

На правах рукописи



МАКСИМОВ ДАНИИЛ АЛЕКСАНДРОВИЧ

**РАЗЛИЧИЯ В ХРОМОСОМНОЙ ЛОКАЛИЗАЦИИ БЕЛКА SUUR В
РАЗВИТИИ *DROSOPHILA MELANOGASTER* КОРРЕЛИРУЮТ С
АКТИВНОСТЬЮ ГЕНОВ И СОСТОЯНИЕМ ХРОМАТИНА**

03.01.07 – молекулярная генетика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата биологических наук

Новосибирск – 2014

Работа выполнена в лаборатории геномики Федерального государственного бюджетного учреждения науки Института молекулярной и клеточной биологии Сибирского отделения Российской академии наук, г. Новосибирск.

Научный руководитель:

Белякин Степан Николаевич

кандидат биологических наук, заведующий лабораторией геномики, Федеральное государственное бюджетное учреждение науки Институт молекулярной и клеточной биологии СО РАН, г. Новосибирск

Официальные оппоненты:

Высоцкая Людмила Васильевна

доктор биологических наук, профессор кафедры цитологии и генетики Новосибирского государственного университета, г. Новосибирск

Орищенко Константин Евгеньевич

кандидат биологических наук, заведующий сектором клеточных технологий Института цитологии и генетики СО РАН, г. Новосибирск

Ведущая организация:

Федеральное государственное бюджетное учреждение науки Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, г. Новосибирск

Защита состоится: _____ 2014 г. в ____ часов на заседании диссертационного совета Д 003.074.01 по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук при Федеральном государственном бюджетном учреждении науки Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук (630090, Новосибирск, пр. академика Лаврентьева, 8/2, тел. (383)-363-90-45) Тел: +7- 952-916-7858. e-mail: kokoza@mcb.nsc.ru С диссертацией можно ознакомиться в библиотеке и на сайте Института молекулярной и клеточной биологии СО РАН www.mcb.nsc.ru

Автореферат разослан « ____ » _____ 2014г.

Ученый секретарь
диссертационного совета,
кандидат биологических наук

Е. Б. Кокоза

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы. Хромосомы высших эукариот состоят из протяженных доменов, служащих для обеспечения правильной активности генов в клетке. Эти домены поддерживают активное или репрессированное состояние расположенных в них генов, что также коррелирует с другими аспектами их функционирования. Так, известно, что репликация активных участков хромосом обычно проходит в начале S-фазы клеточного цикла, тогда как неактивные районы реплицируются в конце S-фазы, однако механизм, лежащий в основе этого явления, остается непонятым (Schubeler *et al.*, 2002; MacAlpine *et al.*, 2004).

Для разных групп организмов была установлена закономерность: районы, проявляющие транскрипционную активность в большинстве клеточных типов, характеризуются повышенной плотностью генов. В участках, обогащенных тканеспецифично экспрессирующимися генами, этот показатель обычно снижен, и в большинстве клеточных типов такие районы формируют неактивные домены (Huvet *et al.*, 2007; Hiratani *et al.*, 2008). Эти наблюдения говорят о том, что особенности расположения генов в геноме и регуляция хроматиновых доменов каким-то образом связаны между собой. При этом следует понимать, что расположение генов во всех клетках организма одинаково, в то время как состояние хроматина может изменяться в разных клеточных типах. Вероятно, сниженная плотность генов в районах, обогащенных тканеспецифичными генами, имеет отношение к регуляции их активности в процессе развития.

Политенные хромосомы слюнных желез личинок *Drosophila melanogaster* содержат набор транскрипционно неактивных участков, характеризующихся недорепликацией, которая, по-видимому, является следствием их поздней репликации (Zhimulev *et al.*, 1982). Недорепликация проявляется в том, что эти районы содержат значительно меньшее число нитей ДНК, чем окружающие их участки политенных хромосом. Длина районов недорепликации может достигать нескольких сотен т.п.н. (Belyakin *et al.*, 2005; Sher *et al.*, 2011).

На небольшой выборке генов было показано, что эти районы могут быть обогащены тканеспецифичными генами, в частности, избирательно активными в семенниках (Belyakin *et al.*, 2005). Таким образом, районы недорепликации в политенных хромосомах *D. melanogaster* проявляют признаки, делающие их похожими на домены, описанные у других организмов. Однако систематический анализ взаимного расположения и транскрипционной специфичности содержащихся в них генов не проводился.

Мутация гена *SuUR* (Suppressor of UnderReplication) ускоряет прохождение репликации в описанных выше неактивных районах политенных хромосом и приводит к их полной политенизации (Belyaeva *et al.*, 1998; Sher *et al.*, 2012). Иммуноокрашивание политенных хромосом показало, что белок SUUR в норме связывается с позднореплицируемыми районами (Makunin *et al.*, 2002). Этот результат был подтвержден позднее в экспериментах по картированию бела SUUR методом DamID в культуре клеток Кс *Drosophila melanogaster* (Pindyurin *et al.*, 2007).

Механизм, в котором задействован белок SUUR, остается неизвестным, однако ясно, что он тесно ассоциирован с транскрипционно неактивными районами хромосом. В культуре клеток Кс белок SUUR является одним из основных компонентов всех трех репрессивных типов хроматина, выделенных на основании картирования 53 хроматиновых белков (Filion *et al.*, 2010).

Хроматин является динамичной структурой и, как его белковый состав, так и транскрипционная активность, могут различаться между типами клеток и стадиями развития. В представленной работе был поставлен ряд вопросов о функционировании и динамике доменов, связанных с белком SUUR. Было неясно, обладают ли эти районы какими-либо особенностями расположения генов, как было описано для других организмов, где и когда эти гены активны, а также насколько эти домены изменяются в процессе развития *D. melanogaster*.

Цели и задачи исследования. Целью данной работы было изучение свойств транскрипционно неактивных районов хромосом, связанных с белком SUUR, в процессе развития *D. melanogaster*. В связи с этим были поставлены следующие задачи:

- 1) Изучить особенности расположения генов в районах недорепликации политенных хромосом слюнных желез личинок дрозофилы
- 2) Исследовать транскрипционную специфичность генов в районах недорепликации
- 3) Получить полногеномные профили связывания белка SUUR в разных типах клеток дрозофилы
- 4) Исследовать динамику связывания белка SUUR в разных типах клеток
- 5) Проверить наличие корреляции между изменением уровня связывания SUUR в разных клеточных типах и активностью генов

Научная новизна. В представленной работе на молекулярном уровне показано, что районы недорепликации в хромосомах слюнных желез дрозофилы обогащены белком SUUR. Установлено, что такие районы обладают сниженной плотностью генов и имеют длинные межгенные

интервалы. Впервые показано, что эти характеристики, наряду с обогащением семенник-специфичными генами, свойственны в первую очередь хроматину, связанному с белком SUUR. Установлено, что белок SUUR распределен в хроматине из разных тканей по-разному, причем наблюдаемая динамика отличается для разных типов хроматина, согласно классификации Filion *et al.* (2010). Показано, что участки хромосом, в которых изменяется уровень связывания белка SUUR в разных типах клеток, содержат преимущественно тканеспецифичные гены, имеющие разный уровень экспрессии между этими типами клеток.

Научно-практическая ценность. Получены данные об экспрессии 14056 генов дрозофилы на десяти разных стадиях развития от эмбриона до имаго обоих полов. Построены профили связывания белка SUUR в эмбрионах, слюнных железах и мозговых ганглиях личинок дрозофилы. Создан алгоритм, позволяющий находить в геноме районы, обладающие низкой плотностью генов, длинными межгенными интервалами и обогащенные семенник-специфичными генами. С использованием этого алгоритма найдены 110 районов, характеризующихся сниженным уровнем политенизации и связыванием белка SUUR. Эти данные могут быть использованы для дальнейшего изучения организации тканеспецифичных генов дрозофилы. Предложен метод, позволяющий оценивать динамику хроматиновых белков в разных клеточных типах.

Основные положения диссертации, выносимые на защиту. В хромосомах слюнных желез личинок дрозофилы белок SUUR связывается с районами, характеризующимися низкой плотностью генов, длинными межгенными интервалами и обогащением семенник-специфичными генами. Профиль связывания белка SUUR отличается в хромосомах эмбрионов, мозговых ганглиев и слюнных железах личинок, а так же в культуре клеток Кс. Эти отличия коррелируют с изменением в распределении модификации H3K27me3 и с активностью генов. Наблюдаемые различия в связывании SUUR отличаются для разных типов хроматина, согласно классификации Filion *et al.* (2010).

Личный вклад автора. Основные результаты получены автором самостоятельно. Работа по молекулярному клонированию векторов pUC18-Dam-Myc-SUUR и pUC18-Dam-Myc проведена С.Н. Беякиным. Работа по получению данных о транскрипции генов дрозофилы на 10 стадиях развития выполнена совместно с С.Н. Беякиным и В.В. Шломой. Разработка алгоритма предсказания районов недорепликации осуществлялась совместно с В.Н. Бабенко, И.Ф. Жимулевым, Е.С. Беяевой и С.Н. Беякиным.

Апробация работы. Основные результаты работы были представлены на следующих конференциях: Международная конференция "Хромосома-2009" (Новосибирск, 2009), Международная Студенческая Конференция (Новосибирск, 2010), ESF-EMBO Conference on System Biology of Drosophila (Пултуск, Польша, 2012).

Публикации. Результаты исследования представлены в 5 статьях и 3 тезисах конференций.

Структура и объем работы. Диссертация изложена на 128 страницах машинописного текста и содержит 23 рисунка. Текст состоит из введения, обзора литературы, описания использованных в работе материалов и методов, результатов, обсуждения и выводов, а так же списка цитируемой литературы, в который входит 176 ссылок.

МАТЕРИАЛЫ И МЕТОДЫ

Линии мух. В работе были использованы линии дрозофил, полученные из Bloomington Stock Center, а так же линия *u w* (фонд лаборатории).

Выделение РНК и анализ экспрессии генов в развитии. Выделение и флюоресцентное мечение РНК из биологических образцов проводили по стандартным методикам, согласно рекомендациям фирм-производителей (Invitrogen, Amersham, QIAGEN). Для анализа экспрессии использовали микрочипы DGRC-2, полученные из Центра Исследования Генома Дрозофилы.

Молекулярное клонирование и процедура DamID. Клонирование и трансформация генетических конструкций проводили по стандартным протоколам, либо согласно рекомендациям производителей (QIAGEN, SibEnzyme, Fermentas, New England Biolabs). Генетические конструкции создавали на основе экспрессионного вектора pUAST. Процедуру DamID проводили по протоколам, описанным ранее, с некоторыми изменениями (Pollack *et al.*, 1999; Greil *et al.*, 2006). Анализ DamID образцов проводили при помощи микрочипов «*D. melanogaster* Gene Expression 4x72K Arrays» (Nimblegen).

Биоинформатический анализ данных. Первичный анализ транскрипции и DamID проводили при помощи программного обеспечения ArrayStar, TIGR Spotfinder и TIGR MIDAS. Дальнейший анализ этих данных, а так же изучение районов недорепликации проводили при помощи программного обеспечения MS Excel и при помощи оригинальных программ, написанных на языке программирования Perl версии 5.12.

РЕЗУЛЬТАТЫ

Связывание белка SUUR с районами недорепликации политенных хромосом слюнных желез *D. melanogaster*. Для получения профиля связывания белка SUUR был использован метод DamID (van Steensel, Henikoff, 2000). Он основан на способности белка Dam (ДНК-аденин-метилтрансфераза *Escherichia coli*) метилировать аденин в последовательности ГАТЦ *in vivo*. Метилированные фрагменты могут быть избирательно амплифицированы и проанализированы на микрочипах. Если Dam ввести в состав химерного белка вместе с белком SUUR, метилирование ДНК будет происходить предпочтительно в сайтах связывания SUUR с хромосомами, что выявляется как локальное усиление сигнала на микрочипах. В качестве отрицательного контроля используют профиль метилирования генома только белком Dam.

Нами были созданы генетические конструкции, кодирующие химерный белок Dam-SUUR и отдельный белок Dam, под контролем минимального промотора гена *hsp70*. Эти конструкции были трансформированы в геном дрозофилы. Из личиночных слюнных желез самок полученных линий были выделены образцы геномной ДНК, которые провели через все этапы процедуры DamID, после чего проанализировали на микрочипах.

Визуальное сравнение полученного профиля с данными о представленности ДНК в слюнной железе из работы С.Н. Белякина и соавторов (Belyakin *et al.*, 2005) выявило хорошее совпадение между зонами повышенного связывания белка SUUR и участками недорепликации ДНК (рис. 1). Парное сравнение всех данных из обоих профилей подтвердило наблюдаемую закономерность – практически все недореплицированные участки проявляют повышенное связывание с SUUR, но не все связанные с этим белком последовательности проявляют признаки недорепликации, что хорошо соотносится с ранее известными фактами (Belyaeva *et al.*, 1998; Makunin *et al.*, 2002; Zhimulev *et al.*, 2003a).

Организация генов в районах недорепликации. Для того чтобы исследовать свойства районов генома, с которыми связывается белок SUUR, мы использовали 52 района недорепликации, описанные ранее (Belyakin *et al.*, 2005). Эти районы являются наиболее характерными участками связывания SUUR, и эффект мутации гена *SuUR* сильнее всего проявляется в репликации именно этих районов хромосом.

Районы недорепликации сравнивали со всем геномом, а так же с их флангами. В качестве флангов считали участки хромосомы, граничащие с каждым районом недорепликации, причем длину фланга определяли так,

чтобы число генов в нем равнялось половине от числа генов в самом районе недорепликации.

Оказалось, что плотность генов в районах недорепликации значительно ниже, чем в их флангах. Так, в районах недорепликации медианная плотность генов составила 70 генов на млн.п.н., а во флангах она оказалась равной 154 генам на млн.п.н. (рис. 2, А). Чтобы исследовать природу низкой генной плотности в районах недорепликации, был проведен анализ длин генов в них. Выяснилось, что районы недорепликации обогащены короткими генами. Причем наибольшую достоверность наблюдали для генов короче 900 п.н. (291 ген в районах недорепликации и 143 во флангах, Рис. 2, Б). Этот результат указывает на то, что снижение генной плотности в районах недорепликации, наряду со снижением средней длины гена, должно быть связано с увеличением длин межгенных интервалов в них. Действительно, оказалось, что районы недорепликации значительно обогащены длинными межгенными интервалами. В частности, число межгенных участков длиннее 8 т.п.н. в районах недорепликации в 2,7 раз больше, чем во флангах (рис. 2, В).

Районы недорепликации обогащены семенник-специфичными генами. Для получения данных о специфике транскрипции в районах недорепликации был проведен анализ экспрессии большинства генов дрозофилы на разных этапах ее развития. Для получения этих данных были

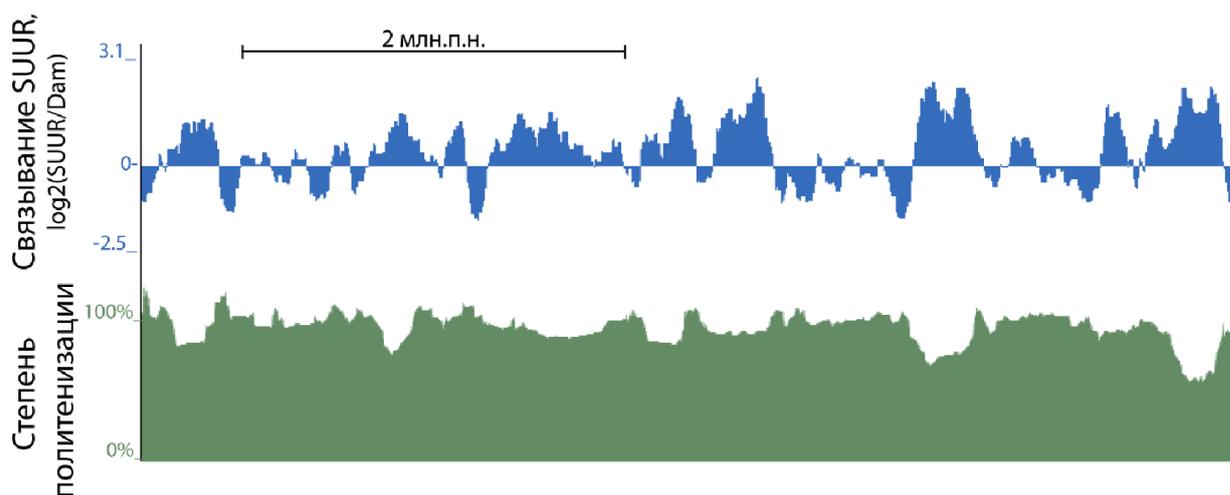


Рис. 1. Сопоставление полученного нами профиля связывания белка SUUR и опубликованных ранее данных о степени политенизации в слюнной железе (Belyakin *et al.*, 2005) для участка хромосомы 2L (геномные координаты 13603352-19299980). Значения верхнего графика выше 0 соответствуют районам связывания SUUR. Высота нижнего графика соответствует степени политенизации ДНК. Значения ниже 100% соответствуют недореплицированным районам.

использованы олигонуклеотидные транскриптомные микрочипы, на которые гибридизовали кДНК, полученную из мух на разных стадиях развития: три эмбриональных стадии (0-1 часов, 5-6 часов, 11-12 часов после откладки), три личиночных стадии (первый, второй и третий личиночные возрасты), стадия предкуколки, стадия куколки, взрослые самки и самцы.

Оказалось, что районы недорепликации обогащены короткими генами, транскрипция которых в развитии начинается в метаморфозе и наиболее активна на стадии взрослых самцов, но не самок (рис. 3, А). Ранее было показано, что такой паттерн экспрессии в самцах формируется за счет активности генов в развивающихся мужских гонадах (Arbeitman *et al.*, 2002). Чтобы проверить, действительно ли наблюдаемое повышение транскрипции этих генов в самцах вызвано усиленной экспрессией в семенниках, были использованы данные об интенсивности транскрипции в 13 различных тканях взрослых мух, полученные с использованием микрочипов другими исследователями (Chintapalli *et al.*, 2007). Сравнение усредненных профилей экспрессии генов в районах недорепликации и их флангах в 13 тканях подтвердило значительное обогащение районов недорепликации генами,

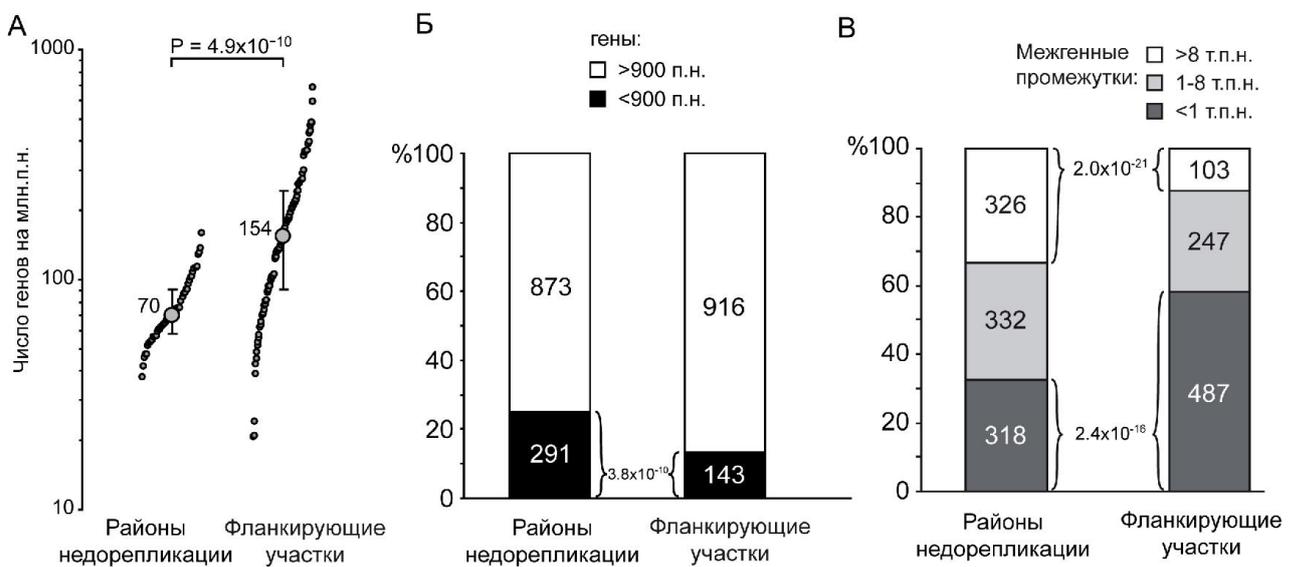


Рис. 2. Сравнение генной плотности, длин генов и длин межгенных интервалов в районах недорепликации и их флангах. А – Значения количества генов на 1 млн.п.н. для каждого района недорепликации и каждого фланга. Большими серыми точками обозначены медианные значения, черными линиями обозначены диапазоны разброса значений для 25-75 перцентилей всех данных. Б – Доли генов короче 900 п.н. в исследуемых районах относительно всех генов. Цифры внутри гистограмм обозначают абсолютные числа генов. В – Доли межгенных интервалов разных длин в районах недорепликации и в их флангах. Размеры выборок обозначены цифрами внутри гистограмм.

экспрессирующимися в семенниках взрослых мух (рис. 3, Б). Анализ распределения длин семенник-специфичных генов в этих районах показал, что большинство из них не превышают 2 т.п.н. в длину (1235 из 1653). При этом доля таких коротких семенник-специфичных генов в районах недорепликации значительно больше, чем в их фланкирующих участках: 41% в районах недорепликации и 17% во флангах.

Предсказание районов недорепликации. Известно, что при увеличении числа копий гена *SuUR* увеличивается и число районов хромосом, проявляющих недорепликацию (Zhimulev *et al.*, 2003a; Zhimulev *et al.*, 2003b). Это означает, что в геноме дрозофилы могут присутствовать и другие районы, обладающие обнаруженными нами геномными характеристиками, но не представленные среди 51 района недорепликации. Мы решили проверить функциональность обнаруженных характеристик и провели полногеномный поиск районов ДНК с похожими свойствами.

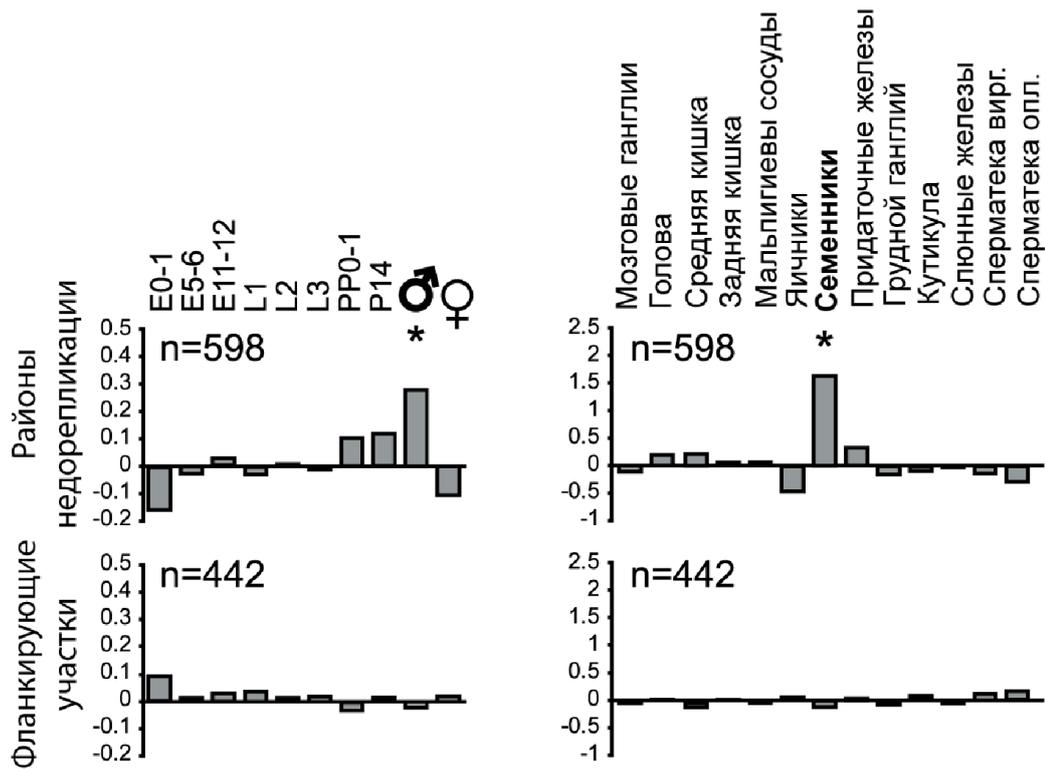


Рис. 3. Усредненные значения экспрессии всех генов, попадающих в районы недорепликации, либо в их фланги, на разных этапах развития (А) и в разных тканях (Б) дрозофилы (Chintapalli *et al.*, 2007). Звездочкой обозначена усиленная экспрессия в семенниках взрослых самцов. По оси ординат – логарифм отношения экспрессии генов в выбранном образце к экспрессии в смеси всех образцов, нормированный на уровень экспрессии всех генов в образце.

Для проведения такого анализа была создана компьютерная программа, которая производит поиск участков генома, где, наряду с низкой плотностью генов, наблюдается достоверное повышение доли длинных межгенных промежутков и семенник-специфичных генов по сравнению с флангами этих участков. С помощью этой программы были обнаружены 110 районов, обладающих указанными свойствами (рис. 4, А). Оказалось, что 41 из 51 экспериментально установленного района недорепликации пересекается с предсказанными участками хотя бы на 50%. Выяснилось, что предсказанные районы достоверно проявляют признаки недорепликации в слюнных железах, хотя и в меньшей степени, чем 51 экспериментально установленный

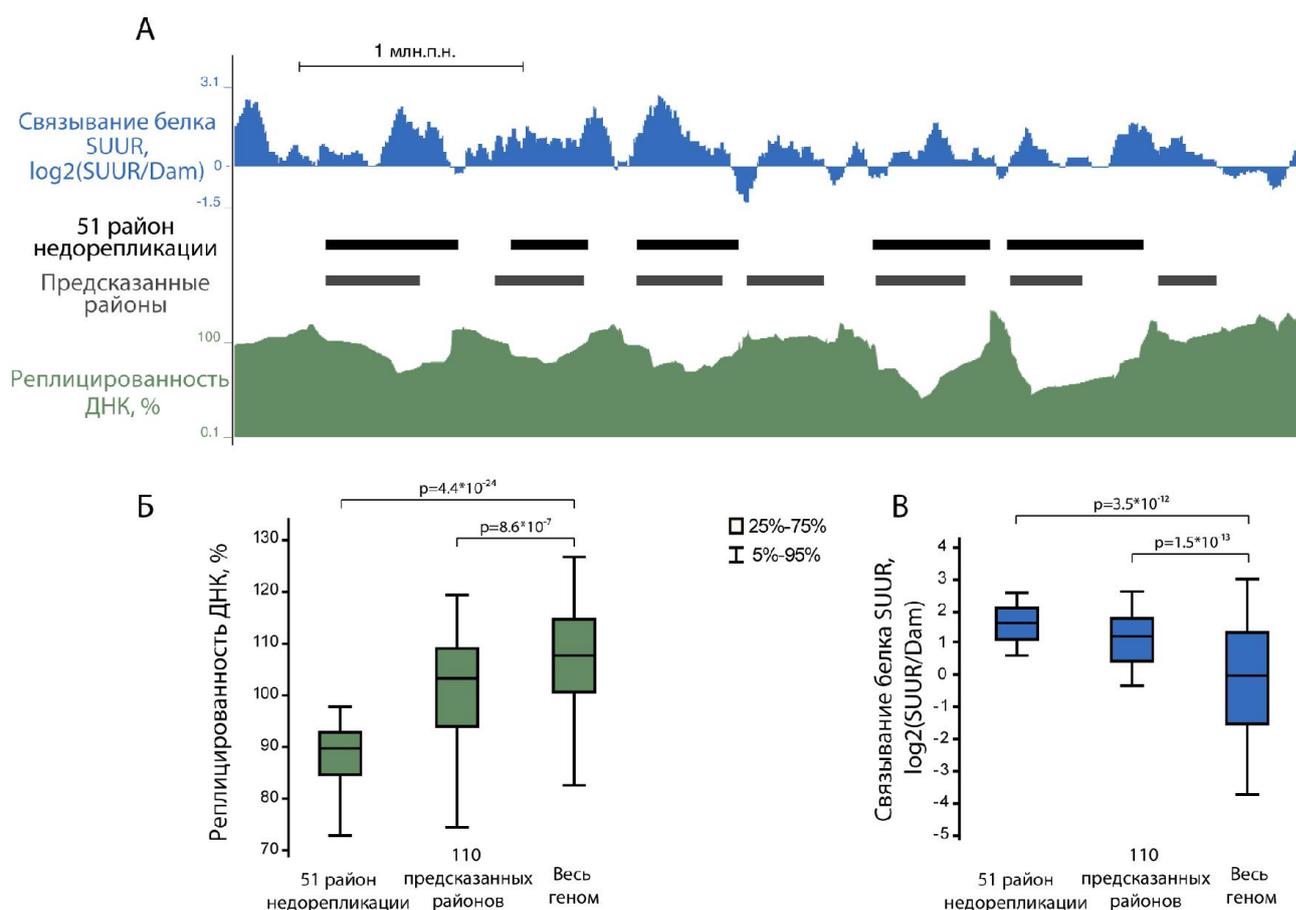


Рис. 4. Степень политенизации и связывание с белком SUUR предсказанных районов. А – профиль связывания SUUR и степень политенизации участка хромосомы 2L (геномные координаты 14046263-18813360). Серыми прямоугольниками обозначены семь предсказанных районов. Видно, что пять из них хорошо перекрываются с экспериментально обнаруженными районами недорепликации (черные прямоугольники). Б, В – распределения значений степени политенизации и связывания SUUR для районов недорепликации, 110 предсказанных районов и всего генома. Черными линиями внутри гистограмм обозначены медианные значения. Достоверность отличий от всего генома определяли по критерию Манна-Уитни.

район (рис. 4, Б). В то же время, белок SUUR связывается с предсказанными районами практически так же сильно, как с 51 контрольным районом недорепликации (рис. 4, В).

Динамика связывания SUUR в развитии. Известно, что белок SUUR имеет тенденцию к колокализации с белками транскрипционно неактивного хроматина (Zhimulev *et al.*, 2003a; Pindyurin *et al.*, 2007). Если предположить, что этот белок участвует в формировании гетерохроматина, можно ожидать, что в разных тканях этот белок может быть связан с разными участками ДНК, так как тканеспецифичные гены, находящиеся в гетерохроматине в одной ткани, могут активироваться в другой. Чтобы проверить это предположение, мы провели анализ связывания SUUR с хромосомами в двух других биологических образцах – в эмбрионах (12-14 часов после откладки) и в мозговых ганглиях личинок третьего возраста. Кроме того, нами были использованы данные картирования SUUR в культуре клеток Кс (Filion *et al.*, 2010).

Полученные профили связывания SUUR оказались в большой степени похожими друг на друга, однако, помимо областей, где этот белок связывается одинаково в разных образцах, присутствовали многочисленные области с отличающимся связыванием (рис. 5, А). Чтобы оценить наблюдаемый паттерн на качественном уровне и связать его с конкретными генами, в каждом образце мы выделили SUUR-связанные и SUUR-обедненные гены (рис. 5, Б). Выяснилось, что количество генов, которые во всех четырех образцах являются SUUR-связанными (220) и SUUR-обедненными (611) достаточно малы. Когда мы разделили эти два набора генов по пяти типам хроматина, к которым они относятся в клетках Кс, выяснилось, что гены, связанные с SUUR во всех образцах, в основном находятся в неактивном «черном» типе хроматина. Напротив, гены, не связанные с SUUR ни в одном из образцов, обычно располагались в активном «желтом» хроматине (рис. 5, В).

Связь динамики белка SUUR с экспрессией генов в развитии. Наибольший интерес для нас представляли гены, которые изменяют свое связывание с SUUR в разных тканях, что может отражать изменение их хроматинового состояния и может быть сопряжено с изменением их активности. Чтобы изучить этот вопрос, для каждой пары образцов были выделены гены, для которых значение связывания SUUR в одном из образцов было положительным и как минимум в четыре раза выше, чем во втором образце. В каждой паре образцов было обнаружено от 295 до 853 таких генов.

Чтобы выяснить, коррелируют ли различия в связывании SUUR с активностью этих генов, мы использовали ранее опубликованные данные по экспрессии всех генов дрозофилы в выбранных нами биологических образцах (<http://www.modencode.org/>, Filion *et al.*, 2010). Оказалось, что для всех пар исследованных образцов наблюдается одинаковая закономерность: при снижении связывания с SUUR гены часто начинают активироваться, а при усилении связывания SUUR их активность уменьшается (рис. 6, А).

Чтобы проследить, как гены, меняющие свое связывание с белком SUUR, связаны с функцией каждого исследованного нами биологического образца, мы провели анализ их генной онтологии (Gene Ontology). Этот

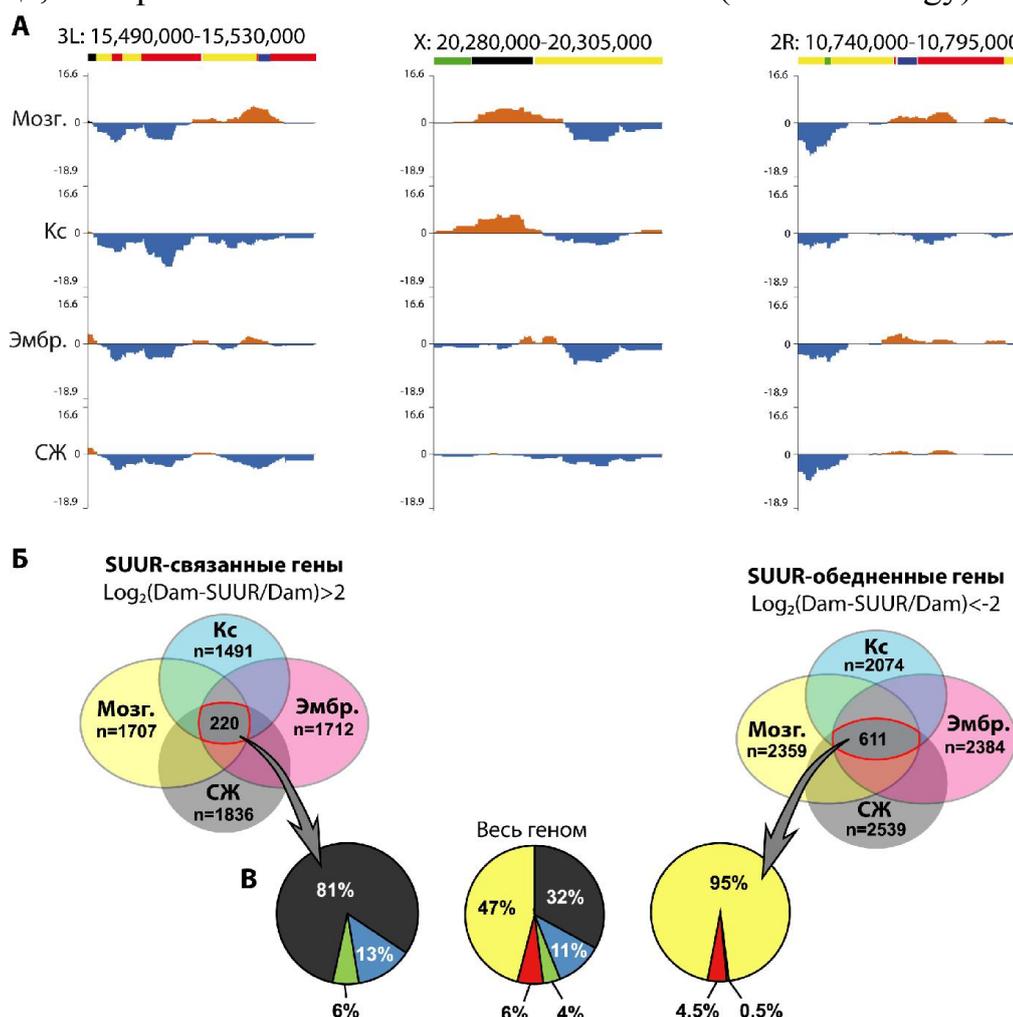


Рис. 5. Различия в связывании белка SUUR с хромосомами мозговых ганглиев личинок (Мозг.), клеток Кс, эмбрионов (Эмбр.) и слюнных желез (СЖ). А – профили связывания SUUR с тремя участками генома. По оси ординат отложены сглаженные значения связывания SUUR, $\log_2(\text{Dam-SUUR}/\text{Dam})$. Значения выше 0 соответствуют участкам связывания с белком SUUR (оранжевый цвет), ниже 0 – районы хромосом, избегаемые SUUR (синий цвет). Б – число генов, связанных и не связанных с SUUR в каждом образце. В – распределение общих связанных и несвязанных генов по типам хроматина в клетках Кс.

биоинформатический анализ позволяет определить, есть ли в изучаемой группе генов обогащение по какой-либо биологической функции (категории). Выяснилось, что гены, связанные с SUUR в одном из образцов слабее, чем в остальных образцах, обычно являются тканеспецифичными и связаны с биологической функцией этого образца. Так, для слюнной железы это были многочисленные гены, имеющие отношение к секретирующей функции этого органа. Для группы генов из мозговых ганглиев обогащение было обнаружено в категориях, связанных с восприятием и передачей информации. Анализ в эмбрионах выявил обогащение по категориям, связанным с развитием и морфогенезом. Для генов из Кс клеток мы так же наблюдали обогащение по ряду категорий, однако проследить функциональную связь этих генов с клетками Кс сложно, поскольку не известны точное происхождение и функция этих клеток.

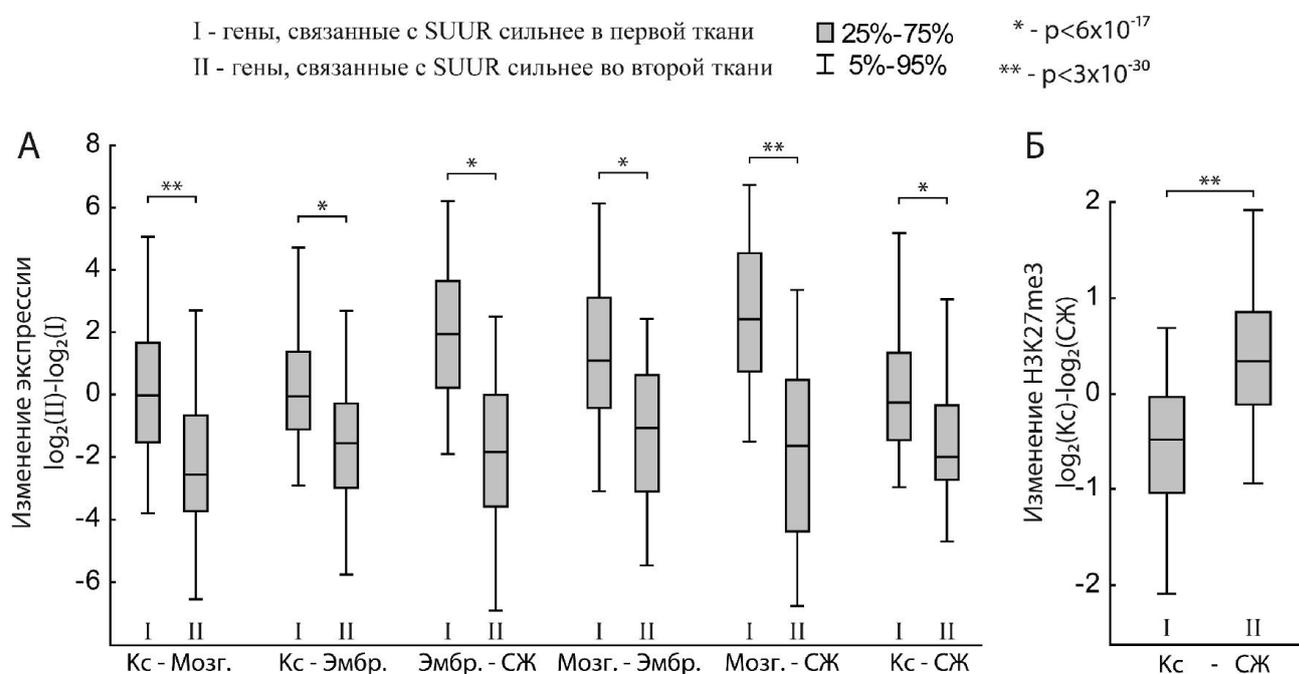


Рис. 6. Изменения экспрессии и обогащения модификацией H3K27me3 генов, меняющих связывание с SUUR в клетках Кс, мозговых ганглиях личинок, эмбрионах и слюнных железах. А – диаграммы распределений изменения уровней экспрессии генов, снижающих (первая диаграмма из пары), либо повышающих (вторая диаграмма) свое связывание с SUUR в образце I по сравнению с образцом II. Во всех образцах можно наблюдать достоверную тенденцию снижения экспрессии при усилении связывания с SUUR. Б – диаграмма изменения уровня H3K27me3 в генах, усиливающих связывание с SUUR в слюнных железах по сравнению с клетками Кс. Достоверность определяли по критерию Манна-Уитни.

Динамика белка SUUR в контексте изменения состояния хроматина в развитии. Поскольку в клетках Кс белок SUUR часто колокализуется с гетерохроматиновыми белками, участвуя в формировании трех репрессивных типов хроматина, можно ожидать, что изменение его связывания будет соответствовать изменению типа хроматина. Чтобы косвенно проследить этот эффект, мы воспользовались данными о распределении маркера неактивного хроматина – модификации H3K27me3 в хромосомах слюнных желез (Sher *et al.*, 2011) и клеток Кс (Filion *et al.*, 2010). Оказалось, что изменение связывания SUUR значительно коррелирует с метилированием H3K27me3. Гены, которые в клетках Кс связаны с белком SUUR сильнее, чем в слюнных железах, обычно имеют повышенный уровень H3K27me3 в клетках Кс. И наоборот, повышенное связывание SUUR в слюнных железах по сравнению с клетками Кс часто соответствует повышенному уровню H3K27me3 в слюнных железах (рис. 6, Б).

Поскольку SUUR является одним из основных компонентов трех из пяти хроматиновых типов, мы решили проследить, зависит ли вариативность связывания SUUR от типа хроматина. Для этого мы разделили все гены, проявляющие отличающийся уровень связывания SUUR в разных образцах по их хроматиновому состоянию в клетках Кс. Оказалось, что доля генов, которые связывают белок SUUR в клетках Кс, но не в остальных образцах, достоверно превышена в «синем» и «зеленом» типах хроматина, и снижена в «красном» типе (рис. 7, А). В то же время, гены, которые не связаны с SUUR в клетках Кс, но связываются с этим белком в других тканях, чаще всего расположены в «красном» типе хроматина, более того, такую картину наблюдали почти для 30% генов в этом типе (рис. 7, А).

Исчезновение белка SUUR из участков, занятых репрессирующими типами хроматина, может свидетельствовать об их переходе в активное состояние. В то же время, появление SUUR в участках, которые ранее были заняты активным «красным» хроматином, может отражать смену их хроматинового состояния на неактивное. Чтобы косвенно проверить это предположение, для пары образцов (слюнные железы и клетки Кс) мы проследили, как изменение распределения модификации H3K27me3 для генов с изменяющимся связыванием белка SUUR коррелирует с отдельными типами хроматина. Оказалось, что все типы, кроме «зеленого», изменяют уровень H3K27me3 совместно с изменением связывания SUUR (рис. 7, Б). Следует отметить, что отсутствие сигнала в «зеленом» хроматине может быть обусловлено выбором маркера неактивного хроматина – модификация H3K27me3 практически отсутствует в этом типе хроматина (Filion *et al.*, 2010).

Известно, что изменение хроматинового состояния участка хромосомы не всегда приводит к изменению экспрессии генов в нем (Schubeler *et al.*, 2001). Поэтому мы проследили, как изменение связи с белком SUUR коррелирует с изменением активности генов в каждом отдельном типе хроматина.

Выяснилось, что гены из «черного» и «синего» типов хроматина способны как активироваться при снижении связывания с SUUR, так и инактивироваться при повышении связывания с этим белком (рис. 7, В). Гены из «красного» и «желтого» типов часто снижают уровень транскрипции

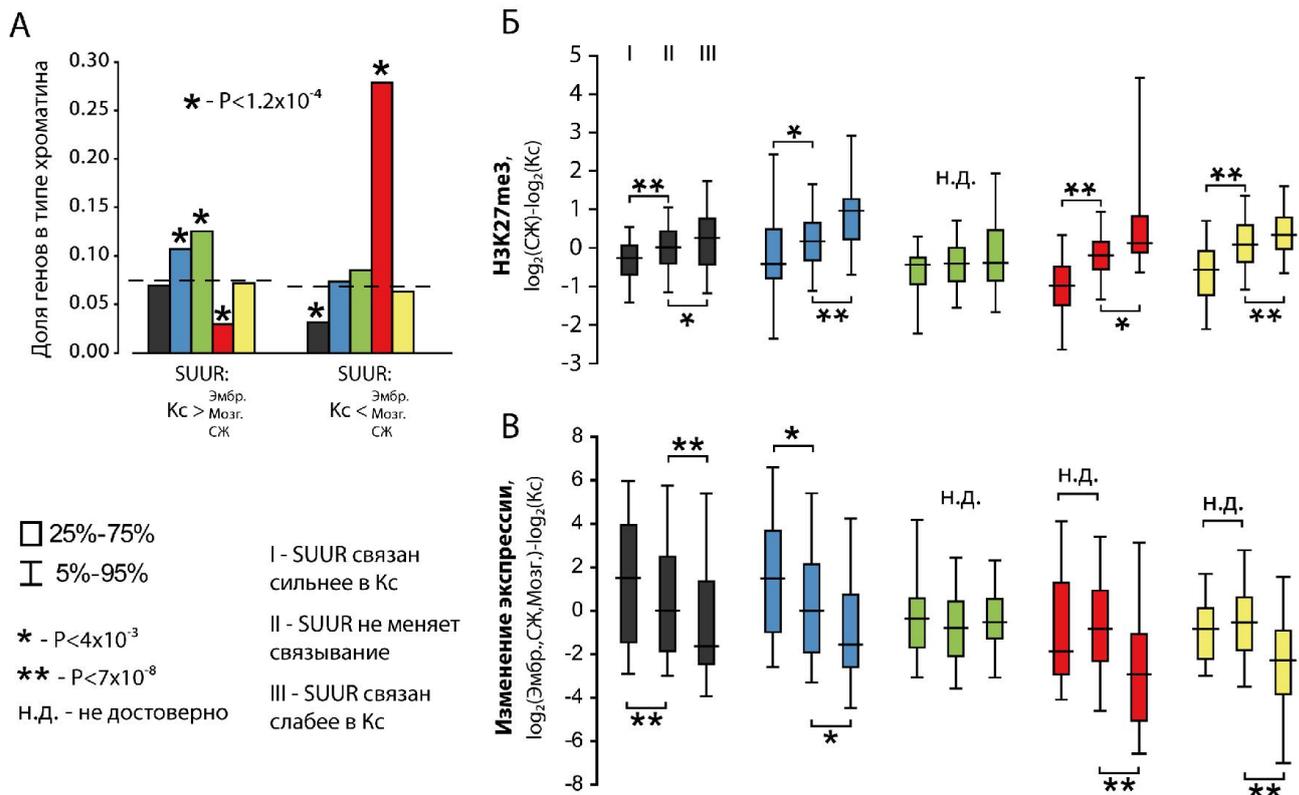


Рис. 7. Корреляция изменения связывания SUUR и сопутствующих изменений активности генов с типами хроматина в клетках Кс. А – доли генов в каждом типе хроматина, снижающих, либо усиливающих связывание с SUUR в клетках Кс по сравнению с эмбрионами, слюнными железами и мозговыми ганглиями. Пунктирными линиями обозначены ожидаемые значения в случае отсутствия зависимости от типа хроматина. Звездочками обозначены достоверно обогащенные (выше пунктирной линии) и обедненные (ниже линии) типы хроматина. Б – изменение уровня H3K27me3 в генах, изменяющих и не изменяющих связывание с SUUR между слюнными железами и клетками Кс. Положительная корреляция между связыванием SUUR и H3K27me3 наблюдается для всех типов хроматина кроме «зеленого». В – аналогичные диаграммы по объединенному изменению экспрессии генов между парами образцов Кс-СЖ, Кс-Эмбрионы, Кс-Мозговые ганглии.

одновременно с усилением связывания с SUUR. По-видимому, обратной ситуации не наблюдается, поскольку в этих двух типах хроматина в клетках Кс практически нет белка SUUR, и дополнительное снижение уровня его связывания не влияет на транскрипцию. Интересно, что гены из «зеленого» типа хроматина практически не проявили связи между транскрипцией и изменением связывания SUUR (рис. 7, В).

ОБСУЖДЕНИЕ

Особенности генетической организации районов, обогащенных белком SUUR. Мы установили, что в хромосомах слюнных желез белок SUUR связывается преимущественно с районами, содержащими короткие семенник-специфичные гены, разделенные длинными межгенными интервалами. Мы также показали, что паттерны связывания SUUR для эмбрионов и мозговых ганглиев личинок отличаются от паттерна для слюнных желез, и эти отличия коррелируют с изменением в распределении других белков хроматина и с транскрипционной активностью генов. Поскольку обнаруженные нами параметры ДНК – длинные межгенные интервалы, низкая плотность и короткий размер генов – идентичны во всех клетках организма, а связывание белка SUUR динамично, то нельзя утверждать, что эти свойства ДНК напрямую определяют силу связывания SUUR.

Единственный обнаруженный нами параметр SUUR-обогащенного хроматина, который может изменяться в разных типах клеток, – это степень активности тканеспецифичных генов. Действительно, в слюнной железе работа семенник-специфичных генов не требуется, поэтому активное связывание с ними белка SUUR неудивительно. Тем не менее, очевидно, что тканеспецифичные гены, содержащиеся в таких районах, активны в каких-то других типах клеток.

Поскольку нами была показана корреляция между активацией генов и снижением связывания с ними белка SUUR, можно предположить, что в тех тканях, где необходима работа тканеспецифичных генов, белок SUUR с ними не связан. Этим можно объяснить наблюдаемую динамику белка SUUR в разных тканях. Такое предположение находит поддержку и в результатах анализа генной онтологии. Кроме того, ранее был показан эффект исчезновения белка SUUR и недорепликации в результате искусственной активации транскрипции в районе, который в норме недореплицирован и связан с SUUR (Koryakov *et al.*, 2012).

В геноме присутствуют гены, работающие специфично в слюнной железе. Изучив расположение таких генов, можно заметить, что они тоже

часто скомпонованы в группы, расположенные в участках ДНК с низкой плотностью генов и длинными межгенными участками. Это, к примеру, кластер генов семейства *Eig71E*, активируемый экдизоном и расположенный в районе 71E, а так же группы генов, ответственных за секреторную функцию слюнной железы – *Sgs3*, *Sgs7*, *Sgs8* (район 68C) и *CG7606*, *Sgs5*, *CG7587* (район 90, B). Все эти гены очень активно работают в слюнной железе, расположены в районах, характеризующихся нормальной степенью политенизации и отсутствием белка SUUR. При этом, согласно нашим данным, в других изученных нами биологических образцах, где эти гены неактивны – в эмбрионах, мозговых ганглиях личинок и клетках Кс – с этими участками активно связывается белок SUUR. Все это говорит о том, что обнаруженные нами характерные участки генома с низкой плотностью генов и длинными межгенными интервалами в первую очередь характеризуются наличием тканеспецифичных генов, активных в одних клетках и репрессированных в других. Связывание же с ними белка SUUR, по-видимому, происходит в тканеспецифичном порядке и имеет отношение к регуляции их экспрессии.

Организация таких тканеспецифичных доменов в геноме представляет значительный интерес. Согласно нашим результатам, у дрозофилы они обладают низкой плотностью генов и обычно окружены участками с высокой плотностью генов. Недавно было установлено, что в этих пограничных зонах обычно расположены активные гены и много активных точек инициации репликации (Sher *et al.*, 2011).

Для генома мыши была показана интересная закономерность, позволяющая проводить аналогии с нашими результатами в контексте динамики хроматина. Оказалось, что в процессе дифференцировки эмбриональных стволовых клеток в нервную ткань, многие участки генома изменяют свое время репликации. При этом координированно изменяется уровень транскрипции, положение в ядре и набор хроматиновых белков таких районов. Было установлено, что большинство участков, которые в обоих изученных образцах реплицируются рано, имеют высокую плотность генов. Остальные участки ДНК, которые хотя бы в одном из образцов реплицируются поздно, имеют значительно сниженную плотность генов (Hiratani *et al.*, 2008). Возможно, такие участки являются функциональным аналогом наблюдаемых нами тканеспецифичных районов у дрозофилы.

Интересен факт увеличения межгенных промежутков в районах недорепликации. Это может быть вызвано тем, что для поддержания таких районов хромосом в компактном, транскрипционно неактивном состоянии необходимы регуляторные последовательности, влияющие не на отдельные

гены, а на район в целом, например, аналогичные элементам LCR (Erner *et al.*, 1981; Aladjem *et al.*, 1995). Кроме того, с межгенных промежутков могут транскрибироваться некодирующие РНК. Недавно было показано, что около 80% генома человека экспрессируются в виде длинных некодирующих РНК, предположительно оказывающих регуляторное воздействие. Большая их часть экспрессируется на очень низком уровне (Hangauer *et al.*, 2013) и не во всех изученных типах клеток (Djebali *et al.*, 2012). Возможно длинные межгенные интервалы у дрозофилы обладают сходными свойствами. Наши результаты совместно с данными о тканеспецифической транскрипции длинных некодирующих РНК у дрозофилы (Brown *et al.*, 2014) могут быть использованы для изучения транскрипционной активности межгенных интервалов.

Причина обогащения выявленных районов короткими генами остается неясной. Большинство коротких генов, которые в слюнных железах связаны с SUUR, транскрибируются в основном в семенниках, причем на очень высоком уровне (Chintapalli *et al.*, 2007). Другим примером сверхактивных коротких генов могут служить уже упоминавшиеся гены семейств *Eig71E* и *Sgs*, специфичные для слюнных желез. Есть свидетельства того, что для поддержания уровня транскрипции на очень высоком уровне выгоднее, чтобы в гене отсутствовали или были максимально уменьшены интроны, так как их транскрипция и сплайсинг занимают дополнительное время и ресурсы (Castillo-Davis *et al.*, 2002; Shabalina *et al.*, 2010). Этим можно объяснить снижение размера таких тканеспецифичных генов.

Таким образом, особенности генетической организации районов, ассоциированных с белком SUUR в разных клетках, могут быть связаны с регуляцией экспрессии тканеспецифичных генов, расположенных в этих районах. Аналогичные характеристики прослеживаются и у других организмов, эволюционно удаленных от *D. melanogaster*, что может отражать сходство принципов регуляции тканеспецифичных генов.

Динамика связывания белка SUUR в развитии и регуляция генов.

Белок SUUR связан с неактивным хроматином, но остается неясным, участвует ли этот белок в его формировании, либо связывание SUUR с таким хроматином происходит уже как следствие. С одной стороны, мутация гена *SuUR* жизнеспособна и фертильна (Belyaeva *et al.*, 1998). Более того, экспрессия генов, в норме связанных с этим белком, не изменяется в случае мутации гена *SuUR* (Sher *et al.*, 2011). С другой стороны, отсутствие белка SUUR влияет на связывание гетерохроматиновых белков и модификации гистонов: H3K27me3 (Sher *et al.*, 2011), HP1 (Koryakov *et al.*, 2006), SU(VAR)3-9 и H3K9me (Koryakov *et al.*, 2011). Более того, SUUR выступает

как слабый модификатор эффекта положения мозаичного типа (Belyaeva *et al.*, 2003). Таким образом, хотя SUUR каким-то образом связан с функционированием неактивного хроматина у дрозофилы, он, по-видимому, не является ключевым регулятором экспрессии генов, с которыми он связывается.

Белковый состав хроматина не всегда соответствует активности генов в нем. Об этом свидетельствуют результаты работы, в которой изучали связывание белка Pc в эмбрионах, имагинальных дисках T3 и в культуре клеток мезодермального происхождения S2 дрозофилы. Выяснилось, что многие гены, связанные с белком Pc в одной ткани и не связанные в другой, вовсе не обязательно изменяют свою транскрипцию. Более того, многие из них оказались специфичными для нервной системы, то есть вообще не должны активироваться в изученных типах клеток (Kwong *et al.*, 2008). По-видимому, для активации таких генов, помимо снижения связывания белка Pc, необходимы дополнительные специфические факторы.

При изучении активности локуса гена β -глобина человека, помещенного в клетки мыши, было сделано схожее наблюдение. Оказалось, что изменение белкового состава в этом локусе и приобретение им свойств активного хроматина не приводило к активации транскрипции гена β -глобина (Schubeler *et al.*, 2000). Транскрипция этого гена появлялась только при наличии специфического энхансера - LCR (Reik *et al.*, 1998).

Таким образом, для активации тканеспецифического гена необходимо, по-видимому, соблюдение двух условий: наличие транскрипционного фактора, активирующего этот ген, и доступность хроматина, допускающая взаимодействие этого фактора с геном. В такой модели белок SUUR служит фактором, связанным с механизмом инактивации районов хроматина, но не транскрипции отдельных генов. Действительно, не все гены, с которыми переставал связываться белок SUUR, активировались, по-видимому из-за того, что в выбранных образцах для них отсутствовали специфические активаторы транскрипции.

Динамика SUUR в разных типах хроматина. Полученные нами данные свидетельствуют об изменчивости состояния хроматина для отдельных участков хромосом в разных типах клеток. Наибольшую динамику в развитии проявили «синий», «зеленый» и «красный» типы хроматина.

Динамика «синего» хроматина логично соотносится с тем фактом, что этот тип характеризуется Pc-опосредованной репрессией, которая характерна для генов, связанных именно с регуляцией развития организма (Simon, 1995; Pirrotta, 1998; Ringrose, Paro, 2004; Boyer *et al.*, 2006). Несколько

неожиданным является динамичность части «зеленого» хроматина, соответствующего HP1-зависимой инактивации. Такой тип репрессии чаще ассоциируется с постоянно инактивированным прицентромерным гетерохроматином (Schotta *et al.*, 2002), но встречается и в эухроматиновых участках хромосом (Fanti *et al.*, 2003). Возможно такие участки могут активироваться в каких-то типах клеток. Согласно неопубликованным данным, полученным в нашей лаборатории, имеет место значительная вариабельность в связывании белка SU(VAR)3-9, участвующем в этом типе инактивации, в хромосомах слюнных желез и мозговых ганглиев дрозофилы.

Среди активных типов хроматина «красный», но не «желтый» проявил признаки транскрипционной инактивации в других образцах по сравнению с клетками Кс. При исследовании распределения ДНКазыI-гиперчувствительных сайтов в эмбрионах дрозофилы был получен схожий результат. Многие участки генома, которые в клетках Кс принадлежат к неактивным типам хроматина, в эмбрионах повышали свою доступность для ДНКазыI, что считается свидетельством их активации. В то же время, участки «красного» хроматина, часто снижали этот показатель, видимо переходя в неактивное состояние (Thomas *et al.*, 2011). Наши данные подтвердили это наблюдение для разных клеточных типов, причем в нашем случае в качестве маркера гетерохроматина выступал белок SUUR, специфичный для неактивного хроматина.

Результаты нашего исследования дают основания полагать, что неактивный «черный» и активный «желтый» типы хроматина являются наиболее стабильными в развитии, во всяком случае, в исследованных образцах. Ранее стабильность «черного» типа была косвенно показана при помощи репортерного анализа (Filion *et al.*, 2010). Так, когда авторы исследовали, как активируются трансгенные конструкции из разных типов хроматина в глазу мухи, они обнаружили, что именно в участках хромосом, соответствующих «черному» типу в клетках Кс, в глазах находилась самая большая доля неактивных трансгенов. То есть по сравнению с другими неактивными типами, именно в «черном» было меньше всего участков, повысивших свою активность в клетках глаза (Filion *et al.*, 2010). Слабая изменчивость «желтого» хроматина, по-видимому, связана с обогащением этого типа повсеместно-экспрессирующимися генами (Filion *et al.*, 2010).

ВЫВОДЫ

1. Районы хромосом слюнных желез, в которых под действием белка SUUR происходит недорепликация ДНК, характеризуются низкой

- плотностью генов, длинными межгенными интервалами и обогащены семенник-специфичными генами.
2. С помощью метода DamID построены полногеномные профили связывания белка SUUR в слюнных железах и мозговых ганглиях личинок, а также в целых эмбрионах дрозофилы. Установлено, что существуют различия в связывании белка SUUR в разных типах клеток. Выявлены дифференциально связанные гены.
 3. Показано, что относительное снижение связывания белка SUUR с хромосомами коррелирует с активацией части генов, связанных с физиологическими функциями исследованных клеточных типов.
 4. Установлено, что разные типы хроматина, согласно классификации Filion *et al.* (2010), демонстрируют различную степень варьирования связывания SUUR в разных клеточных типах. «Черный» и «желтый» хроматиновые типы наименее динамичны, а «синий» и, в особенности, «красный» демонстрируют более сильное варьирование связывания SUUR в разных клеточных типах.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ РАБОТЫ

Статьи:

1. Babenko V.N., Makunin I.V., Brusentsova I.V., Belyaeva E.S., **Maksimov D.A.**, Belyakin S.N., Maroy P., Vasil'eva L.A., Zhimulev I.F. Paucity and preferential suppression of transgenes in late replication domains of the *D. melanogaster* genome // BMC Genomics. 2010. V.11. P.318.
2. Belyakin S.N., Babenko V.N., **Maksimov D.A.**, Shloma V.V., Kvon E.Z., Belyaeva E.S., Zhimulev I.F. Gene density profile reveals the marking of late replicated domains in the *Drosophila melanogaster* genome // Chromosoma. 2010. V. 119(6). P. 589-600.
3. Жимулев И.Ф., Беляева Е.С., Андреева Е.Н., Андреевкова Н.Г., Бабенко В.Н., Белякин С.Н., Болдырева Л.В., Брусенцова И.В., Демаков С.А., Зыков И.А., Кокоза Е.Б., Колесникова Т.Д., **Максимов Д.А.**, Макунин И.В., Пиндюрин А.В., Семешин В.Ф., Хорошко В.А. Интеркалярный гетерохроматин в геноме *Drosophila* // Генетика. 2010. Т. 46(10). С. 1405-1408.
4. Koryakov D.E., Pokholkova G.V., **Maksimov D.A.**, Belyakin S.N., Belyaeva E.S., Zhimulev I.F. Induced transcription results in local changes in chromatin structure, replication timing, and DNA polytenization in a site of intercalary heterochromatin // Chromosoma. 2012. V. 121(6). P. 573-583.
5. **Maksimov D.A.**, Koryakov D.E., Belyakin S.N. Developmental variation of the SUUR protein binding correlates with gene regulation and specific

chromatin types in *D. melanogaster* // Chromosoma. 2014. V. 123(3). P. 253-264.

Тезисы конференций:

1. Белякин С.Н., Бабенко В.Н., **Максимов Д.А.**, Шлома В.В., Квон Е.З., Беляева Е.С., Жимулев И.Ф. Характеристики генной плотности в районах поздней репликации у *Drosophila melanogaster*// Международная конференция «Хромосома 2009», 2009, Новосибирск, Россия. С. 99.
2. **Максимов Д.А.** Организация районов недорепликации в геноме *D. melanogaster*// Международная студенческая конференция «МНСК-2010», 2010, Новосибирск, Россия.
3. **D.A. Maksimov**, D.E. Koryakov, S.N. Belyakin. Developmental variation of the SUUR protein binding correlates with gene regulation and specific chromatin types in *D. melanogaster*// ESF-EMBO Conference on System Biology of *Drosophila*. May 22-25 2012, Pułtusk, Poland.