tarjan (1973), реализованного в пакете программ Networkx 1.9. Сильносвязной компонентой графа называется подграф, в котором существует путь между любыми вершинами подграфа. Таким образом, вершины любого регуляторного контура в графе генных сетей входят в максимальную сильносвязную компоненту этого графа.

Кластеризация вершин и выявление структурных модулей генной сети регуляции циркадного ритма выполнялись с помощью итерационного метода Louvain, который обеспечивает оптимальное разбиение графа на структурные модули, или кластеры (Blondel *et al.*, 2008). Для генерации случайных графов применены методы, реализованные в библиотеке программ Networkx 1.9. Для визуализации графа генных сетей использована система Gephi и Cytoscape.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Реконструкция генной сети Circadian rhythm

Генная сеть циркадного осциллятора реконструирована на основе анализа и систематизации данных, представленных в 279 научных публикациях, а также информации из баз данных SWISS-PROT и Entrez Gene. Вносились данные, полученные при исследовании клеточных линий или тканей и органов мыши, крысы или человека. Реконструированный фрагмент генной сети регуляции циркадного ритма включает 54 гена, 122 белка, 16 небелковых субстанций, 436 реакций и регуляторных событий. Для анализа генной сети построен ее граф (рис. 1).

В табл. 1 представлены основные характеристики реконструированной генной сети Circadian rhythm. Размер графа определяется числом вершин (676) и дуг (936). Диаметр графа (19) соответствует максимальному расстоянию между его вершинами. Средняя степень вершин графа зависит от представления графа генной сети и детальности описания механизмов регуляции. В систему GeneNet в отличие от других вариантов описания генных сетей (KEGG, PathwayDB, REACTOME и другие базы данных) включено описание регуляторных событий, которые имеют меньшее значение степени вершин. Поэтому наблюдаемое значение средней степени вершин графа (2,74) нельзя сразу сравнивать с показателями генных сетей, полученных из других источников. Для сравнения необходимо предварительно генные сети привести к единому представлению.

Поэтому для выявления особенностей графа генной сети Circadian rhythm проведено сравнение его характеристик с характеристиками моделей случайных графов Эрдеша – Реньи (Erdős – Rényi model) и малых миров (Watts – Strogatz model), которые часто используют для структурного моделирования биологических сетей (Erdős, Rényi, 1959; Watts, Strogatz, 1998; Newman, 2003; Barabasi, Oltvai, 2004). На рис. 2 представлено распределение степеней вершин графа генной сети циркадного ритма, а также моделей случайных графов: модель Реньи с тем же количеством узлов и дуг и модель малых миров со средней степенью вершин, как в генной сети. Средняя степень вершин у всех графов



Рис. 1. Граф геновой сети Circadian rhythm. Размер вершин графа пропорционален степени вершины. Наибольшие размеры вершин соответствуют белкам Clock/Bmal1, per1/cry1, per2/cry2, Sirt1.

Таблица 1

Характеристика графа геновой сети Circadian rhythm

Показатель	Значение
Кол-во вершин	676
Кол-во дуг	936
Средняя степень вершин	2,74
Диаметр графа	19
Коэффициент ассоциативности вершин по степени	-0,156
Средний коэффициент кластеризации	0,0072

практически совпадает. Можно видеть, что распределение степеней вершин в графе геновой сети Circadian rhythm отличается от моделей случайных графов. В графе геновой сети присутствуют специфичные для нее неоднородности, обусловленные различными типами вершин: генов, РНК, белков, реакций и регуляторных событий. В частности, граф геновой сети имеет

вершины со степенью более 10, вероятность появления которых в моделях случайных графов очень мала ($p < 0,0005$ для модели случайных графов Реньи и $p < 10^{-6}$ для модели случайных графов малых миров) (рис. 2). Максимальные степени вершин характерны для четырех компонент графа, составляющих основу циркадного осциллятора – транскрипционного фактора

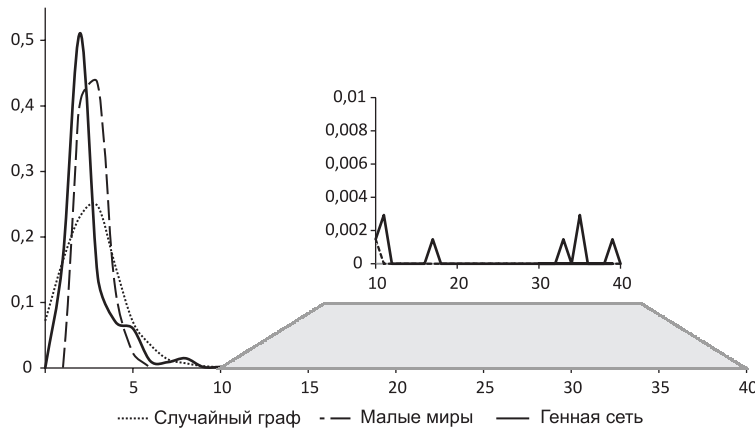


Рис. 2. Распределение степеней вершин графа генной сети *Circadian rhythm* и моделей случайных графов (ось абсцисс – степени вершин, ось ординат – частота распределения числа вершин графа с данной степенью). Сплошная линия – для генной сети, пунктирная – для случайного графа с тем же количеством вершин и дуг (модель случайных графов Реньи), штриховая линия – для модели случайных графов малых миров со средней степенью вершин, как в генной сети.

Clock/Bmal1 (значение степени вершины 39), регуляторов его активности гетеродимерных комплексов *Per1/Cry1* (35) и *Per2/Cry2* (35) и NAD-зависимой деацетилазы *Sirt1* (33). Большое количество связей также у активной формы коактиватора транскрипции *Pgc1a* (17), ТФ *Rev-erb alpha* (11), генов *Arntl* (11) и *Nr1d1* (10). В табл. 2 представлены характеристики некоторых вершин графа генной сети *Circadian rhythm*, соответствующих регуляторным белкам и белковым комплексам.

Показателями, определяющими важность вершин графа генной сети для выполнения различного типа функций, являются степень вершины и различные типы центральности вершин. Следует отметить, что для вершин с максимальными значениями степени значения центральности различного типа также максимальны. Вершины, имеющие большое число дуг, как правило ТФ, связывают регуляторными путями других участников регуляции циркадных часов между собой, так как находятся на перекрестках путей.

Минимальный эксцентриситет (*eccentricity*) в реконструированной генной сети *Circadian rhythm* имеют вершины графа, соответствующие белкам *Clock/Bmal1*, *Err alpha*, *Dec1*, *Dec2*, *Pgc1a 177p538p*, *Per1/Cry1*, *Per2/Cry2* и *Sirt1*. Таким образом, по результатам анализа можно сделать выводы, что центральными вершинами

в генной сети *Circadian rhythm* являются белки *Clock/Bmal1*, *Per1/Cry1*, *Per2/Cry2*, *Sirt1* и *Pgc1a 177p538p*, которые играют важную роль в передаче регуляторных сигналов в генной сети циркадного осциллятора.

Кластерный анализ

На следующем этапе анализа выявляли структурные модули, или кластеры, в графе генной сети *Circadian rhythm*. Близкие друг другу вершины объединены в единый кластер. Каждая вершина может входить только в один кластер. В результате анализа выявлено 18 кластеров с числом вершин более трех, которые были упорядочены в соответствии с числом входящих в них вершин. В табл. 3 представлена биологическая интерпретация некоторых наиболее интересных примеров кластеров.

Самый большой кластер объединяет 28 вершин, формирующих пути превращения (образование комплексов, циклы фосфорилирования и дефосфорилирования) основных негативных регуляторов циркадного осциллятора – белков *Per(1-2)* и *Cry(1-2)* и киназ *Sk1e* и *Sk1d*, которые вносят основной вклад в определение величины его периода. Двадцать вершин кластера 3 образуют путь регуляции транскрипции гена *Per1*, ТФ *Creb*, *GR*, *Dec1* и *Dec2* и посттранскрипционную регуляцию экспрессии его

Таблица 2

Характеристика некоторых вершин в геновой сети Circadian rhythm

Вершина	Степень вершины	Центральность по			Эксцентриситет
		степени	близости	посредничеству	
Clock/Bmal1	39	0,058	0,239	0,452	11
Per1/Cry1	35	0,0519	0,20	0,094	11
Per2/Cry2	35	0,0519	0,20	0,090	11
Sirt1	33	0,049	0,18	0,186	12
Pgc1a 177p538p	17	0,025	0,18	0,11	11
Rev-erb alpha	11	0,016	0,16	0,06	14
<i>Arntl</i>	11	0,016	0,19	0,06	14
<i>Nr1d1</i>	10	0,015	0,19	0,06	14
<i>Per1</i>	8	0,012	0,20	0,075	13
Ror alpha	8	0,012	0,18	0,03	12
Prox1	8	0,012	0,17	0,01	12
<i>Per2</i>	8	0,012	0,20	0,083	12
Cyp7	8	0,012	0,1	0,03	14
Bhlhe40	8	0,012	0,18	0,04	14

Таблица 3

Примеры биологической интерпретации некоторых кластеров графа геновой сети Circadian rhythm

Номер кластера	Кол-во вершин	Биологическая интерпретация
1	28	Образование комплексов Per1/2 и Cry1/2 с киназами CK1e и CK1d, а также циклы фосфорилирования и дефосфорилирования белков Per и киназ
3	20	Экспрессии гена <i>Per1</i>
5	11	Регуляция уровня ТФ Clock/Bmal1 за счет образование комплексов белков Clock и Bmal1 с Dec1, Dec2 и Id2
7	11	Убиквитинирование Per1 и Per2
8	10	Экспрессия гена ТФ Dbp и регуляция этим ТФ транскрипции генов <i>Top1</i> и <i>Cyp3A4</i>
11	9	Экспрессия и регуляция активности белка Pgc1a (ген <i>Ppargc1a</i>)
12	9	Убиквитинирование Cry1 и Cry2
13	8	ТФ Clock/Bmal1 и компоненты, подавляющие его активность различными путями
15	7	Путь синтеза NAD ⁺ и регуляция им активности фермента Parp1
16	7	Образование TF Pparg gamma/Rxr alpha: регуляция транскрипции гена <i>Pparg</i> фактором Себра, который транскрибируется с гена <i>Cebpa</i>
17	7	Регуляция экспрессии гена <i>Ldlr</i> с участием ТФ Hes1 и Hes6
18	7	Убиквитинирование Cry1

RNA-связывающим белком Lark. Кластеры 18 (7 вершин), 12 (9 вершин) и 7 (11 вершин) представляют пути убиквитинирования белков *Per1*, *Per2*, *Cry1* *Cry2*. Кластер 13 (8 вершин) объединяет белки, подавляющие активность ТФ Clock/Bmal1, используя различные механизмы. Кластер 11 объединяет 9 вершин – компонент пути экспрессии гена *Ppargc1a*, кодирующего коактиватор транскрипции *Pgc1a* и регуляцию активности этого белка киназой AMPK, активность которого регулируется, в свою очередь, соотношением АТР/АМР.

Другим примером может служить кластер 9, компоненты которого представляют собой путь синтеза NAD^+ и регуляцию активности фермента *Parp1* соотношением NAD^+/NAM . Таким образом, можно констатировать, что проведенная нами кластеризация вершин графа генной сети Circadian rhythm позволила выявить биологически интерпретируемые группы объектов.

Мотивы

Как уже указано выше, особенность генной сети Circadian rhythm составляет большое количество обратных связей между ее элементами. Мы называем регуляторным мотивом подграф, созданный элементами генной сети, образующими регуляторный контур с обратной связью. В отличие от известного понятия «неслучайные структурные мотивы» в теории графов (Newman, 2003; Kim *et al.*, 2011), регуляторный мотив необязательно является часто повторяю-

щейся структурой в графе, однако важно, что он несет определенную функциональную нагрузку. Для выявления таких мотивов использованы методы поиска регуляторных контуров в графе генной сети. Рассмотрим некоторые из регуляторных мотивов данной генной сети (рис. 3).

Самый короткий регуляторный мотив – основная регуляторная петля циркадного осциллятора, включающая ТФ Clock/Bmal1 гены *Per* и *Cry* и гетеродимер Per/Cry. Здесь отрицательная обратная связь реализуется за счет белок-белковых взаимодействий гетеродимеров Clock/Bmal1 и Per/Cry (рис. 3, а).

Более длинные мотивы (рис. 3, б–ж) образованы петлями как с отрицательными (рис. 3, б, в), так и с положительными обратными связями (рис. 3, е, ж). Отрицательные обратные связи в мотивах б–г реализуются за счет подавления транскрипции гена *Arntl*, кодирующего субъединицу Bmal1 гетеродимера Clock/Bmal1, ТФ Rev-erb alpha, Err alpha, Klf10, в то время как в мотиве д за счет посттранскрипционной модуляции экспрессии Bmal1, осуществляемой miR-142-3p (гены *Rev-erba*, *Erra*, *Klf10* и *Mir142* являются мишенями ТФ Clock/Bmal1).

В мотивах е, ж положительная обратная связь осуществляется за счет активации транскрипции гена *Arntl* ТФ Ror alpha и Ppar alpha, гены которых, в свою очередь, также являются мишенями ТФ Clock/Bmal1. Как видим, центральный положительный регулятор данной генной сети ТФ Clock/Bmal1 участвует во всех

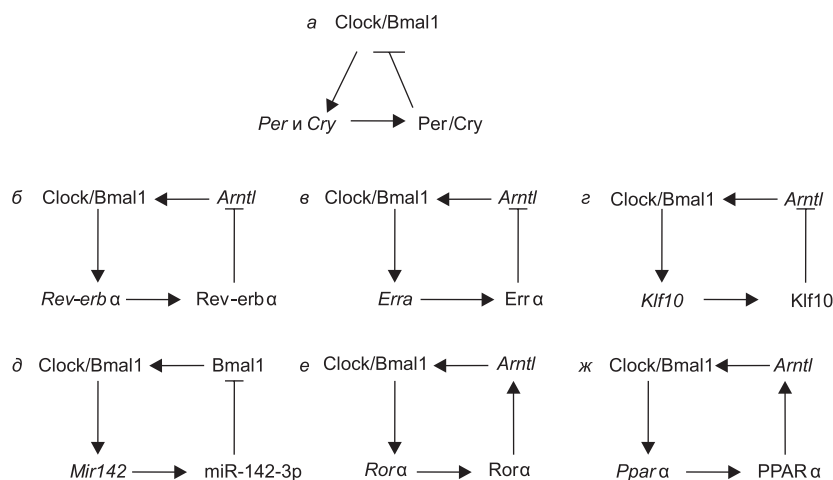


Рис. 3. Примеры регуляторных мотивов в генной сети Circadian Rhythm (комментарии см. в тексте).

этих мотивах, а также имеет максимальную степень вершин в графе, максимальный уровень центральности графа и минимальный эксцентриситет. Можно предположить, что центральная компонента регуляции циркадного ритма должна включать ТФ Clock/Bmal1, который обеспечивает взаимодействие с другими участниками процесса регуляции циркадного ритма. Такое взаимодействие обеспечено большим числом дублирующих путей, из которых наибольший интерес представляют пути с минимальным временем их реализации, которые и определяют главную регуляторную компоненту. В связи с этим для дальнейшего анализа структуры геновой сети Circadian rhythm и выявления центральной компоненты регуляции циркадного ритма мы избрали следующую стратегию:

1. Выделение классов реакций или регуляторных событий с различным временем реализации, например модификация, метаболические реакции, транспорт, процессы транскрипции или трансляции и др. Определение весов дуг в графе для реакций или регуляторных событий различных классов. В настоящей работе мы выделяли только два уровня класса реакций и событий: а) медленные (вес = 1): транспорт, процессы транскрипции или трансляции; б) быстрые (вес = 0,01 \ll 1,0): модификации белков, метаболические реакции и т. д. Использование весов дуг в графе позволяет считать, что время передачи сигнала между объектами в геновой сети пропорционально взвешенному расстоянию между соответствующими вершинами в графе геновой сети (сумма весов дуг пути между вершинами).

2. Поиск всех кратчайших путей, имеющих минимальное взвешенное расстояние от Clock/bmal1 до регуляторных элементов сети: ТФ, корегуляторов транскрипции и мРНК, представленных в геновой сети, считая, что они являются источниками регуляторных сигналов и регуляторных контуров (Clock/Bmal1; Atf4; Atf-5; Bmal1; Cebpa; Clock; Creb; Dbp; Dec1; Dec2; E4bp4; Err alpha; Foxo1; GR; Hes1; Hes6; Hif-1; Hlf; Hltf; Hnf4 alpha; Id2; Klf10; MyoD; Nrf1; Ppar alpha; Ppar alpha/Rxr alpha; Ppar gamma/Rxr alpha; Pparg; Prox1; Rar alpha; RelB; Rev-erb alpha; Ror alpha; Ror gamma; Rxr alpha; Shp; Smad3; Sirt1; Top1; Per1; Per2; Cry1; Cry2; miR-142-3p; miR-419).

3. Поиск всех кратчайших путей, имеющих минимальное взвешенное расстояние от найденных мишеней ТФ Clock/Bmal1 до Clock/Bmal1 (обратные связи).

4. Объединение всех полученных путей в граф и выделение в нем максимальной сильносвязной компоненты. Эта компонента будет объединять все кратчайшие регуляторные контуры (в смысле времени регуляции), связывающие Clock/Bmal1 с другими регуляторами циркадного ритма.

Такая процедура позволила выявить сильносвязную компоненту, соответствующую графу регуляторных контуров Clock/Bmal1, включающую 163 вершины и 188 дуг, объединенных в 27 регуляторных контуров (рис. 4).

Из 27 обратных связей 10 оказались отрицательными, 17 положительными. Выявленные нами регуляторные контуры, в конечном итоге, модулируют уровень экспрессии гена *Arntl*, а также и активность ТФ Clock/Bmal1. Эта модуляция может осуществляться через короткие контуры, как в очевидных случаях, показанных нами на рис. 3, и более сложными путями. Примером может быть положительная обратная связь между экспрессией гена *Per2* и уровнем активности ТФ Clock/Bmal1, которая, на первый взгляд, противоречит описанной выше структуре циркадного осциллятора. Путь реализации событий этой положительной обратной связи, представленных в графе, приведен на рис. 5. Рассмотрим более подробно процессы, обеспечивающие эту обратную связь:

1. Экспрессия гена *Per2* активируется ТФ Clock/Bmal1, ATF4.

2. Нарботанный в результате белок Per2 образует комплекс Per2/Cry2/CK1d.

3. Гетеродимер Per2/Cry2 проникает в ядро и подавляет активность ТФ Clock/Bmal1.

4. В результате этого снижается экспрессия гена *Mir142*, служащего мишенью ТФ Clock/Bmal1.

5. miR-142-3p, продукт гена *Mir-142*, является негативным регулятором экспрессии Bmal1, поэтому снижение его уровня приведет к увеличению уровня Bmal1 и увеличению активности ТФ Clock/Bmal1.

В данном случае контур с положительной обратной связью формируется за счет взаимодействия двух коротких контуров с отрицательной

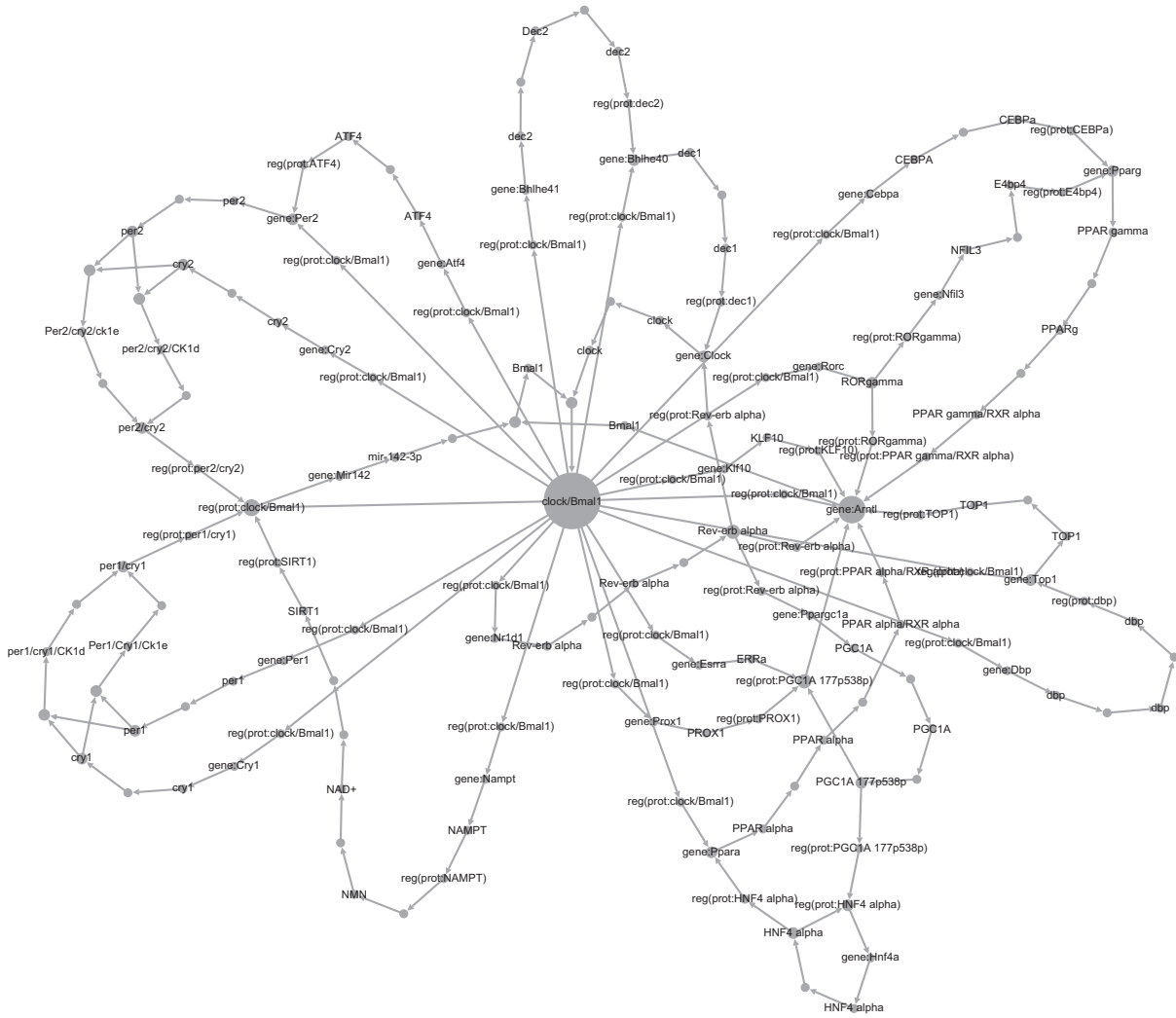


Рис. 4. Сильносвязная компонента графа генной сети Circadian Rhythm, образованная регуляторными контурами, связывающими Clock/Bmal1 с другими регуляторными элементами генной сети.

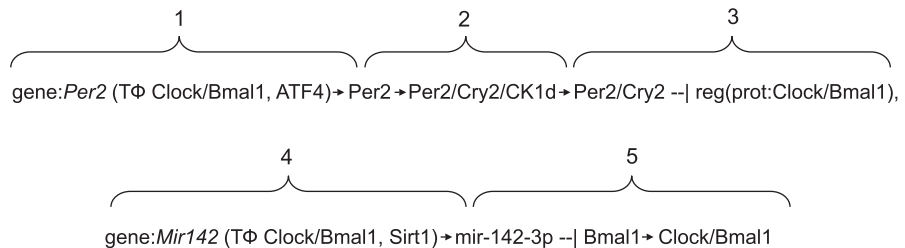


Рис. 5. Пример положительной обратной связи сильносвязной компоненты графа генной сети Circadian Rhythm.

обратной связью (см. рис. 3, а, д). Этот пример показывает продуктивность использованного нами подхода, позволяющего наряду с описанными выше петлями обратных связей выявить дополнительные более сложные регуляторные

контур, которые могут способствовать устойчивости работы циркадного осциллятора. Отметим, что представленные здесь регуляторные контуры и участвующие в них гены, вероятно, можно рассматривать как комплекс

компонент ядра циркадного осциллятора, в то время как группы генов, полученные нами в результате кластерного анализа вершин графа геновой сети, можно представить как элементы, выполняющие отдельные функции в рамках механизма циркадных часов. Так, белки и гены, объединенные в кластерах 1, 7, 12 и 18 (табл. 3) в сумме обеспечивают пути превращения белков Per 1-2 и Cry 1-2, определяющие динамику гетеродимеров Per(1-2)/Cry(1-2), от которой зависит величина периода циркадного осциллятора. Компоненты кластеров 5 и 13 отвечают за снижение активности ТФ Clock/Bmal1 за счет белок-белковых взаимодействий. Кластер 15 – синтез NAD⁺ и регуляция им активности фермента Parp1, одного из модификаторов Clock/Bmal1. Ранее отмечено, что среди прямых мишеней ТФ Clock/Bmal1 гены ТФ и корегуляторов транскрипции представляют значительную группу. И, хотя не все они образуют петли обратных связей с ядром осциллятора, наличие таких генов предполагает возможность вовлечения в циркадную ритмику экспрессии больших массивов генов за счет каскадной регуляции их транскрипции этими факторами.

ЗАКЛЮЧЕНИЕ

Представленные в работе подходы к реконструкции и анализу геновых сетей на основе методов теории графов обеспечивают возможность реконструкции и анализа сложно организованных геновых сетей, включая выявление их модульной организации. Применение этих подходов позволило реконструировать расширенный вариант геновой сети циркадного осциллятора млекопитающих, провести анализ структуры геновой сети и выделить центральную компоненту циркадного осциллятора, которая включает базовые регуляторные контуры, проходящие через ключевой элемент циркадных часов – белок Clock/Bmal1.

Использование кластерного анализа позволило выявить подсистемы, имеющие четкую биологическую интерпретацию и участвующие в функционировании циркадных часов путем взаимодействия с центральной компонентой. Такая структурная модель, включающая как центральную компоненту, так и взаимодействующие с ней функциональные подсистемы,

может стать основой для построения расширенной математической модели динамики геновой сети циркадного осциллятора.

БЛАГОДАРНОСТИ

Работа выполнена при поддержке гранта РФФИ № 14-24-00123.

ЛИТЕРАТУРА

- Подколотная О.А. Молекулярно-генетические аспекты взаимодействия циркадных часов и метаболизма энергетических субстратов млекопитающих // Генетика. 2014. Т. 50. № 2. С. 125–137.
- Albrecht U. Timing to Perfection: The Biology of Central and Peripheral Circadian Clocks // Neuron. 2012. V. 74. P. 246–260.
- Ananko E.A., Podkolodny N.L., Stepanenko I.L., Podkolodnaya O.A., Rasskazov D.A., Miginsky D.S., Likhoshvai V.A., Ratushny A.V., Podkolodnaya N.N., Kolchanov N.A. GeneNet in 2005 // Nucleic Acids Res. 2005. V. 33 (Database issue). P. D425–D427.
- Asher G., Schibler U. Crosstalk between components of circadian and metabolic cycles in mammals // Cell Metab. 2011. V. 13. P. 125–137.
- Barabasi A.-L., Oltvai Z.N. Network biology: understanding the cell's functional organization // Nat. Rev. Genet. 2004. V. 5. P. 101–113.
- Blondel V.D. *et al.* Fast unfolding of communities in large networks // J. Stat. Mechanics. 2008. V. 10008. P. 1–12.
- Erdős P., Rényi A. On random graphs I // Publ. Math. Debrecen. 1959. V. 6. P. 290–297.
- Kim W., Li M. *et al.* Biological network motif detection and evaluation // BMC Systems Biology. 2011. V. 5. P. S5.
- Koschützki D., Schreiber F. Centrality analysis methods for biological networks and their Application to gene regulatory networks // Gen. Regul. Syst. Bio. 2008. V. 2. P. 193–201.
- Morf J., Rey G., Schneider K. *et al.* Cold-inducible RNA-binding protein modulates circadian gene expression posttranscriptionally // Science. 2012. V. 338. P. 379–383.
- Newman M.E.J. The structure and function of complex networks // SIAM Reviews. 2003. V. 45. P. 167–256.
- Reppert M., Weaver D. Molecular analysis of mammalian circadian rhythms // Ann. Rev. Phys. 2001. V. 63. P. 647–676.
- Ripperger J.A., Brown S.A. Transcriptional regulation of circadian clock // Protein Reviews 2010. V. 12. P. 37–78.
- Tarjan R. Enumeration of the elementary circuits of a directed graph // SIAM J. Computing. 1973. V. 2. P. 211–216.
- Virshup D.M., Eide E.J., Forger D. *et al.* Reversible protein phosphorylation regulates circadian rhythms // Cold Spring Harb Symp Quant Biol. 2007. V. 72. P. 413–420.
- Watts D.J., Strogatz S.H. Collective dynamics of 'small-world' networks // Nature. 1998. V. 393 (6684). P. 440–442.
- Yagita K., Tamanini F., van Der Horst G.T., Okamura H. Molecular mechanisms of the biological clock in cultured fibroblasts // Science. 2001. V. 292. P. 278–281.

THE MAMMALIAN CIRCADIAN CLOCK: GENE REGULATORY NETWORK AND THEIR COMPUTER ANALYSIS

O.A. Podkolodnaya¹, N.N. Podkolodnaya^{1,2}, N.L. Podkolodnyy^{1,3}

¹Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia,
e-mail: opodkol@bionet.nsc.ru;

²Novosibirsk National Research State University, Novosibirsk, Russia;

³Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia

Summary

This paper presents the results of the reconstruction and analysis of gene regulatory network of the circadian clock in mammals. Application of graph theory methods makes it possible to analyze the structure of the gene network and identify the central component of circadian clock regulation, which includes the basic regulatory circuits passing through the key element of the circadian clock, the Clock/Bmal1 protein. Cluster analysis has revealed subsystems with clear biological interpretation, which are involved in the functioning of the circadian clock by interacting with the central component. This structural model, which includes the central component and functional subsystems that interact with the central component, can provide grounds for the construction of a mathematical model of the dynamics of the gene network regulating the circadian rhythm.

Key words: circadian clock, gene network, graph analysis methods.