

РОМАНОВ СТАНИСЛАВ ЕВГЕНЬЕВИЧ

**Роль белков CP190 и CG9879 в регуляции генов
дифференцировки сперматоцитов**

Drosophila melanogaster

1.5.7. – генетика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата биологических наук

Новосибирск – 2022

Работа выполнена в лаборатории геномики ФГБНУ «Институт молекулярной и клеточной биологии Сибирского отделения РАН», г. Новосибирск

Научный руководитель: **Лактионов Петр Павлович**
кандидат биологических наук, научный
сотрудник лаборатории геномики
Института молекулярной и клеточной
биологии Сибирского отделения РАН
г. Новосибирск

Официальные оппоненты: **Головнин Антон Клеменсович**
доктор биологических наук, профессор
РАН, заведующий лабораторией
молекулярной генетики дрозофилы
Института биологии гена
Российской Академии Наук
г. Москва

Юдкин Дмитрий Владимирович
кандидат биологических наук, заведую-
щий отделом геномных исследований
Государственного научного центра
вирусологии и биотехнологии «Вектор»
г. Кольцово

Ведущее учреждение: Институт молекулярной биологии
им. В.А. Энгельгардта Российской
академии наук
г. Москва

Защита диссертации состоится: «__» _____ 20__ г. в __ часов на за-
седании Диссертационного совета 24.1.239.01 на базе ФГБНУ «Федераль-
ный исследовательский центр Институт цитологии и генетики Сибирского
отделения РАН» в конференц-зале Института по адресу: 630090, Новоси-
бирск, пр. академика Лаврентьева 10, тел. (383)363-49-06, факс (383)333-12-
78, e-mail: dissov@bionet.nsc.ru.

С диссертацией можно ознакомиться в библиотеке ИЦиГ СО РАН и на сайте
Института <https://www.icgbio.ru/>.

Автореферат разослан «__» _____ 20__ г.

Ученый секретарь
диссертационного совета,
доктор биологических наук

Т.М. Хлебодарова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. На терминальных этапах сперматогенеза дрозофилы в сперматоцитах синхронно активируется более полутора тысяч генов дифференцировки, которые больше нигде не функционируют (White-Cooper, 2010). Ключевую роль в их индукции играют комплекс семенник-специфичных ТВР-ассоциированных факторов tTAF, а также семенник-специфичный комплекс задержки мейоза tMAC (White-Cooper, 2010). Нарушение функции этих комплексов приводит к остановке развития сперматоцитов в профазе мейоза I и полной стерильности самцов. Исследование регуляторных механизмов, ответственных за масштабную активацию генов в сперматоцитах, важно для понимания процессов, лежащих в основе передачи наследственной информации между поколениями.

Степень разработанности темы исследования. С использованием геномных методов показано, что лишь небольшая доля генов дифференцировки является прямыми мишенями tTAF и tMAC (Laktionov et al., 2018). Поэтому, для реконструкции регуляторного генетического каскада, ответственного за активацию генов дифференцировки, необходимо исследовать механизмы, ответственные за индукцию не прямых мишеней tTAF и tMAC.

Изучение генов-мишеней tMAC позволили предположить, что продукт гена *CG9879* необходим для индукции не прямых мишеней транскрипционных активаторов tTAF и tMAC. Кроме того, анализ факторов, гомологичных tMAC, указывает, что CP190 может быть необходим для регуляции генов-мишеней tMAC через дистальные геномные взаимодействия. CP190 является ключевым архитектурным белком дрозофилы и исследование его функции в клетках взрослого организма важно для понимания многих аспектов эпигенетической регуляции. Например, известно, что CP190 устанавливает барьерные элементы на границах доменов H3K27me3 в эмбрионах, но не ясно, участвует ли он в поддержании барьеров на последующих этапах развития.

Целью работы является изучение роли белков CP190 и CG9879 в регуляции активности генов в сперматогенезе *Drosophila melanogaster*. Для достижения этой цели были поставлены следующие **задачи**:

1. Разработать генетическую систему для специфичной инактивации белка CP190 в клетках мужского зародышевого пути дрозофилы.
2. Установить сайты связывания CP190 на хромосомах клеток мужского зародышевого пути на разных этапах дифференцировки. Определить влияние CP190 на экспрессию генов в семенниках и исследовать динамику его связывания с регуляторными элементами генома, ассоциированными с tTAF и tMAC.
3. Исследовать связывание CP190 с хромосомами на границах доменов H3K27me3 в терминально дифференцированных клетках дрозофилы.
4. Исследовать связывание белка CG9879 с генами-мишенями транскрипционных активаторов tTAF и tMAC в клетках мужского зародышевого пути.

5. Получить мутантов с потерей функции гена CG9879 и определить влияние этой мутации на экспрессию генов в семенниках дрозофилы.

Научная новизна и научно-практическая ценность. В данной работе впервые была разработана генетическая система условного спасения мутантов *Sp190*, которая позволила получить взрослых мух, в сперматоцитах которых утрачен белок CP190. При помощи этой системы впервые показана связь CP190 с регуляцией генов дифференцировки, чья активация происходит под действием tMAC. С использованием метода DamID-seq белок CP190 был впервые прокартирован на хромосомах в нескольких типах клеток дрозофилы *in vivo* и была продемонстрирована функциональная связь между CP190 и tMAC в сперматоцитах. Полученные данные позволили показать, что поддержание границ доменов H3K27me3 не является функцией CP190 на поздних этапах развития. В работе также впервые исследована роль белка CG9879 в регуляции транскрипции. Было показано, что CG9879 и tTAF колокализуются на хромосомах сперматоцитов, что указывает на функциональное взаимодействие между ними. Белок CG9879 гомологичен базальному фактору транскрипции TBP, поэтому полученные результаты представляются крайне полезными в силу возрастающего числа исследований, посвященных функции TBP и его паралога в дифференцировке клеток (Mishal and Luna-Arias, 2022).

Основные положения, выносимые на защиту:

1. Белок CP190 *Drosophila melanogaster* изменяет свой паттерн связывания с хромосомами в процессе дифференцировки клеток мужского зародышевого пути и в сперматоцитах локализуется в сайтах связывания семенник-специфичных транскрипционных факторов tMAC и tTAF, участвуя в регуляции tMAC-зависимых генов дифференцировки.

2. Белок CP190 не обязателен для поддержания функции барьеров на границах доменов H3K27me3 в терминально-дифференцированных клетках дрозофилы.

3. Семенник-специфичный транскрипционный фактор CG9879 регулирует в сперматоцитах *D. melanogaster* гены дифференцировки вместе с факторами tTAF и tMAC, однако его потеря слабо влияет на экспрессию генов в семенниках и не приводит к стерильности самцов.

Вклад автора. Большинство результатов, представленных в работе, получены автором лично. Работа по математической обработке данных выполнена совместно с к.б.н. Лактионовым П.П. и к.б.н. Максимовым Д.А. (ИМКБ СО РАН). Иммуноокрашивание и анализ препаратов на конфокальном микроскопе проводил к.б.н. Шлома В.В. (ИМКБ СО РАН).

Публикации. По материалам диссертации опубликованы 3 статьи в изданиях, индексируемых в научных базах данных «Scopus», «Web of Science» и РИНЦ.

Апробация работы. Результаты и материалы работы были представлены на международных конференциях Хромосома (2018), а также Chromosomes and Mitosis (2019). Часть работы была опубликована в сборнике «Методы редактирования генов и геномов» (2020).

Структура диссертации. Диссертация состоит из введения, описания материалов и методов, результатов, обсуждения, а также выводов и списка литературы, в которые входит 242 ссылки. Работа изложена на 130 страницах, содержит 15 рисунков и 5 таблиц.

Благодарности. Работа выполнена в лаборатории геномики ИМКБ СО РАН. Автор выражает глубокую признательность к.б.н. Белякину С.Н. и к.б.н. Лактионову П.П. за поддержку и помощь при проведении всех этапов исследования. Также автор благодарит к.б.н. Шлому В.В. за помощь в проведении экспериментов по иммуноокрашиванию фиксированных препаратов. Автор особенно благодарен к.б.н. Корякову Д.Е. за ценные советы и поддержку на протяжении всей работы.

Работа поддержана грантами РФФИ № 19-34-90108, 17-00-00181 и 13-04-01731.

СОДЕРЖАНИЕ РАБОТЫ

ОБЗОР ЛИТЕРАТУРЫ

Обзор литературы состоит из трёх разделов. В разделе 1.1. описаны стадии развития клеток мужского зародышевого пути в семенниках и подробно обсуждены генетические механизмы регуляции дифференцировки на терминальных этапах сперматогенеза. Особое внимание уделено механизмам действия транскрипционных активаторов tTAF и tMAC. В данном разделе сформулированы гипотезы об участии белков CP190 и CG9879 в масштабной активации генов дифференцировки сперматоцитов. В разделе 1.2. обсуждена функция белка CP190 в эпигенетической регуляции. Особое внимание уделено механизмам действия этого белка и его партнеров в контроле дистальных взаимодействий в геноме. В разделе подробно рассмотрен вопрос об участии CP190 в поддержании структуры хроматина на примере доменов репрессии H3K27me3/Polysomb. Кроме того, обсужден вопрос динамичной активности CP190 в клетках как о возможном механизме эпигенетической регуляции. В разделе 1.3. обсуждена роль гомологов белка CG9879 в транскрипционной регуляции. Подробно описаны механизмы их действия с упором на регуляцию клеточной дифференцировки.

МАТЕРИАЛЫ И МЕТОДЫ

Генетические конструкции. Для картирования белков CG9879 и CP190 в клетках мужского зародышевого пути методом DamID-seq были созданы конструкции *pUAST-attB-hsp70-Stop-Dam-CG9879* (GenBank KY930504) и *pUAST-attB-hsp70-Stop-Dam-CP190* (GenBank MT832326). Конструкция для условного спасения мутантов *Cp190* была заклонирована

в вектор *CP190-mCD8-GFP* (GenBank ON783212). Для удаления *CG9879* методом CRISPR/Cas9 последовательности химерных направляющих РНК были клонированы в плазмиду *pU6-BbsI-chiRNA* (Addgene, #45946).

Трансформация генома дрозофилы. Конструкции для DamID и условного спасения были введены в геном мух с сайтом интеграции *attP40* (BDSC #36304 и #2364) при помощи системы ϕ C31-опосредованной трансформации (Bischof et al., 2007). Так были получены линии $y^1w^{67};hsp70-stop-Dam-CG9879$, $y^1w^{67};hsp70-stop-Dam-CP190$, $y^1w^{67};CP190-mCD8-GFP$. Для удаления *CG9879* при помощи CRISPR/Cas9-опосредованного мутагенеза эмбрионы $y^1, M\{vas-Cas9, U6-tracrRNA\}ZH-2A, w^{1118}; MI04214$ (BDSC #37285 и #51326) трансформировали набором плазмид, кодирующих направляющие РНК. Таким образом были получены линии $y^1w^{67};CG9879^{A1}$ и $y^1w^{67};CG9879^{A2}$.

Линии мух и скрещивания. В экспериментах использовались линии из лабораторного фонда. Трансгенные линии $y^1w^{67};hsp70-stop-Dam-Comr$, $y^1w^{67};hsp70-stop-Dam$ использовали в экспериментах DamID. Линия $y^1w^{67};nanos-Cre$ содержала ген рекомбиназы Cre под управлением тканеспецифичного промотора (Лактионов et al., 2013). В экспериментах по условному спасению использовали линию $y^1w^{67};Ubi-stop-mCD8-GFP$ (Лактионов et al., 2013). Линии с мутациями *Cp190²*, *Cp190³*, *mip40^{EY16520}*, *bam^{delta86}*, *can¹* сводили с трансгенными линиями для экспериментов DamID, RNA-seq или условного спасения. Чтобы получить самцов с Cre-опосредованной рекомбинацией в клетках мужского зародышевого пути, скрещивали самок, несущих конструкцию с *loxP*-сайтами, с самцами, несущими *nanos-Cre*. Для DamID CP190 в личиночных слюнных железах стандартными скрещиваниями были получены линии $y^1w^{67};hsp70-Dam-CP190$ и $y^1w^{67};hsp70-Dam$.

Микроскопия. Гибридизацию РНК в семенниках *in situ* и подбор зондов проводили по методу, описанному в Morris et al. 2009. Для анализа флуоресценции GFP семенники переносили на предметное стекло и сразу исследовали на флуоресцентном микроскопе. Для иммуноокрашивания семенники фиксировали стандартным методом (Solovei and Cremer, 2010), использовали первичные крысиные антитела anti-CP190 (Golovnin et al., 2012) и вторичные антитела A-11077 (Thermo Fisher). Препараты анализировали на конфокальном микроскопе Carl Zeiss LSM 710.

Анализ транскриптома методом RNA-seq. РНК из семенников выделяли при помощи TRIZOL. Пробоподготовку библиотек проводили с использованием набора Illumina Truseq. Эффект делеции *CG9879* исследовали у трехдневных самцов $y^1w^{67}/Y;CG9879^{A1}$, а в качестве контроля использовали семенники y^1w^{67} . Результаты RNA-seq мутантов с делецией *CG9879* доступны в базе данных Gene Expression Omnibus (GEO) по идентификатору GSE97129. Эффект конструкций для условного спасения на экспрессию генов в семенниках исследовали у трехнедельных самцов $y^1w^{67}/Y;CP190-mCD8-GFP/nanos-Cre;Cp190^2, e^1/Cp190^3, e^1$ и $y^1w^{67}/Y;CP190-mCD8-$

GFP/nanos-Cre, и в качестве контроля использовали семенники самцов $y^1w^{67};Ubi-stop-mCD8-GFP/nanos-Cre$. Данные RNA-seq в семенниках со спасением *Cp190* были загружены в базу данных BioProject (PRJNA847720).

DamID-seq. Методология DamID-seq в клетках мужского зародышевого пути описана в Laktionov et al. 2014. Протокол подготовки образцов проводили как описано в Greil et al. 2006 с модификациями, указанными в Maksimov et al. 2016. Биоинформатический анализ результатов проводили при помощи алгоритма DamID-seq (Maksimov et al., 2016). Данные по связыванию CG9879 в клетках мужского зародышевого пути доступны в GEO по идентификатору GSE97182. DamID-seq CP190 в семенниках дикого типа доступен по идентификатору GSE155579.

РЕЗУЛЬТАТЫ

Генетическая система для условного спасения мутантов *Cp190*.

Мутанты *Cp190* не доживают до взрослой стадии. Поэтому чтобы исследовать регуляторную функцию CP190 в клетках зародышевого пути взрослых самцов, мы разработали генетическую систему условного спасения мутантов *Cp190*. Система состоит из двух генетических конструкций (Рисунок 1А). Спасаящая конструкция *CP190-mCD8-GFP* содержит кодирующую последовательность гена *Cp190* под контролем конститутивного промотора гена *Ubi63E*. Вслед за *Cp190* помещен гибридный терминатор транскрипции *HIS3-SV40* и кодирующая последовательность гена химерного мембраносвязанного репортерного белка *mCD8-GFP*. Кодирующая последовательность *Cp190* и терминатор транскрипции фланкированы *loxP*-сайтами узнавания рекомбиназы Cre и образуют спасающую кассету, которую можно удалять путем Cre/loxP-опосредованной рекомбинации в целевых клетках. Вторая конструкция кодирует рекомбиназу Cre под управлением промотора гена *nanos*, который обеспечивает наработку этого белка исключительно в клетках зародышевого пути. При наличии обеих конструкций в геноме мутантов *Cp190* в соматических клетках нарабатывается белок CP190, что обеспечивает спасение мутантного фенотипа (Рисунок 1Б). В то же время, в клетках мужского зародышевого пути спасающая кассета удаляется рекомбиназой, в результате *mCD8-GFP* транскрибируется с промотора *Ubi63E*, что позволяет получить *Cp190*-дефицитные клетки, помеченные GFP.

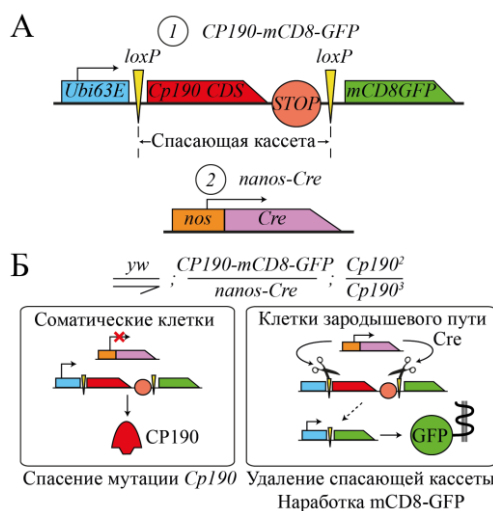


Рисунок 1. Система условного спасения для удаления белка CP190 в клетках мужского зародышевого пути. А) Схема конструкций. Б) Принцип работы системы.

Одной дозы конструкции *CP190-mCD8-GFP* было достаточно для спасения летального фенотипа компаунд-гетерозигот *Cp190²/Cp190³*, у которых не образуется функциональный белок CP190. Самцы с условным спасением (*y^{1w67};CP190-mCD8-GFP/nanos-Cre;Cp190²/Cp190³*) обладали нормальной жизнеспособностью и не демонстрировали видимых нарушений сперматогенеза. Анализ витальных препаратов семенников позволил установить, что *nanos-Cre* обеспечивает продукцию GFP специфично в клетках мужского зародышевого пути начиная со стадии сперматоцитов (Рисунок 2). Это говорит о том, что удаление спасающей кассеты происходит в клетках мужского зародышевого пути эффективно.

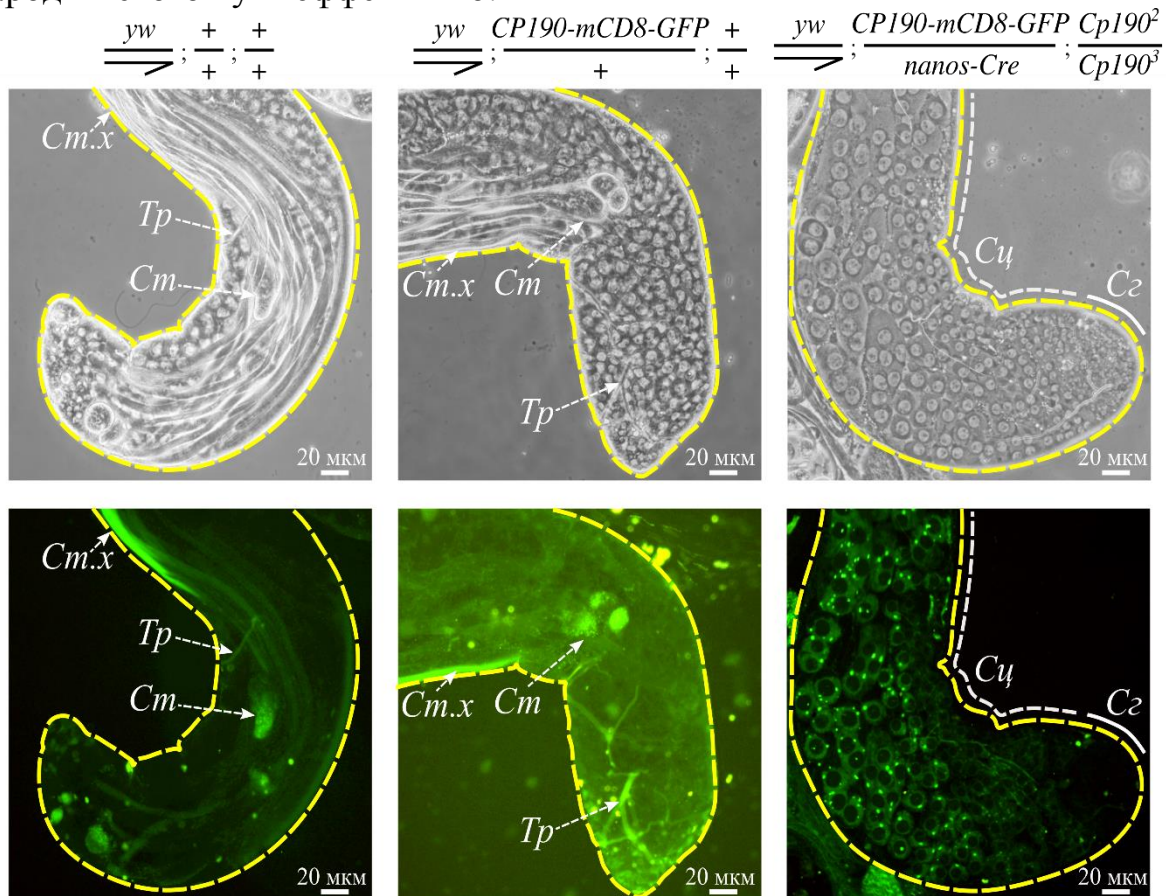


Рисунок 2. Флуоресценция GFP в семенниках у мух дикого типа, мух с одной дозой спасающей конструкции и мух с условным спасением мутации *Cp190*. В верхнем ряду показаны изображения семенников в фазовом контрасте. В нижнем ряду показана флуоресценция GFP. В семенниках с условным спасением флуоресценция GFP наблюдается во всех клетках мужского зародышевого пути начиная с ранних сперматоцитов. В отсутствие *nanos-Cre* в семенниках наблюдается только фоновая автофлуоресценция, характерная для семенников дикого типа. Условные обозначения: *Cm* – головы сперматид, *Cm.x* – хвосты сперматид, *Tr* – дыхательные трубочки, *Cy* – область расположения сперматоцитов, *Cg* – область расположения сперматогоний. Масштаб – 20 мкм.

Система условного спасения значительно снижает уровень эктопической экспрессии *Cp190* в клетках мужского зародышевого пути имаго. Иммуноокрашивание трехнедельных семенников антителами против белка CP190 позволило выявить потерю данного белка в клетках муж-

ского зародышевого пути, при этом только в единичных сперматоцитах белок сохранялся (**Рисунок 3**). Это доказывает, что разработанная система позволяет удалять CP190 в клетках мужского зародышевого пути с высокой эффективностью.

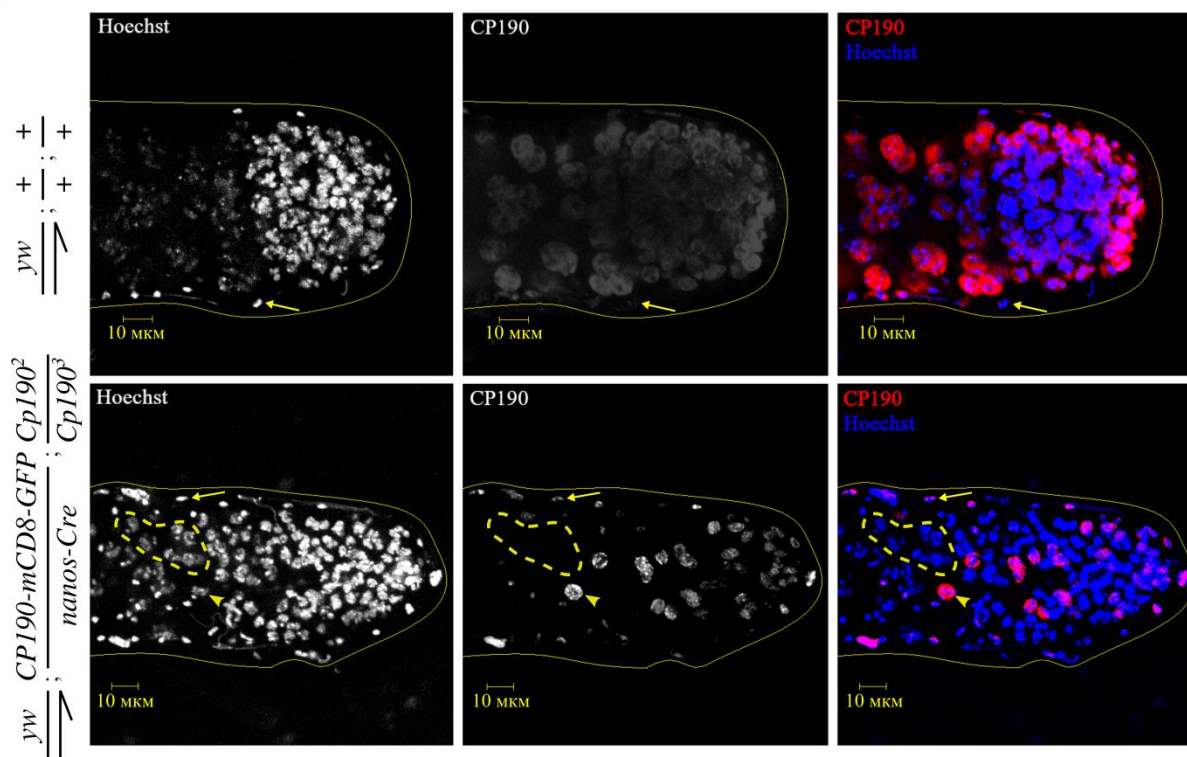


Рисунок 3. Cre-опосредованная рекомбинация приводит к удалению спасающей кассеты и утрате CP190 в клетках мужского зародышевого пути. Конфокальные изображения трехнедельных семенников с иммуноокрашиванием антителами против CP190. В верхнем ряду – семенник *yw*. В нижнем ряду – семенник с условным спасением. Окраску ДНК проводили при помощи Hoechst. Непрерывной желтой линией подчеркнуты границы семенников. Длинной стрелкой показаны соматические клетки оболочки. Пунктирной линией показано скопление CP190-негативных сперматоцитов. Короткой стрелкой показан CP190-позитивный сперматоцит. Масштаб – 10 мкм.

Мутация *Cp190* влияет на экспрессию генов дифференцировки в клетках мужского зародышевого пути. Для поиска генов, чья активность зависит от CP190 в клетках мужского зародышевого пути, мы проанализировали транскриптом трехнедельных семенников методом RNA-seq. Однако мы не могли исключить, что в соматических клетках семенников с условным спасением уровень экспрессии CP190 был повышен по сравнению с нативным уровнем. Поэтому сначала мы оценили, как повышенная экспрессия *Cp190* в соматических клетках влияет на транскриптом целых семенников. Для этого мы исследовали экспрессию генов в семенниках самцов *y¹w⁶⁷;CP190-mCD8-GFP/nanos-Cre;+/+*, в которых функционирует эндогенный белок CP190, а спасающая кассета активна только в соматических клетках (**Рисунок 4А**). Это позволило показать, что оверэкспрессия *Cp190* в соматических клетках приводит к умеренным изменениям транскрипции в сравнении с контролем дикого типа: только 12 генов изменяли экспрессию более чем в 8 раз. При анализе же семенников с условным спасением было

обнаружено 89 таких генов, что является значительно более сильным эффектом. Такое контрастное отличие нельзя объяснить повышенной экспрессией CP190 в соматических клетках семенника, а лишь утратой этого белка в сперматоцитах.

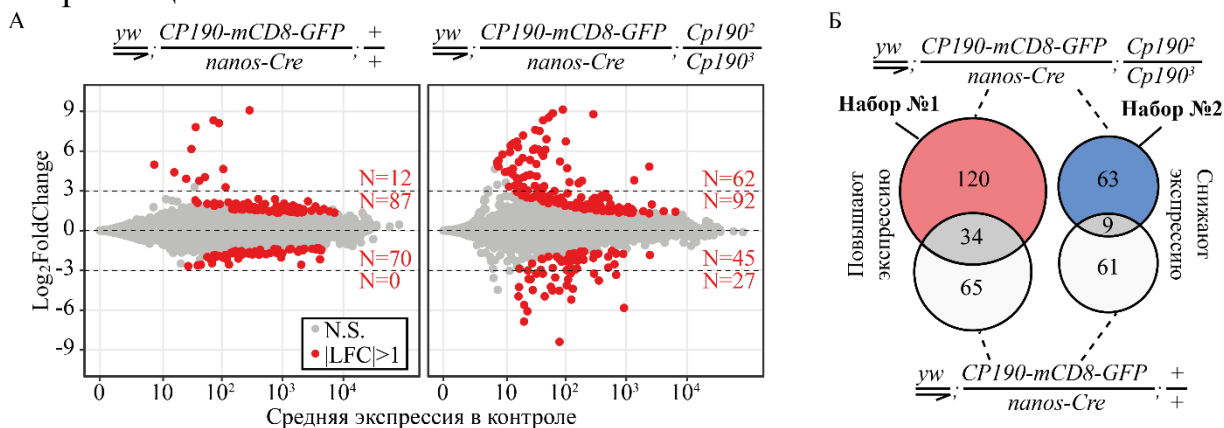


Рисунок 4. Утрата CP190 в клетках мужского зародышевого пути приводит к нарушению активности генов. **А)** Дифференциальная экспрессия генов в семенниках с повышенной экспрессией CP190 в соматических клетках (диаграмма слева) или в семенниках с условным спасением (диаграмма справа) по результатам RNA-seq. По вертикальной оси отложен логарифмированный диапазон изменения экспрессии (Log₂FoldChange, LFC) по сравнению с контролем. По горизонтальной оси отложена нормализованная средняя экспрессия в контроле дикого типа. Гены с достоверным изменением экспрессии (|LFC|>1, P<0.05) помечены красным цветом, их число указано красным текстом. Пунктирными линиями указаны значения LFC ±3 и 0. **Б)** Пересечение между множествами генов, достоверно повысившими и понизившими экспрессию в семенниках с условным спасением или семенниках с повышенной экспрессией CP190 в соматических клетках (красные точки на **Диаграмме А**). Гены, изменившие экспрессию только на фоне условного спасения, обозначены как Набор №1 и Набор №2.

Далее, мы определили гены, чья активность была нарушена в результате утраты CP190 в клетках мужского зародышевого пути (**Рисунок 4Б**). Для этого мы нашли гены, чья активность изменяется на фоне условного спасения в два раза и больше, и исключили из них те, что изменили экспрессию в семенниках с оверэкспрессией. Таким образом было выявлено 183 гена, среди которых 120 увеличивали активность на фоне утраты (Набор №1), а 63 снижали (Набор №2).

В дальнейшем мы хотели определить, какое влияние утрата CP190 в клетках мужского зародышевого пути оказывает на экспрессию генов дифференцировки в сперматоцитах, что требовало от нас выявить паттерн экспрессии CP190-зависимых генов в клетках мужского зародышевого пути. Поэтому мы воспользовались данными транскриптома единичных клеток семенника из базы Fly Cell Atlas, которые позволяют установить, в каких клеточных типах семенника транскрибируются гены (Li et al., 2022) (**Рисунок 5А**). Кластеризация генов по уровням активности в популяциях клеток семенника позволила установить, что большинство генов из Набора №1 транскрибируются специфично в сперматоцитах как гены дифференцировки. В то же время, в Наборе №2 большинство генов функционирует в

специализированных популяциях соматических клетках, и только небольшую их долю можно охарактеризовать как гены дифференцировки сперматоцитов. Таким образом, одной из функций *CP190* в клетках мужского зародышевого пути является регуляция генов дифференцировки сперматоцитов, причем он оказывает на них преимущественно репрессорное воздействие.

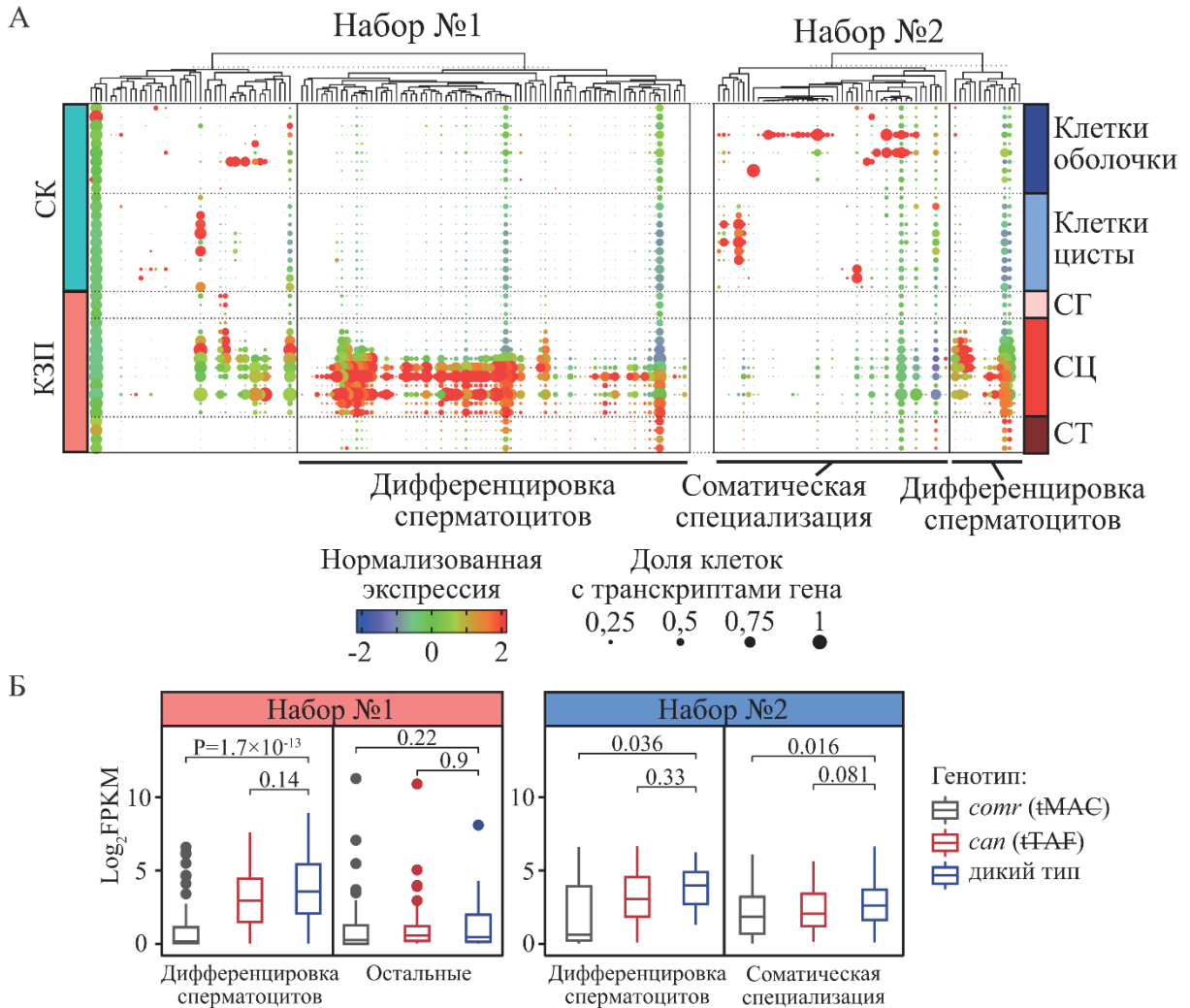


Рисунок 5. Утрата *CP190* в сперматоцитах приводит к нарушению активности tMAC-зависимых генов дифференцировки. А) Экспрессия генов из Наборов №1 и №2 в клеточных популяциях семенника по данным Fly Cell Atlas. Цветными прямоугольниками показана принадлежность популяций клеточным типам. Цвет точек характеризует нормализованный уровень экспрессии гена в популяции клеток. Размер точек описывает долю клеток в популяции, в которых транскрипты гена были обнаружены. В каждом наборе гены были кластеризованы на две группы. В Наборе №1 выделяются гены с высокой активностью в сперматоцитах (гены дифференцировки сперматоцитов). В Наборе №2 кластеры можно охарактеризовать как гены соматической специализации и гены дифференцировки сперматоцитов. СК – соматические клетки, КЗП – клетки зародышевого пути, СГ – сперматогонии, СЦ – сперматоциты, СТ – сперматиды. Б) Экспрессия генов из Наборов №1 и №2 в семенниках дикого типа и мутантов с нарушением функции tTAF (мутанты *can*) или tMAC (мутанты *comr*) по данным RNA-seq. Достоверность отличий между мутантами и диким типом определена в тесте Вилкоксона.

Мутация *Sp190* нарушает экспрессию генов-мишеней tMAC. Чтобы определить, зависит ли активность генов из Наборов №1 и №2, от

транскрипционных активаторов tTAF и tMAC, мы воспользовались полученными ранее данными транскриптомов в семенниках у мутантов с нарушением сперматогенеза (Laktionov et al., 2018) (**Рисунок 5Б**). Эти данные содержали информацию об активности генов в семенниках дикого типа, а также семенниках с остановкой сперматогенеза на стадии сперматоцитов из-за потери функции tTAF (мутация *can*) и tMAC (мутация *comr*). Оказалось, что утрата tTAF почти не влияет на экспрессию CP190-зависимых генов. В то же время, утрата tMAC приводит к снижению их активности, причем наиболее сильно от этого комплекса зависят гены дифференцировки сперматоцитов из Набора №1. Полученные результаты указывают, что утрата CP190 в сперматоцитах нарушает активность tMAC-зависимых генов дифференцировки.

Дифференцировка клеток сопровождается динамичным связыванием CP190 с хроматином. Чтобы выяснить, связывается ли CP190 с регуляторными элементами генома, ассоциированными с tMAC и tTAF, мы воспользовались методом DamID-seq, который позволил нам прокартировать сайты связывания CP190 на хромосомах клеток мужского зародышевого пути (Laktionov et al., 2014). Учитывая соотношение клеток в семенниках, полученный профиль позволил охарактеризовать события, которые происходят в сперматоцитах после индукции генов дифференцировки. А картируя этот белок в клетках зародышевого пути у мутантов *bam*, *tip40* или *can*, мы фиксировали сайты связывания CP190 в сперматогониях и сперматоцитах с утраченной функцией tMAC или tTAF (**Рисунок 6А, Б**). Мы также исследовали локализацию CP190 на политенных хромосомах личиночных слюнных желез, поскольку в этой ткани гены дифференцировки сперматоцитов полностью репрессированы. Сравнение полученных профилей позволяло нам не только охарактеризовать связь между CP190 и активаторами транскрипции в сперматоцитах, но также дало возможность выявить динамику связывания этого белка с хроматином, что важно для понимания функции CP190 в ходе дифференцировки.

С использованием полученных профилей мы определили, связан ли CP190 с генами дифференцировки, которые регулируются двумя комплексами, и посчитали число прямых и непрямых мишеней tTAF и tMAC среди генов, ассоциированных с CP190. Это позволило выявить, что CP190 связан именно с прямыми мишенями (**Рисунок 6В**). В слюнных железах, сперматогониях и сперматоцитах у мутантов *tip40* эта ассоциация нарушена, указывая, что связывание CP190 с этими генами зависит не только от типа клеток, но и от стадии дифференцировки. У мутантов *tip40* комплекс tMAC не связывается с хромосомами, указывая на то, что tMAC необходим для привлечения CP190 к генам дифференцировки сперматоцитов. При этом CP190 чаще ожидаемого встречается вблизи сайтов связывания tTAF и tMAC в клетках мужского зародышевого пути дикого типа и в сперматоцитах мутантов *can*, у которых нарушена функция tTAF, но не в других типах клеток

(Рисунок 6Г). Таким образом, tMAC необходим для связывания CP190 с регуляторными элементами генома, с которыми взаимодействуют tTAF и tMAC. Это указывает на функциональную связь между CP190 и tMAC.

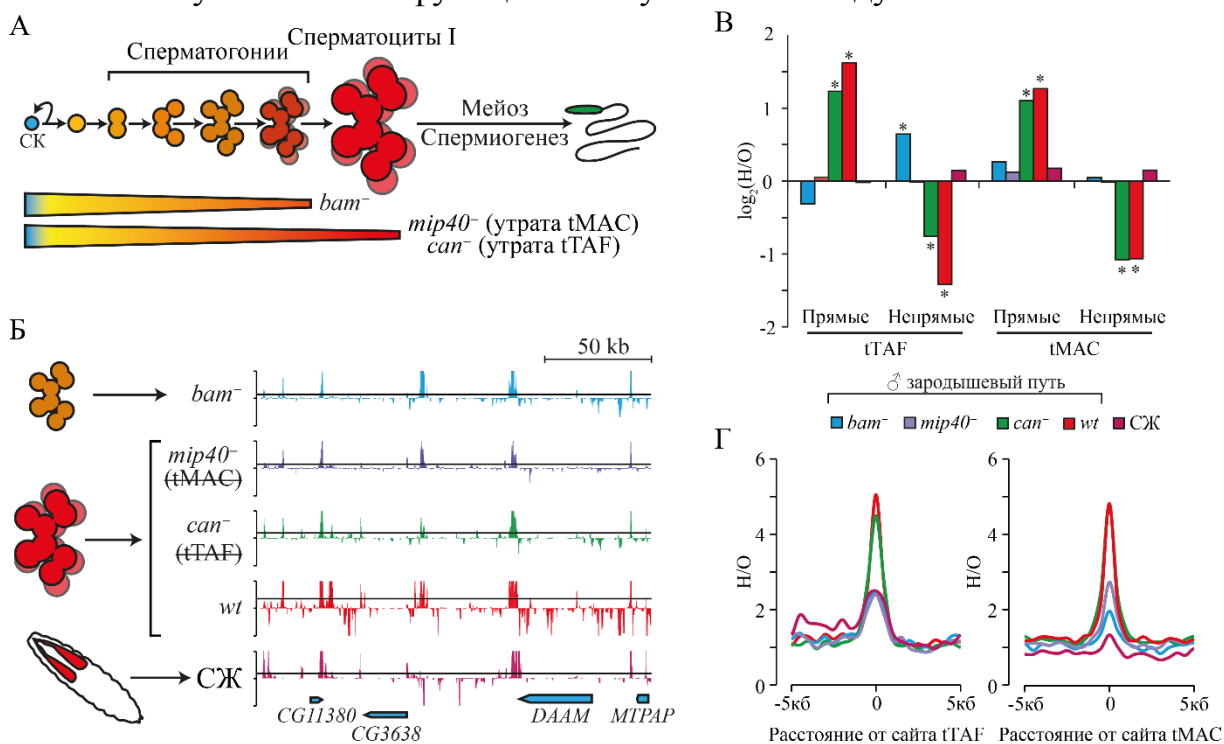


Рисунок 6. В ходе дифференцировки клеток мужского зародышевого пути CP190 перераспределяется к генам дифференцировки сперматоцитов. А) Мутации *bam*, *can* и *mip40* останавливают развитие клеток мужского зародышевого пути на стадии сперматогоний или сперматоцитов. В мутантных семенниках большинство клеток мужского зародышевого пути представлены сперматогониями (мутанты *bam*) или недифференцированными сперматоцитами (мутанты *can* и *mip40*). **Б)** Методом DamID-seq CP190 был прокартирован на хромосомах клеток мужского зародышевого пути дикого типа (*wt*) и мутантов *can*, *mip40* и *bam*, а также на политенных хромосомах личиночных слюнных желез (СЖ). **В)** Среди генов, ассоциированных с CP190 в промоторной области, подсчитывали число прямых и непрямых мишеней tTAF и tMAC. Наблюдаемое число (Н) сравнивали с ожидаемым (О) при помощи теста Хи-квадрат. **Г)** Подсчитывали число сайтов связывания CP190 на указанном расстоянии от предсказанных ранее сайтов связывания tTAF и tMAC. Наблюдаемое число сравнивали с ожидаемым. Условные обозначения: * – P-value < 0.001.

Активная транскрипция и SetDB1, но не связывание CP190, маркирует границы доменов H3K27me3. В данной работе мы впервые прокартировали сайты связывания CP190 в терминально-дифференцированных клетках дрозофилы *in vivo*, включая клетки личиночных слюнных желез. Ранее в этих клетках метка H3K27me3 была прокартирована ранее методом ChIP-chip (Sher et al., 2012). С использованием этих данных мы обнаружили, что распределение CP190 вокруг доменов почти не коррелирует с положением границ (Рисунок 7А). Таким образом, в политенных хромосомах совпадение CP190 с границами доменов H3K27me3 имеет статистически-случайный характер, что позволяет сделать вывод о том, что на полногеномном уровне CP190 не важен для поддержания границ доменов в личиночных

слюнных желез. Мы также обнаружили, что белок SetDB1, который известен как гистонметилтрансфераза, ответственная за модификацию гистона H3 в положении K9, демонстрировал явную предрасположенность к связыванию на границах доменов H3K27me3. Было выявлено, что высокая транскрипция SetDB1-связанных генов служит лучшим маркером границ доменов H3K27me3 (**Рисунок 7Б**), поскольку половина всех границ располагается на расстоянии менее 5 кб от сайтов связывания SetDB1, и уровень транскрипции вблизи пограничных сайтов многократно превышает аналогичный показатель в удаленных от H3K27me3 сайтах SetDB1.

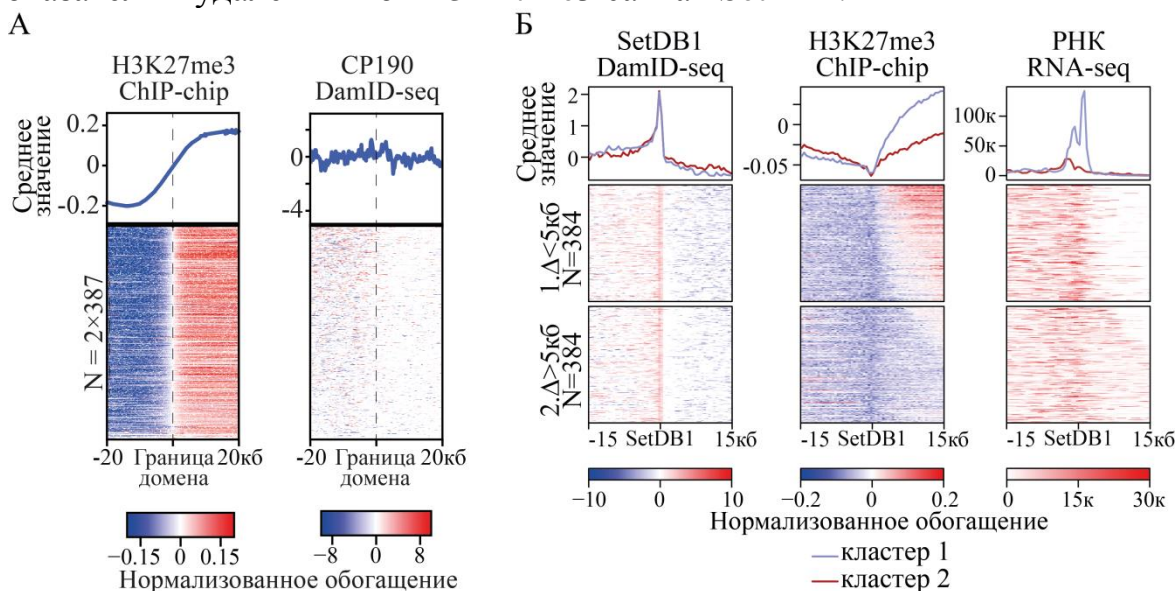


Рисунок 7. CP190 не демонстрирует связи с границами доменов H3K27me3 в политенных хромосомах личиночных слюнных желез. А) Распределение CP190 вблизи границ доменов H3K27me3 случайно. Положения границ определяли с помощью скрытой марковской модели. На тепловой карте показано обогащение локусов вокруг установленных границ доменов заданным белком. Границы доменов являются точкой отсчета, координаты нуклеотидов вне домена отсчитывается со знаком минус, а внутри домена со знаком плюс. Б) Границы доменов H3K27me3 были отсортированы по дистанции до ближайшего сайта связывания SetDB1, и разделены на две группы по дистанции от SetDB1 до границы домена (Δ). Для каждого пограничного сайта SetDB1 было определено обогащение SetDB1, H3K27me3, и покрытие прочтениями в RNA-seq. Сайты связывания SetDB1 взяты за точку отсчета и ориентированы таким образом, чтобы домены H3K27me3 находились справа.

Роль гомолога ТВР CG9879 в регуляции генов дифференцировки мужского зародышевого пути. Чтобы исследовали гипотезу о том, что CG9879 служит промежуточным фактором, ответственным за регуляцию непрямых мишеней tTAF и tMAC, мы использовали DamID-seq для картирования сайтов связывания CG9879 на хромосомах в клетках мужского зародышевого пути дикого типа. Однако оказалось, что этот белок зачастую обнаруживается в регуляторных областях прямых мишеней комплексов (**Рисунок 8А**). При этом исследуемый белок колокализуется с семенник-специфичными транскрипционными активаторами, демонстрируя сильное сродство к сайтам связывания tTAF (**Рисунок 8Б**). В свою очередь, поиск коротких консенсусных последовательностей ДНК в сайтах связывания CG9879

позволил обнаружить обогащение АТ-богатыми мотивами, схожими с ТАТА-боксом (**Рисунок 8В**). В промоторах белок-кодирующих генов этот мотив привлекает белок ТВР, который является ключевым компонентом базального фактора транскрипции TFIIID. Примечательно, что комплекс tTAF является сперматоцит-специфичным паралогом TFIIID, тогда как белок CG9879 гомологичен ТВР. Это наблюдение позволяет предположить, что CG9879 может участвовать в регуляции генов дифференцировки в составе комплекса с tTAF.

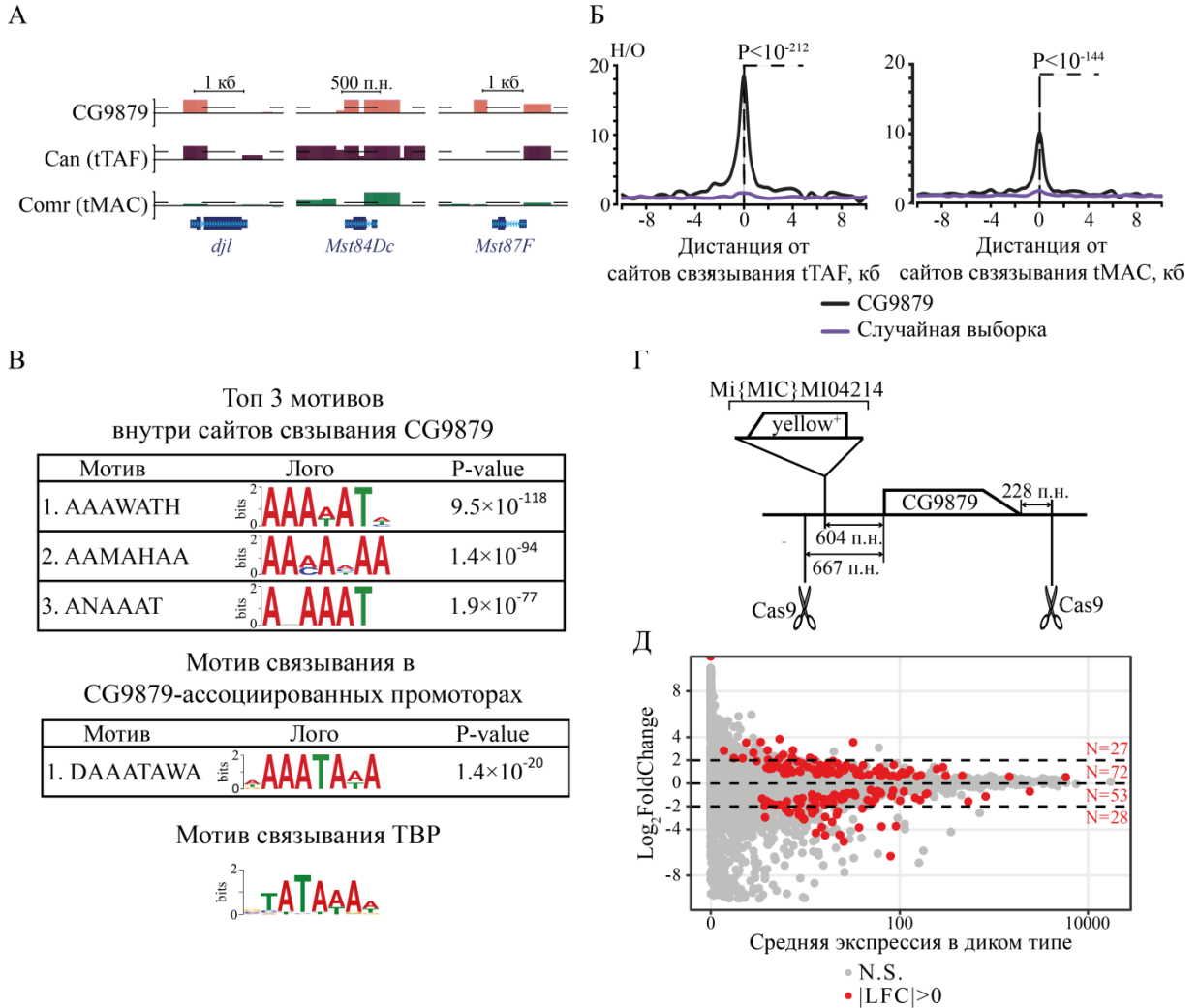


Рисунок 8. CG9879 колокализуется с белками Can и Comr в клетках мужского зародышевого пути. А) DamID-профили белков CG9879, Can (tTAF), Comr (tMAC) в клетках мужского зародышевого пути вблизи генов дифференцировки сперматоцитов *djl*, *Mst84Dc* и *Mst87F*. **Б)** На диаграммах показано отношение наблюдаемой (Н) и ожидаемой (О) частоты пиков CG9879 (черная кривая) и случайного набора геномных локусов (фиолетовая линия) на заданном удалении от сайтов tTAF и tMAC. Достоверность обогащения CG9879 посчитана с помощью биномиального теста. **В)** Сайты связывания CG9879 содержат АТ-богатые мотивы. Для сравнения показан мотив связывания ТВР. **Г)** Для CRISPR/Cas9-опосредованного удаления гена *CG9879* использовали линию мух со встройкой транспозона *MI04214*. Сайты с распознавания химерных направляющих РНК показаны ножницами. **Д)** Делеция *CG9879* приводит к незначительным изменениям транскрипции. На диаграмме показана дифференциальная экспрессия транскриптов в семенниках с делецией *CG9879^{d1}*. Пунктирными линиями показаны граничные значения

$\log_2\text{FoldChange}$ -2, 0 и 2. Число красных точек в интервалах $\log_2\text{FoldChange}$ между граничными значениями показано красным текстом.

Чтобы понять, как CG9879 влияет на экспрессию генов в семенниках, мы удалили кодирующий его ген с помощью системы CRISPR/Cas9 (**Рисунок 8Г**). Удивительно, но мутанты с гомозиготной делецией не демонстрировали никаких морфологических дефектов в семенниках или нарушений фертильности. Более того, анализ транскриптома показал, что только 28 генов значительно снижали экспрессию в семенниках в ответ на утрату CG9879, но ни один из них не был связан этим белком (**Рисунок 8Д**). Мы предположили, что такой незначительный эффект от мутации связан с активностью в сперматоцитах других гомологов ТВР.

ОБСУЖДЕНИЕ

Исследование функции генов в системе условного спасения. Нами была разработана методология для условного спасения мутаций в соматических клетках, что позволило исследовать регуляторные эффекты *Cr190* в клетках зародышевого пути самцов дрозофилы. Высокую эффективность предложенной нами системы подтверждают результаты иммуноокрашивания, согласно которым наработка CP190 происходит только в единичных сперматоцитах. Преимуществом данной системы является то, что целевые клетки, в которых редактируется спасающая конструкция, оказываются помечены репортерным белком GFP. Благодаря этому изучаемую популяцию клеток можно изолировать с использованием сортировки. Принцип работы полученной системы универсален, что позволяет использовать аналогичный подход для исследования других факторов.

Роль белка CP190 в регуляции генов дифференцировки сперматоцитов. В этой работе мы провели подробное исследование эффекта белка CP190 на активность генов в клетках зародышевого пути *D. melanogaster*. Наиболее контрастный эффект, который мы наблюдали в семенниках с условным спасением, заключался в увеличении активности tMAC-зависимых генов дифференцировки сперматоцитов. Вместе с тем, картирование CP190 позволило обнаружить колокализацию данного архитектурного белка с участками генома, ассоциированными с tMAC. Можно предположить, что утрата CP190 приводит к сбоям в коммуникации генов дифференцировки сперматоцитов и их регуляторных элементов. Чтобы проверить это предположение, необходимо дополнить полученные в данном исследовании результаты анализом архитектуры ядра в развивающихся мужских половых клетках.

Динамика связывания CP190 с хроматином в процессе дифференцировки клеток мужского зародышевого пути. До настоящего момента анализ связывания CP190 с хромосомами дрозофилы на полногеномном уровне производили в клеточных культурах или ранних эмбрионах. Было продемонстрировано, что расположение сайтов связывания инсуляторных

белков дрозофилы в основном одинаково для разных типов клеток. Полученные нами данные демонстрируют крайне вариабельный паттерн связывания CP190 с хроматином. Мы предполагаем, что такую динамику можно объяснить несколькими причинами. Во-первых, данные на клетках в культуре могут не отражать должным образом события, происходящие в клетках организма. Во-вторых, установленная нами динамика может объясняться специфичным для клеток мужского зародышевого пути механизмом регуляции CP190.

SetDB1, но не CP190, может участвовать в поддержании доменной организации хромосом. С использованием полученных данных мы хотели проверить гипотезу о том, что CP190 поддерживает границы доменов H3K27me3 на постэмбриональных этапах развития и исследовали колокализацию CP190 с границами этих доменов в политенных хромосомах личиночных слюнных желез. Однако мы не обнаружили обогащения CP190 на границах. Напротив, он зачастую находился внутри доменов. Это означает, что поддержание доменной организации районов, обогащенных H3K27me3, не является функцией CP190. Вместе с тем, в данном исследовании мы впервые показали высокую корреляцию сайтов связывания гистон-метилтрансферазы SetDB1 с положением пограничных элементов триметилирования H3K27. Принимая во внимание исследования на клетках млекопитающих, где продемонстрировано участие SetDB1 в поддержании локальной архитектуры хроматина, наши результаты говорят в пользу высоко консервативной функции SetDB1, которая заключается в регуляции активности барьерных элементов генома.

Избыточная функция ТВР-подобных белков поддерживает стабильность генетической системы сперматогенеза. Полученные нами результаты указывают, что CG9879 может играть роль ТВР в комплексе с tTAF. Об этом говорит то, что в сперматоцитах CG9879 колокализуется с tTAF и связывается с областями генома, обогащенными мотивами, сходными с ТАТА-боксом. В то же время, делеция гена CG9879 привела к очень слабым изменениям в транскриптом семенников и не оказала заметного влияния на сперматогенез. Мы предполагаем, что CG9879 участвует в регуляции ограниченного круга генов дифференцировки сперматоцитов, тогда как остальные ТВР-подобные белки дрозофилы вовлечены в этот процесс в большей степени. Такая избыточность может обеспечивать стабильность генетической системы, необходимой для развития мужских половых клеток дрозофилы.

ВЫВОДЫ

1. Разработана генетическая система условного спасения мутантов *Sp190*. Она позволяет добиться утраты белка CP190 в клетках мужского зародышевого пути, начиная со стадии ранних сперматоцитов, но сохраняет его в соматических клетках.

2. Установлено, что связывание CP190 с хромосомами динамично. По мере дифференцировки клеток мужского зародышевого пути белок CP190 перераспределяется к сайтам связывания tMAC и tTAF и в сперматоцитах регулирует активность tMAC-зависимых генов дифференцировки.

3. Установлено, что в политенных хромосомах личиночных слюнных желез связывание CP190 не коррелирует с границами доменов H3K27me3.

4. Установлено, что белок CG9879 является транскрипционным фактором – семенник-специфичным паралогом TBP. В клетках мужского зародышевого пути белок CG9879 связывается с AT-богатыми мотивами в промоторах прямых мишеней tTAF и tMAC.

5. С использованием системы CRISPR/Cas9 получена полноразмерная делеция гена CG9879. Установлено, что она вызывает лишь незначительные изменения в экспрессии генов в семенниках и не приводит к нарушению фертильности самцов.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи:

1. Laktionov P.P., Maksimov D.A., **Romanov S.E.**, Antoshina P.A., Posukh O.V., White-Cooper H., Koryakov D.E., Belyakin S.N. Genome-wide analysis of gene regulation mechanisms during *Drosophila* spermatogenesis // Epigenetics and Chromatin – 2018. V. 11. – 14. DOI: 10.1186/s13072-018-0183-3. IF=5,465.

2. Kalashnikova D.A.*, Maksimov D.A.*, **Romanov S.E.***, Laktionov P.P., Koryakov D.E. SetDB1 and Su(var)3-9 play non-overlapping roles in somatic cell chromosomes of *Drosophila melanogaster* // J Cell Sci – 2021. V. 134. № 2. – jcs253096. DOI: 10.1242/jcs.253096. IF=5,235.

* - equal contribution.

3. **Romanov S.E.**, Shloma V.V., Koryakov D.E., Belyakin S.N., Laktionov P.P. Insulator protein CP190 regulates expression of spermatocyte differentiation genes in *Drosophila melanogaster* male germline // Molecular Biology – 2022. DOI: 10.1134/S0026893323010120. IF=1,540.

Тезисы конференций:

1. **Romanov S.E.**, Posukh O.V., Maksimov D.A., Belyakin S.N., Laktionov P.P. (2018). *In vivo* DamID mapping uncovered dynamic CP190 chromatin binding landscape in course of *Drosophila* spermatogenesis // International conference Chromosome – Новосибирск: Хромосома, 2018. – С. 68-69.

2. **Romanov S.E.**, Laktionov P.P., Belyakin S.N. (2019) The role of CP190 in genetic regulation of *Drosophila* spermatogenesis // International mini-conference «Chromosomes and Mitosis» – Новосибирск: НИИГУ, 2019. – С. 30.

Публикация в сборнике:

1. **Романов С.Е.**, Лактионов П.П. Получение направленных делеций в геноме *Drosophila melanogaster* с помощью системы CRISPR/Cas9 // Методы редактирования генов и геномов / отв. ред. Закиян С.М., Медведев

С.П., Дементьева Е.В., Власов В.В. – Новосибирск: Издательство СО РАН, 2020. – С. 133-148.