

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
НАУЧНОЕ УЧРЕЖДЕНИЕ  
«ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР  
ИНСТИТУТ ЦИТОЛОГИИ И ГЕНЕТИКИ  
СИБИРСКОГО ОТДЕЛЕНИЯ РОССИЙСКОЙ АКАДЕМИИ  
НАУК»

На правах рукописи

Аркова Ольга Владимировна

**Изучение потенциально функционально значимых  
полиморфизмов ТАТА-боксов  
промоторов генов липидного метаболизма человека**

03.02.07 – генетика

Диссертация на соискание учёной степени

кандидата биологических наук

Научный руководитель

к.б.н. Савинкова Л.К.

Новосибирск – 2017

## Оглавление

Введение .....	6
1. Обзор литературы .....	15
1.1. Характеристика транскрипционного комплекса РНК-полимеразы II.....	15
1.2. Базальные транскрипционные факторы II класса .....	21
1.3. ТАТА-боксы промоторов генов эукариот.....	24
1.4. Взаимодействие ТСБ с ТАТА-боксами.....	29
1.5. Однонуклеотидные полиморфизмы ТАТА-боксов и их ассоциация с наследственными заболеваниями человека .....	37
1.6. Формирование электронной библиотеки и компьютерный анализ неаннотированных однонуклеотидных полиморфизмов ТАТА-боксов промоторов генов липидного гомеостаза человека .....	57
2. Материалы и методы.....	62
2.1. Формирование библиотеки неаннотированных SNPs ТАТА-боксов промоторов генов липидного метаболизма человека.....	62
2.2. Нарботка, выделение и очистка ТСБ.....	63
Выделение ТАТА-связывающего белка человека .....	64
Электрофорез по Лэммли .....	68
Анализ активности белка.....	69
2.3. Получение меченых <sup>32</sup> P олигодезоксирибонуклеотидов.....	69
2.4. Определение кинетических и термодинамических характеристик комплексов ТСБ/ОДН методом задержки ДНК в геле .....	72

2.5. Определение кинетических констант комплексов методом флуоресцентной спектроскопии в режиме реального времени.....	75
2.6. Разработка плазмидных конструкций для определения активности репортерного гена <i>LUC</i> .....	80
Очистка образцов ДНК с помощью гель-фильтрации на Sephadex G100.....	85
Получение репортерных конструкций.....	85
Выделение геномной ДНК .....	86
Получение фрагментов ДНК с использованием ПЦР .....	86
Гидролиз и лигирование фрагментов ДНК.....	87
Приготовление электрокомпетентных клеток <i>E.coli</i> .....	88
Трансформация клеток <i>E.coli</i> .....	88
Подтверждение наличия клонированного фрагмента ДНК с помощью секвенирования.....	90
Очистка плазмидной ДНК центрифугированием в двуступенчатом градиенте хлористого цезия .....	94
Трансформация клеточных линий и измерение активности люциферазы.....	94
3. Результаты и обсуждение.....	96
3.1. Определение кинетических характеристик комплексов ТСБ/ТАТА анцестральных и минорных аллелей с помощью метода EMSA.....	106
3.2. Анализ термодинамических характеристик комплексов ТСБ/ТАТА анцестральных и минорных аллелей .....	115

3.3. Определение в режиме реального времени кинетических характеристик комплексов ТСБ с ТАТА-содержащими олигонуклеотидами, методами остановленной струи и резонансного переноса энергии.....	120
3.4. Изучение влияния анцестральных и минорных аллелей генов LEP и CYP2A6, на экспрессию репортерного гена LUC.....	141
3.5. Анализ возможного фенотипического проявления SNPs в ТАТА – боксах промоторов генов .....	143
Заключение.....	144
Благодарности.....	149
Список литературы.....	150
Приложение 1.....	182
Приложение 2.....	184
Приложение 3.....	185
Приложение 4.....	186
Приложение 5.....	187

## Список сокращений, принятых в работе

SNP - однонуклеотидный полиморфизм (single nucleotide polymorphism)

ТСБ - ТАТА-связывающий белок

PolII - РНК полимеразы II

ПААГ – полиакриламидный гель

TF- транскрипционный фактор

EMSA – метод задержки ДНК в геле (electromobility shift assay)

ФРПЭ – Фёрстеровский резонансный перенос энергии (FRET)

пн - пара нуклеотидов

НК - нуклеиновая кислота

ОДН – олигодезоксирибонуклеотид, содержащий ТАТА-бокс

BRE – элемент ДНК, узнающий транскрипционный фактор ПВ (TFПВ-  
recognition element)

Inr – инициаторный элемент (Initiator)

MTE – элемент «мотив десяти» (Motif ten element)

DPE – нижний промоторный элемент (downstream promoter element)

DCE – нижний кор-элемент (downstream core element)

XCPE1 – X «кор»-промоторный элемент 1 (X core promoter element 1)

## Введение

**Актуальность.** Выяснение молекулярных механизмов генетической предрасположенности к различным заболеваниям человека является одной из фундаментальных проблем генетики, молекулярной генетики и медицины. Одним из наиболее активно развивающихся направлений в этой области является выявление однонуклеотидных полиморфизмов, SNP, кандидатных генов, ассоциированных с тем или иным наследственным заболеванием, и установление связи между определенным вариантом последовательности ДНК и патологией. Оказалось, что кодирующие белки области генома, SNPs в которых приводят к аминокислотной замене и нарушению функции белка, гораздо более изучены, чем регуляторные области, SNPs в которых могут изменять сайты связывания транскрипционных факторов, следствием чего может быть аллель-специфическая экспрессия генов, различия в чувствительности к факторам внутренней и внешней среды, разная предрасположенность к заболеваниям и чувствительность к терапии и т. п. В связи с этим выявление SNPs регуляторных районов генов, в частности, промоторов, и изучение молекулярных механизмов, посредством которых SNPs сайтов связывания транскрипционных факторов могут оказывать влияние на фенотипические признаки, является актуальным и перспективным направлением.

Осуществление проекта «1000 геномов» выявило десятки миллионов SNPs в регуляторных районах генов, что привело к обострению проблемы их аннотации – определению фенотипических проявлений и повышенной предрасположенности к тем или иным наследственным заболеваниям. Очевидно, что без компьютерного анализа такого огромного количества данных не обойтись. Он ускорит поиск биомедицинских маркеров, что будет большим ресурсом в биомедицинских исследо-

ваниях и разработке лекарственных препаратов. В последнее время вопросам аннотации SNPs уделяется все больше внимания, но значительного прорыва не происходит, так как компьютерным методам не хватает надежной экспериментальной верификации, чтобы предотвратить очень дорогостоящую во всех отношениях проверку ложно-положительных прогнозов о связи SNPs с заболеваниями. Важность и актуальность проблемы явилась причиной для проведения нами комплексного компьютерно - экспериментального изучения однонуклеотидных полиморфизмов ТАТА-боксов промоторов генов человека.

Специфические последовательности промоторной ДНК являются кодом, который во многом определяет где, когда и с какой эффективностью будет транскрибироваться и затем экспрессироваться ген. ТАТА-бокс является одним из наиболее изученных промоторных элементов. От последовательности ТАТА-бокса зависит сродство к нему ТАТА-связывающего белка (ТСБ), в составе базального фактора транскрипции TFIID или самостоятельно связывающегося со своим мишеневым участком и запускающим сборку транскрипционного комплекса на ТАТА-содержащих промоторах II класса. Во многих работах с помощью молекулярно-биологических и клинических методов показано, что SNPs в ТАТА-боксе и окружающих его последовательностях ассоциируются с повышенным риском возникновения сложных наследственных заболеваний: сердечно-сосудистых, онкологических, артрита, рассеянного склероза и др. Иногда SNPs являются причиной моногенных заболеваний, таких синдром Жильбера, Коппок-подобная катаракта, бэталассемии различной тяжести и др.

Ранее в секторе молекулярно-генетических механизмов белок-нуклеиновых взаимодействий ИЦиГ были получены экспериментальные данные о зависимости

аффинности ТСБ к ТАТА-боксам некоторых генов эукариот от последовательностей ТАТА-боксов (Савинкова и др., 2007), которые были проанализированы в лаборатории теоретической генетики ИЦиГ СО РАН под руководством Пономаренко М.П. В результате анализа в 2008 г. была разработана модель и уравнение пошагового связывания ТСБ с ТАТА-боксом (Пономаренко и др., 2008), позволяющее прогнозировать сродство ТСБ/ТАТА, основанное на анализе последовательности ДНК ТАТА-боксов. Затем была собрана электронная коллекция SNPs ТАТА-боксов (с фланкирующими последовательностями), ассоциированных с наследственными заболеваниями, которые были подтверждены клинически и молекулярно-биологически, и с помощью разработанного уравнения сделаны прогнозы изменения сродства ТСБ/ТАТА. А в нашем секторе проведена экспериментальная верификация сделанных прогнозов. Эксперименты показали высокую корреляцию экспериментальных результатов с прогнозами, сделанными *in silico*. Коэффициент линейной корреляции,  $r = 0.822$  при  $\alpha < 10^{-7}$  (Savinkova et al., 2013).

Полученные результаты позволили нам приступить к решению следующей задачи - компьютерному формированию библиотеки неаннотированных SNPs ТАТА-боксов промоторов генов человека, отобранных из доступных баз данных, анализу *in silico* последовательностей ДНК ТАТА-боксов с помощью разработанного на основании уравнения пошагового связывания (Пономаренко и др., 2008) Web-сервиса SNP\_TATA Comparator (Расказов и др., 2013), отбору наиболее подходящих примеров, и затем экспериментальному подтверждению *in vitro* и *ex vivo* (на культуре клеток) сделанных прогнозов *in silico* по влиянию SNPs на сродство ТСБ/ТАТА.



**Цель работы** заключалась в анализе *in silico*, *in vitro* и *ex vivo* неаннотированных SNPs ТАТА-боксов промоторов генов человека для выявления потенциально функционально значимых SNPs.

Для достижения цели были поставлены следующие **задачи**:

1. Формирование электронной библиотеки и компьютерный анализ неаннотированных однонуклеотидных полиморфизмов ТАТА-боксов промоторов генов липидного метаболизма человека;
2. Определение кинетических и термодинамических характеристик комплексов ТСБ/ТАТА анцестральных и минорных аллелей с помощью метода EMSA;
3. Анализ термодинамических характеристик комплексов ТСБ/ТАТА анцестральных и минорных аллелей;
4. Определение в режиме реального времени кинетических характеристик комплексов ТСБ/ТАТА методами остановленной струи и резонансного переноса энергии;
5. Получение плазмидных конструкций с анцестральными аллелями и минорными, SNP-содержащими ТАТА-боксами промоторов генов *LEP* и *CYP2A6*, и изучение их влияния на экспрессию репортерного гена *LUC*.
6. Анализ возможного фенотипического проявления SNPs в ТАТА-боксах промоторов генов.

### **Научная новизна**

Впервые в мире проведено компьютерно-экспериментальное исследование, включающее поиск и экстракцию из баз данных неаннотированных SNPs ТАТА-

боксов промоторов генов человека, компьютерный анализ последовательностей с использованием разработанного в ИЦиГ СО РАН Web сервиса SNP\_TATA\_Comparator и экспериментальную верификацию *in vitro* и *ex vivo* на культуре клеток. Экспериментальная верификация *in vitro* с использованием метода задержки ДНК в геле, метода остановленной струи и флуоресцентного резонансного переноса энергии, и *ex vivo* на культуре клеток показала высокую корреляцию опыт/прогноз ( $r=0.89$ ). В работе впервые получены кинетические характеристики комплексов ТСБ с олигодезоксинуклеотидами, идентичными ТАТА-боксам с фланкирующими нуклеотидами анцестральных и минорных аллелей ряда генов липидного метаболизма человека, которые по прогнозам могут быть ассоциированы с повышенным риском возникновения нарушения липидного метаболизма.

### **Основные положения, выносимые на защиту**

1. Экспериментальная верификация показала эффективность использования Web сервиса SNP\_TATA\_Comparator для предсказания новых потенциально функционально значимых мутаций ТАТА-боксов.
2. Мутации ТАТА-боксов, изменяющие сродство ТАТА-связывающего белка, влияют на экспрессию генов, находящихся под регуляцией этих ТАТА-боксов содержащих промоторов.
3. Формирование комплексов ТАТА-связывающего белка с ТАТА-боксами в режиме реального времени происходит как по прямому пути, так и через промежуточные стадии, количество которых возрастает с увеличением концентрации ТАТА-связывающего белка и длины олигонуклеотида.

## **Теоретическая и практическая значимость исследования**

Теоретическая значимость работы заключается в том, что экспериментально доказана возможность использования разработанного компьютерного метода SNP\_TATA\_Comparator для предсказания потенциально значимых мутаций ТАТА-боксов промоторов генов человека. Теоретическое значение работы также заключается в том, что определен механизм реализации мутационных изменений ТАТА-боксов реальных промоторов генов липидного метаболизма человека.

Практическая значимость исследований заключается в том, что полученные данные могут использоваться для доклинической проверки врачами кандидатных SNP-маркеров ТАТА-боксов в предиктивно-превентивной персонализированной медицине.

## **Апробация работы**

По материалам диссертации опубликовано 10 статей в журналах из перечня ВАК. Результаты работы были представлены на научных конференциях: «BGRS-2014» (Новосибирск, 2014), Актуальные проблемы вычислительной и прикладной математики 2015 (АМСА 2015) (Новосибирск, 2015), 7 съезд Российского общества медицинских генетиков (Санкт-Петербург, 2015), «BGRS-2016» (Новосибирск, 2016).

## **Основные публикации**

1. Ponomarenko M., Rasskazov D., Chadaeva I., Sharypova E., Ponomarenko P., Arkova O., Kashina E., Ivanisenko N., Zhechev D., Savinkova L., Kolchanov N. SNP\_TATA\_Comparator: genomewide landmarks for preventive personalized medicine.

// Front Biosci (Schol Ed). 2017. V. 9. N. 2. P.276-306. PMID: 28410120. DOI: 10.2741/S488.

2. Arkova O., Kuznetsov N., Fedorova O., Savinkova L. A real-time study of the interaction of TBP with a TATA box-containing duplex identical to an ancestral or minor allele of human gene LEP or TPI. // Journal of Biomolecular Structure and Dynamics, 2016; 25:1-20. IF = 2.300.

3. Ponomarenko M.P., Arkova O., Rasskazov D., Ponomarenko P., Savinkova L., Kolchanov N. Candidate SNP Markers of Gender-Biased Autoimmune Complications of Monogenic Diseases Are Predicted by a Significant Change in the Affinity of TATA-Binding Protein for Human Gene Promoters. // Front Immunol. 2016; 7:130. IF = 5.695.

4. Chadaeva I.V., Ponomarenko M.P., Rasskazov D.A., Sharypova E.B., Kashina E.V., Matveeva M.Y., Arshinova T.V., Ponomarenko P.M., Arkova O.V., Bondar N.P., Savinkova L.K., Kolchanov N.A. Candidate SNP markers of aggressiveness-related complications and comorbidities of genetic diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. // BMC Genomics. 2016;17 (Suppl 14): S3, IF = 3.867.

5. Arkova O.V., Ponomarenko M.P., Rasskazov D.A., Drachkova I.A., Arshinova T.V., Ponomarenko P.M., Savinkova L.K., Kolchanov N.A. Obesity-related known and candidate SNP markers can significantly change affinity of TATA-binding protein for human gene promoters. // BMC Genomics. 2015;16 Suppl 13. IF = 3.867.

6. Ponomarenko M., Rasskazov D., Arkova O., Ponomarenko P., Suslov V., Savinkova L., Kolchanov N. How to Use SNP\_TATA\_Comparator to Find a Significant Change in Gene Expression Caused by the Regulatory SNP of This Gene's Promoter via

a Change in Affinity of the TATA-Binding Protein for This Promoter. Suslov V, Savinkova L, Kolchanov N. // Biomed Res Int. 2015; 5:359835. IF = 1.579.

7. Arkova O.V., Kuznetsov N.A., Fedorova O.S., Kolchanov N.A., Savinkova L.K. Real Time Interaction between TBP and the TATA box of the Human Triosephosphate Isomerase Gene promoter in the Norm and Pathology. // Acta Naturae. 2014;6(2):36-40, IF = 1.777.

8. Турнаев И.И., Рассказов Д.А., Аркова О.В., Пономаренко М.П., Пономаренко П.М., Савинкова Л.К., Колчанов Н.А. Гипотетические SNP-маркеры, значимо изменяющие оценки сродства ТАТА-связывающего белка к промоторам онкогенов VEGFA, ERBB2, IGF1R, FLT1 KDR, MET – мишеней для химиотерапии. // Мол. биол. 2016. 50(1), 161-73, РИНЦ=0.786.

9. Аркова О.В., Драчкова И.А., Аршинова Т.В., Рассказов Д.А., Суслов В.В., Пономаренко П.М., Пономаренко М.П., Колчанов Н.А., Савинкова Л.К. Прогноз и верификация влияния SNP rs367781716 на взаимодействие ТАТА-связывающего белка с промотором гена ABCA9 человека. // Вавилов. журн. генет. и селекции. 2015, 19(6), 675-681.

10. Аркова О.В., Пономаренко М.П., Аршинова Т.В., Савинкова Л.К. Механизм влияния полиморфизмов кор-промоторов генов пищевого поведения человека на взаимодействие с ТАТА-связывающим белком. Медицинская генетика, 2015, т. 14, №2, с. 10-11.

## **Вклад автора**

Основные результаты получены автором самостоятельно. Автор принимал личное участие в планировании, проведении и обсуждении всех экспериментов, по результатам которых написана диссертация.

Биоинформационный анализ нуклеотидных последовательностей проводился совместно с к.б.н. М.П. Пономаренко. Выращивание химически компетентных клеток и их трансформация проводились совместно с И.А. Драчковой. Оценка взаимодействия ТСБ с ТАТА-боксами с помощью FRET проводилась совместно с к.х.н. Н.А. Кузнецовым. Плазмидные конструкции получены совместно с Е.Б. Шарыповой. Результаты по влиянию плазмидных конструкций на активность гена люциферазы в культуре клеток получены Е.В. Кашиной.

## **Структура и объем диссертации**

Работа состоит из списка сокращений, введения, обзора литературы, материалов и методов, результатов, обсуждения и приложения. Диссертация включает также заключение, выводы и список литературы. Работа изложена на 188 страницах, содержит 21 рисунок, содержит 10 таблиц и 4 схемы, а также 5 приложений. Список цитируемой литературы включает 243 ссылки.

## **1. Обзор литературы**

### **1.1. Характеристика транскрипционного комплекса РНК-полимеразы II**

Нормальное функционирование многоклеточных организмов зависит от правильного сочетания регулирующих экспрессию генов событий, которые обеспечивают необходимый уровень экспрессии в нужное время, в нужном месте и на нужном уровне. В связи с бесчисленными взаимодействиями генетических факторов и факторов внутренней и внешней среды, описание заболеваний является большой проблемой для биомедицинских исследований. Классические генетические заболевания, как правило, возникают из-за изменения в одном гене или аллеле и называются моногенными. Заключаются такие генетические изменения, в основном, в замене одного нуклеотида на другой, вставки нуклеотида или его делеции. Почти все моногенные заболевания наследуются согласно классическим законам Менделя. Примером моногенных заболеваний, возникших в результате однонуклеотидной замены, являются серповидно-клеточная анемия, талассемии, фенилкетонурия, врожденная катаракта и др. Комплексные заболевания, в отличие от моногенных, порой трудно описать и поставить диагноз, так как они возникают в результате сочетанного взаимодействия множества факторов. Известно, что многие комплексные (или многофакторные) заболевания, такие как сердечно-сосудистые и гипертония, болезнь Альцгеймера, рак, сахарный диабет второго типа и др. с повышенной частотой развиваются среди близких родственников. Это свидетельствует о том, что генетические факторы играют главную роль в патогенезе болезней, помимо факторов внутренней и внешней среды и индивидуального поведения.

Использование биоинформатических методов и подходов с возможностью анализа больших объемов данных способствует выявлению и более полной характеристике комплексных заболеваний. Анализ *in silico* позволяет проследить и выбрать из множества факторов внешней среды те, которые влияют на фенотип заболевания, на чувствительность к терапии, выяснить взаимоотношения между ними и дать практические рекомендации для медицины.

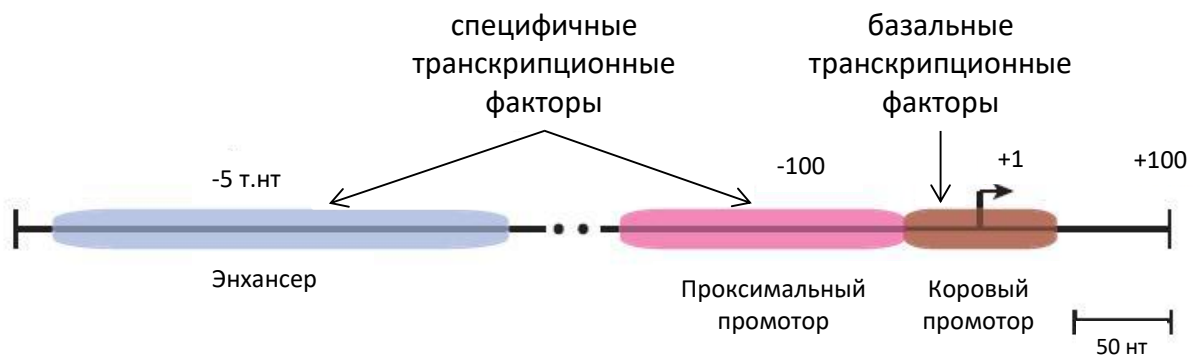
Большинство регуляторных событий происходит на уровне транскрипции и опосредуется взаимодействием транскрипционных факторов с регуляторными участками ДНК, которые сосредоточены в промоторной области гена, расположенной слева от стартового участка транскрипции. Экспрессия большинства генов различается между разными клетками и тканями, на разных стадиях развития организма и при различных физиологических состояниях.

Известно, что РНК-полимераза II (Pol II) работает со всеми белок-кодирующими генами и генами большинства малых ядерных РНК, микро РНК и 5S РНК. Регуляторные элементы генов — это специфические цис-действующие последовательности ДНК, которые узнаются транс-действующими транскрипционными факторами. При этом у многоклеточных эукариот цис-регуляторные элементы обычно делят в зависимости от расстояния от старта транскрипции на «ближние» или промоторные элементы, и дистальные, т.е. действующие на больших расстояниях (Рис. 1). Под промотором гена обычно предполагают совокупность цис-регуляторных элементов, расположенных только вблизи сайта старта транскрипции TSS (transcription start site), и необходимых либо для инициации транскрипции, либо для увеличения частоты инициации транскрипции. Под промоторной областью гена обычно подразумевают набор из корового промотора и проксимальных



промоторных элементов (Allison, 2007) (Рис. 1), расположение которых может сильно различаться. На сегодняшний день под понятием коровый промотор РНК полимеразы II принято учитывать два дополняющих друг друга определения – что это либо самый минимальный участок непрерывной последовательности ДНК, который достаточен для точной инициации транскрипции этой полимеразой (Butler, Kadonaga, 2002; Juven-Gershon et al., 2006), либо это область ДНК, расположенная на расстоянии ~ 40 нуклеотидов слева (вверх) и справа (вниз) от старта транскрипции гена, на которой сконцентрировано переменное количество регуляторных участков (Baumann et al., 2010), таких как ТАТА-бокс, BRE (TFIIВ-recognition element), Inr (Initiator), MTE (Motif ten element), DPE (downstream promoter element), DCE (downstream core element) и XCPE1 (X core promoter element 1) (Juven-Gershon, Kadonaga, 2010).

Исторически так сложилось, что под коровым промотором предполагался транскрипционный элемент, функционирующий повсеместно и взаимодействующий в разных комбинациях с различными энхансерами, находящимися в разнообразном окружении. На сегодняшний день под промоторами подразумевают структурно и функционально разнообразные транскрипционные регуляторные элементы (Juven-Gershon and Kadonaga, 2010), что актуально при анализе активности генов, когда необходимо учитывать наличие всех специфических компонентов корового промотора.



**Рисунок 1.** Цис-действующие регуляторные последовательности ДНК у эукариот. На рисунке указана организация и расположение дистального регуляторного элемента (энхансера), а также проксимального и корового промотора (Ohler, Wassarman, 2010). На рисунке схематично изображено, что с проксимальными промоторными элементами и дистальными регуляторными областями взаимодействуют сиквенс-специфичные транскрипционные факторы, а с коровым промотором взаимодействуют базальные транскрипционные факторы.

По разным источникам в геноме многоклеточных эукариот на белки, участвующие в регуляции транскрипции, приходится от 5 до 10% кодирующих последовательностей. Таких белков в геноме дрожжей закодировано около 300, у *Drosophila* 1000, а у человека количество белков может достигать 3000. Белки, участвующие в регуляции транскрипции, делят на три основных класса:

1. **Белки аппарата базальной транскрипции.** Так как сама по себе полимеразы не может узнавать коровый промотор, то для процесса узнавания требуется наличие общих, или базальных, факторов транскрипции (GTFs - general transcription factors или basal transcription factors), которые функционируют не универсально во всех коровых промоторах (Thomas and Chiang, 2006). Считается, что присутствие общих факторов транскрипции (GTFs) необходимо и может быть до-

статочным условием для обеспечения точной инициации транскрипции *in vitro* (Thomas and Chiang, 2006). Факторы базальной транскрипции включают TFIIA (transcription factor, RNA polymerase II, A), TFIIВ, TFIID, TFIIЕ, TFIIF и TFIIH. TFIID представляет собой мультибелковый комплекс, состоящий из ТСБ (или ТВР - TATA-box binding protein) и ассоциированных с ним факторов (ТВР-associated factors, TAFs). В начале транскрипции на коровом промоторе собирается преинициаторный комплекс транскрипции (preinitiation complex, PIC), состоящий из базальных факторов и РНК-полимеразы II. После окончания сборки PIC, при наличии субстратов (четырёх нуклеозидтрифосфатов) происходит процесс инициации транскрипции, после чего полимеразы переходит в стадию элонгации, удлиняя синтезированную молекулу РНК.

2. **Транскрипционные факторы:** это ДНК-связывающие регуляторные белки, являющиеся ДНК-специфическими и работающие путем связывания с промоторами генов или дистальными регуляторными элементами, и опосредующие транскрипционную активацию или репрессию, специфичную для генов. Активаторы, как правило, стимулируют формирование PIC, но могут также и увеличивать скорость элонгации транскрипции, и обеспечивать множественные раунды транскрипции, а также направлять модификации хроматина (Weake and Workman, 2010).

3. **Транскрипционные коактиваторы и корепрессоры:** К этой группе относят белки, изменяющие транскрипционную активность генов без прямого взаимодействия с ДНК, а только за счет белок-белковых взаимодействий. Коактиваторы и корепрессоры могут привлекать белки с различными энзиматическими активностями в отношении структуры хроматина, но также они и сами могут выполнять эту функцию. Такие белки регулируют связывание и движение транскрипционной ма-

шины через хроматин и состоят из комплексов белков ремоделинга и модификации хроматина (Allison, 2007). Для транскрипции, осуществляемой Pol II, наиболее необходимым и значимым коактиватором является мультибелковый комплекс Mediator (Kim et al., 1994), взаимодействующий с С-концевым доменом Pol II.

Базальный транскрипционный комплекс состоит из РНК-полимеразы II и базальных факторов транскрипции, и сборка базального транскрипционного комплекса на ТАТА-содержащих промоторах начинается с узнавания и связывания ТАТА-бокса ТАТА-связывающим белком (ТСБ). Такое связывание приводит к изгибанию ДНК, причем, чем сильнее изгиб ДНК, тем более активной будет последующая транскрипция (Starr et al., 1995).

После образования комплекса «ТСБ/ТАТА» к нему присоединяется базальный фактор транскрипции TFIIA. Он укрепляет комплекс «ТСБ/ТАТА», к которому затем присоединяется TFIIВ, связываясь с участками, прилегающими к ТАТА-боксу слева и справа. Присоединение TFIIВ регулирует транскрипционную активность (Kim et al., 2005). Затем к субкомплексу «ТАТА-ТСБ-TFIIA-TFIIВ» присоединяется РНК-полимераза II, чаще всего находящаяся в комплексе с TFIIF. Завершает сборку базального комплекса присоединение к образовавшемуся ансамблю базальных факторов TFIIЕ и TFIIН, при добавлении субстратов ведущее к эффективной инициации транскрипции – это последовательный путь сборки PIC (Lagrange et al., 1998). Следует отметить, что путь сборки преинициаторного комплекса до сих пор точно не ясен. Недавно с помощью метода флуоресцентной микроскопии одной молекулы показано (Horn et al., 2016), что активность транскрипции зависит от базальных транскрипционных факторов – ТСБ, IIВ и IIF, которые могут образовывать комплексы между собой. Авторы работы (Horn et al., 2016) показали, что IIВ

может присоединяться к подкомплексу (ТСБ+ДНК) только до добавления ПФ, который препятствует его контакту с подкомплексом. Без труда образуются комплексы между Pol II и ДНК-дуплексами, но для транскрипционной активности этих комплексов необходимы базальные факторы ТСБ, ПВ и частично ПФ.

Известен альтернативный путь сборки – голоферментный. Он заключается в том, что к промотору присоединяется большой голоферментный комплекс, включающий гипофосфорилированную форму РНК-полимеразы II, преассоциированную с переменным количеством факторов, чаще всего с базальными факторами транскрипции TFIIIF, TFIIIE и TFIIH, а также с компонентами мультисубъединичного медиаторного комплекса, хроматин-ремодулирующими белками, ферментами репарации ДНК и процессинга мРНК (Buratowski *et al.*, 1989).

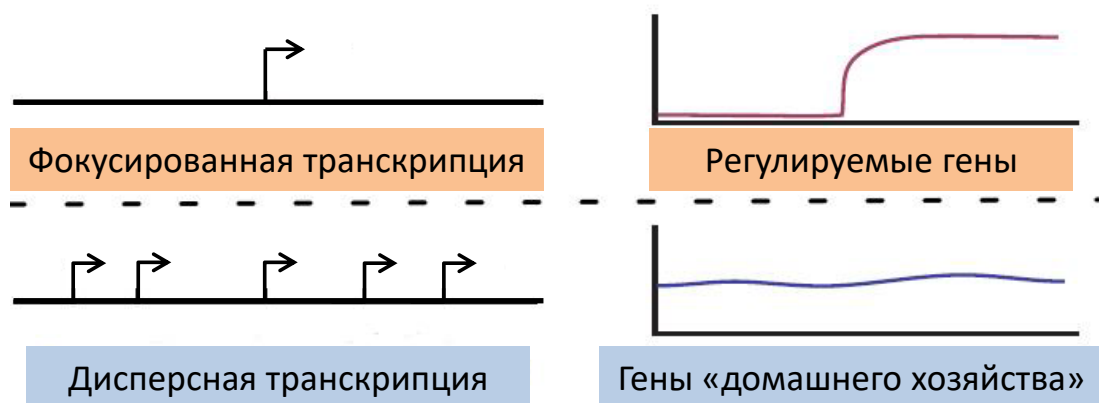
## **1.2. Базальные транскрипционные факторы II класса**

Раньше считалось, что строение базального транскрипционного комплекса инвариантно для разных типов клеток, а сейчас показано существование семейства факторов, родственных ТСБ - TRF (TBP-related factors), и семейства факторов, ассоциированных с ТСБ – TAFs (Malik, Roeder, 2000). Если с ТАТА-направляемых промоторов транскрипцию инициирует РНК-полимераза II в ансамбле с базальными факторами TFIIA, TFIIIB, TFIIID, TFIIIF, TFIIIE и TFIIH, то с DPE-направляемых промоторов транскрипция не может инициироваться этим набором базальных факторов (Maldonado, 1996), так как ТСБ инициирует ТАТА-зависимую и репрессирует DPE-зависимую транскрипцию (Tora, 2010). Поэтому для инициации транскрипции с DPE-направляемых промоторов необходимы NC2 (negative cofactor 2) и Mot1

(АТРаза, которая удаляет ТСБ с ДНК), которые блокируют функцию ТСБ и инициируют DPE-зависимую транскрипцию (Toza, 2010). DPE вместе с Inr работает как заменитель ТАТА-бокса (Veenstra, Wollfe, 2001) и как единая кор-промоторная единица. Таким образом, регуляция транскрипции белок-кодирующих генов обеспечивается не только разнообразием энхансеров, но и разным составом кор-промоторных элементов (Smale, Kadonaga, 2003), существованием ткане- и ген-специфических вариантов базальной транскрипционной машины РНК-полимеразы II.

Транскрипция с корового промотора осуществляется аппаратом базальной транскрипции. Существование множества типов коровых промоторов приводит к наличию разных механизмов базальной транскрипции (Kadonaga, 2012). Так как транскрипция любого белок-кодирующего гена РНК-полимеразой II начинается со сборки базального транскрипционного комплекса на кор-промоторе, то исследование паттернов инициации транскрипции обнаружило два разных способа инициации транскрипции – с одного сильного стартового участка (способ инициации транскрипции называют фокусированным, когда транскрипция начинается с единственного нуклеотида или в пределах кластера нескольких нуклеотидов) или с нескольких слабых сайтов старта транскрипции (диспергированный способ инициации транскрипции), расположенных в области 50-100 нуклеотидов (Smale, Kadonaga, 2003). Регулируемые гены транскрибируются, в основном, с фокусированных промоторов, а конститутивные - с диспергированных. Хотя встречаются и комбинированные промоторы, имеющие много стартовых участков, из которых один наиболее сильный (Juven-Gershon, Kadonaga, 2010).

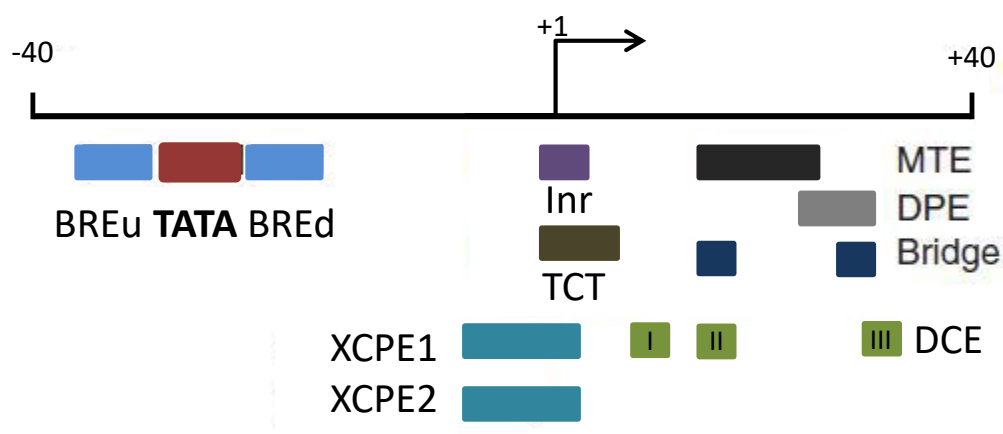
Наиболее древним типом промоторов считаются фокусированные промоторы, которые часто содержат ТАТА-бокс (Juven-Gershon et al., 2008) и являются консервативными от архей до позвоночных. Характерной чертой дисперсных промоторов является сверхпредставленность CpG-островков. Под контролем дисперсных промоторов находятся большинство генов высших эукариот. Например, у позвоночных дисперсные промоторы имеют до 70% генов. Считается, что у генов с регулируемой экспрессией (индуцируемых) есть фокусированные промоторы, а дисперсные промоторы в основном используются генами с конститутивной экспрессией, т. е. гены «домашнего хозяйства» (Рис. 2).



**Рисунок 2.** Фокусированный и дисперсный способ инициации транскрипции (Kadonaga, 2012).

Большая часть исследований механизмов транскрипции РНК-полимеразой II была выполнена на фокусированных промоторах, хотя доля фокусированных промоторов у позвоночных невелика. Это связано с активно развивающимся на сегодняшний день направлением по изучению регулируемых генов, с которыми и ассоциированы фокусированные промоторы, что является биологически значимым. Исследования фокусированных промоторов привело к открытию таких регуляторных

мотивов, как ТАТА-боксы, BREu (upstream TFIIIB recognition element), Inr (initiator), MTE (motif ten element), DPE (downstream promoter element), DCE (downstream core element) и XCPE (Xcore promoter element) (Рис. 3). Дисперсные промоторы обычно лишены мотивов BRE, ТАТА, DPE и MTE (Sandelin et al., 2007; Carninci et al., 2006) в отличие от фокусированных. Скорее всего, существуют фундаментальные различия в механизмах транскрипции с фокусированных и дисперсных промоторов.



**Рисунок 3.** Схема приблизительного расположения ряда коровых промоторных элементов в промоторах (ТАТА-боксы, BREu, Inr, MTE, DPE, DCE, TCT, Bridge и XCPE) (Kadonaga, 2012) в генах, транскрибируемых РНК-полимеразой II.

### 1.3. ТАТА-боксы промоторов генов эукариот

ТАТА-боксы обычно расположены на расстоянии 25-30 нуклеотидов перед стартом транскрипции (Gannon et al., 1979). Однако позднее было обнаружено, что точное положение ТАТА-боксов может варьировать в пределах от 28 до 70 пн перед стартом транскрипции, но наблюдается сильное предпочтение фиксированному месту - 30-31 пн от TSS (Ponjavic et al, 2006; Hahn, 2004). У дрожжей *Saccharomyces cerevisiae* ТАТА-боксы расположены между 70 и 120 пн от TSS (Yang et al., 2007).



ТАТА-бокс (или Goldberg-Hogness box) – это АТ-богатая последовательность длиной 8 пар нуклеотидов, которая является наиболее изученным кор-промоторным элементом, который был первым обнаружен исследователями при изучении кластера гистоновых генов у *Drosophila melanogaster* (Lifton et al., 1978; Goldberg, 1979; Proudfoot, 1979). При этом ТАТА-содержащие промоторы составляют лишь 10–16% генов человека, считываемых РНК-полимеразой II. Из них только около 30% промоторов имеют канонический ТАТА-бокс, что свидетельствует о вариабельности канонической последовательности ТАТА-бокса – ТА-ТАААА – природных промоторов (Hahn et al., 1989). ТАТА-бокс - это богатый А-Т парами элемент с консенсусной последовательностью ТАТАWАAR (W = А или Т, R = А или G) у многоклеточных организмов. Показано, что в контексте разных промоторов одинаковые мутации ТАТА-бокса оказывают различное влияние на активность промоторов (Hsu et al., 2008). Несмотря на такую вариабельность ТАТА-боксов, некоторые ТАТА-содержащие промоторы могут быть очень чувствительны к мутациям ТАТА-боксов. Подтверждением этому являются молекулярно-генетически и клинически подтвержденные однонуклеотидные полиморфные варианты ТАТА-боксов промоторов генов человека, ассоциированные с предрасположенностью к наследственным заболеваниям (Савинкова и др. 2009).

В 1981 году экспериментально было показано, что наличие ТАТА-бокса является необходимым и достаточным условием для специфической инициации РoI II-зависимой транскрипции *in vitro* (Sassone-Corsi et al., 1981). А однонуклеотидные замены (SNP's) в ТАТА-боксах приводят к существенному падению уровня транскрипции гена (Wasylyk et al., 1980).

ТСБ, в составе субъединицы TFIID, распознает ТАТА-боксы и связывается с ним. ТСБ и ТАТА-боксы консервативны от архей до человека (Hernandez, 1993; Reeve, 2003), ТАТА-боксы есть и у растений (впервые обнаружен Molina and Grotewold, 2005). ТАТА-боксы могут функционировать в отсутствие Inr, BRE и DPE мотивов.

Известно о допустимых изменениях в составе последовательности канонического ТАТА-боксы – ТАТАААА. Считается, что не более 20% промоторов генов II класса содержащих ТАТА-боксы имеют именно канонический ТАТА-боксы. Тем не менее, могут быть очень чувствительны к мутациям ТАТА-элементов некоторые ТАТА-содержащие промоторы. Интересно, что одни и те же мутации ТАТА-боксы в контексте разных промоторов оказывают на активность промоторов различное влияние. Например, если рассмотреть аналогичные замены в генах *hsp70* (ТАТА→ТГТА) и *his3* (ТАТА→ТГТА и ТАТА→ТАТГ), то в первом случае мутация ведет к снижению уровня транскрипции примерно в 10 раз, а во втором случае к снижению уровня транскрипции почти в 100 раз (Caron et. al., 1991).

В исследовании (Wobbe et. al., 1990) авторы показали на примере контекста двадцати пяти промоторов как мутации в консенсусе  $T_1A_2T_3A_4A_5A_6A_7$  влияют на активность промотора. Замены  $T_1>C$ ,  $T_3>C$ ,  $A_6>G$ , соответственно, в 1, 3 и 6 положениях, приводили по сравнению с диким типом к снижению активности промоторов до 30%. Мутации, в результате которых последовательности ТАТА-боксов принимали следующий вид - ТАААААА или ТАТГАА приводили к падению активности промоторов до нерегистрируемого уровня. Также сравнивали мутации в последовательностях ТАТNTA и ТАТNAA (N – любой нуклеотид). Например, замена в 5 положении  $A_5$  на Т ( $N=A$ ) снижала транскрипцию до 30%, а любая другая

замена (N=C или G, или T) приводила к увеличению активности в 2 раза. Если продолжить анализ последовательностей, фланкирующих ТАТА-боксы, то можно сказать, что одинаковые мутации в ТАТА-боксах находящихся в различных контекстах приводят к улучшению или ухудшению субстратных качеств. Последовательности, фланкирующие ТАТА-элемент, могут влиять и на аффинность ТСБ к ТАТА. У авторов (Librizzi et al., 1998), например, показано, что изменение фланкирующих последовательностей ТАТА в промоторе гена *MLP* уменьшило константу связывания ТСБ с ТАТА-боксом по сравнению с диким промотором. В работе (Wolner et al., 2001) авторы анализировали два аденовирусных ТАТА-бокса: *MLP* (имеет сильный базальный промотор) и *E4* (имеет слабый базальный промотор, но зато высоко индуцибельный в ответ на действие активаторов). Они брали фланкирующие последовательности *E4* и ставили их вместо фланкирующих последовательностей промотора *MLP*, и соответственно получали промотор *MLP* с высоким ответным уровнем на активаторы и наоборот. Из чего можно сделать вывод, что последовательности, фланкирующие ТАТА-элемент, влияют на уровень базальной транскрипции и на ответ активаторов.

В работе (Faiger et al., 2006) авторы определили некоторые мотивы фланкирующих последовательностей как уникальные для определенных ТАТА-бокса. Ведь фланкирующие последовательности в той же степени влияют на взаимодействие ТСБ с ТАТА-элементом, как и последовательности самих ТАТА – боксов. Поэтому архитектурная сборка ТАТА-бокса с фланкирующими последовательностями имеет важное значение и сильно влияет на уровень базальной и активированной транскрипции. Но пока до конца не ясен точный механизм, приводящий к увеличению активированной транскрипции. Также для гена *MLP* последовательности, фланки-

рующие ТАТА-боксы с разных сторон, являются не равнозначными, – с 5'-стороны фланкирующая последовательность более консервативна и имеет повышенное информационное содержание для ТСБ, чем последовательность, фланкирующая ТАТА-бокс с 3'-стороны (Faiger et al., 2006).

Из всего перечисленного выше становится очевидно влияние полиморфизмов, расположенных в ТАТА-боксе и во фланках, на фенотипические признаки. Таким образом, проявляется их особая значимость в процессе онтогенеза человека, когда идет мощная пролиферация клеток и их перемещение, а наличие SNP's может привести к необратимым последствиям, не говоря про изменения координации во времени и пространстве.

Поэтому очевидна необходимость экспериментально-компьютерного исследования SNPs, находящихся в ТАТА-элементах. Такие исследования могли бы способствовать определению генов-кандидатов комплексных заболеваний человека, таких как гипертония, артрит, онкологические заболевания (Lander, 1996; Chakravarti, 1999); а также можно было бы выяснить роль SNP в развитии этих и других заболеваний и предрасположенности к ним, в различной чувствительности человека к факторам окружающей среды и лекарственным препаратам и др. (Nalushka et al., 1999). Такие работы могли бы выявить функционально неаннотированные SNPs, обладающие потенциальной возможностью влияния на важные физиологические процессы и системы организма человека в целом.

В нашем секторе (Савинкова и др., 2007) были получены аналогичные результаты, показывающие влияние фланкирующих последовательностей на сродство олигонуклеотидов к ТСБ. Для проведения этого эксперимента мы использовали рекомбинантный ТСБ человека. Исследование показало, что сродство ТСБ к олиго-

нуклеотидам, фланкирующим ТАТА-боксы и отличающимся по содержанию АТ-пар, отличается в 25-30 раз. Поэтому можно сделать вывод, что негативный вклад в биологическую активность ТАТА-содержащих промоторов может вносить обогащенность фланков ТАТА-боксов легкоплавкими АТ-парами. Следует отметить, что полученные нами экспериментальные значения находятся в соответствии с разработанным ранее алгоритмом предсказания сродства ТСБ к ДНК исходя из ее нуклеотидной последовательности (Ponomarenko et al., 1999). Используя алгоритм, разработанный на основе данных о взаимодействии ТСБ с ТАТА-содержащими олигонуклеотидами, полученных экспериментально (Ponomarenko et al., 1999), был проведен компьютерный анализ, который показал, что вокруг ТАТА боксов расположены GC-обогащенные участки, имеющие меньшее сродство к ТСБ по сравнению со случайными последовательностями ДНК. Вероятно, такой эволюционный отбор на обогащенность фланков «GC» объясняется тем, что тугоплавкие GC-богатые фланки делают более отличимым ТАТА-бокс для ТСБ, а также локализуют легкоплавкую область ТАТА-элемента, как бы фиксируя локальную конформацию ДНК, которая образуется при связывании с ТСБ.

#### **1.4. Взаимодействие ТСБ с ТАТА-боксами**

Используя словосочетание «взаимодействие ТСБ с ТАТА-боксом», обычно предполагают аффинность их связывания или сродство ( $\delta$ ). В работе использовали формулу изменения сродства, основанную на экспериментально измеренной равновесной константе диссоциации ( $K_D$ )  $\delta = -\ln [K_{D, \text{ТАТА Mut}}] - (-\ln [K_{D, \text{ТАТА}}])$  (Кнорре, Мызина, 1992), в которой от логарифма  $K_D$  ТАТА-бокса с мутацией отнимается логарифм  $K_D$  в норме, т.е. без мутации в ТАТА-боксе. Получается, что чем больше

равновесная константа диссоциации, тем больше необходимо субстрата для достижения равновесия и соответственно аффинность хуже, чем у комплекса с меньшей константой (Savinkova et al., 2013).

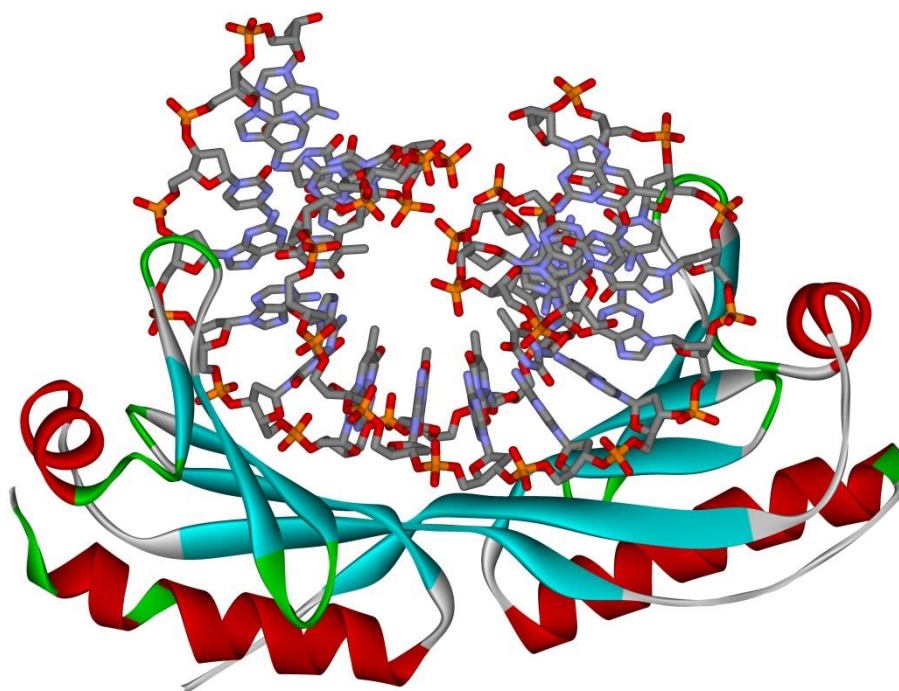
Возможность измерения  $K_D$  появляется благодаря условию выполнения закона действующих масс: 1) лиганд и рецептор сталкиваются благодаря диффузии в верной ориентации и с достаточной энергией; 2) все рецепторы равнодоступны для лиганда; 3) все рецепторы либо свободны, либо связаны с лигандом; 4) ни лиганд, ни рецептор не изменяются после связывания; 5) связывание обратимо.

### **Строение ТАТА-связывающего белка**

Анализ аминокислотной последовательности ТСБ человека, плодовой мушки, дрожжей и *Arabidopsis* показал, что он состоит из филогенетически консервативной С-терминальной части (карбокситерминальный домен- СТД) длиной в 180 аминокислотных остатков, и варибельной N - терминальной части (Hoffman et al., 1990).

Кристаллографический анализ комплексов ТСБ-ТАТА показал, что молекула ТСБ эукариот имеет седлообразную структуру, вогнутая сторона которой служит для специфического связывания с ДНК, а выпуклая для взаимодействия с транскрипционными факторами (Kim et al., 1993; Kim et al., 1993) (Рис. 4). ТСБ состоит из филогенетически консервативного С-концевого домена и варибельного, N – концевого домена (Hoffman et al., 1990). Карбокситерминальный (С-концевой) домен более чем на 80% идентичен у ТСБ разных видов (Patikoglou et al., 1999; Hoffman et al., 1990), он может направлять эффективную и специфическую транскрипцию *in vitro* при добавлении других базальных факторов транскрипции и

РНК-полимеразы II (Meisterernst et al., 1991). Рентгено - структурный анализ результатов футпринтинга и структурного расположения триптических пептидов С-терминального домена (Hernandez, 1993) показали, что он состоит из двух субдоменов, H2 и H2', которые образуют слегка изогнутый антипараллельный бэ́та-лист, формирующий вогнутое ДНК-связывающее седло, и четырех  $\alpha$ -спиралей, лежащих на верхней стороне молекулы. С-терминальный домен контактирует с двуцепочечной ДНК вдоль малой бороздки посредством неполярных и гидрофобных взаимодействий, и вызывает ее локальное расплетание и изгиб спирали. Это создает уникальную конформацию, которая строго необходима для сборки преинициаторного комплекса и коррелирует с эффективностью транскрипции *in vitro* и *in vivo* (Wu et al., 2001). N-концевой домен ТСБ вариабелен по длине и структуре у разных организмов (Horikoshi et al., 1990). Считается, что он обеспечивает видоспецифичность взаимодействия ТСБ.



**Рисунок 4.** Кристаллическая структура комплекса ТСБ человека с ТАТА-элементом (Nikolov et al., 1996).

При изучении взаимодействия ТСБ с ТАТА-боксами промоторов показано, что специфическому связыванию ТСБ с ТАТА-боксом предшествует неспецифическое взаимодействие и скольжение ТСБ вдоль ДНК (Hahn et al., 1989) до встречи с ТАТА-боксом. Затем ТСБ узнает ТАТА-бокс, и в результате интеркаляции фенилаланиновых остатков в малую бороздку ДНК по 5'- и 3'-концам ТАТА-бокса, взаимодействия остатков лизина и аргинина с фосфатными группами ДНК (Starr et al., 1991; Mondal et al., 2014), В-спираль частично разрушается и изгибается, формируются сиквенс-специфические водородные связи между белком и ТАТА-боксом, стабилизирующие комплекс ТСБ/ТАТА (Powell et al., 2002; Kim et al., 2005).

Анализ кристаллов ТСБ/ТАТА и биохимические исследования показали, что деформация ДНК ТАТА-бокса, индуцируемая ТСБ, не зависит строго от последовательности ДНК (Patikoglou et al., 1999). В ряде работ (Davis et al., 1999; Nieb et al., 2007) показано, что в комплексах с консенсусным ТАТА-боксом *AdMLP* ДНК сильно изогнута и анизотропически подвижна в направлении, противоположном изгибу, индуцируемому ТСБ, а ДНК мутантного ТАТА-бокса менее изогнута и подвижна. Исследования позволили (Mondal et al., 2015) предположить, что ДНК ТАТА-бокса регулярно колеблется между несколькими альтернативными структурами, одна из которых более других подходит для связывания ТСБ. Она и используется белком и «подгоняется» далее для образования комплекса ТСБ/ТАТА.

Ввиду важности взаимодействия ТСБ-ТАТА разные авторы изучали влияние разных физиологических условий на процесс комплексообразования. При изучении использовали как полноразмерные белки, так и С-концевые домены ТСБ эукариот (Delgadillo et al., 2009), изучали влияние на процесс связывания ТСБ/ТАТА других



базальных факторов транскрипции (ПВ, ПА) (Gietl et al., 2014; Hieb et al., 2007), TAFs, входящих вместе с ТСБ в состав базального фактора транскрипции TFIID. Проверка влияния температуры, концентрации соли и рН на образование комплексов ТСБ/ТАТА показала, что процесс связывания высоко адаптивен и различается у разных организмов (Gietl et al., 2014).

Hoopes et al. (Hoopes et al., 1992) определили скорость реакции образования комплексов ТСБ *S. cerevisiae* с *AdMLP* и степень конформационного изменения ДНК ТАТА-бокса методом EMSA. Они показали, что ТСБ-ТАТА связывание происходит в две стадии, с образованием двух типов интермедиатов, I1 и I2. Первый тип интермедиатов находится в равновесии со свободными ДНК и белком, и затем превращается в прочно связанный, регистрируемый электрофоретически, конечный комплекс I2. Авторы этой работы положили начало экспериментальному определению различных стадий взаимодействия ТСБ-ТАТА и условий, влияющих на скорость образования комплексов с разными промоторами.

В работе (Hoopes et al., 1992) определили кажущиеся константы образования комплексов ТСБ/ТАТА *AdMLP* при 60 мМ и 80 мМ КСl, равные  $3.4 \cdot 10^5 \text{ M}^{-1} \text{ s}^{-1}$  и  $2.6 \cdot 10^5 \text{ M}^{-1} \text{ s}^{-1}$ , соответственно. Диссоциация комплексов происходит при 60 и 80 мМ КСl со временем полураспада 100 и 65 мин, соответственно. Авторы работы (Hoopes et al., 1992) положили начало определению систематических характеристик взаимодействия ТСБ/ТАТА, экспериментальному определению различных стадий этого взаимодействия и условий, влияющих на скорость образования комплексов с разными промоторами.

Строение транскрипционных комплексов известно, в основном, из биохимических и кристаллографических исследований и рентгеноструктурного анализа. Бла-

годаря этим исследованиям мы представляем структуру комплексов на атомном уровне, но в стационарном режиме. Использование FRET и кинетического метода остановленной струи позволило получить новую информацию о динамике комплексов в процессе их образования и функционирования. При изучении связывания ТСБ/ТАТА с использованием метода остановленной струи с регистрацией FRET сигнала, многими авторами показано, что ТСБ эукариот быстро связывает и одновременно изгибает ДНК (Parkhurst et al., 1996; Masters et al., 2003; Whittington et al., 2008). Изучение кинетики взаимодействия ТСБ/ТАТА позволило обнаружить линейный трехстадийный путь взаимодействия и предположить существование двух промежуточных интермедиатов по пути к конечному комплексу (Parkhurst et al., 1999; Whittington et al., 2008; Delgadillo et al., 2009).

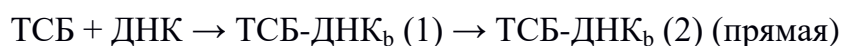
Известно несколько моделей, описывающих механизм взаимодействия ТСБ человека с ТАТА-боксами. Используя видеомикроскопию Tolić-Nørrelykke et al. (Tolić-Nørrelykke et al., 2006) показали, что, по крайней мере, в одном интермедиате ДНК изогнута в меньшей степени, чем в конечном комплексе, за счет того, что только одна часть ТСБ интеркалирует в один конец ТАТА-бокса. Это позволило им предположить поэтапное изгибание ДНК в комплексах, которое другие авторы не подтвердили в используемых ими условиях. Parkhurst et al. (Parkhurst et al., 1996) еще при изучении взаимодействия ТСБ дрожжей *S. cerevisiae* с ТАТА-боксом *AdMLP*, а затем и ТСБ человека, показали, что в режиме реального времени связывание происходит одновременно с изгибом ДНК, образуется два промежуточных комплекса с разными кинетическими характеристиками. Но степень изгиба ДНК в промежуточных и конечном комплексах одинакова (Parkhurst et al., 1999; Delgadillo et al., 2009; Whittington et al., 2008; Masters et al., 2003).

Авторы (Powell et al., 2002) тщательно исследовали термодинамику и кинетику взаимодействия ТСБ *S. cerevisiae* с ТАТА-боксами промоторов *AdML* и *E4* и показали, что термодинамически эти два промотора (ТАТААААААГ и ТАТАТАТА, соответственно) похоже реагируют с ТСБ с образованием двух интермедиатов по пути к финальному комплексу. Образование первого интермедиата (связывание ТСБ с ТАТА и изгибание) происходит для обеих последовательностей через преодоление большого энергетического барьера. Но образование второго интермедиата для *AdMLP* является экзотермическим процессом, а для *E4* – эндотермическим. Образование конечного комплекса для *AdMLP* является высоко эндотермическим процессом, а при образовании комплекса с *E4* происходят незначительные эндотермические изменения. Авторы также отметили разное мольное соотношение интермедиатов: для *E4* второго интермедиата образуется больше в 2.5 раза, чем первого, по сравнению с *AdMLP*. ТСБ узнает вариантную последовательность (ТАТАААСГ) по 3-ступенчатой модели с двумя разветвленными путями. Последовательность имеет драматическое отличие в изгибе спирали ДНК, коррелирующем с транскрипционной активностью. Эта геометрия комплекса в растворе высоко чувствительна к присутствию и концентрации осмолитов, в отличие от комплексов, несущих консенсусные последовательности ТАТА-боксов.

Wittington et al. (Wittington et al., 2008) используя при изучении взаимодействия ТСБ человека с ТАТА-боксом *AdMLP* биофизический метод остановленной струи с регистрацией FRET сигнала и показали, что белок связывается с консенсусным промотором с высокой аффинностью с образованием трех видов комплексов. Одновременно со связыванием происходит сильное изгибание ДНК (до 97°). Структура и функционирование ТСБ/ТАТА комплекса сильно зависит от последовательно-

сти ТАТА-бокса и от вида ТСБ, так как NH<sub>2</sub> –концевые домены определяют видоспецифические различия.

Используя smFRET (одномолекулярный) (Blair et al., 2012) показали, что связывание и изгибание ДНК в комплексе происходит одновременно и частично изогнутые интермедиаты отсутствуют. С консенсусным ТАТА-боксом образуется два типа интермедиатов с ДНК, изогнутой в одинаковой степени, но их образование и диссоциация происходят с разными скоростями. Кинетические данные авторы аппроксимировали следующими моделями:



ТСБ-ДНК<sub>b</sub> (1) – этот интермедиат диссоциирует быстрее в 7 раз, чем движется ко второму состоянию, его можно не регистрировать.

Из экспериментов следует, что молекула ДНК должна подвергаться множественным относительно быстрым переходам “изгиб – исходное состояние”, с остановкой в изогнутом состоянии менее 1 с. Авторы не исключают возможности существования других очень быстрых стадий (менее 30 мс), которые они не регистрируют в используемой системе. Структурные основания для двух кинетических популяций комплексов могут заключаться, например, в связывании ТСБ с ТАТА-боксами в двух противоположных ориентациях.

Другая группа авторов (Nieb et al, 2014) использовала полноразмерный рекомбинантный ТСБ человека и с помощью FRET и конфокальной микроскопии пока-

зала, что ТСБ в режиме реального времени связывает короткие последовательности ТАТА-боксов с высокой специфичностью, одновременно изгибая ДНК. В равновесии находится два типа комплексов при длине ТАТА-боксов 14 пн – несвязанные и связанные с полностью изогнутой ДНК. Но добавление фланкирующих ТАТА-боксов последовательностей увеличивает гетерогенность комплексов – наблюдается неспецифическое кооперативное связывание с множественными участками ДНК, которые конкурируют с ТАТА-боксами. Single molecule FRET эксперименты показали, что взаимодействие может происходить прямым и разветвленным путем, с образованием низкоспецифичных комплексов, которые затем позиционируются в кинетически стабильные (Hieb et al, 2014).

Итак, структурно-функциональные исследования комплексов свидетельствуют о том, что нельзя делать обобщения о строении, кинетике и термодинамике комплексов ТАТА-боксов с ТСБ различных видов, усеченным (С-концевой домен) и полноразмерным белком, комплексами в растворе и в кристаллах, с консенсусными и в разной степени ТАТА-подобными последовательностями, находящимися в разном контексте. До сих пор не исследовалось изменение конформационных свойств ДНК, вызываемых SNPs в регуляторных областях, в частности, ТАТА-боксах конкретных генов человека.

### **1.5. Однонуклеотидные полиморфизмы ТАТА-боксов и их ассоциация с наследственными заболеваниями человека**

Однонуклеотидные полиморфизмы, SNPs, представляют собой наиболее распространенный тип варибельности генома человека. SNPs регуляторных районов

генов и, в частности, кор-промоторов [-70; -20], включающих ТАТА-боксы, могут определять различную чувствительность каждого человека к бактериальной и вирусной инфекциям и лекарственным препаратам, которые используются для их лечения. Регуляторные SNPs связаны с риском развития широко распространенных полигенных (комплексных) заболеваний человека, таких как артрит, гипертония, сердечно-сосудистые заболевания, рак, болезнь Альцгеймера и другие (Baicu et al., 2007; Beerl et al., 2005; Bruehl et al., 2009; Booth et al., 2015).

Изучение регуляторных SNPs (rSNP) дает уникальную информацию не только об устройстве регуляторных районов генов и функциональной значимости отдельных регуляторных элементов. В последнее десятилетие показана роль rSNP в эволюционных процессах, в развитии различных патологических состояний путем изменения уровня экспрессии кандидатных генов (Yokoyama et al., 2016).

На сегодняшний день в БД NCBI dbSNP (Sherry et al., 2001) зарегистрировано свыше 10 млн SNPs большая часть которых пока считается нейтральными функционально, другая же часть SNPs имеет и биологическое и/или медицинское значение (Teufel et al., 2006), и находится как в кодирующей, так и в регуляторной части гена. Работ по изучению полиморфизмов в кодирующих областях генов значительно больше, чем работ посвященных изучению полиморфизмов в регуляторной области гена. Методы изучения SNPs в обеих частях генов основываются на анализе изменения паттерна и/или уровня транскрипции гена, либо через изменение структуры и свойств кодируемого белка (Simoni et al., 2002). Для здоровья человека бывают одинаково значимы как полиморфизмы в кодирующей области гена, так и SNP's, находящиеся в регуляторной области гена. Например, некоторые SNPs в кодирующей части генов метаболизма и репарации ДНК могут влиять на индивиду-

альную восприимчивость к онкологическим заболеваниям и терапевтическому лечению (Bartsch et al., 2007), такие изменения приводят к изменению кодона и соответственно замене одной аминокислоты на другую. К таким SNPs в кодирующей области гена также относятся замены в гене *LIMD1* человека, которые ассоциируются с раком молочной железы и легких (Huggins et al., 2007), в гене аполипопротеина E – это связано с хроническими неинфекционными заболеваниями человека и предрасположенностью к ряду сердечно-сосудистых и нейродегенеративных заболеваний (Hubásek et al., 2003), в гене интерлейкина -1 $\beta$  SNP ассоциируется с раком поджелудочной железы и плохим прогнозом для пациентов (Barber et al., 2000) и др. К полиморфизмам, находящимся не в кодирующей части, можно отнести SNPs интронов. Так в интроне 6 гена триптофаноксигеназы две мононуклеотидные замены ассоциированы с психическими нарушениями у человека (предрасположенность к алкоголизму и наркомании, синдром гиперактивности у детей, синдром Туретта,) (Comings et al., 1996). В гене IL-1Ra (ген рецептора антагониста интерлейкина 1) SNP интрона 2 ассоциируется с одной из разновидностей рака желудка (Machado et al., 2001), SNP в AU-богатой области гена глюкокортикоидного рецептора  $\beta$  – с ревматоидным артритом (Derijk et al., 2001) и др. Такие SNP's вносят свой вклад в фенотипическую изменчивость человека и увеличивают количество заболеваний.

В работе (Савинкова и др., 2009) рассмотрено около 40 полиморфизмов ТАТА-элементов промоторов, которые ассоциированы с тем или иным заболеванием у человека и подтверждены молекулярно-биологически и клинически. Некоторые SNP's снижают транскрипционную активность промоторов, что ведет к повышению риска возникновения: стенокардии (Gordon et al., 1989; Rubin et al., 1991), рака

полости рта в 8-10 раз у курильщиков (Zheng et al., 2001), нейродегенеративных заболеваний (Niemann et al., 2007) и др. Другие же полиморфизмы увеличивают транскрипционную активность промоторов, что также приводит к повышенной чувствительности к стенокардии и легочным заболеваниям, гипертонии (Burgner et al., 2003), а также ассоциируется с различными патологиями в развитии мелких сосудов головного мозга (Zhao et al., 1999) и др. Интересно, что при возникновении *de novo* канонического ТАТА-бокса могут появиться такие болезни как Коппкоподобная катаракта (Brakenhoff et al., 1994) или рак эндометрия (Liu et al., 2001). Кроме того, известны случаи превращения ТАТА-бокса в другой сайт. Например, при наличии SNP -33Т>С ТАТА-бокс превращается в сайт связывания эритроидного фактора транскрипции GATA-1 (Tournamille et al., 1995) и репрессирует ген *DARC* у большинства коренных жителей Западной Африки, что повышает их устойчивость к малярии. В другом случае SNPs -26G>С и -20Т>А разрушают ТАТА-бокс, превращая его в сайт связывания ядерного фактора гепатоцитов HNF-4 (Reijnen et al., 1992). В норме фактор HNF-4 контролирует экспрессию гена фактора свертывания крови IX, полиморфизм же приводит к гемофилии-В Лейдена.

### **Влияние SNP's в ТАТА-боксах на фенотипические признаки**

Стремительно возрастающий в последние годы интерес к выявлению и исследованию SNPs в различных генах вызван тем, что они часто оказываются связанными с разнообразными фенотипическими проявлениями, включающими предрасположенность ко многим заболеваниям. На сегодняшний день активно разрабатываются подходы по созданию прогнозных моделей, которые можно использовать как для выявления лиц с высоким риском к бессимптомным болезням, так и для



дополнительной диагностики и принятия решений, для экспериментального моделирования (Lee et al., 2016).

Создаются и уже разработаны прогнозные модели, учитывающие, например, как физиологические показатели людей, их историю развития, так и наличие в их маркерных генах SNP's (Guo et al., 2015).

Ранее была показана клинически и молекулярно-биологически связь между SNP и ассоциированным с ним заболеванием. В нашем секторе был проведен ряд работ, показывающих, что SNP ТАТА-бокса влияет на взаимодействие ТСБ с ТАТА-боксом (Драчкова и др., 2011; Драчкова и др., 2010). В нашем секторе экспериментально было показано снижение сродства ТСБ к ОДН, несущим полиморфизмы, ассоциированные с аутоиммунными, сердечно-сосудистыми заболеваниями, амиотрофическим латеральным склерозом, неврологическими расстройствами дегенеративного характера, повышенной чувствительностью к бактериальным инфекциям и др., которые были подтверждены молекулярно-биологически и клинически. Во всех рассматриваемых случаях нами было показано снижение сродства ТСБ к ТАТА-содержащим олигонуклеотидам, что согласуется с экспериментальными данными других авторов о снижении мРНК и белка *in vivo* у обследованных пациентов (Таблица 1).

**Таблица 1.** Корреляция срoдства ТСБ к олигонуклеотидам, соответствующим ТАТА-боксам промоторов генов человека, несущим SNPs, ассоциированные с наследственными патологиями (Драчкова и др., 2011), с количеством продукта гена.

Литературные данные		Экспериментальные данные	
Ген, SNP	Содержание фермента	Последовательность олигонуклеотида (5'–3' нити)	Kd, нМ
<i>MBL</i>	Норма	catctatttcTATAgctgcacc	28±13
-35T>C (Boldt et al, 2006)	Снижено содержание фермента в сыворотке крови	catctatttcCATAgctgcacc	160±64
<i>CYP2A6</i>	Норма	tttcaggcagTATAAAggcaaacc	17±4
-48T>G . (Pitarque et al, 2001)	Снижено содержание фермента и синтез РНК <i>in vitro</i> на 50%	tttcaggcagTAGAAAggcaaacc	84±22
<i>SOD1</i>	Норма	aggtctggccTATAAAgtagtcgcgg	40±8
-27A>G (Niemann et al, 2007)	Снижено содержание фермента ~ на 40% <i>in vivo</i>	aggtctggccGTATAAAgtagtcgcgg	170±30
<i>TPI</i>	Норма	cgcggcgctcTATAAAgtgggcagt	7±1
-24T>G (Watanabe et al, 1996)	Снижение содержания мРНК до 50% <i>in vivo</i>	cgcggcgctcTATAGAAgtgggcagt	290±80

В Табл.1 приведены значения равновесных  $K_D$  комплексов ТСБ с ТАТА- боксами ряда промоторов генов человека, содержащих SNPs, ассоциированные с

предрасположенностью к наследственным заболеваниям, полученные в ходе экспериментов.

Один из рассматриваемых генов - манноза-связывающий лектин (*MBL*). *MBL* - белок сыворотки крови, синтезируется в печени и, вместе с другими белками (белки комплемента, сурфактанта легких, фиколина), играет центральную роль во врожденном иммунном ответе (Garred *et al.*, 2006). Ген этого белка расположен на 10 хромосоме (4 экзона и 3 интрона в составе) и представляет собой молекулу букето – подобной структуры - три идентичных полипептидных цепи по 32 кДа, содержащие цистеин-богатые и коллагеновые области, и углевод-связывающие домены, объединены в одну субъединицу. Шесть таких субъединиц образуют *MBL* (Jack *et al.*, 2001). Для рабочего состояния *MBL* сегменты по 6 субъединиц объединяются в структуры более высокого порядка – тетрамеры (Jack *et al.*, 2001; Madsen *et al.*, 1998). Функциональная активность белка состоит в том, что *MBL* является рецептором маннозы и N-ацетилглюкозаминов клеточной поверхности микробов (Торшин и др., 2003). При связывании с маннозой и углеводными остатками, входящими в состав клеточной стенки патогенов, *MBL* может либо прямо опсонизировать патогены, связываясь с их клеточной стенкой и инициируя освобождение провоспалительных цитокинов (Boldt *et al.*, 2006), либо через систему лектинового комплемента способствовать удалению патогенов с помощью комплемент-опосредованного фагоцитоза.

В приведенных исследованиях было показано, что генетически детерминированная вариабельность концентраций *MBL* в сыворотке крови человека влияет на чувствительность к различным инфекциям, а также на предрасположенность к аутоиммунным, метаболическим и сердечно-сосудистым заболеваниям (Sisen,

Minchinton, 2003). Низкие концентрации MBL предрасполагают к повторяющимся инфекциям, при этом варианты гена, кодирующие низкие концентрации сывороточного MBL, довольно широко распространены во всех этнических группах. Поэтому можно предположить, что так как носители таких аллелей есть, то низкие концентрации MBL обеспечивают преимущество хозяину по аналогии с классическим примером серповидно-клеточной анемии (Garred *et al.*, 2006).

При исследовании вариабельности гена *MBL* в бразильских популяциях, было обнаружено 8 общих гаплотипов (включающих полиморфизм в промоторной области и в первом экзоне), которые ассоциируются с прогрессивным снижением концентрации сывороточного MBL: *MBL\*HYPA* > *LYQA* > *LYPA* > *LXPA* >> *HYPD* = *LYPB* = *LYQC* (Boldt *et al.*, 2006). Интересно, что распространены, как и полностью функциональные гаплотипы, так и с полиморфизмами - гаплотип *LYPB*, включающий полиморфизм в ТАТА-боксе (-788 Т> С) и в первом экзоне (1052 G> А), и гаплотип *LYQC*, вызывающий дефицит MBL в гомозиготном состоянии. Также было показано, что полиморфизм в ТАТА-боксе (-788 Т> С) снижает концентрацию MBL в сыворотке крови, что приводит к снижению эффективности иммунной защиты и повышает чувствительность к различным инфекциям, а также к метаболическим и сердечно-сосудистым заболеваниям. Литературных данных о взаимодействии ТСБ с SNP-содержащим ТАТА-боксом гена *MBL* или о влиянии SNP на транскрипцию, нет. Полученные в нашем секторе экспериментальные результаты (Драчкова и др., 2011) свидетельствуют о том, что SNP -35Т>С в ТАТА-боксе этого гена снижает аффинность ТСБ к нему в 5,7 раза. Это вполне согласуется с эффектом снижения количества фермента в сыворотке крови, что может привести к серьезным нарушениям в организме человека, таким как аутоимунные и сосудистые

заболевания, повышенная чувствительность к инфекциям и другое (Sisen, Minchinton, 2003).

Суперсемейство генов цитохрома P450 человека (*CYP 450*) - состоит из 57 функциональных генов и 58 псевдогенов (Zukunft *et al.*, 2005). Элементы семейства находятся в эндоплазматическом ретикулуме печени и внепеченочных тканей и кодируют ферменты, относящиеся к классу монооксигеназ. Цитохромы P450 отвечают за детоксикацию и удаление из организма ксенобиотиков (включая многие лекарственные препараты), опосредуют окислительный метаболизм огромного количества структурно разнообразных экзо- и эндогенных веществ. С помощью изоформ цитохрома P450 осуществляется биосинтез и катаболизм холестерина, стероидных гормонов, так называемых «цитохромных» метаболитов и арахидоновой кислоты, обладающих вазотропным действием. Кластер семейства генов *CYP2* человека расположен на 19 хромосоме и содержит подсемейства *CYP2A*, *CYP2B*, *CYP2F*, *CYP2G*, *CYP2S* и *CYP2T*. Ген *CYP2A6* относится к подсемейству *CYP2A*. *CYP2A6* играет главную роль в метаболизме никотина и кумарина, участвует в детоксикации некоторых лекарств и нитрозаминов, содержащихся в табаке (Pitarque *et al.*, 2005). Кроме печени, *CYP2A6* экспрессируется в стероидо-зависимых тканях, таких как молочные железы, яичники, матка и надпочечники (Higashi *et al.*, 2007).

Цитохром P450<sub>2A6</sub> (*CYP2A6*) человека играет важную роль в активации некоторых проканцерогенов и детоксикации многих лекарств, и является главной никотиноксидазой (Pelkonen *et al.*, 2000). Изобилие полиморфных вариантов гена фермента приводит к большим индивидуальным различиям в активности. В некоторых географических местах популяционные частоты дефектного ТАТА-бокса с мутацией -48T>G оказались очень высокими: в отдельных популяциях Швеции – 5,2%,

Турции – 7,2%, Китая – 15,7% (Pitarque *et al.*, 2001), Японии – 21,3% и Кореи – 22,3% (Yoshida *et al.*, 2003). Такое распределение свидетельствует о больших межэтнических различиях в распространении аллеля. В работе (Pitarque *et al.*, 2001) показано, что экспрессия плазмидных конструкций, содержащих SNP -48T>G в ТА-ТА-боксе промотора гена *CYP2A6*, в клетках В16А2 гепатомы человека составляет 50% от дикого типа и на 50% снижена активность фермента *in vitro*. SNP -48T>G может увеличивать риск возникновения рака легких у курильщиков, потому что снижение активности фермента приводит к снижению скорости метаболизма и выведению из организма потенциальных проканцерогенов табака, (Pitarque *et al.*, 2005). В Табл. 1 видно, что рассматриваемый SNP гена *CYP2A6* приводит к снижению аффинности ТСБ к ТАТА-боксу в 4,9 раза.

Продукт гена *SOD1* (*SOD1* - Cu/Zn) - супероксид дисмутаза кодирует металлофермент, который нейтрализует в клетке свободные радикалы, неконтролируемые концентрации которых могут ее повредить. Для функционирования ферменту необходимы медь и цинк. Хотя продукт гена *SOD1* экспрессируется практически во всех типах клеток организма и относится некоторыми авторами к генам домашнего хозяйства (Skvortsova *et al.*, 2001), наиболее он активен в печени, почках, надпочечниках и селезенке (Kuźma-Kozakiewicz, Kwiecinski, 2009). Ген *SOD1* человека расположен в 21 хромосоме и примерно 10% мутаций в гене ассоциируются с семейным амиотрофическим латеральным склерозом (ALS), поражающим двигательные нейроны человека (Rosen *et al.*, 1993), а 90% мутаций ассоциируются со спорадическим амиотрофическим латеральным склерозом (Kuźma-Kozakiewicz, Kwiecinski, 2009), основная разница между которыми заключается в продолжительности жизни пациентов (46 и 56 лет, соответственно).

Амиотрофический латеральный склероз (ALS) – тяжелое заболевание пока неизвестной полностью этиологии и патогенеза. Пока выявлено 117 различных мутаций гена *SOD1* у больных ALS (Kabashi *et al.*, 2007), что обуславливает большую сложность при выявлении причин, вызывающих повреждение и гибель моторных нейронов, приводящих к тяжелым нейродегенеративным заболеваниям человека. Считается (Kuźma-Kozakiewicz *et al.*, 2009), что аминокислотные замены увеличивают появление популяции молекул белка с частичным фолдингом, что является причиной их агрегации и токсичности. В некоторых работах (Kayatekin *et al.*, 2008; Vonk, Klomp, 2008) была исследована роль баланса Zn и Cu в регуляции стабильности и активности SOD1, авторы пришли к заключению, что в патогенез ALS вовлекается дисбаланс металлов. Другие исследователи склонны считать, что на тяжесть заболевания ALS не влияет уменьшение количества белка SOD1 или присутствие агрегированных и незрелых (не прошедших фолдинг) форм его, а скорее влияет совокупность биофизических и биологических факторов, воздействующих на SOD1, что и делает его токсичным для двигательных нейронов (Vassall *et al.*, 2011). Существует много работ и гипотез, пытающихся пролить свет на причину токсического воздействия SOD1 на двигательные нейроны, но вопрос до сих пор остается открытым. В работе (Niemann *et al.*, 2007) было показано, на примере двух больных семейным амиотрофическим латеральным склерозом, умеренное снижение экспрессии SOD 1 в результате SNP -27A>G в ТАТА-боксе промотора этого гена, не ассоциированное с какими-либо аминокислотными заменами в молекуле белка. Авторами Драчкова и др. (Драчкова и др., 2011) в Таблице 1 было показано снижение аффинности взаимодействия ТСБ с SNP-содержащим ТАТА-боксом в 4,3 раза.

Ген *TPI* (TPI - триозофосфатизомераза) расположен на 12 хромосоме (состоит из 7 экзонов и 6 интронов Valentin *et al.*, 2000) и кодирует фермент, катализирующий взаимопревращение диоксиацетонфосфата и D-глицеральдегид-3-фосфата, которое абсолютно необходимо для гликолитического цикла. TPI экспрессируется во всех типах тканей и клеток и относится к генам «домашнего хозяйства». (Brown *et al.*, 1985; Watanabe *et al.*, 1996; Manco *et al.*, 2009). Функциональная каталитическая активность TPI происходит при формировании его из двух субъединиц с молекулярной массой 27 кДа в гомодимер (Orosz *et al.*, 2006). Промотор гена *TPI* обогащен GC-динуклеотидами и содержит ТАТА- и СААТ-боксы, как большинство генов домашнего хозяйства. В геноме человека есть еще три псевдогена *TPI*, которые потеряли интроны в процессе эволюции и развивались независимо. Различные формы TPI - электрофоретические и хроматографические, были обнаружены во всех тканях человека. Они кодируются одним геном и являются результатом посттрансляционных модификаций белка (Brown *et al.*, 1985). В работах ранее было показано, что мутации, вызывающие дефицит TPI, приводят к снижению активности фермента в эритроцитах до 2-10% от нормы, что в свою очередь приводит к хронической гемолитической анемии, неврологическим расстройствам дегенеративного характера, задержке роста и отставании в развитии, повышенной чувствительности к бактериальным инфекциям, кардиомиопатии, детской смертности и др. (Watanabe *et al.*, 1996; McMullin, 1999). Позже *in vivo* было показано, что снижение активности TPI компенсируется повышением активности других ферментов гликолиза (Orosz *et al.*, 2006). Интересным является вывод, представленный в заключительной работе (Orosz *et al.*, 2010) - неизлечимые последствия дефицита TPI являются про-



явлением скорее конформационных изменений фермента (нарушение белок-белковых взаимодействий), чем каталитической активности.

В качестве примера хочется привести мутацию в ТАТА-боксе -24Т>G, которая была обнаружена у трех из десяти обследованных человек, и та же самая мутация находится в составе гаплотипа 5G-8A -24G. По данным исследователей такая мутация у гетерозигот ассоциируется с умеренным снижением активности ТРІ *in vivo* (26 -50%) (Humphries *et al.*, 1999; Watanabe *et al.*, 1996). В работе (Драчкова и др., 2011) для замены в ТАТА-боксе -24Т>G было получено значительное снижение аффинности ТСБ к ТАТА-боксу - в 41 раз.

В работе (Драчкова и др., 2010) подробно был изучен вопрос о наследственных нарушениях синтеза гемоглобина, вызываемых SNPs в ТАТА-боксе промотора гена β-глобина человека, которые ассоциируются с β-талассемией различной тяжести (Табл. 2). Экспериментально были получены количественные характеристики (K<sub>D</sub>) взаимодействия ТСБ человека с ТАТА-боксами промотора гена β-глобина, несущими SNPs, которые согласуются с литературными данными о снижении синтеза РНК нормального β-глобина у больных β-талассемией.

**Таблица 2.** Влияние SNPs на средство ТСБ к ТАТА-боксам промоторов гена  $\beta$ -глобина человека, и уровень мРНК  $\beta$ -глобина (Драчкова и др., 2010).

Влияние SNPs на уровень мРНК $\beta$ -глобина (литературные данные)			Последовательности олигонуклеотидов (5'-3'нити)	K <sub>D</sub> “ТСБ-ODN” nM
ODN N	SNP	Уровень мРНК $\beta$ -глобина		
1	$\beta$ -глобин (норма) (Fei <i>et al.</i> , 1988)	Норма	cagggctgggCATAAAgtcagggca	21±3
2	-31A > G (Takahara <i>et al.</i> , 1986)	В <i>cos</i> -клетках – в 2 раза меньше мРНК	cagggctgggCGTAAAgtcagggca	120±20
3	-30T > A (Fei <i>et al.</i> , 1988)	<i>In vivo</i> - 8 – 13% мРНК	cagggctgggCAAAAAgtcagggca	92±17
4	-30T > C (Cai <i>et al.</i> , 1989)	Нет данных	cagggctgggCACAAGtcagggca	113±17
5	-29A > G (Antonarakis <i>et al.</i> , 1984)	<i>HeLa</i> и <i>in vivo</i> – 25% мРНК	cagggctgggCATGAAgtcagggca	530±40

6	-28A > C (Poncz <i>et al.</i> , 1982)	<i>In vivo</i> – ча- стичное или полное отсут- ствие мРНК	cagggctgggCATACAagtcagggca	330±50
7	-28A > G (Orkin <i>et al.</i> , 1983)	<i>In vivo</i> – в 10 раз, в <i>HeLa</i> – в 3-5 раз меньше мРНК	cagggctgggCATAGAagtcagggca	60±14
8	-27A > T (Badens <i>et</i> <i>al.</i> , 1999)	<i>HeLa</i> - ~ в 5 раз меньше мРНК	cagggctgggCATAATagtcagggca	63±15

Наиболее распространенным классом моногенных заболеваний в мире являются талассемии  $\alpha$ -,  $\beta$ - и  $\delta$ - различной тяжести, вызванные наследственными нарушениями синтеза гемоглобина (Traeger-Synodinos *et al.*, 2008). Талассемии происходят из-за повреждений (однонуклеотидные замены, вставки, делеции) в структурной части и в регуляторных районах  $\alpha$ -,  $\beta$ - и  $\delta$ - глобиновых генов. В среднем симптомы  $\beta$ -талассемии встречаются у 1 человека на 100000, и гораздо чаще, у 5 - 30% популяций отдельных этнических групп в Средиземноморье, Африке и Южной Азии (Muncie *et al.*, 2009; Galanello *et al.*, 2010). Наследственные изменения SNPs в ТА-ТА-боксах – это нарушения, приводящие к синтезу недостаточного количества цепей нормального  $\beta$ -глобина или полному его отсутствию, в результате чего возникает избыток свободных цепей  $\alpha$ -глобина ( $\alpha_2\beta_2$  – основная структурная единица гемоглобина взрослого человека, Hb A), нарушение эритропоэза и низкий выход (или

полное его отсутствие) гемоглобина взрослого человека (Schechter, 2008).  $\beta$ -талассемии, вызываемые SNPs в ТАТА-боксе промотора гена  $\beta$ -глобина, очень редки.

В литературе описано три основных формы  $\beta$ -талассемии (Muncie *et al.*, 2009; Galanello, Origa, 2010). Тяжелые заболевания, такие как гемолитическая анемия, нарушения в развитии скелета, задержка роста, поражение внутренних органов, желтуха и трансфузионная зависимость (необходимость переливания эритроцитарной массы) с 6-месячного возраста вызывает большая (major)  $\beta^0$ -талассемия (или анемия Кули), вызванная наличием двух дефектных копий генов, унаследованных от обоих родителей (гомозигота по дефектному гену  $\beta$ -глобина), выраженная в сильном понижении или полным отсутствием синтеза нормального  $\beta$ -глобина (Muncie *et al.*, 2009). Такие больные обычно умирают без лечения - в первые 2 года жизни, при лечении доживают до возраста 20 - 30 лет. Гетерозиготы же, носители одного мутантного аллеля, а другого - нормального  $\beta$ -глобина со средним и выше среднего уровнем синтеза цепей нормального  $\beta$ -глобина, имеют умеренно выраженную анемию или бессимптомную анемию с нормальным качеством жизни, и они составляют группу больных малой (minor или trait)  $\beta^+$ -талассемией. Существует также промежуточная форма (intermedia)  $\beta$ -талассемии, больные которой являются также гетерозиготами -  $\beta^+$ -талассемия - но только с остаточным синтезом нормального  $\beta$ -глобина (~10% и выше). Для больных характерны умеренно выраженные симптомы заболевания, иногда с эпизодической трансфузионной зависимостью (Muncie *et al.*, 2009).  $\beta$ -талассемию относят к заболеваниям аутосомно-доминантного типа и при таком типе наследования действие мутантного гена проявляется в гетерозиготном состоянии (Thein, 2008).

Семейство генов  $\beta$ -глобинов (семейство  $\epsilon$ -,  $\gamma$ -,  $\beta$ - и  $\delta$ -генов) человека имеет сходную организацию - 3 экзона и 2 интрона и располагается на 11-ой хромосоме (Schechter, 2008), занимая участок в 60-65 тысяч пар нуклеотидов. Последовательность расположения генов совпадает с порядком их экспрессии в процессе онтогенеза: сначала транскрибируется расположенный в кластере  $\beta$ -глобинов ген эмбриональной  $\epsilon$ -цепи, следующими за ним включаются два гена фетальных  $\gamma$ -цепей, и уже затем функционально активными становятся - гены  $\delta$ - и  $\beta$ -цепей взрослого организма (Schechter, 2008). Существующая программа регуляции транскрипции этого семейства генов осуществляется через проксимальные и дистальные регуляторные элементы ДНК (Levings *et al.*, 2002). Имеются данные о факторах транскрипции, взаимодействующих с индивидуальными промоторами генов  $\beta$ -глобинового кластера на разных стадиях развития организма (Tallack *et al.*, 2010).

В работе (Драчкова и др., 2010) рассмотрены SNPs в ТАТА-боксах гена  $\beta$ -глобина разных людей, являющиеся причиной различной тяжести заболевания крови  $\beta$ -талассемией. В зарубежных работах, наряду с выявлением мутантных ТАТА-боксов промоторов генов больных  $\beta$ -талассемией, получены также относительные характеристики содержания РНК нормальных цепей  $\beta$ -глобинов.

Представленные результаты экспериментов в Таблице 2 по определению влияния SNPs ТАТА-боксов, идентичных ТАТА-боксам больных  $\beta$ -талассемией, на взаимодействие с ТСБ ( $K_D$  комплексов «hТСБ-ODN») показывают, что ТАТА-бокс гена  $\beta$ -глобина здорового человека весьма вариабелен: например, САТАААА вместо канонического ТАТАААА, его иногда называют АТА-бокс, где АТА-бокс является характерной составляющей генов домашнего хозяйства с достаточно постоянным уровнем экспрессии (Fei *et al.*, 1988). В работе (Stewart *et al.*, 2006) было

показано, что последовательности ТАТАААА и САТАААА почти не отличаются при связывании, и что ТАТА-бокс с последовательностью САТАААА более чем в 40 раз менее эффективен в индуцированной транскрипции *in vivo* по сравнению с ТАТАААА (авторы использовали в экспериментах ТСБ дрожжей *S. cerevisiae*). Авторы (Bjornsdottir *et al.*, 2008) пошли дальше, они показали с использованием очищенных из дрожжей базальных факторов транскрипции и матрицы с промотором *CYCI*, что замена ТАТАААА→САТАААА снижает транскрипцию до 40% от нормы. Такую же замену (ТАТАААА→САТАААА) исследовали авторы (Wobbe, Strahl, 1990) и они показали, что транскрипция в *HeLa* снижается в 3 раза при замене в ТАТА-боксе первого Т на С. Из этих примеров видно, что замена Т на С в первой позиции ТАТА-бокса оказывает разное влияние на транскрипцию, вероятно, в силу использования авторами разных экспериментальных условий. В работах (Faiger *et al.*, 2006; Савинкова и др., 2007) рассматривается вклад в сродство ТСБ к олигонуклеотидам последовательностей фланков ТАТА-боксов, которые отличаются по содержанию АТ-пар и количественно могут отличаться в 25–30 раз. В работе (Драчкова и др., 2010) авторы определили сродство ТСБ к ТАТА-боксу гена  $\beta$ -глобина (*ODN1*) здорового человека -  $K_D = 21$  нМ. Там же был рассмотрен полиморфизм SNP -31 А>G (*ODN N2*, Табл. 2) у больной промежуточной  $\beta^+$ -талассемией японки, гетерозиготной по гену  $\beta$ -глобина (Takahara *et al.*, 1986). Такая замена при экспрессии гена в клетках COS приводила к снижению содержания РНК нормального гемоглобина в 2 раза, а отношение  $\beta$ - и  $\alpha$ -цепей равнялось 0.34 в эритроидных клетках пациентки (вместо 0.5), и при этом пациентка не зависела от переливания эритроцитарной массы. Авторы Драчкова и др. (Драчкова и др., 2010)

определили сродство ТСБ к такой ODN -  $K_D = 120$  нМ – уменьшение  $K_D$  почти в 6 раз относительно нормы.

В Таблице 2 представлено сродство ТСБ к ODN с SNP -30T>A (Fei *et al.*, 1988), которое снижалось более чем в 4 раза от нормы –  $K_D = 92$  нМ (ODN N3). Такая замена была встречена у югослава и турка с промежуточной формой  $\beta^+$ -талассемии, с уровнем синтеза РНК нормального  $\beta$ -глобина от 8 до 13%.

Также представлено в Таблице 2 снижение сродства ТСБ до 113 нМ – ODN N4 при SNP -30 T>C, обнаруженной у носителя промежуточной  $\beta^+$ -талассемии (Cai *et al.*, 1989). В работе (Strahs *et al.*, 2003) был приведен пример влияния замены Т на С в третьей позиции в промоторе *AdML* - (ТАТА→ТАСА) – при этом эффективность транскрипции снижалась до 20% от «дикого» типа, что согласуется с измеренным экспериментально снижением сродства в ~5.3 раза.

Для мутации -29 А>G (ODN N 5) (Antonarakis *et al.*, 1984) с фенотипом промежуточной  $\beta^+$ -талассемии, носители которого не зависели от переливания эритроцитарной массы, была получена  $K_D = 530$  нМ (Таблица 2). Интересно, что в работе (Antonarakis *et al.*, 1984) влияние аналогичной замены в  $\beta$ -глобиновом промоторе (САТА → САТG) приводило к снижению связывания ТСБ до 8 -13 % относительно нормы, а синтез РНК в клетках *MEL* достигал 38.5%.

Замена в ТАТА-боксе ТАТАААА → ТАТGААА приводила к снижению транскрипции в *HeLa* до нерегистрируемого уровня (Wobbe *et al.*, 1990), а *in vitro*, с выделенными из дрожжей базальными факторами транскрипции (Vjornsdottir *et al.*, 2008), - до 8% и ниже, что согласуется с полученной характеристикой сродства ТСБ к такому ODN (Таблица 2). Сравнивая эти результаты и данные авторов, открывших этот SNP (Antonarakis *et al.*, 1984) и показавших, что с SNP-содержащего

ТАТА-бокса промотора в *HeLa* и *in vivo* транскрибировалось до 25% РНК  $\beta$ -глобина, можно заключить, что они являются наглядным примером возможностей компенсаторной регуляции экспрессии генов в организме, а также результатом использования разных экспериментальных систем и условий *in vitro*.

У двух братьев 1,5 лет, курдских евреев, больных большой ( $\beta^0$ ) талассемией и зависящих от переливания эритроцитарной массы, была обнаружена SNP -28A>C, приводящая к частичному или полному отсутствию мРНК глобина. Авторы работы (Wobbe *et al.*, 1990) используя ODN с аналогичной мутацией ТАТАААА  $\rightarrow$ ТАТАСАА показали на *HeLa* снижение транскрипции в 20 раз от нормального уровня, и это согласуется с полученными данными в Драчкова и др. (Драчкова и др., 2011), где характеристика взаимодействия ТСБ с таким ODN (N6) -  $K_D = 330$  нМ, а это снижение сродства в 16 раз по сравнению с нормой.

У китайского малыша 14 месяцев, больного умеренно тяжелой формой большой  $\beta^0$ -талассемии, был обнаружен SNP -28A>G (Orkin *et al.*, 1983). Больной оказался гомозиготен по этому аллелю и вскоре возникла необходимость в переливании эритроцитарной массы из-за задержки роста и анемии. Авторы (Orkin *et al.*, 1983) показали, что в *HeLa* синтез нормального  $\beta$ -глобина снижался в 3-5 раз и это совпадает с результатами, полученными *in vivo* - РНК нормального  $\beta$ -глобина было в 10 раз меньше относительно нормы. У других исследователей (Wobbe *et al.*, 1990) транскрипция в *HeLa* при той же замене ТАТАААА  $\rightarrow$ ТАТАГАА снижалась до 5% от нормы. В Таблице 2 представлена  $K_D$  для такого SNP -  $K_D = 60$  нМ, это снижение в 3 раза по сравнению с нормой.

На острове Корсика у нескольких членов семьи с малой ( $\beta^+$ ) талассемией была обнаружена SNP -27A>T (ODN N8) (Badens *et al.*, 1999). У авторов (Драчкова и



др., 2010) наличие такого SNP в ТАТА-боксах привело к уменьшению сродства «ТСБ/ODN» не более чем в 3 раза ( $K_D = 63$  nM), а у исследователей (Wobbe *et al.*, 1990) в HeLa такой SNP (ТАТАААА→ТАТААТАА) приводил к снижению транскрипции примерно в 5 раз относительно нормы – результаты вполне сравнимые.

Поэтому полученные равновесные константы диссоциации ( $K_D$ ) (Таблица 2) комплексов ТСБ с ТАТА-боксами при  $\beta$ -талассемии, свидетельствующие о снижении сродства ТСБ к SNP-содержащим ТАТА-боксам, которое, по литературным данным, согласуется с уменьшением синтеза РНК нормального  $\beta$ -глобина *in vivo*, в cos-клетках и *HeLa*.

### **1.6. Формирование электронной библиотеки и компьютерный анализ неаннотированных однонуклеотидных полиморфизмов ТАТА-боксов промоторов генов липидного гомеостаза человека**

В настоящей работе для исследования была сформирована электронная библиотека из 388 неаннотированных SNPs в кор-промоторных районах 68 генов пищевого поведения человека. В указанном районе для 62 генов с помощью Web сервиса SNP\_TATA\_Comparator было обнаружено 388 SNPs в районе их промоторов, для шести генов - *GALR3*, *MBOAT4*, *NTSR2*, *OPRD1*, *PMCH*, *UCN* - не было обнаружено SNPs. Для 296 из 388 SNPs *in silico* было оценено недостоверное изменение сродства ТСБ к кор-промоторам генов пищевого поведения человека относительно референсного генома человека (hg19). Для каждого из остальных 91 SNPs прогнозировалось достоверное увеличение или уменьшение ТСБ/ДНК-сродства, которым в клетках человека соответствуют избытки или недостатки концентрации белковых продуктов этих генов, потенциально ассоциируемые с патологиями.

Наиболее сильное изменение сродства ТСБ/ТАТА прогнозировалось для генов *NPB*, *PCSK1*, *UBL5*, *PDYN*, *GRP*, *GALR1*, *LEP*, *FTO*, *CRHR1*, *CHRL* и *INS*. Ген *GCG* также относится к системе генов липидного обмена, и был нами отобран, основываясь на предыдущей аналогичной работе в нашей лаборатории. Гены *ABCA9* и *TPI* не относятся непосредственно к генам липидной цепочки, но вносят свой опосредованный вклад. Необходимость их присутствия была обусловлена также наличием примеров для сравнения.

В последнее время высокопроизводительное секвенирование геномной ДНК, в том числе в ходе проекта «1000 геномов» (Colonna et al., 2014), развитие подхода GWAS (Genome-Wide Association Study) для локусов (Welter et al., 2014) и построение полногеномных карт для гаплотипов (International HarMap 3 Consortium et al., 2010), ассоциированных с заболеваниями, открыли возможность учета геномов пациентов в медицинской практике, так называемой постгеномной предиктивно-превентивной персонализированной медицине, PPPM (Mallal et al., 2002; Trovato, 2014). Но десяткам миллионов неаннотированных SNPs соответствует гигантское число ложноположительных (ложноотрицательных) кандидатных SNP-маркеров, отбираемых компьютерными методами для очень дорогостоящих биомедицинских испытаний путем сравнения их частот на различных пациентах с нормой. Это способствует недооценке значимых для медицины SNPs, но имеющих статистически слабый сигнал (Liu, Xuan, 2015), например, из-за их крайне редкой встречаемости и/или их патогенной манифестации лишь в гаплотипах при некоторых сочетаниях с другими SNPs (Kaniwa et al., 2005) и бесполезных затрат на испытания нейтральных «молчащих SNPs» (Yoo et al., 2015). Поэтому предшествующие медицинским испытаниям опыты по проверке отобранных *in silico* кандидатных SNP-маркеров

могли бы исключить из их числа по крайней мере нейтральные «молчащие» SNPs. Ранее на примерах более 70 известных SNP-маркеров заболеваний человека (Ponomarenko et al., 2009), селекционно-ценных признаков животных и растений (Suslov et al., 2010a), а также 146 SNPs единственного функционального ТАТА бокса в геноме ВИЧ-1 (Suslov et al., 2010b) был разработан компьютерный метод оценки достоверности аллельных различий экспрессии гена (Пономаренко и др., 2008), а также подтверждены эти прогнозы на оригинальных экспериментах с использованием гель-ретардации (Electrophoretic Mobility Shift Assay, EMSA) при равновесных (Savinkova et al., 2013) и неравновесных (Drachkova et al., 2014) условиях *in vitro*, а также *in vitro* в режиме «реального времени» с использованием (Драчкова и др., 2012) поверхностного плазмонного резонанса (Surface Plasmon Resonance, SPR) на биосенсоре ProteOn™ XPR36 (Bio-Rad Lab., Inc., США) и переноса флуоресцентной резонансной энергии (Fluorescence Resonance Energy Transfer, FRET) (Arkova et al., 2014) и метода остановленной струи (Stopped-flow) на спектрометре SX20 (Applied Photophysics, Великобритания). Дополнительно, были подтверждены прогнозы данного метода (Пономаренко и др., 2008) на данных независимых экспериментов на генах-паралогах семейства ARF (Auxin Response Factors) арабидопсиса и риса (Mironova et al., 2010), транскриптома арабидопсиса через один час после обработки растения 1 мкМ ауксином (Ponomarenko, Ponomarenko et al., 2015), транскриптома мозга человека (Ponomarenko et al., 2014), а также 68 других опытов разных авторов (Пономаренко и др., 2010). Наконец, в завершение исчерпывающих экспериментальных проверок метода (Пономаренко и др., 2008) был создан общедоступный Web-сервис SNP ТАТА Comparator (Рассказов и др., 2013)

(<http://beehive.bionet.nsc.ru/cgi-bin/mgs/tatascan/start.pl>) для самостоятельного применения пользователями сети Интернет.

Биоинформатическая цель настоящей работы заключалась в поиске, извлечении и аннотации не исследованных однонуклеотидных полиморфизмов, расположенных в кор-промоторных районах (район от -20 до -70 нуклеотида от старта транскрипции) генов, участвующих в формировании пищевого поведения человека и липидного гомеостаза человека.

Как известно, в результате появления новых технологий секвенирования и возрастающих объемов данных о последовательностях геномов человека, возникла задача определения функционально значимых полиморфизмов и отделения их от нейтральных, особенно в регуляторных областях генов. Этой проблеме посвящено пока небольшое количество исследований и достигнуты определенные успехи (Zhao et al., 1999). Для решения этой проблемы в работе использовался разработанный ранее биоинформатический инструмент Web-сервис SNP\_TATA\_Comparator (Рассказов и др., 2013), созданный на основании уравнения равновесного связывания ТСБ/ТАТА за три последовательных шага: скольжение вдоль ДНК в силу неспецифического сродства ТСБ/ДНК ↔ опознавание ТАТА-бокса ↔ стабилизация комплекса ТСБ/ТАТА (Пономаренко и др., 2010), полученного на основании экспериментальных данных нашего сектора и литературных данных:

С помощью этого уравнения были проанализированы последовательности ТАТА-боксов, содержащих SNPs, ассоциированные с повышенным риском возникновения наследственных патологий (Ponomarenko et al., 2009) и сделаны количественные прогнозы изменения сродства ТСБ к таким SNP-содержащим ТАТА-боксам. Проведенная затем в максимально стандартизированных условиях экспериментальная верификация предсказаний сродства ТСБ/ТАТА и его изменения показала, что экспериментальные данные с высокой достоверностью коррелируют с прогнозами: коэффициенты линейной корреляции,  $r = 0.822$  для значений равновесных  $K_D$  ( $-\ln K_D$ ) и  $r = 0.785$  для изменения равновесных  $K_D$  ( $\delta$ ) при SNPs в ТАТА-боксах ( $\delta = -\ln [K_{D, \text{ТАТА Mut}}] - (-\ln [K_{D, \text{ТАТА}}])$ ), при значимости  $\alpha < 10^{-7}$  и  $\alpha < 10^{-3}$ , соответственно (Savinkova et al., 2013). Это позволило нам применить разработанный подход для характеристики неаннотированных полиморфизмов корпромоторных областей генов, ассоциированных с повышенным весом и ожирением, а также сцепленными с ними заболеваниями, и выявления среди них потенциально функционально значимых: используя уравнения равновесного связывания получить *in silico* значения  $K_D$ , характеризующие сродство ТСБ/ТАТА, и оценить изменение сродства в норме и при содержании SNP в районе ТАТА-боксов. В качестве «дикого типа» использовались последовательности промоторов референсного генома человека hg19 (проект 1000 Genomes, dbSNP).

### **Заключение по Обзору литературы**

Количество статей, посвященных изучению влияния полиморфизмов в целом и ТАТА-боксов в частности, на здоровье людей увеличивается с каждым годом. Возможность найти SNP и понять механизм ее влияния по сути открывает дверь в ме-

дицину будущего. Осознание потребности без возможности ее реализовать ведет к накоплению данных, их описанию и систематизации. Современные биоинформатические технологии, как например мета-анализ GWAS, пока не подходят для выявления низкочастотных аллелей рецессивных моделей из-за статистической слабости сигнала. Дополнительный инструмент позволит и выявить и предсказать возможное фенотипическое проявление полиморфизма. В этом плане выявление единичных SNPs классическими методами не теряет своей актуальности.

## **2. Материалы и методы**

### **2.1. Формирование библиотеки неаннотированных SNPs TATA-боксов промоторов генов липидного метаболизма человека**

Формирование электронной библиотеки и компьютерный анализ неаннотированных однонуклеотидных полиморфизмов TATA-боксов промоторов генов липидного гомеостаза человека для прогноза потенциально функционально значимых сделаны при участии Пономаренко М.П.

Электронную библиотеку неаннотированных SNPs генов липидного гомеостаза человека составляли с использованием базы данных dbSNP NCBI (<http://www.ncbi.nlm.nih.gov/SNP/>), которая содержит информацию обо всех известных на сегодняшний день полиморфизмах в геноме человека, а также реферирования оригинальных статей, индексированных в базе данных PubMed NCPI, поскольку в нашей библиотеке есть SNPs, которые до сих пор еще не были внесены в dbSNP. Библиотека включала более 1200 SNPs, включая SNPs генов липидного гомеостаза человека. С использованием Web-сервиса SNP\_TATA\_Comparator (Рассказов и др., 2013), который анализирует независимо референсные и анцестральные (индивиду-

альные) последовательности ТАТА-элементов, были проанализированы все последовательности неаннотированных SNP-содержащих ТАТА-боксов и сделаны прогнозы значимого ( $\alpha < 10^{-7}$ ) изменения сродства ТСБ в результате содержания SNPs. Координаты промоторов всех исследуемых генов в референсном геноме человека (hg19) были взяты из Eukaryotic Promoter Database (EPD) (Dreos et al., 2015) dbSNP, NCBI. Между позициями от  $-70$  до  $-20$  относительно TSS этих промоторов расположен район экспериментально доказанных сайтов ТСБ-связывания (Ponomarenko et al., 2013a, b). Гены со значимым (по прогнозу) изменением сродства были использованы для дальнейшего экспериментального анализа (LEP, ABCA9, GCG, CYP2A6, TPI), с незначимым – исключены.

## 2.2. Нарботка, выделение и очистка ТСБ

Для данной работы использовался полноразмерный ген ТСБ человека, полученный из клеток *E.coli* BL21(DE3) pAR3038. Штамм *E.coli* BL21(DE3) был любезно предоставлен д.б.н. Д.О. Жарковым. Плазмида AR3038 была любезно предоставлена В.Ф. Pugh (США) и содержала полноразмерный ген ТСБ человека с нативной аминокислотной последовательностью.

Приготовление компетентных клеток проводили по следующей методике (Inoue et al., 1990). Со стока ( $-70^{\circ}\text{C}$ ) рассеивали бактерии штрихом на чашке с селективной средой, выращивали при  $37^{\circ}\text{C}$ . Несколько колоний (10-12) диаметром  $\sim 2-3$  мм поместили в 40 мл SOB (Bactotryptone 20 г/л, Vacto Yest extr. 5,5 г/л, NaCl 2 мМ, KCl 10 мМ,  $\text{Mg}^{+2}$  10 мМ) в колбу на 1.0 литр, колбу интенсивно встряхивали (300 r/pm) при  $18^{\circ}\text{C}$  до  $\text{OD}_{600} = 0.6$ .

Далее на холоду 30 мл культуры инкубировали 15 минут при 0°C; центрифугировали при 3 kr/pm 4°C 10 минут, после того как слили супернатант еще раз центрифугировали при 3 kr/pm 4°C 20 секунд. После удаления остатков жидкости осадок суспензировали в 10 мл ТВ (Hepes 10мМ, MnCl<sub>2</sub> 55 мМ, CaCl<sub>2</sub> 15 мМ, KCl 250 мМ, рН = 6.7) буфере. Инкубировали 15 минут при 0°C. Затем опять центрифугировали при 3 kr/pm 4°C 10 минут, сливали супернатант и еще раз центрифугировали при 3 kr/pm 4°C 20 секунд. После удаления остатков жидкости осадок ресуспензировали в 2 мл ТВ. Добавляли DMSO до 3.5% (70 мкл) и инкубировали 15 минут при 0°C. Еще раз добавляли 70 мкл DMSO и после инкубирования 15 минут при 0°C клетки были разаликвачены по 100 мкл и заморожены в жидком азоте.

При трансформации использовали стандартный протокол 60 минут, потом при температуре 42°C 30 секунд (все без встряхивания). Затем поставили в лед на 2 минуты и добавили после этого 400 мкл SOC (Bactotryptone 20 г/л, Vacto Yest extr. 5,5 г/л, NaCl 2 мМ, KCl 10 мМ, Mg<sup>+2</sup> 10 мМ, glucose 10 мМ). Плотнo завинтили крышку, встряхнули горизонтально на бактериальном шейкере при температуре 37°C 30 минут и выселили 50 мкл на чашку с антибиотиком.

### **Выделение ТАГА-связывающего белка человека**

Выращивание бактерий *E.coli* проводилось согласно стандартной методике по В.Ф. Pugh (Pugh, 1995) с некоторыми модификациями:

1. Концентрация IPTG, используемого для индукции *lac*-оперона, была изменена с 0,1 мМ до 1 мМ.
2. Время индукции составило 3 часа вместо 90 минут.



3. Заменено осаждение нуклеиновых кислот полиэтиленимином (PEI) на хроматографию на смоле DEAE-sephacel.
4. Из растворов был убран детергент NP40, необходимый для подавления неспецифической агрегации белков.
5. PEI был заменен на смолу DEAE-sephacel во избежание выпадения ТСБ в осадок вместе с НК на стадии осаждения PEI.
6. Для выращивания бактерий использовались среды LB и 2YT.
7. Для селективного отбора бактерий, содержащих целевую плазмиду, в питательные среды был добавлен антибиотик ампициллин.

Для наработки ТСБ клетки *E.coli* BL21(DE3), несущие рекомбинантную плазмиду, высевали с замороженного стока на чашку с LB-агаром, содержащим 200 мкг/мл ампициллина, и выращивали при 37<sup>0</sup>С 12 часов. 200 колоний со свежей чашки переносили в колбу с 10 мл свежей LB-среды, содержащей 200 мкг/мл ампициллина и 0,4 % глюкозу и инкубировали при 37<sup>0</sup>С (при перемешивании на скорости 280 об/мин) до оптической плотности  $A_{600}=0,5-0,8$ . После инкубации культуру клеток переносили на среду YT в соотношении 1 мл на 70 мл среды, содержащей 200 мкг/мл ампициллина и 0,4 % глюкозу. Культуру выращивали также при 37<sup>0</sup>С при перемешивании на скорости 280 об/мин до оптической плотности  $A_{600}=0,7-0,8$ . После этого температуру понижали до 30<sup>0</sup>С и индуцировали синтез рекомбинантного белка добавлением IPTG до концентрации 1 мМ, одновременно добавляя ампициллин до концентрации 200 мкг/мл. Экспрессию вели 3 часа при 30<sup>0</sup>С при перемешивании на скорости 280 об/мин на качалке. Полученную культуру клеток центрифугировали при 5000g в течение 10 минут при 2<sup>0</sup>С, супернатант

отбирали, а клеточный осадок ресуспендировали в 70 мл охлажденного буфера для лизиса, содержащего 25 мМ Нерес-КОН (рН=7,6), 12,5 мМ  $MgCl_2$ , 0,2 мМ КСl, 10 % глицерин, 0,1 мМ ЕДТА, 1 мМ DTT и 1 мМ PMSF. Полученную смесь центрифугировали в тех же условиях, супернатант отбирали, а полученный осадок ресуспендировали в 10 мл лизисного буфера. Выход клеток со 640 мл культуры составил около 4 грамм. Клетки замораживали и хранили в жидком азоте.

Выделение ТСБ проводилось по 3-х шаговой методике В.Ф. Pugh. *Первый шаг* – удаление НК и ассоциированных с ними белков проводили на колонке со смолой DEAE- sephacel, где за основу был взят метод ионообменной хроматографии. На *второй стадии* проводилось связывание ТСБ с фосфоцеллюлозой и отмывка его высокосолевСБ буфером. На *третьей стадии* проводилось сульфатаммонийное осаждение ТСБ.

*Первая стадия:* Для выделения белка в бактериальную суспензию добавляли лизоцим до концентрации 1,5 мг/мл, инкубировали 15 минут при  $0^{\circ}C$  и озвучивали ультразвуком 15 минут при  $0^{\circ}C$  в режиме: 15 секунд - звук, 40 секунд – пауза. Озвучивание проводилось при 80 Вт мощности до потери вязкости (суммарное время озвучивания составило около 3 минут). Лизированные клетки центрифугировали 20 минут при 14000g при  $4^{\circ}C$ . Супернатант отбирали и разбавляли в 2 раза 0,2 М лизисным буфером. Добавив по каплям при перемешивании 4М КСl и доведя концентрацию КСl до 0,3 М, супернатант наносили на колонку с 40 мл сорбента DEAE-sephacel. В этих условиях нуклеиновые кислоты и ассоциированные с ними белки связываются с DEAE, а ТСБ – нет.

*Вторая стадия:* В проскок добавляли 4М КСl, увеличив концентрацию КСl до 0,35 М, и наносили на колонку с фосфоцеллюлозой P11. Объем колонки составил

26 мл. Промывка колонки осуществлялась 120 мл буфера H0.35, содержащим 20 mM Hepes-KOH (pH=7,6), 2 mM MgCl<sub>2</sub>, 0,35 M KCl, 10 % глицерин, 50 mM EDTA, 1 mM DTT, 1 mM PMSF. Элюирование ТСБ проводилось буфером H1, содержащим те же компоненты, что и буфер H0.35, за исключением того, что KCl находился в концентрации 1 M.

*Третья стадия:* После сбора необходимых фракций производилось высаливание ТСБ насыщенным раствором сульфата аммония: 250 мкл/ 1 мл белка. Осадок собирали центрифугированием при 14000g в течение 120 минут при 4 °С и растворяли в буфере хранения, содержащем 20 mM Трис-HCl (pH 7,6), 2 mM MgCl<sub>2</sub>, 20 % глицерин, 0,1 mM EDTA, 0,2 M глютамат калия, 1 mM DTT, 1 mM PMSF.

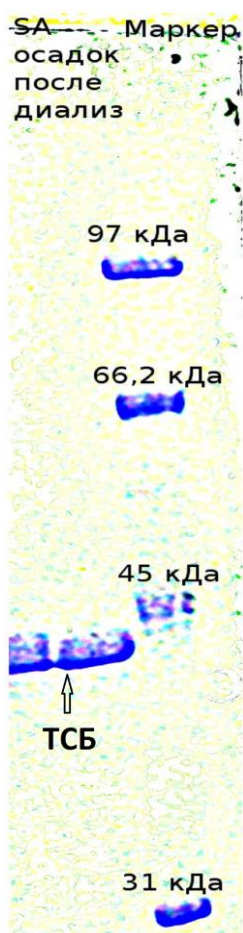
После диализа при 0°С в течение 4 часов ТСБ, растворенный в буфере хранения, центрифугировали при 14000g в течение 5 минут при 10 °С для удаления денатурированного белка. Супернатант, содержащий ТСБ, замораживали.

Концентрация выделенного белка в растворе определялась при помощи метода Брэдфорд, описанного в (Скоупс, 1985). Этот метод даёт допустимо точное значение концентрации белка в пределах от 2 мкг/мл до 120 мкг/мл, так как в этих границах соблюдается линейная зависимость увеличения абсорбции от концентрации. Сам метод основан на реакции красителя (Coomassie Brilliant Blue G-250) и остатками. Связанная форма имеет голубую окраску с максимумом поглощения при 595 нм. Таким образом, увеличение оптической плотности раствора при длине волны, равной 595 нм, пропорционально количеству белка в растворе.

Метод определения концентрации белка по Брэдфорд требует построения стандартной калибровочной кривой перед измерением концентрации ТСБ. Необходимые калибровочные кривые строили по BSA.

## Электрофорез по Лэммли

На каждом этапе выделения и очистки ТСБ проводили электрофоретический анализ белка. Электрофорез проводили в 10 % ПААГ по Лэммли (Laemmli, 1970) (Рис. 5). Для этого к 20 мкл каждого образца добавляли 10 мкл буфера, содержащего: 62,5 мМ Трис-НСl, рН 6.8; 2% ДДС; 10% глицерина; 5%  $\beta$ -меркаптоэтанола и 0,01% бромфенолового синего. После прогрева в течение 5 мин в кипящей водяной бане для денатурации белка образцы разделяли в ступенчатом ПААГ по методу Лэммли (Laemmli, 1970). Состав концентрирующего геля: 125 мМ Трис-НСl, рН 6.8; 0,1% ДДС; 4% акриламид–бисакриламид (29:1); 0,05 % ПСА; 0,15 % ТЕМЕД;



**Рисунок 5.** Результаты электрофореза. Мол.вес ТСБ 38 кДа

разделяющий гель: 375 мМ Трис-НСl, рН 8.9; 0,1% ДДС; 10% акриламид–бисакриламид (30:1); 0,1 % ПСА; 0,15 % ТЕМЕД. Верхний электродный буфер содержал: 25 мМ Трис; 192 мМ глицин; 0,1% ДДС; нижний электродный буфер: 0,05 М Трис-НСl, рН 7.5; 0,1% ДДС. Электрофорез проводился в вертикальных камерах при комнатной температуре и напряжении 10 В/см до тех пор, пока краситель не достигал нижнего края геля.

После завершения электрофореза для визуализации белков гель окрашивали 10 мин раствором, содержащим 45% метанола, 10% уксусной кислоты, 0,25 г (на 100 мл раствора для окрашивания) Coomassie Brilliant Blue G-250. После этого гель отмывали от избытка краски горячей дистиллированной водой.

В качестве одного из вариантов окрашивания гелей проводили окрашивание геля серебром. Гель для обезвоживания оставляли на

ночь в 50% растворе этанола, после чего выдерживали его в течение 30 мин при постоянном помешивании в растворе, содержащем 0,02 М NaOH, 0,3% NH<sub>4</sub>OH и 0,4% AgNO<sub>3</sub>. Затем гель помещали на 10 мин в дистиллированную воду и проявляли в растворе, содержащем 0,01% лимонную кислоту и 0,035% формалин; реакцию останавливали 10% уксусной кислотой.

Концентрацию ТСБ определяли спектрофотометрически относительно раствора с известной концентрацией BSA, растворенном в буфере хранения ТСБ по методу Брэдфорд.

### **Анализ активности белка**

Специфическая, ДНК связывающая активность ТСБ была определена, как описано в (Rashidzadeh et al., 2003).

ТСБ в фиксированной концентрации титровали двуцепочечным ОДН AdML, содержащим канонический ТАТА-бокс, в возрастающей концентрации при 25<sup>0</sup>С 2 часа в 1-кратном буфере связывания. AdML использовали в концентрации от 2 до 192 нМ. Выделен гомогенный по данным электрофореза (Рис. 5), и нативный ТСБ. Анализируя полученные данные, было посчитано, что из 1,4 л культуры было выделено 532 мкг ТСБ в концентрации 0,76 мкг/мкл. Проведенный анализ показал, что ТСБ содержит 38 % активного белка.

### **2.3. Получение меченых <sup>32</sup>P олигодезоксирибонуклеотидов**

В работе использовали олигодезоксирибонуклеотиды (ОДН), синтезированные и дополнительно очищенные электрофорезом в ПААГ ООО «Биосинтез» (Новосибирск). Для получения меченого ОДН цепь метили <sup>32</sup>P. Для этого метку

[ $\gamma$ -<sup>32</sup>P] АТФ (рабочая концентрация 1МБК/мкл) добавляли к олигонуклеотидам в следующих концентрациях: 250, 500 или 750 пМ. Ферментативный перенос метки Р<sup>32</sup> с АТФ на 5' конец ОДН осуществлялся с помощью Т4 полинуклеотидкиназы («SibEnzyme», Новосибирск). Кинирование проводилось в 1-кратном SE-буфере (рН 7.6, 50 мМ Tris-HCl, 10 мМ MgCl<sub>2</sub>, 5 мМ DTT) в течение 1 часа при 37<sup>0</sup>С. Двухцепочечные олигонуклеотиды получали с помощью отжига комплементарных одноцепочечных олигонуклеотидов. Отжиг цепей после кинирования проводили в том же буфере, что и для кинирования (путем нагревания реакционной смеси до 98<sup>0</sup>С и последующем ее медленном, со скоростью 1<sup>0</sup>С/мин, охлаждении до комнатной температуры).

Отожженные дуплексы анализировали в 15% ПААГ в неденатурирующих условиях с последующей радиоавтографией на фосфоримиджере *Molecular Imager PharoFX Plus* (Bio-Rad). Для этого проводили вертикальный электрофорез в 15% ПААГ (при соотношении акриламида к бисакриламиду как 39 к 1) в буфере 1x TBE (10X рН~8.3, 0.89 М Tris base, 0.89 М борной кислоты, 20 мМ EDTA) при комнатной температуре и напряженности поля 15 В/см в течение 2-3 часов (проводя остановку электрофореза по визуальному контролю за пробой, содержащей дуплекс без метки, но с добавлением красителя бромфенолового синего). Гель экспонировали на экране *Imaging Screen-K* (Kodak) для фосфоимиджера. Затем экран сканировали на фосфоимиджере и осуществляли анализ радиоавтографов с помощью программы *Quantity One – 4.5.0* (Bio-Rad). Чтобы отделить 26-мерные двухцепочечные ОДН от неотожженных цепей и от ОДН, в которые метка не вставилась, гель сопоставляли с проекцией, полученной с экрана для фосфоимиджера, с дальнейшим вырезанием участков геля, содержащих дуплексы с меткой [ $\gamma$ -<sup>32</sup>P]АТФ.

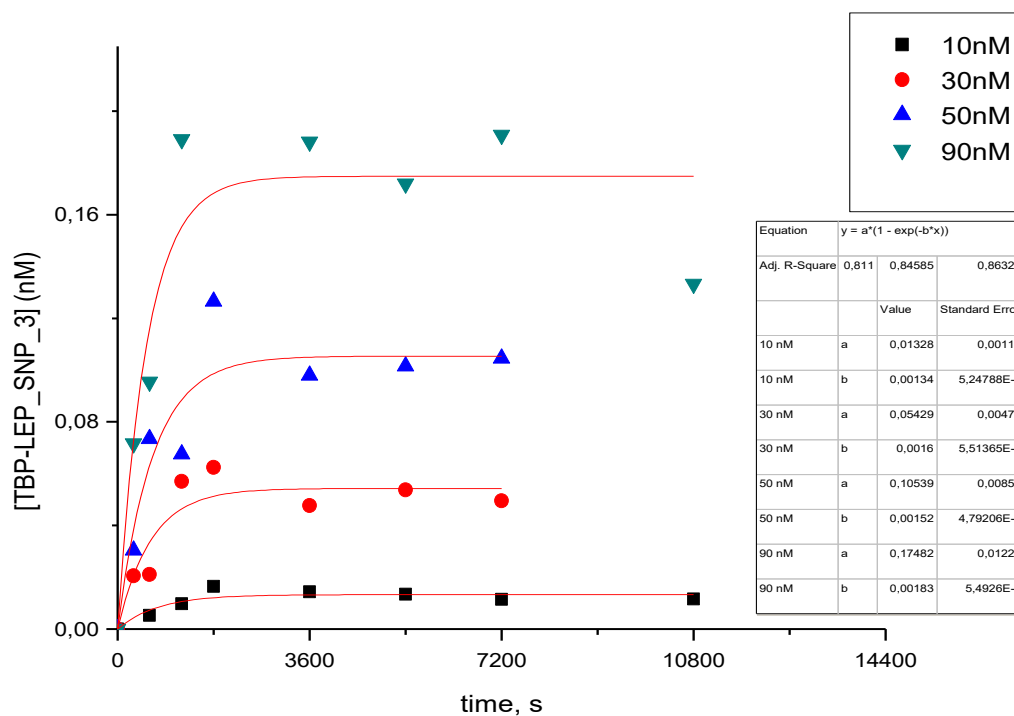
Для элюирования олигонуклеотидов из полиакриламидного геля фрагменты геля с мечеными ОДН помещали в карманы 0,7% агарозного геля и проводили горизонтальный электрофорез в буфере 1x TAE (40 мМ Трис, 20 мМ уксусная кислота, 1 мМ EDTA) при комнатной температуре и напряженности поля 10 В/см в течение 10-15 минут (проводя остановку электрофореза по визуальному контролю за пробой, содержащей дуплекс без метки, но с добавлением красителя бромфенолового синего). После электрофореза производили вырезание полосок геля, расположенных под карманами, разрезали их поперек на фрагменты толщиной около 1 мм. Полученные пластинки анализировали на счетчике Гейгера, отбирая те, которые насчитывали более 400 импульсов в секунду, в центрифужные концентраторы Spin-X с диаметром пор 0,45 мкм. Из агарозного геля меченые ОДН выделяли путем центрифугирования при 14000g 15 минут при комнатной температуре. Небольшие фрагменты агарозы, прошедшие через фильтр вместе с ОДН при дальнейшем переосаждении будут использоваться в качестве соосаждителей. Для осаждения двуцепочечных ОДН добавляли  $MgCl_2$  до 10 мМ, NaAC до 150 мМ и 96% EtOH (2,5 объема реакционной смеси). Инкубацию проводили при  $-20^{\circ}C$  в течение ночи. Затем смесь центрифугировали при комнатной температуре 15 минут при 15000g. Супернатант отбирали, а осадок высушивали. Получившийся осадок растворяли в 1-кратном буфере отжига (10 мМ Hepes-KOH pH 7,6, 20 мМ KCl, 0,1 мМ EDTA) в течение 1 часа.

## 2.4. Определение кинетических и термодинамических характеристик комплексов ТСБ/ОДН методом задержки ДНК в геле

Для определения равновесных констант диссоциации комплексов ( $K_D$ ) ТСБ с ТАТА-содержащими олигонуклеотидами, соответствующими «нормальным» и SNP-содержащим вариантам ТАТА-боксов, использовали традиционный подход, включающий титрование ТСБ в постоянной концентрации (не менее 0,3 нМ и не более 1 нМ) ТАТА-содержащим ОДН в возрастающей концентрации в сочетании с методом изотопного разбавления. Предварительно определяли время установления равновесного состояния по кинетике связывания ТСБ с каждым ОДН.

Эксперименты по изучению кинетики связывания ТСБ с ОДН проводили при постоянной концентрации ТСБ и варьировании концентрации ОДН во времени до установления равновесия при 25°C в буфере связывания следующего состава: 20 мМ Hepes-KOH (pH=7.6), 5 мМ MgCl<sub>2</sub>, 50 мМ KCl, 1 мМ DTT, 100 мкг/мкл BSA, 0.01 % NP-40, 5 % глицерин. Для отделения комплексов "ТСБ/ОДН" от несвязавшегося ОДН использовали метод «задержки ДНК в геле» (EMSA). Электрофорез проводили в 5% ПААГ на трис-глициновом буфере (pH 8.3) при температуре 10° С и напряженности поля 25 В/см в течение 40 мин. Стекла для электрофореза предварительно подвергали силиконизированию. Все пробы наносили одновременно на идущий электрофорез при напряженности поля 10 В/см. Перед электрофорезом с пробами проводили предварительный холостой электрофорез в течение 1 часа при 220 В в охлажденном буфере TGB.





**Рисунок 6.** Кинетика ассоциации ТСБ с ТАТА-содержащим участком промотора *LEP* с SNP rs201381696 (ТСБ-*LEP\_SNP\_3*)

На Рис.6 изображена кинетика ассоциации ТСБ с ТАТА-содержащим участком промотора *LEP* с SNP rs201381696 (кинетические кривые взаимодействия ТСБ/ТАТА для промоторов других генов представлены в Приложении 1 Таблица 1).

Гели высушивали и экспонировали с экраном Imaging Screen-K (Kodak) для фосфоимиджера *Molecular Imager PharosFX Plus* (Bio-Rad). Затем экран сканировали на фосфоимиджере и осуществляли количественный анализ радиоавтографов с помощью компьютерной программы Quantity One – 4.5.0 (Bio-Rad).

Равновесную константу диссоциации комплексов ( $K_D$ ) определяли как концентрацию субстрата, необходимую для образования половины от максимального

количества комплексов.  $K_D$  отражает аффинность или сродство ОДН к ТСБ, поэтому чем больше  $K_D$ , тем меньше аффинность. Анализ данных, полученных методом EMSA, проводился графически с помощью компьютерной программы OriginPro8. Кривая, отображающая равновесное связывание, описывается уравнением:

$$y = P1 \cdot x / (P2 + x),$$

где коэффициент P1 - максимальные концентрации комплекса, а коэффициент P2 =  $K_D$  - равновесная константа диссоциации. Время взаимодействия выбрано на основании предварительных экспериментов, исходя из того, что к этому времени уже наступает равновесие.

Равновесное взаимодействие ТСБ человека с ТАТА-боксами промоторов генов человека описывают следующим образом:



где:

$$K_D = [\text{ТСБ}/\text{ТАТА}] / ([\text{ТСБ}] \cdot [\text{ТАТА}]) = k_{\text{off}} / k_{\text{on}},$$

$K_D$ ,  $M$  - равновесная константа диссоциации комплекса,

$k_{\text{on}}$  ( $\rightarrow$ ),  $M^{-1} \cdot C^{-1}$  - константа скорости связывания ТСБ/ТАТА,

$k_{\text{off}}$  ( $\leftarrow$ ),  $C^{-1}$  - константа скорости диссоциации комплекса ТСБ/ТАТА.

Наблюдая за взаимодействием ТАТА с ТСБ, мы изучаем ответ образовавшейся системы в виде концентрации комплекса ТСБ/ОДН, зависимость образования которого имеет гиперболический характер и представляет собой уравнение изотермы Ленгмюра (оно применимо в тех случаях, когда концентрация ОДН значительно превышает концентрацию связывающего его белка) (Кнорре, Мызина, 1992).

Для определения констант связывания и диссоциации комплексов ТСБ/ОДН снимали зависимости образования комплексов (изотермы связывания) во времени от концентрации ОДН при постоянной концентрации ТСБ (0.3 нМ). Эксперименты проводили при 25<sup>0</sup>С в связывающем буфере следующего состава: 20 мМ HEPES-КОН (рН 7.6), 5 мМ MgCl<sub>2</sub>, 70 мМ KCl, 1 мМ DTT, 100 мкг/мл BSA, 0.01% NP-40 и 5% glycerol. Комплексы ТСБ/ОДН отделяли от несвязавшихся ОДН методом задержки ДНК в геле. Каждый эксперимент по определению кинетики связывания обычно состоял из 32 реакций связывания (8 временных точек с 4 концентрациями ОДН). Все 4 реакции связывания наносили на гель одновременно. Константы скоростей,  $k_a$  и  $k_d$ , определяли с помощью программы GraphPad Prism 5 software (Equation: Association kinetics (two ligand concentrations)) на основе зависимости изотерм связывания ТСБ/ОДН от концентрации ОДН при 0,3 нМ ТСБ. Из полученных экспериментальных величин  $k_a$  и  $k_d$  оценили величину кажущейся равновесной константы диссоциации ( $K_D = k_d / k_a$ ), характеризующую аффинность ТСБ/ОДН, а также время полураспада комплексов ( $t_{1/2} = \ln 2 / k_d$ ) и изменение свободной энергии Гиббса ( $\Delta G = -RT \ln K_A$ , где R – универсальная газовая константа, T – абсолютная температура,  $K_A = k_a / k_d$ ).

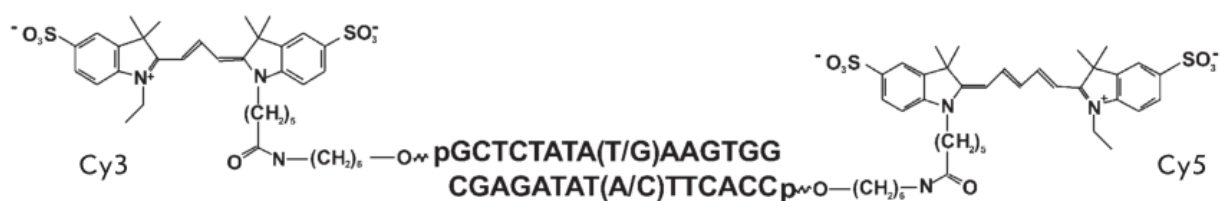
## **2.5. Определение кинетических констант комплексов методом флуоресцентной спектроскопии в режиме реального времени**

Для наблюдения связывания ТСБ с ОДН и определения основных характеристик процесса было проведено несколько экспериментов с ОДН, мечеными различными флуорофорами (Приложение 1; Приложение 2) - парами красителей

(Cu3/Cu5, TAMRA/FAM). Связывание ТСБ с ДНК-дуплексами, соответствующими нормальному ТАТА-боксу и ТАТА-боксу с SNP анализировали методом «остановленной струи» (stopped-flow) на одноименном спектрофотометре SX20 (Applied Photophysics, Великобритания). Мертвое время прибора равно 1,38 мс. Для проведения экспериментов были выбраны stopped-flow флуориметры, так как они обеспечивают минимальную задержку в измерении флуоресценции растворов после смешивания. Для правильной постановки экспериментов были предварительно проведены работы по изучению оптимальных условий для связывания ТСБ/ТАТА (Приложение 3). В качестве контроля проводили следующие эксперименты: 1) ОДН без ТСБ, 2) ОДН без ТАТА с ТСБ.

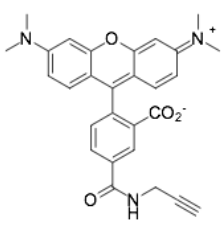
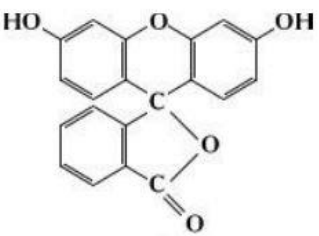
Вначале был проведен эксперимент по связыванию ТСБ с ОДН длиной 15 и 26 п.н., меченных по 5'-концам цианиновыми флуорофорами Cu3/Cu5 (Рис. 7), которые были синтезированы и очищены в ООО «НаноТех-С», Новосибирск. Связывание ТСБ с ДНК-дуплексом, соответствующим нормальному ТАТА-боксу промотора гена TPI (gctctAtATAAgtgg, аллель Т и gctctAtAGAAgtgg, аллель G), анализировали методом «остановленной струи». Длина волны возбуждения флуоресценции красителя Cu3 равна 550 нм, напряжение на детекторе – 575 В. Флуоресценцию красителя Cu5 регистрировали на длинах волн более 645 нм при использовании светофильтра rG-645 (Scott, Германия).

Концентрация ДНК-дуплексов во всех случаях была равна  $1 \cdot 10^{-7}$  М. Процесс связывания с ДНК-дуплексами изучали с использованием активного ТСБ, концентрация которого возрастала по следующей схеме: для аллеля Т – в 1, 2, 4, 6, 8, 10, 20 раз ( $\cdot 10^{-7}$  М); для аллеля G – в 1, 4, 6, 8, 10, 20, 30, 40 раз ( $\cdot 10^{-7}$  М). Время измерения составляло – 50 с, общее количество точек на одну кривую – 6000.



**Рисунок 7.** Используемые в работе ДНК-дуплексы, идентичные последовательности ТАТА-боксов промотора гена *TPI* дикого типа и содержащего SNP -24T > G, меченные FRET-парой флуорофоров Cy3/Cy5 (Приложение 4).

Олигодезоксинуклеотиды (ОДН) длиной 15 и 26 п.н., меченные по 5'- концам флуоресцентными красителями TAMRA и FAM (Рис. 8), были синтезированы и очищены ("BIOSYN", Новосибирск).

<p><b>TAMRA</b> (TMR, тетраметилродамин)</p>  <p>- ксантеновый краситель, образующий FRET-пару с флуоресцеином.</p>	<p><b>FAM</b> (флуоресцеин, диоксифлуоран)</p> 
--	--

**Рисунок 8.** Структура FRET-пары TAMRA/FAM, использованной для мечення изучаемых дуплексов (Приложение 4).

Последовательности используемых в работе двуцепочечных олигодезоксинуклеотидов приведены в **Таблице 3**. Эту пару флуорофоров мы использовали для более детального изучения изменений структуры ТАТА-содержащих ОДН при взаи-

модействии с ТСБ, так как у TAMRA и FAM квантовый выход и чувствительность детекции на много выше, чем у Cy3/Cy5 (Приложение 4).

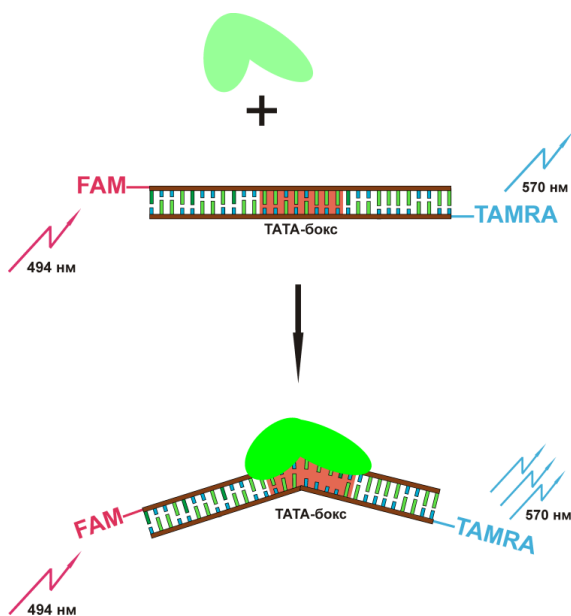
**Таблица 3.** Используемые в работе ДНК-дуплексы, меченные FRET-парой флуорофоров TAMRA/FAM.

<i>Сокращенное название</i>	<i>Последовательность (5' – 3')</i>
<b>Control-15</b>	cctgtccactctcct
<b>LEP-15</b>	ccgcTATAAGAgggg
<b>LEP-SNP-15</b>	ccgcT <b>G</b> TATAAGAgggg
<b>TPI-15</b>	ctcTATATAAgtggg
<b>TPI-SNP-15</b>	ctcTATA <b>G</b> AAgtggg
<b>Control-26</b>	cctgtcccctctccggagctcgagct
<b>LEP-26</b>	atcgggccgcTATAAGAggggccccgc
<b>LEP-SNP-26</b>	atcgggccgcT <b>G</b> TATAAGAggggccccgc
<b>TPI-26</b>	cgcggcgctcTATATAAgtgggcagt
<b>TPI-SNP-26</b>	cgcggcgctcTATA <b>G</b> AAgtgggcagt

Связывание ТСБ с ДНК-дуплексом при Cy3/Cy5 и при TAMRA/FAM проводили при 25<sup>0</sup>С в «буфере связывания», содержащем 20 мМ Hepes-KOH (pH 7.6), 5 мМ MgCl<sub>2</sub>, 50 мМ KCl, 1 мМ DTT, 100 мкг/мл BSA и 5% глицерин.

Длину волны возбуждения флуоресценции красителя FAM устанавливали равной 494 нм. Флуоресценцию красителя TAMRA регистрировали на длинах волн более 570 нм с помощью установки светофильтра OG-570 (Schott Filters, Germany). Регистрацию флуоресценции (Рис. 9) производили при напряжении 430 вольт и температуре 25<sup>0</sup>С, общее количество точек на одну кривую – 8000, причем первые 4000 точек приходились на измерение флуоресценции в течение первой секунды, а оставшиеся 4000 точек – на следующие 100 секунд.

Рабочая концентрация ДНК-дуплексов во всех случаях была равной  $1 \cdot 10^{-7}$  М, концентрация активного ТСБ на начальном этапе была такой же, а потом возрастала пропорционально в 2, 4, 6, 8 и 10 раз. Для анализа полученных данных и построения кинетической модели взаимодействия ТСБ с дуплексами ДНК, а также для расчета констант скоростей всех элементарных реакций использовали программу Dynafit (BioKin, USA) (Kuzmic, 1996). В основе построения кинетической модели взаимодействия ТСБ с дуплексами ДНК использовали принципы построения минимальной кинетической модели (Приложение 5).



**Рисунок 9.** Схематический рисунок, поясняющий экспериментальный подход к изучению излучения флуоресценции в паре красителей TAMRA/FAM.

## 2.6. Разработка плазмидных конструкций для определения активности репортерного гена *LUC*

### 1. Выбор объекта

Для WT (дикий тип) и при наличии SNP в TATA боксах промоторов были разработаны конструкции, содержащие вставки 200-250 пн (из расположенной наиболее близко к TSS части промотора изучаемого гена) в качестве минимального промотора (He et al., 1995) в беспромоторный вектор pGL4.10[luc2] (Рис. 10), несущий ген светлячковой люциферазы.

### 2. Дизайн праймеров

Праймеры выбирались по принципу создания минимального рабочего промотора. Для этого проектировался ампликон из TATA-содержащего участка промотора, фланкированный сайтами рестрикции для HindIII (AAGCCT) и XhoI (CTCGAG) (кроме гена *Leptin* – у него ампликон больше вставки, потому что сайт рестрикции для XhoI уже есть в последовательности промотора, но праймер наиболее подходящий по физико-химическим показателям находится дальше сайта) (Табл. 4).

**Таблица 4.** Дизайн праймеров.

<i>Рестриктаза</i>	<i>Сайт рестрикции</i>
HindIII	5'-A  A G C T T-3' 3'-T T C G A  A-5'
XhoI	5'-C  T C G A G-3' 3'-G A G C T C-5'

Праймеры

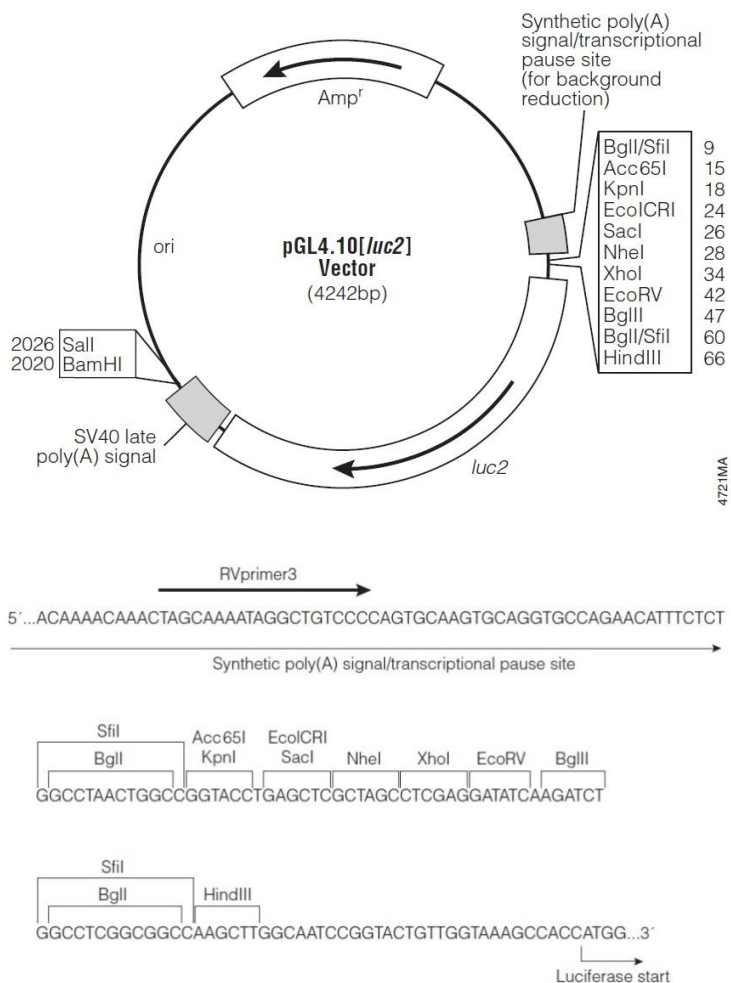
<i>Ген</i>	<i>SNP, аллель</i>	<i>ДНК последовательность</i>
------------	--------------------	-------------------------------



LEP	rs 201381696, A-35g	<i>Содержащая ТАТА-бок</i> atcgggccgc <b>TA/gTAAG</b> aggggcgggc
	Длина ам- пликона 199 пн	<i>ForwardPrimer, 5'-&gt; 3'</i> CAAGGGTGC GCGCGTGG
	Длина встав- ки 158 пн	<i>ReversePrimer Norma, 5'-&gt; 3'</i> CCAAGCTTGCCCGCCCCTCTTATAGCG
		<i>ReversePrimer SNP, 5'-&gt; 3'</i> CCAAGCTT- GCCCGCCCCTCTTACAGCG
<b>Расположение SNP, праймеров и сайтов рестрикции на ДНК</b>		
<p>5' -</p> <p><u>CAAGGGTGC GCGCGTGG</u>CTCCTGGCGCGCCGAGGCCCTCC<u>CTCGAG</u>GCCC  CGCGAGGTGCACACTGCGGGCCCAGGGCTAGCAGCCGCCCGGCACGTGCTA  CCCTGAGGGGCGGGGCGGGAGCTGGCGCTAGAAATGCGCCGGGGCCTGCGG  GGCAGTTGCGCAAGTTGTGATCGGGCCGCT<b>ATAAG</b>AGGGGCGGGC<u>AAGCTT</u>  GG - 3' 3'-</p> <p><u>GCGATATTCTCCCGCCCGTTCGAACC</u> - 5'</p>		
СYP2A6	rs 28399433, T-34g	<i>Содержащая ТАТА-бок</i> tttcaggcag <b>TAT/gAAAG</b> gсааассас
	Длина ам- пликона 261 пн	<i>ForwardPrimer, 5'-&gt; 3'</i> gcctcgagTCCAAACTCCTCAGTTCTACAGC
	Длина встав- ки 261 пн	<i>ReversePrimer Norma, 5'-&gt; 3'</i> ccaagctTGGTTTGCCTTTACTGCCTGA
		<i>ReversePrimer SNP, 5'-&gt; 3'</i> ccaagctTGGTTT- GCCTTTGACTGCCTGA
<p>5'- GGGGTGGTTTGCCTTT<b>ACTGCCTGA</b>AAAAAGAGGGATGGACTTT-  GGCTGATTACATAATTACCTCATTTACCTCCCAACTACATGCCCCAC-  CATGAATCCTGCCTAGCTTGGGACACAGCCAGCAAGGGAGGATAAGGG-  GACCCAGGGCAGCTGAACAGAGAGGGGGCCTCCAGACTAAATCTGTGG-  TACTTCAGGAGGGGTGCCGCAGGGCTGTGGATTTAGGAGGGGGCAACAGA-  TAAGCTGTAGAACTGAGGAGTTTGGAAATATTTGCATAGGGGAGCACTTGG -  3' 3'- <u>CGACATCTTGACTCCTCAAACCT</u> - 5'</p>		

Таким образом было смоделировано, что обратный праймер будет содержать ТАТА-бок, а промоторная область изучаемых генов будет интегрирована в поло-

жение -34 пн от начала транскрипции гена люциферазы. При этом SNP гена *Lep* будет находиться на расстоянии -50 пн от TSS, SNP гена *CYP2A6* - 46 пн.



**Рисунок 10.** Структура беспромоторного вектора pGL4.10 ([promega.com/vectors/](http://promega.com/vectors/))

### Последовательности праймеров

Пары праймеров (Табл. 5) подбирали с использованием NCBI Primer-BLAST. Синтез олигонуклеотидов осуществлялся компанией ООО «БИОССЕТ», Новосибирск.

**Таблица 5.** Последовательности специфических прямых и обратных пар праймеров, использованных в работе для проведения ПЦР. Выделен зеленым сайт рестрикции XhoI, подчеркнут и выделен зеленым сайт рестрикции Hind III, красным цветом обозначена норма и SNP.

Обозначение специфических праймеров	Последовательность 5'→3'	Длина ПЦР-продукта, п.н.
LepFor LepRevN LepRevSNP	CAAGGGTGC GCGCGTGG CC <u>AAGCTT</u> GCCCGCCCCTCTTATAGCG CC <u>AAGCTT</u> GCCCGCCCCTCTTACAGCG	100
<i>CypFor</i> <i>CypRevN</i> <i>CypRevSNP</i>	gcctcgagTCCAAACTCCTCAGTTCTACAGC cc <u>aagctT</u> GGTTTGCCTTTATACTGCCTGA cc <u>aagctT</u> GGTTTGCCTTTGTACTGCCTGA	158
<i>RVprimer3_for</i> <i>pGLrev(2)_rev</i>	TAGCAAAATAGGCTGTCCC CGCTTCATGGCTTTGTG	

### Буферные растворы

TE - 10 mM Трис-ОН, 1 mM ЭДТА

ТАЕ - 40 mM Tris-HCl pH 8.0, 20 mM ацетат натрия, 1 mM ЭДТА

*Приготовление 2YT среды*

На 1 литр среды - 16 г бакто-триптона, 10 г бакто-дрожжевого экстракта, 5 г NaCl, проавтоклавировать.

#### *Приготовление LB среды*

В 1 л воды растворяют 10 г бактотриптона, 5 г дрожжевого экстракта, 10 г NaCl, доводят pH до 7,5 раствором 1 N NaOH и автоклавируют.

#### *Приготовление LB-агара*

В 1 л воды растворяют 10 г бактотриптона, 5 г дрожжевого экстракта, 10 г NaCl, доводят pH до 7,5 раствором 1 N NaOH и автоклавируют.

#### *Приготовление IPTG*

Индуктор ИПТГ (синтетический аналог лактозы, IPTG) – 100 мМ раствор изопропил-бета-D-1-тиогалактопиранозид; 2,38 г ИПТГ растворяют в 100 мл де-ионизированной воды, стерилизуют через фильтр с размером пор 0,22 мкм.

#### *Приготовление красителя (Coomassie Brilliant Blue G-250)*

На 100 мл берется Coomassie Brilliant Blue G-250 – 0,25 г, метанол – 45 мл, H<sub>2</sub>O – 45 мл, ледяная уксусная кислота – 10 мл.

#### *Приготовление красителя бромфенолового синего*

Готовится стоковый раствор - 1% в H<sub>2</sub>O, рабочая концентрация в 1x буфере ~0.005-0.02%.

### **Электрофорез в агарозном геле**

Фрагменты ДНК разделяли в 1-1.5% агарозном горизонтальном геле, приготовленном на буфере 1% TAE с добавлением бромистого этидия (0.01 мкг/мл). В гель наносили образцы, содержащие 0.3-0.5 мкг ДНК, и 0.2% бромфенолового синего. Электрофорез проводили при напряженности электрического поля 5-10 В/см.

## ПЦР

Для проведения ПЦР использовали амплификатор Termocycler (BIO-RAD™ США). Программа: 95° С- 3 мин, 40 циклов: 95° С – 30 сек, 55° С – 30 сек, 72° С - 1 мин.

ПЦР проводили в 20 мкл реакционной смеси, содержащей 75 мМ Трис-НСl, рН 8.9; 20 мМ (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; 0,01 % Tween; 2,5 мМ MgCl<sub>2</sub>; 10% глицерин; 1 U единиц активности Taq ДНК-полимеразы; 0,2 мкМ прямого и обратного праймеров; 0,1 мМ каждого дезоксирибонуклеозидтрифосфата и 50 нг геномной ДНК в качестве матрицы.

## Очистка образцов ДНК с помощью гель-фильтрации на Sephadex G100

Колонку высотой 8 см и диаметром 3 мм заполняли Sephadex G100, предварительно замоченным в дистиллированной воде на 3 часа. В колонку загружали 10 мкл образца ДНК, содержащего 0,01% бромфеноловый синий в качестве маркера, после чего наносили дробно 70 мкл воды и проводили элюцию ДНК добавлением 45 мкл воды. Очищенные продукты амплификации анализировали гель-электрофорезом в агарозном геле.

## Получение репортерных конструкций

Для получения репортерных конструкций ТАТА-содержащие участки промоторов генов *Lep* и *CYP2A6* человека были клонированы в вектор pGL 4.10[luc2] (Promega, США). Этот вектор содержит ген люциферазы светлячков (firefly luciferase), субстратом для которого служит люциферин жуков (beetle luciferin). Для клонирования необходимые фрагменты ДНК были амплифицированы при помощи

ПЦР с использованием праймеров, содержащих сайты рестрикции Xho I и Hind III (последовательности праймеров представлены в таблице 1), в качестве матрицы использовали ДНК клеток линии колоректальной карциномы человека НСТ-116 и ДНК клеток карциномы молочной железы MCF-7.

### **Выделение геномной ДНК**

Экстракция геномной ДНК проводилась по стандартному протоколу фенол-хлороформным методом из клеток линии НСТ-116. К клеточной суспензии добавили 500 мкл лизирующего буфера (50 мМ Tris-HCl pH 8.0, 200 мМ NaCl, 100 мМ EDTA, 1% SDS) и 5 мкл протеиназы К (20 мМ/мл). Гомогенат инкубировали при 55<sup>0</sup>С 2 часа, ц/ф 10 000g 10 минут. Пептиды экстрагировали смесью 1V водонасыщенного фенола и 1/5 V хлороформа, затем 1V смеси фенол: хлороформ в соотношении 1:1, и наконец, 1V хлороформа. Фазы разделяли центрифугированием при 10.000 g в течение 10, 10 и 5 минут, соответственно. К водной фазе добавляли 1V изопропанола, инкубировали 10 минут при комнатной температуре и центрифугировали 15 минут на 15000 g. Осадок промывали 1 мл 75% спирта, центрифугировали 3 минуты на 10000 g. Осадок высушивали 5-7 минут и растворяли в 1-кратном буфере TE.

### **Получение фрагментов ДНК с использованием ПЦР**

Фрагменты ДНК, соответствующие участку промотора, содержащему ТАТА-бокс гена *Lep* человека, получали с помощью полимеразной цепной реакции (ПЦР).

Праймеры, использованные для синтеза фрагмента ДНК, содержащего промоторный район *Lep* человека: *LepFor*, *LepRevN*, *LepRevSNP*. Аналогично праймеры,

использованные для синтеза фрагмента ДНК, содержащего район *Cyp* человека: *CypFor*, *CypRevN*, *CypRevSNP*.

Очистку целевого продукта реакции от соли и невключившихся дезоксирибонуклеозидтрифосфатов проводили с помощью гель-фильтрации на Sephadex G100.

## **Гидролиз и лигирование фрагментов ДНК**

### **Расщепление ДНК эндонуклеазами рестрикции**

Гидролиз полученного ампликона и плазмиды pGL4.10 проводили эндонуклеазами рестрикции XhoI и Hind III (Thermo scientific, UK). Для этого к 0,5 мкг ПЦР-продукта или 4 мкг плазмиды добавляли 46.4 U XhoI и 6.4 U Hind III (Thermo scientific, UK) и инкубировали в течение 2 часов при 37<sup>0</sup>С в буфере R+, содержащем 10 мМ Трис-НСl (рН=8,5), 10 мМ MgCl<sub>2</sub>, 100 мМ KCl, 0,1 мкг/мл БСА. Реакцию ингибировали нагреванием смеси до 65<sup>0</sup>С в течение 15 мин. Полученную смесь очищали на Sephadex G100.

Для анализа наличия ампликона нужной длины проводили электрофорез в агарозном геле на 1X TAE.

Осаждение ампликона проводили путем добавления к нему равного объема изопропанола и через паузу в 10 минут при комнатной температуре центрифугировали при 15000g 20 минут. Далее изопропанол сливали и добавляли 500 мкл 75% EtOH на 10 минут. Затем центрифугировали 5 минут при 10000g, осадок просушивали и растворяли в 20 мкл воды.

Полученные ампликоны встраивали в вектор pGL4.10, содержащий ген люциферазы (*luc2*) (Рис. 10).

### **Лигирование фрагментов ДНК**

Лигирование фрагментов проводили путем добавления к гидролизованному ампликону и плазмиде 100 U T4-ДНК-лигазы (Bioron, Германия) на 1 час при 16°C в буфере, содержащем 50 мМ Трис-НСl (рН=7,8), 10 мМ MgCl<sub>2</sub>, 10 мМ ДТТ, 1 мМ АТФ. Инактивацию фермента проводили в течение 20 минут при 65°C. Лигазную смесь очищали с помощью гель-фильтрации на Sephadex G100 и использовали для трансформации электрокомпетентных клеток *E.coli*.

### **Приготовление электрокомпетентных клеток *E.coli***

Клетки *E.coli* штамма Top10 засеивали в 5 мл питательной среды LB и растили ночную культуру на термостатированной качалке при 37°C и интенсивном помешивании (~ 170 об/мин) в течение 14-18 часов. 100 мкл ночной культуры переносили в 500 мл LB и растили на качалке при 37°C до оптической плотности OD<sub>600</sub>=0,6. Клетки осаждали в стерильных пробирках (Falcon) центрифугированием в течение 5 мин при 3000 g при температуре +4°C. Все дальнейшие процедуры проводились на льду. Осадок ресуспендировали в 50 мкл охлажденной бидистиллированной воды и осаждали центрифугированием в течение 5 мин при 3000 g при 4°C; процедуру повторяли 2 раза. Затем клетки ресуспендировали в 50 мл 10% глицерина и вновь осаждали. К полученным клеткам добавляли 2 объема 20% глицерина, суспендировали, делили на аликвоты по 40 мкл и замораживали на -70 °C.

### **Трансформация клеток *E.coli***

Трансформацию электрокомпетентных клеток *E.coli* проводили методом электропорации. Для этого к 40 мкл клеточной суспензии добавляли 10 мкл очищенной лигазной смеси, тщательно перемешивали и переносили смесь в кювету для элек-



тропорации (ширина щели 1мм). Электропорацию проводили с помощью электропоратора BIO-RAD MicroPulser™, при напряжении 2,0 кВ, время прохождения импульса 3,5-5,0 мс. Смесь быстро переносили в пробирку с 1 мл среды LB и инкубировали без перемешивания 1 час при 37°C. Пробу концентрировали центрифугированием 2 мин 900g, избыточный объем среды сливали и около 50 мкл каждой пробы высевали на чашки Петри (1 г агара на 100 мл LB среды) с селективной средой, содержащей ампициллин (100 мкг/мл). Через 14-17 ч инкубации при 37°C получали отдельные колонии трансформированных клеток.

### **Скрининг колоний на наличие вставки, их полимеразная цепная реакция и электрофорез**

С чашки Петри одиночные колонии скалывали и переносили в 50 мкл подготовленного раствора со средой LB (5мл LB +10 мкл ампициллина (50 мкг/мл)). Эти клетки использовали в качестве матрицы для ПЦР. ПРЦ проводили с праймерами к участкам плазмиды pGL 4.10 (RVprimer3\_for и pGLrev(2)\_rev) в реакционном объеме 10 мкл по программе: 95°C- 2 мин, 35 циклов: 94°C – 15 сек, 50°C – 30 сек, 72°C – 20сек. Каждая реакция ПЦР содержала в качестве матрицы 2 мкл пробы и буфер для ПЦР (25 мМ Трис-НСl pH=7,5; 1,5 мМ MgCl<sub>2</sub>; 70 мМ (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>), а также 0,2 мМ каждого дезоксирибонуклеозидтрифосфата, 0,2 пМ каждого праймера и 0,5 ед. акт. Таq ДНК-полимеразы.

Продукт ПЦР детектировали с помощью электрофореза в 1% горизонтальном агарозном геле. Колонии, длина ПЦР-продукта которых соответствовала наличию вставки, использовали далее.

## **Подтверждение наличия клонированного фрагмента ДНК с помощью секвенирования**

### **Выделение плазмидной ДНК методом щелочного лизиса**

Выделение плазмидной ДНК проводили по модифицированному методу Бирнбойма и Доли (Birnboim, Doly, 1979). Выбранные колонии переносили в пробирки с 5 мл LB, содержащей ампициллин (50 мкг/мл), и выращивали в течение ночи на качалке при температуре 37°C и скорости 170 об/мин.

Ночную культуру осаждали центрифугированием в течение 5 мин при 3000 g и 4°C. Супернатант удаляли. К осадку добавляли 1V раствора 1 (25 mM Трис-HCl, pH=8.0, 50 mM глюкозы, 20 mM ЭДТА, 2 мг/мл лизоцима) и тщательно суспендировали. Сразу же добавляли 2V раствора 2 (0,2 N NaOH, 1% SDS) и аккуратно перемешивали до просветления раствора. После чего добавляли 1,5V 3 M NaAc (pH=4.5-5.0), аккуратно перемешивали до получения белого творожистого осадка и центрифугировали 20 мин при 14000 g и 4°C. Супернатант отбирали, добавляли к нему равный объем изопропанола, тщательно перемешивали и выдерживали при комнатной температуре в течение 10 мин, а затем центрифугировали в течение 10 мин при 10000 g. Супернатант сливали, а осадок промывали 75% этанолом, подсушивали и растворяли в 30 мкл воды. К раствору добавляли 5 мкг РНКазы А и инкубировали 30 мин при 20°C. Очистку плазмидной ДНК проводили с помощью гель-фильтрации на Sephadex G100.

Концентрацию полученной плазмиды определяем с помощью NanoDrop 2000 Spectrophotometer.

## Секвенирование ДНК

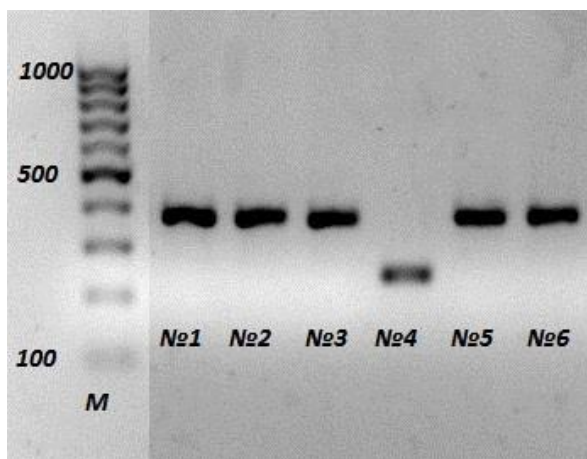
Для подтверждения наличия вставки промотора выделенную плазмиду секвенировали по методу Сэнгера с использованием праймеров *RVprimer3\_for* и *pGLrev(2)\_rev*.

Реакцию Сэнгера проводили в амплификаторе Thermocycler в 30 мкл реакционной смеси, содержащей 800 нг ДНК-матрицы, 2 пмоля праймера, 3 мкл BigDye Terminator Cycle Sequencing Kit v. 3.1 (Applied Biosystems, США), 2 мкл BigDye-буфера (Applied Biosystems, США). Секвенирующую реакцию проводили при следующих условиях: 95°C – 2 мин (прогревание); 95°C – 15 сек, 50°C – 20 сек., 60°C – 2 мин (40 циклов).

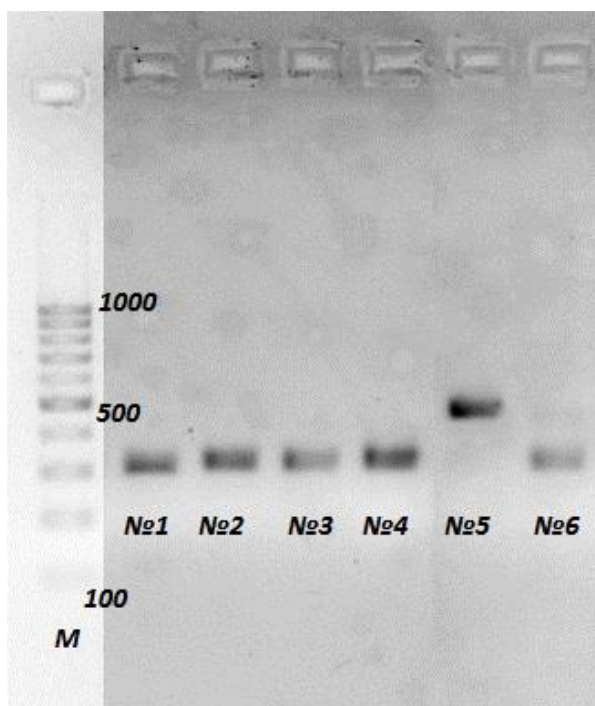
Очистку продуктов реакции проводили переосаждением. Для этого к реакции добавляли равный объем 100% изопропанола и оставляли при комнатной температуре на 10 мин. Пробирки центрифугировали при 12700 g в течение 15 мин. Остатки изопропанола скидывали, осадок подсушивали 10 мин при 37°C. Осадки для анализа передавали в ЦКП «Геномика».

### **Электрофоретический контроль плазмидной ДНК на содержание вставки на 2% агарозном геле с EtBr при 80 V**

*Результаты зафиксированы на трансиллюминаторе BioRad.*

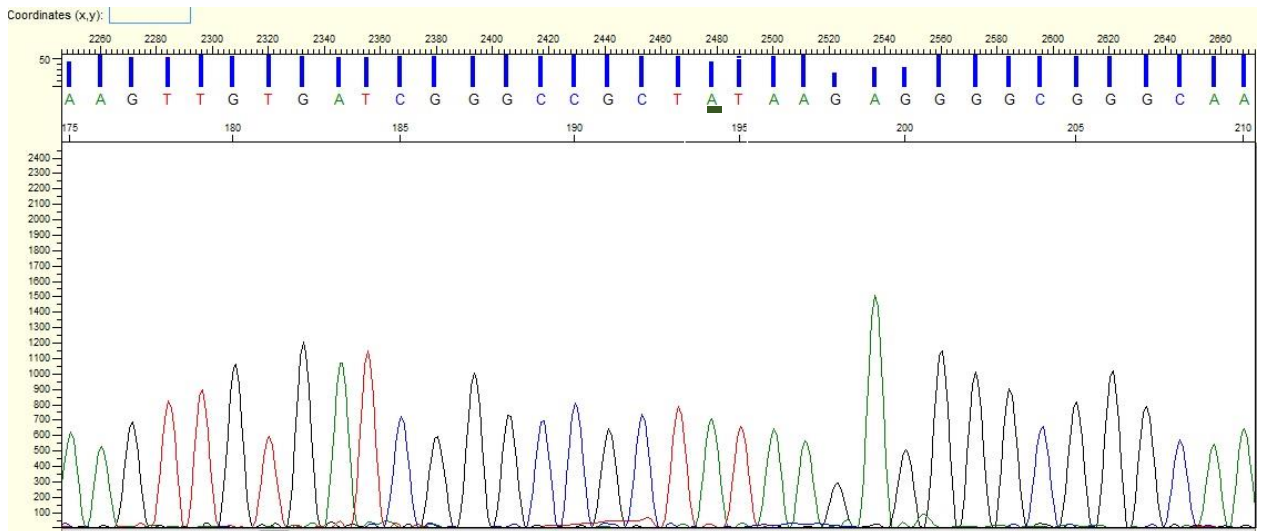


**Рисунок 11. LEP.** Референсный вариант WT. Наличие вставки во всех колониях, кроме №4 (ДНК Маркер 100bp (Медиген, Россия)).



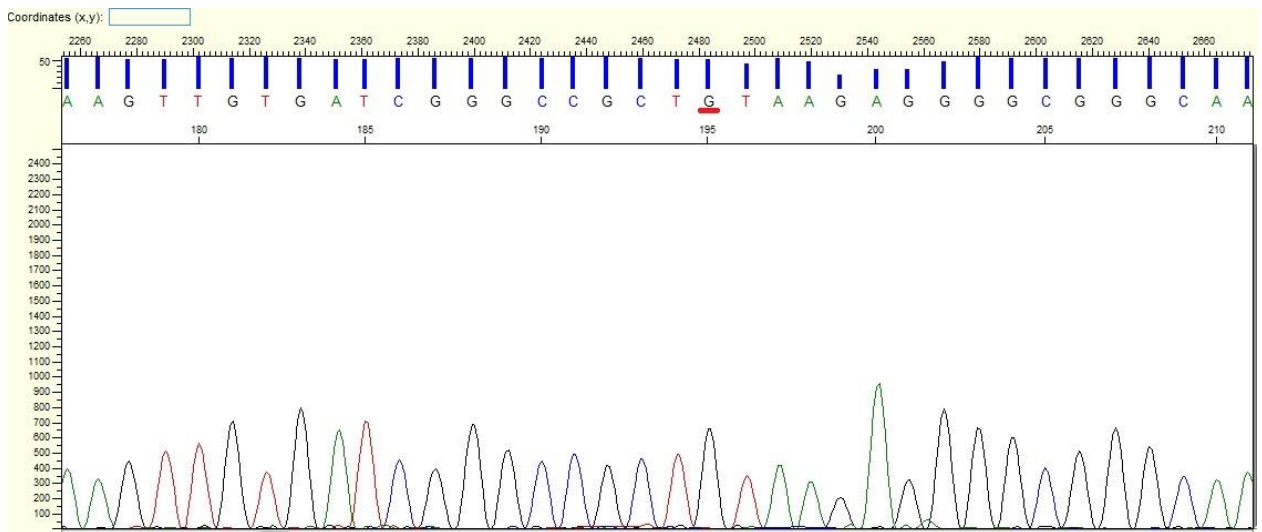
**Рисунок 12. LEP.** Замена А на G в rs201381696 (-35A>G). Наличие вставки только в колонии №5(ДНК Маркер 100bp (Медиген, Россия)).

*Секвеннограммы районов, содержащих однонуклеотидные замены в полученных плазмидных конструкциях*



**Рисунок 13. LEP. Референсный вариант WT.**

№11 AAGTTGTGATCGGGCCGCT**A**TAAGAGGGGCGGGCAA



**Рисунок 14. LEP. Замена A на G в rs201381696 (-35A>G).**

№26 AAGTTGTGATCGGGCCGCT**G**TAAGAGGGGCGGGCAA

## **Очистка плазмидной ДНК центрифугированием в двухступенчатом градиенте хлористого цезия**

Колонии с подтвержденной вставкой промотора наращивали в объеме 500 мл среды и плазмиду выделяли щелочным лизисом. В раствор плазмидной ДНК, выделенной щелочным лизисом добавляли раствор бромистого этидия до концентрации 700 мкг/мл и растворяли в нем кристаллический CsCl в расчете 0,59 г соли на 1 мл ( $\eta=1,38$ ). Затем готовили растворы CsCl в дистиллированной воде с оптической плотностью  $\eta=1,41$  (1,58 г соли на 1 мл воды) и с  $\eta=1,38$  (0,3 г CsCl на 1 мл воды).

Центрифужные пробирки (Beckman Quick-Seal) для ротора Ti-75 заполняли наполовину раствором CsCl с  $\eta=1,41$ . Сверху наслаивали равный объем лизата и уравнивали пробирки раствором CsCl с  $\eta=1,38$ . Отверстие пробирок запаивали и центрифугировали 24 часа при 45 000 об/мин в роторе Ti-75 при 20 °С.

В результате центрифугирования получали две малиново-красные зоны. Нижнюю зону, содержащую плазмидную суперспирализованную ДНК, отбирали через иглу с помощью перистальтического насоса.

Бромистый этидий экстрагировали 4 раза равным объемом изобутилового спирта, после чего пробы разбавляли в 3 раза водой. ДНК осаждали одним объемом изопропанола, осадок ДНК растворяли в 1 мл буфера TE.

## **Трансформация клеточных линий и измерение активности люциферазы**

В работе использовали эукариотическую клеточную линию карциномы молочной железы (MCF-7) и линию колоректального рака (HCT-116). Культивирование клеточных линий осуществлялось на базе Центра клеточных технологий ИЦиГ СО РАН.

Для введения вектора в клетки эукариотических линий использовали реагент ScreenFect (Incella, Germany), который увеличивает проницаемость цитоплазматической мембраны и, таким образом, обеспечивает проникновение в клетку плазмидной конструкции. Трансформацию проводили по стандартному протоколу Dual-Luciferase® Reporter Assay Protocol (Promega). Система включает в себя два вектора: беспромоторный pGL4.10[luc2], в который интегрировали фрагмент ТА-ТА-содержащего промотора генов *LEP* и *CYP* человека и нормировочный вектор pRL-TK. pGL4.10[luc2] содержит ген светлячковой люциферазы, который кодирует белок, обладающий ферментативной активностью, и в присутствии  $Mg^{2+}$  катализирует превращение люциферина в оксильюциферин. При переходе в окисленную форму субстрат испускает флуоресцентный сигнал, который можно детектировать и таким образом оценить количество наработанной люциферазы. Для трансформации брали векторы в соотношении 1:10 (на 0.4 нг RL-TK брали 4 нг исследуемой плазмидной конструкции). После трансформации инкубировали клетки в течение суток, затем обрабатывали 200 мкл Luciferase Assay Reagent II (Promega), смесь пипетировали и проводили детекцию флуоресцентного сигнала с помощью 2030 Multilabel reader VICTOR™ X3. Таким образом, был детектирован флуоресцентный сигнал светлячкового люциферина. Затем добавляли 100 мкл Stop & GloR Reagent (Promega), смесь пипетировали и проводили детекцию флуоресцентного сигнала медузного люциферина аналогичным способом. Для определения относительной экспрессии люциферазы в экспериментальной плазмиде ее активность нормировали на активность медузной люциферазы.

Клеточные линии культивировали в полной среде (F12/DMEM, содержащей 10% FBS) в  $CO_2$ -инкубаторе в условиях:  $+37^{\circ}C$ , 5%  $CO_2$ . Люциферазные репортер-

ные конструкции были получены на основе вектора pGL4.10 [Luc2], Promega, США. ТАТА-содержащие участки промоторов (159 п.о.), WT и минорные аллели генов *LEP* и *CYP2A6* клонировали в беспромоторный вектор pG4.10 по сайтам рестрикции XhoI и Hind III. Точность всех созданных конструкций подтверждали секвенированием по методу Сэнгера. Активность люциферазы измеряли с помощью двойного люциферазного набора для анализа (Dual-Luciferase, Promega, USA) согласно инструкции производителя. Беспромоторный вектор pGL4.10 использовали в качестве отрицательного контроля. Люциферазную активность (условные единицы) выражали как отношение активности светлячковой люциферазы к люциферазе Renila, нормированной на количество клеток. Данные трех и более повторностей каждого эксперимента (каждый эксперимент с тремя репликами) приводили как среднее значение  $\pm$  стандартное отклонение. Статистическую значимость разницы между средними определяли с помощью t-теста Стьюдента ( $p=95\%$ ) (Васильева, 2007).

### 3. Результаты и обсуждение

#### Гены *LEP*, *CYP*, *ABCA9*, *GCG*, *TPI* как гены-мишени

На основании анализа *in silico* для дальнейшего исследования нами были выбраны гены *LEP*, *CYP*, *ABCA9*, *GCG*, *TPI*, продукты которых участвуют в физиологических процессах организма и обмене веществ — метаболизме. Биологическое значение этих процессов заключается в наработке энергии, которая обеспечивает функционирование организма и постоянство его внутренней среды.

Живой организм представляет собой пример ультрастабильной системы, которая осуществляет активный поиск наиболее оптимального и наиболее устойчивого



состояния, что выражается в адаптации, т. е. в удержании переменных показателей организма в физиологических пределах, несмотря на изменения условий существования (Горизонтов, 1976).

Избыточный вес и ожирение влияют на качество и продолжительность жизни человека и вызывают дорогостоящие проблемы со здоровьем. Ожирение ассоциируется с рядом опасных для здоровья человека заболеваний, ведущих к преждевременной смерти. К ним относятся гипертония, сердечно-сосудистые (Lavie et al., 2009), онкологические заболевания (Reeves et al., 2007; Rodriguez et al., 2002), болезни почек, печени и сахарный диабет (Carbone et al., 2013; Chan et al., 1994). В последние годы выявлена связь между избыточным весом и нарушениями репродуктивного здоровья у женщин (Barber et al., 2006; Arroyo et al., 1997; Dravecka et al., 2003), заболеваниями предстательной железы у мужчин (Frederickson et al., 2008), повышенным риском появления детей с врожденными дефектами (Stothard et al., 2009). Во время беременности, у женщин обычно повышается потребление пищи, увеличивается жировая масса и наблюдается резистентность к инсулину. В последние годы увеличивается число женщин, у которых при беременности развивается метаболический дисбаланс с различными последствиями, механизм которого в основном не известен (Zampieri et al., 2015). Показано (Janjgava et al., 2014), что низкие уровни тестостерона являются показателем развития в будущем диабета 2 типа и метаболического синдрома у мужчин. Повышение физиологических уровней тестостерона, как показывает исследование (Janjgava et al., 2014), играет важную роль в уменьшении осложнений диабета. С каждым годом эти заболевания молодеют во всем мире. Несмотря на обнаруженную связь между тяжелыми патологиями и ожирением, молекулярные механизмы, участвующие в этих процессах,

плохо определены. (Nteebe et al., 2014). В большой степени это вызвано тем, что вес тела и энергетический баланс находятся под очень сложным контролем нервной, эндокринной и метаболической систем. С ожирением связано большое количество генов и вклад каждого гена низок. Например, самый высокий вклад для гена рецептора меланокортина 4 (*MC4R*) составляет 1 - 6% (Frederickson et al., 2008). Мутации в этом гене (до 6% людей с тяжелым ожирением, развивающимся с детства) приводят к недостаточности *MC4R*, что является самой распространенной причиной моногенного ожирения у человека. Сложность изучения вызвана еще и тем, что гены, участвующие в регуляции массы тела, участвуют также в регуляции других функций организма. Например, ген мю-опиоидного рецептора (*OPRM1*) участвует в получении удовольствия и процессах нарушения функций опорно-двигательного аппарата (Hall et al., 2010). Бомбесин-подобные пептиды (*BN*) участвуют в регуляции широкого спектра эмоционального поведения, включая обучение, память и питание (Yamada et al., 2002). Исходя из этого, всестороннее изучение генов, участвующих в нарушении метаболических путей и путей использования энергии, механизмов их регуляции, делающих совместный вклад в появление избыточного веса и ожирения, а также депрессии, диабета и других взаимосвязанных патологий, является одной из «горячих» проблем современной системной биологии и медицины.

Многие невропатологические заболевания человека либо вызывают ожирение, либо находятся под его влиянием (Lee and Mattson, 2014). С повышенным уровнем лептина в плазме крови связаны различные сердечно-сосудистые патологии (Zeidan et al., 2003), на развитие которых влияет генетическая предрасположенность, различные экологические факторы, такие как курение, питание и физическая

активность, и системные заболевания человека, такие как гипертония, сахарный диабет, ожирение, или заболевания щитовидной железы (Li et al., 2014). Во многих работах показано, что ожирение делает вклад в нарушение метаболизма, изменение гормонального фона и повышает чувствительность к воспалительным заболеваниям, которые согласованно действуя и способствуя друг другу, приводят к развитию диабета и различных невропатологий: эпилепсии, рассеянного склероза, болезни Альцгеймера и др. (Laughlin et al., 1997; Lee and Mattson, 2014). Сбалансированная энергия гомеостаза, как известно, зависит от нормального функционирования головного мозга, как главного регулятора обмена веществ, который объединяет периферийные и вегетативные сигналы и контролирует пищевое поведение (Lee and Mattson, 2014).

Учитывая сложный характер комплексных заболеваний, в образование которых делает вклад лептин, информация, полученная на разных уровнях исследований, позволит в полной мере изучить особенности этих заболеваний, чтобы лучше понять их патогенез и определить новые мишени для лекарственной терапии.

### **Ген *LEP* (Leptin)**

Открытие лептина и его участия во множестве различных процессов, связанных с пищевым поведением человека, положило начало для понимания регуляции энергетического гомеостаза (Farooqi and O'Rahilly, 2014). Лептин – белок-мономер размером 16 kD, это пептидный гормон, активно секретируемый адипоцитами жировой ткани пропорционально ее массе. Рецессивные мутации в гене лептина связаны с массовым ожирением у мышей и человека (Friedman, 2011). Лептин цирку-

лирует в крови и действует на мозг, который регулирует потребление пищи и расход энергии. Когда масса жира падает, плазменный уровень лептина падает, стимулируя аппетит и подавляя затраты энергии до тех пор, пока жировая масса не восстановится. Когда масса жира увеличивается, количество циркулирующего лептина увеличивается, подавляя аппетит, пока вес не уменьшится (Friedman, 2011). Считалось, что лептин продуцируется главным образом в адипоцитах (Zhang et al., 1994), но затем было показано, что лептин продуцируется также в желудке (Sobhani et al., 2000), сердце (Purdham et al., 2004) и плаценте (Masuzaki et al., 1997). Последние данные показали, что лептин не только модулирует массу тела, но выполняет и другие физиологические функции, в том числе, оказывает нормализующее влияние на сосудистый тонус и артериальное давление, стимуляцию ангиогенеза, кальцификацию клеточных сосудов и др. (Zeidan et al., 2003; Purdham et al., 2004). Многими исследователями показана положительная корреляция между концентрациями циркулирующего лептина и давлением крови (Hall et al., 2010) повышение которого инициирует развитие атеросклероза (Beltowski, 2008).

**Ген ABCA9.** Транспорт различных молекул через липидные мембраны является важной функцией всех живых организмов. К семейству транспортных белков АТФ-связывающей кассеты (ABC – ATP Binding Cassette) относятся белки, многие из которых переносят самые разные соединения: пептиды, холестерин и стероиды, желчные кислоты, ретиноиды, ионы и сахара и т.д. (Dean, Allikmets, 2001). Мутации в генах ABC вызывают различные заболевания, в том числе муковисцидоз, дегенерацию желтого пятна Штаргардта, и нарушения метаболизма липопротеинов и липидов (Oram, Vaughan, 2006). Кроме того, продукты многих генов этого семейства делают вклад в генетические и онкологические заболевания человека, участ-

вуют в формировании лекарственной устойчивости (Dean, Allikmets, 2001). Гены, кодирующие ABC-белки, рассредоточены по всему геному и обладают высокой гомологией по аминокислотному составу среди всех эукариот. В 2002 г. было открыто подсемейство белков-транспортёров ABCA9 макрофагов человека (Piehler et al., 2002). Анализ структуры генома показал, что ген *ABCA9* состоит из 39 экзонов, занимающих область примерно в 85 тыс. п.о. на хромосоме 17q24.2 (Piehler et al., 2002). Предполагается, что наряду с другими ABC-транспортёрами, *ABCA9* участвует в гомеостазе липидов (Ye et al., 2008; Oram, Vaughan, 2006). В работе большой группы авторов (Hedditch et al., 2014) получены интересные и противоречивые результаты о том, что высокие уровни экспрессии *ABCA1*, *ABCA6*, *ABCA8* и *ABCA9* в первичных опухолях статистически значимо ассоциированы со снижением выживаемости больных серозным раком яичников. При исследовании гена *ABCA10*, высоко гомологичного генам *ABCA9* и *ABCA6*, показано, что его экспрессия повышена в опухолях нескольких типов, в том числе в фолликулярной лимфоме (Baecklund et al., 2014), и влияет на исход заболеваний. Hendig с коллегами (Hendig et al., 2010) показали экспрессию 37 генов транспортёров ABC в кожных фибробластах при эластической псевдосаркоме. Семь генов – *ABCA6*, *ABCA9*, *ABCA10*, *ABCB5*, *ABCC2*, *ABCC9* и *ABCD2* – были индуцированы, в то время как экспрессия одного гена, *ABCA3*, была снижена, по сравнению с контрольной группой, по крайней мере в 2 раза.

**Ген *CYP2A6* (cytochrome P450).** *CYP2A6* является одним из участников системы цитохрома P450. Продукт этого гена, никотиноксидаза, участвует в активации проканцерогенов, детоксикации лекарств и формировании воспалительных реакций клетки (Pelkonen et al., 2000). Известно (Minematsu et al., 2003), что структур-

ные изменения в гене *CYP2A6* могут быть ассоциированы с повышенной чувствительностью организма к патологическому влиянию табака, формируя генетическую компоненту в риске развития хронической обструктивной болезни легких (ХОБЛ) (социальное заболевание со сложным механизмом наследования, распространенность и смертность от которого продолжает расти во всем мире (Kokosov, 2002; Markewitz et al., 1999)).

*CYP2A6* фермент имеет несколько ограниченный активный сайт, принимая лишь несколько ксенобиотиков в качестве субстратов. Интерес у научной общест-венности к гену *CYP2A6* значительно возрос после того, как никотино- и некото-рые табачно-специфичные нитрозамины были созданы в качестве высокоаффин-ных субстратов для этого фермента. В последнее время организация и структуры кластера генов *CYP2A* и несколько полиморфных аллелей гена *CYP2A6* были оха-рактеризованы. Были найдены два аллеля с точечной мутации и по крайней мере три различных типа делеции гена, ведущие к недостаточной функции гена. Частоты этих аллелей значительно различаются среди различных этнических групп (Raunio et al., 2001). Было проведено несколько исследований, посвященных рас-смотрению вопроса о взаимосвязи между статусом гена *CYP2A6* и привычки к ку-рению, а также роль *CYP2A6* полиморфизмов в риск развития рака легких. Иссле-дования на европейских популяциях не выявили каких-либо четких ассоциаций между вариантами *CYP2A6* генотипов и курением или предрасположенностью к раку легких, на японской популяции, где был распространен вариант гена *CYP2A6* с обедненными метаболитами, была найдена связь между курением и раком легких (Shimada et al., 1996).

Интерес к вопросу изучения гена *CYP2A6* у широкой общественности возник в 2007 году с появлением статьи по анализу ассоциации SNPs *CYP2A6* и курением (Waga et al., 2007). Так как кумарин-7-гидроксилаза, продукт гена *CYP2A6*, участвует во множестве метаболических реакций, перерабатывает различные лекарственные препараты, и даже никотин, его сложно привязать к какому-то определенному поведению, но авторы однако показали, что люди с одним из вариантов гена *CYP2A6* (группа, названная «психологически открытой для всего нового»), чаще начинают курить, чем люди с другим вариантом этого гена.

У пациентов, страдающих раком легких, был обнаружен полиморфизм в ТА-ТА-боксе промотора гена *CYP2A6* (-48 Т>G). Замена -48 Т>G разрушает ТАТА-бокс гена главной никотиноксидазы (*CYP2A6*) (Pelkonen et al., 2000), и ассоциируется с повышенным риском возникновения воспалительных заболеваний и рака легких у курильщиков - сродство ТСБ/ТАТА снижается в 4.7 раза.

### **Ген *GCG*.**

По бигормональной гипотезе R. Unger 1975 г. в регуляции углеводного обмена ключевыми являются гормоны инсулин и глюкагон (Петунина и др., 2013). Основным продуктом гена *GCG* является глюкагон, который участвует в нескольких этапах метаболизма и представляет собой пептидный гормон, произведенный альфа-клетками эндокринной части поджелудочной железы, также экспрессия глюкагона возможна и в L-клетках кишечника, и в специфических областях головного мозга. Вначале сигнальный пептид отщепляется от препроглюкагона (это предшественник для образования пептидов, включая глюкагон, глюкагонподобные пептиды 1 и 2). Глюкагон экспрессируется для повышения концентрации глюкозы в крови. Его действие противоположно действию инсулина, который снижает уровень глюкозы.

Поджелудочная железа выделяет глюкагон, когда концентрация глюкозы в крови падает слишком низко. Глюкагон способствует преобразованию гликогена в глюкозу, которая высвобождается в кровоток. Высокий уровень глюкозы в крови стимулирует высвобождение инсулина. Инсулин приводит к снижению глюкозы и используется в инсулин зависимых тканях. Таким образом, глюкагон и инсулин являются частью системы обратной связи, которая держит уровни глюкозы в крови на стабильном уровне. Это увеличивает расход энергии и повышается в условиях стресса. Глюкагон принадлежит к семейству нескольких других родственных гормонов.

Существуют генетические маркеры (Мохорт и др., 2013), определяющие риск развития сахарного диабета II типа (80-90% от общего количества пациентов с диагнозом «Диабет»). Среди них есть точечные мутации 2-ого экзона гена рецептора глюкагона, приводящие к нарушению секреции глюкагона (повышение).

Ген глюкагон (*GCG*), кодирует несколько гормонов, важных для энергетического обмена: глюкагон, оксинтомодулин и глюкагон-подобный пептид (GLP) -1 и -2. SNPs в *GCG* могут ассоциироваться с сахарным диабетом 2 типа, ожирением и/или связанными с ним признаками нарушения обмена веществ.

Существует множество описанных примеров SNPs гена *GCG*, вклад их различен, но как правило включаются компенсаторные механизмы организма, нивелирующие возникающие проявления. Самым ярким из описанных полиморфизмов в некодирующей зоне гена *GCG* является SNP rs4664447, расположенный в интроне (сайт сращивания энхансера) (Torekov et al., 2011). В данном исследовании изучали гомозиготных носителей G аллеля rs4664447, что проявлялось в несколько измененных свойствах обмена веществ. Гомозиготные носители имели ~50% от уровня



инсулина в сыворотке крови и самый низкий уровень глюкозы в плазме крови по сравнению с носителями здорового аллеля. Такая замена уменьшила экспрессию *GCG* в результате ошибок сплайсинга, что привело к снижению циркулирующих уровней глюкагона и GLP-1. Но при этом у носителей GG инсулиновая реакция была адекватной на уровень чувствительности к инсулину, что указывает на нормальную функцию бета-клеток с нераспределенной способностью адаптироваться к изменениям в чувствительности к инсулину. Такую интерпретацию данных соотносят с отсутствием ассоциации полиморфизма с сахарным диабетом II типа. Скорее всего, низкий уровень инсулина в сыворотке у гомозиготных носителей можно рассматривать как компенсаторный механизм при низком уровне глюкагона. Действительно, было показано, что эндогенно введенный глюкагон может запустить образование цАМФ в бета-клетках и тем самым секрецию инсулина (Pipeleers et al., 1985). Низкий уровень глюкозы в плазме крови можно также непосредственно объяснить снижением уровня циркулирующего глюкагона, что вызывает уменьшение гликогенолиза и глюконеогенеза. В соответствии с этой концепцией, мыши, нокаутные по *GCG*, показывают более низкое содержание глюкозы в крови и сниженный уровень инсулина в сыворотке крови, а также более высокую чувствительность к инсулину, по сравнению с мышами дикого типа (Hayashi et al., 2009).

**Ген триозофосфатизомеразы (*TPI*)** экспрессируется во всех типах клеток и относится к генам «домашнего хозяйства» (Brown et al., 1985). Во всех тканях человека обнаружены множественные электрофоретические и хроматографические формы *TPI*, но все они кодируются одним геном и являются результатом посттрансляционных модификаций белка (Brown et al., 1985). Молекулярная характеристика аллелей афроамериканцев и европеоидов (Watanabe et al., 1996) позволила

обнаружить SNP (-24T>G) в TATA-бок-подобной последовательности, расположенной на расстоянии 24 пн от старта транскрипции. У людей с мутацией в промоторе *TPI* наблюдается недостаток мРНК, активность фермента составляет 2-10% от нормы, для пациентов характерны нервномышечные расстройства и гемолитическая анемия (Watanabe *et al.*, 1996). Кроме того, недавно показано (Wang *et al.*, 2008), что триозофосфатизомераза при раке желудка может превращать устойчивые к лекарствам клетки в чувствительные, что улучшает прогноз лечения при химиотерапии и делает фермент кандидатной мишенью при разработке новых препаратов для лечения рака желудка.

Во многих работах показано, что SNPs регуляторных областей генов и, в частности, TATA-боксов промоторов, являются причиной возникновения таких моногенных заболеваний, как бэ́та-талассемии разной тяжести, синдром Gilbert, копко-подобная катаракта и др. (Savinkova *et al.*, 2009). SNPs TATA-боксов также делают вклад в повышение риска возникновения различных комплексных заболеваний человека (неврологические и мышечные нарушения, онкологические заболевания, гипертония и др.). Все эти нарушения возникают в результате изменения уровней экспрессии генов и количества синтезируемого белкового продукта, ассоциированных с присутствием SNPs в кор-промоторах, включающих TATA-боксы.

### **3.1. Определение кинетических характеристик комплексов ТСБ/ТАТА анцестральных и минорных аллелей с помощью метода EMSA**

В работе были определены скорости образования и распада ( $k_a$  и  $k_d$ , соответственно) комплексов образованных ТСБ и ODN для референсного и анцестраль-

ных аллелей изучаемых генов – Таблица 6. А также из полученных значений определены времена полураспада комплексов и изменение энергии Гиббса, характеризующее возможность спонтанного протекания реакции взаимодействия ТСБ/ТАТА.

**Таблица 6.** Кинетические характеристики взаимодействия ТСБ/ОДН (определение констант связывания ( $k_a$ ) и диссоциации ( $k_d$ ) комплексов ОДН/ТСБ) для референсных генов (Hg 19) и минорных аллелей.

Ген	5'-3' нить ДНК, 26 п.о.	Норма (Hg 19)/SNP	$k_a(M^{-1}c^{-1})$	$k_d(c^{-1})$	$-\Delta G$ (kcal/mol)	$t_{1/2}$ МИН
	ТАТА-бокс					
<i>LEP</i>	gcТАТАAGag	Hg 19	$(2.4 \pm 0.6) * 10^3$	$(1.8 \pm 0.6) * 10^{-4}$	9.71±0.9	64.17±13
	<u>a</u> сТАТАAGag	-38G>A rs200487063	$(3.5 \pm 0.5) * 10^3$	$(1.5 \pm 0.2) * 10^{-4}$	10.04±1	77±16
	gcТАТАAGtg	-30G>T rs34104384	$(1.1 \pm 0.2) * 10^4$	$(8 \pm 1) * 10^{-4}$	9.73±1	14.44±3
	gcT <u>G</u> ТАAGag	-35A>G rs201381696	$(5.6 \pm 0.8) * 10^3$	$(1.3 \pm 0.2) * 10^{-3}$	9.04±0.9	88.85±21
<i>ABCA9</i>	aТАТАCAAat	Hg 19	$(7 \pm 1) * 10^3$	$(4.6 \pm 0.8) * 10^{-4}$	9.79±0.8	25.11±7
	aaТАТ <u>G</u> CAAa	-37A>G rs367781716	$(2.9 \pm 0.4) * 10^3$	$(4.9 \pm 0.7) * 10^{-4}$	9.23±1.2	23.57±7
<i>GCG</i>	ttТАТАТАCtc	Hg 19	$(7 \pm 1) * 10^4$	$(5 \pm 1) * 10^{-4}$	11.1±1	23.1±5
	ttТАТА <u>C</u> ACT	-41T>C rs183433761	$(3 \pm 1) * 10^4$	$(6 \pm 1) * 10^{-4}$	10.5±0.9	19.25±4
<i>TPI</i>	tcТАТАТАAg	Hg 19	$(2.0 \pm 0.1) * 10^5$	$(6.8 \pm 0.8) * 10^{-4}$	11.5±0.9	16.98±3
	tcТАТА <u>G</u> AAgt	-24T>G rs1800202	$(4.6 \pm 0.5) * 10^3$	$(4.9 \pm 0.6) * 10^{-4}$	9.4±1.3	23.57±3
<i>CYP2A6</i>	gТАТАAAAGg	Hg 19	$(3.0 \pm 0.3) * 10^4$	$(7 \pm 1) * 10^{-4}$	10.4±1	16.5±3
	gТА <u>G</u> AAAGg	-34T>G rs28399433	$(2.2 \pm 0.3) * 10^3$	$(2.1 \pm 0.5) * 10^{-4}$	9.57±1.1	55±13

*LEP*. В Табл. 6 представлены значения  $k_a$  и  $k_d$ , характеризующие скорости образования и распада комплексов ТСБ с ТАТА-боксом в норме и при содержании SNPs, ассоциированных с нарушением липидного обмена. В Таблице 6 также представлены рассчитанные значения  $K_D$ , характеризующие аффинность ТСБ к ОДН и времена полураспада комплексов, характеризующие их прочность. Из результатов, представленных в Таблице 6, видно, что сделанные прогнозы влияния SNPs на сродство ТСБ/ТАТА, подтверждены экспериментально. Полученные экспериментальные значения  $-\ln K_D$  и  $\Delta$  хорошо согласуются с прогнозами. Из представленных результатов можно видеть, что SNP -38G>A в кор-промоторе гена *LEP* приводит к увеличению скорости образования комплексов ТСБ/ТАТА на 50%, SNP -30G>T приводит к увеличению скорости образования комплексов почти в 5 раз, а SNP -35A>G – к увеличению в 24 раза.

Скорость распада комплексов для -38G>A повысилась на 20%, для -30A>T уменьшилась в 4.4 раза, а для -35A>G повысилась в 14 раз. Аффинность ( $K_D$ ) ТСБ к ODNs с SNP -38G>A увеличилась на 50%, с SNP -30A>T (рис. 15) - увеличилась примерно на 12%, а для SNP -35A>G аффинность снизилась в 2.8 раза (для WT  $K_D$  = 80 nM, а для SNP -35A>G = 230 nM). При этом минимальное время полу-жизни комплексов, 8.9 мин, наблюдалось для SNP -35A>G, что вызвано увеличением скорости их распада в 14 раз относительно нормы и очень низкой аффинностью. Следует отметить, что SNP -38G>A и SNP -30A>T, расположены во фланках ТАТА-бокса, и только SNP -35A>G вносит изменение в последовательность самого ТАТА-бокса, что и приводит к снижению аффинности до 230 nM. Ранее (Savinkova et al., 2013) мы показали, что SNP -21C>T, расположенный во фланке ТАТА-элемента промотора гена тканевого фактора (*TF*), приводит к увеличению аффинности в 2.7



Количественные изменения экспрессии генов, вызываемые полиморфизмами в регуляторных районах генов, играют важную роль в восприимчивости человека к различным заболеваниям, условиям окружающей среды и лекарственной терапии. Исследования многих лабораторий направлены на выяснение того, как эти изменения последовательности ДНК реализуются на молекулярном уровне. Мы экспериментально показали, что влияние SNPs, ассоциированных с повышенным риском возникновения ожирения и сопутствующих ему патологий, на взаимодействие ТСБ с ТАТА-боксами, приводящее к изменению аффинности взаимодействия ТСБ/ТАТА на молекулярном уровне реализуется через изменение скоростей образования и распада комплексов ТСБ/ТАТА.

**Ген *ABCA9*.** Экспериментальные оценки изменения кинетических характеристик связывания ТСБ с промотором референсного и минорного аллелей гена *ABCA9* представлены в таблице, в которой показано, что скорость образования,  $k_a$ , комплекса ТСБ с минорным вариантом ТАТА-бокса снижается в 2,4 раза, тогда как скорость его распада,  $k_d$ , определяющая время полужизни ( $t_{1/2}$ ) этого комплекса, изменяется незначительно: с 25 мин в норме (референсный аллель) до 23,6 мин для минорного аллеля. В свою очередь, свободная энергия Гиббса ( $\Delta G$ ), которая характеризует спонтанное протекание реакции взаимодействия ТСБ с промотором гена *ABCA9* человека, уменьшается для минорного аллеля и свидетельствует о дестабилизации комплекса ТСБ с этим промотором вследствие ухудшения последовательности ТАТА-бокса и снижения сродства между ними.

В качестве экспериментальной проверки этого потенциально кандидатного SNP-маркера были измерены методом гель-ретардации (EMSA) скорости образо-

вания ( $k_a$ ) и распада ( $k_d$ ) комплексов ТСБ с олигонуклеотидами, идентичными аллелям  $-37T$  (норма) и  $-37C$  (недостаток синтеза белка гена *ABCA9*). Установлено, что константа скорости образования комплексов,  $k_a$ , в 2,4 раза ниже нормы. Прогноз с помощью Web-сервиса SNP\_TATA\_Comparator (Рассказов и др., 2013) относительного изменения кажущейся константы диссоциации ( $K_D = k_d/k_a$ ) и экспериментальное значение этой величины совпали в пределах точности эксперимента и расчетов. В работе также измерены время полураспада и свободная энергия Гиббса комплекса ТСБ с промотором *ABCA9*, в контексте которых обсуждаются в последней главе вероятностные фенотипические проявления SNP-маркера rs367781716.

**Ген GCG.** Влияние SNP  $-41T>C$  (rs183433761) гена *GCG* на взаимодействие ТСБ с ТАТА-боксами, приводит к изменению аффинности взаимодействия ТСБ/ТАТА и происходит это через изменение скоростей образования и распада комплексов ТСБ/ТАТА. Экспериментально измеренная скорость образования комплекса падает в 2,3 раза, а скорость распада комплекса возрастает в 1,2 раза. Получается, что комплексы ТСБ/ТАТА при наличии полиморфизма образуются медленнее и распадаются быстрее, чем при нормальном аллеле. Из этого заключения вполне ожидаемо, что время жизни комплексов должно уменьшиться, что и представлено из экспериментально полученных расчетов – время полужизни комплексов ТСБ/ТАТА при замене  $-41T>C$  гена *GCG* уменьшается в 1,2 раза, в абсолютном значении с 23 до 19 минут. А изменение свободной энергии Гиббса ( $-\Delta G$ ), которая характеризует спонтанное протекание реакции взаимодействия ТСБ с ТАТА-содержащим ОДН, уменьшается для минорного аллеля с 11,1 до 10,5 ккал/моль, что свидетельствует о дестабилизации комплекса ТСБ с таким ОДН, так как изме-

няется в худшую сторону последовательность ТАТА-бокса, что и ведет к снижению сродства ТСБ/ТАТА.

**Ген *TPI*.** SNP -24T>G в ТАТА-боксе гена *TPI* (Watanabe et al., 1996), приводит к 43-кратному уменьшению скорости ассоциации ТСБ/ТАТА и небольшому снижению (на 30%) скорости диссоциации. Период полураспада комплексов возрастает в 1,4 раза. Когда ТАТА-бокс не изменен, период полураспада составляет 17 минут, а при наличии SNP в ТАТА-боксе период полураспада составляет 23 мин. Эта мутация приводит к уменьшению  $-\Delta G_0$  на 2,1 ккал/моль. Существуют люди с дефицитом по *TPI*, имеющие недостаток мРНК, у них активность фермента в эритроцитах и других клетках было 2-10% от нормы; пациенты имели неврологические и мышечные расстройства и гемолитическую анемию. Наблюдались и другие носители гетерозиготных аллелей, которые имели умеренное снижение активности *TPI* в естественных условиях на 26-50% (Watanabe et al., 1996). Было показано, что SNP -24T>G (очень редко встречается, его распространенность составляет менее 1%) (Patikoglou и др., 1999), приводит к самым значительным связанным с ним расстройствам - 43-кратное снижение скорости ассоциации ТСБ/ТАТА и наибольшее снижение сродства среди всей рассмотренной группы генов. При таком уменьшении, вероятность образования комплекса понижается. Скорости диссоциации уменьшается в 1,4 раза, а период полураспада комплексов возрастает в 1,4 раза. Как видно из Таблицы 6, объединение ТСБ с одним из очень хороших последовательностей ТАТА-бокс, ТАТАТА, в котором пятый "Т" заменяется на "G", сопровождается наибольшим уменьшением свободной энергии связывания для рассмотренной группы генов, на 2,1 ккал/моль. Вполне возможно, что эта энергия исполь-



зуются для введения конформационных изменений и укрепления низкоаффинных ТСБ/ТАТА комплексов.

**Ген *CYP2A6*.** В гене *CYP2A6* при замене -34Т>G rs28399433 падают скорости и ассоциации комплекса ТСБ/ТАТА (в 13,6 раза) и диссоциации его (в 3,5 раз). Из уменьшения скорости ассоциации ТСБ/ТАТА с 30 до 2,2 М<sup>-1</sup>с<sup>-1</sup> получается, что комплексы образуются медленнее, но дольше взаимодействуют. При практически неизменной свободной энергии время полужизни комплекса увеличивается в 3,33 раза. Хочется обратить внимание, что снижение аффинности в данном случае ( $K_D$  увеличивается в 4 раза с 25 до 100 нМ) также как и в полиморфизме гена *TPI* приводит к увеличению жизни комплекса ТСБ/ТАТА. Если механизм происходящего процесса схож, то можно ожидать что в случае -34Т>G rs28399433 гена *CYP2A6* будет снижение мРНК и появление соответствующих фенотипических признаков.

Анализируя уравнение свободной энергии системы Гиббса, становится понятным, что, так как внутренняя энергия и энтальпия, с одной стороны, и энтропия - с другой являются функциями состояния системы, то и свободная энергия является функцией состояния системы. В нашем случае происходит при рассматриваемых полиморфизмах в генах *ABCA9*, *GCG*, *TPI*, *CYP2A6* и rs201381696 гена *LEP* изменение свободной энергии Гиббса в сторону уменьшения предпочтения в протекании реакции по кинетическим причинам из-за увеличения энтальпии на проведение реакции взаимодействия ТСБ/ТАТА. В случаях rs200487063 и rs34104384 гена *LEP* изменения свободной энергии реакции взаимодействия ТСБ/ТАТА происходит из-за увеличения энтропии системы.

Количественные изменения в экспрессии генов, вызванные наличием полиморфизмов в регуляторных областях играют важную роль в человеческой восприимчивости к различным заболеваниям, условиям окружающей среды и лекарственной терапии. Впервые в нашем секторе было экспериментально показано, что последствия полиморфизмов, связанных с повышенным риском развития различных наследственных патологий человека, реализуются через взаимодействие между ТСБ и ТАТА-боксом, что на молекулярном уровне происходит через изменения скоростей ассоциации/диссоциации в связывании ТСБ/ТАТА (Drachkova et al., 2014). Значения свободной энергии Гиббса были отрицательными, указывая, что ТСБ/ТАТА связывание было спонтанным. Поскольку ТСБ/ТАТА связывание происходит при самых стандартных условиях реакции, при той же температуре, снижение  $\Delta G_0$  значения указывают, что ТСБ/ТАТА комплекс дестабилизируется при наличии SNPs.

Хорошее согласие между значениями  $K_D$ , характеризующими сродство ТСБ/ТАТА, полученные в условиях равновесия (Savinkova et al., 2013) и выведенные из кинетических констант, позволяют предположить, что кинетические параметры взаимодействия между ТСБ и ТАТА-содержащими ODN, были определены правильно. Все полученные характеристики взаимодействия ТСБ/ТАТА находятся в пределах диапазона значений, характерных для ТСБ/ТАТА взаимодействий.

Очевидно, что ТСБ связывается с консенсусным и полиморфными вариантами ТАТА-бокса *in vivo* и *in vitro* при различных условиях, с разными TAFs и транскрипционными факторами. Тем не менее, зная корреляцию между степенью (Drachkova et al., 2014) изменения скоростей ассоциации и диссоциации комплексов ТСБ и SNP-содержащих ТАТА-боксов *in vitro*, а также степенью тяжести со-

путствующих заболеваний, мы можем прогнозировать вклад изучаемого SNP, еще не ассоциированного с конкретным заболеванием, в систему, где исследуемый ген наиболее ярко выражен. Это свидетельствует о том, что полученные результаты адекватно отображают молекулярные взаимодействия, происходящие в живом организме.

### 3.2. Анализ термодинамических характеристик комплексов ТСБ/ТАТА анцестральных и минорных аллелей

**Табл. 7.** Экспериментальная верификация прогнозируемых *in silico* значений сродства ( $K_D$ ) ТСБ/ОДН и влияние на него верифицированных SNPs.

hg19	5'-3' нить ДНК, 26 п.о.	Аллель (WT, rs)	Прогноз		Эксперимент		
			$-\ln[K_D]$	$\delta$	$K_D$ , nM	$-\ln[K_D]$	$\delta$
<i>LEP</i>	atcgggccgcTATAAGaggggcgggc	WT	19.43		80±20	16.34	
	atcgggcc <b>a</b> cTATAAGaggggcgggc	-38G>A rs200487063	19.85	0.42	40±10	17.03	0.69
	atcgggccgcTATAAG <b>t</b> ggggcgggc	-30A>T rs34104384	19.7	0.27	70±20	16.47	0.13
	atcgggccgcT <b>G</b> TAAAGaggggcgggc	-35A>G rs201381696	18.22	-1.21	230±10	15.29	-1.05
<i>ABCA9</i>	gctcagaaaTATACAAataattgcaa	WT	19.66		70±20	16.47	
	gctcagaaaTAT <b>G</b> CAAataattgcaa	-37A>G rs367781716	18.73	-0.93	170±30	15.59	-0.88
<i>GCG</i>	cactgctttTATATA <b>C</b> tctccagctc	WT	20.89		8±2	18.64	
	cactgctttTATA <b>C</b> Actctccagctc	-41T>C rs183433761	20.28	-0.61	18±4	17.83	-0.81
<i>CYP2A6</i>	tttcaggcagTATAAAGgcaaacac	WT	20.03		25±4	17.5	
	tttcaggcagTA <b>G</b> AAAGgcaaacac	-34T>G	18.53	-1.5	100±30	16.12	-

		rs28399433					1.38
<i>TPI</i>	cgcggcgctcTATATAAgtgggcagt	WT	20.11		3±0.3	19.62	
	cgcggcgctcTATAGAAgtgggcagt	-24T>G	19.08	-1.03	21±4	17.68	-1.94
		rs1800202					
$r = 0.57$		$r = 0.89$					
$\alpha < 0.05$		$\alpha < 0.005$					

В Табл. 7 приведены количественные прогнозы ( $\ln K_D$ ) сродства ТСБ/ТАТА анцестральных аллелей и величина их изменений ( $\delta$ ) относительно референсного аллеля. В таблице также приведены экспериментально полученные значения сродства ТСБ/ТАТА и его изменений ( $\delta$ ). Как можно видеть их таблицы, наблюдается достаточно высокое соответствие прогнозов и экспериментальных данных. Коэффициент линейной корреляции ( $r = 0.89$ ).

Мы определили *in silico* все неаннотированные SNPs промотора гена лептина (*LEP*) человека, полученные в результате выполнения проекта 1000 Genomes project (Colonna et al., 2014). Поиск проводился в районе от [-20 до -70] от старта транскрипции, где располагаются все экспериментально определенные ТАТА-боксы. Мы обнаружили SNP маркер дефицита лептина, rs201381696, который был клинически связан с индуцированным диетой ожирением и стеатозом (Lee, Mattson, 2014; Zhang et al., 2015), и два предполагаемых SNPs rs200487063 и rs34104384, для избытка лептина, которые коррелируют с индуцированной ожирением гипертонией, что было определено ранее (Tümer и др 2007; Beltowski, 2008). Экспериментально мы показали, что сродство ( $K_D$ ) ТСБ к промотору референсного гена равнялось 80 нМ. В случае аллеля -35A>G (rs201381696), сродство было ниже почти в 3 раза,  $K_D = 230$  нМ. Оно надежно ( $Z = 17$ ,  $\alpha < 10^{-7}$ ) указывает на дефицит

лептина у носителей этого SNP. Величина сродства ТСБ в случае аллелей -38G>A (rs200487063) и -30A>T (rs34104384), была выше на 50% и 11 %, соответственно. В последнем случае разница находится на границе достоверности.

Экспериментально охарактеризован, не ассоциированный с каким-либо нарушением здоровья человека, SNP rs367781716, -37A>G, промотора к альтернативному старту транскрипции в позиции chr17:66985252 (минус-цепь) гена *ABCA9*, который был детектирован экспериментально (Piehler et al., 2002), но не был до сих пор документирован в базе данных GENECODE v.19 (Harrow et al., 2012). С помощью Web-сервиса (Рассказов и др., 2013) было предсказано для этого SNP rs367781716 достоверное снижение сродства ТСБ к промотору. Полученный с помощью Web-сервиса прогноз относительного изменения в 2,4 раза кажущейся равновесной константы диссоциации  $K_D$  комплекса ТСБ с минорным аллелем -37G промотора гена *ABCA9* приведен в Таблице 7. Данные независимых опытов *in vitro* и *in vivo* (Pugh, 2000; Stewart, Stargell, 2001; Mogno et al., 2010) указывают на то, что такому снижению сродства ТСБ/промотор соответствует недостаточность у пациента-носителя аллеля -37A>G продукта *ABCA9*, транспортера АТФ - связывающей кассеты А9. В свою очередь, этот *in silico* прогноз риска недостаточности экспрессии транспортера, предположительно участвующего в липидном гомеостазе, в случае аллеля -37A>G позволяет предположить rs367781716 в качестве кандидатного SNP-маркера для нарушений липидного обмена.

Охарактеризован экспериментально SNP rs183433761, -41T>C, промотора к альтернативному старту транскрипции в позиции chr2:162152278 гена *GCG*. Полученный в результате выполнения проекта 1000 Genomes project SNP, - полиморфизм был охарактеризован *in silico* с помощью того же Web-сервиса, и показал на

уменьшение сродства ( $\delta = -0.61$ ) между ТСБ и ТАТА-боксом при его наличии. Относительное изменение кажущейся равновесной константы диссоциации  $K_D$  ( $K_D$  возросла с 8 до 18 нМ) комплекса ТСБ с минорным аллелем -41С промотора гена *GCG* в 2,25 раза – также свидетельствует о снижении сродства ( $\delta = -0.81$ ). Такое изменение предположительно может внести вклад в снижение экспрессии глюкогона и, как следствие, - к нерегулируемому и не постоянному уровню концентрации глюкозы в крови.

Полиморфизм -34Т>G rs28399433 в ТАТА-боксе гена *CYP2A6* из системы цитохрома P450, найденный в проектах 1000 геномов, ExAC и GO-ESP, привел к существенному снижению сродства ТСБ/ТАТА в 4 раза ( $K_D$  возросло с 25 до 100 нМ), что было весьма ожидаемо исходя из биоинформатических прогнозов ( $\delta = -1.5$ ). Экспериментальная  $\delta = -1.38$  указывает на существенное ухудшение сродства, что в дальнейшем можно будет использовать для маркирования замены. Так как рядом с исследуемой заменой -34Т>G расположен уже описанный полиморфизм (-48 Т>G) (Pelkonen et al., 2000), приводящий к снижению сродства ТСБ/ТАТА в тех же порядках - в 4.7 раза, то можно предполагать, что исследуемая замена -34Т>G будет также вносить вклад в разрушение ТАТА-бокса гена главной никотиноксидазы и ассоциироваться с повышенным риском возникновения воспалительных заболеваний и рака легких.

Замена -24Т>G rs1800202 в гене *TPI* приводит к снижению сродства в 7 раз ( $K_D$  возросла с 3 до 21 нМ), при прогнозируемом  $\delta = -1.03$  против экспериментального  $\delta = -1.94$ . Такое резкое ухудшение сродства обращает внимание на последовательность ОДН и замену в ней с целью дальнейшего исследовательского моделирования, и способствует закреплению такого полиморфизма именно в таком окру-

жении в качестве маркера. Так как фермент триозофосфатизомераза относится к «генам домашнего хозяйства» и участвует в гликолитическом пути, то логично ожидать, что нарушение метаболизма триозофосфатов может привести к избыточной продукции метаболитов гликолиза. Нарушения в ТАТА-боксе гена триозофосфатизомеразы приводят к снижению мРНК и появлению некоторых болезней (Watanabe et al., 1996).

Анализируя Таблицу 7 прогнозируемых *in silico* значений сродства  $K_D$  ТСБ/ОДН и  $K_D$ , полученных экспериментально, можно увидеть, что экспериментальная оценка величины  $K_D$  совпала с прогнозом *in silico* в пределах точности используемых расчетов и измерений:  $\delta$ , характеризующая изменение значения  $-\ln K_D$  комплексов ТСБ с референсным и минорным аллелями, равна значимым величинам по прогнозу и в результате экспериментальной верификации. Это означает, что с использованием двух тест-систем, *in silico* и *in vitro*, было установлено, что неаннотированные ранее рассматриваемые SNPs достоверно нарушают связывание ТСБ с промоторами генов человека на одном из первых этапов инициации транскрипции этих генов (Ponomarenko et al., 2013a, b). Согласно результатам независимых опытов *in vitro* и *in vivo* (Pugh, 2000; Stewart, Stargell, 2001; Mogno et al., 2010), такому изменению сродства ТСБ к промотору соответствующего гена у пациента, «маркированного» изучаемым аллелем – соответствует недостаточность либо избыток кодируемого этим геном белка.

### **3.3. Определение в режиме реального времени кинетических характеристик комплексов ТСБ с ТАТА-содержащими олигонуклеотидами, методами остановленной струи и резонансного переноса энергии**

#### *Связывание ТСБ с олигонуклеотидами, мечеными флуорофорами Су3 и Су5.*

Используя рекомбинантный ТСБ человека и флуоресцентно меченные FRET-парой флуорофоров Су3- и Су5-дуплексы, идентичные ТАТА-боксу промотора гена триозофосфатизомеразы (*TPI*) здорового человека и ТАТА-боксу, содержащему SNP - 24Т>G, методом «остановленной струи» в режиме реального времени исследован механизм процесса образования комплекса ТСБ/ТАТА. Впервые показано, что образование комплексов ТСБ с нормальным ТАТА-боксом происходит в 5.5 раза быстрее, а диссоциация – в 31 раз медленнее, чем в случае ТАТА-бокса, содержащего SNP. Для промотора гена *TPI* дикого типа требуются меньшие концентрации ТСБ, чем для SNP-содержащего ТАТА-бокса, ассоциированного с неврологическими и мышечными нарушениями, кардиомиопатией и другими заболеваниями. Из полученных результатов следует, что ТСБ при образовании комплекса одновременно связывает и изгибает ДНК (Arkova et al., 2014).

Кинетика связывания ДНК-дуплексов с ТСБ, представленная на Рис. 16 и 17, свидетельствует о том, что образование комплекса ТСБ/ТАТА приводит к росту интенсивности флуоресценции Су5. Рост интенсивности FRET-сигнала обусловлен изгибанием ДНК-дуплекса в комплексе с ТСБ, за счет чего происходит сближение остатков флуорофоров Су3 и Су5. Анализ кинетических кривых ДНК-дуплексов показал, что изгибание дуплекса, содержащего ТАТА-боксы дикого типа, происходит при меньших концентрациях ТСБ, чем в случае аллеля G ТАТА-бокса (Рис. 16 и Рис. 17). На основании полученных данных нами предложен кинетический меха-



низ связывания ТСБ с нормальным ТАТА-боксом гена *TPI* и с ТАТА-боксом, содержащим SNP, описываемый одностадийной схемой:

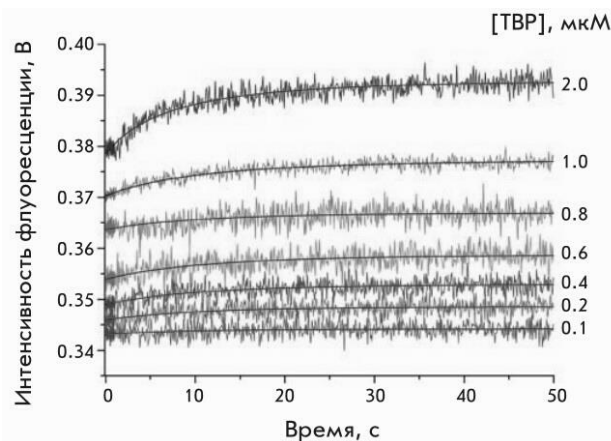
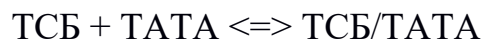


Рис. 16. Кинетика связывания и изгибания ДНК-дуплекса, идентичного последовательности ТАТА-боксов гена *TPI* дикого типа (Arkova et al., 2014).

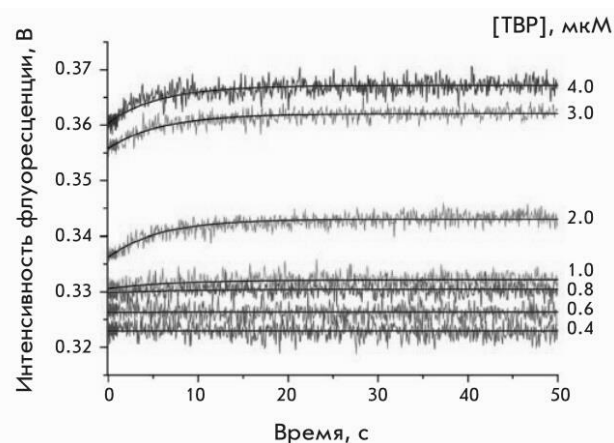


Рис. 17. Кинетика связывания и изгибания ДНК-дуплекса, идентичного последовательности содержащего SNP ТАТА-боксов промотора гена *TPI*. (Arkova et al., 2014).

Константы скоростей прямой и обратной реакции представлены в Таблице 8. Видно, что образование комплексов ТСБ с ТАТА-боксом дикого типа происходит в 5.5 раза быстрее ( $1.1 \times 10^6 \text{ M}^{-1} \text{ c}^{-1}$ ), чем с аллелем G ( $0.2 \times 10^6 \text{ M}^{-1} \text{ c}^{-1}$ ), а диссоциация комплексов ТСБ/ТАТА – в 31 раз медленнее ( $2.8 \times 10^{-3} \text{ c}^{-1}$  у дикого типа и  $8.9 \times 10^{-2} \text{ c}^{-1}$  в случае аллеля G). Необходимо отметить, что подобное различие в константах скорости образования и распада комплекса ТСБ/ТАТА приводит к отличию значений равновесных констант диссоциации в 150 раз ( $2.7 \times 10^{-9} \text{ M}$  в норме и  $0.4 \times 10^{-6} \text{ M}$  при наличии мутации). Различие в значениях констант диссоциации ( $K_D$ ) между нормальным и SNP-содержащим ТАТА-боксом говорит о резком снижении аффинности ТСБ к олигонуклеотидам с измененным ТАТА-боксом.

Полученные данные свидетельствуют о том, что возникшая в ТАТА-боксе G/C-пара делает структуру ДНК более жесткой, что затрудняет связывание ТАТА-боксов

с ТСБ и образование функционального комплекса, имеющего оптимальную конформацию. Из этого следует, что *in vivo* ген триозофосфатизомеразы, содержащий SNP -24t>G в ТАТА-боксе, транскрибируется и экспрессируется менее эффективно. Эти результаты подтверждены клинически (Watanabe et al., 1996).

Сравнение полученных нами и опубликованных данных (Watanabe et al., 1996; Humphries et al., 1999) показывает, что уменьшение в 150 раз сродства ТСБ к SNP-содержащему ТАТА-боксу промотора *TPI* приводит к повышению риска развития ряда заболеваний, связанных с недостатком триозофосфатизомеразы. Вероятно, недостаток *TPI* может компенсироваться другими путями (например, в пентозофосфатном цикле), что следует из различий в реакции больных на недостаток *TPI* в организме (Watanabe et al., 1996; Chang et al., 1993). Несмотря на то, что сродство ТСБ к SNP-содержащему ТАТА-боксу промотора гена *TPI* снижено в 150 раз, активность *TPI* в эритроцитах одних пациентов падает до 3–10% от нормы (Chang et al., 1993), а у некоторых гетерозиготных носителей этого полиморфного аллеля наблюдается умеренное (26–50% от нормы) снижение активности *TPI* (Watanabe et al., 1996).

**Таблица 8.** Кинетика взаимодействия ТСБ с ТАТА-боксом *TPI*.  $k_{on}$  – константа скорости прямой реакции ТСБ/ТАТА;  $k_{off}$  – константа скорости обратной реакции ТСБ/ТАТА;  $K_A$  – константа ассоциации, полученная из кинетических значений ( $k_{on} / k_{off}$ );  $K_D$  – константа диссоциации, полученная из кинетических значений ( $k_{off} / k_{on}$ ).

Константа	ТАТА-бокс, аллель T SNP -24T > G	ТАТА-бокс, аллель G SNP -24T > G
$k_{on}, M^{-1}c^{-1}$	$(1.1 \pm 0.1) \times 10^6$	$(0.2 \pm 0.1) \times 10^6$
$k_{off}, c^{-1}$	$(2.8 \pm 0.1) \times 10^{-3}$	$(8.9 \pm 1.2) \times 10^{-2}$
$K_{ass}, M^{-1}$	$3.7 \times 10^8$	$2.3 \times 10^6$
$K_D, M$	$2.7 \times 10^{-9} = 2.7 \text{ нМ}$	$0.4 \times 10^{-6} = 400 \text{ нМ}$

Следует отметить, что впервые путем регистрации в режиме реального времени взаимодействия ТСБ человека с флуоресцентно мечеными Су3 и Су5 ТАТА-содержащими дуплексами показано, что ТСБ быстро связывает и одновременно изгибает ДНК ТАТА-бокса гена *TPI*. Этот результат согласуется с данными, полученными ранее с использованием полноразмерного ТСБ человека и ТАТА-бокса *AdMLP* с консенсусной последовательностью 5'-GCGTATAAAAGGGC-3', к 5'-концу которой присоединен флуорофор TAMRA, а к 3'-концу – флуоресцеин (Masters et al., 2003), которые свидетельствуют об одностадийном механизме процесса связывания и одновременного изгибания ТАТА-бокса белком ТСБ.

Следует отметить, что в мире подобные исследования проводились с использованием ТСБ разного вида – полной и укороченной форм (С-концевого домена) – и, в основном, только модельного промотора AdML (реже E4) с консенсусной последовательностью ТАТА-бокса. Полученные результаты на реальных промоторах ге-

нов человека улучшили представление о взаимодействии ТСБ/ТАТА, ключевом в инициации и регуляции транскрипции и синтеза белков в клетках эукариот.

**Связывание ТСБ с олигонуклеотидами, мечеными флуорофорами TAMRA и FAM.** Как известно, клеточные процессы высоко динамичны. На сегодняшний день динамика транскрипции и других молекулярных процессов плохо понята из-за отсутствия методов измерения кинетических параметров взаимодействия *in vivo*. Измерение взаимодействия ТСБ с ТАТА-содержащими олигонуклеотидами в режиме реального времени поможет лучше понять начальный этап процесса транскрипции – взаимодействие ТСБ/ТАТА, запускающее сборку транскрипционного комплекса. Для более тонкого наблюдения различий в структуре ДНК при взаимодействии с ТСБ в режиме реального времени мы использовали пару флуорофоров TAMRA и FAM с более высоким квантовым выходом и, соответственно, чувствительностью детекции, чем у Cy3/Cy5.

В представленной работе изучали взаимодействие рекомбинантного полноразмерного ТСБ человека с ТАТА-элементами промоторов реальных генов человека, *LEP* и *TPI*, с их анцестральными и минорными аллелями. Известно, что минорный аллель *TPI* (SNP – 24Т>G) ассоциирован с повышенным риском возникновения таких заболеваний, как прогрессирующие неврологические и мышечные нарушения, рак легкого и др. (Савинкова и др., 2009). Нами исследуется возможное фенотипическое проявление минорного аллеля гена *LEP* относительно анцестрального аллеля референсного генома. В работе использовались двуцепочечные олигодезокирибонуклеотиды (ОДН) длиной 15 и 26 вр, 5'-концы которых метились FRET-парой флуорофоров TAMRA и FAM. Процесс образования комплексов ТСБ с ОДН изу-

чали с помощью метода остановленной струи с регистрацией FRET сигнала в режиме реального времени. Показано, что образование комплексов ТСБ с ОДН, идентичными ТАТА-элементам с фланкирующими последовательностями анцестральных аллелей генов *LEP* и *TPI* человека и минорных, при содержании SNPs (SNP - single nucleotide polymorphism), сопровождается множественными изменениями изгиба ДНК и образованием тройных комплексов ТСБ/ДНК/ТСБ, имеющих разные скорости образования и диссоциации для разных ТАТА-элементов. Впервые для изучения взаимодействия ТСБ/ТАТА использовались последовательности ТАТА-элементов – atcgggccgcTA/GTAAGAggggcgggc и cgcggcgctcTAT/GAAGtgggcagt - промоторов реальных генов *LEP* и *TPI*, соответственно, в окружении реальных фланкирующих нуклеотидов, также влияющих на взаимодействие.

Первичное изменение структуры ДНК при взаимодействии с ТСБ происходит в миллисекундном диапазоне времени, поэтому для изучения этого процесса был использован кинетический метод остановленной струи. Конформационные изменения ДНК-дуплексов, содержащих ТАТА-элементы промоторов генов *LEP* и *TPI*, при взаимодействии с ТСБ были зарегистрированы методом остановленной струи по изменению FRET-сигнала. В качестве неспецифической последовательности ДНК были использованы дуплексы ОДН, не содержащие ТАТА-элемент (Таблица 3). Кроме того, для выяснения роли длины фланкирующей последовательности в работе были использованы ДНК-дуплексы длиной 15 и 26 пар оснований. FRET-красители, TAMRA и FAM, были введены на 5'-концы ОДН (Приложение 3).

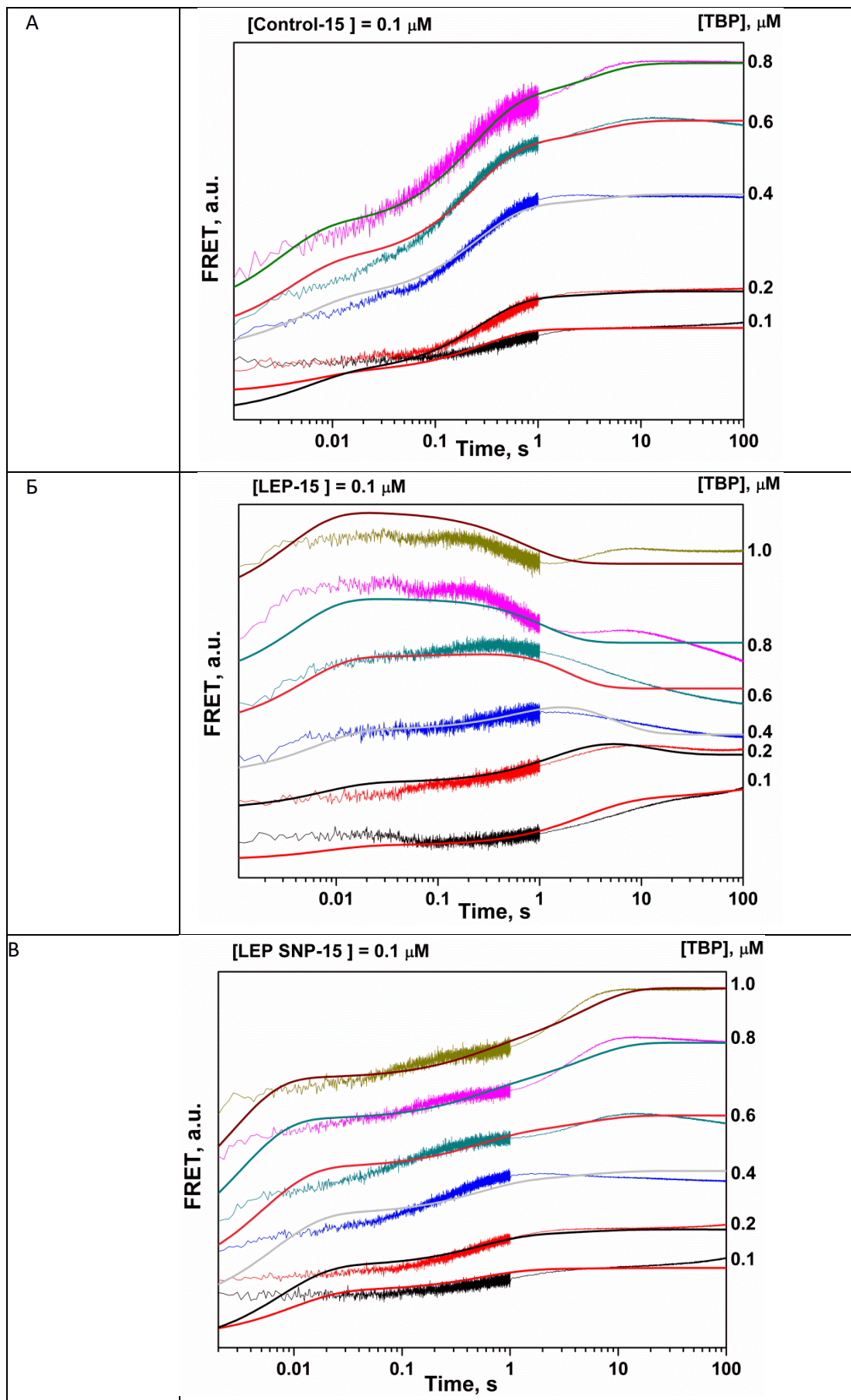
Используемые в работе дуплексы принадлежат генам *TPI* и *LEP*.

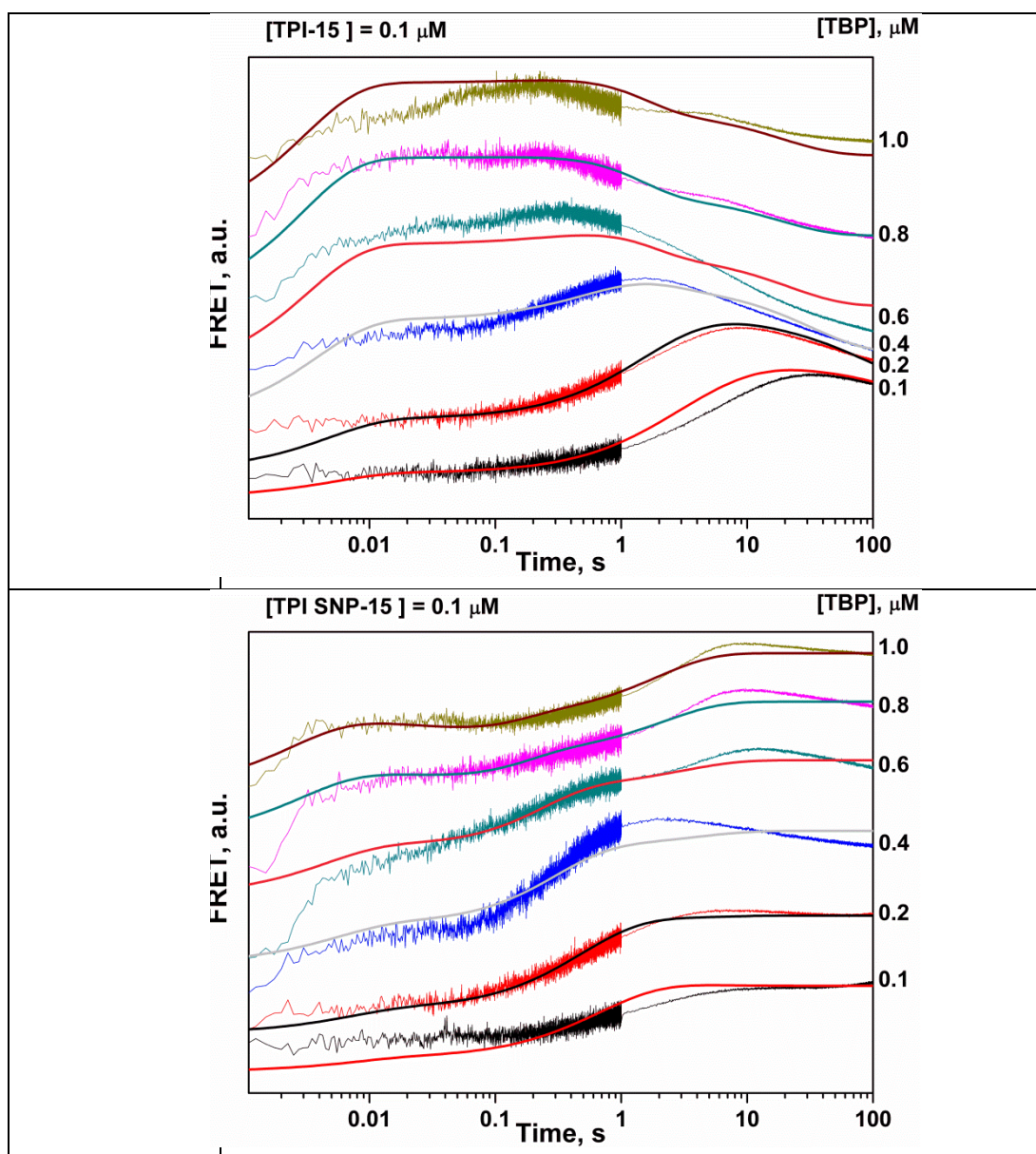
***Взаимодействие ТСБ с олигодезоксирибонуклеотидами длиной 15 бп, идентичными ТАТА-элементам промоторов генов LEP и TPI.***

Кинетика связывания ТСБ с контрольным ДНК-дуплексом, не содержащим ТАТА-бокса, свидетельствует об образовании неспецифического комплекса, в котором происходит сближение пары FRET-красителей, за счет изгибания двойной спирали (Рисунок 18А). Анализ концентрационной серии кинетических кривых показал, что первая фаза роста FRET-сигнала до 1 с описывается минимальной кинетической схемой, в которой на начальном этапе взаимодействия образуется первичный комплекс  $E \cdot S_1$ , который затем превращается в комплекс  $E \cdot S_2$ . Вторая фаза роста характеризует связывание второй молекулы ТСБ с комплексами  $E \cdot S_1$  и  $E \cdot S_2$  (Схема 1). Константы скоростей прямой и обратной реакции представлены в Таблице 9.

**Таблица 9.** Константы скоростей и равновесия, характеризующие взаимодействие ТСБ с 15-ти звенными ДНК-дуплексами.

<b>Constants</b>	<b>Control-15</b>	<b>LEP-15</b>	<b>LEP SNP-15</b>	<b>TPI-15</b>	<b>TPI SNP-15</b>
$k_1, M^{-1}c^{-1}$	$(2.0 \pm 0.3) \times 10^8$	$(2.2 \pm 0.7) \times 10^8$	$(2.2 \pm 0.3) \times 10^8$	$(1.8 \pm 0.3) \times 10^8$	$(2.7 \pm 0.4) \times 10^8$
$k_{-1}, c^{-1}$	120±20	80±30	90±30	130±30	90±10
$K_1, M$	$6.0 \times 10^{-7}$	$3.6 \times 10^{-7}$	$4.1 \times 10^{-7}$	$7.2 \times 10^{-7}$	$3.3 \times 10^{-7}$
$k_2, c^{-1}$	1.0±0.5	1.2±0.6	0.7±0.5	3.8±1.4	2.4±1.6
$k_{-2}, c^{-1}$	1.6±0.8	0.03±0.02	1.0±0.4	0.04±0.02	0.3±0.1
$K_2$	1.6	0.025	1.4	0.01	0.12
$k_3, M^{-1}c^{-1}$	$(9.7 \pm 6.3) \times 10^6$	$(2.9 \pm 0.9) \times 10^6$	$(1.0 \pm 0.7) \times 10^6$	$(2.3 \pm 0.6) \times 10^6$	$(13 \pm 5) \times 10^6$
$k_{-3}, c^{-1}$	7.6±5.5	0.4±0.2	1.2±1.0	0.9±0.6	9.9±5.2
$K_3, M$	$0.78 \times 10^{-6}$	$0.14 \times 10^{-6}$	$1.2 \times 10^{-6}$	$0.4 \times 10^{-6}$	$0.76 \times 10^{-6}$
$k_4, c^{-1}$		0.9±0.7	0.05±0.02	0.17±0.05	1.1±0.3
$k_{-4}, c^{-1}$		0.02±0.01	0.14±0.04	0.02±0.01	0.16±0.04
$K_4$		0.02	2.8	0.12	0.15

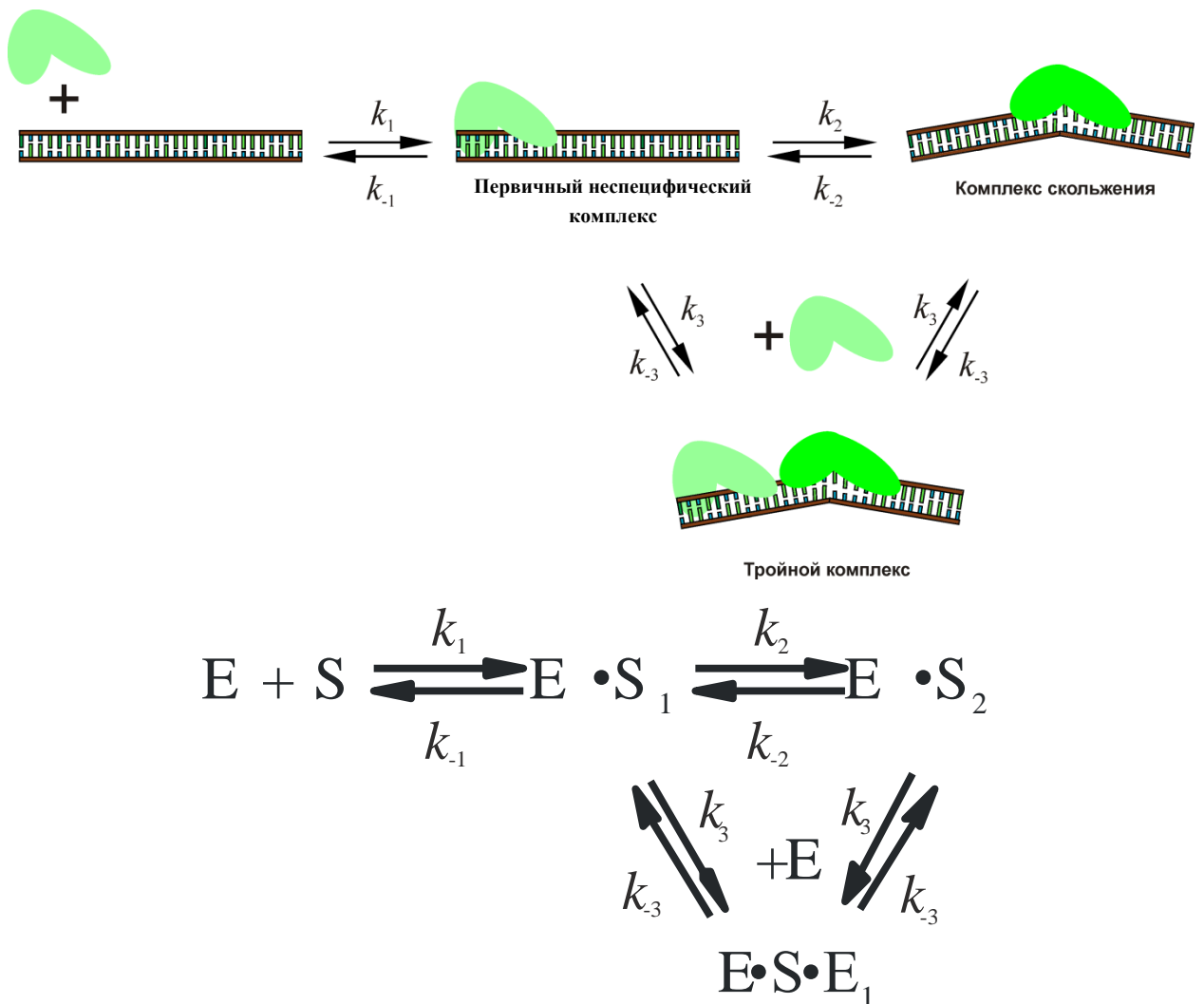




**Рисунок 18.** Кинетика связывания и изгибания коротких 15-ти звенных ДНК-дуплексов. (А) Образование неспецифического комплекса. (Б) ТАТА-элемент анцестрального аллеля гена *LEP* референсного генома человека Hg19. (В) ТАТА-элемент минорного аллеля гена *LEP* (SNP - 35A>G). (Г) ТАТА-элемент анцестрального аллеля гена *TPI* референсного генома человека Hg19. (Д) ТАТА-элемент минорного аллеля гена *TPI* (SNP -24T>G).



**Схема 1.** Кинетический механизм взаимодействия ТСБ с ДНК-дуплексом, не содержащим ТАТА-элемента.



Кинетика связывания ТСБ с ДНК-дуплексами, содержащими ТАТА-элемент (Рисунок 18Б-18Д) значительно отличается от контрольных кривых, что свидетельствует о регистрации специфических взаимодействий при формировании комплекса. Первая фаза роста FRET-сигнала завершается в течение 10 мс, после этого происходит медленное увеличение сигнала. Необходимо отметить, что вторая фаза роста FRET-сигнала имеет выраженную концентрационную зависимость. Медленный процесс, сопровождающийся ростом FRET-сигнала примерно до 50 с при соотно-

шении ТСБ:ДНК = 1:1 ускоряется примерно в 100 раз при соотношении ТСБ:ДНК = 10:1 (заканчивается за 0,5 с) (Arkova et al., 2016). После второй фазы роста на кинетических кривых при высоких концентрациях ТСБ происходит уменьшение FRET-сигнала с последующим ростом и повторным падением, что свидетельствует о происходящих конформационных изменениях дуплекса в процессе взаимодействия с ТСБ.

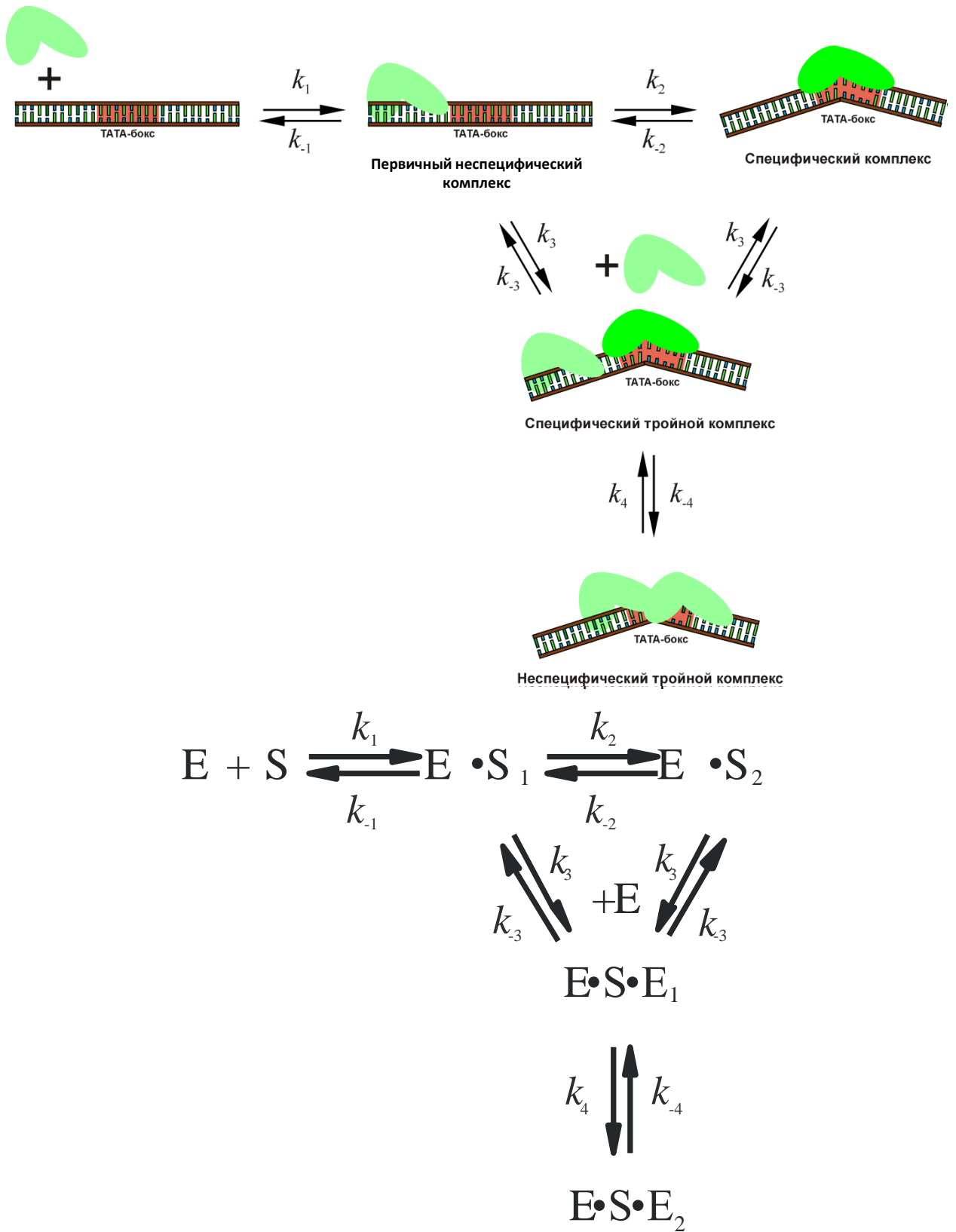
Сложный профиль кинетических кривых свидетельствует о многостадийном механизме образования комплекса ТСБ с ДНК-дуплексами, содержащими ТАТА-элемент. На основании анализа кинетических кривых с помощью программы DynaFit была определена минимальная кинетическая модель (Приложение 2), представленная на Схеме 2, удовлетворительно описывающая полученные результаты (Таблица 9). Согласно Схеме 2, специфическое связывание ТСБ с ДНК-дуплексами, содержащими ТАТА-элемент, проходит через образование промежуточного первичного комплекса  $E \cdot S_1$ . Интересно отметить, что рост FRET-сигнала при образовании первичного комплекса  $E \cdot S_1$  свидетельствует о сближении концов модельных дуплексов, что может произойти за счет изгибания ДНК. Более того, последующая трансформация первичного комплекса  $E \cdot S_1$  в специфический комплекс  $E \cdot S_2$  приводит к дальнейшему росту FRET-сигнала за счет изгиба ДНК. Однако связывание второй молекулы ТСБ приводит к уменьшению FRET-сигнала, и свидетельствует о том, что в тройных комплексах  $E \cdot S \cdot E_1$  и  $E \cdot S \cdot E_2$  происходит либо экранирование флуорофоров белковой молекулой, либо релаксация изгиба ДНК, приводящая к увеличению расстояния между флуорофорами.

Интересно отметить, что константа диссоциации  $k_1$  первичного комплекса  $E \cdot S_1$  имеет схожую величину для всех ДНК-дуплексов (Таблица 9). В то же время кон-

станта диссоциации  $k_2$  специфического комплекса  $E \cdot S_2$  значительно отличается для использованных ДНК-дуплексов и соответствует аффинности ТСБ к ТАТА-элементам. На основании значения константы  $k_2$  специфичность ТСБ к ДНК-дуплексам увеличивается в ряду Control-15 < LEP SNP-15 < TPI SNP-15 < LEP-15 < TPI-15.

Сравнение константы диссоциации  $k_3$  тройного комплекса  $E \cdot S \cdot E_1$  показывает, что комплекс более стабилен для LEP-15 и TPI-15, чем для Control-15, LEP SNP-15 и TPI SNP-15. Константы равновесия  $K_4$  между тройным комплексом  $E \cdot S \cdot E_1$  и комплексом  $E \cdot S \cdot E_2$  имеют близкие значения для TPI SNP-15 и TPI-15, однако отличаются примерно в 100 раз для LEP-15 и LEP SNP-15.

**Схема 2.** Кинетический механизм взаимодействия ТСБ с ДНК-дуплексами, содержащими ТАТА-элемент.



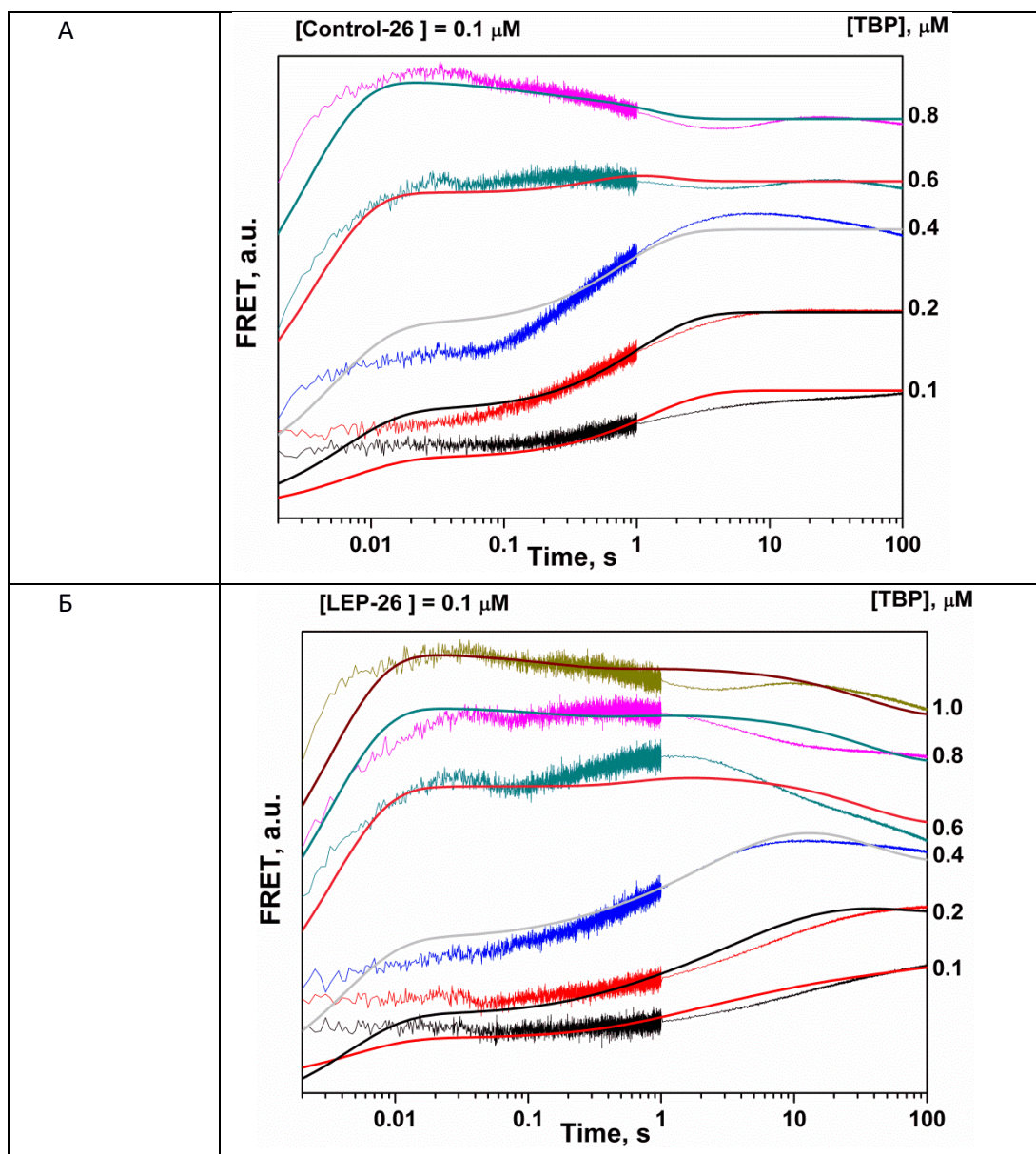
Взаимодействие ТСБ с олигодезоксирибонуклеотидами длиной 26 bp, идентичными ТАТА-элементам промоторов генов *LEP* и *TPI*.

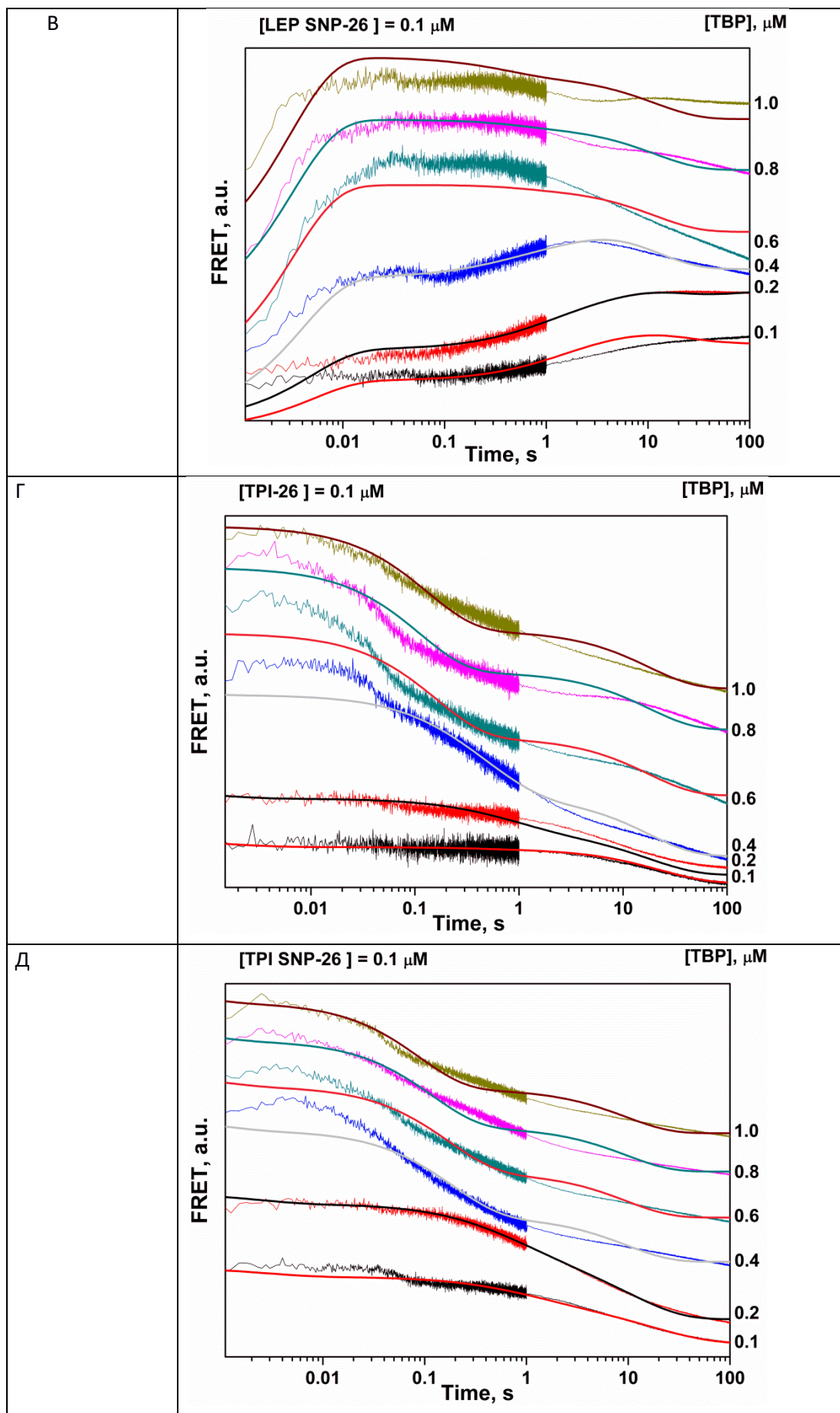
Кинетические кривые взаимодействия ТСБ с более длинным 26-ти звенным ДНК-дуплексом, не содержащим ТАТА-элемент представлены на Рисунке 19А. Интересно отметить, что профиль отличается от кривых, полученных для короткого 15-ти звенного контрольного дуплекса, однако последовательность фаз роста и падения FRET-сигнала совпадает с LEP-15 и TPI-15 (Рис. 18А, 18Б, 18Г и 19А). Эти данные свидетельствуют о том, что длина дуплекса является одним из важных факторов, значительно влияющих на стадии взаимодействия ТСБ и ДНК, как отмечалось и в работе (Hieb et al., 2014).

Интересно отметить, что кинетические кривые, полученные для 26-ти звенных дуплексов, содержащих ТАТА-элемент гена *LEP* (LEP-26 и LEP SNP-26) имеют похожий профиль с контрольным дуплексом Control-26. В то время как дуплексы, содержащие ТАТА-элемент гена *TPI* значительно отличаются и от Control-26 и от LEP-26 и LEP SNP-26 (см. Рис. 19). Возможно, это вызвано тем, что дуплекс TPI более АТ-богатый, чем LEP, и имеет более подвижную конформацию. Возможно, ввиду более высокого сродства ТСБ к ТАТА-боксу *TPI* ( $K_D = 4.8$  nM) по сравнению с *LEP* ( $K_D = 80$  nM) (Drachkova et al., 2014; Arkova et al., 2015), происходит «залипание» ТСБ на ТАТА-боксе, что уменьшает количество актов «связывания-диссоциации» в процессе образования энергетически устойчивого комплекса.

Анализ кинетических кривых проводили по Схеме 2, полученные константы скорости и равновесия представлены в Таблице 10. Сравнение данных Таблиц 9 и 10 показывает, что формирование первичного комплекса  $E \cdot S_1$  фактически не зависит от длины дуплекса и его состава.

Образование специфического комплекса  $E \cdot S_2$  проходит более эффективно в случае присутствия ТАТА-элементов. По сравнению с контрольным дуплексом Control-26, константа диссоциации  $k_2$  комплекса  $E \cdot S_2$  примерно в 3 и 10 раз меньше для LEP-элемента и TPI-элемента, соответственно. Так же как и в случае коротких 15-ти звенных дуплексов формирование тройного комплекса  $E \cdot S \cdot E_1$  и его трансформация в комплекс  $E \cdot S \cdot E_2$  проходит более эффективно в случае ДНК-дуплексов, содержащих ТАТА-элемент.





**Рисунок 19.** Кинетика связывания и изгибания коротких 26-ти звенных ДНК-дуплексов. (А) Образование неспецифического комплекса. (Б) ТАТА-элемент анцестрального аллеля гена *LEP* референсного генома человека Hg19. (В) ТАТА-элемент минорного аллеля гена *LEP* (SNP - 35A>G). (Г) ТАТА-элемент анцестрального аллеля гена *TPI* референсного генома человека Hg19. (Д) ТАТА-элемент минорного аллеля гена *TPI* (SNP -24T>G)

**Таблица 10.** Константы скорости и равновесия, характеризующие взаимодействие ТСБ с 26-ти звенными ДНК-дуплексами.

<i>Constants</i>	<i>Control-26</i>	<i>LEP-26</i>	<i>LEP SNP-26</i>	<i>TPI-26</i>	<i>TPI SNP-26</i>
$k_1, M^{-1}c^{-1}$	$(2.0 \pm 0.2) \times 10^8$	$(1.8 \pm 0.2) \times 10^8$	$(2.0 \pm 0.2) \times 10^8$	$(6.4 \pm 3.3) \times 10^8$	$(2.5 \pm 1.4) \times 10^8$
$k_{-1}, c^{-1}$	120±10	150±20	140±20	70±50	140±80
$K_1, M$	$6.0 \times 10^{-7}$	$8.3 \times 10^{-7}$	$7.0 \times 10^{-7}$	$1.1 \times 10^{-7}$	$5.6 \times 10^{-7}$
$k_2, c^{-1}$	0.64±0.22	0.31±0.05	0.46±0.14	1.3±0.6	1.3±1.1
$k_{-2}, c^{-1}$	0.58±0.05	0.08±0.05	0.18±0.10	0.10±0.08	0.05±0.03
$K_2$	0.9	0.26	0.39	0.08	0.04
$k_3, M^{-1}c^{-1}$	$(6.3 \pm 5.6) \times 10^6$	$(8.8 \pm 6.3) \times 10^6$	$(6.8 \pm 3.2) \times 10^6$	$(10.5 \pm 5.4) \times 10^6$	$(9.7 \pm 6.2) \times 10^6$
$k_{-3}, c^{-1}$	0.8±0.3	0.05±0.04	0.7±0.3	0.29±0.22	0.33±0.21
$K_3$	$0.13 \times 10^{-6}$	$0.006 \times 10^{-6}$	$0.10 \times 10^{-6}$	$0.028 \times 10^{-6}$	$0.034 \times 10^{-6}$
$k_4, c^{-1}$	1.9±0.8	7.0±0.8	4.3±3.0	0.054±0.03	0.06±0.03
$k_{-4}, c^{-1}$	3.2±1.3	0.03±0.01	0.07±0.05	0.02±0.01	0.018±0.011
$K_4$	1.7	0.004	0.016	0.37	0.3

Нами впервые были определены конформационные переходы ТАТА-содержащих ДНК-дуплексов, идентичных анцестральным и минорным аллелям реальных генов *LEP* и *TPI* человека, при образовании комплексов с ТСБ в режиме ре-



ального времени. Полученные данные свидетельствуют о том, что формирование специфического комплекса  $E \cdot S_2$  является ключевой стадией распознавания ТАТА-элемента в ДНК-дуплексах длиной 15 и 26 п.о. и проходит через образование промежуточного первичного комплекса  $E \cdot S_1$ . При повышении концентрации ТСБ и длины дуплексов наблюдается увеличение гетерогенности конформационных переходов ДНК («изгиб — исходное состояние»). Формирование тройного комплекса  $E \cdot S \cdot E_1$  и его трансформация в комплекс  $E \cdot S \cdot E_2$  объясняется способностью ТСБ связываться с неспецифическими участками ДНК и, по-видимому, может иметь регуляторный характер *in vivo*.

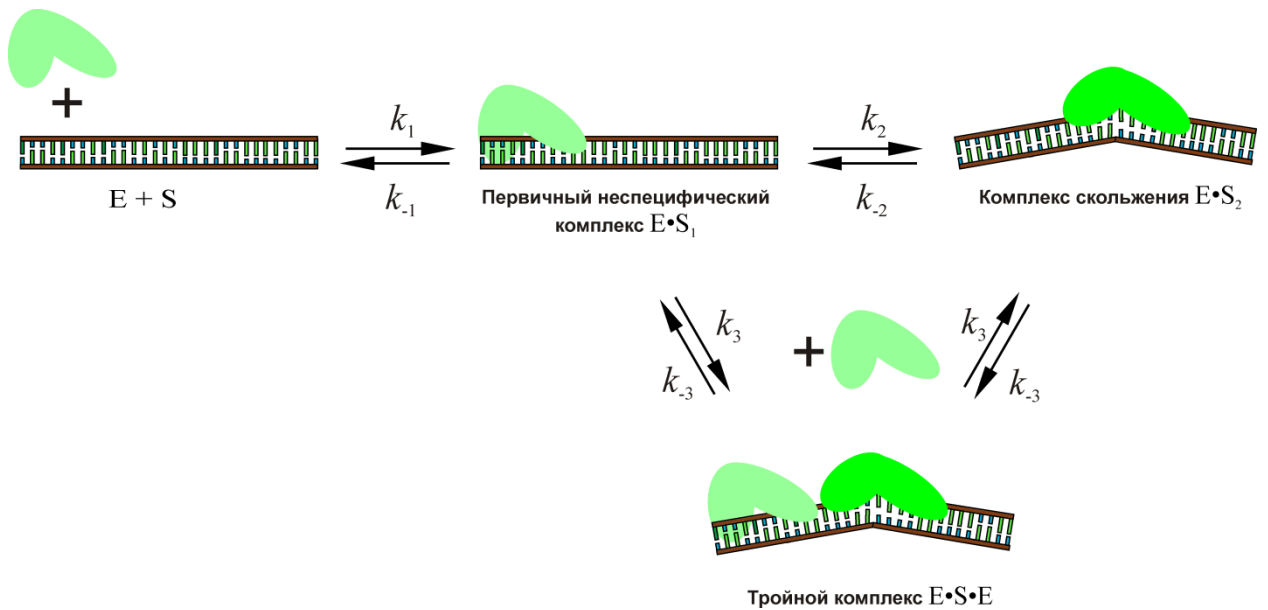
Как видно из Таблицы 10, константа диссоциации  $k_1$  первичного комплекса ТСБ с ДНК, имеет схожую величину для всех ДНК-дуплексов, что свидетельствует об отсутствии их специфичности (Control-26). Также примерно одинаковы скорости образования и диссоциации первичных, неспецифических комплексов ( $k_1, M^{-1} s^{-1}$  и  $k_{-1}, s^{-1}$ ).

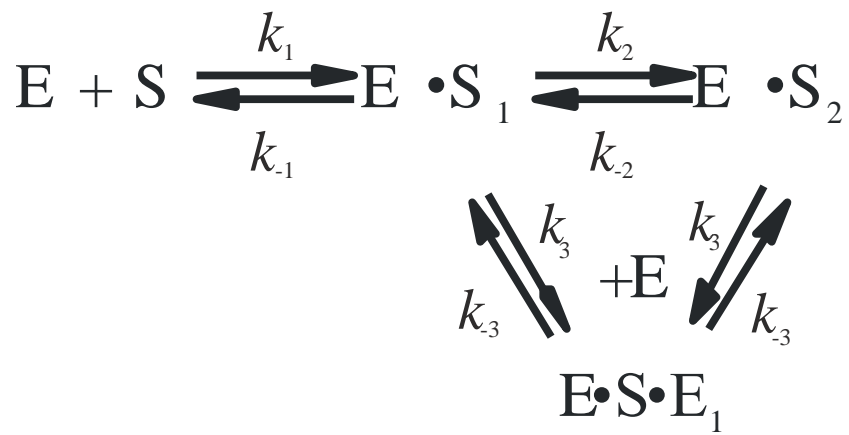
Константа диссоциации  $k_2$ , характеризующая аффинность ТСБ/ТАТА промежуточного специфического комплекса примерно в 3 и 10 раз меньше для LEP-элемента и TPI-элемента, соответственно. Формирование специфического комплекса происходит более эффективно в случае ДНК-дуплексов, содержащих ТАТА-элемент. На основании значений константы  $K_2$ , специфичность ТСБ к ДНК-дуплексам увеличивается в ряду Control-15 < LEP SNP-15 < TPI SNP-15 < LEP-15 < TPI-15.

Итак, нами впервые в режиме реального времени определены конформационные переходы ТАТА-содержащих ДНК-дуплексов, идентичных анцестральным и минорным аллелям реальных генов *LEP* и *TPI* человека, при образовании комплек-

сов с ТСБ. Сложный профиль кинетических кривых свидетельствует о многостадийном механизме образования комплекса ТСБ с ДНК-дуплексами, содержащими ТАТА-элемент. Полученные данные свидетельствуют о том, что формирование специфического комплекса ТСБ/ТАТА происходит через образование промежуточных комплексов. При повышении концентрации ТСБ и длины дуплексов наблюдается увеличение гетерогенности конформационных переходов ДНК («изгиб — исходное состояние») (Watanabe et al., 1996; Friedman, 2011).

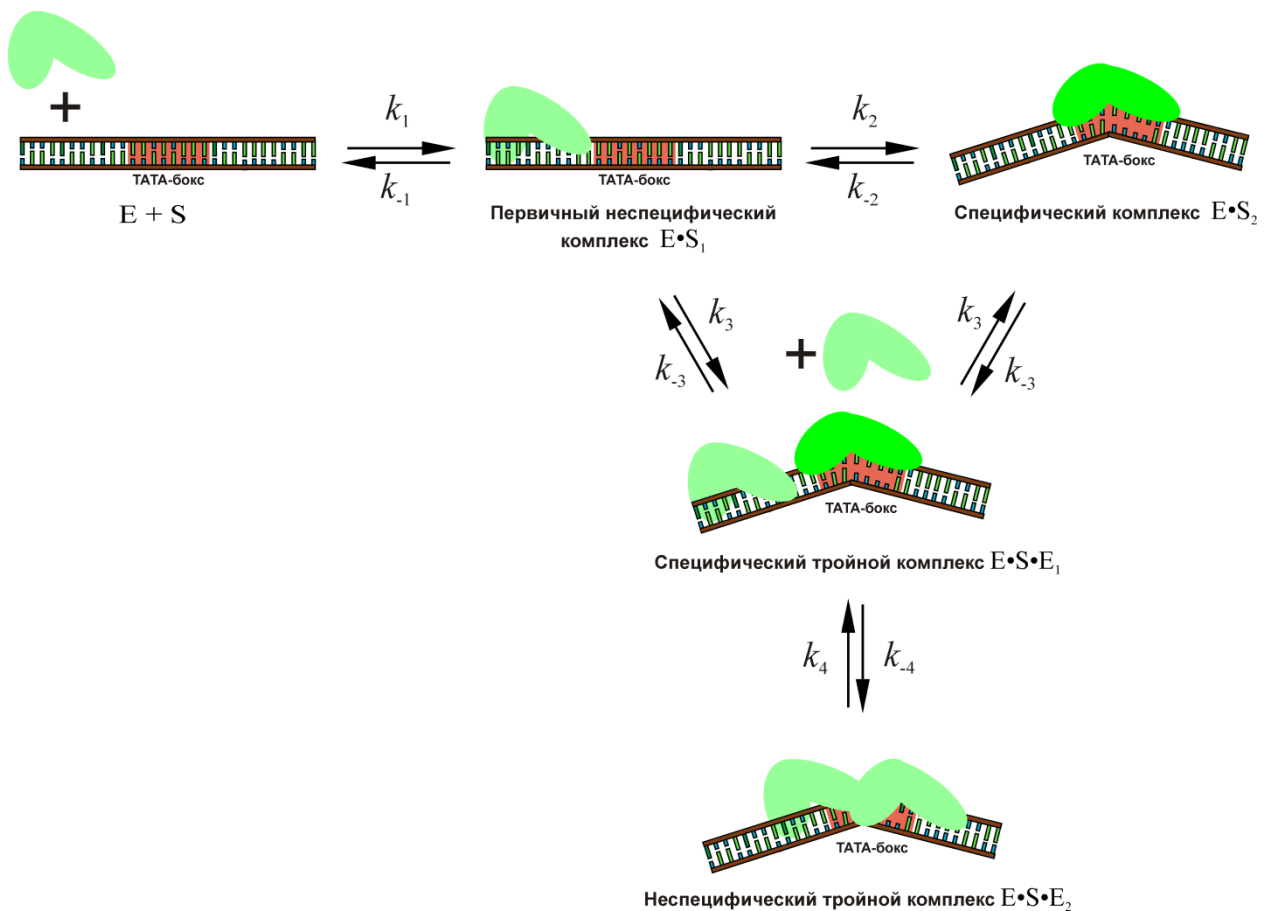
**Схема 3.** Кинетический механизм взаимодействия ТСБ с ДНК-дуплексом, не содержащим ТАТА-элемент.

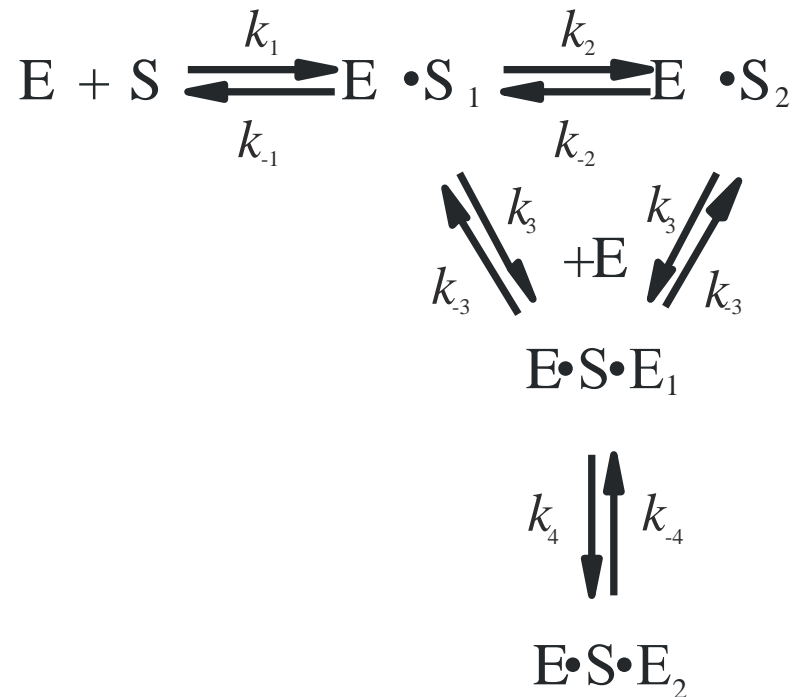




На основании проведенных опытов нами была предложена схема взаимодействия ТСБ/ТАТА.

**Схема 4.** Кинетический механизм взаимодействия ТСБ с ДНК-дуплексами, содержащими ТАТА-элемент.





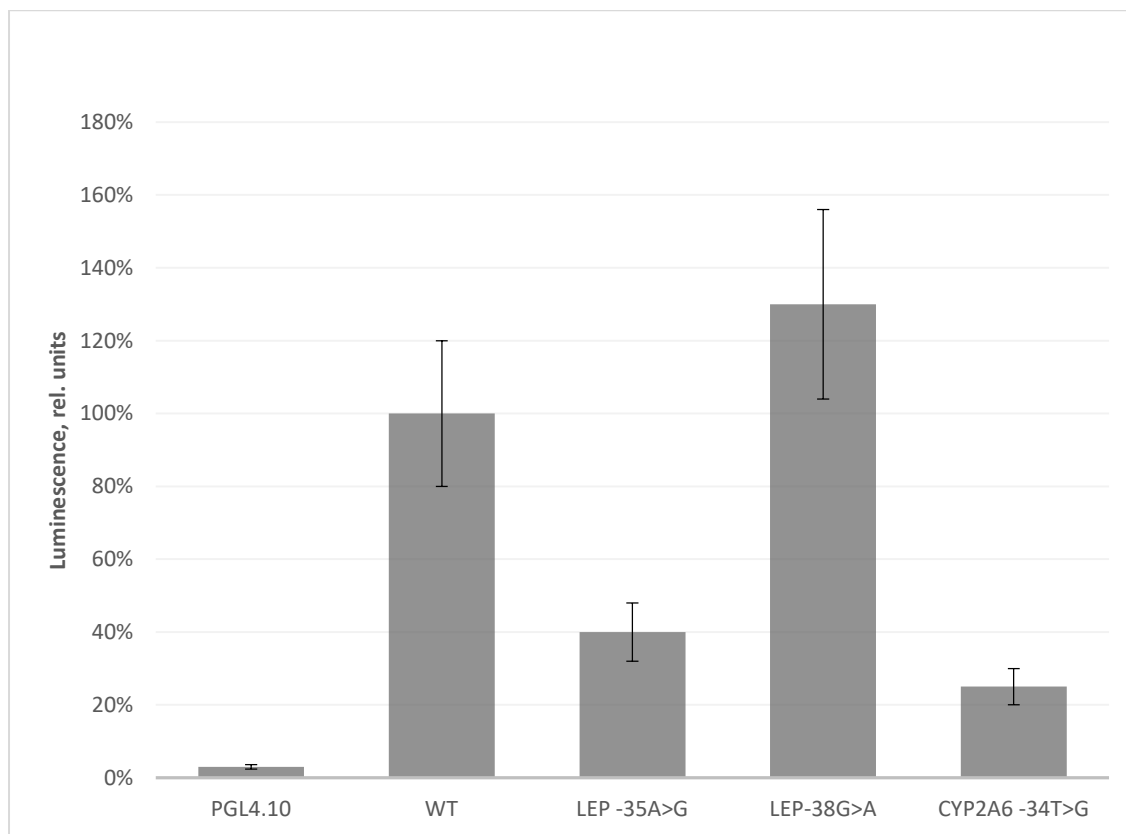
Итак, впервые определены конформационные переходы ТАТА-содержащих ДНК-дуплексов, идентичных анцестральным и минорным аллелям реальных генов *LEP* и *TPI* человека, при образовании комплексов с ТСБ в режиме реального времени. Полученные данные свидетельствуют о том, что формирование специфического комплекса  $\text{E} \cdot \text{S}_2$  - ключевой стадии распознавания ТАТА-элемента в ДНК-дуплексах длиной 15 и 26 п.о. - проходит через образование неспецифического  $\text{E} \cdot \text{S}_1$ . При повышении концентрации ТСБ и длины дуплексов наблюдается увеличение гетерогенности конформационных переходов ДНК («изгиб — исходное состояние»). На основании полученных схем можно предположить, что образование специфических комплексов в наших условиях может быть результатом связывания с ОДН как одной молекулы ТСБ, так и двух (возможно, и трех молекул при длине ОДН, равной 26 bp). Формирование тройного комплекса  $\text{E} \cdot \text{S} \cdot \text{E}_1$  и его трансформа-

ция в неспецифический комплекс  $E \cdot S \cdot E_2$  объясняется также способностью ТСБ связываться с неспецифическими участками ДНК и, по-видимому, имеет регуляторный характер *in vivo*.

### **3.4. Изучение влияния анцестральных и минорных аллелей генов *LEP* и *CYP2A6*, на экспрессию репортерного гена *LUC***

Для экспериментального изучения влияния выявленных SNPs ТАТА-боксов на транскрипцию репортерного гена *LUC* использовали эукариотические линии клеток MCF-7 и HCT116 (аденокарцинома молочной железы и колоректальная карцинома человека, соответственно). ТАТА-содержащие участки промоторов, WT и минорные аллели генов *LEP* и *CYP2A6* клонировали в беспромоторный вектор pG4.10 по сайтам рестрикции XhoI и Hind III.

Высота серых столбиков (Рис. 20) и их погрешности соответствуют среднему значению оценок и их стандартных отклонений, рассчитанных по меньшей мере в трех независимых измерениях.



**Рисунок 20.** Культурально-клеточная верификация некоторых маркерных SNPs генов *LEP* (-35A>G rs201381696 и -38G>A rs200487063) и *CYP2A6* (-34T>G rs28399433) на клеточных линиях человека MCF-7 и HCT-116, трансфицированных pGL 4.10 вектором, содержащим репортерный ген *LUC* (слева – контроль для WT в виде вставки беспромоторного вектора pG4.10; следующий показатель, содержащий WT промотор, – контрольный для остальных, SNP-содержащих ТАТА-боксов).

На графике показаны 95% доверительные интервалы для люминисценции *LUC* при разных вставках. Люминисценция для pGL 4.10 значимо ниже, чем люминисценция для WT промотора, поэтому дальнейшие сравнения следует проводить с WT промотором, а не с pGL 4.10. Далее люминисценция как для *LEP* -35A>G, так и для *CYP2A6* -34T>G значимо ниже, чем люминисценция для WT промотора, откуда

можно сделать вывод, что наличие таких полиморфизмов ухудшает связывание ТСБ/ТАТА комплекса и снижает последующую транскрипцию. Данное заключение совпадает с предсказанным *in silico* эффектом, а также с экспериментально полученными данными методом EMSA.

Для *LEP* -38G>A люминисценция выше, чем для WT. Различие статистически значимо на уровне 10%. Это означает, что -38G>A rs200487063 «улучшает» промотор увеличивая сродство ТСБ к ТАТА-боксу, что ведет к увеличению транскрипции с *LUC*. Такой результат был предсказан *in silico*, а также совпадает с экспериментально полученными данными методом EMSA.

### **3.5. Анализ возможного фенотипического проявления SNPs в ТАТА – боксах промоторов генов**

Как известно, генетическая вариабельность вносит большой вклад в заболевание, чем эпидемиологическая (Lee et al., 2016), к которой относятся различные факторы окружающей среды. Генетические варианты как предикторы имеют четыре преимущества над клиническими: 1) они остаются неизменными всю жизнь; 2) они могут быть измерены; 3) их можно объединять и использовать вместе с другими параметрами для предсказания развития заболевания; 4) с их помощью можно предсказать пожизненный риск пациенту в отличие от клинических факторов, которые могут использоваться в данный момент.

Необходимо иметь в виду, что предсказательная модель не является диагностическим инструментом, с ее помощью можно оценивать вероятность развития заболевания в будущем. Прогностические модели лучше строить для каждого конкретного случая, основываясь на сочетании генетической, эпидемиологической и

клинической вариабельности, так как вместе они создают уникальную для каждого пациента характеристику (Zhao et al., 2016).

### Заключение

Трудно переоценить роль SNPs: они могут использоваться для эволюционных исследований, картирования генов и локусов комплексных заболеваний. Всестороннее изучение механизмов влияния конкретных SNPs на здоровье человека делает более эффективным лечение различных заболеваний, позволит определить, будет ли человек воспринимать данное лекарство или реагировать лишь на его побочные эффекты. Фармацевты смогут разрабатывать индивидуальные лекарственные препараты для пациентов с аналогичными клиническими симптомами и фенотипами заболеваний.

В последние годы системная медицина, возникшая из системной биологии, становится краеугольным камнем в изучении комплексных или многофакторных заболеваний. Комплексные заболевания характеризуются гетерогенными клиническими проявлениями. Они являются результатом нарушения взаимодействия генетического материала пациента и факторов внутренней и внешней среды. Несмотря на огромный поток исследований, наше понимание этиологии комплексных заболеваний остается далеко не полным. Для достижения наиболее полного понимания последствий происходящих нарушений необходим комплексный подход, включающий рассмотрение всех возможных взаимодействий, происходящих *in vivo*. Это включает и моделирование событий *in vitro*, происходящих с самых первых шагов экспрессии гена – в частности, взаимодействия ТСБ/ДНК, запускающего сборку



транскрипционного комплекса. Авторы (Rhee and Pugh, 2011) считают, что внутренняя аффинность связывания транскрипционного фактора с ДНК не является определяющим фактором взаимодействия *in vivo*. Игрет роль комбинация таких факторов, как кооперативный эффект, локальная концентрация и конкуренция транскрипционных факторов за сайты связывания с другими транскрипционными факторами и нуклеосомами. Вклад этих составляющих может сильно различаться для разных сайтов ДНК. Проведенные нами ранее исследования (Savinkova et al., 2013; Drachkova et al., 2014) показали, что, несмотря на сложность картины взаимодействия ТСБ/ДНК *in vivo*, огромное количество регуляторных событий, происходящих *in vivo*, взаимодействия ТСБ/ДНК *in vitro* адекватно отражают эти события: предсказанное снижение аффинности ТСБ к аллельным вариантам ТАТА-боксов и подтвержденное экспериментально, коррелирует с изменением количества выхода продуктов генов. В случае генов бэта – глобинов это приводит к возникновению моногенного заболевания - различной тяжести бэта-талассемий, в других случаях повышает риск возникновения гипертонии, рака, сердечно-сосудистых и других комплексных заболеваний.

При сердечно-сосудистых заболеваниях наряду с генетической предрасположенностью большое значение имеет нарушение веса, приводящее к ожирению. Распространенность ожирения продолжает расти, по данным ВОЗ более 50% европейцев в настоящее время классифицируется с избыточным весом (Holvoet, 2008). В результате этого возникает легочная и сердечная недостаточность и повышается вероятность возникновения атеросклероза. Учитывая сложный характер комплексных заболеваний, информация, полученная на разных уровнях исследований, поз-

волит в полной мере изучить особенности этих заболеваний, чтобы лучше понять их патогенез и определить новые мишени для лекарственной терапии.

В работе сделан анализ SNPs кор-промоторных районов [-70; -20], содержащих ТАТА-боксы генов *LEP*, *ABCA9*, *GCG* и еще 61 гена пищевого поведения. Анализ кор-промотора сделан с использованием Web-сервиса SNP\_TATA\_Comparator (Рассказов и др., 2013), <http://beehive.bionet.nsc.ru/cgi-bin/mgs/tatascan/start.pl>, который находится в свободном доступе для анализа влияния мутаций в районе [-70; -20] промоторов генов человека на экспрессию этих генов, включая оценки сродства ТСБ/ДНК *in vitro* (Ponomarenko et al., 2009). Web-сервис создан на основании экспериментальной верификации (Savinkova et al., 2013; Drachkova et al., 2014; Arkova et al., 2014; Драчкова и др., 2012) ранее разработанного метода (Рассказов и др., 2013) для оценки сродства ТСБ к ТАТА-боксам промоторов и определения границ его применимости (Suslov et al., 2010; Suslov et al., 2010; Ponomarenko, 2015; Ponomarenko et al., 2010). Полученные с использованием Web-сервиса SNP\_TATA\_Comparator прогнозы влияния SNPs кор-промотора генов *LEP*, *ABCA9*, *GCG*, *TPI* и *CYP2A6* на взаимодействие ТАТА-связывающего белка подтверждены нами экспериментально с использованием классического метода EMSA. Получены кинетические характеристики ( $k_a$  и  $k_d$ ) образования комплексов ТСБ/ODN, а также рассчитанные значения сродства ТСБ/ТАТА,  $K_D$ , и его изменения ( $\delta$ ). Получены расчетные  $t_{1/2}$  и  $-\Delta G$ .

Для трех рассмотренных SNPs ТАТА-боксов промотора гена *LEP* rs200487063, rs34104384 и rs201381696 было получено совпадение прогнозов *in silico* и полученных экспериментально характеристик связывания ТСБ/ТАТА (Рис. 21). Подтвер-

ждение прогнозов для анализируемых аллелей было получено также через влияние полиморфизмов гена *LEP* на экспрессию репортерного гена *LUC*.



**Рисунок 21.** Ожирение, связанное с дефицитом лептина, и его коррекция экзогенным лептином (Farooqi, O'Rahilly, 2014).

На основании полученных нами компьютерно-экспериментальных результатов (Табл. 7), а именно корреляции прогнозов и экспериментальных значений *in vitro* и *ex vivo*, можно предположить, что у носителей SNP -34Т>G ТАТА-бокса промотора гена *LEP* и SNP -35 А>G ТАТА-бокса промотора гена *CYP2A6* вероятность развития заболеваний, в которых участвуют эти гены, выше, чем у носителей других исследованных нами полиморфизмов. Для более точного предсказания необходима клиническая характеристика пациента и эпидемиологические сведения.

Поэтому исследуемые SNPs предлагаются в качестве потенциальных кандидатов SNP-маркеров для отслеживания нарушений липидного обмена. Это является экспериментально-компьютерным прогнозом данной работы, который может быть верифицирован по биомедицинским стандартам и протоколам.

### **Выводы**

1. Проведенное экспериментальное исследование показало эффективность использования разработанного в ИЦиГ СО РАН Web сервиса SNP-TATA-Comparator для предсказания новых потенциально функционально значимых мутаций TATA-боксов промоторов генов человека.
2. Всестороннее экспериментальное изучение связывания TATA-связывающего белка с олигонуклеотидами, идентичными TATA-боксам референсных и минорных аллелей ряда генов липидного метаболизма показало, что минимальное увеличение сродства, на 14%, оказывает мутация -30A>T (rs34104384) TATA-бокса промотора гена *LEP*, максимальное, в 2 раза, мутация -38G>A (rs200487063) TATA-бокса этого гена. Максимальное уменьшение сродства в 7 раз происходит при мутации -24T>G (rs1800202) TATA-бокса гена *TPI*. Экспериментальные данные коррелируют с прогнозами, сделанными с помощью разработанного Web сервиса ( $r = 0.89$ ).
3. Изменение наследственной информации в результате мутаций в TATA-боксе реализуется через изменение скоростей образования и распада комплексов TATA-связывающего белка с TATA-боксами промоторов генов липидного метаболизма. Скорости образования изменяются от 0.2 до 14 раз. Мутация -35A>G (rs201381696) промотора гена *LEP*, нарушающая последовательность TATA-бокса, приводит к максимальному увеличению скорости распада комплексов - в 7 раз.

4. Образование комплексов ТАТА-связывающего белка с меченными FRET-парой флуорофоров TAMRA и FAM олигонуклеотидами, идентичными ТАТА-боксам промоторов референсных генов *LEP* и *TPI* и их минорным аллелям, в режиме реального времени, может происходить как по прямому пути, так и через промежуточные стадии, вероятность возникновения которых возрастает с увеличением длины олигонуклеотидов и концентрации ТАТА-связывающего белка.

5. На культуре клеток человека НСТ116 показана хорошая корреляция влияния ( $r=0.89$ ) мутаций ТАТА-боксов на сродство ТАТА-связывающего белка к ТАТА-боксам и на экспрессию репортерного гена *LUC* относительно «дикого типа» (WT): мутация -38G>A (rs200487063) ТАТА-боксов промотора гена *LEP*, повышающая сродство в 2 раза, в линии клеток НСТ116 повышала экспрессию репортерного гена *LUC* в 1.5 - 1.8 раза, мутация -35A>G (rs201381596), снижающая сродство в 3 раза, снижала экспрессию в среднем в 2.5 - 3 раза; мутация -34T>G (rs28399433) ТАТА-боксов промотора гена *CYP2A6*, снижающая сродство в 4 раза, также в 4 раза снижала экспрессию репортерного гена *LUC*.

### Благодарности

Хочу поблагодарить своего научного руководителя к.б.н. Савинкову Л.К. за настойчивость и терпение при руководстве выполнением данной работы. За помощь в работе к.б.н. М.П. Пономаренко, И.А. Драчкову, к.х.н. Н.А. Кузнецова, а также Е.Б. Шарыпову и Е.В. Кашину. Хочу поблагодарить также за помощь, как в работе, так и за психологическую поддержку к.б.н. Н.П.Бондарь, к.б.н. Л.О.Брызгалова и к.б.н. Н.И.Ершова.

### Список литературы

1. Васильева Л.А. Статистические методы в биологии, медицине и сельском хозяйстве. – Новосибирск: ИЦиГ СО РАН, НГУ, 2007. – С. 127.
2. Горизонтов П.Д. Гомеостаз. – М.: Медицина, 1976. – С. 464.
3. Драчкова И.А., Аршинова Т.В., Пономаренко П.М., Меркулова Т.И., Колчанов Н.А., Савинкова Л.К. Влияние полиморфизмов ТАТА-боксов промотора гена бета-глобина человека, ассоциированный с бета-талассемией, на взаимодействие ТАТА-связывающего белка //Информационный вестник ВОГИС. – 2010. – Вып. 14. – № 4. – С. 698-705.
4. Драчкова И.А., Аршинова Т.В., Пономаренко П.М., Меркулова Т.И., Савинкова Л.К., Колчанов Н.А. Изменение сродства ТАТА-связывающего белка к олигонуклеотидам, соответствующим ТАТА-боксам промоторов генов человека, несущим полиморфизмы, ассоциированные с наследственными заболеваниями //Вавиловский журнал генетики и селекции. – 2011. – Вып. 15. – № 3. – С. 439-447.
5. Драчкова И.А., Шеховцов С.В., Пельтек С.Е., Пономаренко П.М., Аршинова Т.В., Пономаренко М.П., Меркулова Т.И., Савинкова Л.К., Колчанов Н.А. Изучение взаимодействия ТВР человека с ТАТА-элементом промотора гена NOS2A с использованием метода поверхностного плазмонного резонанса //Вавиловский журнал генетики и селекции. – 2012. – Вып. 16. – № 2. – С. 391-396.
6. Кнорре Д.Г., Мызина С.Д. Биологическая химия. – М.: Высшая школа, 1992. – С. 105.
7. Мохорт Т.В., Забаровская З.В., Шепелькевич А.П. Клиническая эндокринология. – Минск: Выш. шк., 2013. – С. 415.

8. Петунина Н.А., Трухина Л.В., Синицына Е.И., Шестакова М.В. Глюкагон и  $\alpha$ -клетки – новая терапевтическая мишень в лечении сахарного диабета //Сахарный диабет. – 2013. – № 3. – С. 35–40.
9. Пономаренко П.М., Пономаренко М.П., Драчкова И.А., Лысова М.В., Аршинова Т.В., Савинкова Л.К., Колчанов Н.А. Прогноз изменения аффинности ТАТА-связывающего белка к ТАТА-боксам в результате полиморфизмов ТАТА-боксов промоторов генов человека //Молекуляр. биология. – 2009. – No. 43. – P. 512-520.
10. Пономаренко П.М., Савинкова Л.К., Драчкова И.А., Лысова М.В., Аршинова Т.В., Пономаренко М.П., Колчанов Н.А. Пошаговая модель связывания ТВР/ТАТА бокс позволяет предсказать наследственные заболевания человека по точечным полиморфизмам //Doklady Biochemistry and Biophysics. – 2008. – Vol. 419. – P. 88-92.
11. Рассказов Д.А., Гунбин К.В., Пономаренко П.М., Вишневский О.В., Пономаренко М.П., Афонников Д.А. SNP\_TATA\_COMPARATOR: Web-сервис применения уравнения равновесия ТВР/ТАТА-комплекса в сравнительной оценке SNPs промоторов генов, связанных с болезнями человека //Вавиловский журнал генетики и селекции. – 2013. – Vol. 17. – No. 4/1. – P. 599-606.
12. Савинкова Л.К., Драчкова И.А., Пономаренко М.П., Лысова М.В., Аршинова Т.В., Колчанов Н.А. Взаимодействие рекомбинантного ТАТА-связывающего белка с ТАТА-боксами промоторов генов млекопитающих //Экологическая генетика. – 2007. – Vol. 5. – No. 2. – P. 44-49.
13. Савинкова Л.К., Пономаренко М.П., Пономаренко П.М., Драчкова И.А., Лысова М.В., Аршинова Т.В., Колчанов Н.А. Полиморфизмы ТАТА-боксов промоторов

- генов человека и ассоциированные с ними наследственные патологии //BIOCHEMISTRY-MOSCOW+. – 2009. – Vol. 74. – No. 2. – P. 117-129.
14. Скоупс Р. Методы очистки белков. – М.: Мир, 1985. – С. 54.
15. Торшин И.Ю., Громова А.О., Никонов А.А. Гены и цереброваскулярная патология (ассоциативные исследования) //Ж. Неврологии и психиатрии. – 2009. – Т. 109. – № 5. – С. 77-85.
16. Allison L.A. Fundamental Molecular Biology. – Wiley-Blackwell.: 1st edn., 2007. – P.101-130.
17. Altshuler D.M., Gibbs R.A., Peltonen L. et al. International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations //Nature. – 2010. – Vol. 467. – No. 7311. – P. 52-58.
18. Antonarakis S.E., Irkin S.H., Cheng T.C. et al. Beta-thalassemia in American Blacks: novel mutations in the “TATA” box and an acceptor splice site //Proc.Natl. Acad. Sci. USA. – 1984. – V. 81. – No. 4. – P. 1154–1158.
19. Arkova O.V., Ponomarenko M.P., Rasskazov D.A., Drachkova I.A. et al. Obesity-related known and candidate SNP markers can significantly change affinity of TATA-binding protein for human gene promoters //BMC Genomics. – 2015. – Suppl. 13.
20. Arkova O., Kuznetsov N., Fedorova O., Savinkova L. A real-time study of the interaction of TBP with a TATA box-containing duplex identical to an ancestral or minor allele of human gene LEP or TPI //J Biomol Struct Dyn. – 2016. – No. 25. – P.1-12.
21. Arkova O.V., Kuznetsov N.A., Fedorova O.S., Kolchanov N.A., Savinkova L.K. Real-Time Interaction between TBP and the TATA Box of the Human Triosephosphate Isomerase Gene Promoter in the Norm and Pathology //Acta Naturae. – 2014. – Vol. 6. – No. 2. – P. 36-40.



22. Arnaud E., Barbalat V., Nicaud V., Cambien F. et al. Polymorphisms in the 5' regulatory region of the tissue factor gene and the risk of myocardial infarction and venous thromboembolism: the ECTIM and PATHROS studies. *Etude Cas-Témoins de l'Infarctus du Myocarde. Paris Thrombosis case-control Study //Arterioscler Thromb Vasc Biol.* – 2000. – Vol. 20. – No.3. – P. 892-8.
23. Arroyo A., Laughlin G.A., Morales A.J., Yen S.S. Inappropriate gonadotropin secretion in polycystic ovary syndrome: influence of adiposity //*J Clin Endocrinol Metab.* – 1997. – Vol. 82. – No. 11. – P. 3728–33.
24. Badens C., Jassim N., Martini N. et al. Characterization of a new polymorphism, IVS-I–108 (TC), and a new beta-thalassemia mutation, –27 (AT), discovered in the course of a prenatal diagnosis //*Hemoglobin.* – 1999. – Vol. 23. – No. 4. – P. 339–344.
25. Baecklund F., Foo J.N., Bracci P., Darabi H. et al. A comprehensive evaluation of the role of genetic variation in follicular lymphoma survival //*BMC Med. Genet.* – 2014. – No. 15. – P.113.
26. Baicy K., London E.D., Monterosso J., Wong M.L., Delibasi T., Sharma A., Licinio J. Leptin replacement alters brain response to food cues in genetically leptin-deficient adults //*Proc Natl Acad Sci U S A.* – 2007. – Vol. 104. – No. 46. – P.18276-9.
27. Barber M.D., Powell J.J., Lynch S.F., Fearon K.C., Ross J.A. A polymorphism of the interleukin-1 beta gene influences survival in pancreatic cancer //*Br J Cancer.* – 2000. – Vol. 83. – No. 11. – P. 1443-7.
28. Barber T.M., McCarthy M.I., Wass J.A., Franks S. Obesity and polycystic ovary syndrome //*Clin Endocrinol (Oxf).* – 2006. – Vol. 65. – No. 2. – P. 137–45.

29. Bartsch H., Dally H., Popanda O., Risch A., Schmezer P. Genetic risk profiles for cancer susceptibility and therapy response //Recent Results Cancer Res. – 2007. – Vol. 174. – P. 19-36.
30. Baumann M., Pontiller J., Ernst W. Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: an overview //Mol Biotechnol. –2010. – Vol. 45. – No. 3. – P. 241-7.
31. Beerl M.S., Silverman J.M., Davis K.L., Marin D., Grossman H.Z., Schmeidler J., Purohit D.P., Perl D.P., Davidson M., Mohs R.C., Haroutunian V. Type 2 diabetes is negatively associated with Alzheimer's disease neuropathology //J Gerontol A Biol Sci Med Sci. – 2005. – Vol. 60. – No. 4. – P. 471-5.
32. Beltowski J. Central vs. peripheral leptin excess in the pathogenesis of obesity-associated hypertension //J Hypertens. – 2008. – Vol. 26. – No. 4. – P. 827–828.
33. Birnboim H.C., Doly J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA //Nucleic Acids Res. – 1979. – Vol. 7. – No. 6. – P. 1513-23.
34. Bjornsdottir G., Myers L.C. Minimal components of the RNA polymerase II transcription apparatus determine the consensus TATA box //Nucl. Acids Res. – 2008. – Vol. 36. – No. 9. – P. 2906-2916.
35. Blair R.H., Goodrich J.A., Kugel J.F. Single-molecule fluorescence resonance energy transfer shows uniformity in TATA binding protein-induced DNA bending and heterogeneity in binding kinetics //Biochemistry. – 2012. – Vol. 51. – No. 38. – P. 7444-55.
36. Boldt A.B., Culpi L., Tsuneto L.T., de Souza I.R., Kun J.F., Petzl-Erle M.L. Diversity of the MBL2 Gene in Various Brazilian Populations and the Case of Selection at the

- Mannose-Binding Lectin Locus //Human Immunology. – 2006. – Vol. 67. – No. 9. – P. 722-734.
37. Booth T., Royle N.A., Corley J., Gow A.J. et al. Association of allostatic load with brain structure and cognitive ability in later life //Neurobiol Aging. – 2015. – Vol. 36. – No. 3. – P. 1390-9.
38. Brakenhoff R.H., Henskens H.A., van Rossum M.W., Lubsen N.H., Schoenmakers J.G. Activation of the gamma E-crystallin pseudogene in the human hereditary Cop-pock-like cataract //Hum Mol Genet. – 1994. – Vol. 3. – No. 2. – P. 279-83.
39. Brown J.R., Daar I.O., Krug J.R., Maquat L.E. Characterization of the functional gene and several processed pseudogenes in the human triosephosphate isomerase gene family //Mol. Cell. Biol. – 1985. – Vol. 5. – No. 7. – P. 1694-1706.
40. Bruehl H., Wolf O.T., Sweat V., Tirsi A., Richardson S., Convit A. Modifiers of cognitive function and brain structure in middle-aged and elderly individuals with type 2 diabetes mellitus //Brain Res. – 2009. – No. 1280. – P. 186-94.
41. Buratowski S., Hahn S., Guarente L., Sharp P.A. Five intermediate complexes in transcription initiation by RNA polymerase II //Cell. – 1989. – Vol. 56. – No. 4. – P. 549-61.
42. Burgner D., Rockett K., Ackerman H., Hull J. et al. Haplotypic relationship between SNP and microsatellite markers at the NOS2A locus in two populations //Genes Immun. – 2003. – Vol. 4. – No. 7. – P. 506-14.
43. Butler J.E., Kadonaga J.T. The RNA polymerase II core promoter: a key component in the regulation of gene expression //Genes Dev. – 2002. – Vol. 16. – No. 20. – P. 2583-2592.

44. Cai S.P., Zhang J.Z., Doherty M., Kan Y.W. A new TATA box mutation detected at prenatal diagnosis for beta-thalassemia //Am. J. Hum. Genet. – 1989. – Vol. 45. – No. 1. – P.112-114.
45. Carbone A., Bigoloni A., Galli L., Spagnuolo V., Gianotti N., Lazzarin A., Castagna A. Glucose tolerance in HIV-1 treated patients who switched from boosted-protease inhibitors to etravirine //AIDS. – 2013. – Vol. 27. – No. 16. – P. 2661-3.
46. Carninci P., Sandelin A., et al. Genome-wide analysis of mammalian promoter architecture and evolution //Nat Genet. – 2006. – Vol. 38. – No. 6. – P. 626-35.
47. Caron C., Rousset R., Béraud C., Moncollin V., Egly J. M., and Jalinot P. Functional and biochemical interaction of the HTLV-I Tax1 transactivator with TBP //EMBO J. – 1993. – Vol. 12. – No. 11. – P. 4269–4278.
48. Caron C., Goudemand J., Marey A., Beague D. et al. Are haemostatic and fibrinolytic parameters predictors of preeclampsia in pregnancy-associated hypertension? //Thromb Haemost. – 1991. – Vol. 66. – No. 4. – P. 410-4.
49. Chakravarti A. Population genetics--making sense out of sequence //Nat Genet. – 1999. – Vol. 21. – 1 Suppl. – P. 56-60.
50. Chan J.M., Rimm E.B., Colditz G.A., Stampfer M.J., Willett W.C. Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in men //Diabetes Care. – 1994. – Vol.17. – No. 9. – P. 961–9.
51. Chang M.L., Artymiuk P.J., Wu X., Hollán S., Lammi A., Maquat L.E. Human triosephosphate isomerase deficiency resulting from mutation of Phe-240 //Am J Hum Genet. – 1993. – Vol. 52. – No. 6. – P. 1260-9.
52. Colonna V., Ayub Q., Chen Y., Pagani L., Luisi P., Pybus M., McVean G.A. 1000 Genomes Project Consortium //Genome Biology. – 2014. – No. 15.

53. Comings D.E., Gade R., Muhleman D., Chiu C., Wu S. et al. Exon and intron variants in the human tryptophan 2,3-dioxygenase gene: potential association with Tourette syndrome, substance abuse and other disorders //Pharmacogenetics. – 1996. – Vol. 6. – No. 4. – P. 307-18.
54. Davis N.A., Majee S.S., Kahn J.D. TATA box DNA deformation with and without the TATA box-binding protein //Journal of Molecular Biology. – 1999. – Vol. 291. – P. 249-265.
55. Dean M., Allikmets R. Complete characterization of the human ABC gene family //J. Bioenerg. Biomembr. – 2001. – Vol. 33. – No. 6. – P. 475-479.
56. Delgadillo, R.F., Whittington, J.E., Parkhurst, L.K., Parkhurst, L.J. The TATA-binding protein core domain in solution variably bends TATA sequences via a three-step binding mechanism //Biochemistry. – 2009. – Vol. 48. – P. 1801–1809.
57. Derijk R.H., Schaaf M.J., Turner G., Datson N.A. et al. A human glucocorticoid receptor gene variant that increases the stability of the glucocorticoid receptor beta-isoform mRNA is associated with rheumatoid arthritis //J Rheumatol. – 2001. – Vol. 28. – No. 11. – P.2383-8.
58. Drachkova I., Savinkova L., Arshinova T., Ponomarenko M., Peltek S., Kolchanov N. The mechanism by which TATA-box polymorphisms associated with human hereditary diseases influence interactions with the TATA-binding protein //Human Mutation. – 2014. – No. 35. – P. 601–608.
59. Dravecka I., Lazurova I., Kraus V. Obesity is the major factor determining an insulin sensitivity and androgen production in women with anovulatory cycles //Bratisl Lek Listy. – 2003. – Vol. 104. – No. 12. – P. 393–9.

60. Dreos R., Ambrosini G., Périer R.C., Bucher P. The Eukaryotic Promoter Database: expansion of EPD new and new promoter analysis tools //Nucleic Acids Res. – 2015. – Vol. 43(Database issue). – D 92-6.
61. Faiger H., Ivanchenko M., Cohen I., Haran T.E. TBP flanking sequences: asymmetry of binding, long-range effects and consensus sequences //Nucleic Acids Res. – 2006. – Vol. 34. – No. 1. – P. 104-19.
62. Farooqi I.S., O'Rahilly S. 20 years of leptin: human disorders of leptin action //J Endocrinol. – 2014. – Vol. 223. – No. 1. – P. 63-70.
63. Fei Y.J., Stoming T.A., Efremov G.D. et al. Beta-thalassemia due to a T–A mutation within the ATA box //Biochem. Biophys. Res. Commun. – 1988. – Vol.153. – No. 2. – P. 741–747.
64. Frederickson R., Hagglund M., Olszewski P.K., Stephansson O., Jacobsson J.A., Olszewska A.M. et al. The obesity gene, FTO, is of ancient origin, up-regulated during food deprivation and expressed in neurons of feeding-related nuclei of the brain //Endocrinology. – 2008. – Vol. 149. – No. 5. – P. 2062–71.
65. Friedman J. M. Leptin and the regulation of body weight //Keio J Med. – 2011. – Vol. 6. – No. 1. – P. 1–9.
66. Galanello R., Origa R. Beta thalassemia //Orphanet J. of Rare Diseases. – 2010. – 5:11.
67. Gannon F., O'Hare K., Perrin F., LePennec J.P. et al. Organisation and sequences at the 5' end of a cloned complete ovalbumin gene //Nature. – 1979. – Vol. 278. – 428-434.
68. Garred P., Larsen F., Seyfarth J., Fujita R., Madsen H.O. Mannose-binding lectin and its genetic variants //Genes and Immunity. – 2006. – Vol. 7. – P. 85–94.

69. Gietl A., Holzmeister P., Blombach F., Schulz S., von Voithenberg L.V., Lamb D.C., Werner F., Tinnefeld P., Grohmann D. Eukaryotic and archaeal TBP and TFB/TF(II)B follow different promoter DNA bending pathways //Nucleic Acids Research. – 2014. – Vol. 42. – P. 6219-6231.
70. Goldberg M.L. Sequence Analysis of Drosophila Histone Genes. Ph.D. – Thesis, Stanford University, CA: 1979. – 159 pp.
71. Gordon D.J., Rifkind B.M. High-density lipoprotein--the clinical implications of recent studies //N Engl J Med. – 1989. – Vol. 321. – No.19. – P. 1311-6.
72. Guo Y.I., Qian Y., Gong Y.I., Pan C., Shi G., Wan H.. A predictive model for the development of chronic obstructive pulmonary disease //Biomed Rep. – 2015. – Vol. 3. – No. 6. – P. 853-863.
73. Hahn S. Structure and mechanism of the RNA polymerase II transcription machinery //Nat Struct Mol Biol. – 2004. – Vol. 11. – No. 5. – P. 394-403.
74. Hahn, S., Buratowski, S., Sharp, P., Guarente, L. Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences //Proceedings of the National Academy of Sciences of the United States of America. – 1989. – Vol. 86. – P. 5718-5722.
75. Hall J.E., da Silva A.A., do Carmo J.M., Dubinion J., Hamza S., Munusamy S. et al. Obesity-induced hypertension: role of sympathetic nervous system, leptin and melanocortins //J Biol Chem. – 2010. – Vol. 285. – No. 23. – P. 17271–6.
76. Halushka M.K., Fan J.B., Bentley K., Hsie L., Shen N., Weder A., Cooper R., Lipshutz R., Chakravarti A. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis //Nat Genet. – 1999. – Vol. 22. – No. 3. – P. 239-47.

77. Harrow J., Frankish A., Gonzalez J.M., Tapanari E. et al. GENCODE: the reference human genome annotation for The ENCODE Project //Genome Res. – 2012. – Vol. 22. – No. 9. – P. 1760-1774.
78. Hayashi Y., Yamamoto M., Mizoguchi H., Watanabe C., Ito R. et al. Mice deficient for glucagon gene - derived peptides display normoglycemia and hyperplasia of islet {alpha}-cells but not of intestinal L-cells //Mol Endocrinol. – 2009. – Vol. 23. – No. 12. – P. 1990-9.
79. He Y., Chen H., Quon M.J., Reitman M. The mouse obese gene. Genomic organization, promoter activity, and activation by CCAAT/enhancer-binding protein alpha //J Biol Chem. – 1995. – Vol. 270. – No. 48. – P. 28887–28891.
80. Hedditch E.L., Gao B., Russell A.J., Lu Y. et al. ABCA transporter gene expression and poor outcome in epithelial ovarian cancer //J. Natl. Cancer Inst. – 2014. – Vol. 106. – No. 7. – P. 149.
81. Hendig D., Langmann T., Kocken S., Zarbock R. et al. Gene expression profiling of ABC transporters in dermal fibroblasts of pseudoxanthoma elasticum patients identifies new candidates involved in PXE pathogenesis //Lab. Invest. – 2008. – Vol. 88. – No. 12. – P. 1303-1315.
82. Hernandez N. TBP, a universal eukaryotic transcription factor? //Genes & Development. – 1993. – Vol. 7(7B). – P. 1291-308.
83. Hernandez N. TBP, a universal eukaryotic transcription factor? //Genes Dev. – 1993. – Vol. 7. – No. 7B. – P. 1291-308.
84. Hieb A.R., Gansen A., Böhm V., Langowski J. The conformational state of the nucleosome entry-exit site modulates TATA box-specific TBP binding //Nucleic Acids Research. – 2014. – Vol. 42. – No. 12. – P. 7561-76.



85. Hieb A.R., Halsey W.A., Betterton M.D., Perkins T.T., Kugel J.F., Goodrich J.A. TFIIA changes the conformation of the DNA in TBP/TATA complexes and increases their kinetic stability // *Journal of Molecular Biology*. – 2007. – Vol. 372. – P. 619-32.
86. Higashi E., Fukami T., Itoh M. et al. Human CYP2A6 Is Induced by Estrogen via Estrogen Receptor // *Drug Metab Dispos*. – 2007. – Vol. 35. – No. 10. – P. 1935-41.
87. Hoffman A., Sinn E., Yamamoto T., Wang J. et al. Highly conserved core domain and unique N terminus with presumptive regulatory motifs in a human TATA factor (TFIID) // *Nature*. – 1990. – Vol. 346. – P. 387-390.
88. Holvoet P. Relations between metabolic syndrome, oxidative stress and inflammation and cardiovascular disease // *Verh K Acad Geneesk Belg*. – 2008. – Vol. 70. – No. 3. – P. 193-219.
89. Hoopes B.C., LeBlanc J.F., Hawley D.K. Kinetic analysis of yeast TFIID-TATA box complex formation suggests a multi-step pathway // *Journal of Biological Chemistry*. – 1992. – Vol. 267. – P. 11539-11547.
90. Horikoshi M., Yamamoto T., Ohkuma Y., Weil P.A., Roeder R.G. Analysis of structure-function relationships of yeast TATA box binding factor TFIID // *Cell*. – 1990. – Vol. 61. – P. 1171-1178.
91. Horn A.E., Kugel J.F., Goodrich J.A. Single molecule microscopy reveals mechanistic insight into RNA polymerase II preinitiation complex assembly and transcriptional activity // *Nucleic Acids Res*. – 2016. – Vol. 44. – No. 15. – P. 7132-7143.
92. Hsu J.Y., Juven-Gershon T., Marr M.T. 2nd, Wright K.J., Tjian R., Kadonaga J.T. TBP, Mot1, and NC2 establish a regulatory circuit that controls DPE-dependent versus TATA-dependent transcription // *Genes Dev*. – 2008. – Vol. 22. – No. 17. – P. 2353-8.

93. Hubáček J.A., Pitha J., Adámková V., Skodová Z., Lánská V., Poledne R. Apolipoprotein E and apolipoprotein CI polymorphisms in the Czech population: almost complete linkage disequilibrium of the less frequent alleles of both polymorphisms //Physiol Res. – 2003. – Vol. 52. – No. 2. – P. 195-200.
94. Huggins C.J., Gill M., Andrulis I.L. Identification of rare variants in the hLIMD1 gene in breast cancer //Cancer Genet Cytogenet. – 2007. – Vol. 178. – No. 1. – P. 36-41.
95. Humphries A., Ationu A., Wild B., Layton D.M. The consequence of nucleotide substitutions in the triosephosphate isomerase (TPI) gene promoter //Blood Cells Mol Dis. – 1999. – Vol. 25. – No. 3-4. – P. 210-217.
96. Humphries A., Ationu A., Lalloz M.R., Layton D.M. Ancestral origin of variation in the triosephosphate isomerase gene promoter //Hum Genet. – 1999. – Vol. 104. – No. 6. – P. 486-91.
97. Inoue H., Nojima H., Okayama H. High efficiency transformation of Escherichia coli with plasmids //Gene. – 1990. – Vol. 96. – No. 1. – P. 23-8.
98. Jack D.L., Klein N.J., Turner M.W. Mannose-binding lectin: targeting the microbial world for complement attack and opsonophagocytosis //Immunological Reviews. – 2001. – Vol. 180. – P. 86-99.
99. Janjgava S., Zerekidze T., Uchava L., Giorgadze E., Asatiani K. Influence of testosterone replacement therapy on metabolic disorders in male patients with type 2 diabetes mellitus and androgen deficiency //Eur J Med Res. – 2014. – No. 19. – P. 56.
100. Juven-Gershon T., Hsu J.Y., Kadonaga J.T. Perspectives on the RNA polymerase II core promoter //Biochem Soc Trans. – 2006. – Vol. 34. – Pt. 6. – P. 1047-1050.

101. Juven-Gershon T., Hsu J.Y., Theisen J.W., Kadonaga J.T. The RNA polymerase II core promoter -the gateway to transcription //Current opinion in cell biology. – 2008. – Vol. 20. – No. 3. – P. 253-259.
102. Juven-Gershon T., Kadonaga, J.T. Regulation of gene expression via the core promoter and the basal transcriptional machinery //Developmental biology. – 2010. – Vol. 339. – No. 2. – P. 225-229.
103. Kabashi E., Valdmanis P.N., Dion B.P., Rouleau G.A. Oxidized/Misfolded superoxide dismutase-1: the cause of all amyotrophic lateral sclerosis //Ann Neurol. – 2007. – Vol. 62. – P. 553-559.
104. Kadonaga J.T. Perspectives on the RNA polymerase II core promoter. Wiley interdisciplinary reviews //Developmental biology. – 2012. – Vol. 1. – No. 1. – P. 40-51.
105. Kaniwa N., Kurose K., Jinno H. et al. Racial variability in haplotype frequencies of UGT1A1 and glucuronidation activity of a novel single nucleotide polymorphism 686C> T (P229L) found in an African-American //Drug Metabolism and Disposition. – 2005. – Vol. 33. – No. 3. – P. 458-465.
106. Kayatekin C., Zitzewitz J.A., Matthews C.R. Zinc binding modulates the entire folding free energy surface of human Cu,Zn superoxide dismutase //J Mol Biol. – 2008. – Vol. 384. – No. 2. – P. 540-55.
107. Kim J. L., Nikolov D. B., Burley S. K. Co-crystal structure of TBP recognizing the minor groove of a TATA element //Nature. – 1993. – Vol. 365. – P. 520-527.
108. Kim T., Barrera L., Zheng M., Qu C., Singer M.A., Richmond T.A., Wu Y., Green R.D., Ren B. A high-resolution map of active promoters in the human genome //Nature. – 2005. – Vol. 436. – P. 876-880.

109. Kim T.H., Barrera L.O., Zheng M., Qu C. et al. A high-resolution map of active promoters in the human genome //Nature. – 2005. – Vol. 436. – No. 7052. – P. 876-880.
110. Kim Y., Geiger J. H., Hahn S., Sigler P. B. Crystal structure of a yeast TBP/TATA-box complex //Nature. – 1993. – Vol. 365. – P. 512-520.
111. Kim Y.J., Bjorklund S., Li Y., Sayre M.H., Kornberg R.D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminalrepeat domain of RNA polymerase II //Cell. – 1994. – Vol. 77. – No. 4. – P. 599-608.
112. Kokosov AN. Chronic bronchitis and obstructive lung disease: analytical assay //Ter Arkh. – 2000. – Vol. 72. – No. 3. – P. 75-7.
113. Kuźma-Kozakiewicz M., Kwieciński H. The genetics of amyotrophic lateral sclerosis //Neurol Neurochir Pol. – 2009. – Vol. 43. – No. 6. – P. 538-549.
114. Kuzmic P. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase //Anal Biochem. – 1996. – Vol. 237. – No. 2. – P. 260-73.
115. Laemmli U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4 //Nature. – 1970. – Vol. 227. – No. 5259. – P. 680-5.
116. Lagrange T., Kapanidis A.N., Tang H., Reinberg D., Ebright R.H. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB //Genes Dev. – 1998. – Vol. 12. – No. 1. – P. 34-44.
117. Lander E.S. The new genomics: global views of biology //Science. – 1996. – Vol. 274. – No. 5287. – P. 536-9.

118. Laughlin G. A., Dominguez C. E., Yen S. S. Nutritional and endocrine-metabolic aberrations in women with functional hypothalamic amenorrhea //J Clin Endocrinol Metab. – 1998. – Vol. 83. – No. 1. – P. 25–32.
119. Lavie C.J., Milani R.V., Ventura H.O. Obesity and cardiovascular disease: risk factor, paradox, and impact of weight loss //J Am Coll Cardiol. – 2009. – Vol. 53. – No. 21. – P. 1925-32.
120. Lee E.B., Mattson M.P. The neuropathology of obesity: insights from human disease //Acta Neuropathol. – 2014. – Vol. 127. – No. 1. – P. 3–28.
121. Lee Y.H., Bang H., Kim D.J. How to Establish Clinical Prediction Models //Endocrinol Metab (Seoul). – 2016. – Vol. 31. – No. 1. – P. 38-44.
122. Levings P.P., Bungert J. The human beta-globin locus control region //Eur J Biochem. – 2002. – Vol. 269. – P. 1589-99.
123. Li S.S., Kendall C.W., de Souza R.J., Jayalath V.H. et al. Dietary pulses, satiety and food intake: a systematic review and meta-analysis of acute feeding trials //Obesity (Silver Spring). – 2014. – Vol. 22. – No. 8. – P. 1773-80.
124. Librizzi M.D., Brenowitz M., Willis I.M. The TATA element and its context affect the cooperative interaction of TATA-binding protein with the TFIIB-related factor, TFIIB70 //J Biol Chem. – 1998. – Vol. 273. – No. 8. – P. 4563-8.
125. Lifton R.P., Goldberg M.L., Karp R.W., Hogness D.S. The organization of the histone genes in *Drosophila melanogaster*: functional and evolutionary implications //Cold Spring Harb Symp Quant Biol. – 1978. – Vol. 42. – Pt. 2. – P. 1047-51.
126. Liu Z., Wong J., Tsai S.Y., Tsai M.J., O'Malley B.W. Sequential recruitment of steroid receptor coactivator-1 (SRC-1) and p300 enhances progesterone receptor-

- dependent initiation and reinitiation of transcription from chromatin //Proc Natl Acad Sci U S A. – 2001. – Vol. 98. – No. 22. – P. 12426-31.
127. Liu C., Xuan Z. Prioritization of Cancer-Related Genomic Variants by SNP Association Network //Cancer Inform. – 2015. – No. 14. – P. 57–70.
128. Machado J.C., Pharoah P., Sousa S., Carvalho R. et al. Interleukin 1B and interleukin 1RN polymorphisms are associated with increased risk of gastric carcinoma //Gastroenterology. – 2001. – Vol. 121. – No. 4. – P. 823-9.
129. Madsen H.O, Satz M.L, Hogh B., Svejgaard A., Garred P. Different molecular events result in low protein levels of manna-binding lectin in populations from Southeast Africa and South America //J Immunol. – 1998. – Vol. 161. – No. 6. – P. 3169-75.
130. Maldonado E., Shiekhattar R., Sheldon M., Cho H. et al. A human RNA polymerase II complex associated with SRB and DNA-repair proteins //Nature. – 1996. – Vol. 381. – No. 6577. – P. 86-9.
131. Malik S., Roeder R.G. Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells //Trends Biochem Sci. – 2000. – Vol. 25. – No. 6. – P. 277-83.
132. Mallal S., Nolan D., Witt C., Masel G., Martin A.M., Moore C., Christiansen F.T. Association between presence of HLA-B 5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir //Lancet. – 2002. – No. 359. – P. 727-732.
133. Manco L., Machado P., Lopes D., Nogueira F. et al. Analysis of TPI gene promoter variation in three Sub-Saharan Africa population samples //A. J. Human Biol. – 2009. – Vol. 21. – P. 118-120.

134. Markewitz B.A., Owens M.W., Payne D.K. The pathogenesis of chronic obstructive pulmonary disease //Am J Med Sci. – 1999. – Vol. 318. – No. 2. – P. 74-8.
135. Masters K.M., Parkhurst K.M., Daugherty M.A., Parkhurst L.J. Native human TATA-binding protein simultaneously binds and bends promoter DNA without a slow isomerization step or TFIIB requirement //J Biol Chem. – 2003. – Vol. 278. – No. 34. – P. 31685-90.
136. Masuzaki H., Ogawa Y., Sagawa N. et al. Non-adipose tissue production of leptin: leptin as a novel placenta-derived hormone in humans //Nature Med. – 1997. – Vol. 3. – No. 9. – P. 11029-33.
137. McMullin M.F. The molecular basis of disorders of red cell enzymes //J Clin. Pathol. – 1999. – Vol. 52. – P. 241-244.
138. Meisterernst M., Roeder R.G. Family of proteins that interact with TFIID and regulate promoter activity //Cell. – 1991. – Vol. 67. – P. 557-567.
139. Minematsu N., Nakamura H., Iwata M., Tateno H. et al. Association of CYP2A6 deletion polymorphism with smoking habit and development of pulmonary emphysema //Thorax. – 2003. – Vol. 58. – No. 7. – P. 623-8.
140. Mironova V.V., Omelyanchuk N.A., Ponomarenko P.M., Ponomarenko M.P., Kolchanov N.A. Specific/nonspecific binding of TBP to promoter DNA of the auxin response factor genes in plants correlated with ARFs function on gene transcription (activator/repressor) //Doklady Biochemistry and Biophysics. – 2010. – No. 433. – P. 191-196.
141. Mogno I., Vallania F., Mitra R.D., Cohen B. TATA is a modular component of synthetic promoters //Genome Res. – 2010. – Vol. 20. – No. 10. – P. 1391-1397.

142. Molina C., Grotewold E. Genome wide analysis of Arabidopsis core promoters //BMC Genomics. – 2005. – Vol. 6. – P. 25
143. Mondal M., Choudhury D., Chakrabarti J., Bhattacharyya D. Role of indirect readout mechanism in TATA box binding protein-DNA interaction //Journal of Computer-Aided Molecular Design. – 2015. – Vol. 29. – P. 283-95.
144. Mondal M., Mukherjee S., Bhattacharyya D. Contribution of phenylalanine side chain intercalation to the TATA-box binding protein-DNA interaction: molecular dynamics and dispersion-corrected density functional theory studies //Journal of Molecular Modeling. – 2014. – Vol. 20. – P. 2499.
145. Muncie H.L., JR., Campbell J.C. Alpha and beta thalassemia //A. Fam. Pfyician. – 2009. – Vol. 80. – No. 4. – P.339-344.
146. Niemann S., Broom W.J., Brown R.H. Analysis of a genetic defect in the TATA box of the SOD1 gene in a patient with familial amyotrophic lateral sclerosis //Muscle Nerve. – 2007. – Vol. 36. – No. 5. – P. 704-7.
147. Nikolov D.B., Chen H., Halay E.D., Hoffman A., Roeder R.G., Burley S.K. Crystal structure of a human TATA box-binding protein/TATA element complex //Proceedings of the National Academy of Sciences of the United States of America. – 1996. – Vol. 93. – P. 4862-4867.
148. Nteeba J., Ganesan S., Keating A.F. Impact of obesity on ovotoxicity induced by 7,12-dimethylbenz[a]anthracene in mice //Biol Reprod. – 2014. – Vol. 90. – No. 3. – P. 68.
149. Ohler U., Wassarman D.A. Promoting developmental transcription //Development. – 2010. – Vol. 137. – No. 1. – P. 15-26.



150. Oram J., Vaughan A. ATP-Binding cassette cholesterol transporters and cardiovascular disease //Circ. Res. – 2006. – Vol. 99. – No. 10. – P. 1031-1043.
151. Orkin S.H., Sexton J.P., Cheng T.C. et al. TATA box transcription mutation in beta-thalassemia //Nucleic Acids Res. – 1983. – V. 11. – No. 14. – P. 4727–4734.
152. Orosz F., Olah J., Ovadi J. Reappraisal of triosephosphate isomerase deficiency //Eur J Haematol. – 2010. – Vol. 86. – P. 265-267.
153. Orosz F., Olah J., Ovadi J. Triosephosphate isomerase deficiency fact and doubts //IUBMB Life. – 2006. – Vol. 58. – No. 12. – P. 703-715.
154. Parkhurst K.M., Brenowitz M., Parkhurst L.J. Simultaneous binding and bending of promoter DNA by the TATA binding protein: real time kinetic measurements //Biochemistry. – 1996. – Vol. 35. – No. 23. – P. 7459-7465.
155. Parkhurst, K.M., Richards, R.M., Brenowitz, M., Parkhurst, L.J. Intermediate species possessing bent DNA are present along the pathway to formation of a final TBP-TATA complex //Journal of Molecular Biology. – 1999. – Vol. 289. – No. 5. – P. 1327–1341.
156. Patikoglou G.A., Kim J.L., Sun L., Yang S.H. et al. TATA element recognition by the TATA box-binding protein has been conserved throughout evolution //Genes & Development. – 1999. – Vol. 13. – P. 3217-30.
157. Pelkonen O., Rautio A., Raunio H., Pasanen, M. CYP2A6: a human coumarin 7-hydroxylase //Toxicol. – 2000. – V. 144. – P. 139-147.
158. Piehler A., Kaminski W.E., Wenzel J., Langmann T., Schmitz G. Molecular structure of a novel cholesterol-responsive A subclass ABC transporter, ABCA9 //Biochem. Biophys. Res. Commun. – 2002. – Vol. 295. – No. 2. – P. 408-416.

159. Pipeleers D.G., Schuit F.C., in't Veld P.A., Maes E. et al. Interplay of nutrients and hormones in the regulation of insulin release //Endocrinology. – 1985. – Vol. 117. – No. 3. – P. 824-33.
160. Pitarque M., Rodríguez-Antona C., Sundberg I. Transcriptional Regulation of the Human CYP2A6 Gene //J Pharmacol Exp Ther. – 2005. – Vol. 313. – No. 2. – P. 814-22.
161. Pitarque M., Richter O., Oke B., Berkkan H., Oscarson M., Sundberg I. Identification of a single nucleotide polymorphism in the TATA box of the CYP2A6 Gene: Impairment of its promoter activity //Biochem Biophys Res Com. – 2001. – Vol. 284. – P. 455-460.
162. Poncz M., Ballantine M., Solowiejczyk D. et al. Beta-thalassemia in a Kurdish Jew. Single base changes in the T-A-T-A box //J. Biol. Chem. – 1982. – Vol. 257. – No. 11. – P. 5994–5996.
163. Ponjavic J., Lenhard B., Kai C., Kawai J. et al. Transcriptional and structural impact of TATA-initiation site spacing in mammalian core promoters //Genome Biol. – 2006. – Vol. 7. – No. 8. – R. 78.
164. Ponomarenko M., Mironova V., Gunbin K., Savinkova L. Hogness Box. – Eds S. Maloy, & K. Hughes. Brenner's Encyclopedia of Genetics (2nd ed.). San Diego: Acad. Press, Elsevier Inc. – 2013a. – No. 3. – P. 491-494.
165. Ponomarenko M., Savinkova L., Kolchanov N. Initiation Factors. – S. Maloy, & K. Hughes (Eds.) Brenner's Encyclopedia of Genetics (2nd ed.). San Diego: Acad. Press, Elsevier Inc. – 2013b. – No. 4. – P. 83-85.
166. Ponomarenko M.P., Ponomarenko Y.V., Frolov A.S., Podkolodny N.L., Savinkova L.K., Kolchanov N.A., Overton G.C. Candidate SNP Markers of Chronopatholo-

- gies Are Predicted by a Significant Change in the Affinity of TATA-Binding Protein for Human Gene Promoters //Bioinformatics. – 1999. – Vol. 15. – P. 687-703.
167. Ponomarenko M.P., Suslov V.V., Gunbin K.V., Ponomarenko P.M., Vishnevsky O.V., Kolchanov N.A. Identification of the relationship between variability of expression of signaling pathway genes in the human brain and affinity of TATA-binding protein to their promoters //Vavilov Journal of Genetics and Breeding. – 2014. – No.18. – P. 1219-1230.
168. Ponomarenko P.M., Ponomarenko M.P. Sequence-based prediction of transcription up-regulation by auxin in plants //Journal of Bioinformatics and Computational Biology. – 2015. – Vol. 13. – No. 1.
169. Ponomarenko P.M., Ponomarenko M.P., Drachkova I.A., Lysova M.V., Arshinova T.V., Savinkova L.K., Kolchanov N.A. Prediction of the affinity of the TATA-binding protein to TATA boxes with single nucleotide polymorphisms //Molecular Biology (Moscow). – 2009. – Vol. 43. – No. 3 – P. 512-520.
170. Ponomarenko P.M., Suslov V.V., Savinkova L.K., Ponomarenko M.P., Kolchanov N.A. A precise equation of equilibrium of four steps of TBP binding with the TATA box for prognosis of phenotypic manifestation of mutations //Biophysics (Moscow). – 2010. – No. 55. – P. 358–369.
171. Powell R., Parkhurst K., Parkhurst L. Comparison of TATA-binding protein recognition of a variant and consensus DNA promoters //Journal of Biological Chemistry. – 2002. – Vol. 277. – P. 7776-7784.
172. Proudfoot N.J. Eukaryotic promoters? //Nature. – 1979. – Vol. 279. – No. 5712. – P. 376.

173. Pugh B.F. Control of gene expression through regulation of the TATA binding protein //Gene. – 2000. – Vol. 255. – No. 1. – P. 1-14.
174. Pugh B.F. Purification of the human TATA-binding protein, TBP //Methods Mol Biol. –1995. – No. 37. – P. 359-67.
175. Purdham D.M., Zou M.X., Rajapurohitam V., Karmazyn M. Rat heart is a site of leptin production and action //Am. J. Physiol. Heart. Circ. Physiol. – 2004. – Vol. 287. – No. 6. – P. 2877-84.
176. Rashidzadeh H., Khrapunov S., Chance M.R., Brenowitz M. Solution structure and interdomain interactions of the *Saccharomyces cerevisiae* "TATA binding protein" (TBP) probed by radiolytic protein footprinting //Biochemistry. – 2003. – Vol. 42. – No. 13. – P. 3655-65.
177. Raunio H., Rautio A., Gullstén H., Pelkonen O. Polymorphisms of CYP2A6 and its practical consequences //Br J Clin Pharmacol. – 2001. – Vol. 52. – No. 4. – P. 357-63.
178. Reeve J.N. Archaeal chromatin and transcription //Mol Microbiol. – 2003. – Vol. 48. – No. 3. – P. 587-98.
179. Reeves G.K., Pirie K., Beral V., Green J., Spencer E., Bull D. et al. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study //BMJ. – 2007. – Vol. 335. – No. 7630. – P. 1134.
180. Reijnen M.J., Sladek F.M., Bertina R.M., Reitsma P.H. Disruption of a binding site for hepatocyte nuclear factor 4 results in hemophilia B Leyden //Proc Natl Acad Sci U S A. – 1992. – Vol. 89. – No. 14. – P. 6300-3.
181. Rhee H.S., Pugh B.F. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution //Cell. – 2011. – Vol. 147. – No. 6. – P. 1408-19.

182. Rodriguez C., Calle E.E., Fakhrabadi-Shokoochi D., Jacobs E.J., Thun M.J. Body mass index, height, and the risk of ovarian cancer mortality in a prospective cohort of postmenopausal women // *Cancer Epidemiol Biomarkers Prev.* – 2002. – Vol. 11. – No. 9. – P. 822–8.
183. Rubin E.M., Krauss R.M., Spangler E.A., Verstuyft J.G., Clift S.M. Inhibition of early atherogenesis in transgenic mice by human apolipoprotein AI // *Nature.* – 1991. – Vol. 353. – No. 6341. – P. 265-7.
184. Sandelin A., Carninci P., Lenhard B., Ponjavic J., Hayashizaki Y., Hume D.A. Mammalian RNA polymerase II core promoters: insights from genome-wide studies // *Nat Rev Genet.* – 2007. – Vol. 8. – No. 6. – P. 424-36.
185. Sassone-Corsi P., Corden J., Kédinger C., Chambon P. Promotion of specific in vitro transcription by excised "TATA" box sequences inserted in a foreign nucleotide environment // *Nucleic Acids Res.* – 1981. – Vol. 9. – No. 16. – P. 3941-58.
186. Savinkova L., Drachkova I., Arshinova T., Ponomarenko P., Ponomarenko M., Kolchanov N. An experimental verification of the predicted effects of promoter TATA-box polymorphisms associated with human diseases on interactions between the TATA boxes and TATA-binding protein // *PLoS One.* – 2013. – Vol. 8. – No. 2. – P. 22-23.
187. Schechter A.N. Hemoglobin research and the origins of molecular medicine // *Blood.* – 2008. – Vol. 112. – No. 10. – P. 3927-3938.
188. Sherry S.T., Ward M.H., Kholodov M., Baker J., Phan L., Smigielski E.M., Sirotkin K. dbSNP: the NCBI database of genetic variation // *Nucleic Acids Res.* – 2001. – Vol. 29. – No. 1. – P. 308-11.

189. Shimada T., Yamazaki H., Guengerich F.P. Ethnic-related differences in coumarin 7-hydroxylation activities catalyzed by cytochrome P4502A6 in liver microsomes of Japanese and Caucasian populations //Xenobiotica. – 1996. – Vol. 26. – No. 4. – P. 395-403.
190. Simoni M., Nieschlag E., Gromoll J. Isoforms and single nucleotide polymorphisms of the FSH receptor gene: implications for human reproduction //Hum Reprod Update. – 2002. – Vol. 8. – No. 5. – P. 413-21.
191. Sisen D.P, Minchinton R.M. Impact mannose-binding lectin on susceptibility to infection diseases //Clin Infect Dis. – 2003. – Vol. 37. – No. 1. – P. 1496-505.
192. Skvortsova V.I., Limborska S.A., Slominsky P.A. et al. Sporadic amyotrophic lateral sclerosis associated with the D90A CuZn-superoxide dismutase mutations in Russia //European Journal of Neurology. – 2001. – Vol. 8. – P. 167-172.
193. Smale S.T., Kadonaga J.T. The RNA polymerase II core promoter //Annu Rev Biochem. – 2003. – Vol. 72. – P. 449-79.
194. Sobhani I., Bado A., Vissuzaine C., Buyse M., Kermorgant S. et al. Leptin secretion and leptin receptor in the human stomach //Gut. – 2000. – Vol. 47. – No. 2. – P. 178–183.
195. Starr D.B., Hawley D.B. TFIID binds in the minor groove of the TATA box //Cell. – 1991. – Vol. 67. – P. 1231-1240.
196. Starr D.B., Hoopes B.C., Hawley D.K. DNA bending is an important component of site-specific recognition by the TATA binding protein //J Mol Biol. – 1995. – Vol.250. – No.4. – P. 434-46.

197. Stewart J.J., Fischback J.A., Chen X., Stargell L.A. Non optimal TATA element exhibit diverse mechanistic consequences //J. Biol. Chem. – 2006. – Vol. 281. – No. 32. – P. 22665-22673.
198. Stewart J.J., Stargell L.A. The stability of the TFIIA-TBP-DNA complex is dependent on the sequence of the TATAAA element //J. Biol. Chem. – 2001. – Vol. 276. – No. 32. – P. 30078-30084.
199. Stothard K.J., Tennant P.W., Bell R., Rankin J. Maternal overweight and obesity and the risk of congenital anomalies: a systematic review and meta-analysis //JAMA. – 2009. – Vol. 301. – No. 6. – P. 636–50.
200. Strahs D., Barash D., Qian X., Schlick T. Sequence-dependent solution structure and motion of 13 TATA/TBP (TATA-box binding protein) complexes// Biopolymers. – 2003. – Vol. 69. – P. 216-244.
201. Suslov V.V., Ponomarenko P.M., Efimov V.M., Savinkova L.K., Ponomarenko M.P., Kolchanov N.A. SNPs in the HIV-1 TATA box and the AIDS pandemic. Journal of Bioinformatics and Computational Biolog. – 2010b. – No. 8. – P. 607-625.
202. Suslov V.V., Ponomarenko P.M., Ponomarenko M.P., Drachkova I.A., Arshinova T.V., Savinkova L.K., Kolchanov, N.A. TATA box polymorphisms in genes of commercial and laboratory animals and plants associated with selectively valuable traits //Russian Journal of Genetics. – 2010a. – No. 46. – P. 394-403.
203. Takihara Y., Nakamura T., Yamada H., et al. A novel mutation in the TATA box in a Japanese patient with beta + -thalassemia //Blood. – 1986. – Vol. 67. – P. 547–550.

204. Tallack M.R., Whittington T., Yuen W.S. et al. A global role for KLF1 in erythropoiesis revealed by ChIP-seq in primary erythroid cells //Genome Res. – 2010. – Vol. 20. – P. 1052-63.
205. Teufel A., Krupp M., Weinmann A., Galle P.R. Current bioinformatics tools in genomic biomedical research (Review) //Int J Mol Med. – 2006. – Vol. 17. – No. 6. – P. 967-73.
206. Thein S.L. Genetic modifiers of the  $\beta$ -haemoglobinopathies //British J. Haematol. – 2008. – Vol. 141. – P. 357-366.
207. Thomas M.C., Chiang C.M. The general transcription machinery and general cofactors //Critical reviews in biochemistry and molecular biology. – 2006. – Vol. 41. – No. 3. – P. 105-178.
208. Tolić-Nørrelykke, S.F., Rasmussen, M.B, Pavone, F.S, Berg-Sørensen, K., Oddershede, L.B. Stepwise bending of DNA by a single TATA-box binding protein //Biophysical Journal. – 2006. – Vol. 90. – No. 10. – P. 3694-3703.
209. Tora L., Timmers H.T. The TATA box regulates TATA-binding protein (TBP) dynamics in vivo //Trends Biochem Sci. – 2010. – Vol. 35. – No. 6. – P. 309-14.
210. Torekov S.S., Ma L., Grarup N., Hartmann B., Hainerová I.A. et al. Homozygous carriers of the G allele of rs4664447 of the glucagon gene (GCG) are characterised by decreased fasting and stimulated levels of insulin, glucagon and glucagon-like peptide (GLP)-1 //Diabetologia. – 2011. – Vol. 54. – No. 11. – P. 2820-3.
211. Tournamille C., Colin Y., Cartron J.P., Le Van Kim C. Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals //Nat Genet. – 1995. – Vol. 10. – No. 2. – P. 224-8.



212. Traeger-Synodinos J., Vrettou C., Kanavakis E. Rapid detection of fetal Mendelian disorders: thalassemia and sickle cell syndromes //Methods Mol Biol. – 2008. – Vol. 444. – P. 133-145.
213. Trovato G.M. Sustainable medical research by effective and comprehensive medical skills: overcoming the frontiers by predictive, preventive and personalized medicine //EPMA Journal. – 2014. – No. 5. – P. 14.
214. Tumer N., Erdos B., Matheny M., Cudykier I., Scarpace P.J. Leptin antagonist reverses hypertension caused by leptin overexpression, but fails to normalize obesity-related hypertension //J Hypertens. – 2007. – Vol. 25. – No. 12. – P. 2471–8.
215. Valentin C., Pissard S., Martin K. et al. Triose phosphate isomerase deficiency in 3 French families: two novel null alleles, a frameshift mutation (TPI Alfortville) and an alteration in the initiation codon (TPI Paris) // Blood. – 2000. – V. 96. – No. 3. – P. 1130-1135.
216. Vassall K.A, Stubbs H.R., Primmer H.A. et al. Decreased stability and increased formation of soluble aggregates by immature superoxide dismutase do not account for disease severity in ALS //Proc Natl Acad Sci U S A. – 2011. – Vol. 108. – No. 6. – P. 210-215.
217. Veenstra G.J., Wolffe A.P. Gene-selective developmental roles of general transcription factors //Trends Biochem. Sci. – 2001. – Vol. 26. – No. 11. – P. 665-71.
218. Vonk W.I., Klomp L.W. Role of transition metals in the pathogenesis of amyotrophic lateral sclerosis //Biochem Soc Trans. – 2008. – Vol. 36. – P. 1322-8.
219. Waga C., Iwahashi K. CYP2A6 gene polymorphism and personality traits for NEO-FFI on the smoking behavior of youths //Drug Chem Toxicol. – 2007. – Vol. 30. – No. 4. – P. 343-9.

220. Wang X., Lu Y., Yang J., Shi Y., Lan M. et al. Identification of triosephosphate isomerase as an anti-drug resistance agent in human gastric cancer cells using functional proteomic analysis // *J Cancer Res Clin Oncol.* – 2008. – Vol. 134. – No. 9. – P. 995-1003.
221. Wasylyk B., Derbyshire R., Guy A., Molko D. et al. Specific in vitro transcription of conalbumin gene is drastically decreased by single-point mutation in T-A-T-A box homology sequence // *Proc Natl Acad Sci U S A.* – 1980. – Vol. 77. – No. 12. – P. 7024-7028.
222. Watanabe M., Zingg B.C., Mohrenweiser H.W. Molecular analysis of a series of alleles in humans with reduced activity at the triosephosphate isomerase locus // *Am. J. Hum. Genet.* – 1996. – V. 58. – P. 308-316.
223. Weake V.M., Workman J.L. Inducible gene expression: diverse regulatory mechanisms // *Nature reviews.* – 2010. – Vol. 11. – No. 6. – P. 426-437.
224. Welter D., MacArthur J., Morales J., Burdett T., Hall P. et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations // *Nucleic Acids Res.* – 2014. – No. 42(Database issue). – P. D1001-6.
225. Whittington J.E., Delgadillo R.F., Attebury T.J., Parkhurst L.K., Daugherty M.A., Parkhurst L.J. TATA-binding protein recognition and bending of a consensus promoter are protein species dependent // *Biochemistry.* – 2008. – Vol. 47. – No. 27. – P. 7264–7273.
226. Wobbe C.R., Strahl K. Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription in vitro // *Mol. Cell. Biol.* – 1990. – Vol. 10. – No. 8. – P. 3859-3867.

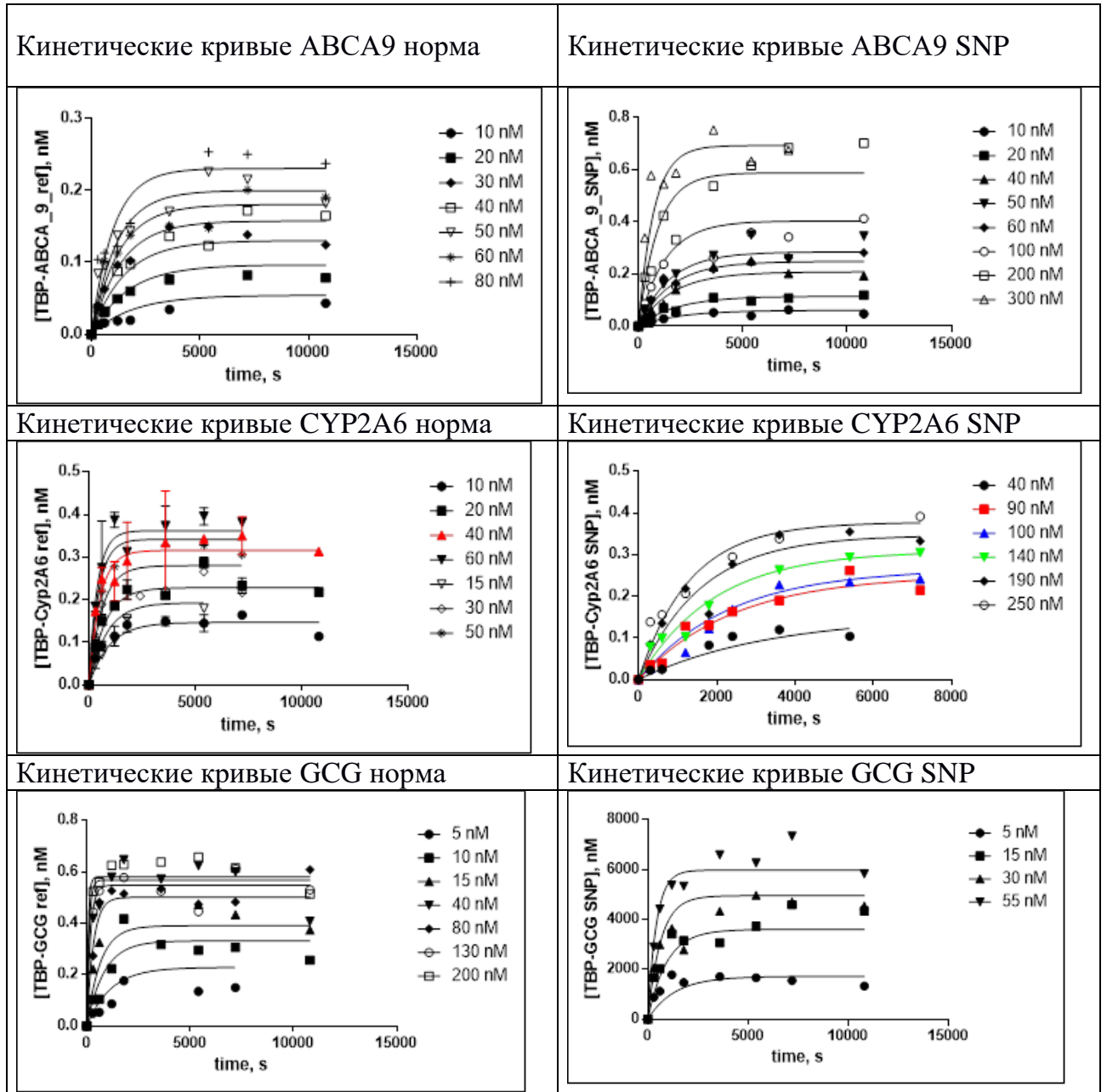
227. Wobbe C. R., Struhl K. Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription in vitro //Mol Cell Biol. – 1990. – Vol. 10. – No. 8. – P. 3859–3867.
228. Wolner B.S., Gralla J.D. TATA-flanking sequences influence the rate and stability of TATA-binding protein and TFIIB binding //J Biol Chem. – 2001. – Vol. 276. – No. 9. – P. 6260-6.
229. Wu J., Parkhurst K.M., Powell R.M., Brenowitz M., Parkhurst L.J. DNA bends in TATA-binding protein-TATA complexes in solution are DNA sequence-dependent //Journal of Biological Chemistr. – 2001. – Vol. 276. – P. 14614-14622.
230. Yamada K., Santo-Yamada Y., Wada E., Wada K. Role of bombesin (BN)-like peptides/receptors in emotional behavior by comparison of three strains of BN-like peptide receptor knockout mice //Mol Psychiatry. – 2002. – Vol. 7. – No. 1. – P. 113-7.
231. Yang C., Bolotin E., Jiang T., Sladek F.M., Martinez E. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters //Gene. – 2007. – Vol. 389. – No. 1. – P. 52-65.
232. Ye D., Hoekstra M., Out R., Meurs I., Kruijt J.K., Hildebrand R.B., Van Berkel T.J.C., Van Eck M. Hepatic cell-specific ATP-binding cassette (ABC) transporter profiling identifies putative novel candidates for lipid homeostasis in mice //Atherosclerosis. – 2008. – Vol. 196. – No. 2. – P. 650-658.
233. Yokoyama J.S., Wang Y., Schork A.J., Thompson W.K., Karch C.M., Cruchaga C., et al. Association between genetic traits for immune-mediated diseases and Alzheimer disease //JAMA Neurol. – 2016. – Vol. 73. – No. 6. – P. 691-7.

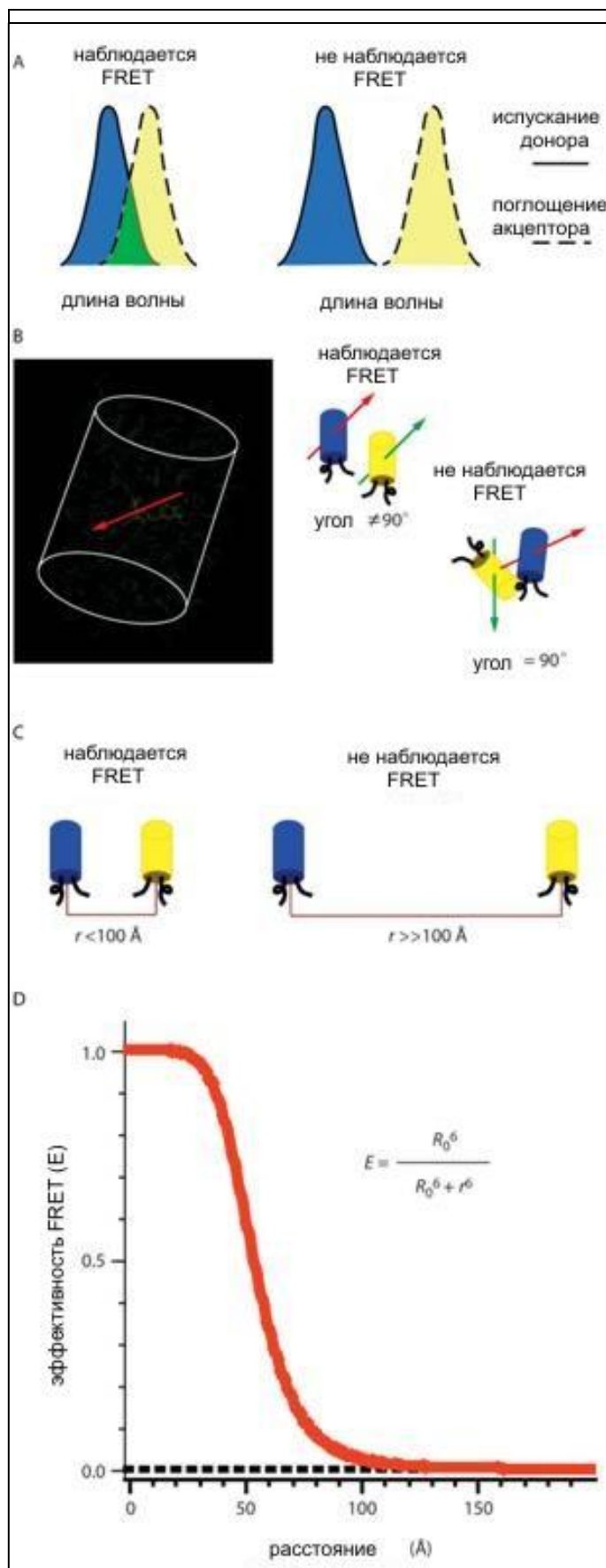
234. Yoo S.S., Jin C., Jung D.K., Choi Y.Y., Choi J.E., Lee W.K., Park J.Y. Putative functional variants of XRCC1 identified by RegulomeDB were not associated with lung cancer risk in a Korean population //Cancer Genetics. – 2015. – No. 208. – P. 19-24.
235. Yoshida R., Nakajima M., Nishimura K. et al. Effects of polymorphism in promoter region of human CYP2A6 gene (CYP2A6\*9) on expression level of messenger ribonucleic acid and enzymatic activity in vivo and in vitro //Clin Pharmacol Ther. – 2003. – V. 74. – No. 1. – P. 69-76.
236. Zampieri T.T., Ramos-Lobo A.M., Furigo I.C., Pedrosa A.B., Buonfiglio D.C., Donato J. Jr. SOCS3 deficiency in leptin receptor-expressing cells mitigates the development of pregnancy-induced metabolic changes //Mol Metab. – 2015. – Vol. 4. – No. 3. – P. 237–45.
237. Zeidan A., Broman J., Hellstrand P., Sward K. Cholesterol dependence of vascular ERK1/2 activation and growth in response to stretch: role of endothelin-1 //Arterioscler Thromb Vasc Biol. – 2003. – Vol. 23. – No. 9. – P. 1528–34.
238. Zhang L., Song H., Ge Y., Ji G., Yao Z. Temporal relationship between diet-induced steatosis and onset of insulin/leptin resistance in male Wistar rats //PLoS One. – 2015. – Vol.10. – No. 2. – e0117008.
239. Zhang Y., Proenca R., Maffei M., Barone M., Leopold L., Friedman J.M. Positional cloning of the mouse obese gene and its human homologue //Nature. – 1994. – Vol. 372. – No. 6505. – P. 425–432.
240. Zhao Y.Y., Liu L., Zhou T., Zhou N.N. et al. A retrospective analysis of the clinicopathological and molecular characteristics of pulmonary blastoma // Onco Targets Ther. – 2016. – No. 9. – P. 6915-6920. – eCollection.

241. Zhao Y.Y., Zhou J., Narayanan C.S., Cui Y., Kumar A. Role of C/A polymorphism at -20 on the expression of human angiotensinogen gene //Hypertension. – 1999. – Vol. 33. – No. 1. – P. 108-15.
242. Zheng Z., Park J.Y., Guillemette C., Schantz S.P., Lazarus P. Tobacco carcinogen-detoxifying enzyme UGT1A7 and its association with orolaryngeal cancer risk //J Natl Cancer Inst. – 2001. – Vol. 93. – No. 18. – P.1411-8.
243. Zukunft J., Lang T., Richter T. et al. A natural CYP2B6 TATA box polymorphism (-82T→C) leading to enhanced transcription and relocation of the transcriptional start site //Mol. Pharmacol. – 2005. – Vol. 67. – P. 1772-1782.

## Приложение 1

Таблица 1. Кинетические кривые взаимодействия ТСБ-ТАТА во времени.





**Рисунок 1.** Флуоресцентный резонансный перенос энергии (ФРПЭ или FRET- Förster resonance energy transfer).

(А) ФРПЭ может протекать лишь тогда, когда спектр излучения донора (сплошная линия) перекрывается со спектром возбуждения акцептора (пунктирная линия). Перекрывающиеся области спектра показаны зеленым цветом.

(В) Дипольный момент хромофора внутри белковой цилиндрической структуры (красная стрелка) имеет определенную ориентацию относительно структуры ФР. Если эмиссия диполя (красная стрелка) донора (синий цилиндр) образует угол 90 градусов с диполем (зеленая стрелка) акцептора (желтый цилиндр), то FRET не может произойти. Угол между диполями, не равный 90 градусам, приводит к появлению FRET.

(С) Резонансный перенос энергии протекает лишь тогда, когда расстояние между хромофорами ( $r$ ) менее, чем 1.8 от Фёрстеровского радиуса ( $R_0$ ), который обычно меньше, чем 100 Å.

(D) Зависимость эффективности FRET ( $E$ ) от расстояния ( $r$ ) для пары донор-акцептор, для которой  $R_0 = 54 \text{ \AA}$  и диполи ориентированы статистически.

(Behlke M.A., *et al.* Fluorescence and Fluorescence Applications // Integrated DNA Technologies. 2005. P. 1-13.)

## Приложение 2

FRET или Фёрстеровский резонансный перенос энергии (ФРПЭ).

(Parkhurst L.J., Parkhurst K.M., Powell R., Wu J., Williams S. Time-resolved fluorescence resonance energy transfer studies of DNA bending in double-stranded oligonucleotides and in DNA-protein complexes // Biopolymers. – 2001-2002. – Vol. 61. – No. 3. – P.180-200.)

Для ФРПЭ необходимо соблюдение нескольких условий:

- спектр флуоресценции донора должен перекрываться со спектром поглощения акцептора;
- донор и акцептор должны находиться не дальше определённого расстояния друг от друга;
- дипольные моменты донора и акцептора должны иметь определённое взаимное расположение в пространстве.

Эффективность ФРПЭ рассчитывается по формуле.

$$E = R_0^6 / (R_0^6 + r^6)$$

здесь  $R_0$  является Фёрстеровским радиусом: таким расстоянием между донором и акцептором, при котором эффективность переноса равна  $1/2$ .

Важным является то, что ФРПЭ происходит на расстояниях, соизмеримых с размерами биологических объектов, таких как белковые глобулы или мембраны клеток. При этом относительная эффективность переноса энергии обратно зависит от расстояния между ФРПЭ-партнёрами. Если две биомолекулы, меченые ФРПЭ-парой находятся на большом расстоянии, при возбуждении донора будет наблю-

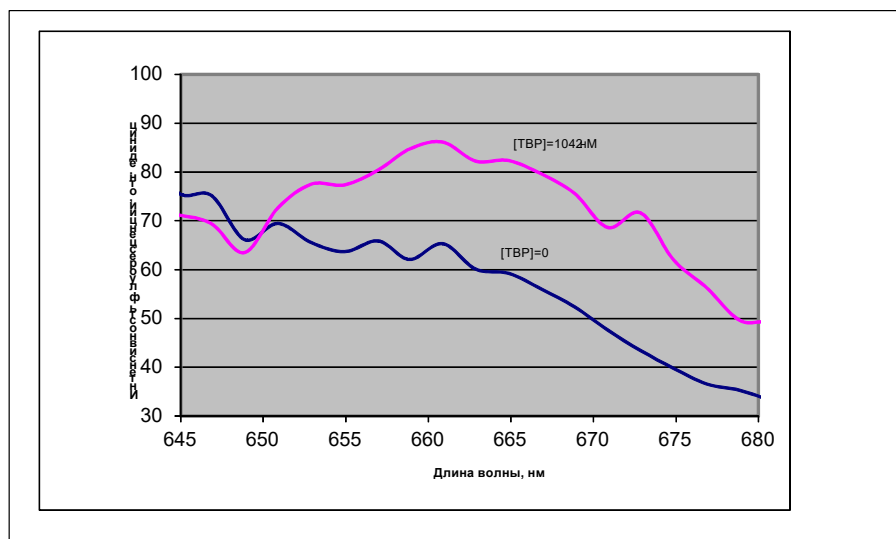


даться только его собственная флуоресценция. В случае, когда молекулы сближены в пространстве, при возбуждении донора будет наблюдаться эмиссия акцептора.

Благодаря своей зависимости от расстояния, ФРПЭ стал своеобразной «молекулярной линейкой», которая позволяет измерять расстояние между молекулами, каждая из которых мечена одним из партнёров переноса.

### Приложение 3

Подбор условий для связывания ТСБ/ТАТА.

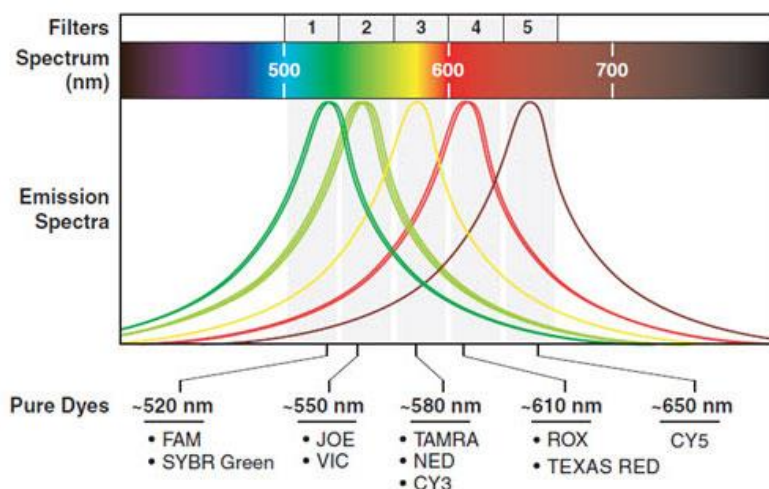


На графике изображена флуоресценция комплекса *ОДН/ТСБ* с оптимальной и нулевой концентрациями ТСБ. Флуоресцентная спектроскопия является одним из самых высокочувствительных методов, позволяющих детектировать очень низкие концентрации веществ (мкМ, нМ и даже пМ) (Векшин Н.Л. Флуоресцентная спектроскопия биополимеров. – Пуццино: Фотон-Век. - 2008). При подборе оптимальных концентраций олигонуклеотидов лучший результат был получен при минимальной концентрации  $1 \cdot 10^{-7}$ М. Также необходимо было подобрать оптимальные concentra-

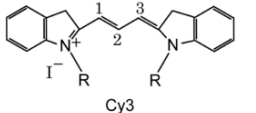
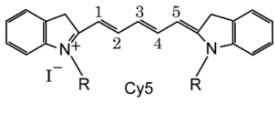
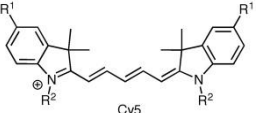
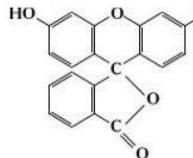
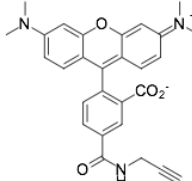
ции белка для связывания олигонуклеотида. Для этого снимались при тех же параметрах значения флуоресценции при возрастающих концентрациях ТСБ (ТСБ менялся от 0 до 1500 нМ в 11 разведениях). При построении концентрационной кривой “флуоресценция - возрастающая концентрация ТСБ” была выбрана лучшая по излучению концентрация ТСБ.

## Приложение 4

**Рисунок 1.** Спектры флуоресцентных красителей ([www.thermofisher.com](http://www.thermofisher.com)).



**Таблица 1.** Характеристики используемых пар флуоресцентных красителей.

Cy3	Cy5	FAM	TAMRA
  <a href="https://upload.wikimedia.org/wikipedia/commons/thumb/9/90/Cy3_Cy5_dyes.gif/300px-Cy3_Cy5_dyes.gif">https://upload.wikimedia.org/wikipedia/commons/thumb/9/90/Cy3_Cy5_dyes.gif/300px-Cy3_Cy5_dyes.gif</a>	 <a href="https://heterocyclist.files.wordpress.com/2011/12/cyanine_struct.jpg">https://heterocyclist.files.wordpress.com/2011/12/cyanine_struct.jpg</a>	 <a href="http://ru.lumiprobe.com/img/p/structure/6-fam-azide.gif">ru.lumiprobe.com/img/p/structure/6-fam-azide.gif</a>	 <a href="http://www.lumiprobe.com/img/p/structure/5-tamra-alkyne.png">www.lumiprobe.com/img/p/structure/5-tamra-alkyne.png</a>
<i>Длина волны возбуждения/излучения в эксперименте, нм</i>			
550	>645	494	>570
<i>Квантовый выход флуоресценции (<math>\Phi &lt; 1</math>)</i>			
0,15	0,27	0,8	0,7

en.wikipedia.org/wiki/Cyanine	en.wikipedia.org/wiki/Cyanine	www.syntol.ru/catalog/zondy-dlya-ptsr-v-realnom-vremeni/fluorestsentye-krasiteli.html	www.syntol.ru/catalog/zondy-dlya-ptsr-v-realnom-vremeni/fluorestsentye-krasiteli.html
-------------------------------	-------------------------------	---	---

## Приложение 5

Кинетическое моделирование.

Для определения *минимальной кинетической модели* взаимодействия ТСБ с ОДН и расчета констант скорости всех элементарных стадий данной модели использовали программу Dynafit (Biokin, США). Эффективность и точность определения параметров для исследуемого процесса зависит от числа параметров и ограничений, налагаемых на обрабатываемые данные. При подборе схемы взаимодействия для полученных данных использовался метод нелинейной регрессии по алгоритму Левенберга-Маркара. Кинетические кривые разбивали на несколько временных диапазонов таким образом, что каждый следующий диапазон полностью включал предыдущий, например от 2 мс до 1, 10, 100 и так далее. Это дает возможность постепенно усложнять и оптимизировать модель взаимодействия белка с субстратом. Обработка кривых в диапазоне **2 мс – 10 секунд**, как правило, отражает **начальные стадии связывания** и дает возможность определить их константы скорости. При обработке следующего по величине диапазона ранее использованный механизм усложнялся, например, к нему добавлялась еще одна равновесная или неравновесная стадия, при этом фиксировались константы, определенные для начальных стадий. После расчета констант скорости добавленных стадий осуществлялось общее корректирование значений констант скорости путем одновременной оптимизации всех параметров системы. Такой пошаговый подход позволил

дискриминировать различные схемы взаимодействия белок-ОДН и определить константы скорости и равновесия для элементарных стадий, входящих в эти схемы.

**Dynafit** – проводит численное интегрирование системы дифференциальных уравнений, которые соответствуют любой кинетической схеме (Kuzmic P. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase //Anal Biochem. – 1996. – Vol. 237. – No. 2. – P. 260-73).

В эту систему дифференциальных уравнений входят:

- все концентрации реагентов
- константы скорости, которые характеризуют процесс превращения интермедиатов в этой цепи.

Если мы видим изменения на кинетической кривой, то тогда мы их пытаемся обсчитать. Для этого берем небольшой временной промежуток от начала и пытаемся описать минимальной схемой. Пусть, например, мы видим, что процесс не описывается одностадийной моделью. Тогда мы пытаемся описать двустадийной моделью. Предположим, модель описала хорошо наши данные. Берем тогда более длинный промежуток от начальной точки и пытаемся описать данные 2-стадийной моделью и т.д.