

Федеральное государственное бюджетное учреждение науки  
**ИНСТИТУТ БИООРГАНИЧЕСКОЙ ХИМИИ**  
*им. академиков М.М. Шемякина и Ю.А. Овчинникова*  
*Российской академии наук*  
*(ИБХ РАН)*

---

*На правах рукописи*

Смирнов Николай Андреевич

**ИССЛЕДОВАНИЕ АКТИВНОСТИ ПОТЕНЦИАЛЬНЫХ  
ИНСУЛЯТОРНЫХ И ЭНХАНСЕРНЫХ ЭЛЕМЕНТОВ ГЕНОМА  
ЧЕЛОВЕКА**

03.01.03 – Молекулярная биология

**АВТОРЕФЕРАТ**

диссертации на соискание ученой степени  
кандидата биологических наук

Москва – 2017

Работа выполнена в лаборатории структуры и функций генов человека Федерального государственного бюджетного учреждения науки Института биоорганической химии имени академиков М. М. Шемякина и Ю. А. Овчинникова Российской академии наук

**Научный руководитель:**

доктор биологических наук

Акопов Сергей Борисович

**Официальные оппоненты:**

Чуриков Николай Андреевич, д.б.н., профессор, заведующий лабораторией организации генома Федерального государственного бюджетного учреждения науки Института молекулярной биологии имени В.А.Энгельгардта Российской академии наук (ИМБ РАН),

Артамонова Ирина Игоревна к.б.н., старший научный сотрудник отдела вычислительной биологии Федерального государственного бюджетного учреждения науки Института общей генетики им. Н.И. Вавилова Российской академии наук (ИОГен РАН).

**Ведущая организация:**

Федеральное государственное бюджетное учреждение науки Институт биологии гена Российской академии наук (ИБГ РАН).

Защита состоится «24» мая 2017 г. в 10 часов на заседании диссертационного совета Д 002.019.01 при Федеральном государственном бюджетном учреждении науки Институте биоорганической химии им. академиков М. М. Шемякина и Ю. А. Овчинникова Российской академии наук по адресу: 117997, г. Москва, ул. Миклухо-Маклая, д.16/10.

С диссертацией можно ознакомиться в библиотеке Федерального государственного бюджетного учреждения науки Института биоорганической химии им. академиков М. М. Шемякина и Ю.А. Овчинникова Российской академии наук, а также на **сайте [www.ibch.ru/about/disser/theses](http://www.ibch.ru/about/disser/theses)**.

Автореферат разослан «\_\_» \_\_\_\_\_ 2017 года.

Ученый секретарь Диссертационного совета,

доктор физико-математических наук

  
В. А. Олейников

## **Характеристика работы**

### **Актуальность проблемы**

Регуляция транскрипции является процессом, играющим основополагающую роль в развитии организмов, так как функциональность генов определяется не только их продуктами, но, также, специфическим паттерном их экспрессии. Большое количество элементов вовлечено в процесс регуляции, как на транскрипционном, так и на пост-транскрипционном уровне. Элементы, которые участвуют в регуляции генов на транскрипционном уровне, включают в себя промоторы, энхансеры, сайленсеры, инсуляторы. В настоящее время установлено, что сложные функциональные и пространственные взаимосвязи между геномными регуляторными элементами, в конечном счете, определяют функциональную идентичность клетки.

Исследования генома, проводившиеся в последние 20 лет, в основном были нацелены на интенсивное изучение и идентификацию белок-кодирующих генов и только отчасти на изучение сложной регуляторной сети, обеспечивающей их экспрессию. Использование методов сравнительной геномики показало, что большая часть консервативных и недавно приобретенных в ходе эволюции участков генома представлена некодирующими последовательностями, играющими структурную и/или функциональную роль. Недавние полногеномные исследования показали, что некодирующие элементы играют важнейшую роль в предрасположенности к возникновению заболеваний.

Несмотря на важную роль, которую играют регуляторные элементы в экспрессии генов, степень изученности их функций в данный момент является недостаточной и большая часть регуляторных элементов не охарактеризована. Основной причиной этого является то, что количество регуляторных элементов в несколько раз превышает количество генов, а также, то, что появление методик для высокоэффективного полногеномного картирования регуляторных элементов стало возможным относительно недавно, в то время как средне- и высокоэффективные подходы для картирования генов, и промоторов являются отработанными и широко распространенными.

К сожалению, функциональный анализ регуляторных элементов на уровне генома сопряжен с рядом технических трудностей и, несмотря на то, что использование полногеномных методов позволило накопить огромные массивы данных, следует учитывать, что картирование большого числа функциональных элементов, выполненное с применением методик, основанных на секвенировании нового поколения, сопряжено с большим количеством ошибок. Поэтому исчерпывающий функциональный анализ регуляторных элементов на уровне отдельных сопряженных сегментов с последующим интегрированием полученных данных представляется более предпочтительным.

## Цели и задачи работы

Целью диссертационной работы является функциональный анализ *цис*-регуляторных элементов (энхансеры, инсуляторы) в локусе *FXYD5-COX7A1* хромосомы 19 человека. Регуляторные элементы для анализа были выбраны на основании данных ChIP-seq проекта ENCODE (Encyclopedia of DNA Elements), а также из элементов, идентифицированных ранее с использованием метода позитивно-негативной селекции (Акоров et al. 2006, Дидыч и соавт. 2009). Был проведен также детальный функциональный анализ двунаправленного промотора генов *PSENN* и *U2AF1L4* и энхансера 12, расположенного во втором интроне гена *U2AF1L4*, ранее идентифицированного в нашей лаборатории (Chernov et al., 2008). Проанализирован эффект *транс*-активации промотора энхансером с использованием системы экспрессии репортерного гена.

В ходе работы были поставлены и решены следующие задачи:

Используя методы сравнительной геномики выявить консервативные области в энхансере 12 и разделить последовательность энхансера на перекрывающиеся фрагменты для их дальнейшего изучения. Для идентификации области, отвечающей за энхансерную активность, провести функциональный анализ полученных перекрывающихся фрагментов энхансера 12, используя систему экспрессии репортерного гена. С помощью EMSA (Electrophoretic Mobility Shift Assay) анализа провести функциональный анализ ортологичных последовательностей приматов для подтверждения функциональной роли консервативных областей энхансера 12.

Используя методику 5'-RACE (Rapid amplification of cDNA ends), определить точки начала транскрипции промотора генов *PSENN* и *U2AF1L4*, расположенного в непосредственной близости от энхансера 12. Используя метод экспрессии репортерного гена, провести анализ активности промотора генов *PSENN* и *U2AF1L4*.

С помощью количественной ОТ-ПЦР (Полимеразная Цепная Реакция с Обратной Транскрипцией) определить уровень и тканеспецифичность экспрессии генов *PSENN* и *U2AF1L4*. Определить консервативные области промотора генов *PSENN* и *U2AF1L4*, используя методы сравнительной геномики, и с помощью сайт-специфического мутагенеза выявить области, являющиеся функциональными.

Создать систему, предназначенную для анализа энхансер-блокирующей активности фрагментов ДНК *in vitro*, в том числе плазмиду pGL4EPV, содержащую в регуляторной области гена люциферазы светлячка энхансер и промотор вируса SV40.

Клонировать в плазмиду pGL4PV, содержащую в регуляторной области гена люциферазы светлячка промотор вируса SV40, десять последовательностей потенциальных энхансеров, выбранных на основании данных ChIP-seq проекта ENCODE. Полученными

векторными конструкциями провести транзистентные трансфекции клеточных линий HeLa и HepG2 с последующим анализом энхансерной активности анализируемых фрагментов.

Клонировать в вектор pGL4EPV между энхансером и минимальным промотором SV40 последовательности четырех потенциальных инсуляторов, обнаруженных нами ранее методом позитивно-негативной селекции в локусе 19q13.12, и 5 последовательностей, выбранных на основании данных ChIP-seq проекта ENCODE, а также контрольную последовательность 5'-HS4 инсулятора кур. Полученными векторными конструкциями провести транзистентные трансфекции клеточных линий HeLa, CHO и HepG2 с последующим анализом энхансер-блокирующей активности анализируемых фрагментов.

Создать систему, предназначенную для анализа промоторной и клеточной специфичности активации промотора энхансером в *транс*-положении при транзистентных трансфекциях. С использованием полученной системы изучить способность клеточных и вирусных энхансеров активировать различные промоторы в *транс*-положении. Используя метод фиксации конформации хромосомом 3C (Chromosome Conformation Capture), продемонстрировать физическую сближенность промотора с энхансером при *транс*-активации.

### **Научная новизна и практическая значимость работы**

В работе с использованием подходов *in vivo* и *in vitro* проведен детальный анализ энхансера, расположенного во втором интроне гена *U2AF1L4* в локусе *FXYD5-COX7A1* хромосомы 19 человека. Показано, что данный элемент имеет консервативные области и сохраняет свою функциональность в эволюционном ряду приматов.

Показано, что область между генами *PSENE1* и *U2AF1L4* на хромосоме 19 человека является истинным двунаправленным промотором. Продемонстрирована согласованная экспрессия генов *U2AF1L4* и *PSENE1* в нескольких клеточных линиях. Определены консервативные области и показана их функциональная роль в работе промотора. В связи с тем, что регуляция генов *U2AF1L4* и *PSENE1* не изучена и не определены элементы и механизмы, отвечающие за их регуляцию, полученные данные являются важным и необходимым этапом для дальнейшего изучения данных генов.

Получена система для функционального анализа регуляторных последовательностей ДНК, основанная на транзистентной трансфекции и двойной люциферазной детекции. С использованием полученной системы проведен функциональный анализ энхансерной, сайленсерной и энхансер-блокирующей активности последовательностей, расположенных в локусе *FXYD5-COX7A1* хромосомы 19 человека.

Полученные данные позволили расширить количество известных валидированных инсульторов и энхансеров, а также выявили некоторые особенности данных элементов, такие как зависимость от ориентации и тканеспецифичность. Новые проверенные элементы могут нести в себе как фундаментальную значимость, являясь основой для дальнейшего изучения функциональной роли данных элементов в геноме, так и практическую ценность для использования в генной терапии и для защиты трансгенных конструкций от эффекта положения.

Получена система для количественной оценки эффекта *транс*-активации промотора энхансером. С использованием данной системы установлено, что способность активировать промотор *в транс*-положении при транзистентных трансфекциях не ограничена лишь вирусными энхансерами, как было показано ранее, данное свойство также характерно и для геномных регуляторных элементов. Показано, что активация промотора энхансером *в транс*-положении зависит от промотора и типа клеток и с помощью 3С анализа продемонстрировано, что данный эффект включает в себя физический контакт промотора с энхансером.

Данная диссертационная работа является частью проекта, направленного на создание исчерпывающей функциональной карты *цис*-регуляторных элементов в локусе *FXYD5-COX7A1* длиной 1 млн. п.н хромосомы 19 человека.

### **Апробация работы**

По теме диссертационной работы опубликованы 4 статьи в отечественных и зарубежных научных журналах. Результаты работы были представлены на российских и международных конференциях.

### **Структура диссертации**

Диссертационная работа изложена на 119 страницах и состоит из введения, обзора литературы, экспериментальной части, результатов и их обсуждения, основных выводов и списка цитируемой литературы, включающего 275 ссылок. Диссертация содержит 9 таблиц и 22 рисунка.

## Содержание работы

### Функциональная диссекция энхансерного элемента, расположенного во втором интроне гена *U2AF1L4* человека

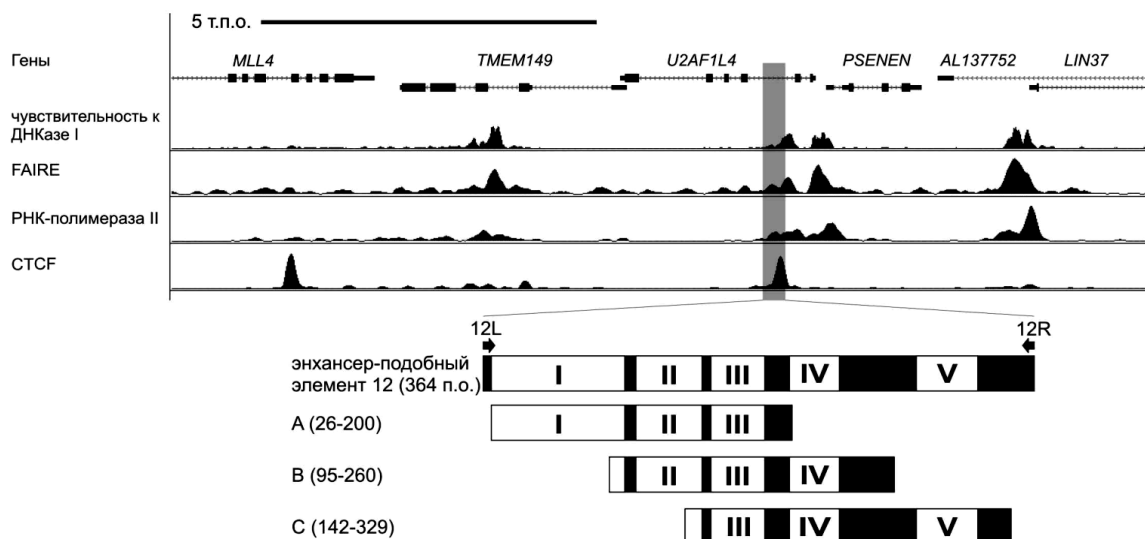
Ранее в нашей лаборатории был идентифицирован ряд коротких (200–500 п.н.) геномных последовательностей, проявляющих энхансерную активность в составе ретровирусной конструкции, интегрированной в геном клеток HeLa (Chernov et al. 2008). Обнаруженные последовательности были картированы в полигенном участке 19q13.12 хромосомы 19 человека длиной 1 млн п.н. Также была проверена их энхансерная активность в составе вектора pGL3PV с использованием системы двойной люциферазной детекции в клетках линии HeLa (Chernov et al. 2008).

Мы провели функциональный анализ одного из наиболее сильных обнаруженных энхансерных элементов (энхансер 12), расположенного во втором интроне гена *U2AF1L4* (Рис. 1), используя систему транзientной экспрессии репортерного гена.

#### *Анализ активности фрагментов энхансера 12*

Для того, чтобы определить наиболее консервативные участки внутри энхансера, нами было проведено выравнивание последовательностей энхансера 12 человека и ортологичных ему последовательностей семи видов приматов. Затем были выбраны три перекрывающихся фрагмента А, В и С для того, чтобы определить, какой из участков энхансера 12 отвечает за его активность (Рис. 1). Выбранные фрагменты клонировали в вектор pGL3PV, содержащий ген люциферазы светлячка под контролем промотора SV40.

Для характеристики активностей энхансера 12 и его фрагментов была проведена серия транзientных трансфекций клеток HeLa полученным набором плазмид (Рис. 2). Согласно полученным данным, наибольшую активность проявляет средний фрагмент В, усиливая активность промотора SV40 в два раза, что указывает на наличие в этом фрагменте участка, ответственного за энхансерную активность. Примечательно, что фрагмент С, а, следовательно, и полноразмерный энхансер, содержат участок связывания белка CTCF. Известно, что CTCF в некоторых случаях взаимодействует с сайленсерами и может снижать активность пары промотор–энхансер, поэтому его присутствие может приводить к некоторому снижению активности полноразмерной последовательности энхансера и его фрагмента С. Фрагмент А не обладает энхансерной активностью в данной системе, поэтому с учетом перекрытия данного фрагмента с фрагментом В можно предполагать, что ядро энхансерного элемента расположено в 3' области фрагмента В.



**Рис. 1.** Расположение энхансера 12 относительно ближайших генов на хромосоме 19 человека. Внизу показаны полноразмерный энхансер 12 и его перекрывающиеся делеционные фрагменты (А, В и С), для которых проведен функциональный анализ энхансерной активности в клеточной линии HeLa (Рис. 2). Римскими цифрами обозначены консервативные участки внутри элемента.

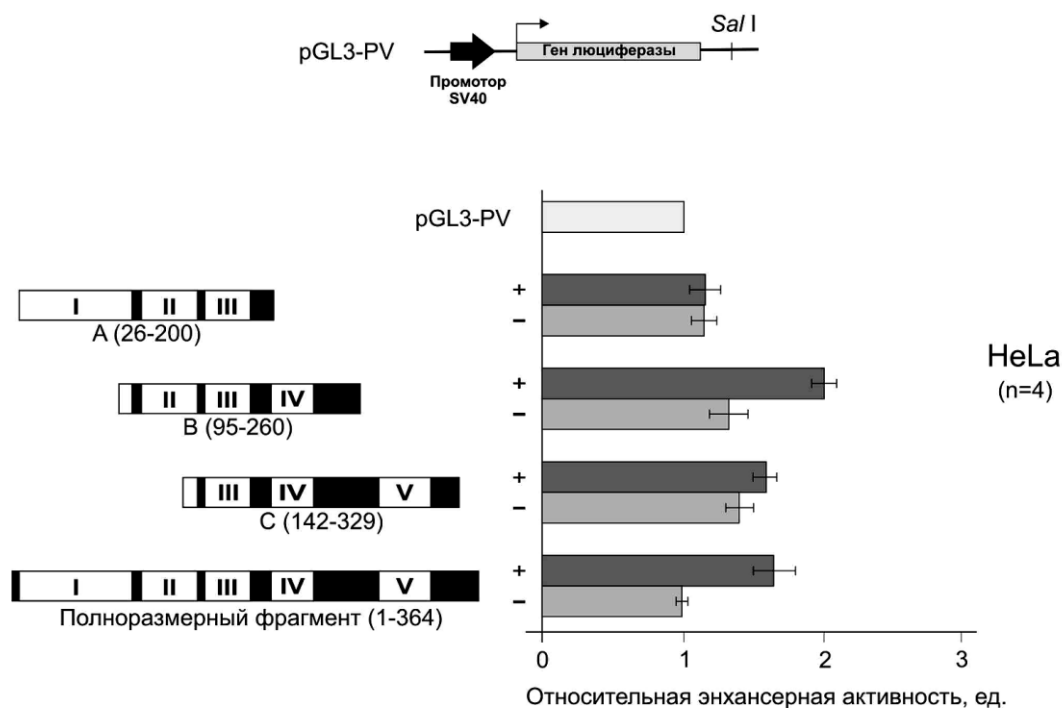
Стоит отметить, что активность проанализированных последовательностей зависит от их ориентации относительно промотора, что не характерно для классических энхансеров. Так полноразмерный энхансер 12, а также его фрагменты В и С проявляли наибольшую активность, находясь в прямых ориентациях относительно промотора (Рис. 2).

### ***Анализ функциональной консервативности энхансера 12***

Чтобы показать функциональную консервативность энхансера 12 в ряду приматов, последовательности энхансера 12 человека и трех приматов (*Hylobates lar*, *Macaca mulatta*, *Saimiri sciureus*) были клонированы в pGL3 вектор, содержащий ген люциферазы светлячка под контролем SV40 промотора. Полученными плазмидными векторами были поставлены транзientные котрансфекции клеточной линии HeLa.

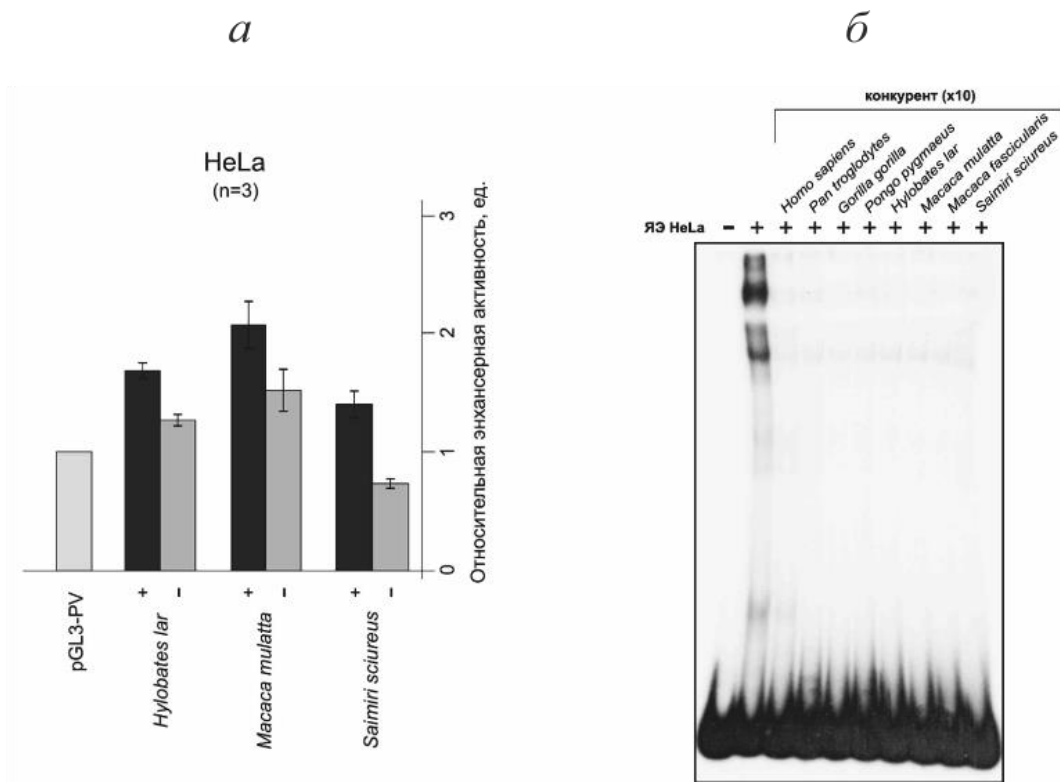
Согласно полученным данным (Рис. 3а) относительная активность энхансера 12 в ряду приматов в прямой и обратной ориентации отличаются незначительно, что свидетельствует о функциональной консервативности ортологичных форм энхансера 12 в клетках человека, также видно, что исследуемый энхансер обладает выраженной зависимостью от ориентации. Относительная энхансерная активность в данном эксперименте является незначительной, что может быть обусловлено индивидуальными особенностями использованной линии клеток HeLa, так как при первичном анализе данный энхансер проявлял более высокую активность.





**Рис. 2.** Определение энхансерной активности в клеточной линии HeLa для полного энхансера 12 и его фрагментов в прямой (+) и обратной (-) ориентациях с использованием системы двойной люциферазной детекции («Promega»). Римскими цифрами обозначены консервативные участки.

Известно, что мутации, затрагивающие сайты связывания транскрипционных факторов, могут приводить к нарушению ДНК-белковых взаимодействий и, следовательно, к изменению или полной утрате активности регуляторного элемента. Для проверки этой гипотезы была проверена способность последовательностей приматов взаимодействовать с белками ядерного экстракта клеток HeLa с помощью сдвига электрофоретической подвижности. Добавление немеченых последовательностей генома приматов в реакцию связывания энхансера 12 человека с белками ядерного экстракта HeLa приводит к ослаблению интенсивности зон торможения, что указывает на способность последовательностей приматов взаимодействовать с теми же белками, что и энхансер 12 (Рис. 3б). Таким образом, нуклеотидные замены, делеции и инсерции в исследованных последовательностях приматов не влияют на их взаимодействие с белками *in vitro*. Можно предполагать, что консервативные области энхансера 12 и области с наименьшим числом межвидовых отличий ответственны за связывание с белками ядерного экстракта клеток HeLa и, вероятно, за проявление энхансерной активности исследуемой области генома человека.



**Рис. 3.** Функциональный анализ последовательностей приматов, ортологичных энхансеру 12. (а) энхансерная активность последовательностей гиббона (*Hylobates lar*), макаки резус (*Macaca mulatta*) и саймири (*Saimiri sciureus*) в прямой (+) и обратной (-) ориентациях в клетках HeLa. (б) EMSA анализ способности последовательностей приматов, гомологичных энхансеру 12, конкурировать с энхансером 12 за связывание с белками ядерного экстракта HeLa (ЯЭ HeLa).

Выявление короткого участка, ответственного за функциональную активность энхансера 12, позволит в дальнейшем идентифицировать связывающиеся с ним транскрипционные факторы, а также изучить механизмы регуляции данного энхансера.

### Изучение двунаправленного промотора генов *PSENER* и *U2AF1L4*

В связи с тем, что проанализированный энхансер 12 расположен во втором интроне гена *U2AF1L4*, нами было выдвинуто предположение, согласно которому промотор данного гена активируется энхансером 12. Для проверки этой гипотезы мы провели детальное изучение промоторной области двух близкорасположенных генов *PSENER* и *U2AF1L4*.

### *Точки инициации транскрипции и тканевая специфичность транскрипции генов *PSENER* и *U2AF1L4**

С помощью процедуры 5'-RACE SMART мы определили сайты инициации транскрипции (TSS-transcription start sites) для генов *PSENER* и *U2AF1L4* в четырех клеточных

линиях различного происхождения (HeLa, A549, NGP-127 и Jurkat). Было показано, что транскрипция в обоих направлениях характеризуется диффузным распределением TSS.

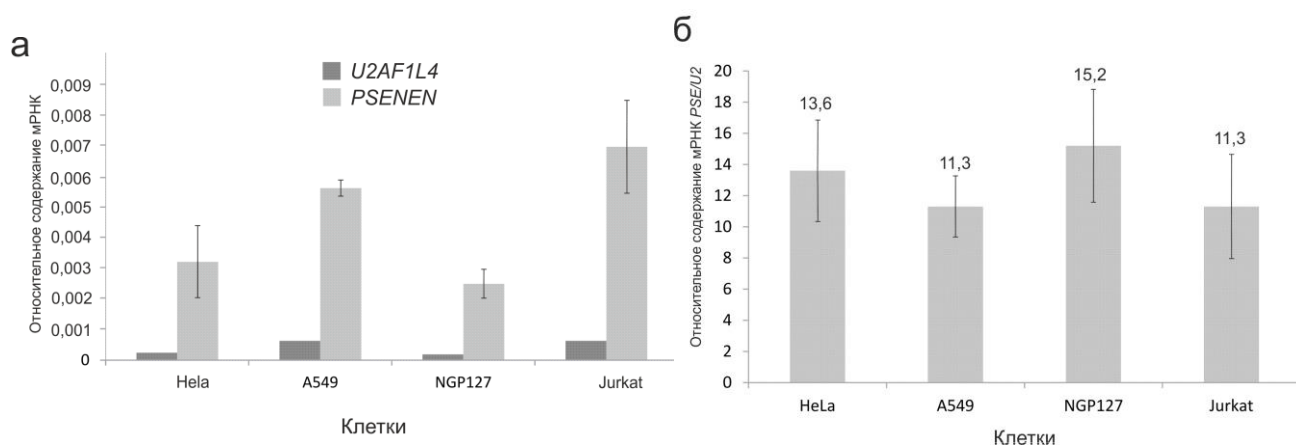
С помощью количественной ОТ-ПЦР мы оценили содержание мРНК *PSENE1* и *U2AF1L4* в тех же четырех клеточных линиях. Содержание мРНК нормализовали к среднему геометрическому содержанию мРНК *GAPDH*, *RPLP0* и генов 18S рРНК. Было показано, что содержание мРНК обоих генов характеризуется умеренной тканевой специфичностью и изменяется в 2-3 раза в различных клеточных линиях (Рис. 4а).

Стоит отметить тот факт, что во всех клетках содержание мРНК гена *PSENE1* от 10 до 15 раз выше, чем мРНК *U2AF1L4*. Однако уровень мРНК фактора *U2AF1L4* может быть несколько заниженным из-за наличия сплайс форм, не обнаруживаемых с помощью использованных в работе праймеров. В то же время, относительное содержание мРНК двух генов во всех клетках очень близко (Рис. 4б), что объясняется их согласованной транскрипционной регуляцией. Полученные данные указывают на то, что транскрипция обоих генов находится под контролем общего двунаправленного промотора.

#### **Анализ активности промоторной области генов *PSENE1* и *U2AF1L4***

Несмотря на то, что область длиной 238 п.н., содержащая минимальный промотор гена *PSENE1*, была идентифицирована ранее, ее промоторная активность в направлении гена *U2AF1L4* не изучалась.

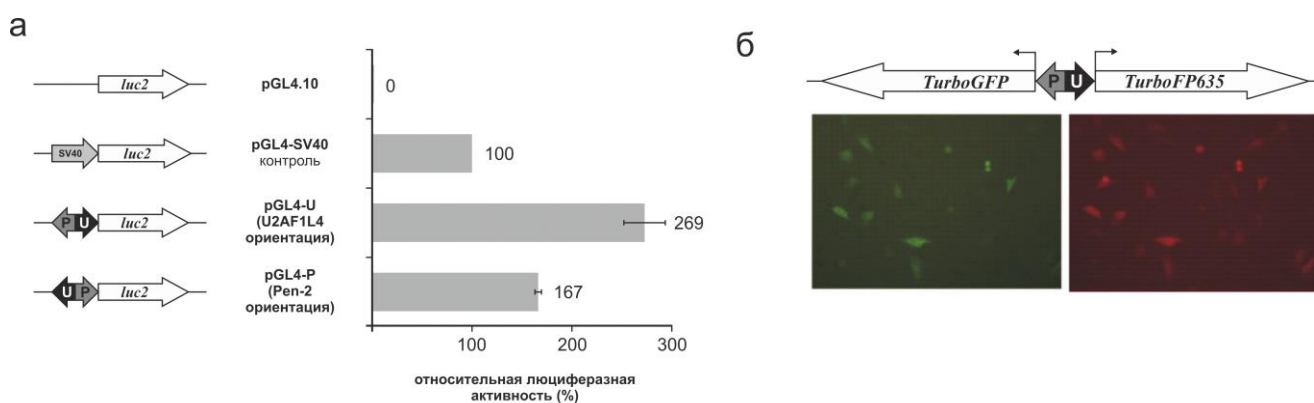
Мы клонировали в двух направлениях фрагмент длиной 269 п.н. (U2P-промотор), захватывающий фрагменты первых экзонов генов *PSENE1* и *U2AF1L4*, в плазмиду рGL4.10 на место промотора с 5'-стороны от гена люциферазы. Таким образом была получена плазида рGL4-P, содержащая фрагмент промотора в *PSENE1* ориентации и рGL4-U, содержащая тот же фрагмент в *U2AF1L4* ориентации (Рис. 5а).



**Рис. 4.** (а) содержание мРНК генов *PSENE1* и *U2AF1L4* в четырех клеточных линиях; (б)-отношение количеств мРНК *PSENE1* и *U2AF1L4* в четырех линиях клеток.

Клонированная промоторная область включала большую часть обнаруженных в четырех клеточных линиях TSS. Полученные конструкции трансфицировали в клетки HeLa и определяли относительную активность люциферазы (Рис. 5а).

Согласно полученным данным, активность промотора в обеих ориентациях в клетках HeLa была в 1,6 раза выше в направлении *U2AF1L4*, чем в направлении гена *PSENEEN*, и в обоих случаях выше активности минимального промотора SV40.



**Рис. 5.** (а) активность промоторной области, расположенной между генами *PSENEEN* и *U2AF1L4*, в клетках линии HeLa, трансфицированных не содержащей промотора плазмидой pGL4.10, содержащей промотор SV40 плазмидой pGL4.10-SV40, плазмидой pGL4.10-U, содержащей промоторную область *PSENEEN/U2AF1L4*, клонированную в направлении гена *U2AF1L4*, и плазмидой pGL4.10-P, содержащей промоторную область *PSENEEN/U2AF1L4*, клонированную в направлении гена *PSENEEN*; (б) одновременная экспрессия флуоресцентных белков TurboGFP и TurboFP635 под управлением промотора *PSENEEN/U2AF1L4* в клетках линии HeLa.

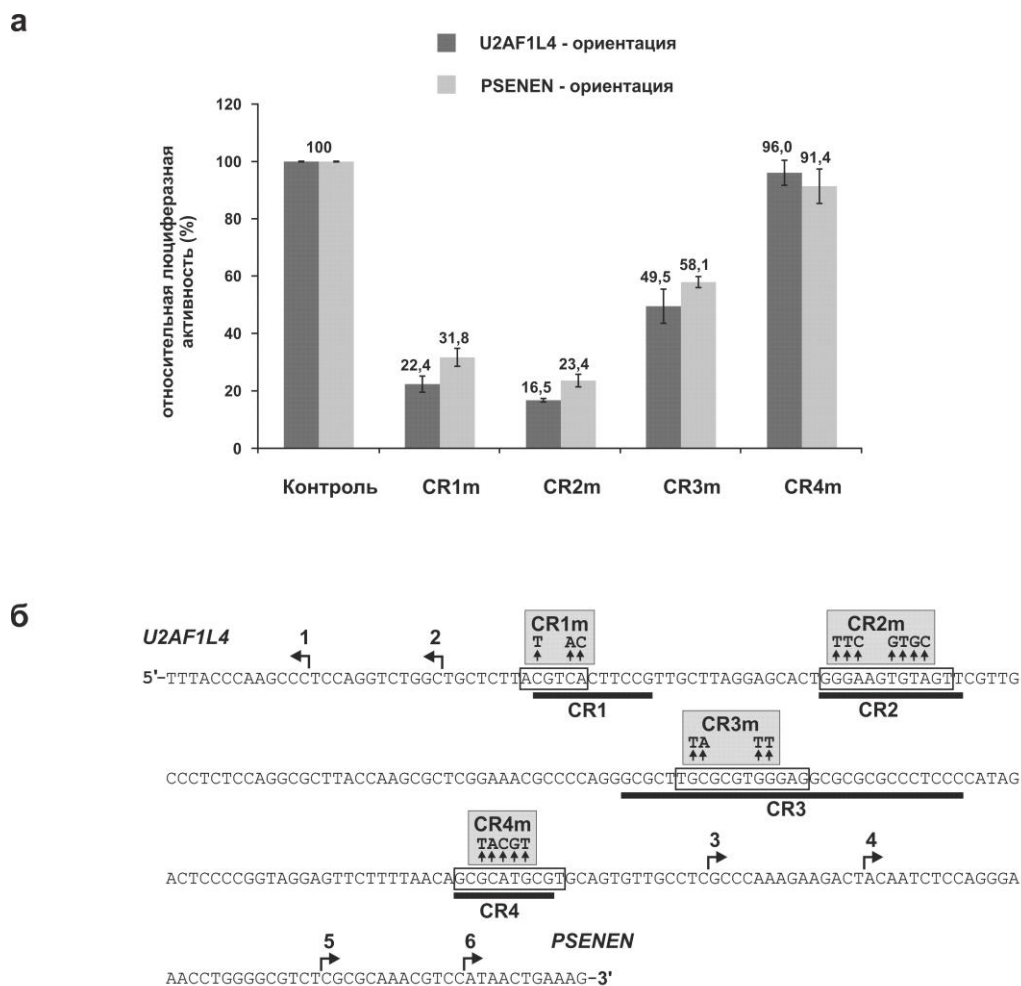
Для того, чтобы продемонстрировать, что U2P-промотор может одновременно управлять экспрессией двух генов, тот же фрагмент длиной 269 п.н. клонировали между генами флуоресцирующих белков TurboGFP и TurboFP635. Полученную конструкцию трансфицировали в клетки HeLa, и проводили визуализацию экспрессии белков с помощью флуоресцентной микроскопии (Рис. 5б). Согласно полученным данным, оба гена экспрессировались одновременно.

Согласованное регулирование транскрипции генов *PSENEEN* и *U2AF1L4* также подтверждается данными GeneCards (<http://www.genecards.org>). Таким образом, согласно полученным данным, промотор между генами *PSENEEN* и *U2AF1L4* является истинным двунаправленным промотором.

### **Функциональный анализ консервативных областей промотора генов *PSENEEN* и *U2AF1L4***

Мы сравнили нуклеотидные последовательности промотора *U2AF1L4/PSENEEN* и семи паралогов предполагаемых промоторов малых и больших субъединиц гена U2AF. Пять из

восемью промоторов содержат последовательности, очень близкие к тандемному повтору  $(TGCGCG)_2$  с варибельной второй частью. Промотор гена *U2AF1* содержит три такие последовательности. Известно, что с данной последовательностью связываются характерные для двунаправленных промоторов транскрипционные факторы NRF1 и GABP.



**Рис. 6.** (а) активность мутантных форм промотора *PSENE1/U2AF1L4* в клетках HeLa. Активность промоторной области дикого типа (контроль) принята за 100%. (б) промоторная область генов *PSENE1/U2AF1L4*, в которой стрелками показаны введенные в последовательность мутации внутри консервативных областей (CR1-4).

Анализ последовательности промотора *PSENE1/U2AF1L4* позволил выявить четыре консервативные области (CR1-4), связывающих, по данным ChIP-seq проекта ENCODE, ряд транскрипционных факторов, характерных для двунаправленных промоторов. Влияние данных областей на транскрипцию в направлении *U2AF1L4* ранее не было изучено.

Для того чтобы выяснить роль консервативных последовательностей в активности промотора *PSENE1/U2AF1L4*, мы ввели мутации в предполагаемые сайты связывания факторов транскрипции (Рис. 6б).

Мутантные промоторы были клонированы в вектор pGL4.10. Плазмиды, содержащие вставку в ориентации *PSENER* или *U2AF1L4*, были приготовлены для всех мутантов. Полученные конструкции трансфицировали в клетки HeLa. Плазмиды, содержащие промотор дикого типа *PSENER/U2AF1L4*, использовали в качестве контроля (Рис. 6а). Согласно полученным данным, мутации в области CR1 или CR2 приводят к снижению активности промотора *PSENER/U2AF1L4* в обеих ориентациях на 70%, а в CR3 на 40-50%. Мутация в области CR4 не оказывает существенного влияния на активность промотора *PSENER/U2AF1L4*. Полученные результаты подтверждают гипотезу о том, что транскрипция генов *PSENER* и *U2AF1L4* управляется двунаправленным промотором и общими транскрипционными регуляторными элементами для обоих направлений.

### **Оценка энхансер-блокирующей активности инсуляторов с помощью транзientной трансфекции**

В рамках проекта по созданию функциональной карты валидированных регуляторных элементов в локусе *FXYD5-COX7A1* хромосомы 19 человека длиной 1 млн п.н. мы провели поиск новых, а также анализ ранее идентифицированных *cis*-регуляторных элементов.

#### ***Выбор фрагментов ДНК***

Основываясь на данных иммунопреципитации хроматина с последующим массированным секвенированием (ChIP-seq) проекта ENCODE, мы выбрали шесть характерных фрагментов ДНК, обладающих свойствами потенциальных энхансеров, и пять фрагментов со свойствами потенциальных инсуляторов. Отбор основывался на способности фрагментов связывать факторы транскрипции Р300 и CTCF, их сверхчувствительности к расщеплению ДНК-азой I в составе хроматина, а также на состоянии их хроматина.

Кроме этого, были отобраны четыре фрагмента ДНК, для которых имелось экспериментальное подтверждение их энхансер-блокирующей активности методом позитивно-негативной селекции.

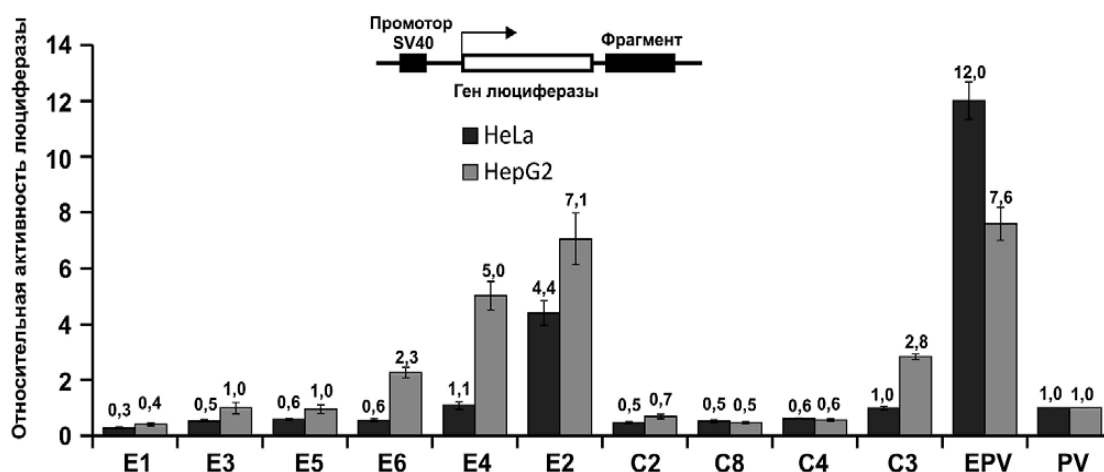
#### ***Энхансерная активность геномных фрагментов***

Для проверки энхансерной и сайленсерной активностей, выбранные фрагменты ДНК были клонированы в плазмиду pGL4PV на место энхансера, с 3'-стороны от репортерного гена люциферазы (Рис. 7). Полученные конструкции трансфицировали в клетки HeLa и HepG2.

Из шести последовательностей со свойствами потенциальных энхансеров (Е1-Е6, см. Рис. 7) способность усиливать активность промотора SV40 в использованной системе в клетках HepG2 проявили три, причем только один из них (Е2) обладал активностью и в клетках HeLa. Из четырех последовательностей, отобранных по связыванию фактора транскрипции CTCF

(потенциальных инсуляторов), один (С3) проявил относительно невысокую энхансерную активность в клетках HepG2, остальные фрагменты, как и ожидалось, не проявили энхансерной активности. Примечательно, что фрагмент С3, кроме связывания CTCF, обладал также свойствами потенциального энхансера.

Таким образом, из семи потенциальных энхансеров в использованной системе ожидаемую активность проявили четыре (~60%), что свидетельствует о существенной, хотя и не полной корреляции между их свойствами, использованными для отбора, и их энхансерной активностью в использованной системе.

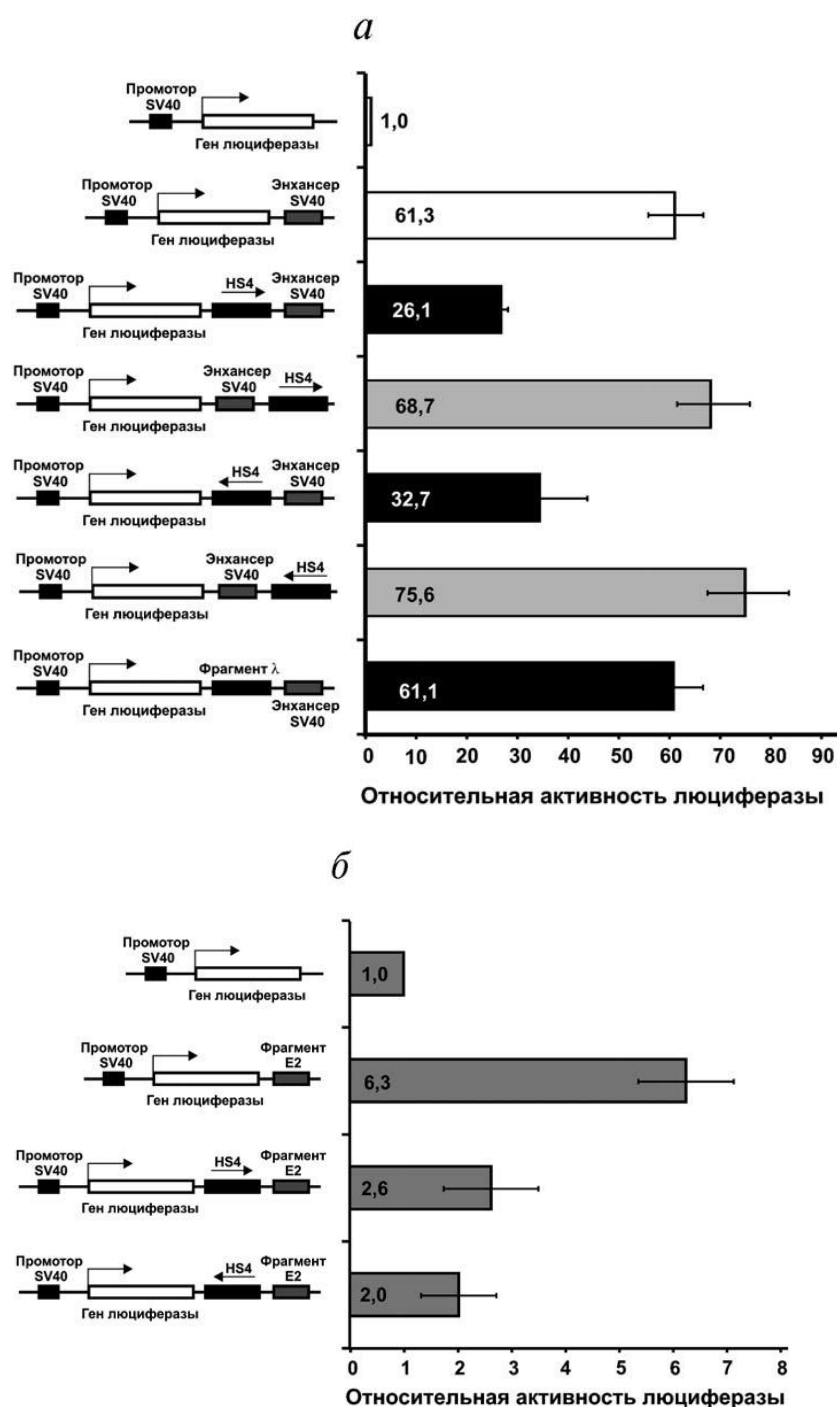


**Рис. 7.** Энхансерная активность фрагментов ДНК из генома человека в клетках HeLa и HepG2. Фрагменты были отобраны в качестве потенциальных энхансеров (E1–E6) или потенциальных инсуляторов (C2–C8). Фрагменты ДНК были клонированы в плазмиду pGL4PV, как показано в верхней части рисунка. Контрольные плазмиды: PV – pGL4PV (содержит промотор SV40 и не содержит энхансера); EPV – pGL4EPV (содержит промотор и энхансер SV40). Активность люциферазы после трансфекции клеток плазмидой pGL4PV принята за единицу

### *Энхансер-блокирующая активность инсулятора cHS4 из бета-глобинового локуса кур*

Для того, чтобы подтвердить пригодность системы для анализа энхансер-блокирующей активности инсуляторов, мы проверили активность хорошо охарактеризованного инсулятора cHS4 из бета-глобинового локуса кур. Для этого полноразмерная последовательность инсулятора cHS4 длиной 1,2 тыс. п.н. была клонирована в обеих ориентациях между энхансером и промотором SV40, а также между промотором SV40 и энхансером E2, наиболее сильным идентифицированным нами энхансером (см. Рис. 7). Параллельно были приготовлены контрольные конструкции, в которых для проверки сайленсерной активности cHS4 был клонирован вне пары промотор-энхансер и конструкции, не содержащие инсулятора.

Конструкции были линейаризованы и трансфицированы в клетки HeLa, после чего определена активность люциферазы (Рис. 8).



**Рис. 8.** Энхансер-блокирующая активность инсулятора HS4 из бета-глобинового локуса кур в клетках HeLa по отношению пары промотор–энхансер SV40 (а) и пары промотор SV40–энхансер E2 (б). За единицу принята активность люциферазы после трансфекции клеток плазмидами, содержащими промотор SV40 и соответствующий энхансер.

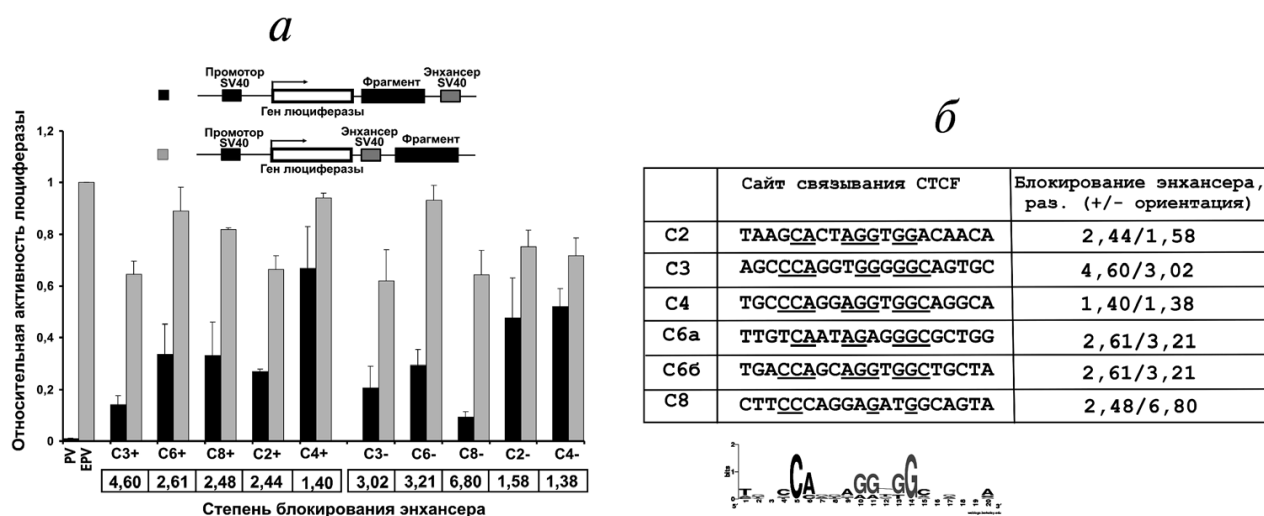
Как видно из Рис. 8а, инсулятор cHS4 способен понижать активность пары промотор–энхансер SV40 в 2-3 раза, причем, только в том случае, когда он расположен между энхансером и промотором. Эффект подавления активности отсутствует при клонировании между энхансером и промотором контрольного фрагмента ДНК фага лямбда. Аналогичный результат



получен для пары промотор SV40 – энхансер E2. В обоих случаях энхансер-блокирующая активность слабо зависела от ориентации последовательности инсулятора cHS4. Таким образом было показано, что полученный набор конструкций пригоден для проверки и анализа энхансер-блокирующих инсуляторных элементов.

### Проверка энхансер-блокирующей активности геномных фрагментов

Для проверки способности отобранных геномных фрагментов ДНК блокировать активацию промотора близлежащим энхансером, последовательности потенциальных инсуляторов были клонированы между промотором и энхансером вируса SV40 в вектор pGL4. Контрольные конструкции, содержащие те же фрагменты на месте энхансера, либо вне пары энхансер-промотор, позволяли оценить энхансерную или сайленсерную активности фрагментов, наличие которых могло бы помешать определению энхансер-блокирующей активности (см. схему на Рис. 9). Для того, чтобы избежать взаимодействия между энхансером и промотором в кольцевой форме плазмиды, полученные конструкции линейаризовали и трансфицировали в клетки млекопитающих, после чего измеряли активность репортерного гена люциферазы с использованием двойной люциферазной системы.



**Рис. 9.** Энхансер-блокирующая активность фрагментов ДНК из генома человека в клетках HeLa. (а) Проверка энхансер-блокирующей активности (темные столбцы) и сайленсерной активности (светлые столбцы) в обеих ориентациях. «+» – ориентация фрагмента совпадает с ориентацией в геноме, «-» – ориентация противоположна ориентации в геноме. Активность люциферазы после трансфекции клеток плазмидой pGL4EPV, содержащей энхансер и промотор SV40, принята за единицу; (б) выравнивание участков связывания белка CTCF из фрагментов геномной ДНК человека. Во втором столбце приведена степень блокирования энхансера соответствующих фрагментов в обеих ориентациях относительно промотора. Подчеркнуты нуклеотиды, совпадающие с консенсусной последовательностью сайта связывания CTCF.

Из пяти геномных фрагментов, отобранных по свойствам, соответствующим инсуляторам, четыре проявили явную способность ослаблять действие энхансера SV40 на промотор SV40 в клетках HeLa только в том случае, когда последовательности были помещены между этими элементами (Рис. 9). Возможным исключением является последовательность С4, инсуляторные свойства которой почти не регистрируются, несмотря на сильное связывание CTCF и состояние хроматина, соответствующее инсулятору.

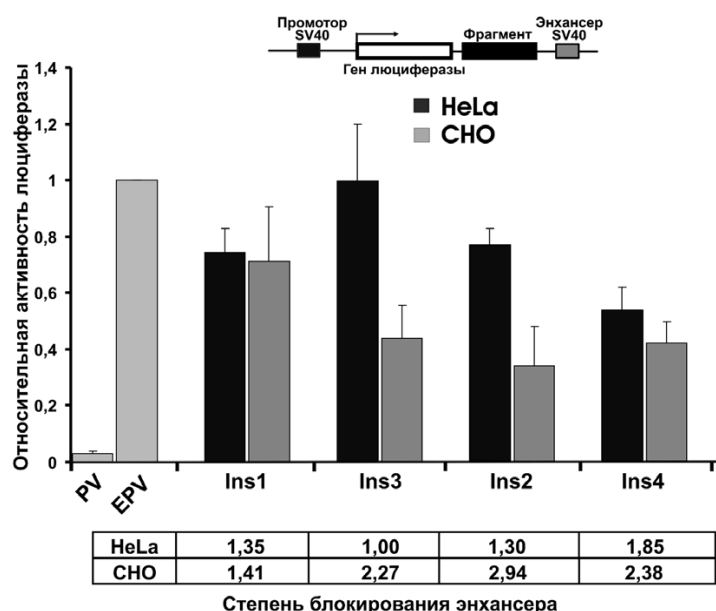
Согласно полученным данным степень ослабления действия энхансера на промотор с учетом поправки на сайленсерную активность колеблется от ~2 (фрагмент С2) до более чем 7 (фрагмент С8) раз. При этом, среди фрагментов присутствуют как те, активность которых сильно зависит от ориентации относительно промотора (С8), так и те, активность которых практически не зависит от ориентации (С3, С6).

Интересно, что фрагмент С3, не относящийся к инсуляторам по структурным особенностям хроматина, но способный связывать CTCF с высокой эффективностью, проявил высокую энхансер-блокирующую активность, тогда как фрагмент С4 с характеристиками инсулятора и также высоким связыванием CTCF, такой активности не проявил.

#### ***Энхансер-блокирующая активность фрагментов, отобранных при помощи позитивно-негативной селекции***

Четыре ранее отобранные с использованием позитивно-негативной селекции последовательности инсуляторов (Ins 1-4) (Акопов et al. 2006) были клонированы в плазмиду pGL4EPV2 между энхансером и промотором SV40. Поскольку отбор потенциальных энхансер-блокирующих фрагментов исходно проводился в клетках CHO, полученные конструкции были линейаризованы и трансфицированы в клетки CHO, и, для сравнения, в клетки HeLa.

Согласно полученным данным, три из четырех (75%) отобранных фрагментов заметно (в 2-3 раза) ослабляли активность пары промотор-энхансер SV40, при этом их активность не зависела от ориентации относительно промотора (Рис. 10). С другой стороны, энхансер-блокирующая активность двух элементов (Ins2 и Ins3) наблюдалась только в клетках CHO и практически отсутствовала в клетках HeLa, а активность Ins4 была одинаковой в этих типах клеток. Интересно также, что энхансер-блокирующая активность элементов Ins2 и Ins3, связывающих фактор транскрипции CTCF, зависит от типа клеток, тогда как активность Ins4, для которой связывания CTCF не наблюдалось, не зависит. Для подтверждения такой корреляции, однако, требуются дальнейшие исследования.



**Рис. 10.** Энхансер-блокирующая активность потенциальных инсуляторов, отобранных из генома человека методом позитивно-негативной селекции, по отношению к паре промотор-энхансер SV40. Приведены данные для клеточных линий HeLa и CHO. За единицу принята активность люциферазы после трансфекции клеток плазмидами, содержащими промотор SV40 и соответствующий энхансер.

### Активация промотора энхансером в *транс* положении

В конце 1980-х годов было показано, что при транзитной трансфекции энхансер может активировать промотор, расположенный на другой плазмиде, когда эти элементы удерживаются вместе с помощью нековалентных связей (Dunaway and Droge 1989; Muller-Strom et al., 1989). В дальнейшем было показано, что эффект активации в *транс*-положении может наблюдаться при совместной трансфекции плазмиды, содержащей энхансер и плазмиды, содержащей промотор и репортерный ген, то есть при проникновении в клетку нескольких копий плазмид, содержащих энхансер и промотор. В таком эксперименте сильный энхансер цитомегаловируса активировал промотор цитомегаловируса в *транс*-положении (то есть, энхансер расположен на отдельной плазмиде котрансфицируемой с геной конструкцией, содержащей промотор и репортерный ген) (D'Aiuto et al., 2006).

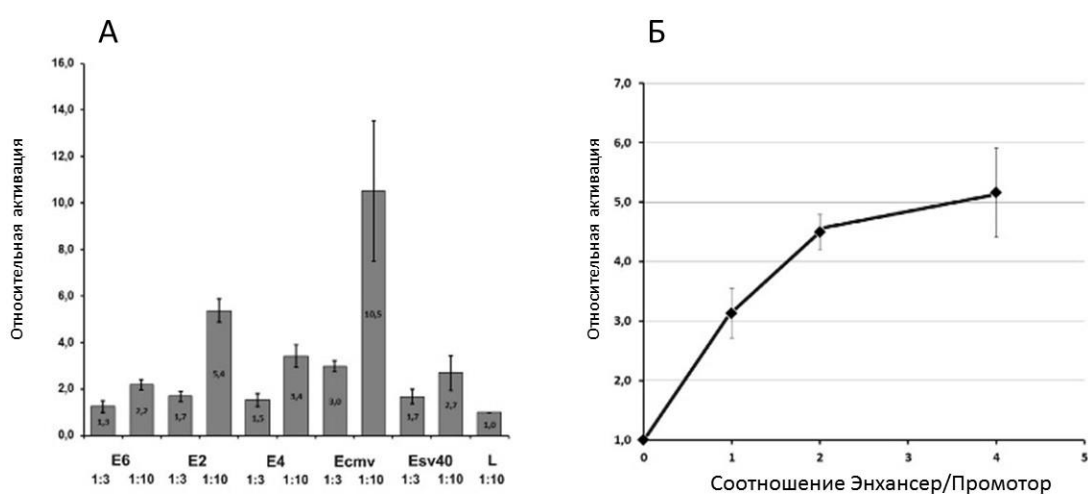
В данной работе мы исследовали промоторную и клеточную специфичность активации промотора энхансером в *транс*-положении, а также механизм этого взаимодействия.

### Активация промотора энхансером в *транс*-положении

Для проверки эффекта *транс*-активации нами были выбраны проверенные ранее геномные энхансеры E2, E4 и E6, а также хорошо изученные вирусные энхансеры цитомегаловируса (Cmv) и SV40 (Esv40). В качестве промоторов были выбраны двунаправленный промотор U2P генов *PSENEN/U2AF1L4*, минимальный промотор

цитомегаловируса, а также промотор гена тимидинкиназы вируса простого герпеса человека. Промоторы U2P и CMV были клонированы в вектор pGL4, в результате чего были получены плазмиды Pu2p и Pcmv. Промотор HSV-*tk* изучали в составе нормировочной плазмиды pRL-TK.

Чтобы нивелировать эффект маскирования *транс*-активации, мы не проводили нормирования люциферазной активности по референсной плазмиде pRL-TK. Вместо этого мы предположили, что эффективность трансфекции в различных экспериментах остается постоянной. Это предположение подтверждается относительно небольшой стандартной ошибкой среднего, полученной при измерениях относительной люциферазной активности (Рис. 11 и 12).



**Рис. 11.** Активация промотора энхансером в *транс*-положении. Активность энхансера была нормирована, на активность отрицательного контроля (L), содержащего фрагмент ДНК фага лямбда. (А) Активация промотора гена тимидинкиназы (Ptk) pRL-TK плазмиды в клетках HepG2 различными котрансфицируемыми энхансерами (описание в тексте). L - отрицательный контроль, содержащий фрагмент ДНК фага лямбда. 1:3, 1:10 - мольное соотношение промотор:энхансер. (Б) Влияние соотношения энхансер:промотор на активацию промотора Pcmv энхансером Ecmv при совместной трансфекции в клетки HeLa.

Результаты экспериментов по совместной трансфекции клеток HepG2 с использованием HSV-*tk* промотора (Ptk) и испытуемых энхансеров показаны на рисунке 11а. Как видно, наиболее выраженная активация (более 10 раз), наблюдается в случае совместной трансфекции конструкциями, содержащими энхансер CMV и промотор CMV при соотношении энхансер:промотор равном 10:1. Тем не менее, другие энхансеры, в том числе вирусный энхансер SV40 и три энхансера из генома человека, также способны усиливать транскрипцию промотора Ptk в клетках HepG2 в 2-5 раз. Степень *транс*-активации зависит от соотношения энхансер-промотор (Рис. 11б). В клетках HeLa активация выходит на плато при соотношении энхансер:промотор более 4.

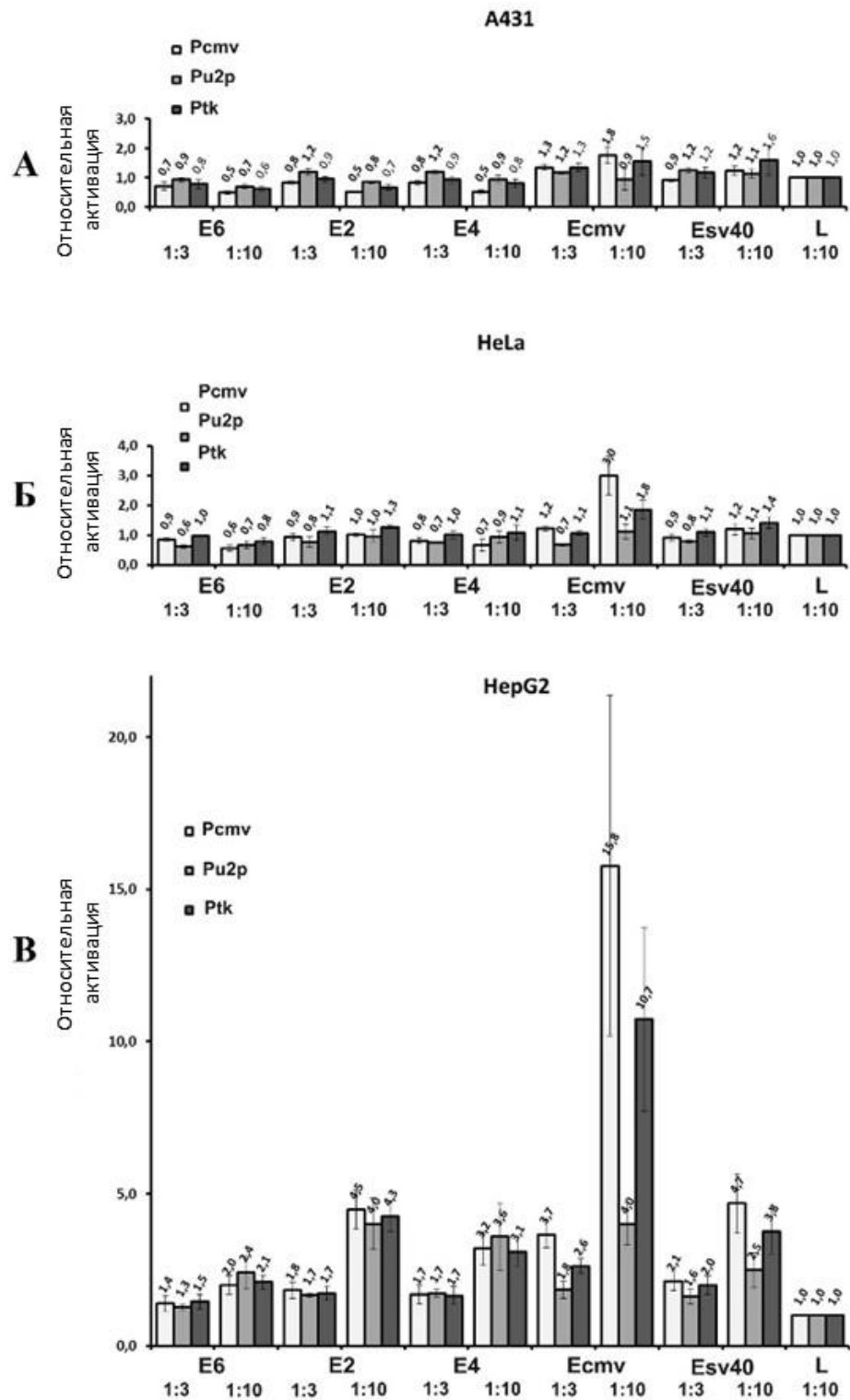
### ***Промоторная и клеточная специфичность***

Эффект *транс*-активации промотора энхансером при транзientной трансфекции продемонстрировал значительную степень специфичности по отношению к промотору в клетках HepG2 (Рис. 12). Наибольшую способность к *транс*-активации обоих вирусных промоторов (Pcmv и Ptk) проявили вирусные энхансеры Ecmv и Esv40 (рис. 12в). В то же время, клеточные энхансеры E2 и E4 имели значительно более низкую промоторную специфичность. Эта особенность энхансеров также была продемонстрирована и в клетках HeLa (Рис. 12б).

Согласно полученным данным, активность проанализированных энхансеров зависит от типа клеток (рис. 12). Так, максимальная активация наблюдалась в клетках гепатоцеллюлярной карциномы (HepG2), менее значительный эффект был обнаружен в клетках аденокарциномы шейки матки (HeLa) и практически не обнаруживался в клетках эпидермальной карциномы (A431).

Клеточную специфичность эффекта *транс*-активации можно объяснить тем, что в стандартных условиях липофекции количество экзогенных энхансеров и промоторов, которые попадают в клетку, значительно превышает количество клеточных промоторов или энхансеров отдельно взятого типа. Такое количество экзогенных регуляторных элементов может связывать транскрипционные факторы, необходимые для активации энхансера или промотора и, таким образом, нарушать клеточные регуляторные сети.

Представленные данные указывают на то, что действие энхансеров (и, возможно, промоторов и других регуляторных элементов) в составе трансфицируемых конструкций на транскрипцию генов клетки-хозяина должно учитываться при интерпретации экспериментальных данных, полученных с использованием транзientных трансфекций.



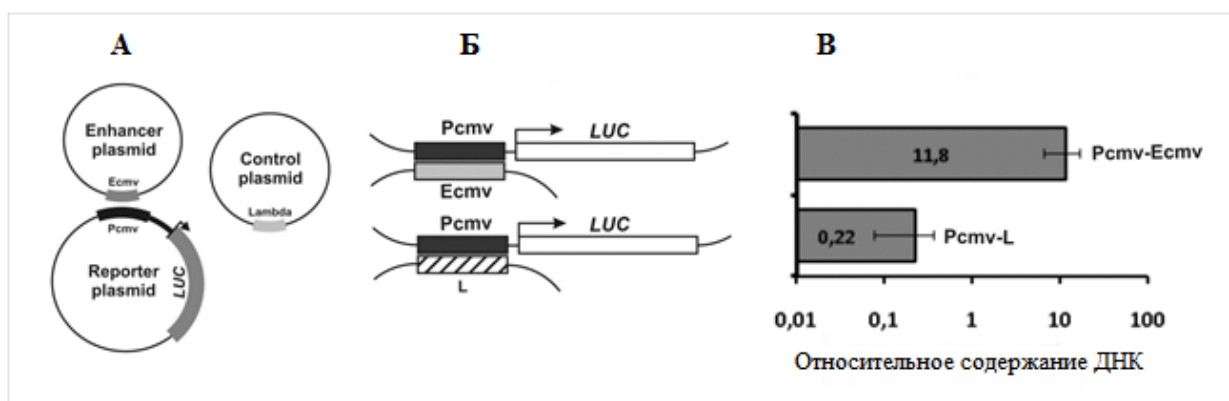
**Рис. 12.** Специфичность активации энхансером промотора в *транс*-положении в зависимости от типа клеток и промотора. L - отрицательный контроль, содержащий фрагмент ДНК фага лямбда. 1:3, 1:10 - мольное соотношение промотор: энхансер.

### **Взаимодействие промотора с энхансером**

Модель дистальных (геномных) взаимодействий энхансера с промотором при котором происходит физическое сближение и контакт между энхансером и промотором, является на данный момент наиболее распространенной.

Для того, чтобы установить, что при активации энхансером промотора в *транс*-положении регуляторные элементы, расположенные в различных плаزمиды, сближаются, мы применили метод Chromosome Conformation Capture (3C), совмещенный с ПЦР в реальном времени.

Для этого была проведена котрансфекция клеток HepG2 смесью, содержащей эквимольные количества трех плазмид: плазмиды, содержащей ген люциферазы светлячка под контролем минимального промотора CMV (Pcmv), плазмиды, содержащей энхансер CMV (Ecmv) и плазмиды (L), содержащей аналогичный энхансеру по длине нейтральный фрагмент ДНК фага лямбда (Рис. 13а). После сшивки, очистки и лигирования ДНК, частоту событий лигирования последовательностей промотор-энхансер и промотор-контрольный фрагмент определяли с помощью количественной ПЦР с использованием праймеров, подобранных для детектирования соответствующих продуктов лигирования. На Рис. 13б схематически представлены предполагаемые комплексы, обнаруживаемые с помощью соответствующих пар праймеров. Рис. 13в представляет относительное содержание продуктов лигирования Pcmv-Ecmv и Pcmv-L, показанных как среднее из двух независимых трансфекций с последующим 3С.



**Рис. 13.** Сближение регуляторных элементов при активации промотора энхансером выявляемое с использованием метода фиксации конформации хромосом (3С) в клетках HepG2. (А) - схематическое представление котрансфицируемых плазмид; (Б) - предполагаемые комплексы, обнаруживаемые с использованием соответствующих пар праймеров; (В) - относительное содержание продуктов лигирования. Pcmv, Ecmv - CMV промотор и энхансер, соответственно. L - отрицательный контроль, содержащий фрагмент ДНК фага лямбда. LUC - репортерный ген люциферазы светлячка. Относительное содержание ДНК отражает частоту событий лигирования соответствующих регуляторных элементов.

Согласно полученным данным (Рис. 13) содержание ДНК (и соответствующая частота лигирования) Pctv и Ectv была в 50 раз выше, чем у Pctv и контрольной последовательности фага лямбда. Эти данные показывают, что промотор и энхансер CMV специфично сближаются при котрансфекции в клетки HepG2, тогда как сближения промотора и контрольной последовательности в тех же условиях не происходит.

Мы полагаем, что метод изучения энхансерной активности в *транс*-положении с использованием совместной трансфекции и репортерного гена может быть удобной моделью для изучения многих важных свойств дистальных взаимодействий энхансера с промотором, включающих клеточную специфичность, специфичность промотора/энхансера, роли различных транскрипционных факторов и влияние структуры хроматина и мутаций в регуляторных элементах на эффект активации.

## **Выводы**

1. Проведен функциональный анализ энхансера, расположенного в локусе *FXYD5-COX7A1* хромосомы 19 человека и гомологичных ему последовательностей приматов. Определена консервативная область, отвечающая за энхансерную активность. При помощи метода EMSA показана функциональная консервативность гомологичных энхансеру последовательностей приматов при связывании с белками.

2. Определены точки начала транскрипции мРНК генов *PSENE1* и *U2AF1L4*. Продемонстрирована согласованная экспрессия данных генов в нескольких клеточных линиях. Определены консервативные области и показана их функциональная роль в работе промотора. Таким образом показано, что область между генами *PSENE1* и *U2AF1L4* на хромосоме 19 человека является истинным двунаправленным промотором.

3. Проведен функциональный анализ энхансерной, сайленсерной и энхансер-блокирующей активности последовательностей, расположенных в локусе *FXYD5-COX7A1* хромосомы 19 человека.

4. Показано, что активация промотора энхансером в *транс*-положении возможна не только для вирусных энхансеров, но и для геномных регуляторных элементов. Активность энхансера в *транс*-положении зависит от промотора и типа клеток и, как было показано с помощью 3С анализа, включает в себя физическую сближенность промотора с энхансером.



## **Основные результаты диссертации изложены в следующих работах**

### **Публикации в научных журналах**

*Дидыч Д.А., Смирнов Н.А., Котова Е.С., Акопов С.Б., Николаев Л.Г., Свердлов Е.Д.* «Функциональная диссекция энхансерного элемента, расположенного во втором интроне гена *U2AF1L4* человека» *Биохимия*, 2011. том. 76, вып. 8, стр. 951-957

*Смирнов Н.А., Дидыч Д.А., Акопов С.Б., Николаев Л.Г., Свердлов Е.Д.* «Оценка энхансер-блокирующей активности инсуляторов при помощи транзientной трансфекции». *Биохимия*, 2013, том 78, вып. 8, стр. 1141 – 1150

*Didych D.A., Shamsutdinov M.F., Smirnov N.A., Akopov S.B., Monastyrskaya G.S., Uspenskaya N.Y., Nikolaev L.G., Sverdlov E.D.* «**Human PSENE1 and U2AF1L4 genes are concertedly regulated by a genuine bidirectional promoter**» *Gene* 515 (2013) 34–41

*Smirnov N.A., Akopov S.B., Didych D.A., Nikolaev L.G.* «**In trans promoter activation by enhancers in transient transfection**» *Gene* 603 (2017) 15–20

### **Материалы российских и международных конференций**

*Смирнов Н.А.* «Анализ энхансер-подобного элемента, расположенного во втором интроне гена *U2AF1L4*.» *VII международная научная конференция студентов, аспирантов и молодых ученых «Ломоносов 2010»*, 2010, Москва, Россия. стр. 209.

*Смирнов Н.А., Дидыч Д.А., Акопов С.Б., Николаев Л.Г., Свердлов Е.Д.* «Оценка уровня активности энхансер-блокирующих последовательностей в системе транзientной трансфекции» *XXIV зимняя молодежная научная школа «Перспективные направления физико-химической биологии и биотехнологии»*, 2012, ИБХ РАН, Москва Россия.

*Smirnov N.A., Didych D.A., Kotova E.S., Akopov S.B., Nikolaev L.G., Sverdlov E.D.* «**Functional dissection of enhancer-like element located within second intron of *U2AF1L4* gene**» *The 4th international IMBG conference for young scientists «Molecular biology: advances and perspectives»*, 2011, Kyiv, Ukraine.

*Смирнов Н.А., Дидыч Д.А.* «Селекция и анализ инсуляторов в локусе 19q13.12 хромосомы 19 человека.» *V международная школа молодых ученых по молекулярной генетике «Непостоянство генома»*, 2012 Москва-Звенигород, Россия. стр. 61.

*Смирнов Н.А., Дидыч Д.А., Акопов С.Б., Николаев Л.Г., Свердлов Е.Д.* «Селекция и анализ энхансер-блокирующих элементов в протяженных локусах генома человека» *международная научно-практическая конференция «Постгеномные методы анализа в биологии, лабораторной и клинической медицине III»*, 2012, Казань, Россия. стр. 94.

*Смирнов Н.А., Дидыч Д.А., Котова Е.С., Акопов С.Б., Николаев Л.Г., Свердлов Е.Д.* «Функциональный анализ энхансер-подобного элемента, расположенного во втором

**интроне гена U2AF1L4»** *Научная конференция по биоорганической химии и биотехнологии «X чтения памяти академика Юрия Анатольевича Овчинникова», 2011 Москва-Пушино, Россия. стр. 60.*

**Дидыч Д.А., Смирнов Н.А., Акопов С.Б., Николаев Л.Г., Свердлов Е.Д. «Участок ДНК, расположенный между генами U2AF1L4 и PSENEN (PEN2) проявляет свойства двунаправленного промотора»** *Научная конференция по биоорганической химии и биотехнологии «X чтения памяти академика Юрия Анатольевича Овчинникова», 2011 Москва-Пушино, Россия. стр. 20.*