

Санкт-Петербургский Государственный университет

Биологический факультет

Кафедра генетики и биотехнологии

Данчурова Анастасия Александровна

Идентификация белков, взаимодействующих с теломерными последовательностями в  
ДНК и РНК.

Магистерская выпускная квалификационная работа.

Работа выполнена в Научной лаборатории биологии амилоидов

Санкт-Петербургского Государственного университета

(зав. лаб – проф. Юрий Олегович Чернов)

Научный руководитель:

к.б.н., с.н.с. Аксёнова Анна Юрьевна

Куратор:

к.б.н., н.с. Сопова Юлия Викторовна

Санкт-Петербург

2018

## Содержание

Список сокращений .....	3
Введение.....	4
1. Обзор литературы .....	5
1.1. Структура и функции теломер .....	5
1.2. Репликация теломерных областей .....	10
1.3. Структура хроматина теломерных областей .....	13
1.4. Транскрипция теломерных областей и РНК TERRA.....	14
1.5. Теломерные последовательности вне теломер и их функции.....	14
1.6. Подходы к изучению теломерных последовательностей.....	18
2. Цель работы .....	20
3. Задачи работы.....	20
4. Материалы и методы.....	21
4.1. Материалы.....	21
4.1.1 Штаммы, использованные в работе .....	21
4.1.2. Среды, использованные в работе .....	22
4.1.3. Плазмиды, использованные в работе.....	22
4.1.4. Олигонуклеотиды, использованные в работе .....	23
4.2. Выделение белков, взаимодействующих с РНК, содержащей человеческие теломерные повторы в дрожжевой системе.....	25
4.3. Масс-спектрометрия.....	27
4.4. Метод ПЦР .....	27
4.5. Рестрикция плазмид UIRL-lacO-25-12-8 и UIRL-lacO-25-12-5 рестриктазой SmaI .....	28
4.6. Дрожжевая трансформация .....	28

4.7. Бактериальная трансформация.....	28
4.8. Гель-фильтрация SUMO-LacI.....	28
4.9. Анализ взаимодействия SUMO-LacI с ДНК, содержащей теломерные повторы и последовательность <i>lacO</i> . ....	29
4.10. Описание модельной системы для выделения белков, взаимодействующих с человеческими теломерными повторами в ДНК в дрожжевой системе. ....	29
4.11. Фрагментация хроматина .....	30
4.12. Анализ структуры белка <i>in silico</i> .....	31
5. Результаты.....	32
5.1. Выделение белков, взаимодействующих с РНК, содержащей человеческие теломерные повторы в дрожжевой системе.....	32
5.2. Анализ структуры белка Ten1 <i>in silico</i> .....	35
5.3. Выделение белков, взаимодействующих с теломерными последовательностями в ДНК.....	37
5.3.1. Очистка SUMO-LacI и анализ условий взаимодействия SUMO-LacI с <i>lacO</i> ..	38
5.3.2. Конструирование штаммов <i>S.cerevisiae</i> AAY51 и AAY52. ....	42
5.3.3. Фрагментация хроматина.....	43
6. Обсуждение результатов .....	45
7. Выводы.....	49
8. Благодарности.....	50
8. Список литературы .....	51

## Список сокращений

ALT – alternative telomere lengthening (альтернативное удлинение теломер)

BIR – breakage-induced replication (индуцированная двунитевыми разрывами репликация)

ITSs – interstitial telomeric sequences (внутренние, интерстициальные теломерные последовательности)

ITL – interstitial telomere loop (внутренняя теломерная петля)

MALDI - matrix assisted laser desorption/ionization (матрично-активированная лазерная десорбция/ионизация, МАЛДИ)

PBS - Phosphate buffered saline (Натрий фосфатный буфер)

SILAC - stable isotope labeling with amino acids in cell culture (маркировка (метка) стабильными изотопами аминокислот в клеточной культуре)

STARs - subtelomeric antisilencing regions (субтеломерные анти-сайленсинговые регионы)

TAS – telomeric associated sequences (последовательности, ассоциированные с теломерами)

TPE – telomeric position effect (позиционный эффект теломер)

TTI – targeted telomere insertion (таргетная вставка теломерных последовательностей)

QTP – quantitative telomeric chromatin isolation protocol (количественный метод выделения теломерного хроматина)

БСА - бычий сывороточный альбумин

ИПТГ - изопропил-β-D-1-тиогалактопиранозид

## Введение

Теломеры представлены нуклеопротеиновыми комплексами, сформированными концевыми участками линейных хромосом и взаимодействующими с ними белками, основной функцией которых является защита концов хромосом от систем рекомбинации и репарации двунитевых разрывов ДНК. Последовательность ДНК теломер у большинства видов эукариот представлена короткими гуанин-богатыми тандемными повторами, которые также обнаруживаются во внутренних частях хромосом и известны как интерстициальные теломерные последовательности (от англ. *interstitial telomeric sequences, ITSs*). Теломерные последовательности ассоциированы с областями с повышенной геномной нестабильностью, а также в силу высокого содержания гуанина и особенностей хроматиновой укладки являются сложными матрицами для работы систем репликации и репарации ДНК. Теломерные последовательности на концевых и внутренних участках хромосом вовлечены в взаимодействия с различными белками, участвующими в регуляции транскрипции проксимальных и дистальных генов и других внутриядерных процессах. Изучение теломер и теломерных последовательностей является непростой задачей для исследователей. В последнее время накапливается все больше данных о разнообразных белках, вовлеченных в процессы функционирования теломерных последовательностей в геноме, но многие участники и детали этих процессов до сих пор остаются мало изученными. Разработка методов для эффективного выделения и идентификации белков, вовлеченных в функционирование теломерных последовательностей, является одной из важных задач современной науки. Основой для таких методов могут служить относительно простые и удобные для генетических манипуляций дрожжевые системы. Подобные методы были разработаны и использованы в данной работе для выделения и идентификации белков, взаимодействующих с теломерными последовательностями в ДНК и РНК.

# 1. Обзор литературы

## 1.1. Структура и функции теломер

Хромосомы эукариот представлены линейными молекулами, концевые участки которых клетке необходимо уметь отличать от двунитевых разрывов ДНК, которые являются одними из самых серьезных повреждений в геноме. Двунитевые разрывы необходимо распознать и исправить при помощи систем репарации. В то же время, концевые участки хромосом, наоборот, необходимо защищать от репаративных клеточных систем, сохраняя таким образом целостность хромосом. В большой степени стабильное состояние концевых участков хромосом обеспечивается теломерами – особыми нуклеопротеиновыми комплексами, сформированными концевыми участками хромосом и взаимодействующими с этими участками белками и РНК (Blackburn, 1991). Структура теломер большинства видов (за исключением, например, *Drosophila melanogaster*, чьи теломерные последовательности представлены ретротранспозонами (Silva-Sousa et al., 2012)) в целом является консервативной. Двухцепочечная ДНК, входящая в состав теломер у большинства организмов представлена тандемными повторами консервативной последовательности. Для млекопитающих характерна консенсусная последовательность 5'-TTAGGG-3', а у дрожжей *Saccharomyces cerevisiae*, которые являются популярным модельным организмом для изучения биологии теломер, она представлена вырожденными повторами с консенсусом C1-3A/TG1-3 (Wellinger & Zakian., 2012). На конце теломер присутствует гуанин-богатый 3'-одноцепочечный выступ, или G-выступ, размер которого у человека составляет примерно 50-500 нуклеотидов (de Lange, 2009), а у дрожжей 3'-выступ значительно меняет свою длину на протяжении клеточного цикла, составляя примерно 12-14 нуклеотидов большую часть времени и увеличиваясь до 100 нуклеотидов в позднюю S-фазу (Largivee et al., 2004; Frank et al., 2006). G-выступ способен встраиваться в теломерную двухцепочечную ДНК и формировать своеобразную «кэпирующую» структуру теломер, известную как T-петля (Griffith et al., 1999). Вытесненная цепь ДНК и новообразовавшаяся при встраивании G-выступа двухцепочечная структура формируют D-петлю (Рис. 1). Так как теломерная ДНК содержит большое количество остатков гуанина, она способна формировать особые плоские молекулярные структуры с неканоническими хугстиновскими G-G взаимодействиями – известные как G-квадруплексы (Рис.1) (Paeschke et al., 2005; Tang et al., 2008). T-петля и G-квадруплексы

являются сложными и стабильными молекулярными образованиями, и их формирование поддерживается и регулируется белками, входящими в состав теломер (Lipps & Rhodes, 2009).

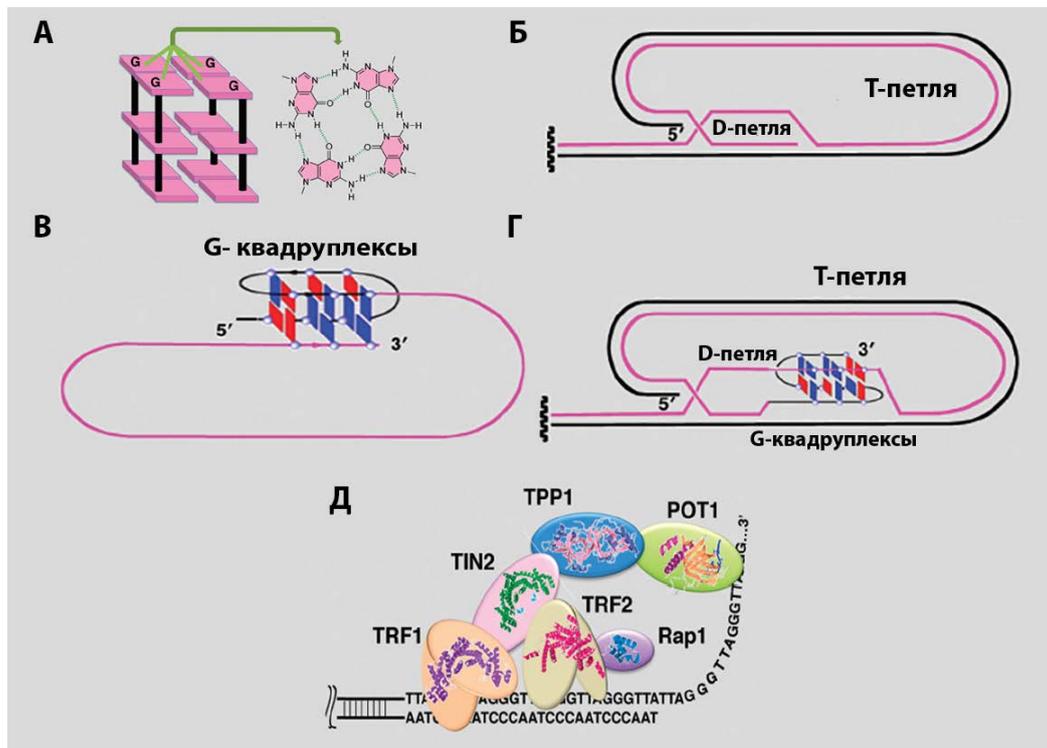


Рисунок 1. Структура концевой части хромосом: А) Схематичное изображение структуры G-квадруплексов. Б) Схематичное изображение Т-петли. В) Модель G-квадруплексов. Г) Схематичное изображение формирования G-квадруплекса при образовании Т-петли. Д) Схематичное строение связанного с теломерами комплекса шелтерина, состоящего из шести белков: TRF1, TRF2, TIN2, TPP1, POT1, RAP1 (из Xu, 2011 с изменениями).

У млекопитающих с теломерной ДНК взаимодействует комплекс шелтерин, состоящий из шести белков: TRF1, TRF2, TIN2, RAP1, POT1 и TPP1. Белки TRF1 и TRF2 содержат Myb-домен и в виде гомодимеров связывают повторы 5'-TTAGGG-3' в двуцепочечной теломерной ДНК, а с 3'-одноцепочечным выступом взаимодействуют гетеродимеры POT1-TPP1 (Рис. 2). Для TRF2 показана роль в формировании Т-петли (Griffith, 1999; Doksanı et al., 2013) и участие в сигнальных процессах системы репарации двунитевых разрывов (Karlseder et al, 1999; Bradshaw, 2005). POT1 также играет роль в формировании Т-петли и подавлении сигнальных процессов систем репарации в теломерах (Hockemeyer et al., 2006). TIN2 взаимодействует с TPP1, участвуя в загрузке димеров POT1-TPP1 на ДНК теломер (Hu et al., 2017) и с TRF1 и

TRF2 и служит, таким образом, связующим звеном в формировании полного шелтеринового комплекса (Ye et al., 2004). Шелтериновый комплекс участвует в регуляции структуры и длины теломер и защищает их от систем репарации (de Lange, 2005).

У дрожжей *Saccharomyces cerevisiae* шелтериновый комплекс отсутствует. С двуцепочечной ДНК, содержащей дрожжевые теломерные повторы, связывается комплекс Rap1-Rif1-Rif2, который участвует в регуляции длины и стабильности теломер (Conrad et al., 1990). Rap1 и белки Rif1 и Rif2 являются негативными регуляторами увеличения длины теломер (Kyrgion et al., 1992; Kaizer et al., 2013) и формируют регулярно расположенные на ДНК комплексы, количество которых коррелирует с длиной теломерного региона (Levy & Blackburn, 2004). Также Rap1 привлекает к теломерам белки Sir2, Sir3, Sir4 и Ku (Yku70–Yku80). Нарушение функционирования Sir3 и Sir4 приводит к укорочению теломер (Palladino et al., 1993). Белки Sir также вовлечены в регуляцию правильной локализации теломер в клеточном ядре (Taddei et al., 2004). Белки Ku вовлечены в процессы репарации посредством негомологичного соединения концов, а в области теломер играют роль в регуляции формирования структуры хроматина (Dewar & Lydall, 2012) и способны взаимодействовать с теломеразной РНК (Chan et al., 2008). Гуанин-богатый 3'-выступ у дрожжей связывается белком Cdc13 (cell division control protein 13) (Nugent et al., 1996), который наряду с белками Stn1 и Ten1, входит в состав комплекса CST. Несмотря на то, что между компонентами шелтерина и CST комплекса не прослеживается гомологии, в клетках *S. cerevisiae* CST является функциональным аналогом шелтерина млекопитающих: выполняет протекторную функцию, а также участвует в регуляции репликации теломер (Bianchi & Shore, 2008).

У млекопитающих, в частности, у человека, CST комплекс также был обнаружен (Martin et al., 2007) – в его состав входят ортологичные дрожжевым белки TEN1 и STN1, а также белок CTC1, не являющийся ортологичным дрожжевому Cdc13 и значительно отличающийся от него по свойствам и функциям. CTC1 неспецифически связывает одноцепочечную ДНК, однако аффинность этого связывания возрастает по мере увеличения размера ДНК (Price et al., 2010), в то время как для Cdc13 давно известна специфичность к 11-буквенному консенсусу dGTGTGGGTGTG (Mitton-Fry et al., 2002). Большая часть комплексов CST в ядрах клеток млекопитающих не ассоциирована с теломерами (Hom & Wuttke, 2017), и предполагается, что этот

комплекс у млекопитающих вовлечен во многие другие, не связанные с теломерами, процессы (Miyake et al., 2009). Это было показано в некоторых работах, где изучалась роль комплекса CST в разрешении проблем репликации (Stewart et al., 2018), таких как задержка репликационной «вилки» как в области теломер, так и вне их (Wang et al., 2012; Stewart et al., 2012). CST млекопитающих привлекается взаимодействующим с STN1 белком-компонентом шелтерина TPP1 в область теломер, где он действует в роли позитивного регулятора праймной активности ДНК-полимеразы альфа (Price et al., 2010). Комплекс CST поддерживает и стабилизирует одноцепочечный выступ и облегчает процессы репликации (Rice & Skordalakes, 2016). Комплекс CST очень сильно похож на RPA, другой гетеротримерный комплекс, который связывает одноцепочечную ДНК при репликации и репарации двунитевых разрывов, а также вовлечен в сигнальные процессы систем репарации и регуляцию клеточного цикла (Zou et al., 2006). Было показано, что Cdc13 совместно с компонентами комплекса RPA участвует в защите теломерной ДНК от действия экзонуклеаз (Greetham et al., 2015).

Также, как и компоненты комплекса RPA компоненты комплекса CST млекопитающих содержат домены с ОБ-укладкой (от англ. OB-fold domain, oligonucleotide/oligosaccharide-binding fold domain, домены взаимодействия с олигонуклеотидами\олигосахаридами) (Bryan et al., 2013; Lei et al., 2004; Zhong et al., 2012), которые являются весьма разнообразными по последовательности, однако сохраняют структуру, состоящую из бета-бочонка и нескольких экспонированных альфа-спиралей. Структурная организация OB-fold домена весьма пластична и подвержена эволюционной дивергенции. В дрожжевом комплексе CST Cdc13 и Stn1 также содержат OB-fold домены (Gao et al., 2007). Причем, в Cdc13 было обнаружено несколько OB-fold доменов, которые отвечают за димеризацию и взаимодействия с ДНК и белками (Cohn, 2013). Этот домен обнаруживается в широком спектре белков, связанных с системой репарации ДНК, и способен связывать одноцепочечную ДНК, РНК и олигопептиды (Flynn & Zou, 2011).

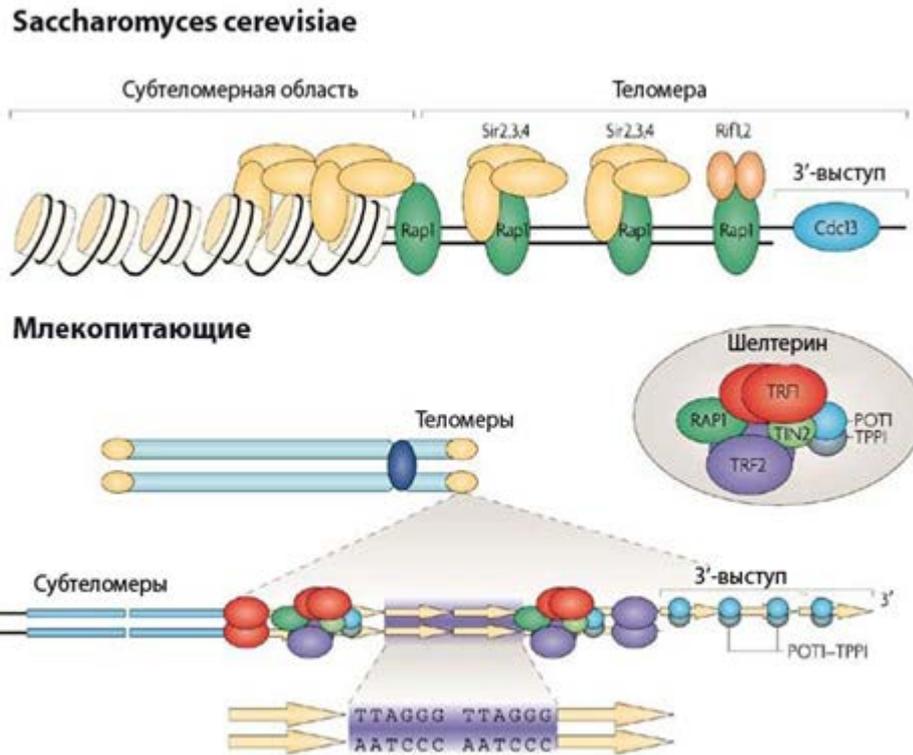


Рисунок 2. Структура теломер у дрожжей *S.cerevisiae* и млекопитающих (из Blasco, 2007 с изменениями).

Субтеломерные области дрожжей, как и у многих других организмов, также представлены повторяющимися последовательностями. Эти повторы являются вырожденными и были названы TAS – последовательностями, ассоциированными с теломерами (telomere associated sequences). У *S. cerevisiae* известны два класса TAS-элементов: Y' и X. Y' следует сразу за последовательностью теломер в количестве 0-4 тандемных копий. Y' бывает двух размеров: длинный (6.7 kb) и короткий (5.2 kb). X-элемент обнаруживается на всех дрожжевых хромосомах, более гетерогенен по размеру и находится проксимальнее к центромерам от Y', если он присутствует (Wellinger et al., 2012). В субтеломерных областях генома *S. cerevisiae* обнаруживаются также последовательности, соответствующие человеческим теломерным повторам - 5'-TTAGGG-3'. Обнаружены они были благодаря открытию жизненно важного дрожжевого белка Tbf1, который с высокой аффинностью связывает последовательность 5'-TTAGGG-3' (Brigati et al., 1993). Эти последовательности известны как STARs (от англ. subtelomeric antisilencing regions), субтеломерные анти-сайленсинговые регионы. В то же время в субтеломерных областях располагаются последовательности, которые приводят к сайленсингу транскрипции близких к ним

регионов. Взаимодействие антагонистических эффектов сайленсеров и анти-сайленсеров в этих областях придает своеобразный квазинеустойчивый паттерн экспрессии генов, расположенных в субтеломерах (Power et al., 2011). Было показано, что STARs способны снижать эффект сайленсинга в теломерах, лишенных X и Y'-элементов. Когда субтеломерные последовательности X и Y' присутствуют, их сайленсинговый эффект оказывается доминирующим над антагонистическим эффектом STARs (Fourel et al., 1999). Tbf1 содержит Myb-домен и имеет высокую степень гомологии с белками TRF1 и TRF2 млекопитающих (Bilaud et al., 1996). В экспериментах с последовательностями STARs было показано, что Tbf1 способен выполнять протекторную функцию в теломерах (Ribaud et al., 2012). Считается, что Tbf1 эволюционно происходит от предкового теломер-связывающего белка, который раньше играл главную структурную роль в формировании теломер. В таком случае, последовательности STARs, с которыми Tbf1 специфически взаимодействует, можно рассматривать как следы перехода от предковой Tbf1-регуляторной системы теломер к Rap1-системе. Этот переход, во всей видимости, был связан с мутацией в гене теломеразной РНК *TLC1*, которая привела к способности TLC1 генерировать повторы TG1-3 и с появлением белка Rap1 с дублированным ДНК-связывающим доменом Myb/SANT, который способен к более эффективному связыванию с последовательностями именно этих повторов (Lue, 2010). Субтеломерные регионы подвержены рекомбинационным перестройкам и содержат потенциальные сайты ориджинов репликации. В результате рекомбинационных процессов гомологичных субтеломерных областей разных хромосом может происходить увеличение размера субтеломер, и они оказываются способными к поддержанию концевых участков хромосом в функциональном состоянии в отсутствие теломеразы (Lundblad & Blackburn, 1993).

## **1.2. Репликация теломерных областей**

Большая часть теломерной ДНК (за исключением самой терминальной области) реплицируется в результате обычной полуконсервативной репликации ДНК. Ориджины репликации в области теломер являются поздно активируемыми, механизм задержки зависит от белка Rif1. Репликация теломерной ДНК представляет из себя непростую задачу для репликативного аппарата в силу высокого содержания гуанина (и, как следствие, формирование уже упомянутых G-квадруплексов) и “кэпирующих”

петель, поэтому для успешной репликации необходимо привлечение специализированных геликаз (Pif1 и Rrm3 у *S. cerevisiae* (Paeschke et al. 2011, Makovets et al., 2004)) и регуляторных белков, взаимодействующих с теломерами. Компоненты CST комплекса Stn1 и Cdc13 взаимодействуют с субъединицами репликативной ДНК-полимеразы альфа, а Cdc13 также способен связывать Est1, регуляторную субъединицу теломеразы, что позволяет обеспечить взаимодействие между ДНК-полимеразой и теломеразой (Grandin et al., 2000). Более того, компоненты комплекса CST способны к расплетению G-квадруплексов, устраняя таким образом физические препятствия на пути полимеразы (Bhattacharjee et al., 2017). С-цепь теломерной ДНК является матрицей для синтеза лидирующей цепи, а G-цепь - матрицей для синтеза отстающей цепи с использованием фрагментов Оказаки. Репликация С-цепи проходит до конца, приводя к образованию “тупых” концов, однако практически сразу происходит деградация С-цепи и формирование G-выступа на 3'-конце (Рис. 3), необходимого субстрата для работы теломеразы, связывания теломерных белков и формирования Т-петли. Укорочение матричной С-цепи теломер приводит к укорочению теломер при каждом последующем репликативном цикле. В процесс деградации матричной лидирующей С-цепи вовлечены нуклеазы, ДНК-геликазы и другие белки (Mimitou & Symington, 2009). Комплекс Mre11-Rad50-Xrs2 (MRX) играет важную роль в процессинге 5'-конца, так как при его отсутствии G-выступы у теломер *S. cerevisiae* становятся короче и не способны удлиняться в S-фазу, как это происходит в норме (Larrivee et al., 2004). При этом, у комплекса Mre11 отсутствует 5'-3'-экзонуклеазная активность, которая требуется для создания теломерного 3'-выступа. В последнее время считается, что роль комплекса MRX заключается в привлечении необходимых экзонуклеаз к концевым участкам хромосом. Например, была показана роль нуклеазы Sae2 и геликазы Sgs1, в отсутствие которых теломеры укорачиваются, но не нарушается работа теломеразы (!Bonetti et al., 2009). Обнаружилось, что укорочение теломер в клетках *sae2Δ / sgs1Δ* частично подавляется сверхэкспрессией экзонуклеазы Exo1 (Bonetti et al., 2009) (Рис. 3).

При репликации G-цепи на самом терминальном участке хромосомы остается недореплицированный участок вследствие того, что праймаза не способна синтезировать праймер на этом участке хромосомы. Таким образом на этой цепи за счет неполной репликации уже формируется небольшой G-выступ, однако затем происходит дополнительная деградация С-цепи (Pfeiffer & Lingner, 2013). Теломераза

*S. cerevisiae* предпочтительно удлиняет около 7% самых коротких теломер в поздней S-фазе (Teixeira et al., 2004). В процессы определения самых коротких теломер вовлечены белки Tel1 (ортолог белка ATM у млекопитающих), MRX-комплекс (Sabourin et al., 2007) и Tbf1 (Arneric & Lingner, 2007). На культурах клеток HeLa и легочной аденокарциномы H1299 показано, что степень восстановления теломерных областей теломеразой составляет около 70-100% и наблюдается на обеих цепях теломерной ДНК (Zhao et al., 2009). Несмотря на то, что отстающей G-цепи не должно быть свойственно укорачиваться (Рис. 3), она также подвергается частичной деградации, а затем удлиняется теломеразой (Pfeiffer & Lingner, 2013). Теломераза представляет собой рибонуклеиновый комплекс, в состав которого входит обратная транскриптаза TERT (Est2 у *S. cerevisiae*), и некодирующая РНК TERC (TLC1 у *S. cerevisiae*), выполняющая роль матрицы для образования теломерных повторов.

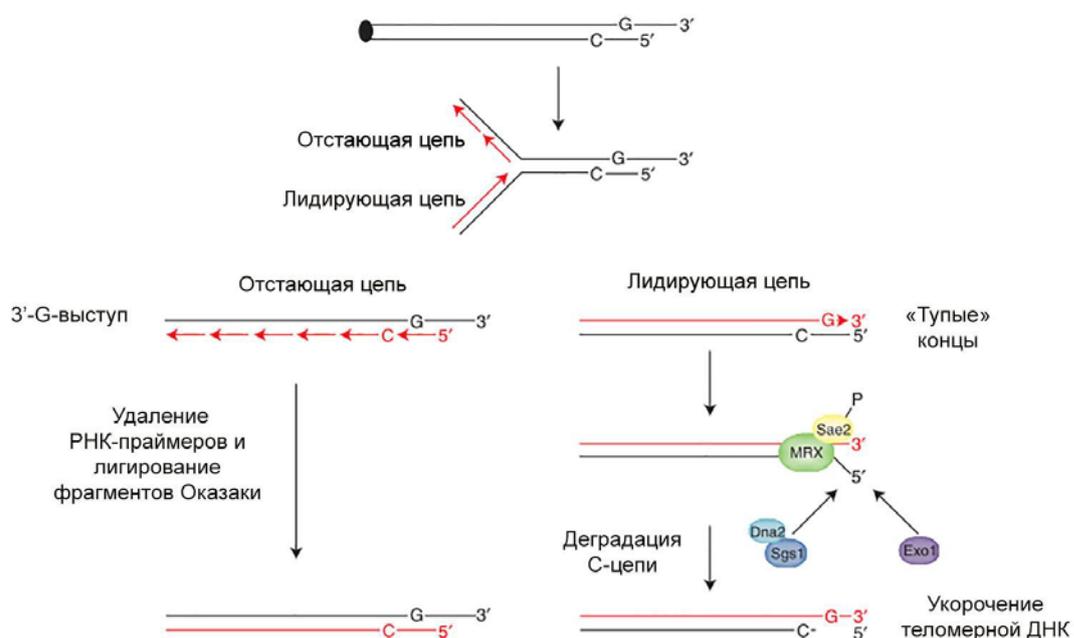


Рисунок 3. Схема репликации теломер (из Pfeiffer & Lingner, 2013 с изменениями). G-богатая цепь служит матрицей для синтеза отстающей цепи, C-богатая цепь становится матрицей для синтеза лидирующей цепи. Предполагаемый интермедиат с "тупыми" концами при процессинге матричной C-цепи и новосинтезируемой цепи подвергается деградации C-цепи при участии комплекса MRX и экзонуклеаз Sae2, Sgs1 и Exo1.

### 1.3. Структура хроматина теломерных областей

Изучение теломерных последовательностей *in vitro* и *in vivo* показало, что они имеют наименьшую афинность к связыванию нуклеосом (Ichikawa et al., 2015). Однако, у высших эукариот теломерные последовательности всё же упакованы в нуклеосомы, более того, размер спейсера между ними оказывается уменьшенным по сравнению с канонической гетерохроматиновой укладкой, но при этом нуклеосомы теломер оказываются более чувствительными к действию нуклеаз по сравнению с обычным гетерохроматином (Ellahi et al., 2015). Для теломерного хроматина характерна повышенная степень компактности, так как практически десятикратная компактизация и формирование защитных надмолекулярных структур является одной из стратегий защиты теломер компонентами шелтеринового комплекса (Sfeir & de Lange, 2014; Bandaria et al., 2016). Существуют разногласия в данных о модификациях хроматина теломер: с одной стороны, есть свидетельства о наличии модификаций, характерных для гетерохроматина. В теломерах мышей было обнаружено большое количество гетерохроматиновых модификаций гистонов H3K9me3 и H4K20me3, и был снижен статус ацетилирования гистонов H3 и H4 (Blasco, 2007; Garcia-Cao et al., 2003). При этом для хроматина человеческих теломер в недавних исследованиях было отмечено присутствие маркеров эухроматина (H2BK5me1 и H3K4me3) (Rosenfeld et al., 2008; O'Sullivan et al., 2010; Cubiles et al., 2018). У дрожжей *S. cerevisiae* в теломерах отсутствует нуклеосомная укладка хроматина, а в субтеломерных областях структура гетерохроматина поддерживается за счет взаимодействия белков Sir4 и Rap1, Sir4 привлекает белки Sir2 и Sir3, последний из которых является NAD-зависимой гистоновой деацетилазой. За счет последовательного деацетилирования гистонов H3 и H4 происходит распространение гетерохроматиновой укладки и так называемого теломерного позиционного эффекта (от англ. Telomere position effect, TPE) (Saveliev et al., 2003; Pisano et al., 2008). Давно обнаруженный у дрожжей TPE обеспечивается укладкой хроматина, зависит от белка Sir3 и приводит к сайленсингу генов, расположенных на близком расстоянии от теломерных областей (Gottschling et al., 1990; Ellahi et al., 2015). На другом виде дрожжей, *Schizosaccharomyces pombe*, было показано, что структура хроматина теломер остается стабильной и наследуется даже при полном отсутствии теломерных повторов (Sadaie et al., 2003), что свидетельствует о роли субтеломерных последовательностей и белковых факторов в формировании и поддержании структуры хроматина теломер.

#### **1.4. Транскрипция теломерных областей и РНК TERRA**

Несмотря на конденсированную структуру хроматина теломер и присутствие в области теломер маркеров гетерохроматина, с субтеломерных областей идет транскрипция, в результате которой образуется особый класс некодирующих РНК TERRA (РНК, содержащая теломерные повторы, от англ. telomeric repeat-containing RNA) (Wellinger & Zakian, 2012). TERRA обнаружены у широкого числа представителей эукариот: дрожжей, растений, позвоночных, включая млекопитающих, что подтверждает наличие у них консервативных функций. В состав обнаруженных РНК TERRA дрожжей *S. cerevisiae* входят последовательности, комплементарные как X-, так и Y'-элементам и теломерным повторам. В транскрипции TERRA у млекопитающих участвует ДНК-полимераза II, у дрожжей - ДНК-полимераза альфа. Полученные транскрипты размером от 100 до 1200 нуклеотидов (до 9000 нуклеотидов у млекопитающих (Azzalin et al., 2007)) подвергаются полиаденилированию и 5'-кэпированию, что увеличивает стабильность транскриптов (Poggio et al., 2010). У млекопитающих в теломерных областях обнаружены сайты начала транскрипции TERRA, в регуляции которых участвуют CpG-островки, факторы реорганизации хроматина, CTCF и когезины (Deng et al., 2012). TERRA способна к образованию G-квадруплексов: как самостоятельно, так и при формировании ДНК-РНКовых гибридов (Balk et al., 2013; Xu et al., 2010). Среди человеческих белков, с которыми взаимодействует TERRA, были обнаружены следующие: фермент процессинга мРНК SRRT/ARS, компонент гистоновой ацетилтрансферазы NuA4 MORF4L2 и фактор ремоделинга нуклеосом ARID1A, которые влияют на количество TERRA в клетке (Scheibe et al., 2013). Также обнаружено, что онкоген TLS/FUS взаимодействует с G-квадруплексами в РНК TERRA и теломерной ДНК и вовлечен в регуляцию длины теломер (Takahama et al., 2013; Kondo et al., 2018). К тому же, уровень образования TERRA в клетках зависит от стадии клеточного цикла, состояния хроматина и длины теломер (Maicher et al., 2014). TERRA способствует аккумуляции теломеразы в области коротких теломер (Cusanelli et al., 2013) и вовлечена в удлинение длины теломер, не зависящее от активности теломеразы (Maicher et al., 2012).

#### **1.5. Теломерные последовательности вне теломер и их функции**

Теломерные последовательности присутствуют не только на концах хромосом, они также обнаруживаются и во внутренних областях генома и носят название ITSs

(от англ. *interstitial telomeric sequences*) – внутренние (интерстициальные) теломерные последовательности. ITSs условно делят на гетерохроматиновые, которые могут достигать до нескольких сотен тысяч п.н., обычно локализованные в области центромер, короткие ITSs, представленные отдельными теломерными гексамерами, разбросанными по геному (Ruiz-Hererra et al., 2008) и соединенные ITSs (от англ. *fusion ITSs*), образовавшиеся в ходе хромосомных перестроек (Bolzan, 2017). Первыми были обнаружены гетерохроматиновые ITSs в геноме китайского хомячка, они составляют около 5% генома и формируют протяженные блоки с гетерохроматиновой укладкой (Faravelli et al., 2002). Затем гетерохроматиновые ITSs были обнаружены у некоторых приматов, грызунов, сумчатых, рептилий, амфибий, рыб и птиц. Эти области ассоциированы с повышенной хрупкостью (ломкостью) хромосом и большим количеством рекомбинационных перестроек (Sanchez et al., 2010). Протяженные гетерохроматиновые ITSs считаются остатками больших геномных перестроек в эволюции кариотипов. Например, так называемые Робертсоновские перестройки, когда происходит транслокация целого плеча и остаются функциональными области теломер, могут играть роль в видообразовании. Поэтому ITSs активно используют в изучении эволюции отдельных групп живых организмов (Rovatsos et al., 2015; Chirino et al., 2017, Mazzoleni et al., 2017). У человека не были обнаружены протяженные гетерохроматиновые ITSs, но присутствуют соединенные ITSs, например, находящаяся в 2q13 второй хромосомы человека последовательность, которая осталась после слияния предковых хромосом (Ijdo et al., 1991). Происхождение всех типов ITSs до сих пор остается проблемой, требующей дальнейшего изучения. Одна из гипотез предполагает, что короткие теломерные повторы могли быть вставлены системами репарации двунитевых разрывов, в которые вовлечена теломераза (Nergadze et al., 2007). Другие предложенные механизмы формирования коротких ITSs – это индуцированная двунитевыми разрывами репликация (от англ. *breakage-induced replication*, BIR) и таргетное встраивание теломерных последовательностей (от англ. *targeted telomeric insertion*, TTI) (Marzec et al., 2015).

На дрожжевых клетках показано, что системы пострепликативной репарации и гомологичной рекомбинации вовлечены в экспансию, то есть в увеличение в количестве единиц тандемных повторов (Aksenova et al., 2015). Взаимодействие этих систем между собой и с системой репликации приводит к полиморфизму длины

повторов (Aksenova et al., 2015). На ДНКовых матрицах, богатых повторами часто происходит задержка, или отставание, репликационной «вилки», а также вставка дополнительных копий повторов, что может быть потенциальным механизмом полиморфизма длины повторяющихся последовательностей в ITSs (Mondello et al, 2000). Таргетное встраивание теломерных последовательностей, или ТТИ, - это вставка теломерной ДНК в различные внутренние сайты хромосом, которая часто наблюдается в клетках, преодолевающих предел Хэйфлика (ограничение числа делений вследствие укорочения теломер) не за счет активации теломеразы, а за счет механизма альтернативного удлинения теломер (от англ. ALT (alternative lengthening of telomeres), в который вовлечены некоторые компоненты гомологичной рекомбинации (Muntoni & Reddel, 2005). Механизм ALT обнаруживается в клетках, мутантных по гену *EST1* (регуляторная субъединица теломеразы) у дрожжей, в клетках млекопитающих с недостаточной активностью теломеразы (Niida et al., 2000; Muntoni & Reddel, 2005) и в клетках некоторых типов опухолей. ALT активируется примерно в 15% раковых опухолей, и большинство из них являются агрессивными опухолями с высокой смертностью (Dilley & Greenberg, 2015). Для них, помимо гетерогенных по длине теломер и увеличенного уровня рекомбинационных процессов в теломерах, также характерно наличие кольцевых экстрахромосомных молекул ДНК, содержащих теломерные повторы и образование телец APB (ALT-associated PML (promyelocytic leukemia) bodies). APB-тельца состоят из теломерной ДНК, белков шелтеринового комплекса, факторов репарации ДНК и ремоделинга хроматина. Механизм возникновения ALT связан не только с белками теломер, но и с мутациями в генах, кодирующих белки-ремоделлеры хроматина, например, ATRX и DAXX, играющих важную роль в организации гетерохроматиновой структуры теломер (O'Sullivan & Almouzni, 2014; Episkopou et al., 2014; Lewis et al., 2010). В клетках с активированным ALT происходит таргетное встраивание теломерной ДНК, регулируемое компонентами ядерного рецептора NR2C/F (TR2, TR4, COUP-TF1, COUP-TF2, EAR2) (Marzec et al., 2015; Aeby & Lingner, 2015), что потенциально может являться важным механизмом появления коротких ITSs. Изучение экспансии повторяющихся последовательностей может помочь выяснить механизмы различных заболеваний, сопряженных с экспансиями и других, не теломерных, повторов, таких как синдром ломкой X-хромосомы, атаксия Фридерикса и многих других (Mirkin, 2007; Shishkin et al., 2009).

Показано, что компоненты шелтерина TRF2 и RAP1 способны взаимодействовать и с ITSs у млекопитающих (Simonet et al. 2011; Yang et al., 2011; Martinez et al., 2010; Garrobo et al, 2014). Недавние исследования показали, что TRF2 участвует в позитивной регуляции транскрипции компонента комплекса polycomb белка PCGF3, взаимодействуя с промоторной областью, которая формирует G-квадруплекс (Purohit et al., 2018). Более того, TRF2 способствует формированию внутренних петель (interstitial telomere loop, ITL), аналогичных T-петле в теломерах, и связывается с ламинами A/C (Wood et al., 2014), которые образуют ламину, внутреннюю выстилку клеточного ядра. Синдром Хатчинсона-Гилфорда, заболевание, известное также как прогерия, или синдром раннего старения, который вызывается доминантными мутациями в генах ламин этого типа, сопряжен с уменьшенным количеством ITL и дисфункцией теломер (Wood et al., 2015). Связывание TRF2 с компонентами ламин на внутренней стороне ядерной оболочки объясняет, например, наблюдение непосредственной близости теломерных последовательностей к внутренней ядерной мембране при мейозе (Scherthan et al., 2000; Ilicheva et al., 2015).

Петли, образованные теломерными последовательностями и ITSs составляют важную компоненту трёхмерной укладки хроматина в ядре, которая в свою очередь является важным аспектом регуляции функционирования всех процессов в геноме (Gonzalez-Suarez & Gonzalo, 2009). Таким образом, ITL могут опосредовать теломерную регуляцию дистанцированных от теломер областей генома. Сейчас уже известно, что обнаруженный в дрожжах позиционный эффект теломер, действие которого способно распространяться на несколько тысяч нуклеотидов вглубь генома, также присутствует и у млекопитающих, более того, протяженность этого эффекта может достигать почти сотен тысяч нуклеотидов. Например, при мышечной дистрофии Ландузи-Дежерина (Лице-лопаточно-плечевой мышечной дистрофии, ЛЛПМД) транскрипция генов DUX4 и SORBS2, ассоциированных с патогенезом заболевания, напрямую регулируется длиной теломер (Stadler et al., 2013) посредством механизма, названного авторами TPE-OLD (позиционный эффект теломер на больших расстояниях) (Robin et al., 2015). В то же время есть гены, находящиеся в еще более значительном отдалении от концевых участков хромосом, но транскрипция которых тем не менее регулируется теломерами (Lou et al., 2009) и, так как гены, находящиеся проксимальнее к теломерам, не зависят от теломерной регуляции, считается, что она может осуществляться при формировании внутренних теломерных петель (Shay,

2018). Эти механизмы также вовлечены в регуляцию транскрипции компонента теломеразы TERT (Kim et al., 2016).

### 1.6. Подходы к изучению теломерных последовательностей

Изучение молекулярных механизмов функционирования теломерных последовательностей в теломерных и во внутрихромосомных сайтах является непростой задачей. Удобными и потому популярными моделями для изучения подобных процессов, как уже было отмечено выше, являются дрожжевые модели, в которых проще проводить генетические манипуляции и анализировать наблюдаемые эффекты. Дрожжевая система для изучения функционирования теломерных последовательностей была сконструирована ранее в лаборатории профессора С.М. Миркина (Университет Тафтса, США): она представляет собой штаммы дрожжей *S. cerevisiae*, в одном из которых есть вставка дрожжевой теломерной последовательности в искусственный интрон репортерного гена, нарушающая его экспрессию (Aksenova et al., 2013) (Рис.4).

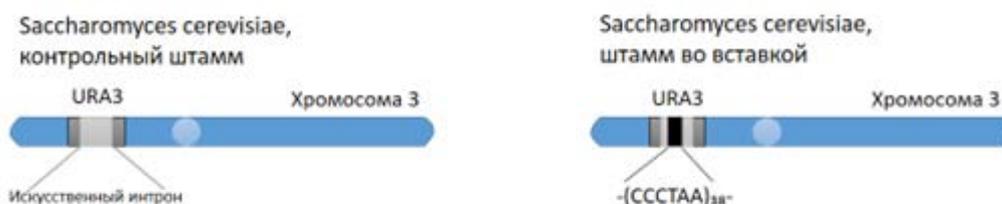


Рисунок 4. Дрожжевая система для изучения влияния человеческих теломерных последовательностей на функционирование генома. Исходная система (Aksenova et al., 2013) содержала не человеческие, а дрожжевые теломерные повторы.

В предыдущих исследованиях на этой модели было показано, что вставка дрожжевых теломерных последовательностей в кодирующую часть репортерного гена в G-ориентации  $-(TG\overline{TGTGGG})_n-$  приводит к увеличению мутагенеза в областях, фланкирующих повторы, микросателлитной нестабильности в дрожжевой системе и крупным хромосомным перестройкам (Aksenova et al., 2013), а также, при вставке дрожжевых теломерных повторов в C-ориентации  $-(CCCACACA)_n-$ , повторы подвержены экспансиям, то есть увеличению в количестве единиц тандемных повторов (Aksenova et al., 2015). Как уже было освещено выше, в этих работах было показано, что процессы экспансии теломерных повторов связаны с работой систем рекомбинационной репарации и пострепликативной репарации, которые также

вовлечены в процесс альтернативного удлинения теломер у клеток с активированным ALT. Использование этой же модели для изучения поведения человеческих теломерных последовательностей, которые, как уже было описано выше, также присутствуют в геноме *S. cerevisiae* и известны как субтеломерные анти-сайленсинговые последовательности (STARs), также показало их нестабильность и склонность к экспансиям. Для того, чтобы выяснить механизмы, лежащие в основе нестабильности теломерных последовательностей, в данной работе была проведена разработка и оптимизация методов выделения и идентификации белков, взаимодействующих с человеческими теломерными последовательностями в ДНК и РНК.

Обнаружение новых белков, вовлеченных в процессы функционирования теломерных последовательностей в геноме является эффективным подходом для выявления механизмов функционирования теломер и теломерных последовательностей.

## **2. Цель работы**

Разработка систем для выделения и идентификации белков, взаимодействующих с теломерными последовательностями в РНК и ДНК.

## **3. Задачи работы**

1. Разработка и оптимизация метода для выделения и идентификации белков, взаимодействующих с РНК, содержащей человеческие теломерные повторы в дрожжевой системе;
2. Выделение белков, взаимодействующих с РНК, содержащей теломерные последовательности.
3. Разработка и оптимизация метода для выделения и идентификации белков, взаимодействующих с теломерными повторами в ДНК в дрожжевой системе.
4. Идентификация выделенных белков методом масс-спектрометрии.

## 4. Материалы и методы

### 4.1. Материалы

#### 4.1.1 Штаммы, использованные в работе

В работе использовали штаммы дрожжей *Saccharomyces cerevisiae*, ведущие свое происхождение от штамма S288C. Генотипы штаммов, использованных в работе, а также их названия и источник происхождения указаны в Таблице 1.

Таблица 1. Использованные в работе штаммы *S. cerevisiae* и их генотипы.

Название штамма	Генотип штамма	Происхождение штамма
SMY710	<i>MATa leu2-1 trp1-63 ura3-52 his3-200 ade2Δ::KanMX4</i>	Aksenova et al., 2013
SMY758	<i>MATa leu2-1 trp1-63 ura3-52 his3-200 ade2Δ::KanMX4 III(75423-75715)::URA3-Int-(CCCTAA)38</i>	Предоставлен научным руководителем Аксеновой А.Ю.
SMY803	<i>MATa leu2-1 trp1-63 ura3-52 his3-200 ade2Δ::KanMX4 III(75423-75715)::URA3-Int</i>	Aksenova et al., 2013
ААУ11	<i>MATa leu2-1 trp1-63 ura3-52 his3-200 ade2Δ::KanMX4 III(75423-75715)::URA3-Int-lacO</i>	Предоставлен научным руководителем Аксеновой А. Ю.
ААУ51	<i>MATa leu2-1 trp1-63 ura3-52 his3-200 ade2Δ::KanMX4 III(75423-75715)::URA3-Int-(CCCTAA)36-lacO</i>	Сконструирован в ходе работы
ААУ52	<i>MATa leu2-1 trp1-63 ura3-52 his3-200 ade2Δ::KanMX4 III(75423-75715)::URA3-Int-(TTAGGG)36-lacO</i>	Сконструирован в ходе работы

#### 4.1.2. Среды, использованные в работе

В данной работе были использованы следующие среды: YPD жидкая (2% пептон, 1% дрожжевой экстракт, 2% глюкоза), YPD Агар (2% пептон, 1% дрожжевой экстракт, 2% глюкоза, 2% агар), SC -Тр, LB (1% триптон, 0,5% дрожжевой экстракт, 1% NaCl), LB Агар (1% триптон, 0,5% дрожжевой экстракт, 1% NaCl, 1,5% Агар), LB Агар с ампициллином (2,5 мкг/мл).

#### 4.1.3. Плазмиды, использованные в работе

В работе использовались плазмиды: UIRL (*pISL-UR-Intron-A3-TRP1-ISR*) (Aksenova et al, 2013), UIRL-lacO-25-12-8 (плазмида, сконструированная на основе UIRL, содержащая *lacO* и -(CCCTAA)<sub>36</sub>- в интроне *URA3*), UIRL-lacO-25-12-5 (плазмида, сконструированная на основе UIRL, содержащая *lacO* и -(TTAGGG)<sub>36</sub>- в интроне *URA3*), UIRL-25-12-8 (плазмида, сконструированная на основе UIRL, содержащая -(CCCTAA)<sub>36</sub>- в интроне *URA3*), UIRL-lacO9 (плазмида, сконструированная на основе UIRL, содержащая *lacO* в интроне *URA3*) (Рис. 5.). Все плазмиды были сконструированы научным руководителем Аксеновой А. Ю.

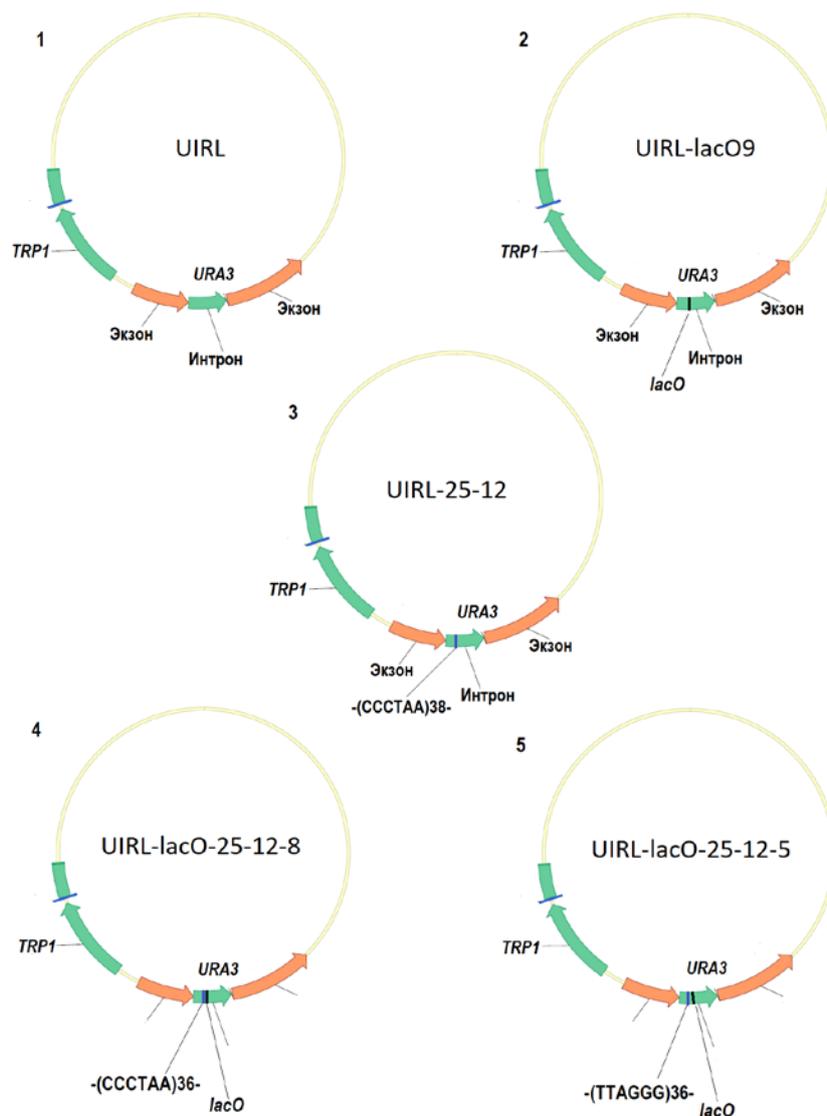


Рисунок 5. Карты плазмид, использованных в работе. 1) UURL (*pISL-UR-Intron-A3-TRP1-ISR*) (Aksenova et al, 2013), 2) UURL-lacO9 (плазмида, сконструированная на основе UURL, содержащая *lacO* в интроне *URA3*), 3) UURL-25-12 (плазмида, сконструированная на основе UURL, содержащая *-(CCCTAA)38-* в интроне *URA3*) 4) UURL-lacO-25-12-8 (плазмида, сконструированная на основе UURL, содержащая *lacO* и *-(CCCTAA)36-* в интроне *URA3*), 5) UURL-lacO-25-12-5 (плазмида, сконструированная на основе UURL, содержащая *lacO* и *-(TTAGGG)36-* в интроне *URA3*).

#### 4.1.4. Олигонуклеотиды, использованные в работе

В таблице 2 приведены использованные в работе олигонуклеотиды, места их “отжига” на кассете *URA3* приведены на Рис. 6.

Таблица 2. Олигонуклеотиды.

Олигонуклеотид	Последовательность	Использовали для:
RIP1	TCGAGCAATTGGGACCGTGCAAT TCTTCTTACAGTTAAATGGG- биотин	выделения белков, взаимодействующих с РНК, содержащей теломерные повторы в гене <i>URA3-Int</i>
RIP1_URA35	GGTTCSTTTGTTACTTCTTCTGCC GCCTGCTTCAAACCGC-биотин	
UURL2	TACAGATCAGTCAATATAGGAGG TT	совместно с UURL1 для проверки длины повтора в вставке в штаммах SMY758 и SMY803
A36a-R	AGGGTCGTTGCCTTCTGGTGTT	совместно с UURL1 для проверки вставки в штаммах AAY51 и AAY52
UURL1	GAAGTAACAAAGGAACCTAGAG GGTA	получения ПЦР-продуктов для исследования взаимодействия LacI- <i>lacO</i> ; проверки длины повтора в вставке при трансформации и контроля присутствия последовательности <i>lacO</i>
lacOR	ССААТААТАТCGTGGTTАТТАСА GATCAGT	
URA3INTF	TTTTCACTCTCCCАТАACCTCСТА ТА	
URA3INTR	GCCCСТАТТТАТТССААТААТАТ GT	

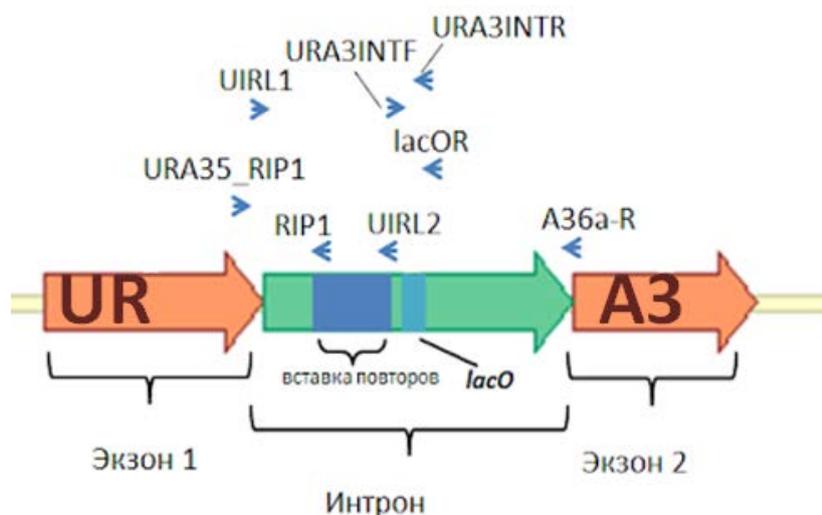


Рисунок 6. Карта “отжига” олигонуклеотидов на фрагменте кассеты *URA3-Int*. Кассета (размером ~3 Kb) предназначена для в третью хромосому около ориджина репликации *ARS306*. Содержит репортерный ген *URA3* (обозначен оранжевым цветом), искусственный интрон из гена *ACT1* (обозначен зеленым цветом), ген *TRP1* (на рисунке не отмечен) и фланкирующие последовательности третьей хромосомы (на рисунке не отмечены). В интрон вставлены человеческих теломерных повторов в разной ориентации (обозначены темно-синим цветом) и последовательность *lacO* (обозначена голубым цветом).

#### 4.2. Выделение белков, взаимодействующих с РНК, содержащей человеческие теломерные повторы в дрожжевой системе.

Для выделения белков, взаимодействующих с РНК, содержащей человеческие теломерные повторы использовали штаммы *S. cerevisiae* SMY758 и SMY803, которые выращивали в YPD до достижения оптической плотности ~0,6 (длина волны = 600 нм), затем лизировали в буфере РЕВ (20 мМ Трис-НСl (рН = 7.6), 140 мМ КСl, 5 мМ MgCl<sub>2</sub>, 0,5 мМ дитиотрейтол, 0,5 мМ фенилметилсульфоксид, 1% Тритон X-100, 5 мкМ леупептин, 5 мкМ, пепстатин А, 1 мМ бензамидин) при помощи шариков Silica Lysing Beads Lysing Matrix В (MP Biomedicals) и прибора FastPrep-24 Classic (MP Biomedicals) (30 сек 6.5М/сек, 1,5 мин при + 4°C).

Центрифугировали клеточный лизат, и полученный супернатант лизата обрабатывали ДНКазой (DNase TURBO + Dnase TURBO Inactivation Reagent (Thermo

Fisher Scientific)). Затем эту суспензию инкубировали с биотинилированными олигонуклеотидами (RIP1, RIP1\_URA35, концентрация – 100 пмоль/мкл) в соотношении 1:100 к объему пробы в течение 1,5 часов при 37°C. Затем инкубировали, перемешивая, с шариками с нашитым стрептавидином Dynabeads MyOne Streptavidin T1 (Thermo Fisher Scientific) в течение 1,5 часов при 4°C. Так как олигонуклеотиды содержат биотин в своем составе, они прочно и специфично связываются со стрептавидином, нашитым на магнитные шарики.

После инкубации, на магните отбирали и сливали фракцию не связавшихся белков, промывали пять раз буфером для отмывки (20 мМ Трис-НСl (рН = 7.6), 150 мМ КСl, 5 мМ MgCl<sub>2</sub>, 0,1% Твин-20 (затем сменили на NP-40)). После чего, добавив к буферу додецилсульфат натрия до 0,1%, проводили элюцию белков в объеме 50 мкл при комнатной температуре в течение 5 минут.

Элюат анализировали при помощи гель-электрофореза по Лэммли, затем вырезали интересующие полосы и подготавливали пробы для масс-спектрометрического анализа. Сначала в течение ночи отмывали фрагменты геля дистиллированной водой, затем несколько раз промывали буфером для отмывки (50% ацетонитрил в 10 мМ NH<sub>4</sub>НСO<sub>3</sub>). Отбирали жидкость и высушивали фрагменты геля. Затем добавляли к фрагментам геля 10 мМ дитиотрейтол в 10 мМ NH<sub>4</sub>НСO<sub>3</sub> и инкубировали 15 мин при 56°C. Затем, для алкилирования цистеиновых остатков, добавляли по 5 мкл 55 мМ йодацетамида в 50 мМ растворе бикарбоната аммония. После чего удаляли избыток йодацетамида, повторно добавляя дитиотрейтол. Проводили трипсинолиз в геле, добавляя трипсин в концентрации 10 нМ в соотношении 1:50 к субстрату по массе. Инкубировали в течение ночи при 37°C. Затем добавляли 10% трифторуксусную кислоту для дезактивации трипсина и 50% ацетонитрил. После трипсинолиза фрагменты геля подвергали ультразвуку для выхода пептидов из геля. Полученный супернатант подвергали масс-спектрометрическому анализу.

Также, помимо анализа отдельных вырезанных полос геля после электрофореза, проводили трипсинолиз всей смеси элюированных белков. Для этого добавляли к раствору бикарбонат аммония до концентрации 50 мМ, восстанавливали дисульфидные связи в белках, добавляя дитиотрейтол до концентрации 45 мМ, затем, для алкилирования цистеиновых остатков, добавляли по 1 мкл 100 мМ йодацетамида в 50 мМ растворе бикарбоната аммония на каждый наномоль белка. Затем удаляли избыток

йодацетамида, повторно добавляя дитиотрейтол. Проводили трипсинолиз, добавляя трипсин в концентрации 10 нМ в соотношении 1:50 к субстрату по массе. Инкубировали в течение ночи при 37°C. Затем добавляли 10% трифторуксусную кислоту для дезактивации трипсина. Дополнительно пробы очищали при помощи носиков ZipTip Pipette Tips (EMD Millipore), содержащих специальную хроматографическую колонку C18 для очистки и концентрирования пептидов и белков размером от 3 до 100 кДа. После чего использовали высокоэффективную жидкостную хроматографию в сочетании с тандемной масс-спектрометрией (ВЭЖХ-МС) для идентификации полученных пептидов на масс-спектрометре с матрично-активированной лазерной десорбцией/ионизацией (МАЛДИ, MALDI) Bruker Ultraflexxtreme.

#### **4.3. Масс-спектрометрия**

Эксперименты по масс-спектрометрической идентификации выделенных белков проводились сотрудниками ресурсного центра "Развитие молекулярных и клеточных технологий" Санкт-Петербургского Государственного университета на масс-спектрометре с матрично-активированной лазерной десорбцией/ионизацией (МАЛДИ, MALDI) Bruker Ultraflexxtreme. Принцип метода состоит в лазерной ионизации пептидов, полученных в результате трипсинолиза выделенных белков. Затем измеряется время пролета ионизированного пептида в вакууме, которое определяет отношение массы к заряду этого пептида. Эта информация позволяет определить пептид по базам данных, которые встроены в обеспечение масс-спектрометрического прибора. В случае анализа полной фракции белков без предварительного разделения в геле использовали высокоэффективную жидкостную хроматографию в сочетании с тандемной масс-спектрометрией (ВЭЖХ-МС) на том же приборе.

#### **4.4. Метод ПЦР**

Проводили по стандартному протоколу (Lorenz, 2012), меняя параметры в зависимости от свойств используемых праймеров. Температура плавления ДНК составляла 95 °С, затем повторяли 30-40 циклов чередования температуры плавления, температуры отжига праймеров и температуры элонгации, затем пробы охлаждали до 4°C.

Смесь для 1 реакции ПЦР (25 мкл) содержала 5 мкл 5X буфера для ПЦР-реакции GoTaq Reaction Buffer Green (Promega Corporation), 200 нмоль смеси дезоксирибонуклеозидтрифосфатов, по 12,5 пмоль праймеров, 0,25 мкл полимеразы (смесь Taq-полимеразы и «Фурии» (Бигль) 9:1).

#### **4.5. Рестрикция плазмид UIRL-lacO-25-12-8 и UIRL-lacO-25-12-5 рестриктазой SmaI**

Проводили при оптимальной температуре действия (37°C) рестриктазы SmaI (FastDigest, Fermentas) в течение 1-1,5 ч. в буфере FastDigest (Thermo Fisher Scientific).

#### **4.6. Дрожжевая трансформация**

Для интеграции кассеты *URA3-Int* в геном дрожжей проводили интегративную трансформацию. Интегративная трансформация осуществлялась по протоколу дрожжевой трансформации с использованием ацетата лития (Gietz & Schiestl, 2007). Для трансформации штамма SMY710 использовали выделенные из геля ПЦР-продукты после рестрикции SmaI (FastDigest, Fermentas) плазмид UIRL-lacO-25-12-5 и UIRL-lacO-25-12-8 (содержащие повторы в разной ориентации и последовательность *lacO*). Наличие вставки проверяли методом ПЦР с использованием праймеров UIRL1 и A36a-R, длину повтора в вставке проверяли методом ПЦР с использованием праймеров UIRL1 и lacOR и секвенированием.

#### **4.7. Бактериальная трансформация**

Смешивали 100 мкл клеток *E.coli* XL-10 Gold с 1 мкл плазмиды, инкубировали 20 минут на льду, затем подвергали тепловому шоку – в течение 50 секунд нагревали пробирки до 42°C, затем перемещали на три минуты на лёд, затем инкубировали ещё три минуты при комнатной температуре. После этого добавляли четырехкратный избыток жидкой среды LB и инкубировали в течение часа при 37°C и перемешивании на скорости 165-180 rpm. После чего высевали на чашки со средой LB Агар с ампициллином (2,5 мкг/мл) по 200 мкл суспензии и растили в течение ночи при 37°C.

#### **4.8. Гель-фильтрация SUMO-LacI**

От научного руководителя Аксеновой А.Ю. были получены фракции рекомбинантного белка SUMO-LacI, наработанного из *E.coli* и очищенного на колонке

His-Trap HP (GE Healthcare Life Sciences). Белок содержит SUMO-таг, который в данной работе решено было оставить. Эти белковые фракции сначала были сконцентрированы на колонках (Pierce Protein Concentrators PES, 10K MWCO, 0.5 mL, Thermofisher Scientific) до концентрации 1,8 мг/мл, затем были очищены методом гель-фильтрации для дальнейшего изучения связывания SUMO-LacI с ДНК.

Для проведения гель-фильтрации раствор белка SUMO-LacI, перенесенный в 50 мМ бикарбонатный буфер, наносили на колонку Superdex 75 10/300 GL (GE Healthcare Life Sciences), использовали прибор NGC Chromatography Systems (Bio-Rad). Белок элюировали фосфатным буфером (PBS) фракциями по 500 мкл. Очищенный белок LacI затем использовался в дрожжей системе для выделения белков, взаимодействующих с теломерными последовательностями в ДНК.

#### **4.9. Анализ взаимодействия SUMO-LacI с ДНК, содержащей теломерные повторы и последовательность *lacO*.**

Инкубировали SUMO-LacI и ПЦР-продукты, полученные с плазмид UIRL, UIRL-*lacO*-25-12, UIRL-25-12 в молярном соотношении SUMO-LacI к ПЦР-продукту 1:4 в течение 3 часов при комнатной температуре. Затем смешивали 10 мкл смеси с буфером для нанесения Orange DNA Loading Dye 6X (Thermo Fisher Scientific) и проводили электрофорез в 6%-ном полиакриламидном геле в буфере TBE.

#### **4.10. Описание модельной системы для выделения белков, взаимодействующих с человеческими теломерными повторами в ДНК в дрожжевой системе.**

Система для выделения белков, взаимодействующих с теломерными повторами в ДНК основана на бактериальной системе взаимодействия LacI-*lacO*, которая регулирует метаболизм лактозы (Рис. 7). LacI является белком-репрессором, который, связываясь с регуляторным участком ДНК *lacO*, препятствует транскрипции лактозного оперона. Прочное и специфическое взаимодействие LacI-*lacO* используется нами для выделения белков, взаимодействующих с теломерными повторами в дрожжевой ДНК. Схема системы для выделения и идентификации белков, взаимодействующих с теломерными последовательностями, представлена на рисунке 12.

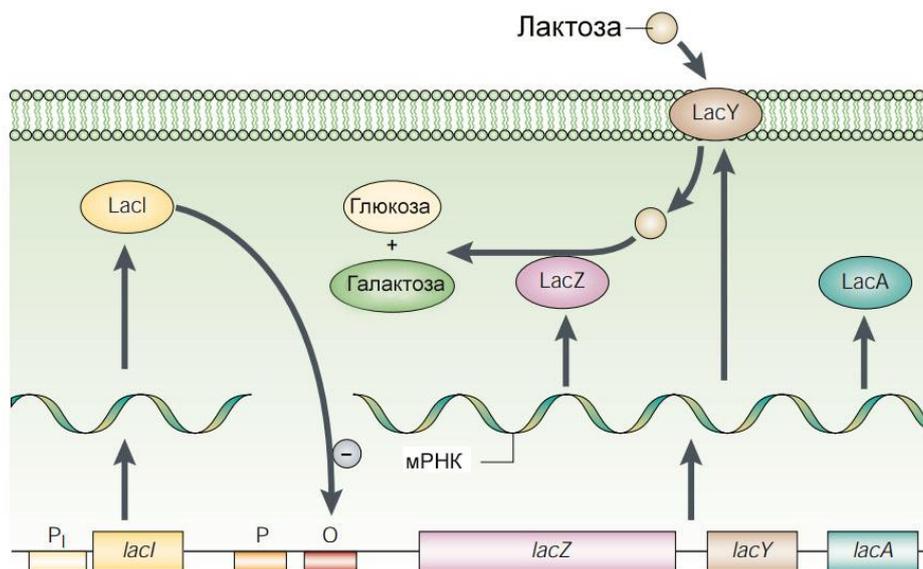


Рисунок 7. Схема регуляции лактозного оперона *E. coli*. Гены *lacZ*, *lacY* и *lacA* транскрибируются в составе одного оперона (промотор оперона обозначен буквой P). Образовавшаяся мРНК содержит открытые рамки считывания трёх белков, вовлеченных в метаболизм лактозы: бета-галактозидаза LacZ, пермеаза LacY и галактозид-О-ацетилтрансфераза LacA. Оперон регулируется репрессором LacI, который связывает промотор (место связывания обозначено буквой O). Взаимодействию репрессора с ДНК препятствуют индукторы, например, лактоза или ее искусственные аналоги (ИПТГ) (по Shuman & Silhavy, 2003 с изменениями).

#### 4.11. Фрагментация хроматина

Для подбора условий фрагментации использовали штаммы *S. cerevisiae* SMY758 и SMY803. Клетки выращивали в жидкой YPD до достижения оптической плотности ~0,6 (длина волны = 600 нм), затем лизировали в буфере РЕВ (20 мМ Трис-НСl (рН = 7.6), 140 мМ КСl, 5 мМ MgCl<sub>2</sub>, 0,5 мМ дитиотрейтол, 0,5 мМ фенилметилсульфоксид, 1% Тритон X-100, 5 мкМ леупептин, 5 мкМ, пепстатин А, 1 мМ бензамидин) при помощи шариков Silica Lysing Beads Lysing Matrix В (MP Biomedicals) и прибора FastPrep-24 Classic (MP Biomedicals) (30 сек 6.5 М/сек, 1,5 мин при + 4°C). Супернатант отбирали и проводили фрагментацию на ультразвуковом гомогенизаторе Bandelin SONOPULS HD 2070 с параметрами: 20 сек ультразвука\20 сек пауза, 6\*10% циклов, 70% мощности.

#### **4.12. Анализ структуры белка *in silico***

Последовательности для Ten1 *S. cerevisiae* и его гомологов были получены в базе данных UniProt (<http://www.uniprot.org/>).

Для предсказания вторичной структуры белка Ten1 использовались онлайн-версия программы PredictProtein Open (<https://www.predictprotein.org/>) и онлайн-версия программы PHYRE2 (Kelley et al., 2015). Для моделирования 3D структуры использовалась онлайн-версия программы PHYRE2.

## 5. Результаты

### 5.1. Выделение белков, взаимодействующих с РНК, содержащей человеческие теломерные повторы в дрожжевой системе.

Идентификация белков, взаимодействующих с РНК, содержащей человеческие теломерные повторы может помочь в выяснении механизмов, лежащих в основе наблюдаемых эффектов при вставке теломерных повторов. Недавно было показано, что вставка теломерных повторов внутрь дрожжевого репортерного гена приводит к снижению количества зрелой мРНК репортерного гена, при этом увеличивая количество пре-РНК (Aksenova et al., 2015). Выделение и идентификация белков, вовлеченных в процессы, обуславливающие наблюдаемые эффекты, поможет понять детали механизмов этих процессов.



Рисунок 8. Схема эксперимента для выделения и идентификации белков, взаимодействующих с теломерными последовательностями в РНК.

В данной работе мы выделяли белки, связывающие РНК, содержащую человеческие теломерные повторы, используя биотинилированные праймеры, комплементарные к последовательности РНК, содержащей человеческие теломерные повторы, и шарики с нашитым стрептавидином Dynabeads MyOne Streptavidin T1 (Thermo Fisher Scientific) (Рис. 8). Полученные белки анализировали при помощи гель-электрофореза по Лэммли, и затем вырезали интересующие нас фрагменты геля, обрабатывали их трипсином и анализировали методом масс-спектрометрии. Также были проанализированы пептиды, которые получают после трипсинолиза раствора, содержащего тотальную фракцию белков, взаимодействующих

с теломерными последовательностями в РНК, полученными после элюции с магнитных шариков. Для этого раствор после трипсинолиза полной фракции белков идентифицировали при помощи высокоэффективной жидкостной хроматографии в сочетании с тандемной масс-спектрометрией (ВЭЖХ-МС) (Рис. 8).

На данном этапе выяснилось, что при анализе всего раствора белков, полученного после элюции, используемый нами детергент (Твин-20) препятствует масс-спектрометрическому анализу белков, давая дополнительный сигнал, который препятствует определению пиков, соответствующих пептидам при анализе на масс-спектрометре. Детергент был заменен на NP-40 (который тоже не является идеальным для использования при масс-спектрометрическом анализе, но удаляется дополнительной очисткой эффективнее, чем Твин-20). Для дополнительной очистки проб использовали колонки HiPPR Detergent Removal Resin (Thermo Fisher Scientific) для удаления детергента из образцов, в ресурсном центре “Развитие молекулярных и клеточных технологий” использовали метод удаления детергентов при помощи этилацетата (Yeung & Stanley, 2010). Однако в первом случае не удалось избавиться от детергента полностью, а второй метод критически снижал количество пептидов в пробах.

Анализ белков из вырезанных из геля полос (Рис. 9) после электрофореза по Лэммли также был проведен при дополнительной очистке проб при помощи носиков ZipTip Pipette Tips (EMD Millipore), содержащих специальную хроматографическую колонку C18 для очистки и концентрирования пептидов и белков размером от 3 до 100 кДа. Нам удалось проанализировать некоторые пробы: так, например, в пробе № 3 (Рис. 9) с высокими уровнями значимости были идентифицированы белки Ten1 (белок регуляции длины теломер 1) и EF1-alfa (фактор элонгации альфа). Фактор элонгации является белком, связывающим РНК, поэтому мы ожидали его обнаружить. Ten1 был обнаружен именно в том штамме, в котором есть вставка человеческих теломерных повторов, тогда как в полосе контрольного штамма, которая соответствует белкам той же массы (проба № 1, рис. 9), его обнаружить не удалось. Этот белок участвует в регуляции длины теломер, защищает концевые участки, формируя комплексы с белками Cdc3p и Stn1p, является гомологом человеческого белка Ten1, который входит в состав комплекса CST.

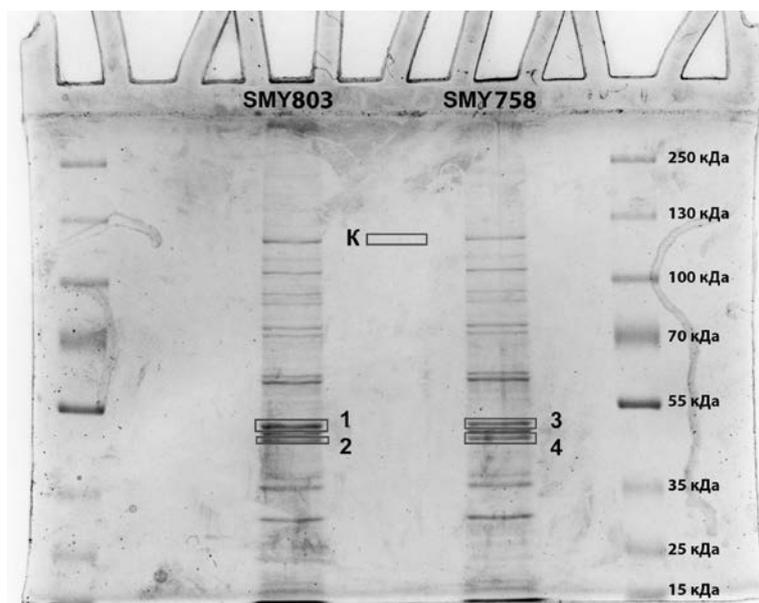


Рисунок 9. Фотография геля после электрофореза по Лэммли проб, полученных в результате эксперимента по выделению белков, связывающих РНК, содержащую человеческие теломерные повторы. SMY803 – обогащенная фракция белков, связывающих человеческие теломерные последовательности в РНК, выделенная из штамма *S. cerevisiae*, используемого в качестве контроля (без вставки), SMY758 – та же фракция, но выделенная из штамма SMY758 (со вставкой человеческих теломерных повторов в репортерном гене). К – вырезанная полоса пустого геля (контроль), 1-4 – вырезанные фрагменты геля для анализа белков методом масс-спектрометрии.

В пробах № 2 и № 4 в обоих штаммах была обнаружена фосфоглицераткиназа 1 (Pgk1), которая является белком, наиболее часто идентифицируемым методом масс-спектрометрии, и, дополнительно, в пробе № 4 с высокой значимостью был обнаружен продукт гена *BUD5* - Bud site selection protein 5. Этот белок является фактором ГТФ-ГДФ обмена малой ГТФазы Bud1/Rsr1, участвует в определении паттерна почкования и ориентации веретена деления. Также с низкой значимостью был определен белок Msc6, про который известно очень немного, но показано, что у мутантов по гену *MSC6* наблюдаются дефекты при мейотической рекомбинации сестринских хроматид (Thompson & Stahl, 1999).

Также с низким уровнем значимости (ниже 50) были обнаружены другие белки. Некоторые из них появлялись неоднократно при разных повторах экспериментов.

Наиболее часто встречающиеся при масс-спектрометрическом анализе белки приведены в таблице 3.

Таблица 3. Список белков, определенных методом масс-спектрометрии с низким уровнем значимости во фракции выделенных белков, взаимодействующих с теломерными последовательностями в РНК.

Имя белка в базах данных	Значимость
YL054_YEAST	48,5
EIS1_YEAST	42,2
ATM1_YEAST	42
RPC2_YEAST	41
SEC16_YEAST	39,6
YRA1_YEAST	39,37
ATR_YEAST	36,8
BRR2_YEAST	34,8

## 5.2. Анализ структуры белка Ten1 *in silico*

Последовательность белка Ten1 *S. cerevisiae* была проанализирована при помощи онлайн-версии программы PredictProtein Open (Рис. 10).

Предсказанные элементы вторичной структуры Ten1 представляют собой бета-складчатые слои и три потенциальные альфа-спирали, при этом в предсказанных бета-структурах видно регулярное чередование экспонированных и скрытых бета-структур. Также Ten1 содержит области с потенциальной неупорядоченной структурой (Рис. 10 и рис. 11).

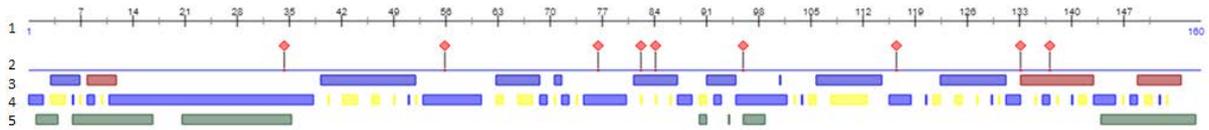


Рисунок 10. Результат предсказания вторичной структуры белка Ten1 *S. cerevisiae* при помощи программы PredictProtein Open. На шкале обозначены номера аминокислот в последовательности, на второй строке представлены мишени для модификаций, на третьей строке синие цилиндры обозначают элементы бета-складчатой укладки, красные цилиндры - предсказанные альфа-спирали. На четвертой строке синие цилиндры обозначают экспонированные элементы укладки, а желтые - скрытые внутри белковой глобулы элементы. На последней строке темно-зелеными цилиндрами отмечены области с вероятными неупорядоченными структурами.

3D-структура Ten1 *S. cerevisiae* была предсказана при помощи онлайн-версии программы PHYRE2 (Kelley et al., 2015) (Рис. 12). Этот веб-сервис позволяет эффективно предсказывать структуру разных фрагментов последовательности (включая неупорядоченные структуры), а затем располагать предсказанные с достаточной значимостью фрагменты в трёхмерной структуре.



Рисунок 11. Элементы вторичной структуры белка Ten1 *S. cerevisiae*, предсказанные программой PHYRE2. 1 - последовательность белка Ten1, 2 - элементы вторичной структуры, 3 - значимость предсказания элементов вторичной структуры, 4 - предсказание регионов с неупорядоченной структурой, 5 - значимость предсказания упорядоченности структуры.

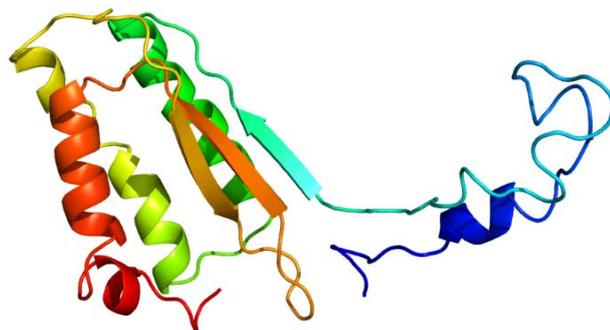


Рисунок 12. Трехмерная структура белка Ten1 *S. cerevisiae*, предсказанная программой PHyre2.

### 5.3. Выделение белков, взаимодействующих с теломерными последовательностями в ДНК

Система для выделения белков, взаимодействующих с теломерными последовательностями в ДНК включает в себя использование штаммов дрожжей *S. cerevisiae*, несущих в интроне репортерного гена *URA3* последовательность *lacO* и фланкирующие последовательности человеческих теломерных последовательностей, а также белка LacI, который с высокой аффинностью связывает последовательность *lacO*. В связи с этим, эксперимент по разработке системы выделения и идентификации белков, взаимодействующих с теломерными последовательностями в ДНК, включал в себя конструирование штаммов *S. cerevisiae* с последовательностями *lacO* и теломерных повторов, очистку белка LacI и проверку его связывания с ДНК (Рис. 13). Полученные штаммы выращивали до оптической плотности 0,6 и выделяли хроматин для последующей фрагментации. Фрагментированный хроматин предполагалось использовать для осаждения при помощи смолы с иммобилизованным белком LacI.

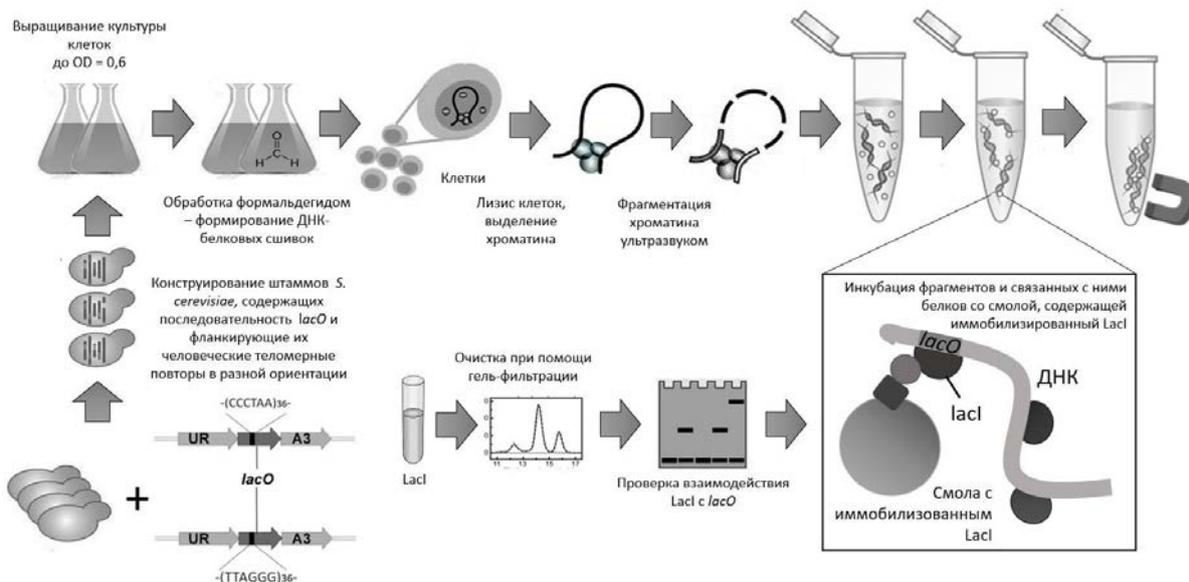


Рисунок 13. Схема эксперимента по выделению и идентификации белков, взаимодействующих с теломерными последовательностями в ДНК.

### 5.3.1. Очистка SUMO-LacI и анализ условий взаимодействия SUMO-LacI с *lacO*

ДНК-связывающие белки играют важную роль в регуляции экспрессии генов, поэтому изучение белков, взаимодействующих с человеческими теломерными повторами в ДНК может помочь детально изучить процессы, обуславливающие свойства повторяющихся последовательностей в геноме. Схема эксперимента приведена на рисунке 12.

Для выделения белков, взаимодействующих с человеческими теломерными повторами использовалась система, основанная на взаимодействии LacI-*lacO*. Поэтому необходимо было доочистить имеющийся выделенный из *E. coli* рекомбинантный белок SUMO-LacI, а затем проверить наличие его взаимодействия с ДНК. Сначала были проанализированы имеющиеся фракции после очистки на колонке His-Trap (Рис. 14)

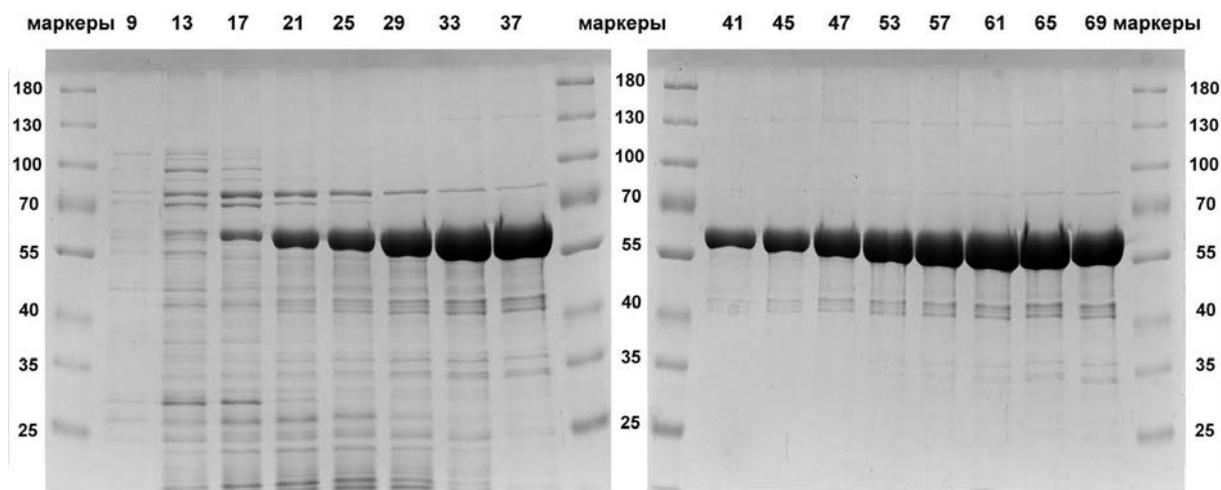


Рисунок 14. Фотографии акриламидных гелей (10%) после электрофореза по Лэмбли. 9, 13, 17, 21, 25, 29, 33, 37, 43, 45, 49, 53, 57, 61, 65, 69 - номера фракции SUMO-LacI после очистки на колонке His-Trap HP (GE Healthcare Life Sciences).

Затем мы сконцентрировали фракции 53-69 на колонках (Pierce Protein Concentrators PES, 10K MWCO, 0.5 mL, Thermofisher Scientific).

Провели гель-фильтрацию сконцентрированных проб 53-69 с более высокой концентрацией LacI (Рис.14). Отобрали полученные фракции А2-А9 и объединили их.

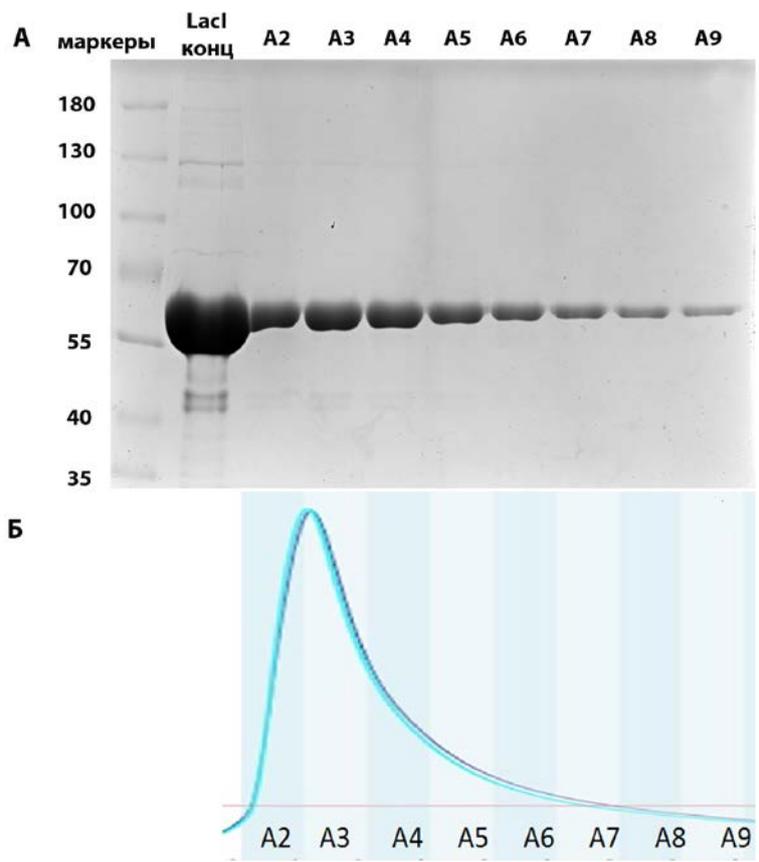


Рисунок 15. А - Фотография акриламидного геля (10%) после гель-фильтрации проб SUMO-LacI. Б - пример хроматограммы гель-фильтрации сконцентрированных проб SUMO-LacI. “LacI конц” – сконцентрированный белок до гель-фильтрации, А2-А9 – исследованные фракции, полученные после гель-фильтрации (объемом по 500 мкл).

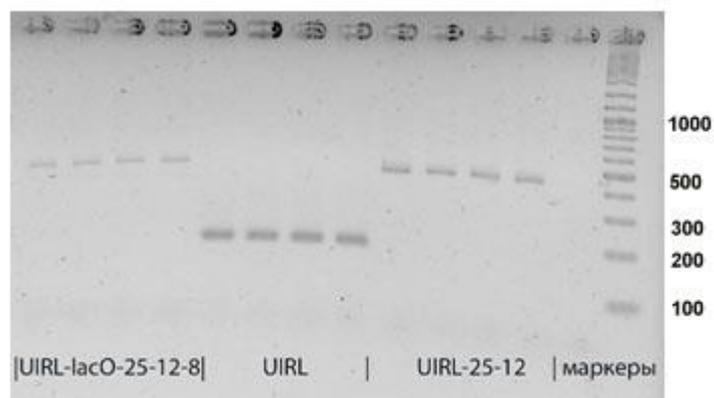


Рисунок 16. Фотография агарозного геля (1%) после ПЦР с плазмид UIRL-lacO-25-12-8 (содержит *lacO* и -(CCCTAA)<sub>36</sub>-), UIRL (не содержит *lacO* и повторы) и UIRL-25-12 (содержит повтор -(CCCTAA)<sub>36</sub>-) с использованием праймеров UIRL1 и lacOR.

Для того, чтобы подтвердить способность SUMO-LacI взаимодействовать с ДНК и подобрать условия для их эффективного взаимодействия, мы анализировали электрофоретическую подвижность очищенного и сконцентрированного SUMO-LacI после инкубации с продуктами ПЦР, содержащими *lacO* и контролями, содержащими теломерные повторы (Рис. 16). Были подобраны оптимальные условия инкубации: молярное соотношение SUMO-LacI к ПЦР-продукту = 1:4, инкубация в течение 3 часов при комнатной температуре в 10 мМ Tris 5 мМ MgCl<sub>2</sub> (Рис. 17).

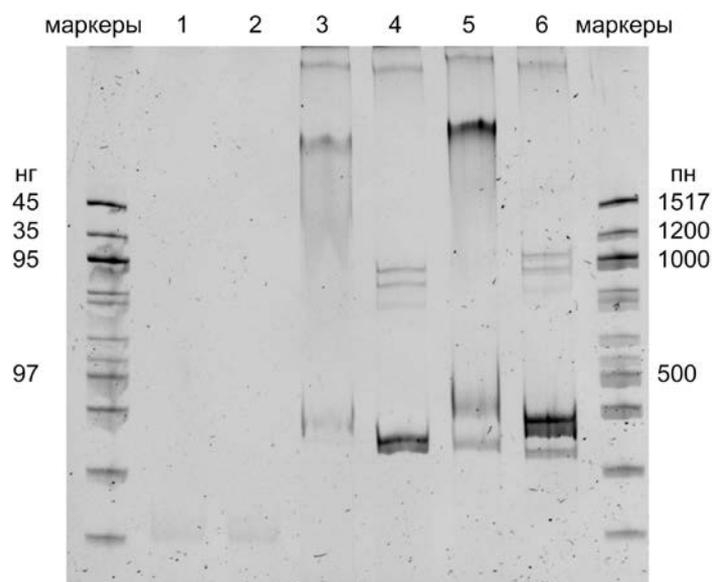


Рисунок 17. Фотография геля после анализа взаимодействия SUMO-LacI и ПЦР-продуктов: 1 - UIRL и LacI, 2 - UIRL и БСА, 3 - UIRL-25-12-8 и LacI, 4 - UIRL-25-12-8 и БСА, 5 - UIRL-lacO-25-12-8 и LacI, 6 - UIRL-lacO-25-12-8 и БСА. Во всех пробах молярное отношение ПЦР-продукта к белку 1:4, инкубация проводилась в течение 3 часов при комнатной температуре.

На геле (Рис. 17) на 5 дорожке можно наблюдать связывание белка SUMO-LacI с ПЦР-продуктом плазмиды UIRL-lacO-25-12-8, тогда как в соответствующем контроле (дорожка 5), в которую был добавлен бычий сывороточный альбумин в той же концентрации, связывания не наблюдается. Однако, в дорожке 4 также наблюдается взаимодействие белка LacI с ПЦР-продуктом плазмиды UIRL-25-12, но в меньшей степени. Дорожки 1 и 2 оказались малоинформативны, так как концентрация ПЦР-продукта плазмиды UIRL оказалась недостаточной. Однако видно, что разница между 1 и 2 дорожкой отсутствует.

### 5.3.2. Конструирование штаммов *S.cerevisiae* AAY51 и AAY52.

Для создания дрожжевой системы для выделения и идентификация белков, взаимодействующих с теломерными последовательностями в ДНК, были сконструированы дрожжевые штаммы, имеющие в интроне репортерного гена *URA3* вставку человеческих теломерных повторов в разной ориентации и последовательность *lacO*, необходимую для специфического выделения связывающих ДНК белков за счет взаимодействия *LacI-lacO* (Рис. 18). Для этого использовали фрагменты, полученные после рестрикции плазмид UIRL-*lacO*-25-12-8 и UIRL-*lacO*-25-12-5 рестриктазой *SmiI* (FastDigest, Fermentas), была проведена трансформация штамма *S.cerevisiae* SMY710. Отбирали клоны на селективной среде SC -Trp, так как кассета *URA3-Int* содержит *TRP1*. Затем при помощи ПЦР с использованием праймеров UIRL1 и a-36R были отобраны клоны со вставкой, а затем длина повтора в ставке была проверена с использованием праймеров UIRL1 и *lacOR* и методом секвенирования.

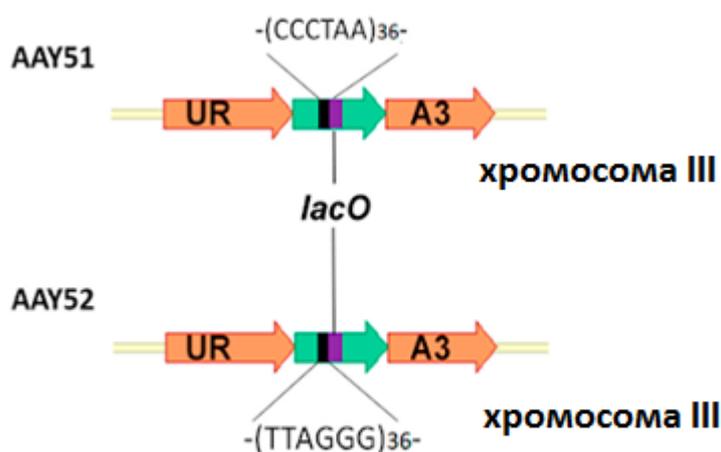


Рисунок 18. Схематичное изображение взаимного расположения вставок человеческих теломерных повторов в разной ориентации и последовательности *lacO* в интроне репортерного гена *URA3* в третьей хромосоме в сконструированных штаммах AAY51 и AAY52.

Сконструированные штаммы предполагается использовать для выделения белков, взаимодействующих с теломерными повторами в ДНК. Мы провели иммобилизацию белка SUMO-LacI на смоле NHS Mag Sepharose (GE Healthcare Life Sciences) и

планируем использовать ее для селективного осаждения фрагментов ДНК, содержащих последовательность *lacO*, и связанных с ними белков из экстрактов штаммов *S.cerevisiae* ААУ51 и ААУ52. Селективное обогащение фрагментами ДНК, содержащих вставки *lacO* и фланкирующих их последовательностей человеческих теломерных повторов будет происходить за счет взаимодействия белка LacI с вставкой *lacO*. Полученные при этом осаждении белки будут идентифицированы методом высокоэффективной жидкостной хроматографии в сочетании с тандемной масс-спектрометрией (ВЭЖХ-МС). В качестве контроля планируется использовать штамм *S. cerevisiae* ААУ11.

### 5.3.3. Фрагментация хроматина.

Для того, чтобы эффективно выделить именно те белки, которые взаимодействуют с теломерными повторами в ДНК, фрагменты ДНК должны быть размером около 500 пн. Подбор условий для фрагментации хроматина производился на штаммах SMY758 и SMY803. Фрагментацию проводили на ультразвуковом гомогенизаторе Bandelin SONOPULS HD 2070 с параметрами: 20 сек ультразвука\20 сек пауза, 6\*10% циклов, 70% мощности. После 10, 20, 30, 40 циклов фрагментации отбирали аликвоты, которые затем инкубировали в течение 3 часов при 65°C с протеиназой К (Proteinase K, Helicon). Затем инкубировали в течение 15 минут при 37°C с РНКазой А (RNase A, Thermo Fisher Scientific) и проверяли размер фрагментов при помощи электрофореза (Рис.19). Несмотря на то, что в рамках данной работы не удалось оптимизировать этап фрагментации хроматина для получения фрагментов со средней длиной 500 пн, полученные фрагменты были успешно использованы для селективного осаждения на смоле с иммобилизованным белком LacI. Эксперименты проводились научным руководителем Аксеновой А.Ю. и в рамках данной работы не обсуждаются.

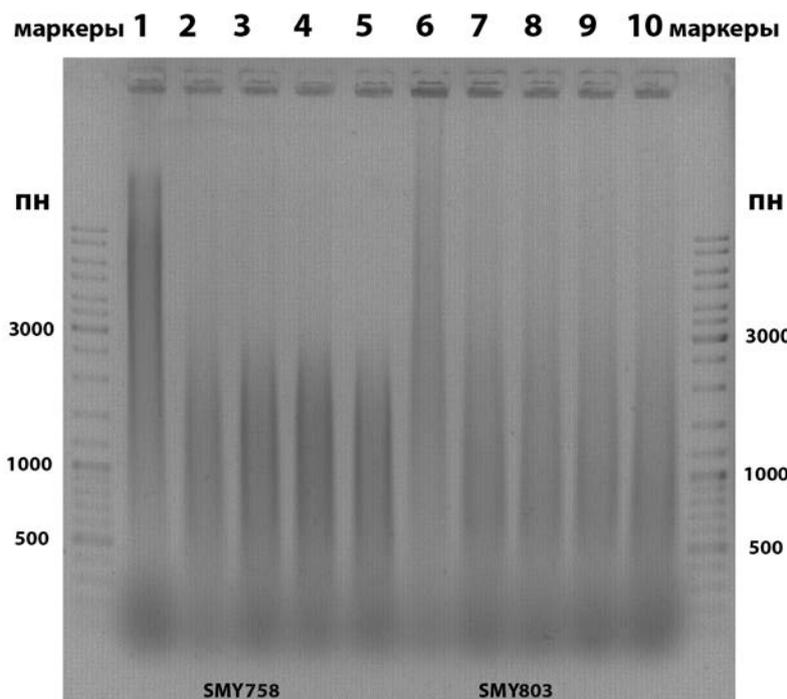


Рисунок 19. Пример агарозного геля (1%) после фрагментации хроматина прибором Bandelin SONOPULS HD 2070. «контроль» - хроматин до фрагментации, 10, 20, 30, 40 – число циклов 20 сек ультразвука\20 сек паузы, после которых были отобраны аликвоты.

## 6. Обсуждение результатов

Теломеры и теломерные последовательности играют важную роль в обеспечении стабильности генома. Функционирование теломерных последовательностей тесно взаимосвязано с белками, специфически взаимодействующими с теломерными повторами. Открытие новых участников функционирования теломерных последовательностей в геноме имеет важное фундаментальное и прикладное значение. Во-первых, эволюция теломер и теломерных последовательностей тесно связана с появлением линейных хромосом у эукариот. Выявление новых белков, участвующих в регуляции теломер и теломерных последовательностей, поможет дополнить уже имеющиеся знания об эволюции организмов и понять, как меняются белки и выполняемые ими функции, как это взаимосвязано с функциями теломер (Lue, 2018). Во-вторых, дисфункции теломер и теломерных последовательностей ассоциированы с различными патологическими состояниями, включая опухолеобразование (Meena et al., 2015), прогерия и некоторые другие. Теломерные последовательности на концевых и внутренних сайтах хромосом вовлечены в сложные и регуляторные сети, которые обеспечивают регуляцию экспрессии генов и правильную трехмерную организацию хроматина в ядре (Shay, 2018). Выяснение деталей этих взаимодействий поможет как в разработке терапии для заболеваний, так и в понимании природы клеточного старения. Во многих лабораториях разрабатываются и активно используются различные методы выделения и идентификации белков, вовлеченных в функционирование теломерных последовательностей.

Среди современных методов идентификации новых белков, взаимодействующих с теломерными последовательностями можно назвать маркировку (мечение) стабильными изотопами аминокислот в клеточной культуре (от англ. SILAC - stable isotope labeling with amino acids in cell culture). Этот популярный количественный метод протеомного анализа основан на использовании питательной среды с аминокислотами, которые содержат в своем составе “тяжелые” изотопы, часто  $^{15}\text{N}$  и  $^{13}\text{C}$ . Разница в сигнале от “меченых” и не “меченных” аминокислот анализируется методом масс-спектрометрии. Метод затем был адаптирован для выделения белков, взаимодействующих с ДНК (Mittler et al, 2009) и стал активно использоваться для идентификации белков, взаимодействующих с различными специфическими последовательностями. В частности, он был использован для выявления белков, взаимодействующих с теломерными последовательностями в ДНК и позволил

обнаружить, что белок млекопитающих HOP1 способен специфически взаимодействовать с теломерной ДНК (Karpej et al., 2013). Этот же подход использовался для выявления новых белков, связывающих РНК TERRA (Scheibe et al., 2013), для обнаружения новых белков, взаимодействующих с теломерными последовательностями в ДНК у *Neurospora crassa*, грибка, теломеры которого очень похожи на теломеры позвоночных животных (Casas-Vila et al., 2015) также использовали похожий подход, но уже без использования изотопов.

Кроме того, для идентификации теломерных белков используются методы, основанные на иммунопреципитации теломерного хроматина (Q-TIP) антителами к TRF2 (Grolimund et al., 2013). Все эти методы являются довольно эффективными, однако непростыми для использования. Большинство из них предполагает работу с изотопами, что сильно осложняет использование методики в лабораториях некоторых стран. В методах, основанных на иммунопреципитации, выделение новых белков основано на их взаимодействии с уже известными белками, связывающими теломерные последовательности, что делает эти методики неэффективными для объективного обнаружения новых белков. Идентификация выделенных белков в системах, которые разрабатывались и оптимизировались в данной работе, также основана на масс-спектрометрическом анализе. Однако выделение белков, взаимодействующих с теломерными последовательностями в РНК и ДНК, не включает в себя использование изотопов и специфичных к уже известным теломер-связывающим белкам антител. Более того, использование относительно более простой и дрожжевой системы делает эти системы удобными для адаптации этого метода к различным модификациям.

Среди белков, взаимодействующих с теломерными последовательностями в РНК, был обнаружен компонент комплекса CST Ten1. Этот комплекс играет главную роль в функционировании теломер у дрожжей *S. cerevisiae*, и в его состав помимо Ten1 входят белки Stn1 и Cdc13. Все компоненты комплекса являются ДНК-связывающими белками, для Stn1 и Cdc13 показано наличие доменов с ОВ-укладкой, однако структура Ten1 остается мало изученной. Домены с ОВ-укладкой являются доменами с низкой консервативностью по последовательности и преимущественно состоят из бета-бочонка и нескольких экспонированных альфа-спиралей, которые формируют домен связывания пептидов и одноцепочечных ДНК и РНК.

Ten1 взаимодействует с одноцепочечной ДНК, однако для него не показана способность связывания с теломерными последовательностями в РНК, поэтому, если

эти результаты будут подтверждены, возможно будет говорить о дополнительных функциях Ten1 при взаимодействии с РНК, содержащей теломерные последовательности. Для Ten1 произвели биоинформатический анализ структуры. Примечательно, что Ten1 *S. cerevisiae* имеет слабую степень гомологии с белками Ten1 других организмов, даже, например, дрожжей *Shizosaccharomyces pombe*. При использовании онлайн-версии программы PredictProtein Open были предсказаны элементы вторичной структуры белка Ten1 (Рис.10). В его структуре предсказано наличие бета-складчатых элементов и несколько альфа-спиралей, которые могут представлять собой определенный тип ОВ-укладки. Чередование экспонированных и скрытых бета-структур Ten1 характерно для структуры бета-бочонка. Также наличие последовательностей с предположительной неорганизованной структурой косвенно может подтверждать способность связывания РНК, потому что подобные регионы характерны и очень важны для РНК-связывающих белков (Calabretta & Richard, 2015). С использованием ресурса RHYRE2 была предсказана трехмерная структура белка Ten1, которая строилась на основе множественного выравнивания (Рис. 12). Структура предсказана с низким доверительным интервалом (65,3 и меньше), что говорит о низкой гомологии Ten1 со всеми известными белками в базах данных. Также RHYRE2 делает дополнительное предсказание вторичной структуры и так же, как и PredictProtein Open показывает потенциальное наличие участков с неорганизованной структурой и схожие паттерны бета-слоев и альфа-спиралей (Рис. 10 и рис. 11).

Помимо Ten1 с низкой значимостью были обнаружены некоторые белки, которые могут быть также вовлечены во взаимодействие с теломерными последовательностями. В частности, обнаруженный белок Yga1 играет важную роль в формировании стабильных комплексов мРНК и их экспорте из ядра (Zenklusen et al., 2002). Был обнаружен белок Mec1 (ATR\_YEAST в таблице 3), участвующий в сигнальных процессах репарации ДНК, который ассоциирован с теломерами и участвует в процессах рекомбинации и функционирования теломер. Также в пробах с низкой значимостью обнаруживается РНК-геликаза Brr2, компонент сплайсосомы. Эти результаты подтверждают, что метод позволяет выделять белки, вовлеченные в процессы функционирования РНК, содержащей теломерные последовательности и при дальнейшей оптимизации позволит с высокой эффективностью выделять и идентифицировать новые компоненты функционирования теломерных последовательностей в геноме.

Дальнейшая оптимизация метода выделения и идентификации белков, взаимодействующих с теломерными последовательностями в ДНК, позволит получить удобный метод, не основанный на предположениях или дополнительном знании об уже существующих теломер-связывающих белках.

## 7. Выводы

1. В ходе работы разработан и оптимизирован метод выделения и идентификации белков, взаимодействующих с теломерными последовательностями в РНК.
2. Отработана идентификация выявленных белков, взаимодействующих с человеческими теломерными последовательностями в РНК, методом масс-спектрометрии. Среди идентифицированных белков, взаимодействующих с человеческими теломерными последовательностями в РНК, обнаружен белок Ten1, участвующий в регуляции длины теломер.
3. Разработан метод выделения и идентификации белков, взаимодействующих с теломерными повторами в ДНК. Показано специфическое взаимодействие SUMO-LacI с последовательностями, содержащими теломерные повторы и *lacO* в модельной дрожжевой системе.

## **8. Благодарности**

Хотелось бы выразить благодарность своему научному руководителю Аксеновой Анне Юрьевне за бесценные советы и чуткое руководство, куратору работы Соповой Юлии Викторовне за помощь и советы в самых непростых ситуациях, возникавших в ходе работы. Хочется поблагодарить Чернова Юрия Олеговича, Рубеля Александра Анатольевича и всех сотрудников Лаборатории Биологии амилоидов за создание дружественной атмосферы и постоянную поддержку и готовность помочь. Особую благодарность хочется выразить всем сотрудникам ресурсного центра “Развитие молекулярных и клеточных технологий” Санкт-Петербургского Государственного университета за помощь в проведении важных экспериментов данной работы.

## 9. Список литературы

1. Aeby, E. & Lingner, J., 2015. ALT telomeres get together with nuclear receptors. *Cell*, 160(5), pp.811–813.
2. Aksenova, A.Y., Greenwell, P., Dominska, M., Shishkin, A.A., Kim, J.C., Petes, T.D. & Mirkin, S.M., 2013. Genome rearrangements caused by interstitial telomeric sequences in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 110(49), pp.19866–19871.
3. Aksenova, A.Y., Han, G., Shishkin, A.A., Volkov, K.V. & Mirkin, S.M., 2015. Expansion of Interstitial Telomeric Sequences in Yeast. *Cell Reports*, 13(8), pp.1545–1551.
4. Arnerić, M. & Lingner, J., 2007. Tell kinase and subtelomere-bound Tbf1 mediate preferential elongation of short telomeres by telomerase in yeast. *EMBO Reports*, 8(11), pp.1080-1085.
5. Azzalin, C.M., Reichenbach, P., Khoriauli, L., Giulotto, E. & Lingner, J., 2007. Telomeric repeat-containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science*, 318(5851), pp.798–801.
6. Balk, B., Maicher, A., Dees, M., Klermund, J., Luke-Glaser, S., Bender, K. & Luke, B., 2013. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nature Structural and Molecular Biology*, 20(10), pp.1199–1206.
7. Bandaria, J.N., Qin, P., Berk, V., Chu, S. & Yildiz, A., 2016. Shelterin protects chromosome ends by compacting telomeric chromatin. *Cell*, 164(4), pp.735–746.
8. Bhattacharjee, A., Wang, Y., Diao, J. & Price, C.M., 2017. Dynamic DNA binding, junction recognition and G4 melting activity underlie the telomeric and genome-wide roles of human CST. *Nucleic Acids Research*, 45(21), pp.12311–12324.
9. Bianchi, A. & Shore, D., 2008. How Telomerase Reaches Its End: Mechanism of Telomerase Regulation by the Telomeric Complex. *Molecular Cell*, 31(2), pp.153–165.
10. Bilaud, T., Koering, C.E., Binet-Brasselet, E., Ancelin, K., Pollice, A., Gasser, S.M. & Gilson, E., 1996. The telobox, a Myb-related telomeric DNA binding motif found in proteins from yeast, plants and human. *Nucleic Acids Research*, 24(7), pp.1294-1303.
11. Blackburn, E.H., 1991. Structure and function of telomeres. *Nature*, 350(6319), pp.569–573.
12. Blasco, M.A., 2007. The epigenetic regulation of mammalian telomeres. *Nature Reviews Genetics*, 8(4), pp.299–309.

13. Bolzán, A.D., 2017. Interstitial telomeric sequences in vertebrate chromosomes: Origin, function, instability and evolution. *Mutation Research - Reviews in Mutation Research*, 773, pp.51–65.
14. Bonetti, D., Martina, M., Clerici, M., Lucchini, G. & Longhese, M.P., 2009. Multiple pathways regulate 3' overhang generation at *S. cerevisiae* telomeres, *Molecular cell*, 35, pp. 70–81
15. Bradshaw, P.S., Stavropoulos, D.J. & Meyn, M.S., 2005. Human telomeric protein TRF2 associates with genomic double-strand breaks as an early response to DNA damage. *Nature genetics*, 37(2), pp.193–197.
16. Brigati, C., Kurtz, S., Balderes, D., Vidali, G. & Shore, D., 1993. An essential yeast gene encoding a TTAGGG repeat-binding protein. *Molecular and cellular biology*, 13(2), pp.1306–1314.
17. Bryan, C., Rice, C., Harkisheimer, M., Schultz, D.C. & Skordalakes, E., 2013. Structure of the Human Telomeric Stn1-Ten1 Capping Complex. *PLoS ONE*, 8(6).
18. Calabretta, S. & Richard, S., 2015. Emerging Roles of Disordered Sequences in RNA-Binding Proteins. *Trends in biochemical sciences*, 40(11), pp.662-672.
19. Casas-Vila N., Scheibe, M., Freiwald, A., Kappei, D. & Butter F., 2015. Identification of TTAGGG-binding proteins in *Neurospora crassa*, a fungus with vertebrate-like telomere repeats. *BMC Genomics*. 16, 965.
20. Chan, A., Boulé, J.B. & Zakian, V.A., 2008. Two pathways recruit telomerase to *Saccharomyces cerevisiae* telomeres. *PLoS Genetics*, 4(10):e1000236.
21. Chirino, M.G., Dalíková, M., Marec, F.R. & Bressa, M.J., 2017. Chromosomal distribution of interstitial telomeric sequences as signs of evolution through chromosome fusion in six species of the giant water bugs (Hemiptera, Belostoma). *Ecology and Evolution*, 7(14), pp.5227–5235.
22. Cohn, M., 2013. OB fold contributes to telomere maintenance. *Structure*, 21(1), pp.3–4.
23. Conrad, M.N., Wright, J.H., Wolf, A.J. & Zakian, V.A., 1990. RAP1 protein interacts with yeast telomeres in vivo: Overproduction alters telomere structure and decreases chromosome stability. *Cell*, 63(4), pp.739–750.
24. Cubiles, M.D., Barroso, S., Vaquero-Sedas, M.I., Enguix, A.1., Aguilera, A., Vega-Palas, M.A., 2018. Epigenetic features of human telomeres. *Nucleic Acids Research*, 46(5), pp.2347–2355.

25. Cusanelli, E., Romero, C. & Chartrand, P., 2013. Telomeric Noncoding RNA TERRA Is Induced by Telomere Shortening to Nucleate Telomerase Molecules at Short Telomeres. *Molecular Cell*, 51(6), pp.780–791.
26. Deng, Z., Wang, Z., Stong, N., Plasschaert, R., Moczan, A., Chen, H.S., Hu, S., Wikramasinghe, P., Davuluri, R.V., Bartolomei, M.S., Riethman, H. & Lieberman, P.M., 2012. A role for CTCF and cohesin in subtelomere chromatin organization, TERRA transcription, and telomere end protection. *EMBO Journal*, 31(21), pp.4165–4178.
27. Dewar, J.M. & Lydall, D., 2012. Similarities and differences between “uncapped” telomeres and DNA double-strand breaks. *Chromosoma*, 121(2), pp.117–130.
28. Dilley, R.L. & Greenberg, R.A., 2015. ALternative Telomere Maintenance and Cancer. *Trends in Cancer*, 1(2), pp.145–156.
29. Doksani, Y., Wu, J.Y., de Lange. T. & Zhuang X., 2013. XSuper-resolution fluorescence imaging of telomeres reveals TRF2-dependent T-loop formation. *Cell*, 155(2), pp.345–356.
30. Ellahi, A., Thurtle, D.M. & Rine, J., 2015. The chromatin and transcriptional landscape of native *Saccharomyces cerevisiae* telomeres and subtelomeric domains. *Genetics*, 200(2), pp.505–521.
31. Episkopou, H., Draskovic, I., Van Beneden, A., Tilman, G., Mattiussi, M., Gobin, M., Arnoult, N., Londoño-Vallejo, A. & Decottignies, A., 2014. Alternative Lengthening of Telomeres is characterized by reduced compaction of telomeric chromatin. *Nucleic Acids Research*, 42(7), pp.4391–4405.
32. Faravelli, M., Azzalin, C.M., Bertoni, L., Chernova, O., Attolini, C., Mondello, C. & Giulotto, E., 2002. Molecular organization of internal telomeric sequences in Chinese hamster chromosomes. *Gene*, 283(1–2), pp.11–16.
33. Flynn, L.R. & Zou, L., 2011. Oligonucleotide/Oligosaccharide-Binding (OB) Fold Proteins: A Growing Family of Genome Guardians. *Crit Rev Biochem Mol Biol.*, 45(4), pp.266–275.
34. Fourel, G., Revardel, E., Koering, C.E. & Gilson, E., 1999. Cohabitation of insulators and silencing elements in yeast subtelomeric regions. *The EMBO Journal*, 18(9), pp.2522–2537.
35. Frank, C.J., Hyde, M. & Greider, C.W., 2006. Regulation of Telomere Elongation by the Cyclin-Dependent Kinase CDK1. *Molecular Cell*, 24(3), pp.423–432.

36. Gao, H., Cervantes, R.B., Mandell, E.K., Otero, J.H. & Lundblad, V., 2007. RPA-like proteins mediate yeast telomere function. *Nature Structural and Molecular Biology*, 14(3), pp.208–214.
37. García-Cao, M., O'Sullivan, R., Peters, A.H., Jenuwein, T. & Blasco, M.A., 2004. Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nature Genetics*, 36(1), pp.94–99.
38. Garrobo, I., Marión, R.M., Domínguez, O., Pisano, D.G. & Blasco, M.A., 2014. Genome-wide analysis of in vivo TRF1 binding to chromatin restricts its location exclusively to telomeric repeats. *Cell Cycle*, 13(23), pp.3742–3749.
39. Gietz, R.D., Schiestl, R.H., 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, 2(1), pp.31-34.
40. Gonzalez-Suarez, I. & Gonzalo, S., 2009. Crosstalk between chromatin structure, nuclear compartmentalization, and telomere biology. *Cytogenetic and Genome Research*, 122(3–4), pp.202–210.
41. Gottschling, D.E., Aparicio, O.M., Billington, B.L. & Zakian, V.A., 1990. Position effect at *S. cerevisiae* telomeres: Reversible repression of Pol II transcription. *Cell*, 63(4), pp.751–762.
42. Grandin, N., Damon, C. & Charbonneau, M., 2000. Cdc13 cooperates with the yeast Ku proteins and Stn1 to regulate telomerase recruitment. *Molecular and cellular biology*, 20(22), pp. 8397-8408.
43. Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H. & de Lange, T., 1999. Mammalian telomeres end in a large duplex loop. *Cell*, 97(4), pp.503–514.
44. Greetham, M., Skordalakes, E., Lydall, D., Connolly, B.A., 2015. The Telomere Binding Protein Cdc13 and the Single-Stranded DNA Binding Protein RPA Protect Telomeric DNA from Resection by Exonucleases. *Journal of molecular biology*, 427(19), pp.3023-3030.
45. Grolimund, L., Aeby, E., Hamelin, R., Armand, F., Chiappe, D., Moniatte, M. & Lingner, J., 2013. A quantitative telomeric chromatin isolation protocol identifies different telomeric states. *Nature Communications*. 4, 284.
46. Hockemeyer, D., Daniels, J.P., Takai, H. & de Lange, T. 2006. Recent Expansion of the Telomeric Complex in Rodents: Two Distinct POT1 Proteins Protect Mouse Telomeres. *Cell*, 126(1), pp.63–77.
47. Hom, R.A. & Wuttke, D.S., 2017. Human CST Prefers G-Rich but Not Necessarily Telomeric Sequences. *Biochemistry*, 56(32), pp.4210–4218.

48. Hu, C., Rai, R., Huang, C., Broton, C., Long, J., Xu, Y., Xue, J., Lei, M., Chang, S. & Chen, Y., 2017. Structural and functional analyses of the mammalian TIN2-TPP1-TRF2 telomeric complex. *Cell Research*, 27(12), pp.1485–1502.
49. Ichikawa, Y., Nishimura, Y., Kurumizaka, H. & Shimizu, M., 2015. Nucleosome organization and chromatin dynamics in telomeres. *Biomolecular Concepts*, 6(1), pp.67–75.
50. Ijdo, J.W., Baldini, A., Ward, D.C., Reeders, S.T. & Wells, R.A., 1991. Origin of human chromosome 2: an ancestral telomere-telomere fusion. *Proceedings of the National Academy of Sciences of the United States of America*, 88(20), pp.9051–9055.
51. Ilicheva, N.V., Podgornaya, O.I. & Voronin, A.P., 2015. Telomere Repeat-Binding Factor 2 Is Responsible for the Telomere Attachment to the Nuclear Membrane. *Advances in Protein Chemistry and Structural Biology, Elsevier Inc*, 101, pp.67-96.
52. Kaizer, H., Connelly, C.J., Bettridge, K., Viggiani, C. & Greider, C.W., 2015. Regulation of telomere length requires a conserved N-terminal domain of RiF2 in *Saccharomyces cerevisiae*. *Genetics*, 201(2), pp.573–586.
53. Karlseder, J., Broccoli, D. & Dai, Y., 1999. p53- and ATM-Dependent Apoptosis Induced by Telomeres Lacking TRF2. *Science*, 283, pp.0–4.
54. Kappei, D., Butter, F., Benda, C., Scheibe, M., Draškovič, I., Stevense, M, Novo, C.L., Basquin, C., Araki, M., Araki, K., Krastev, D,B., Kittler, R., Jessberger, R., Londoño-Vallejo, J.A., Mann, M. & Buchholz, F., 2013. HOT1 is a mammalian direct telomere repeat-binding protein contributing to telomerase recruitment. *The EMBO Journal*, 32(12), pp.1681-1701.
55. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. & Sternberg, M.J., 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6), pp. 845-858.
56. Kim, W., Ludlow, A.T., Min, J., Robin, J.D., Stadler, G., Mender, I., Lai, T.P., Zhang, N., Wright, W.E. & Shay, J.W., 2016. Regulation of the Human Telomerase Gene TERT by Telomere Position Effect—Over Long Distances (TPE-OLD): Implications for Aging and Cancer. *PLoS Biology*, 14(12), pp.1–25.
57. Kim, J.C. et al., 2017. The role of break-induced replication in large-scale expansions of (CAG)<sub>n</sub>•(CTG)<sub>n</sub> repeats. *Nat Struct Mol Biol*, 24(1), pp.55–60.
58. Kondo, K., Mashima, T., Oyoshi, T., Yagi, R., Kurokawa, R., Kobayashi, N., Nagata, T. & Katahira, M., 2018. Plastic roles of phenylalanine and tyrosine residues of TLS/FUS in complex formation with the G-quadruplexes of telomeric DNA and TERRA. *Scientific Reports*, 8(1), pp.1–11.

59. Kyrion, G., Boakye, K.A. & Lustig, A.J., 1992. C-terminal truncation of RAP1 results in the deregulation of telomere size, stability, and function in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 12(11), pp.5159–5173.
60. De Lange, T., 2005. Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes and Development*, 19(18), pp.2100–2110.
61. Larrivé, M., LeBel, C. & Wellinger, R.J., 2004. The generation of proper constitutive G-tails on yeast telomeres is dependent on the MRX complex. *Genes and Development*, 18(12), pp.1391–1396.
62. Lei, M., Podell, E.R. & Cech, T.R., 2004. Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nature Structural and Molecular Biology*, 11(12), pp.1223–1229.
63. Levy, D.L. & Blackburn, E.H., 2004. Counting of Rif1p and Rif2p on *Saccharomyces cerevisiae* Telomeres Regulates Telomere Length. *Society*, 24(24), pp.10857–10867.
64. Lewis, P.W., Elsaesser, S.J., Noh, K.M., Stadler, S.C. & Allis, C.D., 2010. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proceedings of the National Academy of Sciences*, 107(32), pp.14075–14080.
65. Lipps, H.J. & Rhodes, D., 2009. G-quadruplex structures: *in vivo* evidence and function. *Trends in Cell Biology*, 19(8), pp.414–422.
66. Lorenz, T.C. 2012. Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies. *J Vis Exp*, (63):e3998.
67. Lou, Z., Wei, J., Riethman, H., Baur, J.A., Voglauer, R., Shay, J.W. & Wright, W.E., 2009. Telomere length regulates ISG15 expression in human cells. *Aging*, 1(7), pp.608–621.
68. Lue, N.F., 2010. Plasticity of telomere maintenance mechanisms in yeast. *Trends in biochemical sciences*, 35(1), pp.8-17.
69. Lue, N.F., 2018. Evolving Linear Chromosomes and Telomeres: A C-Strand-Centric View. *Trends in biochemical sciences*, 43(5), pp.314-326.
70. Lundblad, V. & Blackburn, E.H., 1993. An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell*, 73(2), pp.347–360.
71. Maicher, A., Kastner, L. & Luke, B., 2012. Telomeres and disease: Enter TERRA. *RNA Biology*, 9(6), pp.843–849.
72. Maicher, A., Lockhart, A. & Luke, B., 2014. Breaking new ground: Digging into TERRA function. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1839(5), pp.387–394.

73. Makovets, S., Herskowitz, I. & Blackburn, E.H., 2004. Anatomy and Dynamics of DNA Replication Fork Movement in Yeast Telomeric Regions. *Molecular and cellular biology*, 24(9), pp.4019–4031.
74. Martín, V., Du, L.L., Rozenzhak, S. & Russell, P., 2007. Protection of telomeres by a conserved Stn1-Ten1 complex. *Proceedings of the National Academy of Sciences of the United States of America*, 104(35), pp.14038–14043.
75. Martinez, P., Thanasoula, M., Carlos, A.R., Gómez-López, G., Tejera, A.M., Schoeftner, S., Dominguez, O., Pisano, D.G., Tarsounas, M. & Blasco, M.A., 2010. Mammalian Rap1 controls telomere function and gene expression through binding to telomeric and extratelomeric sites. *Nature Cell Biology*, 12(8), pp.768–780.
76. Martínez, P. & Blasco, M.A., 2011. Telomeric and extra-telomeric roles for telomerase and the telomere-binding proteins. *Nature Reviews Cancer*, 11(3), pp.161–176.
77. Marzec, P., Armenise, C., Pérot, G., Roumelioti, F.M., Basyuk, E., Gagos, S., Chibon, F. & Déjardin, J., 2015. Nuclear-Receptor-Mediated Telomere Insertion Leads to Genome Instability in ALT Cancers. *Cell*, 160(5), pp.913–927.
78. Mazzoleni, S., Schillaci, O., Sineo, L. & Dumas, F., 2017. Distribution of Interstitial Telomeric Sequences in Primates and the Pygmy Tree Shrew (Scandentia). *Cytogenetic and Genome Research*, 151(3), pp.141–150.
79. Meena, J., Rudolph, K.L. & Günes C., 2015. Telomere Dysfunction, Chromosomal Instability and Cancer. *Recent results in cancer research*, 200, pp. 61-79.
80. Mimitou, E.P. & Symington, L.S., 2009. DNA end resection: Many nucleases make light work. *DNA Repair (Amst)*, 8, pp. 983–995.
81. Mirkin, S.M., 2007. Expandable DNA repeats and human disease. *Nature*, 447(7147), pp.932–940.
82. Mittler, G, Butter, F. & Mann, M. A SILAC-based DNA protein interaction screen that identifies candidate binding proteins to functional DNA elements. *Genome research*, 19(2), pp.284-293.
83. Mitton-Fry, R.M., Anderson, E.M., Hughes, T.R., Lundblad, V. & Wuttke, D.S., 2002. Conserved structure for single-stranded telomeric DNA recognition. *Science*, 296(5565), pp.145–147.
84. Miyake, Y., Nakamura, M., Nabetani, A., Shimamura, S., Tamura, M., Yonehara, S., Saito, M. & Ishikawa, F., 2009. RPA-like Mammalian Ctc1-Stn1-Ten1 Complex Binds to Single-Stranded DNA and Protects Telomeres Independently of the Pot1 Pathway. *Molecular Cell*, 36(2), pp.193–206.

85. Mondello, C., Pirzio, L., Azzalin, C.M. & Giulotto, E., 2000. Instability of interstitial telomeric sequences in the human genome. *Genomics*, 68(2), pp.111–117.
86. Muntoni, A. & Reddel, R.R., 2005. The first molecular details of ALT in human tumor cells. *Human Molecular Genetics*, 14(SUPPL. 2), pp.191–196.
87. Nergadze, S.G., Santagostino, M.A., Salzano, A., Mondello, C. & Giulotto, E., 2007. Contribution of telomerase RNA retrotranscription to DNA double-strand break repair during mammalian genome evolution. *Genome Biology*, 8(12), R260.
88. Niida, H., Shinkai, Y., Hande, M.P., Matsumoto, T., Takehara, S., Tachibana, M., Oshimura, M., Lansdorp, P.M. and Furuichi, Y., 2000. Telomere maintenance in telomerase-deficient mouse embryonic stem cells: characterization of an amplified telomeric DNA. *Molecular and cellular biology*, 20, pp.4115–4127.
89. Nugent, C.I., Hughes, T.R., Lue, N.F. & Lundblad, V., 1996. Cdc13p: A single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science*, 274(5285), pp.249–252.
90. O’Sullivan, R.J. & Almouzni, G., 2014. Assembly of telomeric chromatin to create ALTerNative endings. *Trends in Cell Biology*, 24(11), pp.675–685.
91. O’Sullivan, R.J., Kubicek, S., Schreiber, S.L. & Karlseder, J., 2010. Reduced histone biosynthesis and chromatin changes arising from a damage signal at telomeres. *Nature Structural and Molecular Biology*, 17(10), pp.1218–1225.
92. Paeschke, K., Simonsson T., Postberg J., Rhodes D. & Lipps H.J., 2005. Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nature Structural and Molecular Biology*, 12(10), pp.847–854.
93. Paeschke, K., Capra, J.A. & Zakian, V.A., 2011. DNA replication through G-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 DNA helicase. *Cell*, 145(5), pp.678-691.
94. Palladino, F., Laroche, T., Gilson, E., Axelrod, A., Pillus, L. & Gasser, S.M., 1993. SIR3 and SIR4 proteins are required for the positioning and integrity of yeast telomeres. *Cell*, 75(3), pp.543–555.
95. Pfeiffer, V. & Lingner, J., 2013. Replication of Telomeres and the Regulation of Telomerase. *Cold Spring Harb Perspect Biol.*, 5(5), pp.1–22.
96. Pisano, S., Galati, A. & Cacchione, S., 2008. Telomeric nucleosomes: Forgotten players at chromosome ends. *Cellular and Molecular Life Sciences*, 65(22), pp.3553–3563.

97. Porro, A., Feuerhahn, S., Reichenbach, P. & Lingner, J., 2010. Molecular Dissection of Telomeric Repeat-Containing RNA Biogenesis Unveils the Presence of Distinct and Multiple Regulatory Pathways. *Molecular and Cellular Biology*, 30(20), pp.4808–4817.
98. Price, C.M., Boltz, K.A., Chaiken, M.F., Stewart, J.A., Beilstein, M.A. & Shippen, D.E., 2010. Evolution of CST function in telomere maintenance. *Cell Cycle*, 9(16), pp.3157–3165.
99. Purohit, G., Mukherjee, A.K., Sharma, S. & Chowdhury, S., 2018. Extratelomeric binding of the telomere binding protein TRF2 at the PCGF3 promoter is G-quadruplex motif dependent. *Biochemistry*, 57(16), pp.2317–2324.
100. Ribaud, V., Ribeyre, C., Damay, P. & Shore, D., 2012. DNA-end capping by the budding yeast transcription factor and subtelomeric binding protein Tbf1. *The EMBO Journal*, 31(1), pp.138-149.
101. Rice, C. & Skordalakes, E., 2016. Structure and function of the telomeric CST complex. *Computational and Structural Biotechnology Journal*, 14, pp.161–167.
102. Robin, J.D., Ludlow, A.T., Batten, K., Gaillard, M.C., Stadler, G., Magdinier, F., Wright, W.E. & Shay, J.W., 2015. SORBS2 transcription is activated by telomere position effect-over long distance upon telomere shortening in muscle cells from patients with facioscapulohumeral dystrophy. *Genome Research*, 25(12), pp.1781–1790.
103. Rosenfeld, J.A., Wang, Z., Schones, D.E., Zhao, K., DeSalle, R. & Zhang, M.Q., 2009. Determination of enriched histone modifications in non-genic portions of the human genome. *BMC Genomics*, 10, pp.1–11.
104. Rovatsos, M., Kratochvíl, L., Altmanová, M. & Johnson Pokorná, M., 2015. Interstitial telomeric motifs in squamate reptiles: When the exceptions outnumber the rule. *PLoS ONE*, 10(8), pp.1–14.
105. Power, P., Jeffery, D., Rehman, M.A., Chatterji, A. & Yankulov, K., 2011. Subtelomeric core X and Y' elements in *S.cerevisiae* suppress extreme variations in gene silencing. *PLoS ONE*, 6(3).
106. Ruiz-Herrera, A., Nergadze, S.G., Santagostino, M. & Giulotto, E., 2009. Telomeric repeats far from the ends: Mechanisms of origin and role in evolution. *Cytogenetic and Genome Research*, 122(3–4), pp.219–228.
107. Sabourin, M., Tuzon, C.T. & Zakian, V.A., 2007. Telomerase and Tel1p preferentially associate with short telomeres in *S. cerevisiae*. *Molecular cell*, 27(4), pp.550-561.

108. Sadaie, M., Naito, T. & Ishikawa, F., 2003. Stable inheritance of telomere chromatin structure and function in the absence of telomeric repeats. *Genes and Development*, 17(18), pp.2271–2282.
109. Sánchez, J., Bianchi, M.S. & Bolzán, A.D., 2010. Relationship between heterochromatic interstitial telomeric sequences and chromosome damage induced by the radiomimetic compound streptonigrin in Chinese hamster ovary cells. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, 684(1–2), pp.90–97.
110. Saveliev, A., Everett, C., Sharpe, T., Webster, Z. & Festenstein, R., 2003. DNA triplet repeats mediate heterochromatin-protein-1-sensitive variegated gene silencing. *Nature*, 422(6934), pp.909–913.
111. Scheibe, M., Arnoult, N., Kappei, D., Buchholz, F., Decottignies, A., Butter, F. & Mann, M., 2013. Quantitative interaction screen of telomeric repeat-containing RNA reveals novel TERRA regulators. *Genome Research*, 23(12), pp.2149–2157.
112. Scherthan, H., Jerratsch, M., Li, B., Smith, S., Hultén, M., Lock, T. & de Lange, T., 2000. Mammalian meiotic telomeres: protein composition and redistribution in relation to nuclear pores. *Molecular biology of the cell*, 11(12), pp.4189–4203.
113. Sfeir, A. & de Lande, T., 2014. Removal of Shelterin Reveals the Telomere End-Protection Problem. *Science*, 593(2012), pp.593–598.
114. Shay, J.W., 2018. Telomeres and aging. *Current Opinion in Cell Biology*, 52, pp.1–7.
115. Shishkin, A.A., Voineagu, I., Matera, R., Cherng, N., Chernet, B.T., Krasilnikova, M.M., Narayanan, V., Lobachev, K.S. & Mirkin, S.M., 2009. Large-scale expansions of Friedreich's ataxia GAA repeats in yeast. *Molecular Cell*, 35(1), pp.82–92.
116. Shuman, H.A. & Silhavy, T.J., 2003. The art and design of genetic screens: *Escherichia coli*. *Nature Reviews Genetics*, 4(6), pp.419-31.
117. Silva-Sousa, R., López-Panadés, E. & Casacuberta, E., 2012. *Drosophila* telomeres: An example of co-evolution with transposable elements. *Repetitive DNA*, 7, pp.46–67.
118. Simonet, T., Zaragosi, L.E., Philippe, C., Lebrigand, K., Schouteden, C., Augereau, A., Bauwens, S., Ye, J., Santagostino, M., Giulotto, E., Magdinier, F., Horard, B., Barbry, P., Waldmann, R. & Gilson, E., 2011. The human TTAGGG repeat factors 1 and 2 bind to a subset of interstitial telomeric sequences and satellite repeats. *Cell Research*, 21(7), pp.1028–1038.
119. Stadler, G., Rahimov, F., King, O.D., Chen, J.C., Robin, J.D., Wagner, K.R., Shay, J.W., Emerson, C.P. Jr. & Wright, W.E., 2013. Telomere Position Effect (TPE) Regulates

DUX4 in Human Facioscapulohumeral Muscular Dystrophy (FSHD). *Nat Struct Mol Biol*, 20(6), pp. 671–678.

120. Stewart, J.A., Wang, F., Chaiken, M.F., Kasbek, C., Chastain, P.D. 2nd, Wright, W.E. & Price, C.M., 2012. Human CST promotes telomere duplex replication and general replication restart after fork stalling. *EMBO Journal*, 31(17), pp.3537–3549.

121. Stewart, J.A., Wang, Y., Ackerson, S.M. & Schuck, P.L., 2018. Emerging roles of CST in maintaining genome stability and human disease. *Front Biosci (Landmark Ed)*, 23, pp. 1564-1586.

122. Taddei, A., Hediger, F., Neumann, F.R. & Gasser, S.M., 2004. The Function of Nuclear Architecture: A Genetic Approach. *Annual Review of Genetics*, 38(1), pp.305–345.

123. Takahama, K., Takada, A., Tada, S., Shimizu, M., Sayama, K., Kurokawa, R. & Oyoshi, T., 2013. Regulation of telomere length by G-quadruplex telomere DNA- and TERRA-binding protein TLS/FUS. *Chemistry and Biology*, 20(3), pp.341–350.

124. Tang, J., Kan, Z.Y., Yao, Y., Wang, Q., Hao, Y.H. & Tan, Z., 2008. G-quadruplex preferentially forms at the very 3' end of vertebrate telomeric DNA. *Nucleic Acids Research*, 36(4), pp.1200–1208.

125. Teixeira, M.T., Arneric, M., Sperisen, P. & Lingner, J., 2004. Telomere length homeostasis is achieved via a switch between telomerase- extendible and -nonextendible states. *Cell*, 117(3), pp.323-335.

126. Thompson, D.A. & Stahl, F.W., 1999. Genetic control of recombination partner preference in yeast meiosis. Isolation and characterization of mutants elevated for meiotic unequal sister-chromatid recombination. *Genetics*, 153(2), pp.621-641.

127. Tommerup, H., Dousmanis, a & de Lange, T., 1994. Unusual chromatin in human telomeres. *Molecular and cellular biology*, 14(9), pp.5777–85.

128. Wang, F., Stewart, J.A., Kasbek, C., Zhao, Y., Wright, W.E. & Price, C.M., 2012. Human CST has independent functions during telomere duplex replication and C-strand fill-in. *Cell Reports*, 2(5), pp.1096–1103.

129. Wellinger, R.J. & Zakian, V.A., 2012. Everything you ever wanted to know about *Saccharomyces cerevisiae* telomeres: Beginning to end. *Genetics*, 191(4), pp.1073–1105.

130. Wood, A.M., Rendtlew Danielsen, J.M., Lucas, C.A., Rice, E.L., Scalzo, D., Shimi, T., Goldman, R.D., Smith, E.D., Le Beau, M.M. & Kosak, S.T., 2014. TRF2 and lamin A/C interact to facilitate the functional organization of chromosome ends. *Nature Communications*, 5, pp.1–9.

131. Wood, A.M., Laster, K., Rice, E.L. & Kosak, S.T., 2015. A beginning of the end: new insights into the functional organization of telomeres. *Nucleus*, 6(3), pp.14–17.
132. Xu, Y, Suzuki, Y., Ito, K. & Komiyama, M., 2010. Telomeric repeat-containing RNA structure in living cells. *Proceedings of the National Academy of Sciences*, 107(33), pp.14579–14584.
133. Xu, Y., 2011. Chemistry in human telomere biology: structure, function and targeting of telomere DNA/RNA. *Chemical Society Reviews*, 40(5), p.2719.
134. Yang, D., Xiong, Y., Kim, H., He, Q., Li, Y., Chen, R. & Songyang, Z., 2011. Human telomeric proteins occupy selective interstitial sites. *Cell Research*, 21(7), pp.1013–1027.
135. Yeung, Y.G. & Stanley, E.R., 2010. Rapid detergent removal from peptide samples with ethyl acetate for mass spectrometry analysis. *Current protocols in protein science*, 16(16.12).
136. Ye, J.Z., Donigian, J.R., van Overbeek, M., Loayza, D., Luo, Y., Krutchinsky, A.N., Chait, B.T. & de Lange, T., 2004. TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. *Journal of Biological Chemistry*, 279(45), pp.47264–47271.
137. Zenklusen, D., Vinciguerra, P., Wyss, J.C. & Stutz, F., 2002. Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. *Molecular and cellular biology*, 22(23), pp.8241-8253.
138. Zhao, Y., Sfeir, A.J., Zou, Y., Buseman, C.M., Chow, T.T., Shay, J.W. & Wright, W.E., 2009. Telomere extension occurs at most chromosome ends and is uncoupled from fill-in in human cancer cells. *Cell*, 138(3), pp.463-475.
139. Zhong, F.L., Batista, L.F., Freund, A., Pech, M.F., Venteicher, A.S. & Artandi, S.E., 2012. TPP1 OB-fold domain controls telomere maintenance by recruiting telomerase to chromosome ends. *Cell*, 150(3), pp.481–494.
140. Zou, Y., Liu, Y., Wu, X. & Shell, S.M., 2011. Functions of human replication protein A (RPA): from DNA replication to DNA damage and stress responses. *J Cell Physiol.*, 208(2), pp.267–273.