

Выявление и анализ динамических паттернов суточной экспрессии генов млекопитающих

О.А. Подколотная¹, Н.Н. Твердохлеб^{1, 3}, Н.Л. Подколотный^{1, 2}✉

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук, Новосибирск, Россия

³ Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

Целью исследования было выявление и анализ паттернов суточной динамики экспрессии генов, различающихся по форме кривой. Можно ожидать, что сходство паттернов суточной экспрессии генов (формы кривой) является отражением синхронизации экспрессии генов общими внешними и внутренними сигналами или участия в сходных биологических процессах. Разные сигналы, имеющие суточную динамику (свет, активность, питание, стресс, температура и т. д.), могут воздействовать на разные уровни регуляции экспрессии, что может проявляться в различной форме паттернов суточной экспрессии генов. Работа выполнена с использованием экспериментальных данных по экспрессии генов на уровне трансляции (профилирование рибосом) в печени и почках мыши (GSE67305 и GSE81283). Для выявления генов с суточным ритмом экспрессии был использован однофакторный дисперсионный анализ. Предложен подход к выявлению сходных по форме кривых суточной динамики экспрессии генов на основе кластерного анализа. Расстояние между генами рассчитывалось путем выравнивания фаз и поиска максимальной по циклическому сдвигу кросс-корреляции между паттернами суточной экспрессии этих генов. Данный подход позволил выявить гены, имеющие не только паттерны экспрессии с одним максимумом (синусоидальные, асимметричные со смещением влево или вправо, импульсные), но и сложные комбинированные сигналы с несколькими экстремумами. В результате впервые выявлены группы генов, объединенных по сходству формы кривой суточной экспрессии, без учета их фазовых характеристик. Функциональный анализ обогащения терминами геновой онтологии групп генов с резко различающимися паттернами суточной экспрессии (синусоидальными и импульсными) в почках и печени мыши показал, что группа генов с синусоидальным паттерном суточной экспрессии в большей мере ассоциирована с регуляцией циркадного ритма и метаболизма. Группа генов с импульсным паттерном суточной экспрессии в значительной степени связана с защитными функциями организма, требующими формирования быстрого ответа. Показана информативность анализа динамических паттернов кривых суточной динамики экспрессии генов для функционального описания генов. Выделенные динамические паттерны суточной экспрессии имеют большое значение для дальнейшего изучения сложной циркадной регуляции, синхронизации и взаимодействия биологических процессов с суточной динамикой в организме млекопитающих.

Ключевые слова: циркадный ритм; трансляция; геновые онтологии; тканеспецифичность; биологические процессы; динамические паттерны суточной экспрессии.

Detection and analysis of dynamic patterns of diurnal expression of mammalian genes

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

¹ Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

² Institute of Computational Mathematics and Mathematical Geophysics, SB RAS, Novosibirsk

³ Novosibirsk State University, Novosibirsk, Russia

The purpose of the study is to identify and analyze patterns of the diurnal dynamics of the expression of genes that differ in the shape of the curve. It can be expected that the similarity of the patterns of daily expression (shape of the curve) of genes is a reflection of the synchronization of gene expression by common external and internal signals or participation in similar biological processes. Different signals that have daily dynamics (light, activity, nutrition, stress, temperature, etc.) can affect different levels of expression regulation, which can be manifested in various forms of patterns of daily gene expression. In our research, we used experimental data on gene expression at the level of translation (ribosome profiling) in the liver and kidney of a mouse (GSE67305 and GSE81283). To identify genes with a daily rhythm of expression, we used a one-way analysis of variance. To identify similar-in-shape curves of the daily dynamics of gene expression, we propose an approach based on cluster analysis. The distance between the genes was calculated by aligning the phases and finding the maximum cross-correlation between the patterns of the daily expression of these genes by the cyclic shift. This approach allowed us to identify genes that have not only expression patterns with a single maximum (sinusoidal, asymmetrical, shifted to the left or right, pulsed), but also complex composite signals with several extremes. As a result, the groups of genes united by the similarity of the shape of the daily expression curve without regard to their phase characteristics were identified. GO enrichment analysis of groups of genes with sharply different patterns of daily expression (sinusoidal and pulsed) in the mouse kidneys and liver showed that the group of genes with a sinusoidal pattern was more associated with regulation of circadian rhythm and metabolism. The group of genes with a pulsed pattern is largely associated with the protective functions of the organism, which require the quick response. Thus, our studies have confirmed the effectiveness of the proposed approach to the analysis of the diurnal dynamics of gene expression. The identified dynamic patterns of diurnal expression are important for the

further study of complex circadian regulation, synchronization and interaction of biological processes with diurnal dynamics in mammals.

Key words: circadian rhythm; translation; GO enrichment analysis; tissue specificity; biological processes; dynamic patterns of diurnal gene expression.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Подколотная О.А., Твердохлеб Н.Н., Подколотный Н.Л. Выявление и анализ динамических паттернов суточной экспрессии генов млекопитающих. Вавиловский журнал генетики и селекции. 2018;22(8):1055-1062. DOI 10.18699/VJ18.450

HOW TO CITE THIS ARTICLE:

Podkolodnaya O.A., Tverdokhle N.N., Podkolodny N.L. Detection and analysis of dynamic patterns of diurnal expression of mammalian genes. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2018;22(8):1055-1062. DOI 10.18699/VJ18.450 (in Russian)

Ритмические изменения с периодом, близким к 24 часам, наблюдаются практически на всех уровнях функционирования организмов млекопитающих – от молекулярно-генетических до сложных физиологических процессов, согласованное протекание которых обусловлено различными синхронизирующими сигналами и механизмами. К настоящему времени выявлены проявления органо-/тканеспецифичности суточной ритмики экспрессии генов и показано, что более сорока процентов белоккодирующих генов млекопитающих проявляют суточные осцилляции как минимум в одном органе на уровне мРНК (Zhang et al., 2014; Laing et al., 2015).

Циркадная регуляция функционирования генетических и метаболических сетей, путей передачи сигналов и т. д. осуществляется не только на уровне транскрипции, но и на уровне трансляции, посттранскрипционной и посттрансляционной модификации, транспорта и деградации мРНК и белков (Mendoza-Viveros et al., 2017). Нарушение циркадных ритмов может быть как причиной, так и следствием различных расстройств, включая метаболический синдром, воспалительные заболевания, ментальные расстройства, рак и др. (Sulli et al., 2018).

На сегодняшний день описаны механизм и различные факторы, определяющие суточные изменения в экспрессии генов (Flôres et al., 2016; Atger et al., 2017; Sulli et al., 2018). Основой функционирования этого механизма, безусловно, служат молекулярно-генетические циркадные осцилляторы, присутствующие практически в каждой клетке организма. У млекопитающих главный водитель циркадного ритма, сформированный нейронами супрахиазматических ядер гипоталамуса, синхронизирует ритм всей сети периферических циркадных осцилляторов (Ripperger, Brown, 2010). Однако физическая активность, прием пищи, температура и другие внешние и внутренние физиологические сигналы, имеющие суточную динамику, также могут формировать циклические изменения в уровнях экспрессии генов в клетках периферических органов. Циклическая повторяемость этих сигналов приводит к формированию и закреплению механизмов поддержки такого рода суточных изменений, отличных от циркадных часов.

Выявление среди генов с суточной динамикой различных групп генов, участвующих в регуляции биологических процессов, протекающих с суточной периодичностью, расширяет возможности исследования регуляции физиологических процессов и последствий их нарушения. При анализе данных по динамике суточной экспрессии генов, как правило, внимание сосредоточено на анализе

сходства и различия их фазовых характеристик (время суток с максимальным уровнем экспрессии) (Janich et al., 2015; Подколотный и др., 2017; Castelo-Szekely et al., 2017). Однако паттерны суточной экспрессии генов могут отличаться не только фазовыми характеристиками, но и формой кривой, описывающей циклическую динамику экспрессии.

Целью данной работы было выявление паттернов суточной динамики экспрессии генов, различающихся по форме кривой независимо от их фазы, и анализ групп генов, имеющих сходные динамические паттерны экспрессии.

Можно ожидать, что сходство паттернов суточной экспрессии (формы кривой) генов является отражением синхронизации экспрессии генов общими внешними и внутренними сигналами или участия в сходных биологических процессах. Это особенно актуально для исследования суточных ритмов, в формировании которых играют важную роль различные внешние и внутренние сигналы, имеющие суточную динамику, в частности свет, активность, питание, стресс, температура и т. д.

В работе предложен метод выявления сходных по форме кривых суточной динамики экспрессии генов. На этой основе выявлены группы генов, различающиеся динамическими паттернами суточной экспрессии. Результаты анализа обогащения терминами генной онтологии (Gene Ontology, здесь и далее – GO) выделенных групп генов с резко различающимися паттернами суточной экспрессии в почках и печени мыши подтвердили эффективность предложенного подхода к выявлению динамических паттернов кривых суточной динамики экспрессии генов, которые могут иметь большое значение для дальнейшего изучения сложной циркадной регуляции, синхронизации и взаимодействия биологических процессов с суточной динамикой в организме млекопитающих.

Материалы и методы

Использованы данные, полученные методом профилирования плотности рибосом на мРНК в печени и почках мыши, представленные в работах (Janich et al., 2015; Castelo-Szekely et al., 2017), которые доступны в базе данных GEO (печень: GSE67305 и почки: GSE81283). Этот подход позволяет получить информацию о транскрибуемых мРНК (транслятом) и их трансляционной активности в определенный момент времени *in vivo* (Ingolia, 2014). В настоящее время транслятом рассматривается как более точная, чем транскриптом, оценка уровня экспрессии и, как следствие, функциональности генов. Это связано

с тем, что корреляция данных протеома и транскриптома достаточно высока, и ее значение существенно выше корреляции протеома и транскриптома (Smircich et al., 2015).

В эксперименте мыши содержались в условиях светотемнового режима (12 ч свет/12 ч темнота) и свободного доступа к пище. Материал собирали каждые 2 ч в течение 24-часового суточного цикла (12 временных точек (ZT0–ZT22)), где ZT0 – включение света; ZT12 – выключение света). Каждой временной точке соответствовали две независимые реплики эксперимента (два животных, у каждого забирали печень и почки), для которых были получены значения плотности профилирования рибосом на мРНК анализируемых генов, использованные в нашем исследовании. Таким образом, массив данных включал 10980 генов, экспрессирующихся в печени со значениями плотности рибосом на мРНК, и 12622 гена – в почках.

При первичном анализе используемых экспериментальных данных мы выявили точки в кривых суточной динамики экспрессии некоторых генов, в которых разброс значений в двух репликах существенно выше, чем в других точках кривой. Эти точки можно рассматривать как выбросы. Пример такой кривой суточной экспрессии гена *Slc19a2* (*Thiamine transporter 1*) в печени мыши с экстремальным значением разброса плотности рибосом на мРНК в точке ZT12 приведен на рис. 1.

Необходимо отметить, что причиной существенных вариаций в дисперсии, имеющих, скорее всего, случайный характер, могут быть не только ошибки измерений, но и биологическая вариабельность, свойственная природе данных.

Для выявления генов с суточным ритмом экспрессии использовался однофакторный дисперсионный анализ на данных, в которых удалены временные точки с экстремальными выбросами экспрессии. Если суммарная по всем временным точкам средняя внутригрупповая дисперсия измерений уровня экспрессии была ниже, чем общая дисперсия сигнала с уровнем значимости $p < 0.001$, то считалось, что гены обладают ритмичной экспрессией. Такой жесткий критерий был взят для отбора генов с ярко выраженной воспроизводимостью паттерна суточной экспрессии, чтобы уменьшить влияние случайных вариаций в измерениях на результаты анализа ритмичных генов.

Предварительно для анализа были отобраны гены, экспрессия которых воспроизводится в репликах. Используемый нами подход позволил выделить гены, имеющие не только простые паттерны экспрессии с одним максимумом, но и сложные, ежесуточно воспроизводимые паттерны экспрессии со многими экстремумами.

Для выявления экстремальных выбросов в экспрессии генов мы используем следующую процедуру. Для каждого гена проводится дисперсионный анализ с целью проверки гипотезы о равенстве всех средних значений экспрессии этого гена μ_i в каждой группе i , соответствующей измерению в конкретной временной точке.

Нулевая гипотеза $H_0: \mu_1 = \mu_2 = \dots = \mu_{12}$. Альтернативная гипотеза: $H_a: \mu_i \neq \mu_j$ для некоторых i и j . Для проверки нулевой гипотезы используется статистика:

$$F = \frac{\sum_j n_j (\bar{X}_{.j} - \bar{X}_{..})^2 / (k-1)}{\sum_i \sum_j (\bar{X}_{ij} - \bar{X}_{.j})^2 / (N-k)}, \quad (1)$$

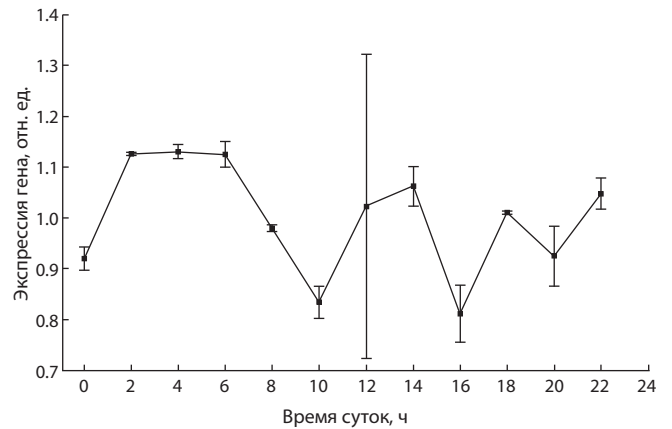


Рис. 1. Пример паттерна суточной экспрессии гена *Slc19a2* (*Thiamine transporter 1*) в печени мыши с экстремальным значением разброса плотности рибосом на мРНК в точке ZT12.

которая имеет F распределение с степенями свободы $k-1$ и $N-k$. Здесь N – общее число измерений во всех группах; k – число групп; n_j – число наблюдений в j -группе; $X_{ij} - i$ реплика ($i = [1, n_j]$) наблюдений экспрессии гена в j -группе ($j = 1, \dots, k$); $\bar{X}_{.j} = \sum_i \bar{X}_{ij} / n_j$ – внутригрупповое среднее для j -группы; $\bar{X}_{..} = \sum_j \bar{X}_{.j} / k$ – общее среднее по всем группам. В нашем случае $N = 24$, $k = 12$, $n_j = 2$.

Пусть в результате получен уровень значимости α_0 . Далее удаляем из анализа наблюдения последовательно одну из временных точек и проводим дисперсионный анализ для $k-1$ групп. Таким образом, получаем 12 значений уровня значимости критерия F, рассчитанных по 11 временным точкам ($\alpha_i, i = [1, 12]$). Если для некоторой временной точки i выполняется условие $\alpha_i < \alpha_0/k$, то временная точка i считается выбросом и удаляется из анализа. Для гена *Slc19a2* (*Thiamine transporter 1*), паттерн суточной экспрессии в печени мыши которого представлен на рис. 1, $\alpha_0 = 0.054$, т. е. нулевая гипотеза H_0 сохраняется и считается, что паттерн суточной экспрессии гена не воспроизводится. При удалении точки ZT12 с экстремальным значением разброса плотности рибосом на мРНК уровень достоверности, рассчитанный по оставшимся 11 точкам, равен: $\alpha_{12} = 8.45 \times 10^{-07}$, или с учетом поправки на множественное тестирование уровень значимости равен $1 - (1 - \alpha_{12})^{12} \approx 12\alpha_{12} = 0.00001$, т. е. показывает с высокой степенью значимости воспроизводимостью в репликах паттерна суточной экспрессии этого гена. Таким образом, выявлено, что 5 % генов содержат «выброс» в данных по экспрессии в печени и 3 % – в почках. В результате из всего массива данных удалено примерно 0.42 % измерений в печени и 0.25 % в почках.

Для выявления устойчивых типов паттернов суточной экспрессии среди генов мы применяли метод иерархического кластерного анализа с алгоритмом вычисления расстояний между кластерами по методу дальнего соседа или полной связи, реализованный в пакете MATLAB (программа linkage с опцией ‘complete’) (Statistics and Machine Learning Toolbox, 2018). Термин «паттерн суточной экспрессии» в данной работе относится к форме сигнала

ла, без учета его суточной фазы. Поэтому при проведении кластерного анализа расстояние между i и j генами рассчитывалось на основе выравнивания фаз и поиска максимальной по циклическому сдвигу кросс-корреляции между паттернами суточной экспрессии этих генов:

$$\text{dist}_{ij} = 1 - R_{ij}, \quad (2)$$

где $R_{ij} = \max_{\tau} \sum_t x_i(t) \cdot x_j(t-\tau)$; $x_i(t)$ – зависимость от времени t экспрессии i -го гена; $x_j(t-\tau)$ – циклический сдвиг фазы на время τ вектора суточной экспрессии j -го гена.

Для выделения групп генов с синусоидальными и импульсными паттернами суточной экспрессии применялся критерий на основе оценки значимости максимальной по циклическому сдвигу кросс-корреляции паттерна суточной экспрессии гена с соответствующими эталонными сигналами:

$$R_i = \max_{\tau} \text{Corr}(x_i(t), \text{signal}(t-\tau)). \quad (3)$$

В качестве эталонного синусоидального сигнала использовался $\text{signal}(t) = \cos(2\pi t/24)$, а в качестве импульсного сигнала – $\text{signal}(t) = [1, 0.3, 0, 0, 0, 0, 0, 0, 0, 0, 0.3]$, где $t = [0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22]$.

Паттерн суточной экспрессии гена считался подобным эталонному сигналу, если уровень значимости максимальной по циклическому сдвигу кросс-корреляции $p\text{-value} < 0.001$.

Анализ обогащения терминами GO групп генов проведен с использованием биоинформационного ресурса DAVID (Huang et al., 2009), который чаще всего применяется для такого рода анализа. Достоверность оценивали с помощью критерия Бенджамини–Хохберга с поправкой на множественность выбора ($p\text{-value} < 0.05$).

Результаты и обсуждение

Выявление динамических паттернов суточной экспрессии генов

На первом этапе работы были выявлены гены, демонстрирующие ритмические изменения экспрессии с периодом 24 ч, оцениваемые по величине показателя профилирования рибосом, в печени и почках. Количество ритмичных генов составило 4906 из 10980 генов, экспрессирующихся в печени, и 4455 из 12622 в почках.

При более детальном рассмотрении кривых, описывающих суточную экспрессию этих генов, мы обнаружили большое разнообразие форм кривых. Кластерный анализ позволил выявить различные паттерны ритмичной экспрессии, имеющие характерную форму: синусоидальные сигналы, асимметричные сигналы со смещением влево или вправо, импульсные сигналы и сигналы с комбинацией этих паттернов (сложные сигналы).

На рис. 2 представлены примеры кластеров с разными типами паттернов суточной экспрессии в печени мыши, выявленных методами кластерного анализа: синусоидальный (а), импульсный (б), асимметричный с левой и правой асимметрией (в, г), сложносоставной (композиционный) (д, е) типы паттернов.

Следует отметить, что гены, имеющие сложносоставной, или композиционный, паттерн суточной экспрессии, очень различаются между собой, и поэтому полученные кластеры включают, как правило, либо отдельные гены (см. рис. 2), либо малое их число.

Гены со схожими паттернами суточной экспрессии могут иметь различающиеся параметры суточного рит-

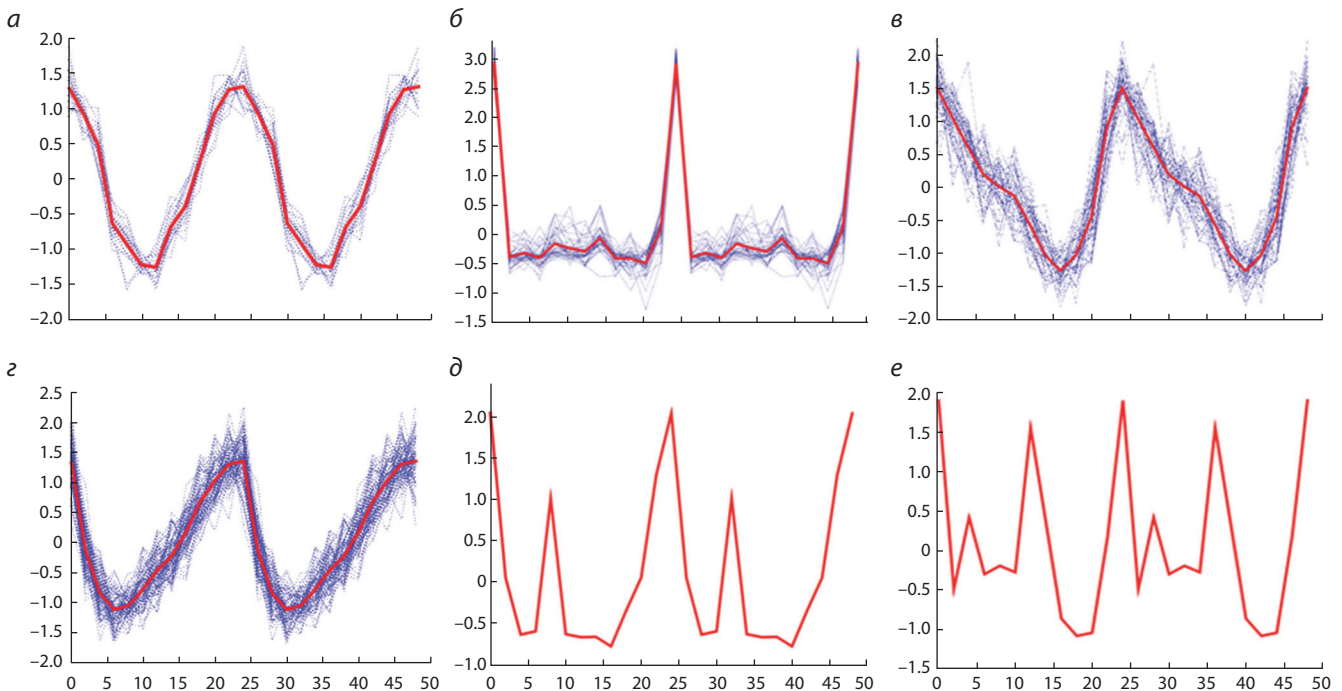


Рис. 2. Примеры кластеров генов с различными типами паттернов суточной экспрессии в печени мыши, выявленных методами кластерного анализа.

а – синусоидальный паттерн суточной экспрессии; б – импульсный паттерн суточной экспрессии; в, г – паттерны суточной экспрессии с левой и правой асимметрией соответственно; д, е – гены со сложными (составными) паттернами суточной экспрессии.

ма. В этом состоит основное отличие нашего подхода от общепринятых подходов к поиску и анализу групп генов с циркадной динамикой, в частности выявления групп коэкспрессирующихся генов, групп генов, имеющих повышенный уровень экспрессии в определенный момент времени, групп генов со сходными фазами сигналов и т. д. (Janich et al., 2015; Подколотный и др., 2017; Castelo-Szekely et al., 2017). И это обуславливает также различие в биологической интерпретации групп генов со схожими паттернами суточной экспрессии.

Динамический паттерн суточной экспрессии (формы кривой) генов формируется под воздействием синхронизации экспрессии генов общими внешними и внутренними сигналами, имеющими суточную динамику (свет, активность, питание, стресс, температура и т. д.). Разные сигналы могут проявляться в различной форме паттернов суточной экспрессии генов. Поэтому можно ожидать, что гены, имеющие сходные паттерны суточной экспрессии (форму сигнала), будут иметь сходную функцию.

В настоящей работе мы рассмотрим группы генов с двумя полярными типами экспрессионных паттернов, а именно синусоидальными и импульсными.

Из общего числа генов с ритмичной экспрессией синусоидальный паттерн экспрессии в печени выявлен у 766, в почках – у 508. Генов с импульсным паттерном экспрессии в печени найдено 188, в почках – 64.

Необходимо отметить, что синусоидальный паттерн суточной экспрессии имеет максимально сглаженную форму среди всех типов ритмичных паттернов, тогда как импульсный сигнал имеет максимальную скорость изменения экспрессии и наиболее резкий фронт. Эти особенности важны при интерпретации групп генов с такими формами экспрессии.

Прежде всего возникает вопрос, насколько выявленные группы генов уникальны для каждого органа. В группе ритмичных генов с синусоидальным паттерном экспрессии общими для двух органов оказались 123 гена. В группе с импульсным паттерном экспрессии общих для двух органов ритмичных генов было только 9 (см. таблицу). Здесь и далее будут использованы следующие названия групп генов: L(only)COS (Liver only cosine) – гены с синусоидальным паттерном экспрессии только в печени; K(only)COS (Kidney only cosine) – гены с синусоидальным паттерном экспрессии только в почках; intersectCos (intersect cosine) – гены с синусоидальным паттерном экспрессии в почках и печени; L(only)IMP (liver only impulse) – гены с импульсным паттерном экспрессии только в печени; K(only)IMP (Kidney only impulse) – гены с импульсным паттерном экспрессии только в почках; intersectImp (intersect impulse) – гены с импульсным паттерном экспрессии в почках и печени.

Функциональный анализ обогащения терминами генной онтологии групп генов с различными паттернами суточной экспрессии

Чтобы получить представление о функциональных характеристиках выявленных нами групп генов, мы сосредоточили внимание на оценке обогащенности этих групп терминами GO категорий Cellular Component и Biological Process, а также перепредставленности генов

Число генов с различными паттернами (синусоидальный и импульсный), ритмичных только в печени, только в почках, а также в печени и почках одновременно

Тип паттерна	Печень и почки	Печень	Почки
Синусоидальные	123 intersectCos	643 L(only)COS	385 K(only)COS
Импульсные	9 intersectImp	177 L(only)IMP	55 K(only)IMP

KEGG Pathway в анализируемых группах, позволяющей ассоциировать их с определенными путями (см. Приложение)¹.

Группа генов с синусоидальным паттерном экспрессии в обоих органах получила название intersectCos. Эта группа характеризовалась обогащением терминами категории Cellular Component, которые свидетельствуют о преимущественно внутриклеточной локализации их белковых продуктов, как в мембранных, так и немембранных органеллах, в том числе в ядре и структурах цитоскелета.

Согласно результатам анализа обогащения терминами GO категории Biological Process, в группе генов intersectCos перепредставлены термины, связанные с регуляцией циркадного ритма, циркадной регуляцией транскрипции, ответом на стероидные гормоны, в частности глюкокортикоиды, метаболизмом липидов и окислительно-восстановительными процессами. Отметим также, что путь Circadian rhythm базы данных KEGG оказался единственным, чьи гены достоверно перепредставлены в группе intersectCos. Паттерн суточной экспрессии генов этой группы наиболее близок к классическому синусоидальному с периодом 24 ч. Можно предполагать, что как в печени, так и в почках суточная ритмичность экспрессии генов группы intersectCos, вероятнее всего, определяется прямым или косвенным регулирующим влиянием эндогенного осциллятора, например циркадного. Это тем более вероятно, что основные гены циркадного осциллятора, а именно *Arntl*, *Cry1*, *Nr1d1*, *Per1*, *Per2*, *Per3*, *Nr1d2*, *Npas2*, *Rorc*, входят в данную группу. А большинство биологических процессов, характерных для группы генов intersectCos, известны как находящиеся под контролем циркадных часов (Yan, 2015). В то же время нельзя исключить влияния других типов эндогенных осцилляторов, таких как регулируемые доступностью пищи (food entrained oscillator) (Flóres et al., 2016). Кроме того, условия проведения эксперимента не исключают, что ритмичность экспрессии некоторых генов данной группы в определенной степени может определяться и внешними ритмичными сигналами, такими как световой режим, режим питания.

В группу генов с синусоидальным паттерном экспрессии в печени (L(only)COS) входили 643 гена. Эта группа характеризовалась обогащением терминами GO катего-

¹ Приложение см. по адресу: <http://www.bionet.nsc.ru/vogis/download/pict-2018-22/appx17.pdf>

рии Cellular Component, которые свидетельствуют о локализации их белковых продуктов в мембранных органеллах, в частности эндоплазматическом ретикулуме, митохондриях, аппарате Гольджи, а также во внеклеточных экзосомах.

Анализ обогащения терминами GO категории Biological Process показал, что в группе генов L(only)COS перепредставлены термины, связанные с метаболическими процессами широкого спектра биологических субстанций, включая липиды, жирные кислоты, аминокислоты, NAD, углеводы, и их регуляцией, а также апоптозом и ответом на инсулин.

Гены двух путей базы данных KEGG – Metabolic pathways и Carbon metabolism – перепредставлены в группе L(only)COS.

Группа ритмичных генов с синусоидальным паттерном экспрессии только в почках, получившее название K(only)COS, включала 385 генов и характеризовалась обогащением такими терминами категории Cellular Component, как внеклеточные экзосомы, эндоплазматический ретикулум, фокальные контакты.

Анализ обогащения GO терминами категории Biological Process показал, что в группе генов K(only)COS перепредставлены термины, резко отличающиеся от характерных для вышеописанных групп. Большая часть из них связана со сборкой и организацией клеточных компонентов, например актинового цитоскелета, адгезией, движением клеток или субклеточных компонентов. Кроме того, можно отметить термины, связанные с ответом на гипоксию.

Статистически значимого обогащения группы K(only)COS генами путей базы данных KEGG не наблюдается. Отметим, что общим для групп генов L(only)COS и K(only)COS, в отличие от группы intersectCos, является отсутствие терминов, ассоциированных с регуляцией циркадного ритма. В то же время нам удалось выявить тканеспецифичные особенности этих двух групп ритмичных генов. Прежде всего, различия касаются категории Cellular Component. Здесь при наличии общих терминов (*extracellular exosome* и *endoplasmic reticulum*) для группы генов печени характерны термины *mitochondrion*, *Golgi apparatus*, а для генов почек – *focal adhesion* и *cell junction*. Прослеживаются различия и для терминов категории Biological Process. Для генов печени характерны термины, связанные с процессами метаболизма, ответом на инсулин, апоптозом, тогда как для генов почек – термины, связанные с адгезией, движением клеток или субклеточных компонентов, ответом на гипоксию и стероидные гормоны.

Такие различия соответствуют функциональным особенностям двух органов. Печень играет особую роль в гомеостазе энергии, поэтому суточная ритмичность метаболических процессов именно в этом органе наиболее выражена и определяется не только циркадными часами, но и ритмами питания (Kornmann et al., 2007; Ribas-Latre, Eckel-Mahan, 2016). Нарушения пищевого поведения (голодание/переедание) являются мощными сигналами, опосредуемыми метаболическими или гормональными изменениями, способными перенастраивать периферические циркадные часы, в частности, в печени. Мутации ге-

нов циркадного осциллятора приводят к метаболическим нарушениям и развитию метаболических заболеваний, сопровождающихся нарушением чувствительности к инсулину (Zhou et al., 2014). Значительная часть энергии в организме продуцируется в печени митохондриями через окислительное фосфорилирование за счет метаболизма липидов и углеводов с наработкой АТФ. С этой же органеллой связан собственный путь апоптоза клетки. Отметим, что циркадные часы и режим питания могут участвовать не только в транскрипционной, но и в трансляционной регуляции экспрессии генов в печени (Atger et al., 2017).

Процессы, которые оказались ритмичными в группе генов K(only)COS, важны для функционирования почек. Так, термины, связанные со сборкой и организацией клеточных компонентов, адгезией, движением клеток или субклеточных компонентов, могут быть отражением адаптации клеточных структур нефрона, например подоцитов, прежде всего к ритмичным изменениям показателей гидростатического давления в капиллярах клубочков, которое подвержено интенсивным суточным колебаниям (Firsov, Bonny, 2018). Эти показатели определяются огромным множеством системных и локальных факторов. Любой из них может находиться под контролем как центральной, так и локальной системы циркадных часов, определяющих их ритмичность (Firsov, Bonny, 2018). Одним из сигналов, способных перенастраивать циркадные часы почек, является уровень кислорода как в крови, так и в почечной ткани, который ритмично изменяется в течение суток, достигая максимума в первой половине активной фазы (Adamovich et al., 2017). Связь между циркадными часами и уровнем кислорода в почке опосредуется транскрипционным фактором HIF1, активность которого индуцируется гипоксией. Следует отметить, что плотность рибосом на мРНК гена *Hif1a* ритмично изменяется в почках в течение суток и воспроизводится в репликах.

Мутация этого гена отменяет регулирующее действие уровня кислорода на циркадные часы почек (Adamovich et al., 2017). Реципрокная регуляция циркадных часов и ответа на гипоксию происходят за счет транскрипционной регуляции, а также путем белок-белковых взаимодействий (Wu et al., 2017). Таким образом, ответ на гипоксию в почках – ритмичный процесс, регулируемый как на системном, так и на локальном уровне.

Использованный нами подход позволил выявить специфическое обогащение групп генов L(only)COS и K(only)COS терминами GO категории Biological Process, что позволяет предполагать наличие тесной взаимосвязи между соответствующими процессами и генами с ритмическим типом экспрессии. Можно полагать, что активность этих процессов тоже ритмична.

Группа генов с импульсным паттерном экспрессии, общим для двух органов, получила название intersectImp и включала всего 9 генов. Возможно, малое число генов данной группы не позволило получить статистически значимых результатов анализа с использованием системы DAVID. Единственное, что можно отметить, это перепредставленность термина *Chaperone* функциональной категории UP_KEYWORDS ($p = 0.0015$), где 4 из 9 генов оказались шаперонами. Безусловно, мы не можем делать

каких-либо заключений на основе этих данных, однако напомним о роли шаперонов в поддержании циркадной ритмики через механизм деградации циркадных белков по убиквитин-протеасомному пути (Ripperger, Brown, 2010). Следует отметить, что суточная ритмичность экспрессии шаперонов может зависеть не только от циркадных часов, но и от циклов питания и изменения температуры тела (Reinke et al., 2008).

Группа генов с импульсным паттерном экспрессии в печени, получившая название L(only)IMP, включала 177 генов. В этой группе отмечена перепредставленность терминов *extracellular exosome*, *respiratory chain* GO категории Cellular Components. Результаты анализа обогащения терминами категории Biological Process показали перепредставленность генов с аннотацией терминами, связанными с защитными реакциями клетки и организма (например, *acute inflammatory response*, *acute-phase response*, *blood coagulation*, *regulation of complement activation*, *regulation of humoral immune response*, *fibrinolysis*, *blood coagulation*). Другая группа терминов связана с промежуточным метаболизмом (в частности, *respiratory electron transport chain*, *ATP metabolic process*, *generation of precursor metabolites and energy*, *oxidative phosphorylation*). Это в значительной степени подтверждается перепредставленностью генов путей *Complement and coagulation cascades*, *Oxidative phosphorylation* и *Metabolic pathways*, описанных в базе данных KEGG.

Первая группа терминов связана с защитной функцией организма. Поскольку основная задача циркадных часов – синхронизация всех процессов, протекающих в организме для адаптации его к ритмически изменяющимся внешним воздействиям, очевидно, что защитные функции и циркадные часы должны быть взаимосвязаны. Действительно, эти две системы имеют многочисленные взаиморегулирующие связи (Cermakian et al., 2014). Печень служит источником многих компонентов защитной системы организма, в том числе острофазных белков, белков системы комплемента и ряда регуляторов системы активации комплемента, являющихся важными компонентами иммунитета, белков свертывающей и противосвертывающей системы крови. Многие из этих белков секреторные и кодируются генами группы L(only)IMP.

Вторая группа терминов, связанная с оксидативным фосфорилированием, соответствует процессам, происходящим в митохондриях. Суточная ритмичность многих митохондриальных процессов подтверждена при исследованиях как в тканях печени, так и на изолированных митохондриях (de Goede et al., 2018). Целостность механизма молекулярных циркадных часов, и центральных, и периферических, необходима для полноценной регуляции митохондриального дыхания в клетках печени (de Goede et al., 2018). Кроме того, было показано, что внешние воздействия – состав пищи, режим кормления/голодание и др. – могут быть дополнительными регуляторами ритмичности этих процессов в печени (de Goede et al., 2018).

Группа генов с импульсным паттерном экспрессии в почках, названная K(only)IMP, включала 55 генов. В этой группе GO анализ не выявил обогащения терминами ни одной из категорий при заданном уровне значимости критерия Бенджамини–Хохберга.

Заключение

Суточные ритмы биологических процессов могут отражаться в суточной синхронизации различных участников этого процесса, число находящихся в активной фазе которых в течение суток может не меняться. Поэтому подходы, связанные с выявлением групп генов, находящихся в активной фазе в определенное время суток, с последующим функциональным GO анализом этих групп (Janich et al., 2015; Подколотный и др., 2017; Castelo-Szekely et al., 2017) позволяют выявлять цикличность только тех процессов, которые существенно меняют число генов-участников рассматриваемого процесса в течение суток. Отсутствие детального описания стадий процессов с указанием генов, участвующих в них, затрудняет проведение анализа циркадной регуляции большинства биологических процессов.

Однако паттерны суточной экспрессии генов различаются не только фазовыми характеристиками, но и формой кривой, описывающей циклическую динамику экспрессии. Сходство формы кривой суточной динамики экспрессии генов может быть отражением участия в сходных процессах продуктов этих генов.

В настоящей работе предложен подход к выделению различных типов форм кривых суточной динамики экспрессии генов независимо от их фазы. Такой динамический паттерн суточной экспрессии генов можно рассматривать как дополнительную функционально важную характеристику генов. Это актуально именно для исследования суточной динамики экспрессии генов, которые могут реагировать на различные синхронизирующие сигналы: свет, физическую, эмоциональную или умственную активность, питание, стресс, изменение температуры и т. д. Разные сигналы могут воздействовать на различные уровни регуляции экспрессии, что может проявляться в различной форме паттернов суточной экспрессии генов.

Предложенный нами подход позволил выделить группы генов, сходные по форме кривой суточной экспрессии, без учета их фазовых характеристик. Выявлены группы, которые имеют не только паттерны экспрессии с одним максимумом (синусоидальные, асимметричные со смещением влево или вправо, импульсные), но и сложные, ежесуточно воспроизводящиеся, с несколькими максимумами.

Функциональный анализ обогащения терминами GO групп генов с резко различающимися паттернами суточной экспрессии (синусоидальными и импульсными) в почках и печени мыши показал, что группа генов с синусоидальным паттерном суточной экспрессии как в почках, так и в почках в большей мере ассоциирована с регуляцией циркадного ритма и метаболизма. В то же время нам удалось выявить тканеспецифичные особенности групп генов с синусоидальным паттерном суточной экспрессии, которые оказались связаны с функциональными особенностями двух органов. Группа генов с импульсным паттерном суточной экспрессии в значительной степени связана с защитными функциями организма, требующими формирования быстрого ответа. В формирование таких ответов существенный вклад должны вносить механизмы, способные обеспечить значительное увеличение экспрессии генов за короткое время. Одним из таких механизмов может быть регуляция эффективности трансляции генов (Janich et al., 2015; Castelo-Szekely et al., 2017).

Таким образом, наши исследования подтвердили эффективность предложенного подхода к выявлению и анализу кривых суточной динамики экспрессии генов. Выделенные динамические паттерны суточной экспрессии имеют большое значение для дальнейшего изучения сложной циркадной регуляции, синхронизации и взаимодействия биологических процессов с суточной динамикой в организме млекопитающих.

Благодарности

Разработка методов анализа данных и программного обеспечения выполнена в рамках государственного задания ИВМиМГ СО РАН по проекту 0315-2016-0005. Подготовка данных, проведение расчетов и биологическая интерпретация были проведены при поддержке бюджетного проекта № 0324-2018-0017.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Список литературы / References

Подколodный Н.Л., Тverdokhleб Н.Н., Подколodная О.А. Анализ циркадного ритма биологических процессов в печени и почках мыши. Вавиловский журнал генетики и селекции. 2017;21(8): 903-910. DOI 10.18699/VJ17.311.
[Podkolodnyy N.L., Tverdokhleб N.N., Podkolodnaya O.A. Analysis of the circadian rhythm of biological processes in mouse liver and kidney. Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding. 2017;21(8):903-910. DOI 10.18699/VJ17.311. (in Russian)]
Adamovich Y., Ladeux B., Golik M., Koeners M.P., Asher G. Rhythmic oxygen levels reset circadian clocks through HIF1 α . Cell Metab. 2017;25(1):93-101. DOI 10.1016/j.cmet.2016.09.014.
Atger F., Mauvoisin D., Weger B., Gobet C., Gachon F. Regulation of mammalian physiology by interconnected circadian and feeding rhythms. Front. Endocrinol. (Lausanne). 2017;8:42. DOI 10.3389/fendo.2017.00042. eCollection 2017.
Castelo-Szekely V., Arpat A.B., Janich P., Gatfield D. Translational contributions to tissue specificity in rhythmic and constitutive gene expression. Genome Biol. 2017;18(1):116.
Cermakian N., Westfall S., Kiessling S. Circadian clocks and inflammation: reciprocal regulation and shared mediators. Arch. Immunol. Ther. Exp. (Warsz). 2014;62(4):303-318. DOI 10.1007/s00005-014-0286-x. Epub 2014 Apr 1.
de Goede P., Wefers J., Brombacher E.C., Schrauwen P., Kalsbeek A. Circadian rhythms in mitochondrial respiration. J. Mol. Endocrinol. 2018;60(3):R115-R130. DOI 10.1530/JME-17-0196. Epub 2018 Jan 29.
Firsov D., Bonny O. Circadian rhythms and the kidney. Nat. Rev. Nephrol. 2018;14(10):626-635. DOI 10.1038/s41581-018-0048-9.
Flôres D.E., Bettilyon C.N., Yamazaki S. Period-independent novel circadian oscillators revealed by timed exercise and palatable meals. Sci. Rep. 2016;6:21945. DOI 10.1038/srep21945.
Huang D.W., Sherman B.T., Lempicki R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 2009;4:44-57.

Ingolia N.T. Ribosome profiling: new views of translation, from single codons to genome scale. Nat. Rev. Genet. 2014;15(3):205-213. Epub 2014; Jan 28. DOI 10.1038/nrg3645.
Janich P., Arpat A.B., Castelo-Szekely V., Lopes M., Gatfield D. Ribosome profiling reveals the rhythmic liver transcriptome and circadian clock regulation by upstream open reading frames. Genome Res. 2015;25(12):1848-1859.
Kornmann B., Schaad O., Bujard H., Takahashi J.S., Schibler U. System-driven and oscillator-dependent circadian transcription in mice with a conditionally active liver clock. PLoS Biol. 2007;5(2):e34. DOI 10.1371/journal.pbio.0050034.
Laing E.E., Johnston J.D., Möller-Levet C.S., Bucca G., Smith C.P., Dijk D.J., Archer S.N. Exploiting human and mouse transcriptomic data: Identification of circadian genes and pathways influencing health. Bioessays. 2015;37(5):544-556. DOI 10.1002/bies.201400193.
Mendoza-Viveros L., Bouchard-Cannon P., Hegazi S., Cheng A.H., Pastore S., Cheng H.M. Molecular modulators of the circadian clock: lessons from flies and mice. Cell Mol. Life Sci. 2017;74(6):1035-1059. DOI 10.1007/s00018-016-2378-8.
Reinke H., Saini C., Fleury-Olela F., Dibner C., Benjamin I.J., Schibler U. Differential display of DNA-binding proteins reveals heat-shock factor 1 as a circadian transcription factor. Genes Dev. 2008;22(3):331-345. DOI 10.1101/gad.453808.
Ribas-Latre A., Eckel-Mahan K. Interdependence of nutrient metabolism and the circadian clock system: Importance for metabolic health. Mol. Metab. 2016;14;5(3):133-152. DOI 10.1016/j.molmet.2015.12.006.
Ripperger J.A., Brown S.A. Transcriptional Regulation of Circadian Clocks. In: Albrecht U. (Ed.) The Circadian Clock. N. Y.: Springer, 2010;37-78.
Smircich P., Eastman G., Bispo S., Duhagon M.A., Guerra-Slompo E.P., Garat B., Goldenberg S., Munroe D.J., Dallagiovanna B., Holecz F., Sotelo-Silveira J.R. Ribosome profiling reveals translation control as a key mechanism generating differential gene expression in *Trypanosoma cruzi*. BMC Genomics. 2015;16:443. DOI 10.1186/s12864-015-1563-8.
Statistics and Machine Learning Toolbox™ User's Guide. 2018. The Math Works, Inc. https://uk.mathworks.com/help/pdf_doc/stats/stats.pdf
Sulli G., Manoogian E.N.C., Taub P.R., Panda S. Training the circadian clock, clocking the drugs, and drugging the clock to prevent, manage, and treat chronic diseases. Trends Pharmacol. Sci. 2018;39(9):812-827. DOI 10.1016/j.tips.2018.07.003. Epub 2018 Jul 27.
Wu Y., Tang D., Liu N., Xiong W., Huang H., Li Y., Ma Z., Zhao H., Chen P., Qi X., Zhang E.E. Reciprocal regulation between the circadian clock and hypoxia signaling at the genome level in mammals. Cell Metab. 2017;25(1):73-85. DOI 10.1016/j.cmet.2016.09.009.
Yan Q., Cellular Rhythms and Networks: Implications for Systems Medicine. (Springer Briefs in Cell Biology). Springer, 2015. DOI 10.1007/978-3-319-22819-8.
Zhang R., Lahens N.F., Balance H.I., Hughes M.E., Hogenesch J.B. A circadian gene expression atlas in mammals: implications for biology and medicine. Proc. Natl. Acad. Sci. USA. 2014;111(45): 16219-16224. PMID 25349387. DOI 10.1073/pnas.1408886111.
Zhou B., Zhang Y., Zhang F., Xia Y., Liu J., Huang R., Wang Y., Hu Y., Wu J., Dai C., Wang H., Tu Y., Peng X., Wang Y., Zhai Q. CLOCK/BMAL1 regulates circadian change of mouse hepatic insulin sensitivity by SIRT1. Hepatology. 2014;59(6):2196-206. DOI 10.1002/hep.26992.