

УДК 575.162

## ГЕНЫ, КОНТРОЛИРУЮЩИЕ ПИЩЕВОЕ ПОВЕДЕНИЕ И МАССУ ТЕЛА ЧЕЛОВЕКА, И ИХ ФУНКЦИОНАЛЬНЫЕ И ГЕНОМНЫЕ ХАРАКТЕРИСТИКИ

© 2014 г. Е.В. Игнатьева<sup>1-3</sup>, Д.А. Афонников<sup>1-3</sup>, Е.И. Рогаев<sup>2</sup>,  
Н.А. Колчанов<sup>1,3</sup>

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

<sup>2</sup> Центр нейробиологии и нейрогенетики мозга, Новосибирск, Россия;

<sup>3</sup> Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

Поступила в редакцию 29 сентября 2014 г. Принята к публикации 27 октября 2014 г.

С целью систематизации информации о генах, участвующих в регуляции массы тела и пищевого поведения, была сформирована компиляция, включающая 424 гена, полученных (а) по данным из экспериментальных и обзорных статей, (б) из базы OMIM, (в) по данным мета-анализа экспериментов по полногеномному поиску ассоциаций. Четыре гена из компиляции (*BDNF*, *MC4R*, *PCSK1*, *POMC*) подтверждены всеми тремя источниками данных и рассматриваются как наиболее значимые в системе регуляции массы тела (приоритет 1). Выявлены группы, включающие 3 и 29 генов, подтвержденных двумя из трех источников данных (приоритет 2). Идентифицированы метаболические и сигнальные пути, участвующие в регуляции массы тела, которые можно считать потенциальными мишенями для фармакологических воздействий. Обнаружены районы хромосом человека, содержащие близкорасположенные гены из компиляции, содержащиеся в числе других генов, внесенные в компиляцию только по данным мета-анализа экспериментов по полногеномному поиску ассоциаций (*ETV5*, *MIR148A*, *NFE2L3*, *TMEM160*), что может помочь интерпретировать функции этих генов. К числу двенадцати генов из компиляции, наименее толерантных к мутациям, отнесены гены *LRPI*, *LRP5*, *RAI1*, *FASN*, *LYST*, *RPTOR*, *DGKD*, *LRPIB*, *NCOA1*, *ADCY3*. Компиляция может быть полезна как источник информации о генах-кандидатах, значимых для оценки риска развития ожирения и разработки фармакологических подходов к коррекции избыточной массы тела.

**Ключевые слова:** пищевое поведение, регуляция массы тела, локализация в геноме, толерантность к мутациям.

### ВВЕДЕНИЕ

Регуляция массы тела – сложный фенотипический признак, который контролируется как генетическими факторами, так и факторами среды. Генетические факторы, провоцирующие ожирение, достаточно сложны. На долю моногенных форм заболевания приходится всего 5 % от всех наблюдаемых случаев в человеческих популяциях. Наиболее изученными локусами, мутации которых связаны с моногенными формами патологии, являются гены, кодирующие лептин (*LEP*) и его рецептор (*LEPR*), рецептор

меланокортина типа 4 (*MC4R*), проопиомеланокортин (*POMC*), пробелок конвертаза субтилизин/кексин типа 1 (*PCSK1*), целенаправленный гомолог-1 гена дрозофилы (*SIMI*) (Blakemore *et al.*, 2010; Zegers *et al.*, 2012).

Обнаружение новых генов, нарушение которых связано с повышенной массой тела, позволяет выявлять биохимические и сигнальные пути, а также механизмы их регуляции, контролирующие рассматриваемый фенотипический признак. Такие гены могут рассматриваться как гены-кандидаты, используемые

для оценки риска развития патологии, а также как потенциальные мишени для фармакологических воздействий. Выявление генов, ассоциированных с патологиями, осуществляется с использованием различных стратегий, включая семейный анализ, исследование генов-кандидатов, а также полногеномный анализ ассоциаций (GWAS) (Hejblum *et al.*, 2011). Однако каждый из подходов имеет свои ограничения: 1) семейный анализ позволяет выявлять только варианты с высокой пенетрантностью; 2) результаты, полученные в ассоциативных исследованиях генов-кандидатов на разных популяционных выборках, не всегда подтверждают друг друга; 3) данные экспериментов GWAS с трудом поддаются интерпретации ввиду слабой изученности функциональной роли многих генов. Таким образом, несмотря на обилие информации, полученной различными экспериментальными методами, в том числе с использованием полногеномных подходов, данные о генетических предпосылках развития ожирения еще не достаточно полны.

Основой для формирования списка генов, потенциально задействованных в развитии ожирения, может быть рассмотрение физиологических систем, контролирующих массу тела, включая базовый метаболизм, регулируемый нервной, эндокринной и иммунной системами.

Важнейшей системой организма, от которой зависит масса тела, является система регуляции пищевого поведения, функционирующая при участии белковых продуктов генов, экспрессируемых как в мозге (Olszewski *et al.*, 2008), так и в периферических органах и тканях: желудке, кишечнике, поджелудочной железе, жировой ткани. Центральное звено системы составляют нейроны аркуатных ядер гипоталамуса, секретирующие нейропептид Y (NPY) и агутиподобный белок (AgRP), а также альфа-меланоцитстимулирующий гормон ( $\alpha$ -MSH), который образуется из проопиомеланокортина (POMC) под действием прогормон-конвертаз (PCSK1 и PCSK2) (Yeo, Heisler, 2012). Активность нейронов аркуатных ядер гипоталамуса контролируется гормонами (лептином, инсулином, грелином, полипептидом YY (PYY), глюкокортикоидами, адренкортикотропином, кортикотропин-релизинг гормоном), нейромедиаторными системами мозга (серотонергиче-

ская, дофаминергическая, адреналиновая, ГАМК-ергическая), а также нейротрофическими факторами BDNF и др. (Yeo, Heisler, 2012; Maniam *et al.*, 2012).

Неотъемлемый этап исследований генетических основ предрасположенности к заболеваниям – теоретическая оценка потенциального влияния конкретных нуклеотидных замен на уровень экспрессии гена или функцию белка (Ponomarenko *et al.*, 2002; Choi *et al.*, 2012) либо возможной роли отдельных генов в развитии патологического процесса (Masoudi-Nejad *et al.*, 2012; Smedley *et al.*, 2014). Petrovski с соавт. (2013) была предложена мера Residual Variation Intolerance Score (RVIS), которая характеризует толерантность гена к мутациям. Отрицательный показатель RVIS свидетельствует о том, что ген находится под давлением стабилизирующего отбора, а положительный указывает на то, что стабилизирующий отбор ослабевает и, наоборот, возможен движущий либо балансирующий отбор. При сравнении выборки генов из базы OMIM со всеми остальными генами человека была выявлена достоверная корреляция между связью гена с заболеванием и пониженным (относительно среднего значения, рассчитанного для полногеномной выборки генов) значением RVIS. Таким образом, мера RVIS представляется нам удобным критерием оценки возможной роли генов в развитии патологии.

Целями нашей работы было формирование компиляции генов, участвующих в регуляции пищевого поведения и массы тела, определение их функциональных и геномных характеристик, а также выявление генов с пониженной толерантностью к мутациям на основе показателя RVIS, что указывает на значимость таких генов в развитии патологии. Была сформирована компиляция, включающая 424 гена, четыре из которых (*BDNF*, *MC4R*, *PCSK1*, *POMC*) имели наивысший приоритет. Выявлены метаболические и сигнальные пути, участвующие в регуляции массы тела. Обнаружены участки генома человека, содержащие близкорасположенные гены из компиляции. С использованием показателя RVIS установлено, что содержащиеся в компиляции гены *LRP1*, *LRP5*, *RAI1*, *FASN*, *LYST*, *RPTOR*, *DGKD*, *LRPIB*, *NCOA1*, *ADCY3*, *ZNF608*, *INSR* наименее толерантны к мутациям, и это указывает на повышенную значимость

мутантных вариантов этих генов в развитии патологий, в частности ожирения.

## МАТЕРИАЛЫ И МЕТОДЫ

Гены были внесены в компиляцию на основании трех информационных источников (табл. 1). Первым источником, именуемым в дальнейшем *Публикации*, были научные статьи (включая как экспериментальные, так и обзорные), в которых охарактеризована роль генов в регуляции пищевого поведения у человека либо у близких видов (мыши и крысы). Если в статье была приведена информация о гене млекопитающего, в компиляцию вносился гомологичный ген человека с соответствующей ссылкой на источник данных.

Вторым источником была база OMIM. Поиск по базе OMIM проводился по ключевым словам «obesity», либо «hyperphagia», либо «anorexia». В результате поиска из базы OMIM получен список из 333 генов, именуемый в дальнейшем *OMIM*. Список генов *OMIM* включал гены двух категорий. К первой, именуемой в дальнейшем *OMIM-allelic variant*, отнесены 73 гена, имеющие хотя бы одно из вышеперечисленных клю-

чевых слов в поле «allelic variants». Ко второй категории, именуемой *OMIM-all text*, отнесены 260 генов, имеющие перечисленные выше ключевые слова только в поле «all text». Поскольку для генов, отнесенных к категории *OMIM-all text*, информация об аллельных вариантах, ассоциированных с изменением пищевого поведения, отсутствовала, гены этой группы можно рассматривать как участвующие в регуляции пищевого поведения и массы тела и выявленные в базе OMIM методом Text Mining.

Третий источник составили статьи, посвященные результатам мета-анализа данных экспериментов по полногеномному поиску ассоциаций (GWAS). В компиляцию были включены гены человека, для которых в результате мета-анализа была показана ассоциация с повышенным индексом массы тела с достоверностью  $p < 5,0 \times 10^{-8}$  (полногеномный уровень). В этот список, именуемый в дальнейшем *GWAS-мета-анализ*, были включены 48 генов человека. При этом 39 генов из 48 были выявлены на основе мета-анализа данных GWAS, полученных с использованием популяционных выборок населения европеоидного происхождения. Кроме того, в проанализированных нами статьях были

Таблица 1

Информационные источники, на основе которых была сформирована компиляция генов пищевого поведения и регуляции массы тела

Информационный источник	Краткое обозначение информационного источника	Кол-во генов	Кол-во публикаций либо поисковый запрос
Научные публикации, о функциональной причастности генов к регуляции пищевого поведения База данных OMIM	<i>Публикации</i>	83	33 (11 обзорных статей, 22 статьи, описывающие данные экспериментов)
	<i>OMIM-allelic variants</i>	73	‘Search: ‘hyperphagia’ OR ‘obesity’ OR ‘anorexia’ (Records with: gene map locus; Prefixes: +, *; Search in: allelic variants)
	<i>OMIM-all text</i>	260	‘Search: ‘hyperphagia’ OR ‘obesity’ OR ‘anorexia’ (Records with: gene map locus; Prefixes: +, *; Search in: all text) <sup>1</sup>
Научные публикации, описывающие, результаты мета-анализа данных GWAS	<i>GWAS-мета-анализ</i>	48	6

<sup>1</sup> Гены, полученные в результате запроса по полю allelic variants, были исключены из списка OMIM-all text.

представлены данные анализа популяционных выборок африканского, восточно-азиатского, северо-американского и австралийского происхождения.

Функциональную аннотацию генов и выявление обогащенных терминов Gene Ontology и метаболических путей осуществляли с помощью Интернет-инструмента DAVID (Huang *et al.*, 2009). Расположение генов на хромосомах исследовали на основе данных таблицы gene2refseq из базы EntrezGene. Для всех генов генома человека были определены координаты их центров, вычисленные как среднее арифметическое между позициями начала и конца каждого гена на хромосоме. Близкорасположенные гены выявляли, используя позиции центров генов. Проводили поиск геномных районов, содержащих центральные позиции четырех генов в пределах 2 Мб.

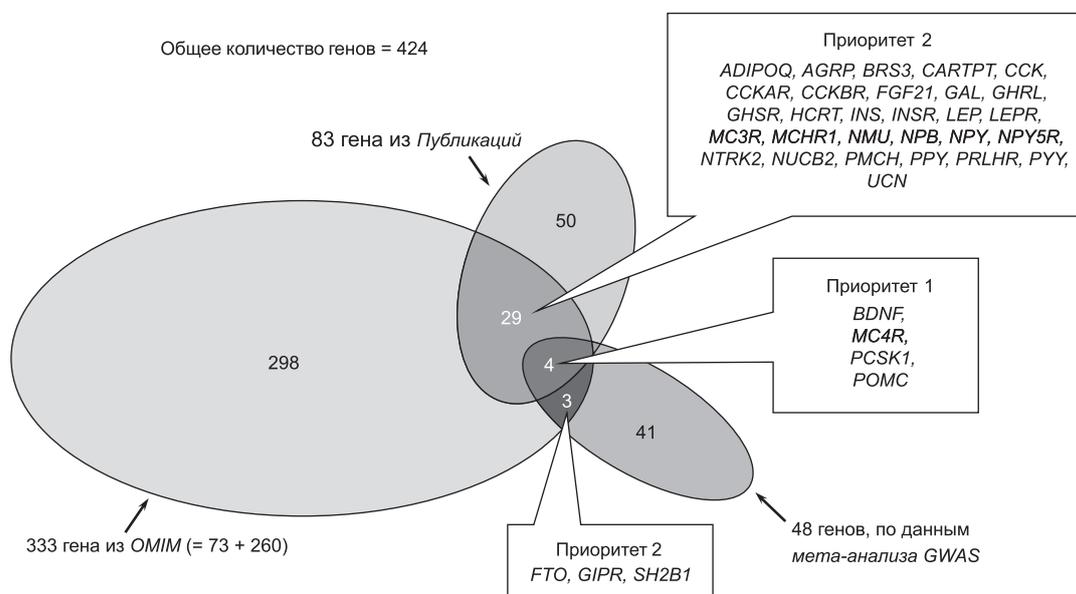
Чтобы оценить толерантность генов к мутациям, использовали скор RVIS. Значения RVIS для 16 956 генов человека (Dataset S2) были экстрагированы из публикации Petrovski с соавт. (2013). В этом исследовании RVIS определялся на основе соотношения количества замен с частотой встречаемости минорного аллеля выше

0,1 ( $MAF > 0,1$ ) к общему количеству замен в кодирующих частях генов. Значение RVIS, равное нулю, соответствует среднему значению этой величины для всех генов из выборки, проанализированной авторами, т. е. практически среднему значению по геному. У генов с положительным RVIS количество замен с  $MAF > 0,1$  превышает ожидаемое, а у генов, имеющих отрицательный RVIS, обнаружено пониженное (по сравнению со среднегеномным) количество замен с  $MAF > 0,1$  (Petrovski *et al.*, 2013).

## РЕЗУЛЬТАТЫ

### Информационное содержание компиляции

С использованием трех информационных источников *Публикации*, *OMIM*, *GWAS-мета-анализ* (см. раздел Материалы и методы и табл. 1) сформирована компиляция объемом 424 гена (рис. 1). 83 гена были описаны в публикациях, 333 гена внесены в компиляцию на основе запросов к базе *OMIM*, и 48 генов внесены по данным информационного источника *GWAS-мета-анализ*. Списки генов, полученных из трех информационных источников, частично



**Рис. 1.** Количество генов, внесенных в компиляцию на основе информационных источников *Публикации*, *OMIM* и *GWAS-мета-анализ* (представление в форме диаграммы Венна). Гены, выявленные по данным всех трех источников (приоритет 1) либо по данным из двух источников (приоритет 2), обозначены на выносках, исходящих из соответствующих областей диаграммы.

перекрывались (рис. 1). Обнаружено, что четыре гена (*BDNF*, *MC4R*, *PCSK*, *POMC*) имеют свидетельства о причастности к регуляции пищевого поведения либо массы тела из всех трех информационных источников (*Публикации*, *OMIM*, *GWAS-мета-анализ*). Следовательно, эти гены можно рассматривать как наиболее значимые в системе регуляции массы тела (на рис. 1 обозначены как Приоритет 1).

Выявлены также два подмножества генов, принадлежащие двум из трех списков. Три гена (*FTO*, *GIPR*, *SH2B1*) содержались на пересечении списков *OMIM* и *GWAS-мета-анализ*. Еще 29 генов отмечены на пересечении списков *Публикации* и *OMIM*. На рис. 1 эти два подмножества генов обозначены как Приоритет 2.

### Функциональная аннотация генов

С использованием системы DAVID был охарактеризован набор клеточных функций белков, кодируемых генами из компиляции. Около четверти генов (24,8 %) кодируют рецепторы клеточной поверхности; 23,9 % генов кодируют белки, проаннотированные GO термином *receptor binding activity*, т. е. сигнальные молекулы (гормоны, нейропептиды и т. д.); 15 % генов кодируют транскрипционные регуляторы, и, наконец, 7 % генов кодируют белки с киназной активностью (рис. 2). Около четверти генов из компиляции составили гены с очень гетерогенными функциями, эта группа на рис. 2 имеет обозначение Other.

С использованием Интернет-инструмента DAVID были выявлены метаболические и сигнальные пути, неслучайно часто ( $p < 10^{-2}$ ) представленные, т. е. перепредставленные, в аннотации генов из компиляции. Уровень перепредставленности выше двух имели 15 путей из базы KEGG, 7 путей из REACTOME и 8 путей из BIOCARTA (рис. 3).

### Локализация генов в геноме

Расположение генов из компиляции в геноме человека представлено на рис. 4. Обнаружено одиннадцать геномных районов протяженностью 2 Мб, включающих центральные позиции четырех близкорасположенных генов. Четыре из одиннадцати выявленных геномных района

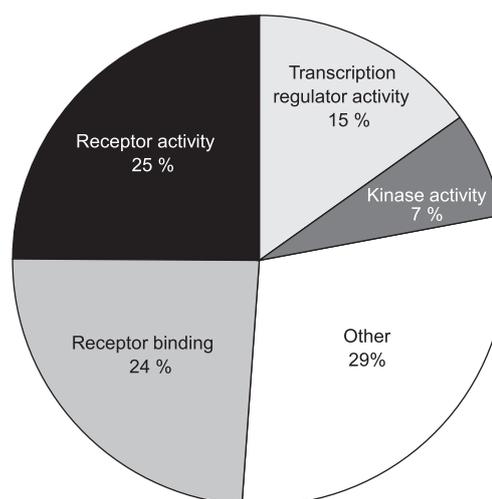
длиной 2 Мб располагались отдельно друг от друга на хромосомах 7, 11, 19, 20.

Кроме того, в трех ситуациях (хромосомы 3, 17, 19), выявленные нами районы протяженностью 2 Мб перекрывались между собой. На хромосоме 3 на перекрывании таких участков содержится пятерка близкорасположенных генов (*IGF2BP2*, *ETV5*, *HRG*, *KNG1*, *ADIPOQ*). На хромосоме 17 найдена семерка генов (*HCRT*, *STAT3*, *PPY*, *PYY*, *SLC4A1*, *CRHR1*, *MAPT*). На хромосоме 19 имелась пятерка генов (*TMEM160*, *FGF21*, *GYS1*, *LHB*, *CPT1C*).

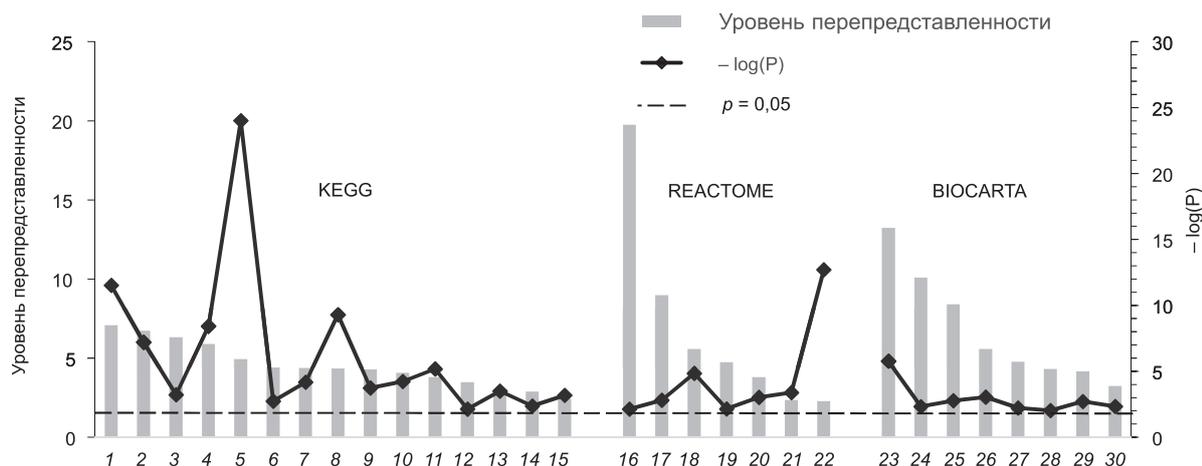
Таким образом, с использованием этого критерия выявлено семь групп близкорасположенных генов (их локализация выделена на рис. 4 овалами).

### Выявление генов, наименее устойчивых к мутациям, на основе величины RVIS

При ранжировании в соответствии со значениями RVIS были выбраны 12 генов с наиболее низким скором (табл. 2). Данные гены обладали значением  $RVIS < -2,13$ . Это, согласно S. Petrovski с соавт. (2013), означает, что эти гены попадают в число 1,49 % наименее толерантных к мутациям генов генома человека.



**Рис. 2.** Распределение генов из компиляции по основным категориям базы Gene Ontology (подраздел *Molecular functions*). Приведены GO термины и доли генов, имеющих в аннотации данные GO термины, от общего количества генов в компиляции.



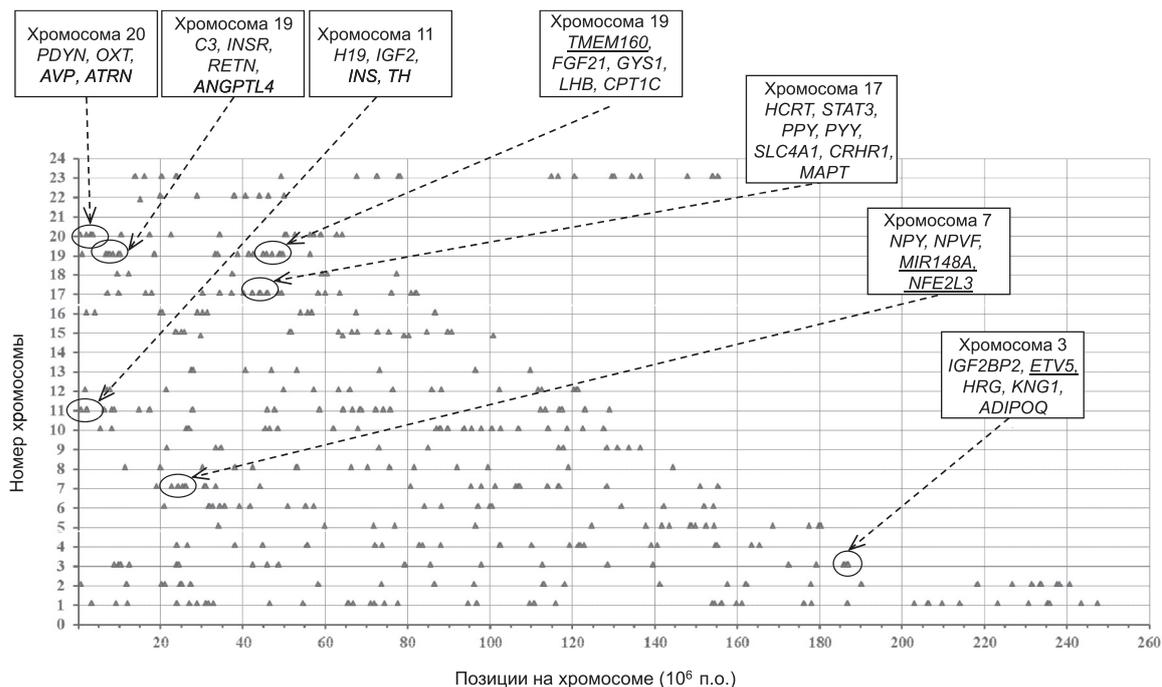
**Рис. 3.** Метаболические и сигнальные пути из баз KEGG, REACTOME и BIOCARTA, перепредставленные в аннотации генов из компиляции более чем в два раза и уровнем достоверности  $p \leq 10^{-2}$ .

1 – Adipocytokine signaling pathway; 2 – Type II diabetes mellitus; 3 – Maturity onset diabetes of the young; 4 – PPAR signaling pathway; 5 – Neuroactive ligand-receptor interaction; 6 – Aldosterone-regulated sodium reabsorption; 7 – NOD-like receptor signaling pathway; 8 – Insulin signaling pathway; 9 – Acute myeloid leukemia; 10 – Pancreatic cancer; 11 – Toll-like receptor signaling pathway; 12 – mTOR signaling pathway; 13 – Neurotrophin signaling pathway; 14 – Progesterone-mediated oocyte maturation; 15 – Calcium signaling pathway; 16 – Mitochondrial Uncoupling; 17 – Proteins Signal attenuation; 18 – Signaling by Insulin receptor; 19 – Regulation of beta-cell development; 20 – Hormone biosynthesis; 21 – Signalling by NGF; 22 – Signaling by GPCR; 23 – Visceral Fat Deposits and the Metabolic Syndrome; 24 – Reversal of Insulin Resistance by Leptin; 25 – Role of PPAR-gamma Coactivators in Obesity and Thermogenesis; 26 – IL-6 signaling pathway; 27 – Insulin Signaling Pathway; 28 – Role of ERBB2 in Signal Transduction and Oncology; 29 – Signal transduction through IL1R; 30 – Mechanism of Gene Regulation by Peroxisome Proliferators via PPARa(alpha).

**Таблица 2**

Двенадцать генов из компиляции, имеющих наиболее низкое значение RVIS, что указывает на низкую толерантность к мутациям в кодирующей части

Символ гена	Название гена	Источник данных	RVIS
<i>LRP1</i>	low density lipoprotein receptor-related protein 1	<i>OMIM-all text</i>	-7,28
<i>LRP5</i>	low density lipoprotein receptor-related protein 5	<i>OMIM-all text</i>	-3,72
<i>RAI1</i>	retinoic acid induced 1	<i>OMIM-all text</i>	-3,68
<i>FASN</i>	fatty acid synthase	<i>OMIM-all text</i>	-3,39
<i>LYST</i>	lysosomal trafficking regulator	<i>OMIM-all text</i>	-3,04
<i>RPTOR</i>	regulatory associated protein of MTOR, complex 1	<i>GWAS-мета-анализ</i>	-2,58
<i>DGKD</i>	diacylglycerol kinase, delta 130kDa	<i>OMIM-all text</i>	-2,34
<i>LRP1B</i>	low density lipoprotein receptor-related protein 1B	<i>GWAS-мета-анализ</i>	-2,29
<i>NCOA1</i>	nuclear receptor coactivator 1	<i>OMIM-all text</i>	-2,19
<i>ADCY3</i>	adenylate cyclase 3	<i>GWAS-мета-анализ</i>	-2,17
<i>ZNF608</i>	zinc finger protein 608	<i>GWAS-мета-анализ</i>	-2,16
<i>INSR</i>	insulin receptor	<i>OMIM-allelic variants, Публикации</i>	-2,14



**Рис. 4.** Локализация генов из компиляции на хромосомах человека. Гены обозначены треугольниками так, что их позиции по оси  $OX$  соответствуют координатам центров генов на хромосоме, а по оси  $OY$  – номеру хромосомы. Овалами выделены либо четыре отдельно расположенные района хромосом протяженностью  $\leq 2 \times 10^6$  п. о., включающие центральные позиции четырех генов из компиляции (хромосомы 7, 11, 19, 20), либо три участка (хромосомы 3, 17, 19), где такие районы длиной  $\leq 2 \times 10^6$  п. о. перекрываются, что позволяет выявить пять, семь и еще пять близкорасположенных генов соответственно. В прямоугольных выносках приведены названия генов. Гены, внесенные в компиляцию только на основании информационного источника *GWAS-мета-анализ*, подчеркнуты.

## ЗАКЛЮЧЕНИЕ

В данной работе представлена компиляция генов, регулирующих пищевое поведение и массу тела, сформированная с использованием трех информационных источников (см. табл. 1). Выявлены четыре гена (*MC4R*, *PCSK*, *POMC*, *BDNF*) с высоким приоритетом, которые содержались во всех трех информационных источниках (см. рис. 1). Белки, кодируемые генами *MC4R*, *PCSK*, *POMC*, экспрессируются в аркуатных ядрах гипоталамуса и выполняют ключевую функцию в системе передачи сигнала насыщения. Ген *BDNF* кодирует нейротрофический фактор мозга, активация которого в вентромедиальных ядрах гипоталамуса снижает аппетит (Yeo, Heisler, 2012).

Значительную долю генов в компиляции составляли гены, кодирующие рецепторы клеточной поверхности, и белки, способные связываться с рецепторами, а также белки с ки-

назой активностью (см. рис. 2). Это наблюдение отражает тот факт, что пищевое поведение контролируется обширной системой нейронов, взаимодействующих между собой посредством сигнальных веществ (нейропептидов, гормонов и т. п.), которые связываются с рецепторами на поверхности клетки и передают сигналы в цитоплазму посредством каскадов сигнальной трансдукции (Olszewski *et al.*, 2008). В компиляции выявлена также значительная доля генов с регуляторными функциями, а именно, транскрипционных регуляторов (63 гена, т. е. 15 % выборки). Известно, что каждый транскрипционный регулятор контролирует активность большой каскады генов (Merkulova *et al.*, 2013), поэтому данная группа генов перспективна для выбора потенциальных мишеней фармакологического воздействия, управляя активностью которых можно препятствовать развитию патологии.

Гены из компиляции задействованы в большом разнообразии метаболических и сигнальных путей, что отражает сложную природу механизмов регуляции массы тела (см. рис. 3). Мы рассматриваем выявленные пути как возможные точки приложения эффектов фармакологических препаратов, которые могут быть разработаны в будущем с целью коррекции массы тела.

При исследовании распределения генов в геноме обнаружены участки, содержащие близкорасположенные гены (рис. 4). Можно предполагать, что в некоторых из описанных нами случаев близкое расположение генов подразумевает наличие координированной регуляции экспрессии. Координированная регуляция может обеспечиваться за счет формирования соответствующей 3D структуры хроматина, функционирования так называемых транскрипционных фабрик, а также районов хроматина с барьерной функцией (Razin *et al.*, 2011; Wang *et al.*, 2012). Проверка этой гипотезы требует более детального анализа с привлечением дополнительных экспериментальных данных.

В пределах трех из выявленных нами районов близкорасположенных генов находились гены (*ETV5*, *MIR148A*, *NFE2L3*, *TMEM160*), внесенные в компиляцию только на основе информационного источника *GWAS-мета-анализ* (подчеркнутые символы генов на рис. 4). Данное наблюдение поможет сформировать гипотезы относительно механизмов функционирования генов, обнаруженных в экспериментах *GWAS*, роль которых в регуляции массы тела охарактеризована пока недостаточно.

Исследование другой геномной характеристики, сора *RVIS* (Petrovski *et al.*, 2013), позволило выявить гены с наименьшей толерантностью к мутациям (см. табл. 2), что указывает на повышенный риск развития патологий, в частности ожирения, в случае обнаружения мутаций в этих генах. Среди двенадцати генов с самым низким скором (*RVIS* < -2,13) обнаружены два гена, кодирующие белки с транскрипционными регуляторными функциями (*RAI1*, *ZNF608*). Наше наблюдение создает мотивацию к дальнейшим исследованиям генов-мишеней транскрипционных факторов *RAI1* и *ZNF608* на основе технологий *Chip-Seq* и т. п.

Выполненные в настоящей работе систематизация данных о генах, регулирующих пищевое поведение и массу тела, а также их анализ позволят расширить знания о механизмах регуляции этих фенотипических признаков. Результаты работы могут быть полезны при разработке более эффективных подходов персонализированной медицины на всех этапах, включая профилактику (генотипирование с целью оценки риска патологии), диагностику (выявление генов, несущих повреждающие мутации), а также выбор средств фармакологического воздействия.

## БЛАГОДАРНОСТИ

Работа выполнена при поддержке РФФ (проект № 14-24-00123).

## ЛИТЕРАТУРА

- Blakemore A.I., Froguel P. Investigation of Mendelian forms of obesity holds out the prospect of personalized medicine // *Ann. N. Y. Acad. Sci.* 2010. V. 1214. P. 180–189.
- Choi Y., Sims G.E., Murphy S. *et al.* Predicting the functional effect of amino acid substitutions and indels // *PLoS One.* 2012. V. 7. No. 10. P. e46688.
- Herrera B.M., Keildson S., Lindgren C.M. Genetics and epigenetics of obesity // *Maturitas.* 2011. V. 69. No. 1. P. 41–49.
- Huang da W. *et al.* Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources // *Nat. Protoc.* 2009. V. 4. No. 1. P. 44–57.
- Maniam J., Morris M.J. The link between stress and feeding behavior // *Neuropharmacology.* 2012. V. 63. No. 1. P. 97–110.
- Masoudi-Nejad A., Meshkin A., Haji-Eghrari B., Bidkhorri G. Candidate gene prioritization // *Mol. Genet. Genomics.* 2012. V. 287. No. 9. P. 679–698.
- Merkulova T.I., Ananko E.A., Ignat'eva E.V., Kolchanov N.A. Regulatory transcription codes in eukaryotic genomes // *Genetika.* 2013. V. 49. No. 1. P. 37–54.
- Olszewski P.K., Cedernaes J., Olsson F. *et al.* Analysis of the network of feeding neuroregulators using the Allen Brain Atlas // *Neurosci. Biobehav. Rev.* 2008. V. 32. No. 5. P. 945–956.
- Oshchepkov D.Y., Vityaev E.E., Grigorovich D.A., Ignat'eva E.V., Khlebodarova T.M. SITECON: a tool for detecting conservative conformational and physicochemical properties in transcription factor binding site alignments and for site recognition // *Nucleic Acids Res.* 2004. V. 32. P. W208–W212.
- Petrovski S., Wang Q., Heinzen E.L., Allen A.S., Goldstein D.B. Genic Intolerance to Functional Variation and the Interpretation of Personal Genomes // *PLoS Genet.* 2013. V. 9. No. 8. P. e1003709.

- Ponomarenko J.V., Merkulova T.I., Vasiliev G.V. *et al.* rSNP\_Guide, a database system for analysis of transcription factor binding to target sequences: application to SNPs and site-directed mutations // *Nucleic Acids Res.* 2001. V. 29. No. 1. P. 312–316.
- Razin S.V., Gavrilov A.A., Pichugin A. *et al.* Transcription factories in the context of the nuclear and genome organization // *Nucleic Acids Res.* 2011. V. 39. No. 21. P. 9085–9092.
- Smedley D., Köhler S., Czeschik J.C. *et al.* Walking the interactome for candidate prioritization in exome sequencing studies of Mendelian diseases // *Bioinformatics.* 2014. V. 30. pii: btu508 ah.
- Wang J., Lunyak V.V., Jordan I.K. Genome-wide prediction and analysis of human chromatin boundary elements // *Nucleic Acids Res.* 2012. V. 40. No. 2. P. 51–529.
- Yeo G.S., Heisler L.K. Unraveling the brain regulation of appetite: lessons from genetics // *Nat. Neurosci.* 2012. V. 15. No. 10. P. 1343–1349.
- Zegers D., Van Hul W., Van Gaal L.F., Beckers S. Monogenic and complex forms of obesity: insights from genetics reveal the leptin-melanocortin signaling pathway as a common player // *Crit. Rev. Eukaryot. Gene Expr.* 2012. V. 22. No. 4. P. 325–343.

## HUMAN GENES CONTROLLING FEEDING BEHAVIOR OR BODY MASS AND THEIR FUNCTIONAL AND GENOMIC CHARACTERISTICS: A REVIEW

E.V. Ignatieva<sup>1-3</sup>, D.A. Afonnikov<sup>1-3</sup>, E.I. Rogaev<sup>2</sup>, N.A. Kolchanov<sup>1,3</sup>

<sup>1</sup> Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia,  
e-mail: eignat@bionet.nsc.ru;

<sup>2</sup> Center for Brain Neurobiology and Neurogenetics, Novosibirsk, Russia;

<sup>3</sup> Novosibirsk National Research State University, Novosibirsk, Russia

### Summary

The goals of this study were to create a compilation of genes controlling human body weight and feeding behavior and to summarize functional and genomic information on these genes. Information on 424 human genes was obtained from scientific publications, OMIM and meta-analysis of GWAS data. Four genes (*BDNF*, *MC4R*, *PCSK1*, and *POMC*) were confirmed by all three data sources; thus, these genes have the highest priority (No. 1). Genes of other two groups (3 and 29 genes) were confirmed by two of three data sources; thus having priority No. 2. Pathways important for body mass regulation were revealed, and they may be candidate pharmacological targets for obesity treatment. Regions of human chromosomes containing closely located genes from the compilation were revealed. Some groups of closely located genes included genes (*ETV5*, *MIR148A*, *NFE2L3*, and *TMEM160*) confirmed by GWAS meta-analysis only. This finding may be helpful in the identification of their functions. Use of Residual Variation Intolerance Score (RVIS) revealed genes with decreased tolerance to functional genetic variation: *LRP1*, *LRP5*, *RAI1*, *FASN*, *LYST*, *RPTOR*, *DGKD*, *LRP1B*, *NCOA1*, and *ADCY3*. The compilation can be used in genotyping for pathology risk estimation and for designing new pharmacological approaches for treatment of human obesity.

**Key words:** feeding behavior, genomic location, regulation of body mass, tolerance to functional genetic variation.