

РОССИЙСКАЯ АКАДЕМИЯ НАУК
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ ЦИТОЛОГИИ И ГЕНЕТИКИ СИБИРСКОГО ОТДЕЛЕНИЯ РОССИЙСКОЙ
АКАДЕМИИ НАУК
ИЦиГ СО РАН

УДК 577.21

№ госрегистрации 01201058864

УТВЕРЖДАЮ
Директор
академик РАН Н. А. Колчанов

(подпись)

“15” мая 2012 г.

М.П.

ОТЧЕТ

О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

В рамках федеральной целевой программы «Научные и научно-педагогические кадры инновационной России» на 2009-2013 годы

Проведение научных исследований коллективами научно-образовательных центров в области физико-химической молекулярной и клеточной биологии

по теме:

**МОЛЕКУЛЯРНЫЕ МЕХАНИЗМЫ СТРЕСС-СПЕЦИФИЧЕСКОГО КОНТРОЛЯ
ЭКСПРЕССИИ ГЕНОВ ЖИВОТНЫХ**

(Государственный контракт № 02.740.11.0705 от 5 апреля 2010 г., шифр заявки «2010-1.1-142-043-002»)

(промежуточный, этап № 5)

Наименование этапа: «Кодирующий потенциал мРНК генов стрессового ответа»

Руководитель проекта

д.б.н., проф.

_____ Г. М. Дымшиц

подпись, дата

Новосибирск 2012

СПИСОК ИСПОЛНИТЕЛЕЙ

Научный руководитель, д.б.н., проф.	_____ подпись, дата	Дымицкий Г. М. (раздел 1, введение, заключение)
Исполнители		
зав. отд. молекулярной генетики, д.б.н., проф.	_____ подпись, дата	Меркулова Т.И. (раздел 1)
зав. лаб. генетики стресса, д.б.н., проф.	_____ подпись, дата	Раушенбах И. Ю. (раздел 3)
зав. лаб. генной инженерии, к.б.н., доц.	_____ подпись, дата	Кочетов А. В. (раздел 2)
д.б.н., с.н.с.	_____ подпись, дата	Грунтенко Н.Е. (раздел 3)
к.б.н., с.н.с.	_____ подпись, дата	Васильев Г.В. (разделы 1)
к.б.н., с.н.с.	_____ подпись, дата	Левицкий В.Г. (раздел 1)
к.б.н., н.с.	_____ подпись, дата	Антонцева Е. А. (раздел 1)
к.б.н., н.с.	_____ подпись, дата	Ощепков Д. Ю. (раздел 1)
к.б.н., н.с.	_____ подпись, дата	Богомолова Е. В. (раздел 3)
к.б.н., с.н.с.	_____ подпись, дата	Адоньева Н.В. (раздел 3)
к.б.н., н.с.	_____ подпись, дата	Лаухина О.В. (раздел 3)
к.б.н., н.с.	_____ подпись, дата	Карпова Е.К. (раздел 3)
к.б.н., н.с.	_____ подпись, дата	Иванов М. К. (раздел 1)
м.н.с.	_____ подпись, дата	Глушков С.А. (раздел 1)
к.б.н., м.н.с.	_____ подпись, дата	Волкова О.А. (раздел 2)
н.с. к.б.н.	_____ подпись, дата	Брызгалов Л.О. (раздел 1)

<i>н.с. к.б.н.</i>	_____	<i>Пахарукова М. Ю.</i>
	подпись, дата	(раздел 1)
<i>н.с. к.б.н.</i>	_____	<i>Климова Н. В.</i>
	подпись, дата	(раздел 1)
<i>н.с. к.б.н.</i>	_____	<i>Брагин А. Г.</i>
	подпись, дата	(раздел 2)
<i>н.с.</i>	_____	<i>Федосеева Л.А.</i>
	подпись, дата	(раздел 1)
<i>м.н.с. к.б.н.</i>	_____	<i>Сметанина М.А.</i>
	подпись, дата	(раздел 1)
<i>м.н.с. к.б.н.</i>	_____	<i>Герасимова С.В.</i>
	подпись, дата	(раздел 2)
<i>н.с. к.б.н.</i>	_____	<i>Еришов Н.И.</i>
	подпись, дата	(раздел 1)
<i>аспирант</i>	_____	<i>Иванисенко Т.В.</i>
	подпись, дата	(раздел 2)
<i>аспирант</i>	_____	<i>Горелова В.В.</i>
	подпись, дата	(раздел 2)
<i>аспирант</i>	_____	<i>Ермаков А.А.</i>
	подпись, дата	(раздел 2)
<i>аспирант</i>	_____	<i>Брагин А.О.</i>
	подпись, дата	(раздел 2)
<i>студент</i>	_____	<i>Никулин П.С.</i>
	подпись, дата	(раздел 2)
<i>студент</i>	_____	<i>Аркова О.В.</i>
	подпись, дата	(раздел 2)
<i>студент</i>	_____	<i>Сайк О.В.</i>
	подпись, дата	(раздел 2)
<i>студент</i>	_____	<i>Медведев К.Е.</i>
	подпись, дата	(раздел 3)
<i>студент</i>	_____	<i>Емцева М.В.</i>
	подпись, дата	(раздел 3)
<i>Нормоконтролер</i>	_____	<i>Амайзер Н.Н.</i>
	подпись, дата	

РЕФЕРАТ

Отчет 59 с., 13 рис., 4 табл., 67 источников.

СИГНАЛЫ ЭКСПРЕССИИ, СТРЕСС, ТРАНСЛЯЦИЯ, мРНК, ИММУНОПРЕЦИПИТАЦИЯ ХРОМАТИНА, SOLID, МЕТОД CHIP-SEQ, FOXA, ЮВЕНИЛЬНЫЙ ГОРМОН, Д2-РЕЦЕПТОРЫ, 20-ГИДРОКСИЭКДИЗОН, ДОФАМИН

Целью проекта является выявление молекулярных механизмов, контролирующих экспрессию генов животных на транскрипционном и посттранскрипционном уровне в норме и в условиях стресса, а также исследование генетических моделей патологических состояний. В рамках проекта планируется использовать современные методы количественного анализа паттернов экспрессии генов эукариот, а также ряд уникальных генетических моделей. За отчетный период были получены следующие результаты:

1. Проведён компьютерный анализ ранее полученных данных по иммунопреципитации хроматина *in vivo* (ChIP-Seq) с антителами против транскрипционных факторов FoxO3A, FoxA2, связанных с ответом на ксенобиотики и онкогенезом. Обработаны данные контрольных экспериментов с использованием антител против суммарного IgG мыши и гистона H3. С помощью экспериментальной верификации проведена оптимизация параметров четырёх программ распознавания сайтов связывания FoxA2 - PWM, SiteGA, MonoChipmunk и DiChipmunk. Такая оптимизация позволила провести высокодостоверный анализ локализации сайтов связывания FoxA2 внутри пиков ChIP-Seq, а также подробно изучить профиль полногеномного распределения сайтов связывания транскрипционного фактора FoxO3A для взрослых и новорождённых особей мышей.
2. Проведено исследование роли механизма реинициации трансляции в синтезе альтернативных изоформ белков человека. Обнаружено, что при определенной конфигурации 5'-НТП, содержащей лидерные рамки считывания, с мРНК человека с высокой степенью вероятности могут дополнительно считываться дополнительные N-концевые изоформы белков. По сравнению с аннотированными полноразмерными вариантами, эти изоформы часто характеризуются иной субклеточной локализацией, а также отсутствием сайтов посттрансляционных модификаций, что свидетельствует об их функциональных различиях. Известно, что эффективность реинициации трансляции зависит от статуса фосфорилирования фактора инициации трансляции eIF2 α , который – в свою очередь – изменяется в ответ на стресс. В рамках этой части работы было предсказано около 300 генов человека, с мРНК которых могут дополнительно считываться укороченные с N-конца изоформы

аннотированных белков, причем эффективность их синтеза может изменяться в условиях стресса.

3. В отчетный период (1) осуществлен иммуногистохимический анализ *corpus allatum* молодых и половозрелых самок дикого типа *D. melanogaster* с использованием кроличьих антител против DD2R; (2) осуществлен иммуногистохимический анализ жирового тела молодых и половозрелых самок *D. melanogaster* дикого типа и с генетической абляцией части клеток *corpus allatum*; (3) исследовано влияние 20Э на экспрессию гена *DD2R*. Впервые показано, что онтогенетический характер ингибирующего влияния ДА на синтез и деградацию ЮГ у *Drosophila* обусловлен переключением на геномном уровне, вызывающем изменение числа дофаминовых Д2-подобных рецепторов в железе, синтезирующей ЮГ, и ткани, синтезирующей ферменты его деградации; Обнаружено, что это переключение обусловлено изменением титра 20-гидроксиэкдизона после оплодотворения самок.

Полученные результаты используются в учебном процессе (курсы «Основные молекулярно-генетические процессы», «Новейшие молекулярно-генетические технологии»). За период пятого этапа защищены три диссертации на степень кандидата биологических наук и опубликовано 4 статьи в рецензируемых журналах.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	7
ОСНОВНАЯ ЧАСТЬ	14
1. Сравнительное изучение паттернов полногеномного распределения транскрипционных факторов семейства FOX в печени экспериментальных животных в норме и под действием гепатоканцерогенных соединений.....	27
2. Исследование сайтов альтернативной инициации трансляции и их вклада в кодирующий потенциал мРНК генов стрессового ответа	40
3. Изучение онтогенетического паттерна экспрессии гена DD2R в тканях, метаболизирующих ювенильный гормон, у самок дрозофилы дикого типа и с генетической абляцией части клеток железы (<i>corpus allatum</i>), в которой гормон синтезируется	47
ЗАКЛЮЧЕНИЕ.....	52
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	54

ВВЕДЕНИЕ

Цель проекта заключается в выявлении молекулярных механизмов, контролирующих экспрессию генов животных на транскрипционном и посттранскрипционном уровне в норме и в условиях стресса. Для выполнения этой работы планируется использовать современные методические подходы и перспективные генетические модели, что позволит исследовать регуляцию экспрессии генов, связанных с развитием социально-значимых патологий (канцерогенез, гипертоническая болезнь).

В настоящее время методы и подходы молекулярной биологии активно используются для исследования фундаментальных аспектов организации сложных биологических систем. Одна из ключевых задач заключается в реконструкции генных сетей и регуляторных контуров, определяющих адаптацию организма животных к стрессовым условиям различной природы: эта информация необходима не только для развития существующих представлений о молекулярно-генетических механизмах адаптации к неблагоприятным факторам, но и для получения новых данных о причинах канцерогенеза и некоторых заболеваний, связанных с дисфункцией регуляторных систем вследствие мутаций или воздействия ксенобиотиков. Эти исследования относятся к актуальным направлениям развития молекулярной биологии и необходимы для разработки новых методов диагностики и лечения патологических состояний человека. В то же время, стресс-специфическая экспрессия генов представляет собой модель, изучение которой может дать новую ценную информацию фундаментального характера. В рамках проекта планируется исследовать следующие модели:

(1) Транскрипционный контроль экспрессии генов печени экспериментальных животных в норме и при воздействии онкогенных ксенобиотиков. Печень представляет собой ключевой орган детоксификации ксенобиотиков. Система генов, контролирующая детоксификацию, сложна и изучена недостаточно полно. Кроме того, многие ее компоненты одновременно являются ключевыми регуляторами базовых клеточных процессов, таких как дифференцировка, пролиферация и апоптоз. Вследствие этого, происходящие при взаимодействии с определенными веществами сбои в регуляторных системах клеток печени могут приводить к развитию онкологических заболеваний. Планируется провести комплексное исследование влияния ряда гепатоканцерогенных соединений на транскриптом печени и полногеномное распределение важных для ее функции факторов транскрипции, что позволит очертить круг генов, ответственных за развитие опухолей печени под действием этих соединений и построить схему регуляторных событий, ведущих к данной патологии.

На предыдущих этапах проекта нами были отработаны методики проведения экспериментов по иммунопреципитации хроматина с последующим секвенированием пула выделенных ДНК на системе SOLiD, а также разработан метод предсказания сайтов связывания

транскрипционных факторов группы FOX. Кроме этого, были отработаны процедуры выделения препаратов РНК для микрочипового анализа транскриптома и препаратов ядерных белковых экстрактов. На втором этапе проекта нами была проведена иммунопреципитация хроматина семенников самцов крыс с антителами против транскрипционного фактора SF1. Создан задел для исследования этих образцов методом ChIP-Seq. С использованием биочипов RatRef-12 BeadChip (Illumina), содержащих 22523 уникальные олигонуклеотидные пробы, проведен анализ изменений транскриптома печени крысы при введении животным двух аминоазосоединений: гепатоканцерогенного для крыс 3'-МеДАБ и неканцерогенного для них ОАТ. Выявлено 1052 и 442 гена, уровень экспрессии которых в печени достоверно изменялся более чем в 2 раза в ответ на введение 3'-МеДАБ и ОАТ, соответственно. При этом экспрессия 408 генов изменялась преимущественно в ответ на канцерогенный 3'-МеДАБ, но не на ОАТ, 74 гена демонстрировали избирательную реакцию на ОАТ. С использованием базы Gene Ontology показано, что около 20% генов, избирательно реагирующих на гепатоканцерогенный 3'-Ме-ДАБ, представлено генами, связанными с клеточным циклом и апоптозом. На третьем этапе проекта нами была проведена иммунопреципитация хроматина ткани печени самцов крыс двухмесячного возраста с обработкой канцерогенами и без, а так же новорождённых с использованием антител против транскрипционных факторов FoxO3A, FoxA2, HNF4. Получены и отсекувенированы (SOLiD) специфические библиотеки.

На четвертом этапе проекта для исследования транскрипционного контроля генов, вовлеченных в онкогенез, проведена иммунопреципитация хроматина тканей печени, почки и надпочечников самцов крыс шестимесячного возраста с использованием антител против транскрипционных факторов SF1 и LRH. Из данного материала получены специфические библиотеки для системы SOLiD. Качество полученного материала проверено тестовым лигированием – «WFA» на системе SOLiD, проводится секвенирование библиотек на системе SOLiD. Изучалась связь между гепатоканцерогенностью азосоединений ОАТ и 3'-МеДАБ для мышей линии GR и одним из ранних событий воздействия этих факторов – их влиянием на ДНК-связывающую активность транскрипционного фактора FoxA. При введении мышам однократно в возрасте 12 дней 3'-МеДАБ вызывал опухоли печени у самцов и самок в 100% случаев, наравне с известным мышинным гепатоканцерогеном ОАТ. При этом ОАТ снижал ДНК-связывающую активность FOXA в печени в 2,5-3 раза и, как следствие, снижал на 40% уровень глюкокортикоидной индукции печень-специфического гена TAT. В отличие от этого, 3'-МеДАБ не изменял ни активность FOXA белков, ни степень глюкокортикоидной индукции TAT.

(2) Изучение механизмов транскрипционного контроля экспрессии генов ренин-ангиотензин-альдостероновой системы крысы в норме и при стрессе. Эта часть

работы связана с исследованием генетического контроля функционирования одной из важнейших физиологических систем, нарушения в которой приводят к развитию гипертонической болезни. Для проведения исследований планируется использовать уникальную генетическую модель, разработанную в ИЦиГ СО РАН – лабораторную линию крыс НИСАГ, характеризующихся увеличенным кровяным давлением и наследственной склонностью к развитию гипертонии в ответ на стресс. Исследования на молекулярно-биологическом уровне позволят выявить регуляторные механизмы, нарушения в которых отвечают за развитие этой патологии. Планируется провести изучение реакции на стресс ключевых генов симпато-адреналовой и ренин-ангиотензиновой систем у крыс гипертензивной линии НИСАГ разных возрастов в почках, сердце и различных отделах мозга. Полученные данные будут использованы для развития представлений о механизмах возникновения артериальной гипертонии, связанной со стрессом.

На предыдущих этапах проекта нами были отработаны некоторые методы и процедуры для проведения высокочувствительного анализа уровня транскрипции индивидуальных генов: для этого были разработаны процедуры измерения 5'-эндо/экзонуклеазной активности препаратов ДНК-полимеразы Taq. Проведен анализ эффективности использования различных вариантов олигонуклеотидов со структурными модификациями. На втором этапе проекта был разработан метод характеристики флуоресцентно-меченных олигонуклеотидных зондов, использование которого позволяет проводить оценку степени мечения олигонуклеотидов флуорофорами и гасителями флуоресценции. Это позволило увеличить чувствительность и воспроизводимость количественной ПЦР в реальном времени за счёт использования охарактеризованных препаратов флуоресцентно-меченных олигонуклеотидов и ДНК- и РНК-зависимых ДНК-полимераз, оптимизации состава реакционной смеси и «горячего старта». На третьем этапе проекта нами была изучена экспрессия генов ренин-ангиотензиновой системы (РАС), кодирующих ангиотензиноген, ренин, ангиотензин-превращающий фермент, рецепторы ангиотензина II первого и второго типа и рецептор ренина/проренина, которую определяли в тканях почки, надпочечников, гипоталамуса и продолговатого мозга у молодых (1,5 мес) и взрослых (4-7 мес) крыс гипертензивной линии НИСАГ и нормотензивной линии WAG. Сделано заключение о том, что на формирование гипертензивного статуса крыс линии НИСАГ может влиять состояние РАС в первые недели их жизни, у взрослых животных за поддержание гипертонии отвечают другие системы.

На четвертом этапе проекта была исследована экспрессия ключевого гена биосинтеза катехоламинов тирозин-гидроксилазы (ТН) в тканях надпочечников, гипоталамуса и продолговатого мозга у молодых (1.5 мес) и взрослых (4-7 мес) крыс гипертензивной линии

НИСАГ в сравнении с нормотензивной линией WAG. Мы не выявили различий в базальном уровне мРНК *Th* в надпочечниках крыс WAG и НИСАГ в возрасте 1.5 и 4 месяца, но у крыс в возрасте 7 месяцев наблюдалось небольшое (23%), но статистически достоверное снижение уровня мРНК *Th* в надпочечниках крыс НИСАГ по сравнению с крысами WAG. Повышение экспрессии гена *Th* в этом органе у обеих линий крыс при стрессе водной депривации не достигало статистически достоверного уровня. В гипоталамусе, но не в продолговатом мозге, у молодых крыс НИСАГ уровень мРНК *Th* оказался втрое повышен по сравнению с таковым у нормотензивных крыс WAG. К возрасту 4 месяцев повышенный вдвое уровень мРНК *Th* наблюдался в продолговатом мозге крыс НИСАГ, однако таких различий в гипоталамусе не было. Необычно, что стресс водной депривации вызвал повышение содержания мРНК *Th* в гипоталамусе крыс НИСАГ втрое, в то время как у крыс WAG этот показатель вообще не изменилось. Это наблюдение подтверждает повышенную реактивность симпатической нервной системы крыс НИСАГ. К возрасту 7 мес содержание мРНК *Th* в мозговых структурах крыс НИСАГ, как и в надпочечниках, оказывается даже ниже, чем у крыс WAG. Сделано заключение о том, что на формирование гипертензивного статуса крыс НИСАГ может влиять состояние ренин-ангиотензиновой системы в первые недели их жизни, у взрослых животных за поддержание высокого уровня артериального давления отвечают другие системы.

(3) Изучение сигналов, локализованных в мРНК и контролирующих общую и стресс-специфическую трансляционную активность матриц. В рамках проекта планируется провести исследования молекулярных механизмов стрессового ответа, действующих не только на уровне контроля транскрипции, но и на уровне трансляции мРНК. Известно, что в условиях различных видов стресса происходит приостановка трансляции основного пула клеточных мРНК и экспрессируются преимущественно стресс-специфические гены. Сигналы, определяющие такую селективную трансляцию, в настоящее время малоизучены. Планируется провести исследование, направленное на выявление характеристик эукариотических мРНК, связанных с селективной трансляцией в стрессовых условиях. Знание таких характеристик позволит предсказывать мРНК, трансляция которых связана с адаптацией к неблагоприятным условиям, что – в свою очередь – будет способствовать выявлению соответствующих генов.

На предыдущих этапах проекта нами были созданы репрезентативные выборки негомологичных нуклеотидных последовательностей функциональных районов мРНК эукариотических генов. На втором этапе был проведен компьютерный анализ структурно-функциональной организации эукариотических матриц. Впервые исследована структурная характеристика мРНК – размер поверхности молекулы, доступной для взаимодействия (accessible surface area, ASA), отражающая способность участков молекулы к

взаимодействию с другими макромолекулами. Обнаружено, что в молекуле РНК параметр ASA контекстно-зависим. Впервые показано, что альтернативные стартовые кодоны характеризуются эволюционной консервативностью, что свидетельствует об их функциональной значимости. Проведен компьютерный анализ мРНК *Mus musculus*, характеризующихся различной эффективностью трансляции в норме и при стрессе. Выдвинуто предположение о том, что реинициация трансляции, характеристики которой зависят от статуса фосфорилирования фактора eIF2 α , является одним из ключевых механизмов стресс-специфического контроля экспрессии генов млекопитающих;

(4) Исследование роли контроля экспрессии генов дофаминовых (ДА) рецепторов в поддержании гормонального статуса в нормальных и стрессирующих условиях. Установлено, что D2-подобный рецептор (D2R) у млекопитающих является ключевым модулятором двигательной функции. У человека D2R оказался важнейшей терапевтической мишенью для лечения брадикинезии, возникающей в результате болезни Паркинсона. Однако пока практически ничего не известно о том какую роль контроль экспрессии генов рецепторов дофамина играет в регуляции гормонального статуса организма в стрессирующих условиях. Проведение подобных исследований на человеке невозможно, так как требует фармакологического изменения уровней стресс-связанных гормонов и проведения генетических экспериментов. Поскольку показано, что ДА-подобные рецепторы высоко-консервативны у млекопитающих и дрозофилы, в рамках настоящего проекта использована созданная в ИЦиГ СО РАН уникальная модель (серия линий дрозофилы с мутациями, вызывающими изменения в уровне различных стресс-связанных гормонов, в том числе дофамина), на которой установлен механизм взаиморегуляции гормонов стресса в контроле репродуктивной функции и устойчивости к действию стрессоров. Эта модель высокоперспективна еще и потому, что соответствующие регуляторные контуры у насекомых устроены значительно проще, чем у млекопитающих, и это позволяет реконструировать их организацию и эволюцию. Выявленные закономерности можно будет использовать как модель для коррекции гормонального дисбаланса, возникающего при стрессе у млекопитающих.

На предыдущих этапах проекта нами было показано, что ксенобиотическая активация или инактивация D2-подобных рецепторов изменяют интенсивность ответа щелочной фосфатазы на стресс: стресс-реактивность фермента снижена у самок, кормленных бромокриптином (агонистом D2-подобных рецепторов дофамина), и повышена у самок, кормленных метаклопрамидом (антагонистом D2-подобных рецепторов дофамина). Щелочная фосфатаза регулирует содержание предшественника дофамина – тирозина. Таким образом, показано, что изменение функциональной активности D2-подобных рецепторов влияет на метаболизм

дофамина. Также было показано, что негативная регуляция активности щелочной фосфатазы дофамином опосредуется D2-подобными рецепторами: снижение числа D2-подобных рецепторов повышает активность фермента. На основе данных, представленных в выводе 5 и 6 можно заключить, что нами впервые показано, что D2-подобные рецепторы дофамина играют существенную роль в развитии стресс-реакции насекомых. На втором этапе проекта изучено влияние ксенобиотической активации D2-подобных рецепторов на метаболизм ювенильного гормона в нормальных и стрессирующих условиях. Исследован метаболизм ювенильного гормона самок дрозофилы, у которых экспрессия гена *DDR2* специфически снижена в железе, производящей ювенильный гормон (с помощью тканеспецифичной РНК-интерференции на линиях трансгенных мух). Подтверждена гипотеза, выдвинутая нами ранее, согласно которой синтез и деградация ювенильного гормона у насекомых находятся под общим контролем и уровень деградации гормона может служить индикатором уровня его синтеза. Впервые установлено, что негативная регуляция синтеза ювенильного гормона дофамином опосредуется D2-подобными рецепторами. Обнаружено, что ксенобиотическая активация D2-подобных рецепторов вызывает снижение деградации ЮГ и этот эффект имеет онтогенетический характер. На третьем этапе проекта нами было обнаружено наличие D2-подобных рецепторов в жировом теле и *corpus allatum* и подтвержден сделанный нами на первом этапе реализации проекта вывод о том, что негативная регуляция синтеза ЮГ дофамином опосредуется D2-подобными рецепторами.

На четвертом этапе проекта был осуществлен иммуногистохимический анализ яичников половозрелых самок *D. melanogaster* с использованием кроличьих антител против рецептора дофамина 2R (*DD2R*) и измерен титр 20-гидроксиэкдизона (20Э) у самок дрозофилы со сниженной экспрессией гена *DD2R* в железе, синтезирующей ювенильный гормон. Впервые продемонстрировано, что в фолликулярных клетках яичников (месте синтеза 20Э) половозрелых самок дрозофилы отсутствуют D2-подобные рецепторы. Впервые показано, что у самок дрозофилы с повышенным в результате снижения числа *DD2R* в *corpus allatum* титром ЮГ, уровень 20Э резко повышен. Таким образом подтверждено, что дофамин регулирует уровень 20Э опосредованно, через систему метаболизма ЮГ.

Решение проблем, которым посвящен настоящий проект, необходимо для развития признанного правительством Российской Федерации приоритетного направления «Живые системы», а также для ряда критических технологий в рамках этого направления. Вовлечение студентов Новосибирского государственного университета и аспирантов Института цитологии и генетики СО РАН в исследования в рамках данного проекта и использование его результатов в преподавании биологических дисциплин, таких как молекулярная биология, новейшие методы молекулярно-генетических исследований и др. сыграет важную роль в

формировании у студентов целостного представления о молекулярных механизмах контроля экспрессии генов. Таким образом, в рамках НИР планируется проведение комплексного исследования молекулярных механизмов контроля экспрессии генов в стрессовых условиях, действующих на транскрипционном и трансляционном уровнях. Выполнение НИР должно обеспечивать достижение научных результатов мирового уровня, подготовку и закрепление в сфере науки и образования научных и научно-педагогических кадров, формирование эффективных и жизнеспособных научных коллективов.

Основная цель пятого этапа проекта — изучение кодирующего потенциала мРНК стрессового ответа. Задачи этапа включают следующие виды работ:

1. Сравнительное изучение паттернов полногеномного распределения транскрипционных факторов семейства FOX в печени экспериментальных животных в норме и под действием гепатоканцерогенных соединений.
2. Исследование сайтов альтернативной инициации трансляции и их вклада в кодирующий потенциал мРНК генов стрессового ответа
3. Изучение онтогенетического паттерна экспрессии гена DD2R в тканях, метаболизирующих ювенильный гормон, у самок дрозофилы дикого типа и с генетической абляцией части клеток железы (*corpus allatum*), в которой гормон синтезируется.

ОСНОВНАЯ ЧАСТЬ

Развитие и жизнедеятельность многоклеточных животных обеспечивается строго контролируемым и специфическим уровнем экспрессии генов в различных органах и тканях. Целью проекта является выявление молекулярных механизмов, регулирующих экспрессию генов животных на транскрипционном и посттранскрипционном уровне в норме и в условиях стресса, а также исследование генетических моделей патологических состояний. В рамках проекта планируется использовать современные методы количественного анализа паттернов экспрессии генов эукариот, а также ряд уникальных генетических моделей. В рамках отчетного (пятого) этапа главным образом была проведена работа по изучению кодирующего потенциала мРНК генов стрессового ответа, однако также была получена новая интересная информация по другим направлениям работы.

Исследование молекулярных основ канцерогенеза требует применения современных высокопроизводительных методов анализа геномов. В рамках проекта такие методы были применены для изучения форм рака, формирующихся при обработке лабораторных животных ксенобиотиками. При этом основное внимание было уделено влиянию канцерогенов на паттерн экспрессии регуляторных генов, связанных с развитием опухолей – и, в частности, на выявление сайтов связывания специфических транскрипционных факторов. Известно, что канцерогенез характеризуется нарушением баланса процессов пролиферации, апоптоза и дифференцировки. За поддержание такого баланса в высокодифференцированной клетке отвечают регуляторные белки - транскрипционные факторы, которые обеспечивают специфику протеома определенного типа клеток. Примером таких белков является семейство транскрипционных факторов FOXA. Белки этого семейства (FoxA1, FoxA2, FoxA3) экспрессируются в печени, поджелудочной железе и легких, способствуют декомпактизации хроматина и контролируют экспрессию множества генов-мишеней. В печени обнаружено более 100 генов-мишеней FoxA, продукты которых регулируют процессы дифференцировки, клеточного метаболизма, поддержания фенотипа клеток. Вовлеченность факторов FoxA в индукцию опухолей исследована нами на модели воздействия на организм различных гепатоканцерогенных веществ, прежде всего азосоединений, которые обладают выраженной тканевой, видовой и линейной специфичностью.

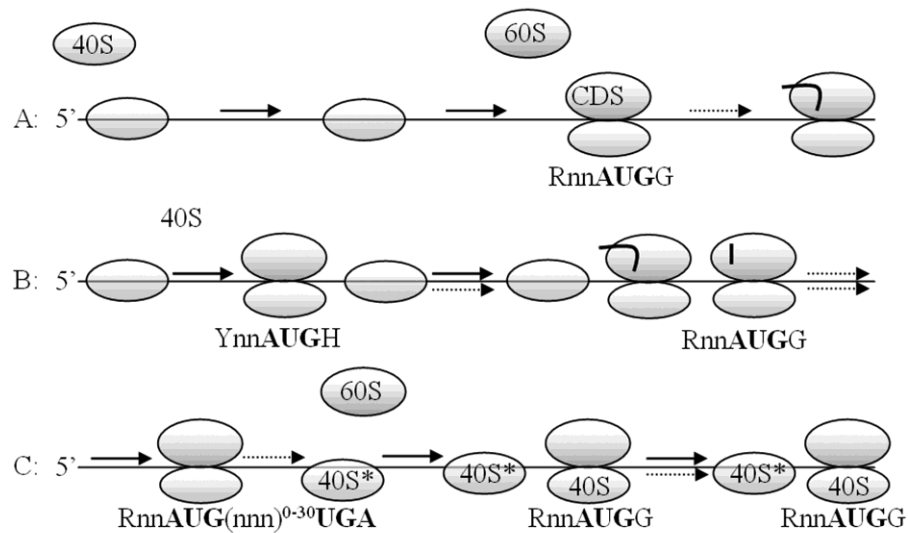
Точная локализация сайтов связывания транскрипционных факторов в регуляторных районах генов является ключевым пунктом для понимания особенностей регуляции процессов дифференцировки, пролиферации и апоптоза в тканях многоклеточного организма. Для изучения профилей распределения сайтов связывания транскрипционных факторов в масштабе всего генома, а не отдельных генов, применяются различные варианты метода иммунопреципитации хроматина *in vivo* (ChIP). Этот метод является уникальным по своим

возможностям инструментом для изучения взаимодействия между ДНК-связывающими белками (транскрипционными факторами, гистонами и др.) и ядерной ДНК, и, соответственно, для выявления реально занятых в данном типе клеток сайтов связывания. Наиболее совершенной на данный момент модификацией этого метода является ChIP-Seq. Суть метода: на первом этапе проводится обработка образца ткани или суспензии клеток формалином, приводящая к образованию ковалентных ДНК-белковых сшивок. Далее фиксированный ядерный хроматин выделяется и подвергается ультразвуковому дроблению на фрагменты с максимумом распределения по длинам в районе 250-500 п.н. На следующем этапе с помощью специфических антител к интересующему нас белку проводится селективная иммунопреципитация ДНК-белковых комплексов. Ковалентные связи ДНК-белок гидролизуются, ДНК из преципитированных комплексов выделяется и исследуется. Данные о полногеномном профиле распределения сайтов связывания получают путём секвенирования всего пула выделенной ДНК с помощью приборов массового параллельного секвенирования второго поколения – SOLiD фирмы ABI либо GA analyzer фирмы Illumina (метод ChIP-Seq). Метод ChIP-Seq обладает принципиальными преимуществами перед другими методами и другими вариантами иммунопреципитации хроматина. Во-первых, ChIP-Seq даёт многократно больший объём информации, во-вторых, результатом является картина полногеномного, а не только промоторного распределения сайтов связывания, в третьих, полученный результат свободен от предварительной селекции исходных данных, могущих существенно исказить конечный результат, что неизбежно в случае с микрочипами, несущими фиксированное число заранее выбранных последовательностей. Наконец, результатом ChIP-Seq являются не относительные уровни сигнала, как в случае микрочипов, а конкретные районы последовательностей ДНК, что даёт несравненно больше возможностей для теоретического анализа результатов эксперимента. В результате секвенирования всего пула иммунопреципитированной ДНК, наложения этой информации на референсную геномную последовательность и обработки этих данных специализированным программным обеспечением получают совокупность пиков покрытия, соответствующих районам взаимодействия ДНК-Белок. Завершающей стадией эксперимента является биоинформационный анализ полученных пиков, их достоверности, распределения и изменений в результате тех или иных регуляторных событий. На настоящий момент создано множество программ для анализа результатов ChIP-Seq экспериментов, преимущественно основанных на алгоритмах поиска *de novo* определённых мотивов в последовательностях (методы класса «pattern detection» (Kulakovskiy et al., 2010)). Однако целый ряд возникающих при таком анализе проблем может быть успешно разрешён с применением методов анализа,

основанных на другом принципе - распознавания сайтов связывания транскрипционных факторов «с обучением» (pattern matching).

Исследованием контроля экспрессии генов стрессового ответа на уровне трансляции мРНК. Общие представления. По-видимому, инициация трансляции большинства эукариотических мРНК происходит по механизму “линейного сканирования”, хотя в ряде случаев возможна инициация на внутреннем участке мРНК или шунтирование потока рибосом (Kozak, 2005; Jackson et al., 2010). Согласно модели линейного сканирования, 40S субъединица рибосомы связывается с кепом на 5'-конце мРНК и движется вдоль матрицы до инициаторного кодона AUG. Считается, что нуклеотиды в районе стартового кодона (контекст) важны для его взаимодействия с рибосомой: наибольший вклад вносят пурины в положении -3 и гуанин в положении +4 вокруг AUG (рис. 1 А). Если контекст триплета AUG не оптимален, то часть рибосом не распознает его как стартовый кодон и может начать трансляцию на следующем AUG (механизм leaky scanning или “сканирование с подтеканием”) (Kozak, 2005) (рис. 1 Б). Реинициация трансляции в эукариотических клетках, в отличие от прокариотических, малоэффективна. Показано, что после трансляции очень небольших ORF некоторые 40S субъединицы рибосом могут продолжить сканирование и повторно иницировать трансляцию на расположенной ниже dORF (рис. 1 С). Эффективность реинициации повышается также при увеличении размера участка, лежащего между стоп-кодоном uORF и следующим стартовым кодоном AUG. Предполагают, что это дает 40S субъединице рибосомы больше времени для связывания с недостающими факторами инициации трансляции (Jackson et al., 2010). Таким образом, трансляция с помощью механизма реинициации может осуществляться только при соблюдении определенных условий, ограничивающих размер uORF и ее расположение (рис. 1 D).

В 5'-концевом участке мРНК могут содержаться триплеты AUG, не принадлежащие к основной ORF. Каким образом можно интерпретировать присутствие uAUG в эукариотических мРНК? Если трансляция мРНК осуществляется по механизму линейного сканирования, то uAUG могут взаимодействовать с 40S субъединицами рибосом, на них может начинаться трансляция и с них могут считываться соответствующие белковые продукты. Если трансляция мРНК осуществляется по другому механизму (IRES или шунтирование потока рибосом), то рибосомы могут пропускать uAUG-содержащие участки. В настоящее время нет четких критериев, позволяющих предсказать механизм инициации трансляции по нуклеотидной последовательности мРНК, что не позволяет оценить масштаб использования этих альтернативных механизмов.



D:
 cggcgacagcucagagcagggagggccgcccaccucgcccggccggccggagcgggacgccccaggg
 cccccccccgggcaccccgcgucAUG¹. caa. cgc. cug. gug. gcc. ugg. gac. cca. gcA. UG
²u. cuc. ccc. cug. ccg. ccg. ccg. ccu. gcc. uuu. aaa. ucc. AUG³. gaa. gug. gc
 c. aac. uuc. uac. uac. gag. gcg. gac. ugc. uug. gcu. gcu. gcg. uac. gcc. ggc. aag
 . gcg. gcc. ccc. gcg. gcg. ccc. ccc. gcg. gcc. aga. ccc. ggg. ccg. cgc. ccc. ccc.
 gcc. ggc. gag. cug. ggc. agc. auc. ggc. gac. cac. gag. cgc. gcc. auc. gac. uuc. a
 gc. ccg. uac. cug. gag. ccg. cug. ggc. gcg. ccg. cag. gcc. ccg. gcg. ccc. gcc. ac
 g. gcc. acg. gac. acc. uuc. gag. gcg. gcu. ccg. ccc. gcg. ccc. gcc. ccc. gcg. ccc
 . gcc. ucc. ucc. ggg. cag. cac. cac. gac. uuc. cuc. ucc. gac. cuc. uuc. ucc. gac.
 gac. uac. ggg. ggc. aag. aac. ugc. aag. aag. ccg. gcc. gag. uac. gcc. uac. gug. a
 gc. cug. ggg. cgc. cug. ggg. gcc. gcc. aag. ggc. gcg. cug. cac. ccc. ggc. ugc. uu
 c. gcg. ccc. cug. cac. cca. ccg. ccc. ccg. ccg. ccg. ccg. ccg. ccc. gcc. gag. cuc
 . aag. gcg. gag. ccg. ggc. uuc. gag. ccc. gcg. gac. ugc. aag. cgg. aag. gag. gag.
 gcc. ggg. gcg. ccg. ggc. ggc. ggc. gca. guc. AUG⁴. gcg. gcg. ggc. uuc. ccg. uac. g
 cg. cug. cgc. gcu. uac. cuc. gcc. uac. cag. gcg. gug. ccg. agc. ggc. agc. agc. gg
 g. agc. cuc. ucc. acg. ucc. ucc. ucg. ucc. .

Рисунок 1. Механизмы альтернативной трансляции.

(А) линейное сканирование – базовый вариант инициации трансляции в эукариотических клетках. 40S субъединица рибосомы связывается с кепом и движется линейно в 3'-направлении до первого триплетта AUG. (B) «сканирование с подтеканием» (leaky scanning). Если 5'-проксимальный AUG расположен в субоптимальном контексте, некоторые 40S субъединицы рибосом распознают его и иницируют трансляцию, но остальные продолжают сканирование и могут распознать следующий (следующие) AUG в качестве сайтов инициации. (C) Реинициация трансляции возможна, если 5'-проксимальная рамка считывания мала. В этом случае после терминации трансляции некоторые 40S субъединицы рибосом могут остаться связанными с мРНК и продолжить сканирование, за время которого они связывают eIF2-met-tRNAi-GTP и восстанавливают трансляционную компетентность. Этот механизм позволяет рибосомам пропускать стартовые кодоны в оптимальном контексте и иницировать трансляцию на AUG, удаленных от 5'-конца мРНК. (D) Пример трансляционного контроля синтеза транскрипционного фактора C/EBPβ. AUG¹ – стартовый кодон наиболее протяженной изоформы этого белка, расположенный в субоптимальном контексте (ucAUGc). Некоторые 40S субъединицы рибосомы распознают его, но остальные продолжают линейное сканирование и иницируют трансляцию на AUG², который расположен в другой рамке считывания и с него начинается трансляция маленькой ORF (11 кодонов). Этот стартовый кодон расположен в более оптимальном контексте (agcAUGu) и существенная часть рибосом иницирует на нем трансляцию. Соответственно, AUG³ и AUG⁴, с которых начинается трансляция N-укороченных изоформ этого белка, либо пропускают AUG² в результате «сканирования с подтеканием», либо реиницируют трансляцию после терминации на стоп-кодоне ORF, начинающейся с AUG².

Обозначения: сплошные стрелки – линейное сканирование, прерывистые стрелки – трансляция; R, пурин; Y, пиримидин; H, не G; n, любой нуклеотид; 40S, малая субъединица рибосомы; 40S*, малая субъединица рибосомы в процессе восстановления трансляционной компетентности (то есть, временно не способная распознавать стартовые AUG кодоны); 60S, большая субъединица рибосомы.

Однако представляется достаточно вероятным, что и IRES, и участки шунтирования потока рибосом требуют присутствия в составе 5'-НТП специфических сигналов, в то время как линейное сканирование представляет собой “базовый” механизм инициации трансляции, который работает при наличии кэпа на 5'-конце матрицы. Кроме того, uAUG могут присутствовать в участке мРНК, расположенном ниже IRES или “шунтированного” участка лидерного района, по которому 40S субъединицы рибосом будут двигаться по правилам линейного сканирования. Таким образом, согласно современным представлениям о механизмах инициации трансляции, транслироваться должна существенная часть uORF. В целом, в настоящее время отсутствуют четкие критерии, позволяющие предсказать трансляционный потенциал uORF и их влияние на трансляцию, поэтому эти структурные элементы мРНК эукариот не аннотируют в банках данных нуклеотидных последовательностей и специалисты – экспериментаторы не знают об их существовании.

Потенциальные альтернативные стартовые кодоны трансляции. В подавляющем большинстве эукариотических мРНК, представленных в банках нуклеотидных последовательностей, аннотирован только один стартовый кодон трансляции. Следует, однако, учесть, что границы БКП предсказаны, а не определены экспериментально. Точность предсказания зависит от адекватности использованных критериев и может варьировать. Анализ нуклеотидных последовательностей эукариотических мРНК показал, что они часто содержат стартовый кодон трансляции в субоптимальном контексте и триплеты AUG в составе 5'-НТП (Rogozin et al., 2001; Kochetov, 2005; Bazykin and Kochetov, 2011) (в этой связи сам термин “нетранслируемый” представляется не вполне адекватным, по-видимому, 5'-НТП следует понимать как район мРНК, расположенный выше стартового кодона трансляции основной ORF).

Функциональный смысл субоптимальных стартовых кодонов БКП не очевиден, поскольку их присутствие снижает эффективность трансляции эукариотических мРНК. Считается, что они могут использоваться в механизмах контроля уровня экспрессии генов для предотвращения избыточной продукции регуляторных факторов (Kochetov et al., 1998; Rogozin et al., 2001), поскольку изменение контекста стартового кодона во многих случаях приводит к негативным последствиям. Особенно значимы мутации в ключевых позициях контекста: например, замена G – T в позиции -3 перед стартовым кодоном гена BRCA1 приводит к 70% снижению уровня синтеза белка и сопровождается высоким риском рака груди; замена G на A в позиции +4 мРНК рецептора андрогенов приводит к снижению уровня синтеза на 20 - 50% и сопровождается синдромом частичной нечувствительности к андрогенам. Замена в позиции -3 перед AUG (G заменен на A) в мРНК гена 11 beta-hydroxysteroid dehydrogenase type 1 увеличивает синтез белка, что может отражаться в ассоциации этой мутации с синдромами

ожирения и диабета (Malavasi et al., 2010). В некоторых случаях мутации в минорных позициях контекста также приводят к заболеваниям: например, замена С на Т в позиции -1 мРНК CD40 снижает трансляцию на 13% и сопровождается развитием болезни Грейвса (Jacobson et al., 2005). Вообще, значимость мутаций, изменяющих трансляционную активность мРНК, привлекает все больше внимания. Для описания таких ситуаций был даже предложен специальный термин «translational pathophysiology» (трансляционная патофизиология) (Cazzola and Skoda, 2000); можно также привести недавние обзоры (Chatterjee and Pal, 2009; Le Quesne et al., 2010).

Ранее нами было выдвинуто предположение о том, что при картировании структуры генов и оценке трансляционных характеристик мРНК необходимо учитывать возможность инициации трансляции на нескольких альтернативных стартовых кодонах. Мы предположили, что мРНК, у которых аннотированный стартовый кодон трансляции БКП расположен в субоптимальном контексте, могут содержать дополнительные альтернативные сайты инициации. С целью проверки этой гипотезы проведен сравнительный анализ нуклеотидных последовательностей мРНК эукариот, у которых стартовый кодон трансляции расположен в оптимальном и субоптимальном контексте (Kochetov, 2008). На рис. 2 приведены средние частоты кодонов AUG в начале белок-кодирующих последовательностей мРНК мыши (downstream AUG, dAUG), различающихся по контексту аннотированного стартового кодона.

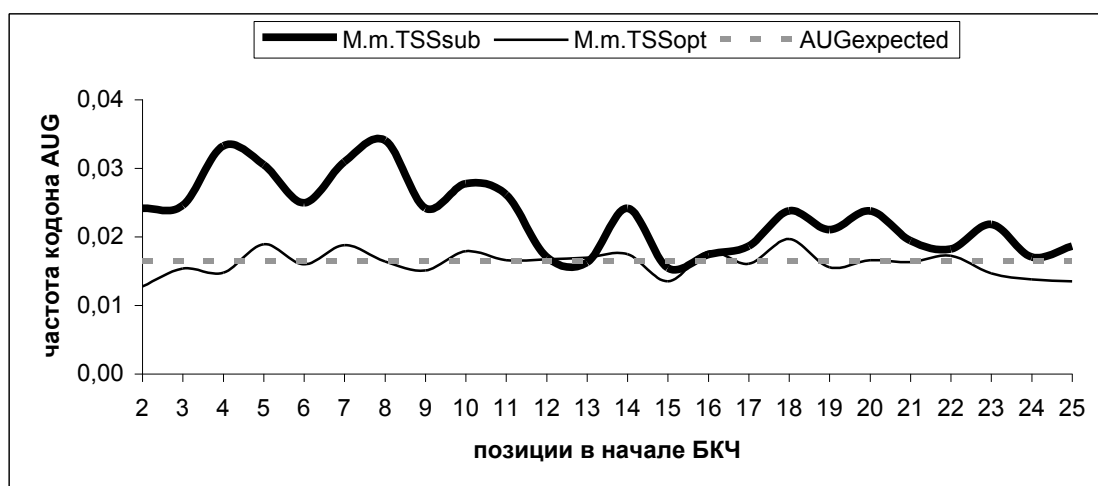


Рисунок 2. Средняя частота кодона AUG в начале БКП мРНК мыши (M. m.), характеризующихся оптимальным (пурин в позиции -3, TSSopt) и субоптимальным (пиримидин в позиции -3, TSSsub) контекстами стартового кодона трансляции. AUGexpected (0.016) – ожидаемая частота кодона, рассчитанная исходя из предположения о случайном присутствии одного из 61 кодонов в позициях белок-кодирующей части (БКП) мРНК.

Можно видеть, что средняя частота dAUG в мРНК с оптимальным контекстом старта трансляции близка к случайной, в то время как в выборке с субоптимальным контекстом старта трансляции эти частоты в 1.5 - 2.5 раза выше. Аналогичная зависимость наблюдалась и

в мРНК других эукариот (Kochetov, 2008). Смысл использования альтернативных стартовых кодонов достаточно очевиден. Они могут применяться для увеличения трансляционной активности мРНК, поскольку добавочный инициаторный кодон будет распознаваться рибосомами, “не заметившими” проксимальный AUG в субоптимальном контексте. Кроме этого, такие укороченные белки, транслирующиеся с альтернативных стартовых кодонов, могут обладать иными функциональными характеристиками. Судя по некоторым данным, альтернативная трансляция может использоваться для контроля экспрессии регуляторных факторов. При этом N-концевые изоформы регуляторного белка могут отличаться по функциональным свойствам. Например, ген транскрипционного фактора человека MDM2 кодирует две изоформы белка (p90 и p75), причем p75 представлен комбинацией из трех N-концевых изоформ, трансляция которых начинается с 1, 62 и 102 кодонов AUG и которые отличаются по выполняемым функциям (Cheng and Cohen, 2007). мРНК важнейшего фактора p53 человека дополнительно кодирует N-укороченный вариант белка, начинающийся с AUG40 и не содержащий транс-активационного домена. Было выдвинуто предположение о возможной роли этой изоформы p53 в качестве транзитного ингибитора полноразмерного варианта белка во время клеточного деления (Courtois et al., 2002). Альтернативная инициация трансляции у AUG кодонов в позициях 1, 17, 33, 69 в мРНК протеинкиназы Sgk1 (Serum- and glucocorticoid-induced kinase 1) мыши приводит к синтезу четырех N-концевых изоформ, различающихся по стабильности, субклеточной локализации и функциям (Arteaga et al., 2007). Аналогично, «сканирование с подтеканием» и альтернативная трансляция позволяют синтезировать N-усеченные варианты рецептора глюкокортикоидов человека (Yudt and Cidlowski, 2001). Согласно нашим оценкам, такой “трансляционный полиморфизм” эукариотических белков может вносить существенный вклад в протеом клеток и органелл (Kochetov, 2008; Vazykin and Kochetov, 2011), сопоставимый со вкладом альтернативного сплайсинга.

Контроль экспрессии генов стрессового ответа на посттранскрипционном уровне: роль механизма реинициации трансляции. Считается, что реинициация трансляции в эукариотических системах малоэффективна и может использоваться в регуляторных целях (рис. 1 С). Действительно, реинициация заметна, если 5'-проксимальная рамка считывания мала (причем чем она чем меньше, тем реинициация эффективнее) (Jackson et al., 2010). Поэтому, в большинстве известных случаев реинициации трансляции, пептид, который кодируют uORF, не выполняет каких-то функций сам по себе. Однако, в некоторых ситуациях специфическая аминокислотная последовательность такого пептида важна, потому что скорость элонгации или терминации трансляции зависит от присутствия определенных факторов, что и регулирует трансляционную активность мРНК и уровень экспрессии

соответствующих генов. У растений трансляционная активность мРНК гена транскрипционного фактора bZIP11 ингибируется при избытке сахарозы, что ведет к изменениям в метаболизме некоторых аминокислот. Показано, что эта репрессия регулируется uORF2, кодирующей пептид из 28 аминокислотных остатков. Этот пептид, по-видимому, обладает специфической конформацией, которая при избытке сахарозы в клетке приводит к остановке транслирующих его рибосом (Rahmani et al., 2009). У *Neurospora crassa* в мРНК генов, контролирующих метаболизм аргинина расположена uORF, кодирующая arginine-attenuator-peptide, который, в присутствии избытка аргинина в клетке, останавливает трансляцию uORF на стадии терминации (Spevak et al., 2010). В мРНК S-аденозилметиониндекарбоксилазы млекопитающих в 5'-НТП расположена маленькая uORF, кодирующая гексапептид MAGDIS. В присутствии поламинов ингибиторный эффект этой uORF изменяется, что позволяет регулировать экспрессию связанных с их метаболизмом ферментов (Ivanov et al., 2010). Однако, в большинстве случаев трансляционный эффект uORF не зависит от аминокислотной последовательности кодируемого ими пептида.

Реинициация трансляции может в принципе использоваться для тонкого контроля экспрессии гена, поскольку имеет определенные преимущества перед «сканированием с подтеканием». Поскольку 40S субъединицы рибосомы после терминации трансляции на стоп-кодоне uORF продолжают сканирование, но восстанавливают трансляционную компетентность не сразу (как минимум должен связаться комплекс eIF2-met-tRNAi-GTP), то такие рибосомы могут пропустить ближайшие к терминатору uORF триплеты AUG и инициировать трансляцию на удаленных сайтах. Более того, поскольку процесс восстановления трансляционной компетентности вероятностный, трансляция может реиницироваться на нескольких таких кодонах. Если 40S субъединицы рибосом пропускают uAUG в результате «сканирования с подтеканием», то при линейном сканировании в 3'-направлении каждый встречный AUG в оптимальном контексте будет восприниматься в качестве сайта инициации трансляции и их пропуск невозможен, поэтому для тонкой настройки трансляционного контроля экспрессии регуляторных генов часто используется комбинация реинициации и «сканирования с подтеканием».

Например, в мРНК транскрипционного фактора SCL содержатся четыре стартовых кодона, перед которыми расположена uORF со стартовым кодоном в субоптимальном контексте. Было показано, что часть 40S субъединиц рибосомы пропускает uAUG с помощью «сканирования с подтеканием» и иницирует трансляцию на первом стартовом кодоне SCL. Однако, инициация трансляции на трех других AUG, расположенных во внутренних районах БКП, осуществляется в результате реинициации трансляции и в ней задействованы те рибосомы, которые транслировали uORF (uORF расположена вблизи от AUG1 SCL и этот

стартовый кодон не распознается рибосомами, терминирующими трансляцию на стоп-кодоне uORF, так как они не успевают восстановить трансляционную компетентность) (Calkhoven et al., 2003). Сходная схема синтеза N-концевых изоформ была обнаружена для важных транскрипционных факторов C/EBP β и C/EBP α (Calkhoven et al., 2000; Wethmar et al., 2010) (рис. 1 D). В мРНК РНКазы H1 человека в 5'-НТП есть uORF, стартовый кодон которой расположен в субоптимальном контексте. С мРНК считываются две N-изоформы белка – митохондриальная (с AUG1) и ядерная (с AUG2). Трансляция на AUG1 инициируется рибосомами, пропустившими uORF («сканирование с подтеканием»), тогда как трансляция на AUG2 реиницируется рибосомами после трансляции uORF. Это дает возможность точного контроля синтеза изоформ РНКазы H1 (митохондриальная изоформа составляет около 10% от ядерной), причем это соотношение важно – при его нарушении клетки гибнут (Suzuki et al., 2010).

Роль реинициации трансляции в стресс-специфическом контроле экспрессии генов эукариот связана с тем, что ее эффективность может зависеть от функциональной активности трансляционных факторов. Например, было показано, что скорость восстановления трансляционной компетентности 40S субъединиц рибосом зависит от статуса фосфорилирования eIF2 α . В 5'-НТП мРНК транскрипционного фактора GCN4 дрожжей расположено несколько uORFs. Рибосомы после трансляции uORF1 способны реиницировать на стартовом кодоне следующей uORF и в нормальных условиях инициация трансляции на белок-кодирующей последовательности низка, поскольку размеры спейсера между стоп-кодомом uORF1 и стартовыми кодонами нижерасположенных uORF достаточны для восстановления трансляционной компетентности. Однако, в условиях стресса происходит фосфорилирование eIF2 α и концентрация его активной формы снижается. Это приводит к увеличению времени, необходимой для восстановления комплекса 40S-eIF2-met-tRNAⁱ-GTP и – в результате – к инициации трансляции на стартовом кодоне GCN4. Этот механизм используется клетками для быстрого ответа на стрессовые условия, поскольку скорость синтеза белка с уже синтезированных и присутствующих в цитоплазме мРНК значительно выше, чем при индукции транскрипции. У млекопитающих аналогичный механизм был найден у ключевых транскрипционных факторов стрессового ответа ATF4 (Vattem and Wek, 2004), ATF5 (Zhou et al., 2008). В мРНК ATF5 в составе 5'-НТП расположены две uORF: короткая 5'-проксимальная uORF1 и более длинная uORF2, которая перекрывается с БКП. При нормальном статусе фосфорилирования eIF2 α рибосомы реиницируют трансляцию на uORF2 и не достигают старта трансляции ATF5, но при стрессе в условиях фосфорилирования eIF2 α ситуация меняется. Эти примеры показывают, что uORF могут использоваться не только

для конститутивного контроля трансляционной активности мРНК, но и способны изменять спектр и количество продуктов трансляции в зависимости от состояния клетки.

На предыдущем этапе проекта нами был проведен анализ мРНК генов млекопитающих, характеризующихся различной эффективностью трансляции в норме и при стрессе. Выдвинуто предположение о том, что реинициация трансляции, характеристики которой зависят от статуса фосфорилирования фактора eIF2 α , является одним из ключевых механизмов стресс-специфического контроля экспрессии генов млекопитающих. На текущем этапе проекта нами было запланировано оценить вклад реинициации трансляции в синтез белков человека, а также связь соответствующих генов с молекулярными механизмами стрессового ответа.

*Исследование молекулярно-генетических механизмов ответа на стрессовые воздействия также проводилось на эволюционно-консервативной модели *Drosophila* (у насекомых ответ на стрессовые факторы сопровождается изменением в балансе ряда гормонов и биогенных аминов, что может рассматриваться в качестве интересной модели эволюционной консервативности механизмов стрессового ответа).*

Известно, что у животных нейромедиатор и нейрогормон дофамин (ДА) опосредует ряд нейральных и эндокринных функций (Blenau and Baumann, 2001; Neckameyer and Leal, 2002; Waddel, 2010; Van Swinderen and Andretic, 2011). Дофаминовые сигнальные пути участвуют в регуляции агрессии, гормонального контроля полового поведения, контроля двигательной активности. Два семейства G-протеин-сопряженных рецепторов, D1-подобных и D2-подобных, опосредуют физиологические функции дофамина. Показано, что эти рецепторы высоко консервативны у млекопитающих, человека в частности, и у дрозофилы (Vallone et al., 2000; Hearn et al., 2002; Draper et al., 2007). Установлено, что у млекопитающих D2-подобный рецептор является ключевым модулятором двигательной функции. У человека D2-подобный рецептор оказался важнейшей терапевтической мишенью для лечения брадикинезии, возникающей в результате болезни Паркинсона (обз: Vallone et al., 2000).

Однако пока практически ничего не известно о том, какую роль контроль экспрессии генов рецепторов дофамина играет в регуляции гормонального статуса организма в стрессовых условиях. Проведение подобных исследований на человеке невозможно, так как требует фармакологического изменения уровней стресс-связанных гормонов и проведения генетических экспериментов. Поскольку показано, что ДА-подобные рецепторы являются высококонсервативными, в рамках настоящего проекта используется уникальная модель, созданная нами ранее. Модель представляет собой набор линий дрозофилы, несущих мутации, каждая из которых вызывает изменения в уровне одного из стресс-связанных гормонов, в том числе дофамина. Эта модель высокоперспективна еще и потому, что соответствующие

регуляторные контуры у насекомых устроены значительно проще, чем у млекопитающих, что позволяет четко определить все наличествующие в них взаимосвязи. Выявленные закономерности могут быть далее использованы как модель для коррекции гормонального дисбаланса, возникающего при стрессе у млекопитающих.

Установлено, что у взрослых насекомых, в том числе у *Drosophila*, ювенильный гормон (ЮГ) и 20-гидроксиэкдизон (20Э), с одной стороны, задействованы в механизме стресс-реакции (Rauschenbach et al., 1995; Hirashima et al., 2000), а с другой - играют роль гонадотропинов (они инициируют синтез желточных белков (ЖБ) в жировом теле (ЖТ) и фолликулярных клетках яичника сразу после вылета имаго, поддерживая его на определенном уровне после оплодотворения, регулируя поглощение желточных белков ооцитами) и детерминируют половое поведение и синтез феромонов (обзоры: Truman and Riddiford, 2007; Toivonen and Partridge, 2009).

Модель гормонального контроля синтеза желточных белков, в которой ювенильный гормон стимулирует синтез (а затем – секрецию и поглощение) желточных белков фолликулярными клетками яичников и одновременно индуцирует экдизон-секретирующие клетки яичника к синтезу или секреции 20Э, который, в свою очередь, инициирует синтез желточных белков в жировом теле была разработана Постлесвейтом и Паркером (Postlethwait and Parker, 1987). Эта модель была дополнена Соллером и соавторами (Soller et al., 1999) в результате анализа данных своих экспериментов по обработке самок *D. melanogaster* экзогенными ЮГ и 20Э. Они полагают, что развитие вителлогенических ооцитов, включая как синтез желточных белков в фолликулярных клетках, так и их поглощение ооцитами, стимулируются ЮГ, тогда как 20Э контролирует превителлогенические стадии развития ооцитов. Согласно их представлениям, у *Drosophila* для нормального течения оогенеза принципиальным является поддержание баланса ЮГ и 20Э (Soller et al., 1999). В наших исследованиях это представление было подтверждено и показано, что фармакологическое нарушение баланса вызывает значительные изменения в ходе оогенеза (Gruntenko et al., 2003). Мы также показали (обз: Gruntenko and Rauschenbach, 2008), что нарушения репродуктивной функции, наблюдающиеся у дрозофилы в неблагоприятных условиях, вызваны происходящим при стрессе изменением баланса ювенильного гормона и 20-гидроксиэкдизона.

Таким образом, очевидно, что для поддержания баланса гонадотропинов, необходимого для нормального размножения, должен существовать механизм, обеспечивающий регуляцию титра одного из гонадотропинов под воздействием изменения в титре другого. Каков же может быть механизм подобной взаимной регуляции ЮГ и 20Э?

Показано, что помимо специфических нейропептидов, обладающих проторакотропным, аллатотропным и аллатостатическим действием (обзоры: Truman and Riddiford, 2007; Toivonen and Partridge, 2009), метаболизм ЮГ и 20Э может регулироваться биогенными аминами, октопамином (ОА) и ДА, влияние которых на метаболизм гонадотропинов обнаружено в экспериментах *in vitro* на различных видах насекомых (Lafon-Cazal and Baehr, 1988; Woodring and Hoffmann, 1994; Kaatz et al., 1994; Rachinsky, 1994; Granger et al., 1996; Hirashima et al., 1999). Однако долгое время оставалось неизвестным, обладают ли эндогенные амины таким же влиянием *in vivo*, является ли это влияние генетически детерминированным, и влияют ли октопамин и дофамин как нейрогормоны на воспроизводительную функцию (оогенез, плодовитость) насекомых. Мы выяснили это, используя фармакологические подходы и генетические модели – линии дрозофилы у которых в результате мутации резко изменен уровень одного из биогенных аминов.

Для оценки влияния изменения уровня биогенных аминов *in vivo* на метаболизм ЮГ, мы использовали мутантные линии *D. melanogaster*: линию $T\beta h^{nM18}$ с нуль мутацией гена *тирамин-β-гидроксилазы*, приводящей к полному отсутствию октопамина у мух этой линии, и линий *D. melanogaster*, с удвоенным в результате мутации гена *N-β-аланилдофаминсинтетазы (ebony)* уровнем дофамина. Мы обнаружили, что эти мутации приводят к существенным изменениям в метаболизме ювенильного гормона. Самки, лишенные октопамина ($T\beta h^{nM18}$), как молодые, так и интенсивно размножающиеся, имеют уровень деградации ювенильного гормона значительно более высокий, чем самки линии-предшественницы и чем особи дикого типа (*Canton S*). В то же время, у молодых самок с удвоенным содержанием дофамина (*ste, ebony*) деградация гормона значительно ниже, а у зрелых выше, чем у дикого типа (обзор: Gruntenko and Rauschenbach, 2009). Мы заключили, что октопамин оказывает ингибирующее влияние на активность ферментов деградации ювенильного гормона как у молодых, так и у половозрелых самок дрозофилы, а дофамин ингибирует активность ферментов деградации ЮГ у молодых самок и стимулирует ее у зрелых самок.

Заметим, что такое же стадии-зависимое влияние дофамина на метаболизм ЮГ у самок тараканов обнаружено Пастором с соавторами (Pastor et al., 1991): дофамин стимулирует синтез ЮГ в железе *corpus allatum (CA)* 2-суточных самок *Blattella germanica*, и ингибирует ее у 6-суточных. Онтогенетический характер регуляции уровня ЮГ дофамином был также продемонстрирован Грангер с соавторами (Granger et al., 1996), показавшими, что этот биогенный амин стимулирует синтез ЮГ в *CA* в первые два дня последней личиночной стадии табачного бражника *Manduca sexta*, но ингибирует его в начале пред-куколичного периода.

Естественно возникает вопрос, каким может быть механизм такого разнонаправленного действия дофамина на метаболизм ЮГ на разных стадиях развития, обнаруженного у насекомых из различных отрядов? Как уже было сказано выше, допаминэргические сигналы опосредуются через трансмембранные рецепторы. Семейство рецепторов дофамина подразделяется на два больших подсемейства: D1-подобные рецепторы (D1 и D5) и D2-подобные рецепторы (D2, D3 и D4) (обзор: Vallone et al., 2000). Показано, что D1-подобные рецепторы связываются со стимулирующими G-белками (Gas), вызывая активацию аденилатциклазы и запуск каскада, приводящего к активации генов-мишеней. D2-подобные рецепторы передают сигнал через ингибирующие G-белки (Gai/o), вызывая ингибирование аденилатциклазы и, соответственно, ингибированию генов-мишеней (обзоры: Missale et al., 1998; Vallone et al., 2000). D2-подобный рецептор *Drosophila* (DD2R) был клонирован и охарактеризован (Hearn et al., 2002). Дрэпер с соавторами (Draper et al., 2007) определили паттерн пространственной экспрессии гена *DD2R* и функционально охарактеризовали мух со сниженным уровнем экспрессии *DD2R*.

Обнаруженное нами изменение влияния ДА на активность ферментов деградации ЮГ с ингибирующего (у молодых самок дрозофилы) на стимулирующее (у половозрелых) и на уровень синтеза гормона со стимулирующего на ингибирующее позволило нам предположить наличие переключения на геномном уровне, приводящего к изменению состава и/или числа D1- и D2-подобных рецепторов в жировом теле (ЖТ) (месте синтеза ферментов деградации ЮГ) и *corpus allatum* (CA – железе, синтезирующей ЮГ) (обзор: Gruntenko and Rauschenbach, 2008). Учитывая, что синтез и деградация ЮГ у самок дрозофилы находятся в противофазе и под общим контролем мы предположили (Gruntenko and Rauschenbach, 2008), что у молодых самок дрозофилы в ЖТ представлены, в основном, D2-подобные рецепторы, а в CA – D1-подобные рецепторы. При половом созревании, после того как ЮГ индуцирует синтез эcdистероидов и начало вителлогенеза (обзор: Toivonen, Partridge, 2009) происходит переключение и в ЖТ синтезируются D1-подобные рецепторы, а в CA – D2-подобные, что приводит к инвертированию влияния ДА на активность ферментов синтеза и деградации ЮГ. Такое переключение может, по-видимому, осуществлять 20Э, титр которого низок у молодых самок дрозофилы и повышается при половом созревании (Handler, 1982). Последнее предположение хорошо согласуется с данными Грангер с соавторами (Granger et al., 1996), показавшими, что эcdистероиды вовлечены в регуляцию стадийспецифического ответа CA на ДА у личинок последнего возраста *Manduca sexta*.

В рамках этапа 5 настоящего проекта мы иммуногистохимически изучили: (1) экспрессию гена *DD2R* в CA молодых и половозрелых самок дикого типа *D. melanogaster*, (2) экспрессию гена *DD2R* в ЖТ молодых и половозрелых самок дикого типа *D. melanogaster*, (3) влияние 20Э

на экспрессию гена *DD2R* в ЖТ самок дикого типа *D. melanogaster* и (4) экспрессию гена *DD2R* в ЖТ молодых и половозрелых самок *D. melanogaster* с генетической абляцией части клеток *CA*.

1. Проведение сравнительного изучения паттернов полногеномного распределения транскрипционных факторов семейства FOX в печени экспериментальных животных в норме и под действием гепатоканцерогенных соединений.

В ходе предыдущих этапов выполнения проекта были получены специфические ChIP-Seq библиотеки для системы SOLiD. Поскольку библиотеки, полученные от животных, обработанных ортоазоаминотолуолом, как оказалось, имели неудовлетворительное качество, для последующего секвенирования на системе SOLiD и теоретического анализа полученных данных были использованы только библиотеки, полученные от новорождённых и взрослых (двухмесячных) мышей.

Для картирования в качестве референсного был использован геном мыши (июль 2007 г.; NCBI37/mm9), с маскированными повторами, доступный для скачивания на сервере Genome Bioinformatics Group of UC Santa Cruz (<http://hgdownload.cse.ucsc.edu/downloads.html#mouse>), общая длина референсной последовательности составила 2654895218 п.н. Поскольку длина прочтений системы SOLiD в использованном нами варианте составляет 35 оснований, то локализация таких коротких прочтений в областях повторяющихся последовательностей является крайне затруднительной и неточной, соответственно, маскирование повторов (удаление районов локализации повторов из референсной последовательности) является необходимым.

В качестве программы для картирования прочтений был использован рекомендованный производителем SOLiD™ BioScope™ Software с настройками по умолчанию. Однако в результате дальнейшей проверки выяснилась необходимость оптимизации параметров работы программы для более эффективного удаления шума и артефактов, вносимых особенностями экспериментальной процедуры. В итоге проверки возможных вариантов наборов параметров был изменен параметр `mapping.scheme.unmapped.35`, который был выбран равным 26.1.0. (по умолчанию – 25.2.0), каждое из значений соответствует параметрам выравнивания [`anchorLength.mismatchAllowed.anchorStart`]. Изменив настройки таким образом, мы снизили количество разрешенных несовпадений с референсной последовательностью и увеличили максимально возможную длину выравниваемого фрагмента, что позволило снизить количество шума в данных, соответственно повысив качество картирования. Однако такое ужесточение настроек привело к заметному уменьшению числа картированных на геном

прочтений. Сводные данные по всем использованным в анализе результатам экспериментов ChIP-Seq приведены в Таблице 1.

Таблица 1. Результаты картирования полученных прочтений результатов экспериментов ChIP-Seq для различных антител на референсный геном.

Антитела	Стадия развития	Общее количество прочтений	Количество картированных на геном прочтений	Процент картированных на геном прочтений
IgG	Мыши взрослые	111027313	5762319	5%
IgG	Мыши новорожденные	141744159	7300503	5%
HistH3	Мыши взрослые	64028347	24971005	39%
HNF4	Мыши взрослые	59331013	4046375	7%
HNF4	Мыши новорожденные	77361352	4719042	6%
FoxO3A	Мыши взрослые	118352700	17921830	15%
FoxO3A	Мыши новорожденные	98713570	13946040	14%
FoxA2	Мыши взрослые	103487891	18159301	18%
FoxA2	Мыши новорожденные	106441572	17397410	16%

Далее в соответствии с рекомендациями производителя с помощью этого же ПО (SOLiD™ BioScore™ Software) производилась конвертация выходного формата файла с результатами картирования (.ma -> .bam), для последующей подачи на вход ПО MACS version 1.4.0 (Zhang et al., 2008). Программа MACS предназначена для проведения процедуры поиска пиков ChIP-Seq (ChIP-Seq peak calling) и является одной из самых широко используемых. Кроме того, она обладает наибольшей точностью в корректном выделении пиков ChIP-Seq и правильном определении их границ (Malone et al., 2011). Программа MACS была использована с рекомендуемыми разработчиками настройками. Кроме того, одной из опций программы, существенно выделяющей её из серии программ для процедуры поиска пиков ChIP-Seq, является возможность использовать контрольные распределения прочтений для снижения уровня шума, вызванного ошибками и неточностями экспериментальной процедуры. Все перечисленные вычислительные процедуры были выполнены на высокопроизводительном вычислительном комплексе ЦКП «Биоинформатика» СО РАН,

основой которого служит вычислительный кластер с 64 двойными блейд-серверами HP BL2×220G6 пиковой производительностью 10.362 ТФлопс.

Полученные результаты картирования прочтений экспериментов ChIP-Seq показывают падение относительной доли картированных прочтений в случае транскрипционных факторов семейства FOX по сравнению с гистоном H3. к подобному результату могут приводить сразу несколько факторов. Во-первых, обычно районы повторов обеднены сайтами связывания транскрипционных факторов, и маскирование повторов не может существенным образом сказаться на результатах анализа. Однако на данном этапе возникает ситуация, существенно осложняющая анализ полногеномного распределения сайтов связывания транскрипционных факторов семейства FOX. Ранее в нашей лаборатории было показано эффективное связывание FOX-белков с микросателлитными повторами, представляющими из себя тандемный повтор тетрауклеотида (TTTG)_n, где n равно трём и более (Брызгалов и др., 2008). Подобные ситуации хоть и редки, но описаны и для некоторых других транскрипционных факторов (Albanese et al., 2001; Gangwal et al., 2010). Соответственно, некоторое количество чтений должно попадать в районы соответствующих микросателлитных повторов и выпадать из картирования на маскированном по повторам референсе. Кроме того, создание библиотек системы SOLiD проводилось с использованием набора SOLiD ChIP-Seq Kit фирмы Applied Biosystems в строгом соответствии с инструкцией к нему. В ходе создания библиотек в данном варианте при выполнении процедуры гидролиза ковалентных ДНК-белковых сшивок применяется нагревание иммунопреципитированного материала при 95°C в течение 15 минут в отличие от нагревания в течение 2 часов при 65°C в стандартной методике. При этом происходит денатурация/ренатурация фрагментов ДНК, что при низких концентрациях ДНК должно приводить к более эффективной ренатурации повторяющихся последовательностей. Поскольку на следующей стадии процесса осуществляется лигирование двуцепочечных фрагментов ДНК с двуцепочечными олигонуклеотидными адаптерами, то стадия денатурации может вызывать селективное увеличение доли повторов относительно уникальных последовательностей в составе полученной SOLiD-библиотеки. Такое смещение представленности повторов относительно их доли в геноме может быть не принципиальным при получении ChIP-Seq SOLiD-библиотек для большинства транскрипционных факторов, но оказаться существенным в случае факторов семейства FOX, способных связываться с определёнными микросателлитными повторами, поскольку для картирования прочтений системы SOLiD обычно применяется референс с маскированными повторяющимися последовательностями.

В ходе первого этапа биоинформатического анализа проводилась оценка полученных ChIP-Seq пиков для всех образцов по форме и особенностям распределения прочтений, что

позволило оценить их качество и внести необходимые изменения в параметры работы программ на последующих этапах анализа. В качестве положительного контроля правильности проведения экспериментальных процедур использовался эксперимент с антителами против метилированного гистона H3 для взрослых мышей. Процент картированных ридов для гистона H3 максимальный, экспериментальные артефакты, выражающиеся в появлении «столбообразных» пиков, незначительны. Пики ChIP-Seq имеют ожидаемую форму. При данном невысоком среднем покрытии (менее чем 3-х кратное в среднем на геном) программой MACS выделено 41163 пика, при этом многие прочтения находятся вне расположения пиков, что связано с повсеместным распределением гистона H3 в составе хроматина. Таким образом, можно заключить, что при проведении экспериментальных процедур не было допущено систематических ошибок. Отрицательным контролем служили эксперименты по иммунопреципитации с использованием антител против суммарного IgG мыши как для новорождённых, так и взрослых мышей. Поскольку данный белок отсутствует в составе хроматина, то результатом иммунопреципитации должны были стать исключительно неспецифические последовательности и артефакты, возникающие в результате процедуры создания библиотек SOLiD в условиях крайне малого количества ДНК. Наблюдаемый результат оказался таким, как и ожидалось. Для ChIP-Seq с IgG мыши основная масса прочтений была локализована в характерных очень узких и очень высоких «столбообразных» пиках, остальные прочтения были достаточно равномерно разбросаны по всему геному. На рис.3 в правой части приведены примеры подобных «столбообразных» артефактов. Аналогичная картина наблюдалась также в результатах экспериментов с антителами против транскрипционного фактора HNF4 и для взрослых, и для новорождённых мышей. Основная часть прочтений концентрировалась в «столбообразных» пиках, из чего нами был сделан вывод об отсутствии всякого специфического связывания этих антител с хроматином клеточных ядер. В ChIP-Seq экспериментах, в отличие от Вестерн-блоттинга, происходит связывание антител не просто с соответствующим белком-транскрипционным фактором, а с комплексом ДНК-белок в составе хроматина, зачастую имеющим изменённую конформацию. Связываться с таким комплексом способна только некоторая часть антител против данного белка, и использованные нами антитела против HNF4 (Santa Cruz, sc-101059) по-видимому, непригодны для постановки ChIP-Seq экспериментов.

Для экспериментов с антителами против транскрипционного фактора FoxO3A и особенно FoxA2 наблюдалась иная картина. Наряду с определённым количеством «столбообразных» артефактных пиков наблюдались ChIP-Seq пики правильной формы и ожидаемого размера. Для эксперимента с антителами против FoxA2 относительная доля чтений, локализованных в зонах между пиками, была заметно меньше, чем для эксперимента с антителами против

FoxO3A, что может быть вызвано недостаточно высокой специфичностью антител против FoxO3A либо неоптимальной их концентрацией на стадии собственно иммунопреципитации. На рис. 3 приведён пример визуализации картирования чтений на одни и те же районы референсного генома в случае ChIP-Seq эксперимента с антителами против FoxA2 для взрослой мыши и аналогичного с антителами против IgG. Хорошо видно наличие в обоих случаях «столбообразного» артефактного пика, меньшего по высоте в случае FoxA2. Для FoxA2 также видны две группы прочтений, сгруппированные в ChIP-Seq пики правильной формы. В случае антител против IgG, являющемся отрицательным контролем, хорошо видно полное отсутствие локализации отдельных прочтений как в области ChIP-Seq пиков, так и в районах между ними, что также подтверждает достаточно высокое качество проведения экспериментальных процедур. Тонкими синими линиями внизу рисунка показан результат работы программы MACS – два выделенных ChIP-Seq пика.

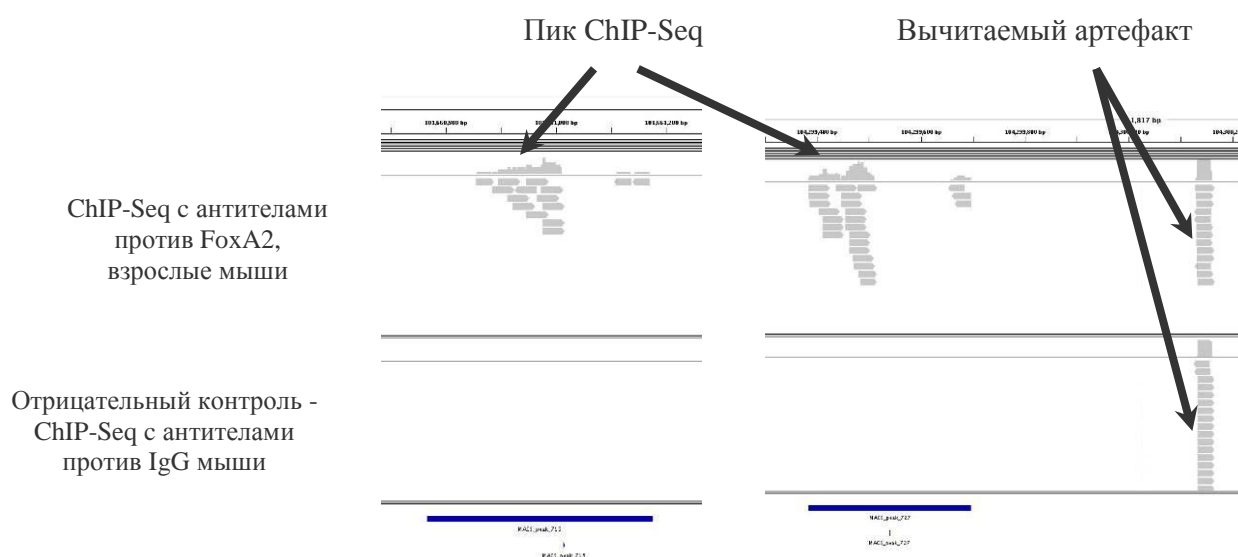


Рисунок 3. Пример визуализации картирования на идентичный район генома мыши прочтений ChIP-Seq эксперимента и выделения пиков ChIP-Seq для взрослых мышей. Снизу – отрицательный контроль, антитела против IgG мыши, сверху - антитела против FoxA2 мыши. В правой части рисунка показаны два пика ChIP-Seq правильной формы. Слева – вычитаемый артефакт, идентичный как для случая IgG, так и для случая FoxA2. Синие линии внизу рисунка – районы пиков ChIP-Seq, идентифицируемые программой MACS.

Программа MACS при настройках по умолчанию достаточно эффективно распознаёт отдельно стоящие артефактные «столбики» и отделяет их от истинных пиков ChIP-Seq. Однако при случайном попадании такого «столбика» в область локализации рассеянной группы отдельных прочтений MACS зачастую идентифицирует такую группу как ChIP-Seq пик большой величины. Аналогично, при наложении «столбика» на истинный пик ChIP-Seq

происходит существенное завышение истинной высоты пика, которая является важной его характеристикой и критерием, активно используемым при анализе. Поэтому на втором этапе нами был проведён повторный анализ выделения пиков ChIP-Seq программой MACS для экспериментов с антителами против FoxA2 как для новорождённых, так и для взрослых мышей с использованием данных ChIP-Seq с антителами против IgG мыши в качестве вычитаемого фона. Использование отрицательного контроля в качестве фонового распределения при повторном анализе данных позволило заметно улучшить качество выделения пиков ChIP-Seq. Важность подобной операции хорошо демонстрируется сравнением наших данных с аналогичными данными по иммунопреципитации хроматина для транскрипционных факторов семейства Fox мыши, взятыми из работы (Wederell et al., 2008), исходные данные находятся в свободном доступе (<http://www.bcgsc.ca/data/ChIP-Seq>) и были повторно проанализированы нами.

Индивидуальные пики ChIP-Seq эксперимента были выделены из исходных данных из работы (Wederell et al., 2008) и сгруппированы в соответствии с их относительной высотой. Далее для каждого из этих пиков извлекалась соответствующая ему последовательность геномной ДНК, с помощью программы monoChIPMunk (Kulakovskiy et al., 2010), созданной для анализа результатов ChIP-Seq экспериментов, в этих последовательностях производился поиск сайтов связывания транскрипционного фактора FoxA. Если считать, что количество прочтений на единицу длины зависит в первую очередь от эффективности связывания транскрипционного фактора, то стоит ожидать, что чем выше пик ChIP-Seq, тем выше аффинность связывания фактора FoxA с этим участком ДНК. Соответственно, логично ожидать в этом районе наличия либо высокоаффинного хорошо предсказываемого сайта связывания, либо кластера из нескольких сайтов связывания, для которого вероятность идентификации хотя бы одного сайта из нескольких также повышена. В действительности же наблюдается довольно парадоксальное распределение (рис. 4). Сначала с ростом высоты пика вероятность нахождения в нём сайта связывания Fox-белков действительно растёт с 70 до 90%, достигая максимума на середине распределения. Однако с дальнейшим ростом высоты пика вероятность идентификации внутри него сайта начинает падать, и у самых высоких пиков снижается вдвое по сравнению с максимальным значением в области пиков средней высоты. Наиболее корректным объяснением наблюдаемого парадокса представляется недостаточно эффективное удаление артефактных пиков, всегда имеющих очень большую высоту, в ходе первичного анализа результатов ChIP-Seq экспериментов авторами данной работы.

Такое предположение подтверждается результатами аналогичного анализа наших данных экспериментов с антителами против FoxA2 как для взрослых мышей (рис. 5), так и для

новорождённых, а также анализом данных из работы (Wallerman et al., 2009) по изучению полногеномного профиля распределения сайтов связывания транскрипционного фактора FoxA2 в клетках гепатомы человека HepG2.

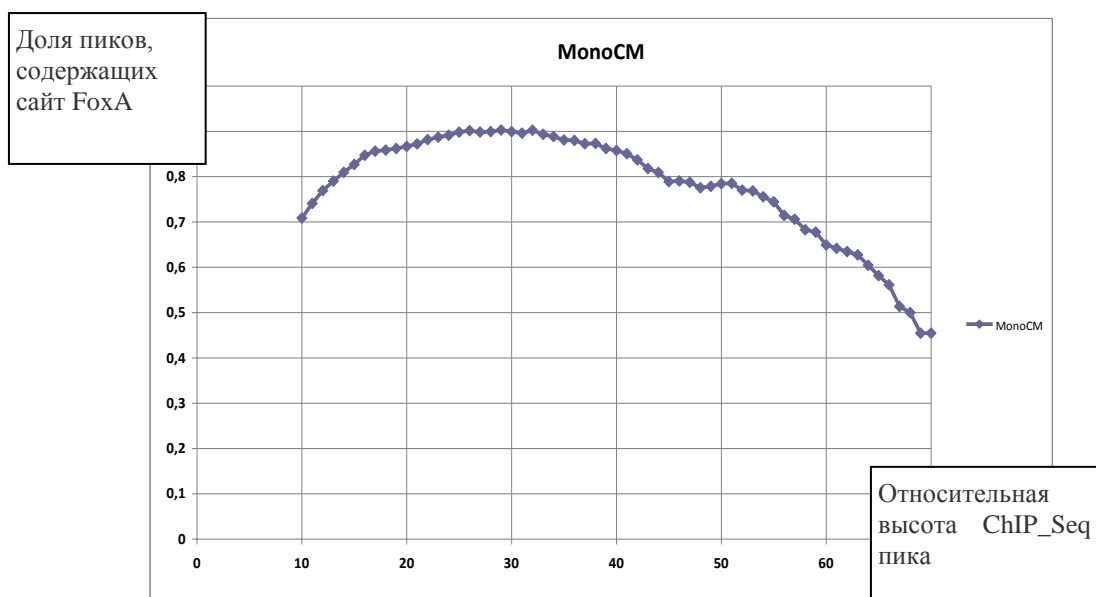


Рисунок 4. Зависимость частоты предсказания программой monoChIPMunk сайта связывания транскрипционного фактора FoxA внутри пика ChIP-Seq от относительной высоты этого пика, анализ по данным из работы (Wederell et al., 2008)

На Рис. 5 хорошо видно, что в данных нашего эксперимента после вычитания фона на всём протяжении распределения пиков ChIP-Seq по высоте не наблюдается значимого падения доли предсказанных сайтов связывания фактора FoxA2 в пиках большой высоты. Общее число выделенных ChIP-Seq пиков при использованных настройках программы MACS составило для новорождённых мышей - 5330 пиков; для взрослых мышей- 7065 пиков со средней длиной 250-300 нуклеотидов.

Как правило, профиль распределения ChIP-Seq пиков для изучаемого транскрипционного фактора в геноме считается соответствующим расположению реальных сайтов связывания этого белка. Однако это верно не для всех случаев. Во-первых, разрешающая способность метода не всегда достаточна для локализации сайта связывания с точностью 100-200 нуклеотидов. Среди пиков ChIP-Seq могут с разной частотой присутствовать пики, вызванные неспецифическим связыванием антител с другими белками либо являющиеся артефактами пробоподготовки. Хорошей иллюстрацией данного положения является работа (Motalebipour et al., 2009).

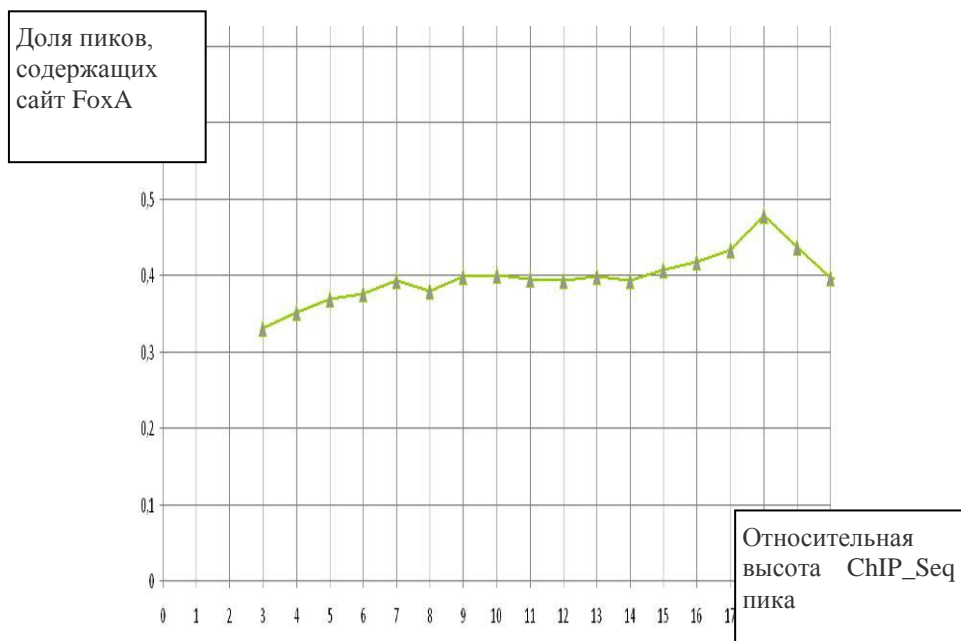


Рисунок 5. Зависимость частоты предсказания программой monoChIPMunk сайта связывания транскрипционного фактора FoxA2 внутри пика ChIP-Seq от относительной высоты этого пика, наши данные, ChIP-Seq с антителами против FoxA2, взрослые мыши.

В данной работе проведён анализ полногеномного распределения транскрипционных факторов FOXA1 и FOXA3 для клеточной линии гепатомы HepG2 человека. Однако только 12,5% (1022 из общего числа 8175) пиков для фактора FOXA1 охарактеризованы как содержащие варианты пятибуквенного мотива TnTTnAc, при том, что даже данный мотив совершенно недостаточен для распознавания эффективного сайта связывания FOXA. В качестве потенциальных мотивов связывания для основной массы пиков в данной работе предложены экзотические варианты CGnnnC, AnnCG, CGC etc. В третьих, технология ChIP-Seq не может различить ситуаций прямого связывания транскрипционного фактора с ДНК в районе специфического сайта и противоположной ситуации – связывание с определённым районом хроматина в результате белок-белкового взаимодействия интересующего нас ТФ с другими белками, непосредственно взаимодействующими с ДНК (piggy back binding), при этом доли прямого и непрямого связывания существенно отличаются для разных транскрипционных факторов (Farnham, 2009).

Использование высокоэффективных программ распознавания сайтов связывания FOX-белков позволяет не только корректно анализировать результаты экспериментов ChIP-Seq, но и жёстко разделять возможные ошибки анализа от возможных ошибок в ходе проведения собственно экспериментальной работы. Поэтому в качестве одной из первоочередных задач на этапе анализа результатов ChIP-Seq экспериментов нами было выбрано тестирование и

оптимизация программ, способных к эффективному распознаванию сайтов связывания белков семейства FOX, а также экспериментальная проверка их работоспособности.

Все эти методы распознавания ССТФ разделяются на два класса: (а) методы распознавания «учителем» (pattern matching), подразумевающие привлечение выборки экспериментально подтвержденных ССТФ; (б) методы распознавания, основанные на поиске мотивов de novo (pattern detection) (Elnitski et al., 2006). Для подхода pattern matching на основе литературных данных собирается выборка последовательностей ДНК обучения (длина каждой 20-50 нуклеотидов, в центральном положении каждой последовательности расположен известный сайт связывания транскрипционного фактора, ССТФ). Для подхода pattern detection достаточно иметь выборку более протяженных последовательностей ДНК, обогащенных сайтами исследуемого типа, при этом в отличие от задачи pattern matching позиции сайтов изначально не известны. Для реализации подхода pattern detection применяются программы поиска мотивов, они дают выборки потенциальных ССТФ, на основе которых можно применять методы распознавания также как и для подхода pattern matching. В результате выборка потенциальных сайтов, каждый из которых представляет собой реализацию найденного мотива, используется как выборка обучения для метода распознавания ССТФ.

В свою очередь, среди методов pattern matching выделяются два типа: (1) канонические, основанные на построении весовой матрицы и (2) неканонические. Канонические методы заключаются в оценке аффинности последовательности потенциального сайта как суммы вкладов отдельных позиций сайта. Для этого в каждой позиции выравнивания в выборке обучения оцениваются вероятности появления моно- или динуклеотидов (вычисляются частоты), затем аффинность потенциального сайта рассчитывается как линейная сумма по вероятностям соответствующих позиций. Неканонические методы учитывают зависимости между удаленными позициями функциональных сайтов. Эти методы очень разнообразны.

Нами были применены методы pattern matching обоих типов (Levitsky et al., 2007): а) динуклеотидные весовые матрицы, в которых применяется протяженный размер окна, настраиваемый с помощью стандартной процедуры оценки точности; б) метод с учетом зависимостей частот динуклеотидов в различных районах ССТФ SiteGA. Далее примененные методы pattern matching будут обозначаться как PWM (а) и SiteGA (б).

В качестве метода класса pattern detection использован метод Chipmunk (Kulakovskiy et al., 2010). Метод использует результат эксперимента ChIP-Seq (выборка нуклеотидных последовательностей и профиль покрытия для каждой нуклеотидной позиции в этой выборке) для поиска мотива в виде частотной матрицы. В процессе поиска учитывается большая вероятность нахождения сайта в областях с большим покрытием. Результатом применения метода Chipmunk является выравнивание потенциальных сайтов, на основе которого строится

описанный выше метод весовой матрицы. В данной работе нами применено два варианта расчета – на основе моно- и динуклеотидных частот. Далее примененные методы *pattern matching* будут обозначаться как Mono- и DiChipmunk.

Для выбора порогов четырех методов распознавания мы составили выборку сайтов 64 ССТФ FoxA, предсказанных этими методами с достаточно низкими порогами в ChIP-Seq локусах (Wederell *et al.*, 2008) с максимальным покрытием локуса не менее 15, перекрывающих районы длины 1000 нуклеотидов выше аннотированных стартов транскрипции генов (промоторные районы генов). Порог максимального покрытия локуса подбирается эмпирически (перебором). Перебор основан на том, что (1) при низких значениях этого порога среди отобранных локусов заметную часть составляют локусы со слабым взаимодействием TF-ДНК и (вероятно) ошибками эксперимента, тогда как (б) при средних и высоких значениях порога эта доля гораздо меньше.

Доля таких «ошибок» определяется с помощью верифицированного заранее (отдельно от рассматриваемого эксперимента ChIP-Seq) метода распознавания ССТФ. Этот метод обучается на выборке экспериментальных подтвержденных ССТФ. Порог метода распознавания определяется либо с помощью процедуры оценки точности (Levitsky *et al.*, 2007), либо с помощью экспериментальной проверки большой выборки предсказанных сайтов. Решающим правилом для выбора порога покрытия является график зависимости доли локусов, содержащих сайты ChIP-Seq (ось Y) от покрытия (ось X). Эта зависимость обычно линейна при малых значениях покрытия, а затем число локусов выходит на плато и уже существенно не изменяется при росте покрытия. Порог определяется как точка перехода линейного роста на плато. Порог 15 был выбран на основе изложенных ранее соображений (зависимость доли локусов, содержащих сайты от покрытия). Промоторные районы были выбраны для поиска потенциальных сайтов, так именно в этих районах генома ожидается присутствие функциональных сайтов. Затем для каждого из 64 олигонуклеотидов составленной выборки проведена экспериментальная проверка аффинности с помощью метода EMSA (Electrophoresis mobility shift assay), пример такой проверки приведён на рис. 6. В каждом случае проверка заключается в оценке конкурентного вытеснения тестируемого олигонуклеотида по сравнению с олигонуклеотидом с известной аффинностью. На основе данных проверки EMSA для каждого из 64 проверенных олигонуклеотидов была определена относительная оценка аффинности «EMSA score».

Анализ распределения аффинностей всех 64 сайтов показал, что нет четкой границы между сайтами с высокой и низкой аффинностью.

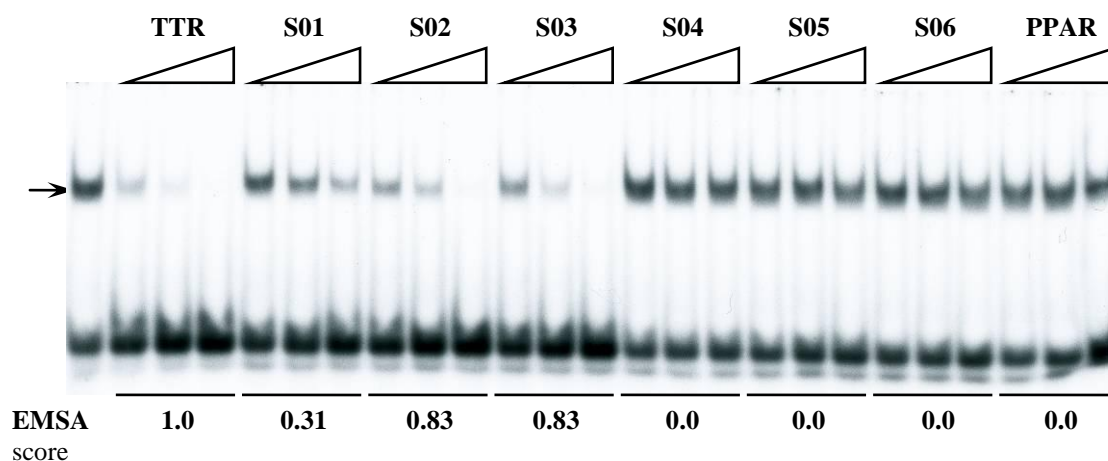


Рисунок 6. Конкуренция олигонуклеотидов в EMSA assay за рекомбинантный белок GST-FoxA2. В качестве пробы использован радиоактивно меченый олигонуклеотид TTR содержащий известный сайт связывания FoxA2. В качестве конкурента к пробам добавлялся возрастающий избыток немеченых олигонуклеотидов S01-S06 (2, 5 и 20 нг соответственно), содержащих предсказанные сайты связывания FoxA. Полоса, соответствующая ДНК-белковому комплексу, отмечена стрелкой. Немеченые олигонуклеотиды TTR и PPAR взяты в качестве положительного и отрицательного (PPAR) контролей. Внизу указан рассчитанный по результатам автордиографии EMSA score.

Выборка 64 сайтов далее была разделена на две группы: (1) 23 потенциальных сайта с «EMSA score» менее 0.25 составили выборку «не подтвержденных» сайтов; (2) оставшиеся 41 сайт («EMSA score» более 0.25) включили в выборку «подтвержденных» сайтов. Пороги методов распознавания PWM, SiteGA, Mono- и DiChipmunk были выбраны как максимальные значения соответствующих функций распознавания для 23 «неподтвержденных» сайтов.

После выбора оптимальных параметров и настроек для каждой из четырех указанных программ они были применены для распознавания реальных сайтов связывания фактора транскрипции FoxA2 в пиках ChIP-Seq эксперимента как для взрослых, так и для новорожденных мышей. Результаты работы данных программ приведены в Таблице 2.

Таблица 2. Результаты анализа ChIP-Seq эксперимента с антителами против FoxA2 для взрослых и новорожденных мышей для четырех программ распознавания сайтов связывания - PWM, SiteGA, Mono- и DiChipmunk.

Программа	Число сайтов FoxA2, Взрослые мыши	Число сайтов FoxA2, Новорожденные мыши
DiChipmunk	1924	2656
SiteGA	2770	3435
MonoChipmunk	2478	3741
PWM	1121	1601

Число сайтов FoxA2, выявленных разными программами, отличается. Наименьшее число сайтов было идентифицировано программой PWM. Программы SiteGA, Mono- и DiChirpmunk обладают примерно сходной эффективностью в выявлении сайтов FoxA2. При этом метод SiteGA, сравнимый по эффективности с обоими вариантами Chirpmunks, был построен на основе той же самой обучающей выборки из 64 сайтов FoxA. Однако наборы сайтов, обнаруженные разными программами, не являются простыми подмножествами и не полностью пересекаются. Этот эффект наглядно изображён на рис. 7. Поскольку применённые для анализа программы используют разные алгоритмы поиска сайтов связывания фактора FoxA2, то они более эффективно выявляют разные сайты, отличающиеся по своей структуре. Одновременное использование методов, относящихся к разным классам (например, SiteGA и DiChirpmunk) позволяет идентифицировать существенно больше сайтов связывания FoxA2 и соответственно ChIP-seq пиков, содержащих эти сайты.

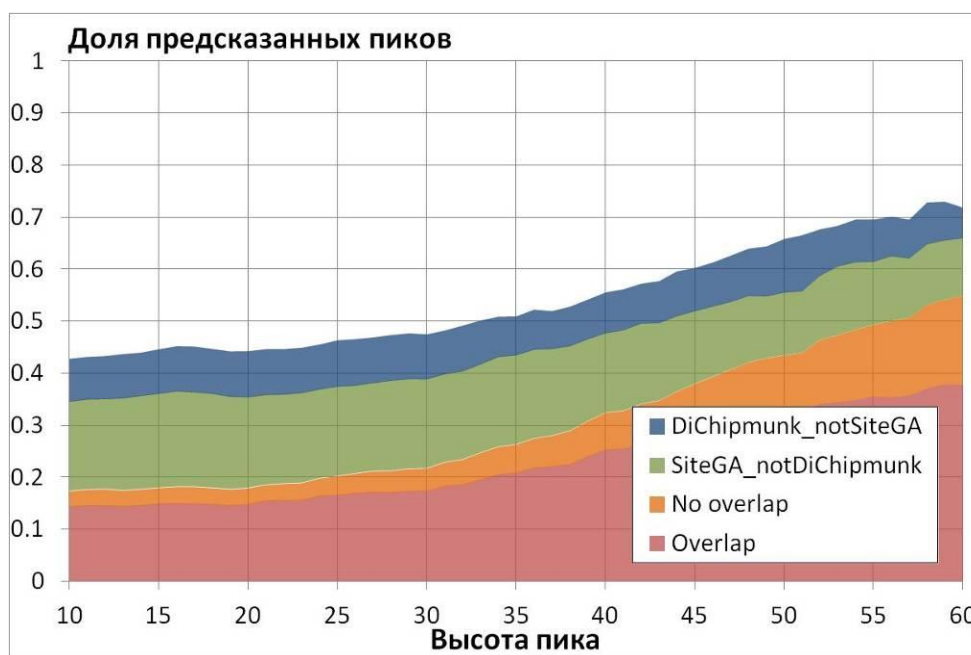


Рисунок 7. Доли сайтов связывания FoxA2, предсказанных программами DiChirpmunk и SiteGA в пиках ChIP-Seq в зависимости от высоты пика для новорождённых мышей. Розовым цветом показана доля сайтов, предсказанных обеими программами. Оранжевым – доля сайтов, обнаруженных в одном и том же пике, но не идентичных. Зелёным – доля сайтов, предсказанных только программой SiteGA, синим – только DiChirpmunk.

Далее был поведён анализ распределения обнаруженных нами в пиках ChIP-Seq сайтов связывания FoxA2 по различным районам генома. Для этого была взята выборка сайтов, полученная с помощью анализа, проведённого комбинацией программ DiChirpmunk и SiteGA. Для новорождённых мышей эта выборка состояла из 4639 сайтов, из них SiteGA дала 3435,

DiChirpmunk 2656 сайтов, 1452 сайтов было идентифицировано обеими программами. Аналогичный анализ для взрослых мышей дал 3219 сайтов, из них SiteGA обнаружила 2270 сайтов, DiChirpmunk 1924 сайтов, 975 сайтов было обнаружено обеими программами. Референсная геномная последовательность была разделена на следующие области:

- 1) Район ближнего промотора (5'-район в 1000 пар нуклеотидов от старта транскрипции)
- 2) Район первого интрона, часто обогащённый регуляторными элементами
- 3) Все остальные интроны, кроме первого
- 4) Экзоны
- 5) Межгенные районы, вся остальная часть референсной геномной последовательности с маскированными повторами.

Результат анализа распределения сайтов связывания FoxA2, найденных в ChIP-Seq пиках, по различным районам генома представлен на рис. 8. Суммарные результаты распределения достаточно близки, хотя имеются и определённые различия. Так, в межгенных районах расположено 44% сайтов FoxA2 у новорождённых мышей и 38% - у взрослых. Обращает внимание большая относительная доля сайтов связывания FoxA2, попадающая в район первого интрона (13,5% и 14,2% соответственно). Также нами был проведён сравнительный анализ доли тех генов, в промоторной области которых были обнаружены сайты FoxA2 внутри пиков ChIP-Seq. Оказалось, что в выборке таких генов у взрослых мышей 72% генов присутствуют как в выборке взрослых, так и новорождённых мышей. Оставшиеся 28% генов имеют связанный с промотором транскрипционный фактор FoxA2 только у взрослых мышей.

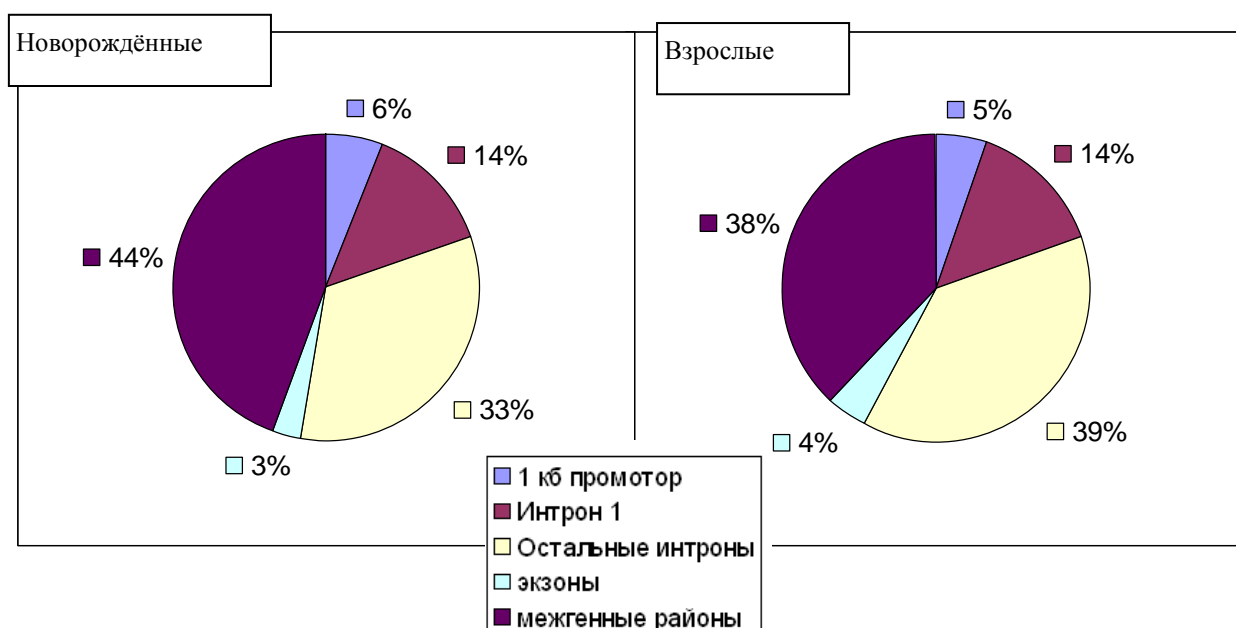


Рисунок 8. Диаграмма процентного распределения локализованных сайтов связывания FoxA2 в различных районах генома по результатам ChIP-Seq эксперимента. Слева – новорождённые мыши, справа – взрослые.

2. Исследование сайтов альтернативной инициации трансляции и их вклада в кодирующий потенциал мРНК генов стрессового ответа

На предыдущем этапе реализации проекта нами было показано, что в мРНК генов *Mus musculus*, трансляция которых активируется во время стресса, часто находятся лидерные рамки считывания (uORF). На основе этого факта мы предположили, что увеличение трансляционной активности при стрессе основано на механизме реинициации трансляции. На данном этапе проекта нами была проверена гипотеза о вкладе механизма реинициации трансляции в синтез альтернативных изоформ белков человека, а также о связи этих белков с молекулярными механизмами стрессового ответа.

Согласно некоторым экспериментальным данным, если в составе 5'-НТП мРНК расположена небольшая открытая рамка считывания (uORF), то после ее трансляции рибосомы сохраняют способность к реинициации. Считается, что эффективность реинициации в целом невелика и зависит от нескольких параметров мРНК: размера uORF и размера участка мРНК между стоп-кодоном uORF и следующим стартовым кодоном трансляции (Kozak, 2005; Jackson et al., 2010). По-видимому, размер uORF имеет значение потому, что за время трансляции небольшой рамки считывания с рибосомы не успевают отделиться некоторые белки – факторы инициации трансляции, присутствие которых позволяет 40S субъединице рибосомы оставаться связанной с мРНК и продолжить линейное сканирование. В процессе сканирования происходит до-сборка инициационного комплекса (связывание с комплексом eIF2-met-tRNA_i-АТФ), что позволяет восстановить способность 40S субъединицы к полноценной инициации трансляции. Предполагается, что такое восстановление требует определенного времени, поэтому если проксимальный кодон AUG расположен вблизи стоп-кодона uORF, то 40S субъединицы рибосом не успеют связать необходимые факторы, будут пропускать такие стартовые кодоны и смогут инициировать трансляцию только на нижерасположенных стартах. Нужно отметить, что скорость восстановления способности рибосом к реинициации трансляции зависит от активности факторов инициации трансляции, в том числе – eIF2 α (на этом принципе основан механизм стресс-специфического контроля экспрессии, детально описанный в обзоре литературы). Важно, что особенности реинициации трансляции и ее роль в контроле экспрессии генов остаются малоизученными. Экспериментально показано всего несколько примеров использования этого механизма для синтеза N-концевых изоформ белков (Kochetov, 2008). В частности, не известно, как тот или иной размер спейсера между uORF и следующей рамкой считывания связан с эффективностью реинициации, какова роль контекста старт- и стоп-кодонов, вторичной структуры мРНК и т.п. В настоящее время возможность реинициации трансляции не

учитывается при аннотации структурно-функциональной организации генов эукариот и предсказании их кодирующего потенциала.

Для этой части работы нами была выделена модельная выборка мРНК человека, блок-схема мРНК которых приведена на рис. 9.

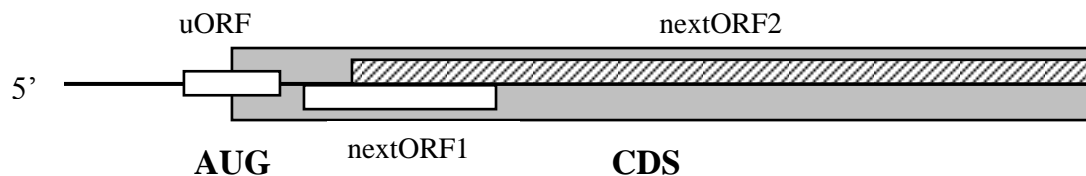


Рисунок 9. Структурно-функциональная организация мРНК человека, отобранных для исследования роли реинициации трансляции в синтезе N-концевых изоформ белков.

Согласно схеме, мРНК должны были содержать один uAUG и uORF должна была перекрываться с аннотированной белок-кодирующей последовательностью. В этом случае рибосомы способны реиницировать трансляцию на нижерасположенных AUG, что может приводить к синтезу укороченных с N-конца изоформ аннотированных белков. Рассматривался кодон AUG, ближайший к стоп-кодону uORF (рамка считывания, начинающаяся с такого потенциального стартового кодона, обозначалась как nextORF). В случае, если следующий за uORF кодон AUG был расположен в той же рамке считывания, что и аннотированная белок-кодирующая последовательность, nextORF представляла собой укороченный с 5'-конца вариант БКП (nextORF2 на рис. 9). В альтернативном случае nextORF будет представлять собой открытую рамку считывания, не связанную с БКП (обычно, небольшую: nextORF1 на рис. 9). В рамках работы планировалось исследовать взаимосвязь между uORF, nextORF, характеристиками спейсера и контекста соответствующих стартовых кодонов AUG.

Также определенные ограничения были наложены на возможную длину спейсера. Согласно экспериментальным данным (полученным в системе трансляции *in vitro*), реинициация наблюдалась при размере uORF 10-13 кодонов и ее уровень заметно снижался (но не до нуля) при увеличении размера uORF до 33 кодонов (Kozak, 2001). В другой работе реинициация наблюдалась при размере uORF равном 24 кодонам, но не при увеличении ее длины до 40 кодонов (Luukkonen et al., 1995). В данной работе были отобраны мРНК, у которых размер uORF варьировал от 12 до 93 нуклеотидов (т.е., от 3 до 30 кодонов).

Из БД EMBL были выделены 56754 карточки с кДНК *Homo sapiens*, из которых 44416 карточек содержали мРНК с полноразмерными белок-кодирующими последовательностями и 5'-НТП, размер которого варьировал от 15 до 1000 нуклеотидов. 43% этой выборки содержали

от одного до нескольких десятков uAUG, из них были отобраны 7220 мРНК с одной uORF. 754 негомологичных мРНК человека соответствовали схеме, изображенной на рис. 9. Эти мРНК рассмотрены в дальнейшем компьютерном анализе.

Найдено, что средний размер uORF составлял 52 ± 23 нуклеотидов (среднее значение и среднеквадратичное отклонение, соответственно). Средний размер спейсера между стоп-кодом uORF и следующим триплетом AUG составлял 68 ± 77 нуклеотидов, что – согласно экспериментальным данным – достаточно для реинициации трансляции (Luukkonen et al., 1995; Kozak, 2001; Calkhoven et al., 2000; 2003). Предполагалось, что одним из преимуществ реинициации перед «сканированием с подтеканием» является возможность доставки рибосом до следующей рамки считывания с пропуском всех триплетов AUG, расположенных в составе uORF («reinitiation shunt» (Kozak, 2001)). Действительно, 17% мРНК в модельной выборке характеризовалась наличием в uORF как минимум двух AUG (включая собственно стартовый кодон БКП). 297 мРНК (39% выборки) содержали nextORF, соответствующую укороченной с N-конца изоформе аннотированного белка.

Контексты стартовых кодонов трансляции uORF, БКП и nextORF приведены в таблице 3.

Таблица 3. Нуклеотиды, расположенные в ключевых позициях контекста стартового кодона трансляции uORF (uAUG), аннотированной белок-кодирующей последовательности (TIS) и nextORF (nextAUG).*

Контекст	uAUG,%	TIS,%	nextAUG,%
R ⁻³	49	80	60
R ⁻³ G ⁺⁴	14	27	18
Y ⁻³	51	20	40
Y ⁻³ W ⁺⁴	39	14	29

*R, пурин (A or G); Y, пиримидин (C or U); W, не G (A or C or U).

Можно было ожидать, что большинство uAUG будут располагаться в субоптимальном контексте, что позволит рибосомам транслировать как uORF, так и белок-кодирующую последовательность (за счет механизма «сканирование с подтеканием» (Kozak, 2005)). Однако, это оказалось верным для 51% uORF с вариантом контекста YnnAUG и для 39% uAUG с более выраженным субоптимальным вариантом контекста YnnAUGW (табл. 3). При этом 14% uAUG были расположены в оптимальном варианте контекста RnnAUGG, что – согласно существующим представлениям – должно практически полностью ингибировать трансляцию с аннотированного стартового кодона. Представляется вероятным, что структура этой группы мРНК предсказана неправильно и граница белок-кодирующей последовательности должна быть сдвинута в 5'- или (скорее) 3'-направлении. Аннотированный стартовый кодон был в

большинстве случаев расположен в оптимальном контексте. nextAUG, следующий за uORF был расположен в оптимальном контексте в 60% случаев (табл. 3).

Далее был проведен систематический анализ корреляций между характеристиками uORF, БКП и nextORF. Можно было предположить, что если часть nextORF, представляющих собой укороченные с N-конца варианты белок-кодирующих последовательностей, синтезируются *in vivo* и являются функционально-значимыми, то в ходе эволюции должны поддерживаться параметры мРНК, способствующие более эффективной реинициации трансляции. В то же время, если nextORF расположены в других рамках считывания и кодируют не связанные с аннотированным белком пептиды, их функциональная значимость намного менее вероятна (хотя и не может быть полностью исключена). Поэтому в анализ были отдельно взяты nextORF1 (out-frame) и nextORF2 (in-frame, рис. 9). Согласно полученным данным, между размером uORF и контекстом nextAUG существует достоверная негативная корреляция, но только в том случае nextORF2, кодирующих N-концевую изоформу белка ($r_k = -0.16$ ($p = 0.007$); $r_p = -0.13$ ($p = 0.026$); в случае nextORF1 $r_k = -0.013$; $r_p = -0.005$, значения не достоверны). Другая зависимость была найдена между размером спейсера между uORF и контекстом nextAUG: показано, что положительная достоверная корреляция существует в том случае, если nextORF кодирует N-концевую изоформу аннотированного белка ($r_k = 0.075$ ($p=0.05$); $r_p = 0.06$ (недостоверно)), в случае out-frame nextORF1 значения недостоверны ($r_k = 0.04$; $r_p = 0.007$). По-видимому, эти закономерности могут отражать функциональную значимость некоторых N-концевых изоформ аннотированных белков и – соответственно – эволюционную оптимизацию структурных элементов мРНК, связанных с реинициацией трансляции: более короткая uORF, более протяженный спейсер и более оптимальный контекст nextAUG способствуют более эффективному синтезу таких изоформ (Luukkonen et al., 1995; Kozak, 2001; 2005; Jackson et al., 2010).

В целом, модельная схема, примененная в данной работе, позволила провести приблизительную оценку того, используются ли uORF и механизм реинициации трансляции в сложной структуре эукариотических мРНК для синтеза N-концевых изоформ белков. В качестве индикатора «применимости» этих механизмов *in vivo* рассматривались взаимосвязи между параметрами uORF, БКП и nextORF (в разных рамках считывания), которые – в ходе эволюционного процесса – могли быть оптимизированы для повышения эффективности функционально-значимых событий реинициации трансляции. Были найдены некоторые статистически достоверные взаимосвязи, которые могут рассматриваться в качестве таких доказательств. Следует, однако, отметить, что значения коэффициентов корреляции довольно низкие, что может быть связано со сложностью формирования выборки для компьютерного анализа (ограниченной репрезентативностью): были взяты только мРНК с одним uAUG,

рассматривались только ближайшие к стоп-кодону uORF триплеты AUG, не брался в расчет контекст nextAUG. В частности, существенным недостатком использованной схемы является рассмотрение только ближайшего к uORF триплета AUG (nextAUG). Собственно одно из достоинств реинициации трансляции в качестве механизма, потенциально позволяющего расширять кодирующий потенциал эукариотических мРНК, заключается в возможности «адресовать» 40S субъединицы рибосом сразу на несколько удаленных нижерасположенных кодонов AUG, причем близкорасположенные стартовые кодоны могут пропускаться (обз: Kochetov, 2008). Однако, собрать репрезентативную выборку мРНК с потенциальными стартовыми кодонами такого типа затруднительно, поэтому была рассмотрена упрощенная ситуация.

В таблице 4 приведены данные о предсказанной субклеточной локализации 297 полноразмерных белков человека и их потенциальных N-концевых изоформ, трансляция которых может происходить с помощью механизма реинициации трансляции (предсказание было сделано с помощью программы *TargetP* (<http://www.cbs.dtu.dk/services/TargetP/>)). Рассматривались три возможных предсказанных локализации: митохондриальная (наличие митохондриального секреторного пептида), секреторная (наличие секреторного пептида, направляющего белок в мембрану или в экстраклеточное пространство), или «другая локализация».

Таблица 4. Сравнительный анализ субклеточной локализации белков человека и их потенциальных N-укороченных трансляционных изоформ (%)

Сигнал	БКП N-изоформа			
	Фракция	МТР	SP	Другая
МТР*	16	3	1	12
SP*	20	1	6	13
Другая*	64	8	3	53
Всего	100	12	10	78

*МТР, митохондриальный секреторный пептид; SP, секреторный пептид.

Можно видеть, что N-концевые изоформы в 25% случаев теряли секреторный характер (происходил переход из МТР или SP-вариантов в «Другую локализацию»), чего можно было ожидать при отсутствии N-концевого участка белка. 9% N-концевых изоформ сохраняли локализацию полноразмерного варианта белка; 11% N-концевых изоформ белков были предсказаны как МТР- или SP-содержащие, в то время как их полноразмерные варианты были отнесены программой *TargetP* к категории «Другая локализация» (табл. 4). В целом, 38%

N-концевых изоформ белков из модельной выборки характеризуются субклеточной локализацией, отличной от их полноразмерных вариантов.

Поскольку N-концевые изоформы белков могут отличаться не только субклеточной локализацией, в рамках этой части работы был также проведен дополнительный анализ их различий. Для этого был выделен N-концевой части полноразмерного белка, отсутствующий у N-концевой трансляционной изоформы (средний размер этого участка составил около 30 аминокислотных остатков; далее были использованы только сегменты, размер которых превышал 9 а.о.) и был проведен анализ их сходства с паттернами, собранными в базах данных ProSite (<http://www.expasy.org/prosite/>) и кластерами ProDom (<http://www.toulouse.inra.fr/prodom.html>). Анализ ProDom был проведен с помощью БД IPR (версия 8.1); сходство оценивалось с помощью однократной итерация модуля blastpgp пакета BLAST с параметрами по умолчанию (e-value = 0.01). Было найдено, что значительная часть сегментов полноразмерных белков, отсутствующих в N-концевых трансляционных изоформах, характеризуются сходством с кластерами ProDom (32) и паттернами ProSite (72 сегмента). Анализ БД ProSite показал, что большая часть найденных паттернов соответствует сайтам посттрансляционной модификации белков (ck2_phospho_site, myristil, pkc_phospho_site, asn_glycosylation), что – в целом – также может говорить о функциональных различиях между полноразмерными белками и их трансляционными N-концевыми изоформами.

Дополнительным свидетельством функциональной значимости N-концевых изоформ белков может служить эволюционная консервативность конфигурации мРНК, способствующей реинициации трансляции. Для оценки такой консервативности было проведено сравнение структуры 297 мРНК *Homo sapiens* и их гомологов из геномов *Mus musculus* и *Rattus norvegicus*. Для этого были выделены мРНК мыши и крысы (в качестве зонда для BLAST был использован участок мРНК человека, содержащий 5'-НТП и 300-нуклеотидный участок белок-кодирующей последовательности). Организация мРНК считалась консервативной, если: (1) позиции uORF, старта трансляции БКП и nextAUG в мРНК человека, мыши и крысы были идентичны; (2) позиции uORF, старта трансляции БКП и nextAUG в мРНК мыши и крысы отличались от мРНК человека, однако общая конфигурация оставалась сходной (то есть, размер uORF варьировал между 12 и 93 нуклеотидами, старт трансляции БКП был расположен внутри uORF и первый AUG в рамке считывания БКП был расположен вблизи позиции nextAUG в соответствующей мРНК человека). Всего были рассмотрены 91 пара мРНК человек-мышь и 61 вариант человек-мышь-крыса. Было найдено, что трансляционная конфигурация мРНК человека была полностью идентичной таковой в 70% мышинных гомологов и 62% мРНК и мыши, и крысы. При более мягких критериях сходства (вариант 2) мРНК человека и мыши были близки в 82% случаях, мРНК человека, мыши и крысы – в 79%

случаев. Следует обратить внимание на тот факт, что реинициация трансляции в общем случае не зависит от аминокислотной последовательности пептида, кодируемого uORF: только конфигурация uORF - старт-кодон БКП – nextAUG имеет значение. Можно предположить, что высокий уровень идентичности трансляционной конфигурации мРНК генов человека и их гомологов из геномов мыши и крысы также говорит в пользу функциональной значимости реинициации трансляции и N-концевых трансляционных изоформ.

Чтобы понять, какие функциональные классы представлены в модельной выборке из 297 белков человека, у которых укороченные с N-конца изоформы могут синтезироваться с помощью реинициации трансляции, был проведен их анализ с помощью Gene Ontology (GO; был использован GOblet Интернет-ресурс (Groth et al., 2004)). Было найдено, что 89 белков из 297 характеризуются терминами GO, связанными с “development”, “regulation of biological processes” и “response to stimuli”, то есть с регуляторными факторами, экспрессия которых строго контролируется. Интересно, что 17 мРНК в этой выборке содержат стартовый кодон uORF в особо оптимальном контексте (RnnAUGG), что позволяет выдвинуть предположение о том, что собственно аннотированный вариант либо не производится вообще, либо синтез имеет место в очень малых объемах. В целом, в составе модельной выборки из 297 мРНК человека были найдены важные регуляторные факторы, для которых – потенциально – могут с помощью реинициации трансляции дополнительно синтезироваться N-концевые трансляционные изоформы. В частности, можно отметить: nuclear receptor ERRG2 (human: AF094518, mouse: NM_011935.2; rat: AB126962.1), thymic stromal lymphopoietin protein TSLP (human: AF338732), caudal type homeobox transcription factor 2 (human: BC014461, mouse: NM_007673.2; rat: NM023963.1), ribosomal protein S6 kinase (human: BC036033; mouse: NM_028259.2), atrophin 1 (human: BC051795; mouse: NM_007881.4; rat: NM017228.2), death associated protein 3 (human: BC107488), thrombopoietin (human: D32047). Если учесть тот факт, что эффективность реинициации трансляции и – связанный с ней процесс выбора стартового кодона – изменяется в зависимости от статуса фосфорилирования фактора инициации eIF2 α , меняющегося при стрессе, то соотношение между аннотированной и укороченной с N-конца изоформы белка также может использоваться в регуляторных целях (например, ATF4 (Vattem and Wek, 2004), ATF5 (Zhou et al., 2008), C/EBP α and β (Calkhoven et al., 2003), SCL (Calkhoven et al., 2000)). В целом, в рамках этой части работы было предсказано около 300 генов человека, с мРНК которых могут дополнительно считываться укороченные с N-конца изоформы аннотированных белков, причем эффективность их синтеза может изменяться в условиях стресса.

3. Изучение онтогенетического паттерна экспрессии гена *DD2R* в тканях, метаболизирующих ювенильный гормон, у самок дрозофилы дикого типа и с генетической абляцией части клеток железы (*corpus allatum*), в которой гормон синтезируется.

В рамках этапа 5 настоящего проекта мы иммуногистохимически изучили: (1) экспрессию гена *DD2R* в *corpus allatum* (*CA*) молодых и половозрелых самок дикого типа *D. melanogaster*, (2) экспрессию гена *DD2R* в жировом теле (*ЖТ*) молодых и половозрелых самок дикого типа *D. melanogaster*, (3) влияние 20-гидроксиэкдизона (20Э) на экспрессию гена *DD2R* в жировом теле самок дикого типа *D. melanogaster* и (4) экспрессию гена *DD2R* в жировом теле молодых и половозрелых самок *D. melanogaster* с генетической абляцией части клеток *CA*.

Одной из гипотез, лежащих в основе настоящего проекта, было наше представление о том, что онтогенетический характер регуляции титра ЮГ дофамином (ДА) (повышение синтеза ювенильного гормона (ЮГ) и снижение деградации гормона при повышении уровня дофамина у молодых самок и - *vice versa* - у половозрелых самок) обусловлен переключением на геномном уровне и изменением числа ингибирующих (Д2-подобных) рецепторов дофамина в железе, синтезирующей ювенильный гормон, и ткани, синтезирующей ферменты его деградации. Для того, чтобы верифицировать гипотезу мы осуществили в рамках этапа 5 иммуногистохимический анализ экспрессии Д2-подобных рецепторов дрозофилы в *CA* и жировом теле молодых и половозрелых самок *D. melanogaster* с использованием кроличьих поликлональных антител против *DD2R*.

Для того, чтобы визуализировать морфологию *CA*, нами была сконструирована линия дрозофилы с генотипом *Aug21>::UAS-RFP*, у особей которой в клетках *CA* специфически экспрессируется красный флюоресцентный белок (RFP), что делает железу хорошо видимой при исследовании с помощью флюоресцентного микроскопа. На рисунке 10 представлены результаты иммуногистохимического анализа экспрессии *DD2R* в клетках *CA* молодых (1-суточных) и половозрелых (5-суточных) самок *Aug21>;Canton S;UAS-RFP*, полученных при скрещивании линий *Canton S* (дикого типа) и *Aug21>::UAS-RFP*.

Хорошо видно, что у половозрелых самок дрозофилы число Д2-подобных рецепторов (зеленый сигнал) в клетках *CA* (красный сигнал) значительно превышает таковое у молодых особей. Этот результат хорошо согласуется со сделанным нами ранее заключением о том, что повышение уровня ДА ингибирует синтез ЮГ у половозрелых самок дрозофилы (Gruntenko and Rauschenbach, 2008).

На рисунке 11 представлены результаты иммуногистохимического анализа экспрессии *DD2R* в клетках *ЖТ* молодых (1-суточных) и половозрелых (6-суточных) самок дикого

типа линии *Canton S*. Ясно видно, что число Д2-подобных рецепторов (зеленый сигнал) в ЖТ молодых самок в десятки раз превышает таковое у половозрелых особей.

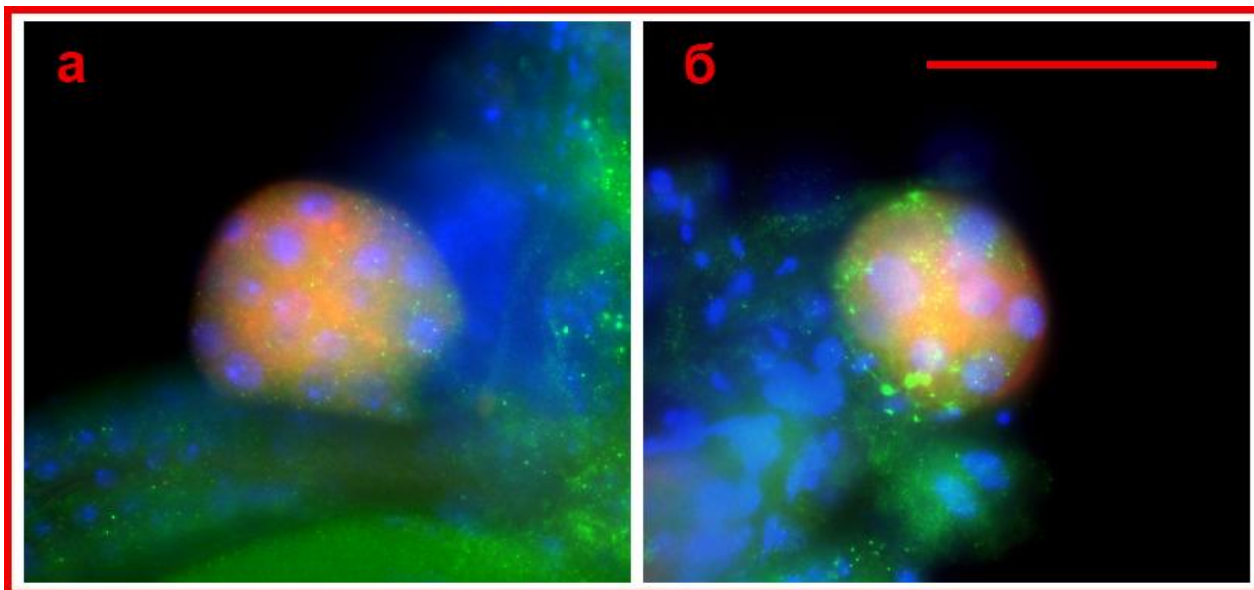


Рисунок 10. Число DD2R (зеленый сигнал) в клетках corpus allatum (красный сигнал) у: [а] 1-суточных и [б] 5-суточных самок *D. melanogaster* линии Aug21>;Canton S;UAS-RFP. Шкала = 50 μ m

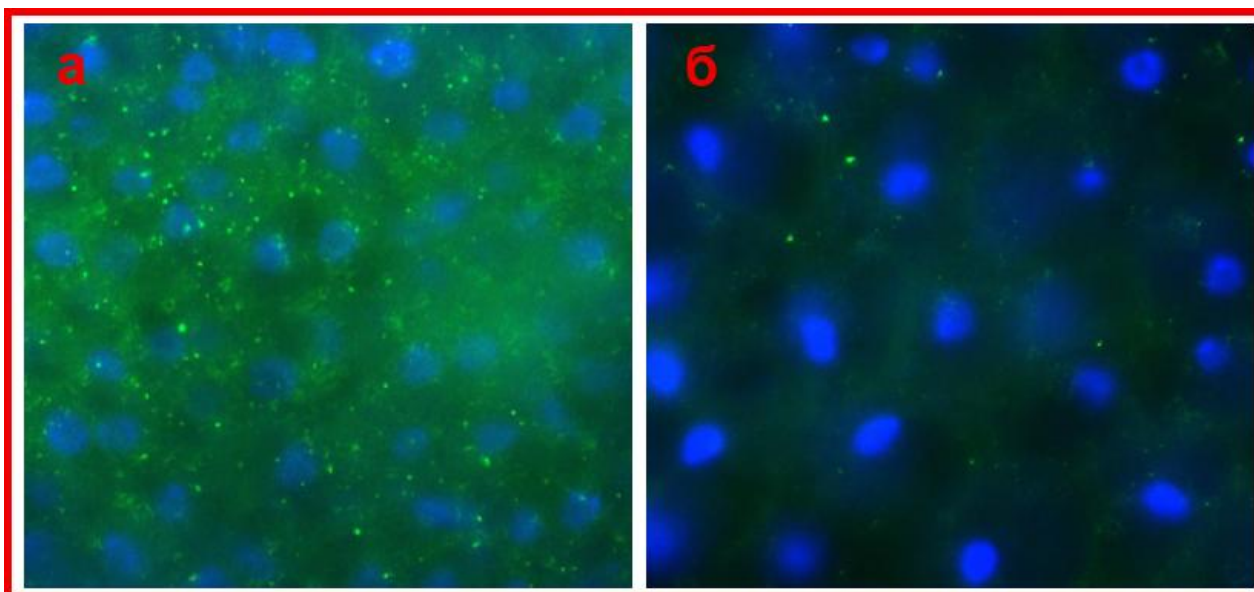


Рисунок 11. Число DD2R (зеленый сигнал) в клетках жирового тела у: [а] 1-суточных и [б] 6-суточных самок *D. melanogaster* линии Canton S. Шкала = 50 μ m

Таким образом, в рамках проекта было впервые продемонстрировано, что онтогенетический характер ингибирующего влияния ДА на синтез и деградацию ЮГ у *Drosophila* обусловлен переключением на геномном уровне, вызывающем изменение числа

дофаминовых D2-подобных рецепторов в железе, синтезирующей ЮГ (СА), и ткани, синтезирующей ферменты его деградации (ЖТ).

Возникает вопрос, каков механизм такого переключения? Мы предположили, что переключение может осуществляться 20-гидроксиэкдизоном, титр которого, как показали Харшмэн с соавторами (Harshman et al., 1999), низок у молодых самок дрозофилы и повышается при половом созревании. Для верификации этой гипотезы было необходимо исследовать влияние 20Э на экспрессию DD2R у молодых самок *Drosophila*.

Самок дикого типа *D. melanogaster* (линии Canton S) сразу после вылета помещали в стаканы, дно и 1 см стенок которых покрывали фильтровальной бумагой пропитанной 0,5 мл питательного раствора, содержащего 0,5% сахарозы, 0,2% дрожжей и 60 мкг 20Э. В контрольной серии 20Э в раствор не добавляли. Через сутки у подопытных и контрольных самок выделяли ЖТ и осуществляли его иммуногистохимический анализ с использованием кроличьих антител против DD2R. Результаты показаны на рисунке 12 (а,б). Хорошо видно, что у молодых (1-суточных) самок, обработанных 20Э (рис. 12 б), число DD2R (зеленый сигнал) резко падает, что свидетельствует в пользу нашей гипотезы.

Харшмэн с соавторами (Harshman et al., 1999) также показали, что без оплодотворения у самок дрозофилы повышения титра 20Э не происходит: у виргинных самок титр 20Э с возрастом не увеличивается. Следовательно, если наша гипотеза верна, паттерн экспрессии DD2R у 6-суточных виргинных самок должен соответствовать таковому у 1-дневных мух. Для того чтобы это проверить, мы осуществили иммуногистохимический анализ ЖТ 6-дневных оплодотворенных и виргинных самок линии Canton S. Результаты показаны на рисунке 12 (в, г). Ясно видно, что паттерн экспрессии DD2R у 6-суточных виргинных самок (рис. 12 г) соответствует паттерну неполовозрелых (1-суточных) самок (рис. 12 а).

Таким образом, в рамках проекта было впервые показано, что онтогенетическое изменение направленности ингибирующего влияния ДА на активность ферментов, деградирующих ЮГ в ЖТ самок дрозофилы, происходит в результате изменения уровня экспрессии гена D2-подобных рецепторов, обусловленного изменением титра 20Э после их оплодотворения.

Имуногистохимический анализ экспрессии гена DD2R в жировом теле молодых и половозрелых самок *D. melanogaster* с генетической абляцией части клеток *corpus allatum*.

В качестве модели использовали гомозиготную трансгенную линию *D. melanogaster* *UAS-reaper::UAS-hid* с рекомбинацией *UAS-reaper* и *UAS-hid* – двух активаторов клеточной смерти и линию *Aug21-Gal4/Cyo::arm-GFP*, несущую конструкцию *Aug21-Gal4* (*Aug21>*), экспрессирующуюся специфически в СА. Скрещивание *Aug21>/Cyo::arm-GFP* с гомозиготами *UAS-reaper::UAS-hid* дает 2 типа потомков: *Aug21>UAS-reaper::UAS-hid*, с генетической абляцией, как мы показали на этапе 2 настоящего проекта, части клеток СА и

снижением синтеза ЮГ, и *Cyo::arm-GFP;UAS-reaper::UAS-hid* (*reaper* и *hid* без драйвера *Aug21-GA14*> не экспрессируются).

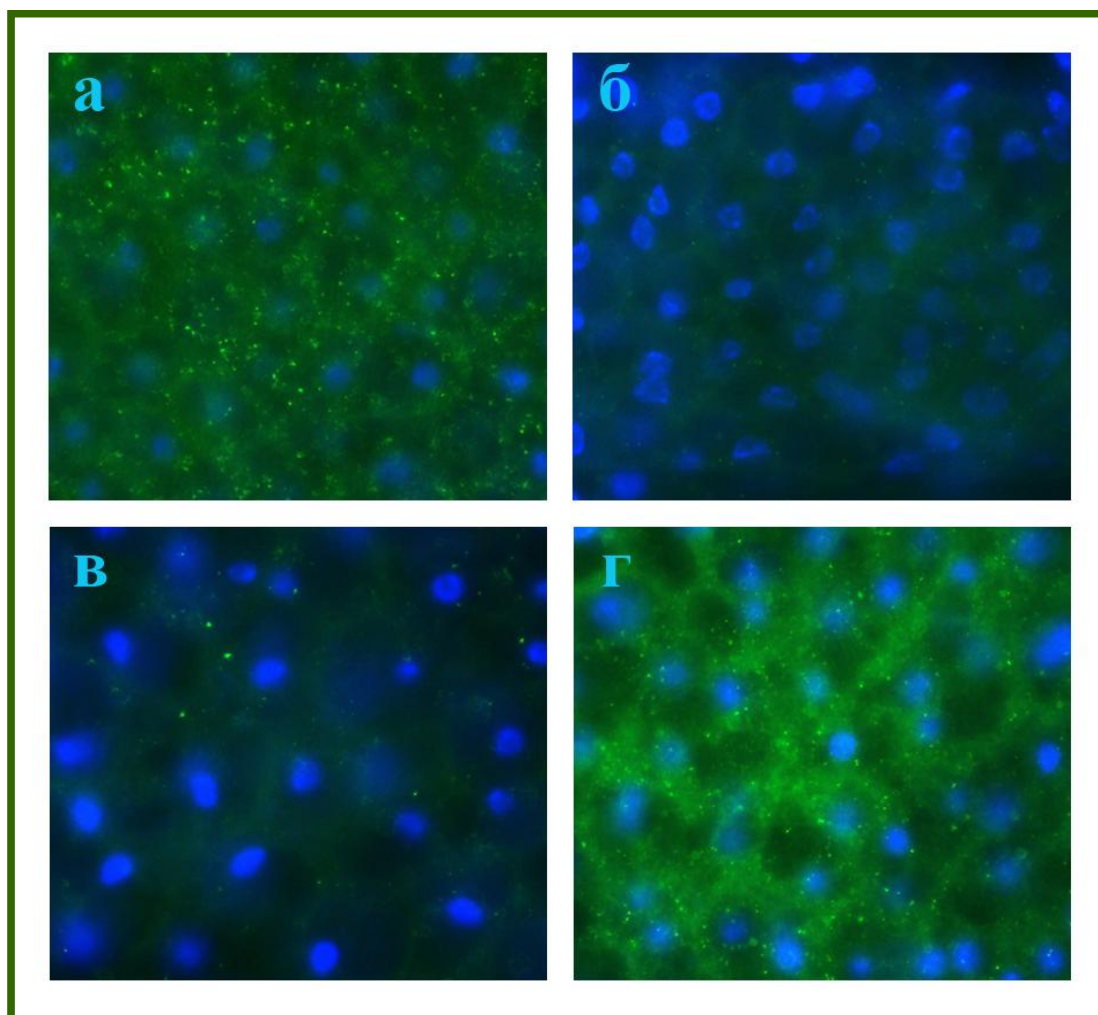


Рисунок 12. Число DD2R (зеленый сигнал) в клетках жирового тела у самок *D. melanogaster* линии Canton S (а) 1-суточных контрольных, (б) – 1-суточных, обработанных 20Э, (в) 6-суточных оплодотворенных, (г) - 6-суточных виргинных. Синий сигнал – ядра клеток, окрашенные DAPI (данные группы из ИЦиГ СО РАН).

На настоящем этапе нами был осуществлен иммуногистохимический анализ уровня экспрессии гена DD2R в жировом теле молодых и половозрелых самок с генотипом *Aug21>UAS-reaper::UAS-hid* - для того, чтобы выяснить, опосредуется ли обнаруженное нами повышение уровня деградации ювенильного гормона (Gruntenko et al., 2010) у этих самок изменением паттерна экспрессии DD2R в ЖТ. На рисунке 13 представлены результаты иммуногистохимического анализа экспрессии DD2R в клетках ЖТ молодых (1-суточных) и половозрелых (6-суточных) самок *Aug21>UAS-reaper::UAS-hid*. Ясно видно, что паттерн экспрессии DD2R в ЖТ самок с генетической абляцией CA (рис. 13) не отличается от такового у самок дикого типа (рис. 11).

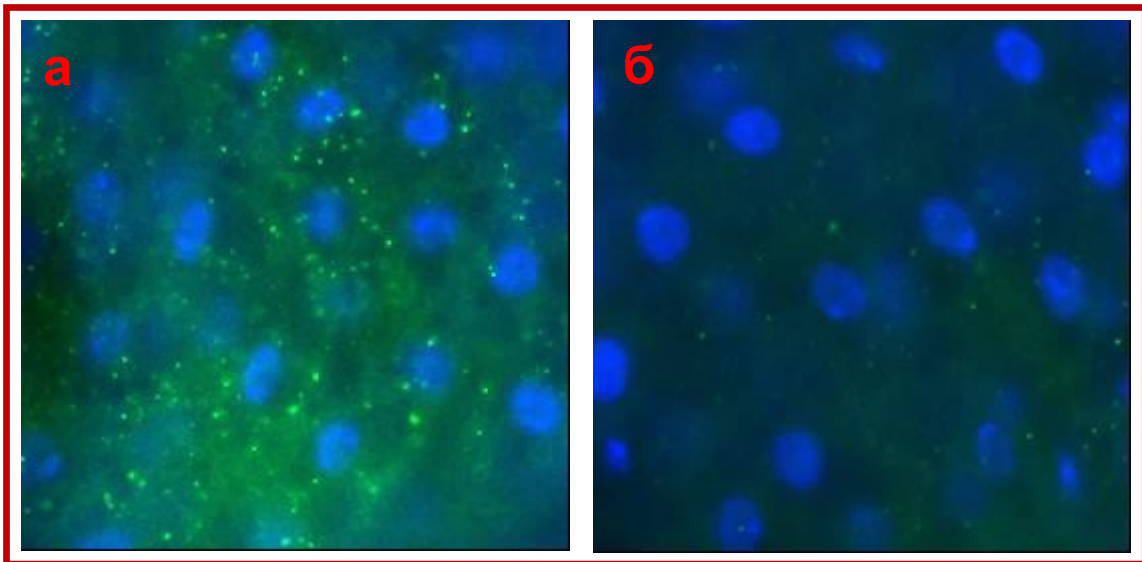


Рисунок 13. Число DD2R (зеленый сигнал) в клетках жирового тела у: [а] 1-суточных и [б] 6-суточных *Aug21>UAS-reaper::UAS-hid* самок *D. melanogaster*. Шкала = 50 μm

Таким образом, дофаминовый сигнальный путь не опосредует влияние уровня синтеза ЮГ на активность деградирующих его ферментов.

ЗАКЛЮЧЕНИЕ

В ходе работ, проведенных на 5-ом этапе реализации проекта:

1. Проведён компьютерный анализ ранее полученных данных по иммунопреципитации хроматина *in vivo* (ChIP-Seq) с антителами против транскрипционных факторов FoxO3A, FoxA2, связанных с ответом на ксенобиотики и онкогенезом. Обработаны данные контрольных экспериментов с использованием антител против суммарного IgG мыши и гистона H3. С помощью экспериментальной верификации проведена оптимизация параметров четырёх программ распознавания сайтов связывания FoxA2 - PWM, SiteGA, MonoChIPmunk и DiChIPmunk. Такая оптимизация позволила провести высокодостоверный анализ локализации сайтов связывания FoxA2 внутри пиков ChIP-Seq, а также подробно изучить профиль полногеномного распределения сайтов связывания транскрипционного фактора FoxO3A для взрослых и новорождённых особей мышей.
2. Проведено исследование роли механизма реинициации трансляции в синтезе альтернативных изоформ белков человека. Обнаружено, что при определенной конфигурации 5'-НТП, содержащей лидерные рамки считывания, с мРНК человека с высокой степенью вероятности могут дополнительно считываться дополнительные N-концевые изоформы белков. По сравнению с аннотированными полноразмерными вариантами, эти изоформы часто характеризуются иной субклеточной локализацией, а также отсутствием сайтов посттрансляционных модификаций, что свидетельствует об их функциональных различиях. Известно, что эффективность реинициации трансляции зависит от статуса фосфорилирования фактора инициации трансляции eIF2 α , который – в свою очередь – изменяется в ответ на стресс. В рамках этой части работы было предсказано около 300 генов человека, с мРНК которых могут дополнительно считываться укороченные с N-конца изоформы аннотированных белков, причем эффективность их синтеза может изменяться в условиях стресса.
3. В отчетный период (1) осуществлен иммуногистохимический анализ *corpus allatum* молодых и половозрелых самок дикого типа *D. melanogaster* с использованием кроличьих антител против DD2R; (2) осуществлен иммуногистохимический анализ жирового тела молодых и половозрелых самок *D. melanogaster* дикого типа и с генетической абляцией части клеток *corpus allatum*; (3) исследовано влияние 20Э на экспрессию гена DD2R. Впервые показано, что онтогенетический характер ингибирующего влияния ДА на синтез и деградацию ЮГ у *Drosophila* обусловлен переключением на геномном уровне, вызывающем изменение числа дофаминовых D2-подобных рецепторов в железе, синтезирующей ЮГ, и ткани, синтезирующей ферменты его деградации; Обнаружено, что это переключение обусловлено изменением титра 20-гидроксиэкдизона после оплодотворения самок.

Таким образом, все задачи, поставленные на данном этапе решены. По результатам работы за период четвертого этапа было опубликовано четыре статьи в ведущих российских и зарубежных журналах (всего за время выполнения проекта были опубликованы 15 статей, из них 7 – в ведущих зарубежных журналах). В исследованиях активное участие принимали молодые ученые, в том числе аспиранты и студенты старших курсов. Полученные результаты уже используются в учебном процессе (курсах молекулярной биологии, компьютерной геномики; спецкурсах «Основные молекулярно-генетические процессы», «Новейшие молекулярно-генетические технологии»). За период четвертого этапа защищены три кандидатская диссертация (всего за время выполнения проекта – 8 кандидатских и одна докторская диссертации).

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Брызгалов Л.О., Ершов Н.И., Ощепков Д.Ю., Каледин В.И., Меркулова Т.И., Выявление генов-мишеней транскрипционных факторов FOXA, связанных с регуляцией пролиферации. Биохимия. 2008. 73. 70 - 75.
2. Albanese V., Biguet N.F., Kiefer H., Bayard E., Mallet J., Meloni R., Quantitative effects on gene silencing by allelic variation at a tetranucleotide microsatellite. Hum. Mol. Genet. 2001. 10. 1785–1792.
3. Arteaga M.F., Alvarez de la Rosa D., Alvarez J.A., Canessa C.M. Multiple translational isoforms give functional specificity to serum- and glucocorticoid-induced kinase 1. Mol. Biol. Cell. 2007. 18. 2072-2080.
4. Bazykin G.A., Kochetov A.V. Alternative translation start sites are conserved in eukaryotic genomes. Nucleic Acids Res. 2011. 39. 567-577.
5. Blenau W., Baumann A. Molecular and pharmacological properties of insect biogenic amine receptors: lessons from *Drosophila melanogaster* and *Apis mellifera*. Arch. Insect Biochem. Physiol. 2001. 48. 13-38.
6. Calkhoven C.F., Muller C., Leutz A. Translational control of C/EBPa and C/EBPb isoform expression. Genes Dev. 2000. 14. 1920-1932.
7. Calkhoven C.F., Muller C., Martin R., Krosi G., Pietsch H. Translational control of SCL-isoform expression in hematopoietic lineage choice. Genes Dev. 2003. 17. 959-964.
8. Cazzola M., Skoda R.C. Translational pathophysiology: a novel molecular mechanism of human disease. Blood. 2000. 95. 3280-3288.
9. Chatterjee S., Pal J.K. Role of 5' - and 3' -untranslated regions of mRNAs in human diseases. Biol. Cell. 2009. 101. 251-262.
10. Cheng T.H., Cohen S.N. Human MDM2 isoforms translated differentially on constitutive versus p53-regulated transcripts have distinct functions in the p53/MDM2 and TSG101/MDM2 feedback control loops. Mol. Cell. Biol. 2007. 27. 111-119.
11. Courtois S., Verhaegh G., North S., Luciani M.G., Lassar P. DeltaN-p53, a natural isoform of p53 lacking the first transactivation domain, counteracts growth suppression by wild-type p53. Oncogene. 2002. 21. 6722–6728.
12. Draper I., Kurshan P.T., McBride E., Jackson F.R., Kopin A.S. Activity is regulated by D2-like receptors in *Drosophila*: an anatomic and functional analysis. Dev. Neurobiol. 2007. 67. 378-393.

13. Elnitski L., Elnitski L., Jin V.X., Farnham P.J., Jones S.J. Locating mammalian transcription factor binding sites: a survey of computational and experimental techniques. *Genome Res.* 2006. 16. 1455–1464.
14. Farnham P.J. Insights from genomic profiling of transcription factors. *Nat. Rev. Genet.* 2009. 10. 605-616.
15. Gangwal K., Close D., Enriquez C. A., Hill C. P., Lessnick S. L. Emergent properties of EWS/FLI regulation via GGAA microsatellites in Ewing's sarcoma. *Genes Cancer.* 2010. 1. 177–187.
16. Granger N.A., Sturgis S.L., Ebersohl R., Geng C., Sparks T.C. Dopaminergic control of corpora allata activity in the larval tobacco hornworm, *Manduca sexta*. *Arch. Insect Biochem. Physiol.* 1996. 32. 449-466.
17. Groth D., Lehrach H., Henning S. GOBlet: a platform for Gene Ontology annotation of anonymous sequence data. *Nucleic Acids Res.* 2004. 32. 313-317.
18. Gruntenko N.E., Chentsova N.A., Andreenkova E.V., Bownes M., Segal D., Adonyeva N.V. and Rauschenbach I.Yu. Stress response in a juvenile hormone deficient *Drosophila melanogaster* mutant apterous56f. *Insect Mol. Biol.* 2003. 12. 353-363.
19. Gruntenko N.E., Rauschenbach I.Yu. Interplay of juvenile hormone, 20-hydroxyecdysone and biogenic amines under normal and stress conditions and its effect on reproduction. *J Insect Physiol.* 2008. 54. 902-908.
20. Gruntenko N., Rauschenbach I. 20-hydroxyecdysone, Juvenile Hormone and Biogenic Amines: Mechanisms of Interaction in Control of *Drosophila* Reproduction under Normal and Stressful Conditions. In: *Ecdysone, structure and function* (ed. G. Smagghe). Springer Science + Business Media B.V. 2009. p.317-334.
21. Gruntenko N.E., Wen D., Karpova E.K., Adonyeva N.V., Liu Y., He Q., Faddeeva N.V., Fomin A.S., Li S., Rauschenbach I.Yu. Altered juvenile hormone metabolism, reproduction and stress response in *Drosophila* adults with genetic ablation of the corpus allatum cells. *Insect Biochem. Mol. Biol.* 2010. 40. 891-897.
22. Handler A.M. Ecdysteroid titers during pupal and adult development in *Drosophila melanogaster*. *Dev. Biol.* 1982. 93. 73-82.
23. Harshman L.G., Loeb A.M., Johnson B.A. Ecdysteroid titers in mated and unmated *Drosophila melanogaster* females. *J. Insect Physiol.* 1999. 45. 571-577.
24. Hearn M., Ren Y., McBride E., Reveillaud I., Beinborn M., Kopin A.A. *Drosophila* dopamine 2-like receptor: molecular characterization and identification of multiple alternatively spliced variants. *Proc. Natl. Acad. Sci. U.S.A.* 2002. 99, 14554–14559.

25. Hirashima A., Raushenbach I.Yu., Sukhanova M. Jh. Ecdysteroids in stress responsive and nonresponsive *Drosophila virilis* lines under stress conditions // Biosci. Biotech. Biochem. 2000. 64. 2657-2662.
26. Hirashima A., Suetsugu E., Hirokado S., Kuwano E., Taniguchi E., Eto M. Effect of octopamine on the activity of juvenile-hormone esterase in the silkworm *Bombyx mori* and the red flour beetle *Tribolium freemani*. Gen. Comp. Endocrinol. 1999. 116. 373-381.
27. Jacobson E.M., Concepcion E., Oashi T., Tomer Y. Efficiency of CD40 gene translation: a case of translational pathophysiology. Endocrinology. 2005. 146. 2684-2691.
28. Jackson R.J., Hellen C.U., Pestova T.V. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell Biol. 2010. 11. 113-127.
29. Ivanov I.P., Atkins J.F., Michael A.J. A profusion of upstream open reading frame mechanisms in polyamine-responsive translational regulation. Nucleic Acids Res. 2010. 38. 353-359.
30. Kaatz H., Eichmuller S., Kreissl S. Stimulatory effect of octopamine on juvenile hormone biosynthesis in honey bees (*Apis mellifera*): physiological and immunocytochemical evidence. J. Insect Physiol. 1994. 40. 856-872.
31. Kochetov A.V., Ischenko I.V., Vorobiev D.G., Kel A.E., Babenko V.N., Kisselev L.L., Kolchanov N.A. Eukaryotic mRNAs encoding abundant and scarce proteins are statistically dissimilar in many structural features. FEBS Lett. 1998. 440. 351-355.
32. Kochetov A.V. AUG codons at the beginning of protein coding sequences are frequent in eukaryotic mRNAs with a suboptimal start codon context. Bioinformatics. 2005. 21. 837-840.
33. Kochetov A.V. Alternative translation and hidden coding potential of eukaryotic mRNAs. BioEssays. 2008. 30. 683-691.
34. Kozak M. Constrains on reinitiation of translation in mammals. Nucleic Acids Res. 2001. 29. 5226-5232
35. Kozak M. Regulation of translation *via* mRNA structure in prokaryotes and eukaryotes. Gene. 2005. 361. 13-37.
36. Kulakovskiy I.V., Boeva VA, Favorov AV, Makeev VJ. Deep and wide digging for binding motifs in ChIP-Seq data. Bioinformatics. 2010. 26. 2622-2623.
37. Lafon-Cazal M., Baehr J.C. Octopaminergic control of *corpora allata* activity in an insect // Experientia. 1988. 44. 895-896.
38. Le Quesne J.P., Spriggs K.A., Bushell M., Willis A.E. Dysregulation of protein synthesis and disease. J. Pathol. 2010. 220. 140-151.
39. Levitsky V.G., Ignatieva E.V., Ananko E.A., Turnaev I.I., Merkulova T.I., Kolchanov N.A., Hodgman T.C. Effective transcription factor binding site prediction using a combination of

- optimization, a genetic algorithm and discriminant analysis to capture distant interactions. *BMC Bioinf.* 2007. 8. 481.
40. Luukkonen B.G.M., Tan W., Schwartz S. Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. *J. Virol.* 1995. 69. 4086–4094.
 41. Malavasi E.L., Kelly V., Nath N., Gambineri A., Dakin R.S., Pagotto U., Pasquali R., Walker B.R., Chapman K.E. Functional effects of polymorphisms in the human gene encoding 11 beta-hydroxysteroid dehydrogenase type 1 (11 beta-HSD1): a sequence variant at the translation start of 11 beta-HSD1 alters enzyme levels. *Endocrinology.* 2010. 151. 195-202.
 42. Malone B.M., Tan F., Bridges S.M., Peng Z. Comparison of four ChIP-Seq analytical algorithms using rice endosperm H3K27 trimethylation profiling data. *PLoS One.* 2011. 6. e25260.
 43. Missale C., Nash S. R., Robinson S. W., Jaber M., Caron M.G. Dopamine receptors: from structure to function. *Physiol. Rev.* 1998. 78. 189-225.
 44. Motallebipour M., Ameer A., Reddy Bysani M.S., Patra K., Wallerman O., Mangion J., Barker M.A., McKernan K.J., Komorowski J., Wadelius C. Differential binding and co-binding pattern of FOXA1 and FOXA3 and their relation to H3K4me3 in HepG2 cells revealed by ChIP-Seq. *Genome Biol.* 2009. 10. R129.
 45. Neckameyer W.S., Leal S.M. Biogenic amines as circulating hormones in insects. In: Plaff D.W., Arnold A.P., Etgen F.M., Fahrbach S.E. and Rubin R.T. (Eds) *Hormones, brain and behavior.* V.3. Academic Press. San Diego. USA. 2002. p. 141-165.
 46. Pastor D., Piulachs M.D., Cassier P., Andre M., Belles X. *In vivo* and *in vitro* study of the action of dopamine on oocyte growth and juvenile hormone production in *Blattella germanica*. *C. R. Acad. Sci. III.* 1991. 313. 207-212.
 47. Postlethwait J.H., Parker J. Regulation of vitellogenesis in *Drosophila* // *Molecular biology of invertebrate development.* Eds O'Connor J. D., Alan R. New York: Liss Inc. 1987. 29-42.
 48. Rachinsky A. Octopamine and serotonin influence on corpora allata activity in honey bee (*Apis mellifera*) larvae. *J. Insect Physiol.* 1994. 40. 549-554.
 49. Rahmani F., Hummel M., Schuurmans J., Wiese-Klinkenberg A., Smeekens S., Hanson J. Sucrose control of translation mediated by an upstream open reading frame-encoded peptide. *Plant Physiol.* 2009. 150. 1356-1367.
 50. Rauschenbach I.Y., Khlebodarova T.M., Chentsova N.A., Gruntenko N.E., Grenback L.G., Yantsen E.I., Filipenko M.L. Metabolism of the juvenile hormone in *Drosophila* adults under normal conditions and heat stress. *J. Insect Physiol.* 1995. 41. 179-189.

51. Rogozin I.B., Kochetov A.V., Kondrashov F.A., Koonin E.V., Milanezi L. Presence of ATG triplets in 5' untranslated regions of eukaryotic cDNAs correlates with a "weak" context of the start codon. *Bioinformatics*. 2001. 17. 890-900.
52. Soller M., Bownes M., Kubli E. Control of oocyte maturation in sexually mature *Drosophila* females. *Dev. Biol.* 1999. 208. 337-351.
53. Spevak C.C., Ivanov I.P., Sachs M.S. Sequence requirements for ribosome stalling by the arginine attenuator peptide. *J. Biol. Chem.* 2010. 285. 40933-40942.
54. Suzuki Y., Holmes J.B., Cerritelli S.M., Sakhuja K., Minczuk M., Holt I.J., Crouch R.J. An upstream open reading frame and the context of the two AUG codons affect the abundance of mitochondrial and nuclear RNase H1. *Mol. Cell Biol.* 2010. 30. 5123-5134.
55. Toivonen J.M, Partridge L. Endocrine regulation of aging and reproduction in *Drosophila*. *Mol. Cell. Endocrinol.* 2009. 299. 39–50.
56. Truman J.W., Riddiford L.M. The morphostatic actions of juvenile hormone. *Insect Biochem. Mol. Biol.* 2007. 37. 761-770.
57. Waddell S. Dopamine reveals neural circuit mechanisms of fly memory. *Trends Neurosci.* 2010. 33. 457-464.
58. Wallerman O., Motallebipour M., Enroth S., Patra K., Bysani M.S., Komorowski J., Wadelius C. Molecular interactions between HNF4a, FOXA2 and GABP identified at regulatory DNA elements through ChIP-sequencing. *Nucleic Acids Res.* 2009. 37. 7498-7508.
59. Wederell E.D., Bilenky M., Cullum R., Thiessen N., Dagpinar M., Delaney A., Varhol R., Zhao Y., Zeng T., Bernier B., Ingham M., Hirst M., Robertson G., Marra M.A., Jones S., Hoodless P.A. Global analysis of in vivo Foxa2-binding sites in mouse adult liver using massively parallel sequencing. *Nucleic Acids Res.* 2008. 36. 4549-4564.
60. Wethmar K., Smink J.J., Leutz A. Upstream open reading frames: molecular switches in (patho)physiology. *Bioessays*. 2010. 32. 885-893.
61. Woodring J., Hoffmann K.H. The effects of octopamine, dopamine and serotonin on juvenile hormone synthesis, in vitro, in the cricket, *Gryllus bimaculatus*. *J. Insect Physiol.* 1994. 40. 797-802.
62. Vallone D., Picetti R., Borrelli E. Structure and function of dopamine receptors. *Neurosci. Biobehav. Rev.* 2000. 24. 125-132.
63. Van Swinderen B., Andretic R. Dopamine in *Drosophila*: setting arousal thresholds in a miniature brain. *Proc. Royal Soc. B.* 2011. 278. 906-913.
64. Vattam A.M., Wek R.C. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc. Natl. Acad. Sci. USA.* 2004. 101. 11269-11274

65. Yudit M.R., Cidlowski J.A. Molecular identification and characterization of a and b forms of the glucocorticoid receptor. *Mol. Endocrinol.* 2001. 15. 1093-1103.
66. Zhang Y., Liu T., Meyer C.A., Eeckhoute J., Johnson D.S., Bernstein B.E., Nusbaum C., Myers R.M., Brown M., Li W., Liu X.S. Model-based Analysis of ChIP-Seq ([MACS](#)) // *Genome Biol.* 2008. 9. R137.
67. Zhou D., Palam L.R., Jiang L., Narasimhan J., Staschke K.A., Wek R.C. Phosphorylation of eIF2 directs ATF5 translational control in response to diverse stress conditions. *J. Biol. Chem.* 2008. 283. 7064-7073.