

Федеральное государственное бюджетное учреждение науки
«Национальный научный центр морской биологии им. А.В.Жирмунского»
Дальневосточного отделения Российской академии наук

На правах рукописи

КАМЕНСКАЯ ДАРЬЯ НИКОЛАЕВНА
СРАВНИТЕЛЬНЫЙ АНАЛИЗ РЕГУЛЯТОРНЫХ
ПОСЛЕДОВАТЕЛЬНОСТЕЙ ПАРАЛОГИЧНЫХ ГЕНОВ ГОРМОНА РОСТА У
ЛОСОСЕВЫХ РЫБ

1.5.7. Генетика

Диссертация на соискание ученой степени
кандидата биологических наук

Научный руководитель:
Доктор биологических наук, профессор
Брыков Владимир Алексеевич

Владивосток – 2023

ОГЛАВЛЕНИЕ

| | |
|-----------------------------------------------------------------------------------------------------|-----------|
| ОГЛАВЛЕНИЕ | 2 |
| СПИСОК СОКРАЩЕНИЙ | 4 |
| ВВЕДЕНИЕ | 6 |
| ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ..... | 14 |
| 1.1. Структура гена | 14 |
| 1.1.1. Общие представления о структуре гена | 14 |
| 1.1.2. Организация генов у эукариот | 15 |
| 1.1.3. Структурная и регуляторная части гена. Структура транскриптов | 17 |
| 1.2. Общие представления о структуре генов гормона роста у позвоночных животных | 20 |
| 1.3. Структура генов гормона роста рыб | 27 |
| 1.4. Эволюция генов гормона роста у позвоночных животных | 35 |
| 1.5. Структура регуляторных участков гена гормона роста позвоночных | 43 |
| 1.5.1. Структура регуляторных участков гена гормона роста у рыб | 52 |
| ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ | 59 |
| 2.1. Объекты исследования | 59 |
| 2.2. Выделение ДНК | 62 |
| 2.3. Амплификация | 62 |
| 2.4. Молекулярное клонирование..... | 64 |
| 2.4.1. Получение компетентных клеток <i>Escherichia coli</i> для молекулярного клонирования..... | 64 |
| 2.4.2. Лигирование и трансформация | 66 |
| 2.5. Идентификация рекомбинантных клонов | 68 |
| 2.6. Секвенирование..... | 69 |
| 2.7. Сборка и выравнивание последовательностей | 69 |
| 2.8. Анализ данных | 70 |
| ГЛАВА 3. РЕЗУЛЬТАТЫ | 71 |
| 3.1. Характеристика регуляторных участков гена гормона роста у гольцов рода <i>Salvelinus</i> | 71 |

| | | |
|----------------------------------|-------------------------------------------------------------------------------|------------|
| 3.2. | Сравнительный анализ паралогичных генов и их промоторов у лососевых рыб | 74 |
| 3.2.1. | Изменчивость в 5'-цис-прилежащем промоторном участке | 74 |
| 3.2.2. | Изменчивость структурных последовательностей генов | 89 |
| 3.3. | Филогения генов гормона роста | 105 |
| ГЛАВА 4. ОБСУЖДЕНИЕ | | 107 |
| 4.1. | Структура промоторных участков гена гормона роста лососевых | 107 |
| 4.2. | Сравнительный анализ изменчивости генов | 115 |
| 4.2.1. | Изменчивость транскрибируемой части гена | 115 |
| 4.2.2. | Изменчивость экзонов | 117 |
| 4.2.3. | Изменчивость интронов | 119 |
| 4.2.4. | Изменчивость регуляторной области | 123 |
| ЗАКЛЮЧЕНИЕ | | 127 |
| ВЫВОДЫ | | 129 |
| СПИСОК ЛИТЕРАТУРЫ | | 131 |
| ПРИЛОЖЕНИЕ I | | 150 |
| ПРИЛОЖЕНИЕ II | | 151 |
| ПРИЛОЖЕНИЕ III | | 152 |
| ПРИЛОЖЕНИЕ IV | | 153 |

СПИСОК СОКРАЩЕНИЙ

- ГГР – ген гормона роста
- ДНК – дезоксирибонуклеиновая кислота
- кДНК – комплементарная ДНК
- млн.л.н. – миллионов лет назад
- мРНК – матричная рибонуклеиновая кислота
- мяРНК – малые ядерные РНК
- об/мин – оборотов в минуту
- п.н. – пара нуклеотидов
- ПЦР – полимеразная цепная реакция
- РНК – рибонуклеиновая кислота
- рРНК – рибосомная РНК
- т.п.н. – тысяча пар нуклеотидов
- тРНК – транспортная РНК
- цАМФ – циклический аденозинмонофосфат
- ЭДТА – этилендиаминтетрауксусная кислота
- АР-2 – Activating protein 2 – активирующий белок 2
- CAT – Chloramphenicol acetyltransferase – ген хлорамфеникол-ацетилтрансферазы
- CRE – cAMP-response element – элемент ответа на циклический АМФ
- CREB – cAMP response element-binding protein – белок, связывающийся с элементом ответа на циклический АМФ
- ERE – Estrogen response element – эстроген-чувствительный элемента
- GRE – Glucocorticoid response element – элемент ответа на глюкокортикоиды
- hCS – Human chorionic somatomammotrophin gene – ген хорионического соматомаммотропина
- hGH-N – Human growth hormone-normal gene – ген гормона роста человека, экспрессируется в клетках гипофиза
- hGH-V – Human growth hormone-variant gene – альтернативный вариант гена гормона роста человека, экспрессируется в плаценте
- hPL – Human placental lactogen – плацентарный лактоген

HNF-3 α – Hepatocyte nuclear factor 3 α – транскрипционный ядерный фактор гепатоцитов 3 α

HRE – Hormone response element – элементы ответа на гормоны

HS – Hypersensitive sites – Сверхчувствительные к ДНКазе I сайты

ITS – Internal Transcribed Spacer – регион внутреннего транскрибируемого спейсера

LCR – Locus control region – область, контролирующая локус

NRE – Negative regulatory element – отрицательные регуляторные элементы

Pre-GH – Precursor growth hormone gene – предшественник гена гормона роста

Pre-PL – Precursor placental lactogen gene – предшественник плацентарного лактогена

Pit-1 – Pituitary-specific positive transcription factor 1 гипофиз-специфичный транскрипционный фактор

RARE – Retinoic Acid Response Element – элемент ответа на ретиноевую кислоту

TBE – Tris/Borate/EDTA-buffer – трис-боратный ЭДТА – буфер

TBP – TATA-binding protein – белок, связывающийся с ТАТА- боксом

TRE – Thyroid Hormone Response Element – элемент ответа на тиреоидные гормоны

ВВЕДЕНИЕ

Актуальность темы исследования. Дупликации генов являются одним из основных способов возникновения нового генетического материала в процессе эволюции (Ohno, 1970; Lynch, Conery, 2000). Дуплицированные гены появляются в результате неравного кроссинговера, дупликации отдельных хромосом, тандемных дупликаций части хромосом или дупликации всего генома (Zhang, 2003; Magadum et al., 2013). В зависимости от значимости выполняемых функций судьба дуплицированных генов может складываться по-разному. Наиболее частый случай – потеря функции в результате накопления мутаций одним из паралогов. Это приводит к тому, что одна из копий становится псевдогеном (Lynch, Conery, 2000, Balakirev, Ayala, 2003). Другой вариант – это субфункционализация, когда оба гена остаются функциональными, но транскрибируются в разных тканях или разделяют между собой функции, которые до дупликации выполнял один ген. Известны случаи, когда дуплицированный ген приобретает новую, существенно отличающуюся функцию (Walsh, 1995; Wagner, 1998). Также считается, что дуплицированные последовательности играют важную роль в геноме эукариот: защищают от вредных мутаций и способствуют генетической устойчивости организма (Krakauer, Nowak, 1999).

В истории всех организмов предполагается этап полногеномной дупликации (Ohno, 1970; Zhang, 2003). В дальнейшем, вероятно, происходил этап вторичной диплоидизации геномов, при котором значительная часть множественных генов была утрачена, а часть дупликаций сохранилась в виде различных генных семейств. Встречаются виды, у которых диплоидизация не завершилась до сих пор. Одна из таких групп – лососевые рыбы.

Лососевые рыбы сформировались после событий автотетраплоидизации и последующей дивергенции. В отряде лососевых выявляются следы четырех этапов дупликации (4R). Последнее из этих событий произошло около 100 млн лет назад (Berthelot et al., 2014; Задесенец, Рубцов, 2018). А квадριвалентные комбинации хромосом в процессе митоза у некоторых видов можно встретить и сейчас. Таким образом, лососевые являются естественными и относительно недавними полиплоидами. Как следствие, многие гены у представителей этой таксономической группы оказались

множественными, в том числе, и представленный двумя копиями ген гормона роста. По-видимому, оба гена гормона роста существуют на протяжении всего времени дивергенции видов в этой группе (25–100 млн лет) (Allendorf, Thorgaard, 1984), и, вероятно, оба функциональны.

Одним из подходов для понимания функциональной значимости, как самих дублированных генов, так и их различных участков является сравнительный анализ фрагментов ДНК у нескольких видов в таксонах. Консервативность последовательностей у представителей одного таксона дает основание считать такие участки важными для функционирования гена. Высокая скорость накопления нуклеотидных замен свидетельствует либо о низкой функциональной значимости участка, либо о влиянии на него положительного отбора (Kondrashov F., Kondrashov A., 2006).

Степень разработанности темы. Нуклеотидные последовательности генов гормона роста известны для большинства групп позвоночных животных. Последовательности генов гормона роста были получены для представителей различных отрядов млекопитающих (Barta et al., 1981; Gordon et al., 1983; Chen et al., 1989), птиц (Arai, Iigo, 2010) и даже для таких древних позвоночных, как миноги (Moriyama et al., 2006). Такой повышенный интерес к строению гена связан не только с важными функциями, которые выполняет гормон роста в организме, но и с особенностями его эволюции. Исследованию молекулярной эволюции гена гормона роста посвящен целый ряд работ (Miller, Eberhardt, 1983; Walker, 1991; Wallis, 1996; Daza et al., 2009). Показано, что современный ген гормона роста, который у большинства позвоночных состоит из пяти экзонов и четырех интронов, получился в результате дубликации небольшого участка гена-предшественника, а в результате дубликации самого гена гормона роста возникли другие гены семейства: пролактин, соматолактин и плацентарный лактоген (Miller, Eberhardt, 1983). У человека и других высших приматов гены гормона роста вместе с плацентарным лактогеном формируют кластер генов, расположенных друг за другом на одной хромосоме (Perez-Maya et al., 2016). Исследователи связывают появление кластера генов с увеличением скорости эволюции гена гормона роста в несколько раз на этапе дивергенции парнокопытных и приматов (Wallis, 1996; Lioupiis et al., 1999). Кластерная организация генов гормона роста и плацентарных лактогенов присутствует только в

геноме высших приматов (Perez-Maya et al., 2016) у всех остальных млекопитающих имеется один ген гормона роста (Das et al., 1996; Maniou et al., 2004; Wallis O., Wallis M., 2001). Исключение составляют птицы (Yuri et al., 2008) и некоторые виды рыб, у которых ген гормона роста представлен двумя несвязанными копиями (Chen et al., 1989; Ver, Daniel, 1992; Devlin, 1993).

Исследования генов гормона роста у лососевых рыб в основном были сосредоточены на создании генетических конструкций, содержащих кДНК гена гормона роста и трансформацией такими конструкциями для получения модифицированных линий рыб с высокой скоростью роста (Du et al., 1992; Devlin et al., 2004; Leggatt et al., 2012). Кроме того, целый ряд работ посвящен описанию полных нуклеотидных последовательностей генов (Male et al., 1992; Du et al., 1993; Devlin, 1993) и анализу интронных участков (Oakley, Phillips, 1999; Phillips et al., 2004). Однако до недавнего времени особенности дублированных генов гормона роста были мало изучены даже у таких экономически важных видов рыб, как лососевые. В работах Паньковой с соавторами (Панькова и др., 2013; Панькова и др., 2017) был проведен сравнительный анализ дивергенции функционально различных участков генов-паралогов в их структурной части, показано с какой интенсивностью действует отбор на гены, возникшие в результате дубликации и выполняющие сходные функции.

Несмотря на такое количество исследований, направленных на изучение структурной части генов гормона роста, очень мало внимания уделяется регуляторным областям, в частности, промоторным участкам. Несколько публикаций посвящено промоторным областям гена гормона роста радужной форели (*Oncorhynchus mykiss*) (Yamada et al., 1993; Argenton et al., 1996; Bernardini et al., 1999), по другим видам лососевых данных нет. Сравнительный анализ сайтов связывания с транскрипционными факторами и путей активации транскрипции позволит оценить, насколько функционально значимы оба гена-паралога.

Цель и задачи исследования. Цель данной работы – изучение дивергенции в промоторных участках паралогичных генов гормона роста *gh1* и *gh2* у лососевых рыб.

Для достижения цели были поставлены следующие задачи:

1. Определить нуклеотидные последовательности 5'-фланкирующих промоторных участков генов *gh1* и *gh2* у четырех представителей гольцов рода *Salvelinus*: *S. malma*, *S. curilus*, *S. taranetzi*, *S. levanidovi*.
2. Провести сравнительный анализ полученных нуклеотидных последовательностей промотора двух генов *gh1* и *gh2* у рыб рода *Salvelinus* с аналогичными последовательностями других представителей семейства Salmonidae.
3. Выявить вероятные цис-регуляторные элементы в исследуемых генах.
4. Оценить уровень дивергенции транскрибируемой части и промоторных последовательностей генов гормона роста лососевых рыб.

Научная новизна. В настоящей работе впервые получены и охарактеризованы промоторные последовательности паралогичных генов гормона роста *gh1* и *gh2* у четырех видов гольцов рода *Salvelinus*. Проведен сравнительный анализ сайтов связывания с транскрипционными факторами, расположенными в промоторной области среди представителей трех родов семейства Salmonidae: *Salvelinus*, *Salmo* и *Oncorhynchus*. Показано, что гены гормона роста у всех исследованных видов содержат одинаковый набор сайтов связывания с факторами транскрипции, формирующими ядро промотора. В промоторе представлены сайты связывания как с тканеспецифичными транскрипционными факторами, так и с другими лигандами. Кроме того, впервые выполненный анализ нуклеотидного разнообразия кодирующей части генов и их промоторных участков позволил оценить уровень дивергенции разных областей в двух паралогичных генах гормона роста у лососевых рыб.

Теоретическое и практическое значение работы. Полученные данные вносят значительный вклад в понимание молекулярных основ организации и транскрипции генов гормона роста у лососевых рыб. Присутствие в промоторной области сайтов связывания не только с тканеспецифичными, но и с другими транскрипционными факторами, а также с рецепторами гормонов и другими лигандами указывает на альтернативные пути регуляции транскрипции. На основе полученных данных о сайтах связывания с транскрипционными факторами в дальнейших исследованиях можно будет установить какие факторы и при каких условиях участвуют в сборке транскрипционного комплекса; происходит ли взаимодействие между транскрипционными факторами или

транскрипционные факторы действуют независимо; присутствуют ли различия в пространственно-временной локализации экспрессии паралогичных генов и зависит ли их активность от стадии жизненного цикла и влияния факторов окружающей среды. Методы, использованные для оценки изменчивости, могут быть применены при изучении других генов. Полученные данные по промоторной последовательности генов гормона роста дополняют уже имеющиеся данные по кодирующей части гена гормона роста гольцов. Такая нуклеотидная последовательность, состоящая из экзонов, интронов и регуляторных областей, может выступать в качестве полноразмерной генетической конструкции для трансформации других видов рыб. Увеличение числа генов гормона роста в геноме позволит получать трансгенные линии рыб с более высокой скоростью роста в условиях аквакультуры.

Методология и методы диссертационного исследования. В данной работе были применены стандартные молекулярно-генетические методы. Разработку праймеров для амплификации фланкирующих участков выполняли с помощью программ Primer Premier 5 (Lalitha, 2000) и Primer 3 (<https://bioinfo.ut.ee/primer3-0.4.0>). Амплификацию промоторных участков генов *gh* осуществляли с помощью полимеразной цепной реакции (ПЦР). Для разделения участков паралогичных генов гормона роста *gh1* и *gh2* очищенные продукты амплификации подвергали молекулярному клонированию. Разделяли и определяли размер полученных фрагментов после ПЦР методом электрофореза в агарозном геле. Автоматическое секвенирование нуклеотидных последовательностей проводили по методу Сэнгера. Для анализа данных использовали актуальные статистические программы. Нуклеотидное разнообразие и дивергенцию оценивали в программе DnaSP 5.10.01 (Librado, Rozas, 2009). Филогенетическую реконструкцию выполняли Байесовским методом (Bayesian Inference, BI) с помощью программы MrBayes v. 3.2.6 (Ronquist et al., 2012) и методом ближайшего соседства (Neighbor-Joining) в программе MEGA 7.0 (Kumar et al., 2016). Для оценки нуклеотидного состава интронов использовали программу MEGA 7.0 (Kumar et al., 2016) и GC content calculator (<https://jamiemcgowan.ie/bioinf/gc.html>).

Основные положения, выносимые на защиту:

1. Паралогичные гены гормона роста *gh1* и *gh2* кроме структурной части, включающей экзоны и интроны, содержат одинаковые регуляторные цис-действующие элементы.
2. В регуляторных участках паралогичных генов гормона роста *gh1* и *gh2* представлены одинаковые сайты связывания с транскрипционными факторами.
3. Регуляторные элементы CRE и Pit-1, обнаружены как в промотроной области, так и в интронах гена гормона роста.
4. Практически идентичное расположение регуляторных элементов и низкий уровень дивергенции указывают на общее происхождение и сохранение функций обоих генов.

Степень достоверности результатов. Достоверность полученных результатов обеспечена современными молекулярно-генетическими методами исследования и статистической обработкой полученных данных с помощью различного программного обеспечения, которое соответствует целям и задачам, поставленным в работе. Использование молекулярного клонирования и анализ большого числа клонов также обеспечивает достоверность результатов, полученных в ходе эксперимента. Публикация результатов в рецензируемых научных журналах подкрепляют их достоверность. Таблицы и рисунки, представленные в работе, подтверждают интерпретацию результатов, научных положений и выводов.

Личный вклад автора. Экспериментальная часть работы была выполнена автором самостоятельно. Автор освоил и применил различные программы для анализа полученных данных и интерпретации результатов. Автор готовил материалы к публикации и представлял результаты на конференциях.

Апробация результатов работы. Результаты работы были представлены на: ежегодных научных конференциях Национального научного центра морской биологии им. А. В. Жирмунского Дальневосточного отделения Российской академии наук (Владивосток, 2015; 2016; 2018); 8 Международной школе-конференции молодых ученых «Системная биология и биоинформатика» (Новосибирск, 2016); 10 Международной научной конференции по биоинформатике регуляции структуры геномов и системной

биологии BGRS/SB-2016 (Новосибирск, 2016); XI открытой юношеской научно-практической конференции «Будущее сильной России – в высоких технологиях» (Владивосток, 2017); Региональной научно-практической конференции студентов, аспирантов и молодых учёных по естественным наукам (Владивосток, 2017); Международной конференции «Научно-технологические разработки в области изучения и мониторинга морских биологических ресурсов» (Владивосток, 2017); X Международном симпозиуме «Современные достижения в популяционной, эволюционной и экологической генетике – МАРЕЕГ» (Владивосток, 2019), XI Международном симпозиуме «Современные достижения в популяционной, эволюционной и экологической генетике – МАРЕЕГ» (Владивосток, 2022); Всероссийской конференции «Морская биология в 21 веке: систематика, генетика, экология морских организмов» (памяти академика Олега Григорьевича Кусакина) (Владивосток, 2022); IV Международной конференции «Современные проблемы биологической эволюции» (Москва, 2022).

Публикации. По теме диссертации опубликовано 12 работ, в том числе 4 статьи в журналах из списка ВАК.

Объем и структура работы. Основной текст диссертации изложен на 149 страницах, состоит из введения, обзора литературы, материалов и методов, результатов, обсуждения, заключения, выводов, списка литературы. Дополнительно представлены четыре приложения. Работа содержит 11 таблиц и 21 рисунок. Список литературы насчитывает 188 наименований, из них 176 на иностранном языке.

Благодарности. Выражаю искреннюю благодарность своему научному руководителю профессору, д.б.н. Брыкову Владимиру Алексеевичу за руководство, советы и поддержку на протяжении многих лет совместной работы. Особую благодарность выражаю к.б.н. Паньковой Марине Владимировне за помощь в освоении молекулярно-генетических методов, обсуждении полученных результатов и подготовке совместных публикаций. Благодарю к.б.н., н.с. Шарину Светлану Николаевну за предварительное ознакомление с текстом диссертационной работы и важные замечания; благодарю к.б.н. н.с. Бондарь Евгению Игоревну за ценные комментарии к некоторым главам диссертационной работы. Глубокую признательность выражаю д.б.н., ст.н.с.

Олейник Алле Геннадьевне за подробное ознакомление с текстом работы, всестороннюю поддержку, ценные замечания и рекомендации.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Структура гена

1.1.1. Общие представления о структуре гена

В начале XX века сложилось представление о гене, как материальной частице, находящейся в хромосоме, способной к саморепродукции и выступающей в роли минимальной единицы рекомбинации, мутирования и генетической функции. Цепочка сцепленных генов представлялась как нитка корпускул, механически связанных с друг другом. Ген, согласно этим представлениям, считался неделимым с помощью кроссинговера (Portin, 1993). Представления школы Томаса Моргана о структуре гена, строились на том, что ген имеет основные свойства хромосом (способность к редупликации, к закономерному распределению в митозе и мейозе), занимает определенный участок хромосомы, является единицей мутации (изменяется как целое), единицей рекомбинации (кроссинговера никогда не наблюдали в пределах гена), единицей функции (все мутации одного гена нарушают одну и ту же функцию). Однако в работах Александра Сергеевича Серебровского и Николая Петровича Дубинина было показано, что гену характерно такое свойство, как «делимость». В 1949 году Мелвин и Кэтлин Грин описали делимость гена *lz* у дрозофилы (Green M., Green K., 1949). Огромный вклад в понимание структуры и функции гена внесли Джордж Бидл и Эдуард Татум, которые в начале 1940-х годов впервые показали, что мутации ауксотрофности у нейроспоры (*Neurospora crassa*) прерывают цепи метаболизма на конкретных этапах (Beadle, Tatum, 1941). При этом аллельные мутации всегда затрагивают один и тот же этап биосинтеза. На основании своих результатов Бидл и Татум сформировали принцип «один ген – один фермент», означавший, что каждый ген контролирует синтез какого-либо определенного фермента. Позже Чарльз Яновский показал (Yanofsky et al., 1964), что не только ген, но и любая его часть могут быть разделены с помощью кроссинговера и кроссинговер может происходить между любой парой нуклеотидов.

Согласно современным представлениям, ген можно определить, как последовательность нуклеотидов ДНК размером от нескольких сотен до миллиона пар нуклеотидов, которая представляет собой единицу наследственной информации и

занимает определенное положение в геноме или хромосоме и контролирует выполнение определенной функции в организме (Portin, 2002).

1.1.2. Организация генов у эукариот

Организация генов у эукариот отличается от организации генов у прокариот. Прокариотам присущ оперонный принцип организации генов. Гены у бактерий объединяются в кластеры таким образом, что ферменты необходимые для осуществления определенного пути биосинтеза, кодируются генами, сцепленными с друг другом. У эукариот количество ДНК, которое содержится в клеточном ядре, значительно больше, чем у прокариот, и увеличивается в процессе усложнения организмов. Гены у эукариот могут быть повторяющимися, в то же время, подавляющее большинство генов неповторяющиеся. Оперонный тип расположения генов у эукариот отсутствует. Гены, контролирующие даже последовательные биохимические реакции, расположены в различных районах хромосомы и даже в разных хромосомах. Например, у дрозофилы гены, кодирующие ферменты, под контролем которых происходит превращение триптофана в бурый глазной пигмент, разбросаны во множестве участков генома (Льюин, 2011). Встречаются примеры и кластерной организации генов. У человека известно несколько типов гемоглобинов. Каждый из них синтезируется в определенном органе и на определенной стадии развития. В геноме гены гемоглобина расположены двумя кластерами: в хромосоме 16 расположены все α – подобные гены, в то время как все β – подобные гены расположены в хромосоме 11. В каждом кластере имеются псевдогены. Интересно, что гены расположены в том порядке вдоль по хромосоме, в какой очередности они включаются в работу в ходе онтогенеза (Orkin, 1995, Baron, 1996). Однако каких-либо данных, свидетельствующих об их функциональной сцепленности или общем контроле по принципу оперонной организации, не получено. Тот факт, что эти гены функционируют в разных тканях и на разных этапах онтогенеза, скорее свидетельствует о независимом контроле экспрессии этих генов. Гены транспортной РНК, которые у дрозофилы оказались множественными, расположены во многих районах хромосом по 1–2 гена в каждом. Однако в одном из районов 16 генов занимают небольшой участок длиной 50 000 п.н., хотя их транскрипция не контролируется одним промотором.

Организация генов, кодирующих 18S и 28S рРНК, у всех эукариот одинаковая. Последовательности генов 18S и 28S рРНК, лидерной последовательности, а также транскрибируемого и нетранскрибируемого спейсеров составляют единицу длиной около 11 000 п.н., которая повторяется несколько сотен раз. Как правило, число копий варьирует в пределах от 100 до 1000: у дрожжей – 140 повторяющихся единиц, у дрозофил – 200–250, у человека – 1250. Лидерная последовательность расположена перед геном 18S и с нее начинается транскрипция. Каждая единица повтора генов рРНК считывается РНК полимеразой I от лидерной последовательности до конца гена 28S рРНК в виде одной молекулы РНК. Участок ДНК между генами 18S и 28S транскрибируется вместе с этими генами и называется транскрибируемым спейсером (ITS). В нём расположены короткие последовательности, которые вырезаются из общего транскрипта в ходе процессинга РНК. У млекопитающих и амфибий в коротком спейсере формируется 5,8S рРНК – небольшая молекула, входящая в состав рибосомы и образующая в ее составе водородную связь с 28S рРНК (Льюин, 2011).

У эукариот длина транскрипта варьирует от 7000 до 8000 п.н. и кодирующая часть составляет 70–80%. В ходе созревания рРНК лидерная последовательность и часть транскрибируемого спейсера, некодирующая 5,8S рРНК, деградируют до нуклеотидов. Транскрибируемые единицы разделяются участком ДНК, называемым нетранскрибируемым спейсером. Его длина варьирует в широких пределах: от 1750 п.н. у дрожжей до 30 000 п.н. у мыши; у дрозофилы он имеет длину 3750–6450 п.н..

Существенная часть генов 28S рРНК содержит инсерции. Каждая из единиц повтора может функционировать независимо от других. У дрозофилы гены, кодирующие 5S рРНК, представлены блоком, состоящим из 160–200 идентичных последовательностей длиной 385 п.н. каждая. Общая длина кластера составляет 60 000–80 000 п.н. Повторяющаяся единица состоит из кодирующей части (~135 п.н.) и спейсера (250 п.н.). Отдельные гены или группы генов в пределах кластера транскрибируются независимо друг от друга: делеции части генов не предотвращают активности остальных. Таким образом, последовательности ДНК не представляют единицы транскрипции и функционируют независимо одна от другой (Жимулев, 1994).

Гены гистонов присутствуют в геноме в нескольких копиях и собраны в tandemно повторяющиеся кластеры. Кластерная организация канонических гистоновых генов характерна для всех многоклеточных. Особенностью генов канонических гистонов является отсутствие в них интронов. Транскрипция этих генов происходит строго во время S-фазы клеточного цикла. Матричная РНК этих генов не полиаденилируется. Стоит отметить, что универсальной организации вариантных гистоновых генов для всех организмов не существует. Они, в свою очередь, не образуют кластеров и разбросаны по всему геному, нередко содержат интроны, транскрибируемая с них РНК полиаденилируется, а транскрипция происходит во время всего клеточного цикла (Brown, 2001; Marzluff et al., 2002). Например, общая длина кластера гистоновых генов у дрозофилы составляет примерно 500 000 п.н.. Отдельные гены в пределах повторенной единицы транскрибируются в противоположных направлениях (т.е. с разных цепей ДНК), что свидетельствует о независимости их функционирования.

Иногда в кластеры объединяются и неповторяющиеся гены. Семь генов, кодирующих белки теплового шока, располагаются во фрагменте ДНК длиной 12 000 п.н.. Для генов этого района также характерна разная направленность транскрипции. Гены, контролирующие развитие крупных частей тела, собраны в кластеры. Такие кластеры называют комплексами (Жимулев, 1994).

1.1.3. Структурная и регуляторная части гена. Структура транскриптов

В отличие от прокариот, у которых единственная РНК полимераз синтезирует все виды РНК, у эукариот три разных РНК-полимеразы транскрибируют РНК с трех разных типов генов. РНК-полимераза I синтезирует 18S, 28S, 5,8S рРНК, РНК-полимераза II считывает мРНК с генов, кодирующие белки и некоторые малые ядерные РНК (мяРНК), РНК-полимераза III транскрибирует гены 5S рРНК, тРНК и остальные мяРНК. Каждый ген состоит из регуляторной части, с которой начинается транскрипция, кодирующей части, в которой записана информация о структуре белка и терминирующей части, в которой завершается транскрипция. Для правильного считывания информации необходимо наличие кодонов инициации и кодонов терминации транскрипции. Обнаружено два типа терминирующих последовательностей, и каждый ген имеет один из

них. Один тип последовательностей может распознаваться непосредственно РНК-полимеразой, другой тип – РНК-полимеразой в ассоциации с ρ -фактором (Жимулев, 2000).

В инициации транскрипции участвует не только РНК-полимераза, но и множество белков, называемых общими факторами транскрипции. Поэтому в регуляторной части помимо промотора, с которым взаимодействует РНК-полимераза П находятся многочисленные регуляторные последовательности, с которыми связываются факторы транскрипции и другие лиганды. Регуляторные последовательности могут находиться как выше, так и ниже точки инициации транскрипции. Ядро промотора включает в себя ТАТА-бокс (ТАТААА) называемый по-другому доменом Хогнесса или доменом Голдберга-Хогнесса, элемент BREu (upstream TFIIIB recognition element), находящийся перед ТАТА-боксом, элемент BREd (downstream TFIIIB recognition element), находящийся после ТАТА-бокса, Inr (Initiator) участок, в котором начинается транскрипция), нижележащий элемент промотора – DPE (downstream promoter element), элемент связанный с репликацией ДНК – DRE (DNA replication-related element). Обычно промоторы содержат различные комбинации из этих элементов. Кроме того в инициации транскрипции важную роль играют и проксимальные элементы. Наиболее известные – это СААТ-бокс (GGCCAATCT) и GC-бокс (GGCCGG) (Buratowski, 1995; Разин и др., 2015).

Хотя промоторные элементы являются ключевыми в осуществлении транскрипции, для максимальной эффективности необходимы удаленные регуляторные элементы, которые могут как активировать (энхансеры), так и подавлять (сайленсеры) транскрипцию. Энхансеры представляют собой специфические участки связывания особых регуляторных белков, усиливающих или активирующих процесс транскрипции. Такое расположение энхансеров дает основание считать, что любой конкретный промотор может находиться под контролем неограниченного числа энхансеров, разбросанных по всему геному. Однако этого не происходит, потому что в геноме существуют так называемые архитектурные элементы, которые поддерживают взаимодействия между удаленными участками генома. Наиболее известные архитектурные элементы – это инсуляторы. Инсуляторы, в зависимости от ряда дополнительных условий, могут либо разделять функциональные домены генома, либо способствовать установлению связи

между удаленными энхансерами и контролируемыми ими промоторами (Разин и др., 2015).

При изучении аденовируса 2 в 1977 году Филипп Шарп (Berget et al., 1977) и Ричард Робертс (Chow et al., 1977) показали, что нуклеотидная последовательность, кодирующая определенный белок, прерывается протяженными некодирующими участками. Они, подобно межгенным спейсерам, не содержат генетической информации и были названы интронами. В свою очередь, кодирующие участки стали называть экзонами. Такой тип мозаичной организации обнаружен у большинства генов эукариот. Число и внутригенная локализация интронов специфичны для каждого гена, это становится очевидным при сравнении организации гомологичных генов у разных видов. Некоторые гены содержат один–два интрона, но часто их значительно больше. Детальное исследование экзонов у десяти наиболее изученных модельных объектов показало, что у эукариот в среднем один ген содержит 3,7 интрона на один тип кодирующего участка ДНК (Deutsch, Long, 1999). Также было показано, что в интронах некоторых генов могут располагаться и другие гены. По результатам проекта «Геном дрозофилы», 8% генов у этого вида локализованы в интронах других генов (Ashburner, 2000).

Длина интрона может быть разной – от нескольких десятков пар нуклеотидов до многих тысяч. Общая длина всех интронов значительно превышает суммарную длину экзонов. Экзоны, как правило, имеют небольшую длину. Разные экзоны в пределах гена не только различаются по составу кодируемых ими аминокислот, но и имеют определенные структурные особенности. Например, в геноме человека обнаружено около 45 тысяч так называемых CpG островков. Это тяжи неметилированной ДНК с высоким содержанием динуклеотидов CpG. Чаще всего они располагаются в регуляторных областях генов, но в генах человека вероятность встретить CpG островки в первых экзонах в 13 раз выше, чем в интронах, и в два раза выше, чем в других экзонах (Antequera, Bird, 1993).

Кроме уникальных последовательностей ДНК, содержащих собственно гены, в геномах организмов встречается значительное количество ДНК-последовательностей, которые представляют собой повторы, различные по своему составу. В свою очередь, некоторые функциональные гены, присутствующие в геноме более чем в одной копии,

ранее представляли собой повторенные последовательности ДНК. Однако, большинство ДНК-повторов не являются генами, и в действительности, функциональное значение многих из них, по-прежнему, неясно (Elder, Turner, 1995). Найдены также последовательности, называемые псевдогенами. Они представляют собой нефункциональные последовательности, которые близки к последовательности активных генов, но имеют ряд нуклеотидных замен, делеций и дупликаций, нарушающих их экспрессию (Vanin, 1985; Balakirev, Ayala, 2003).

1.2. Общие представления о структуре генов гормона роста у позвоночных животных

Гормон роста (соматотропный гормон (СТГ), соматотропин) — один из гормонов передней доли гипофиза. Относится к классу полипептидных гормонов, регулирует рост скелета и участвует в других важных процессах организма. Впервые, гормон роста был выявлен по способности вызывать рост скелета и увеличение массы тела у молодых животных (Harvey, Hull, 1997). Наиболее изученным является гормон роста и вместе с пролактином и соматолактином относится к семейству гипофизарных гормонов. В настоящее время известна полная аминокислотная последовательность и структура генов гормона роста человека, некоторых млекопитающих и ряда других позвоночных животных.

В геноме человека ген гормона роста (hGH) и плацентарный лактоген (hPL), известный также как ген хорионического соматомаммотропина (hCS) образуют кластер из пяти близко расположенных к друг другу генов. В состав этого кластера входят два гена hGH и три гена hPL (Hirt et al., 1987, Barrera-Saldana, 1998). Гены в кластере имеют одинаковую ориентацию и расположены в положении от 5'-концу к 3'-концу в следующем порядке: hGH-N, hPL-1, или hCS-L, hPL-2, или hCS-A, hGH-V, hPL-3, или hCS-B. Гены hPL-2 и hPL-3 экспрессируются в трофобластах плаценты и кодируют один и тот же белковый продукт. Ген hPL-1 отличается от остальных генов, кодирующих хорионический соматомаммотропин, во втором интроне обнаружен измененный сайт сплайсинга, и это может указывать на то, что ген hPL-1 является псевдогеном, хотя производство функциональных продуктов не исключено. Ген hGH-N кодирует

гипофизарный гормон роста, экспрессируется только в специализированных клетках аденогипофиза – соматотрофах. Альтернативный вариант гормона роста кодирует ген hGH-V, который заменяет гипофизарный гормон роста на поздних сроках беременности, а экспрессируется он вместе с генами hPL в плаценте (Hirt et al., 1987, Chen et al., 1989, Perez-Maya et al., 2016).

Гены hGH – hPL расположены на длинном плече 17 хромосомы в сегменте q22–24 (Owerbach et al., 1980, George et al., 1981, Harper et al., 1982). Все пять генов, включая их ближайшие фланкирующие области, были секвенированы (DeNoto et al., 1981; Hirt et al., 1987), и оказались консервативны больше, чем на 95%. Полная нуклеотидная последовательность локуса hGH охватывает 66495 п.н. и содержит последовательности всех пяти генов. Несмотря на то, что по данным рестрикционного анализа гены hGH занимают участок длиной 2,6 т.п.н. каждый, два гена hPL (hPL-2 и hPL-3) находятся на участке длиной 2,9 т.п.н. и ген hPL-3 можно обнаружить во фрагменте 9,5 т.п.н. Все пять генов, входящие в кластер, состоят из пяти экзонов и четырех интронов. Гены разделены межгенными последовательностями длиной от 6000 до 13000 п. н., которые содержат 48 повторяющихся последовательностей типа Alu. Они встречаются в обеих ориентациях, а три Alu последовательности представлены в усеченном виде. Их положения можно рассматривать как исторические ориентиры, отражающие эволюционные события при образовании этого локуса (Chen et al., 1989). Экспрессию генов в разных тканях и в разные периоды времени обеспечивает взаимодействие транскрипционных факторов с различными регуляторными элементами промотора и с последовательностями энхансеров (Barrera-Saldana, 1998). На расстоянии от –15 000 до –32 000 п.н. перед генами hGH и hPL, расположена дистальная область контрольного локуса (LCR), которая играет ключевую роль в регуляции экспрессии генов, входящих в кластер (Jones et al., 1995; Su et al., 2000).

Организация кластера генов GH значительно варьирует среди представителей разных групп приматов, к которым относится и человек. Макака-резус (*Macaca mulatta*), гамадрил (*Papio hamadryas*) и горилла (*Gorilla gorilla gorilla*) имеют шесть генов в кластере (Gonzalez-Alvarez et al., 2006; Rodriguez-Sanchez et al., 2010); человек и шимпанзе пять, а в состав кластера генов орангутана (*Pongo abelii*) входит всего четыре гена, и только три из них являются функциональными. У шимпанзе и орангутана также, как и у

человека, ген GH-N экспрессируется в гипофизе, а продукт гена GH-V можно обнаружить в плаценте. Гены (CSH), подобные генам хорионического соматомаммотропина человека (hCS) у горилл и орангутанов более вариабельны, чем гены GH, но наиболее значимым является изменение количества таких генов в геноме, варьирующих от двух у орангутанов и до четырех у горилл. В гене oCSH-L орангутана также, как и в гене hCL человека, имеется измененный сайт сплайсинга на 5'-конце в интроне 2. А наличие стоп-кодона в рамке считывания в экзоне пять указывает на то, что ген oCSH-L представляет собой псевдоген. Последовательность гена oCSH-B орангутана имеет гомологию с последовательностью генов hCS-A/-B человека, что указывает на то, что данный ген у орангутана функциональный (Perez-Maya et al., 2016).

Тем не менее последовательность гена gCSH-A1 гориллы имеет некие сходства с последовательностью гена oCSH-L орангутана, но существенно отличается от остальных генов gCSH-A2, gCSH-A3 и gCSH-B. В свою очередь авторы (Perez-Maya et al., 2016) отмечают, что гены gCSH-A2, gCSH-A3 и gCSH-B очень похожи между собой. Сходство между этими тремя генами gCSH гориллы может частично отражать события преобразования генома, а также недавние процессы дупликации генов.

У большинства остальных млекопитающих, включая и полуобезьян, ген гормона роста не формирует кластер гомологичных генов и в геноме можно обнаружить только одну копию. Но среди представителей различных отрядов млекопитающих можно встретить виды, у которых обнаружено по два гена гормона роста. У представителей отряда Парнокопытные (*Artiodactyla*) – обыкновенного бегемота (*Hippopotamus amphibius*), жирафа (*Giraffa camelopardalis*) (Manioui et al., 2004) и малого оленёка (*Tragulus javanicus*) (Wallis, Wallis, 2001) найдены два похожих гена GH, которые, несмотря на наличие нуклеотидных замен, кодируют одинаковые белки. По два гена гормона роста встречается у овец (Valinsky et al., 1990) и коз (Yamano et al., 1991). У овец один ген гормона роста представлен одной копией, в то время как другой ген, двумя копиями GH2-N и GH2-Z. Все три гена имеют сходство. У козы по результатам исследования рестрикционных фрагментов обнаружено два фрагмента: первый содержит один ген гормона роста (gGH1), а второй два гена, расположенных тандемным образом (gGH2, gGH3). Гордон с соавторами (Gordon et al., 1983) в своей работе описывают только

один функциональный ген GH, но упоминают о наличии второго гена GH в геноме быка. Авторы считают, что это либо псевдоген GH, либо нефункциональный плацентарный лактоген, потому что продукта данного гена им не удалось обнаружить.

Структура генов гормона роста остается весьма консервативной среди млекопитающих (Barta et al., 1981; Gordon et al., 1983; Das et al., 1996): кодирующая часть состоит из пяти экзонов, разделенных четырьмя промежуточными последовательностями – интронами. Длина интронов варьирует в пределах 200–300 п.н. у разных видов. Например, гены гормона роста быка (Woychik et al., 1982; Gordon et al., 1983) и овцы (Byrne et al., 1987) имеют очень высокое сходство нуклеотидных последовательностей, что можно объяснить их недавней независимой историей и, вероятно, отбором. Общая длина генов у этих двух видов отличается лишь на один нуклеотид: 1793 п.н. – ген гормона роста быка и 1792 п.н. – длина гена гормона роста овцы. Последовательности гена гормона роста у этих видов имеют сходство не только между собой, но и с известными последовательностями генов гормона роста крысы и человека. Причем, высокая степень сходства наблюдается не только в кодирующей области гена, но и во фланкирующих областях. В 5'-фланкирующей области у всех генов имеется последовательность ТАТААА, а также другие консервативные участки, необходимые для инициации транскрипции.

Известные на сегодняшний день нуклеотидные последовательности гена гормона роста различных видов грызунов (Barta et al., 1981; Das et al., 1996; Wallis, Wallis 1995; Liouris et al., 1999) по своей структуре не отличаются от последовательностей генов гормона роста других млекопитающих. Расположение экзонов и интронов в гене аналогично расположению этих структурных частей гена у других видов млекопитающих, а границы экзонов-интронов соответствуют правилу, согласно которому на 5'- и 3'-концах интронов расположены консервативные нуклеотиды, по которым происходит сплайсинг интронов при процессинге мРНК. На 5'-конце интрона это пара нуклеотидов GT, а на 3'-конце это пара нуклеотидов AG (Das et al., 1996). Размер интронов гена GH обычно не превышает 300 п.н., исключение представляет второй интрон (I_B) в гене крысы (Barta et al., 1981). Длина его составляет 718 п.н., что на 530 п.н. больше, чем у мыши (Das et al., 1996). Интересной особенностью интрона В гена GH крысы является более низкое

содержание GC (45,5%), наличие трех длинных богатых AT областей, которые содержат тандемные повторы разной длины, но все эти тандемные повторы заканчиваются длинным поли-А концом. Общая структура этого участка интрона длиной примерно 500 п.н. напоминает дисперсные повторы ДНК, встречающиеся в кластере GH-подобных генов приматов. Но стоит отметить, что, например, интрон В генах гормона роста человека (hGH) имеет длину 200 п.н. и не содержит больших повторяющихся последовательностей (Barta et al., 1981). Не наблюдается и длинных повторяющихся последовательностей в интронах генов европейского кролика (*Oryctolagus cuniculus*) (Wallis, Wallis 1995) и палестинского слепыша (*Spalax ehrenbergi*) (Lioupis et al., 1999). Кластер GH-подобные гены у грызунов не образуют, и ген гормона роста в геноме представлен одной копией.

Ген гормона роста птиц превосходит по длине ген гормона роста млекопитающих. Если средняя длина гена гормона роста млекопитающих примерно 2000 п.н. (1,5 т.п.н – длина гена GH мыши (Das et al., 1996); 1,7 т.п.н. – длина гена GH быка (Gordon et al., 1983); 2000 п.н. – длина гена GH крысы (Barta et al., 1981)), то длина гена гормона роста птиц варьирует в диапазоне от 3000 до 5000 п.н. (3,3 т.п.н. длина гена у мухоловки-пеструшки (*Ficedula hypoleuca*) (Buggiotti et al., 2006); 3,5 т.п.н. длина гена у курицы (Tanaka et al., 1992); 5,2 т.п.н. длина гена гормона роста у утки (*Anas platyrhynchos*) (Kansaku et al., 2008)). Различия в длине генов обусловлены различными длинами интронов. Но несмотря на то, что у разных видов птиц интроны различаются по своей длине, количество их остается такое же, как и в гене гормона роста млекопитающих: четыре интрона разделяют пять экзонов.

Для птиц, представителей отряда воробьинообразные (*Passeriformes*), характерно наличие в геноме двух генов гормона роста GH1A и GH1B. В отличие от гена гормона роста приматов, который в составе кластера генов занимает участок одной хромосомы 17, гены гормона роста воробьинообразных располагаются на разных хромосомах: ген GH1A расположен на хромосоме 1, а ген GH1B – на хромосоме 27 (Arai, Iigo, 2010). Согласно работе Юри с соавторами (Yuri et al., 2008), копия гена гормона роста возникла путем дупликации у предка воробьинообразных. Второй ген присутствует у большинства представителей данного отряда, а экспрессия обоих генов в мозге зебровой амадины

(*Taeniopygia guttata*) указывает на то, что оба гена нормально функционируют (Yuri et al., 2008).

Описание полной нуклеотидной последовательности гена гормона роста среди представителей рептилий и амфибий встречается только для китайской трёхкилевой черепахи (*Chinemys reevesii*) (Liu et al., 2016). Как и у большинства позвоночных животных ген гормона роста китайской трёхкилевой черепахи (*Chinemys reevesii*) состоит из пяти экзонов и четырех интронов. Длина гена составляет 8517 п.н.. Такая разница в размерах, если сравнить с геном гормона роста млекопитающих, связана с большей протяженностью интронов: суммарная длина всех интронов составила 7726 п.н., а длина первого интрона – 2486 п.н.. Это самый длинный первый интрон гена гормона роста из всех описанных среди позвоночных. Но несмотря на разницу в длине, основания, фланкирующие концы интронов, такие же как у млекопитающих: GT на 5'-конце и AG на 3'-конце (Liu et al., 2016). Для морской черепахи (*Chelonia mydas*) (Yasuda et al., 1989), новогвинейского крокодила (*Crocodylus novaeguineae*) (Noso et al., 1995), зеленой игуаны (*Iguana iguana*) (Ávila-Mendoza et al., 2014) и китайского трионикса (*Trionyx sinensis*) (Yang et al., 2020) авторы приводят только аминокислотные последовательности гормона роста, полученные на основе кДНК.

По количеству аминокислот гормон роста рептилий совпадает с гормоном роста человека и других млекопитающих. У всех перечисленных видов гормон роста состоит из 190 аминокислотных остатков. Сходство гормона роста рептилий с гормоном роста млекопитающих составило 75–82%; с гормоном роста птиц 85–89%; с гормоном роста костистых рыб 33–59% (Yasuda et al., 1989; Noso et al., 1995; Ávila-Mendoza et al., 2014; Liu et al., 2016). У амфибий аминокислотная последовательность гена гормона роста известна только для гладкой щпорцевой лягушки (*Xenopus laevis*) (Huang, Brown, 2000) и лягушки-быка (*Rana catesbeiana*) (Kobayashi et al., 1991). У гладкой щпорцевой лягушки ген гормон роста представлен двумя копиями xGH-A и xGH-B. Гены xGH-A и xGH-B кодируют белки-предшественники из 214 и 208 аминокислот, соответственно. Они на 77% идентичны между собой и на 60–70% идентичны с последовательностями гормона роста других видов (Huang, Brown 2000). Гормон роста лягушки-быка состоит из 190 аминокислот и имеет 69% гомологии с гормоном роста морской черепахи; 66% с

гормоном роста курицы; 61% с гормоном роста овцы и 48% с гормоном роста человека (Kobayashi et al., 1991).

Полная нуклеотидная последовательность гена гормона роста была получена для одного из самых древних представителей позвоночных – морской миноги (*Petromyzon marinus*) (Kawauchi et al., 2002; Moriyama et al., 2006). Ген гормона роста морской миноги является самым большим из всех известных генов GH, состоящих из пяти экзонов и четырех интронов. Длина гена GH составляет 13 604 п.н.. Он включает кодирующую и регуляторную часть. В промоторной области имеются сайты связывания с транскрипционными факторами, характерными для генов GH млекопитающих. Информация, которую кодируют каждый экзон, подобна той, что кодируют гены гормона роста млекопитающих и птиц. Первый экзон кодирует 5'-нетранслируемую область (UTR) и первые три аминокислоты сигнального пептида; второй экзон кодирует остальную часть сигнального пептида и N-конец (34 аминокислоты) зрелого белка. Экзоны три и четыре кодируют 42 и 44 аминокислоты зрелого гормона роста соответственно, а пятый экзон кодирует 61 аминокислоту и 3'-нетранслируемую область. Все четыре интрона начинаются с динуклеотида GT и заканчиваются нуклеотидами AG, как и на концах интронов в гене GH млекопитающих. Считается, что именно ген гормона роста, который был обнаружен в геноме бесчелюстных, может являться предшественником гена гормона роста, который встречается у других более эволюционно продвинутых позвоночных животных (Moriyama et al., 2006).

Аналог гена гормона роста позвоночных удалось обнаружить даже в геноме ланцетника. Долгое время считалось, что у бесчерепных отсутствуют гены, кодирующие гипофизарные гормоны (гормон роста, пролактин, аденокортикотропные гормоны) (Holland et al., 2008), однако Ли с соавторами (Li et al., 2014) показали, что у ланцетника (*Branchiostoma japonicum*) присутствует гормон подобный гормону роста позвоночных (GH-like hormone), который является функциональным. Гормон роста ланцетника способен связываться с рецептором гормона роста (GHR), стимулировать экспрессию инсулин-подобного фактора роста (IGF), увеличивать скорость роста и исправлять дефекты морфогенеза у эмбрионов данио-рерио (*Danio rerio*), вызванные нехваткой гормона роста. Последовательность кДНК гормона роста ланцетника состоит из 1646 п.н.,

из которых 627 п.н. приходится на открытую рамку считывания, а 1019 п.н. на 3'-нетранслируемую область. Белок состоит из 208 аминокислот и на 14,5% идентичен другим известным гормонам роста, третичная структура белка близка третичной структуре гормона роста человека. Аналогов пролактина или соматолактина у ланцетника обнаружено не было. Это дает основания считать, что последовательность гена гормона роста является предшественником для остальных генов семейства: пролактина и соматолактина (Li et al., 2014).

1.3. Структура генов гормона роста рыб

Ген гормона роста хрящевых рыб не отличается от генов большинства позвоночных животных. Нуклеотидная и аминокислотная последовательности получены для катрана (*Squalus acanthias*) (Moriyama et al., 2008) и большой голубой акулы (*Prionace glauca*) (Yamaguchi et al., 1989). Ген гормона роста катрана состоит из пяти экзонов и четырех интронов, количество аминокислот, которые кодирует каждый экзон, такое же как у птиц, млекопитающих и костистых рыб. На основании сравнительного анализа аминокислотных последовательностей авторы (Moriyama et al., 2008) показали, что гормон роста катрана и большой голубой акулы ближе к четвероногим, чем к костистым рыбам. Степень гомологии с четвероногими животными составляет 47–68%, с костистыми рыбами 37–52%, соответственно. Такое различие между хрящевыми и костистыми рыбами указывает на быструю молекулярную эволюцию гена гормона роста костистых рыб в сравнении с другими группами позвоночных, у которых эволюция гена гормона роста протекала гораздо медленнее (Yamaguchi et al., 1989; Moriyama et al., 2008). Данные, полученные по хрящевым рыбам, подтверждают, что структура гена гормона роста консервативна и сохраняется на протяжении эволюции большинства позвоночных животных. Считается, что ген, состоящий из пяти экзонов и четырех интронов, отражает структуру гена-предшественника, который дал начало современным генам гормона роста (Moriyama et al., 2006). Однако, среди костистых рыб встречаются виды, у которых в составе гена имеются дополнительные структурные элементы.

Среди костистых рыб можно выделить целые отряды, у которых имеется еще один интрон, и соответственно – еще один экзон, и на основании наличия или отсутствия

дополнительного интрона гены гормона роста рыб по своей структуре могут быть разделены на две группы (Moriyama et al., 2006). К первой группе относятся гены, которые соответствуют генам млекопитающих и птиц, и состоят из пяти экзонов и четырех интронов. Ген с такой структурой встречается у представителей отрядов карпообразных (Cypriniformes) (Chiou et al., 1990; Hong, Schartl, 1993; Panicz et al., 2012) и сомообразных (Siluriformes) (Sekar et al., 2014). Ко второй группе относятся гены, в которых присутствует пятый дополнительный интрон, и состоят они уже из шести экзонов и пяти интронов. Такой ген гормона роста встречается у представителей отрядов лососеобразных (Salmoniformes) (Devlin, 1993; Du et al., 1993), окунеобразных (Perciformes), иглобрюхообразных (Tetraodontiformes) (Venkatesh, Brenner, 1997), камбалообразных (Pleuronectiformes) (Tanaka et al., 1995), цихлообразных (Cichliformes) (Ber, Daniel, 1992), ставридообразных (Carangiformes) (Ohkubo et al., 1996; Almuly et al., 2000). Среди генов, которые имеют дополнительный интрон можно выделить виды с небольшим пятым интроном длиной 70–100 п.н. и виды, у которых размер пятого интрона варьирует в пределах от 200 до 600 п.н.. Короткий интрон встречается у нильской тилляпии (*Oreochromis niloticus*) (Ber, Daniel, 1992), азиатского паралихта (*Paralichthys olivaceus*) (Tanaka et al., 1995) и желтохвоста (*Seriola quinqueradiata*) (Almuly et al., 2000), а интроны длиной более 200 п.н. описаны в генах гормона роста лососеобразных рыб (Salmoniformes) (Yang et al., 1997). Остается пока неясным, произошел ли пятый интрон, который встречается в гене гормона роста тилляпии и лососеобразных от одной и той же предковой последовательности, или возник независимо в разных отрядах. Сравнение нуклеотидных последовательностей пятого интрона гена гормона роста тилляпии и лососевых показало (Yang et al., 1997), что последовательности имеют очень низкую степень гомологии. Эти результаты указывают на то, что эти два типа интронов, возможно, эволюционировали из разных предковых последовательностей в результате событий инсерций (вставок) последовательностей в ген, которые произошли независимо друг от друга (Yang et al., 1997).

Гены гормона роста рыб обладают достаточно высоким уровнем изменчивости по длине. Размер генов гормона роста у рыб может варьировать в достаточно широких пределах: от 1,6 т.п.н. у латеса (*Lates calcarifer*) (Almuly et al., 2000) до более чем 4 т.п.н.

у желтохвоста (*S. quinquerediata*) (Almuly et al., 2000), что в два раза больше генов гормона роста млекопитающих, которые являются консервативными и их размер не превышает 2 т.п.н. (Chen et al., 1989; Ruynänen, Primmer, 2006). Однако сравнение генов гормона роста радужной форели и человека показывает (Chen et al., 1989), что первые четыре экзона гена радужной форели гомологичны экзонам гена гормона роста человека. А последний, пятый экзон, в гене гормона роста человека эквивалентен экзонам V и VI гена гормона роста радужной форели. И такая разница в размерах может быть обусловлена не только присутствием дополнительного интрона в генах гормона роста у представителей некоторых отрядов рыб, но и наличием tandemных повторов различной длины, а также возможных инсерций или делеций. Например, у представителей отряда карпообразных (Cypriniformes) и млекопитающих количество экзонов и интронов одинаковое (5 экзонов и 4 интрона). Однако размер интронов и 3'-нетранслируемых областей у карповых рыб больше, чем у млекопитающих. У рыб в среднем 3'-нетранслируемая область длиной около 500 п.н., а у млекопитающих и птиц около 100 п.н.. Длины первого, третьего и четвертого интронов в гене лососевых рыб больше интронов в гене млекопитающих в 2–2,5 раза (Zhu et al., 1992; Hong, Schartl, 1993).

У бурой фугу (*Fugu rubripes*) длина первого интрона составляет 608 п.н., он в несколько раз больше, чем первый интрон в генах гормона роста других рыб. Связано это с наличием сателлитной последовательности длиной 438 п.н., состоящей из элементов длиной 42 п.н. (TACCTGAGGCTGAATCCACTGTCTTCCTACCTGTСТААСССТ), и их фрагментов. Такой повтор не был найден в интронах гена гормона роста других костистых рыб и, согласно авторам (Venkatesh, Brenner, 1997), эта вставка в первый интрон бурой фугу (*F. rubripes*) произошла не так давно. Во втором и пятом интроне гена гормона роста бурой фугу (*F. rubripes*) встречаются тетра nukлеотидные (СТGT)₂₈ и тринуклеотидные повторы (GAT)₇₉, соответственно. Укороченный вариант тринуклеотидного повтора обнаружен в пятом интроне тилапии (*Oreochromis niloticus*) (Ber, Daniel, 1992), азиатского паралихта (*P. olivaceus*) (Tanaka et al., 1995) и латеса (*L. calcarifer*) (Almuly et al., 2000). Векантеш и Бренер считают (Venkatesh, Brenner, 1997), что этот минисателлитный повтор присутствовал в пятом интроне у общего предка иглобрюхообразных (Tetraodontiformes), камбалообразных (Pleuronectiformes) и окунеобразных (Perciformes).

Повторы различной длины встречаются в интронах гена гормона роста дорады (*Sparus aurata*) (Almuly et al., 2000). Минисателлитные повторы, состоящие из 17 и 22 нуклеотидов, можно обнаружить в первом и третьем интроне соответственно. Различные микросателлитные повторы из ди-, три- и тетрануклеотидов встречаются во всех интронах с первого по пятый. Инвертированные повторы, входящие в состав второго интрона увеличивают его размер до 1747 п.н. (Almuly et al., 2000).

Среди рыб можно встретить примеры, когда размер гена гормона роста изменяется не только за счет инсерций и увеличения числа повторов, но и за счет делеций. У рыб из рода лабео (*Labeo*) ген гормона роста отличается у разных видов (Rajesh, Majumdar, 2007). Связано это с небольшими делециями, которые затронули не только интроны, но и экзоны. Длина транскрипта у всех видов рода лабео (*Labeo*) составляет 630 п.н. (210 аминокислот), исключение составляет роху (*Labeo rohita*), у которого из-за делеции 9 нуклеотидов в четвертом экзоне длина транскрипта сократилась до 621 п.н., что соответствует 207 аминокислотам. Из четырех интронов только третий интрон демонстрирует изменчивость между видами (от 1069 до 1255 п.н.), и связывают это также с небольшими делециями (Rajesh, Majumdar, 2007).

Разное количество интронов, их вариабельность, присутствие повторяющихся последовательностей в гене, могут служить естественным маркером для понимания эволюционных взаимоотношений в различных группах костистых рыб. Пятый интрон, который присутствует у представителей некоторых отрядов костистых рыб, делит их филогенетическое дерево на две группы (Venkatesh, Brenner, 1997). К первой группе относится надотряд Костнопузырные (*Ostariophysii*); к нему принадлежат карпообразные и сомообразные, у которых в составе гена гормона роста всего четыре интрона. Вторая группа включает в себя надотряды протолучеперых (*Protacanthopterygii*) и колючеперых (*Acanthopterygii*), которые имеют общее происхождение, и в составе гена гормона роста присутствует пятый интрон. Поскольку в линии, ведущей к четвероногим, которая дивергировала от костистых рыб в начале эволюции, ген гормона роста имеет только четыре интрона, то, возможно, что пятый интрон был либо приобретён общим предком протолучеперых (*Protacanthopterygii*) и колючеперых (*Acanthopterygii*) после отделения костнопузырных (*Ostariophysii*), либо пятый интрон присутствовал у общего предка и был

утерян в линии ведущей к костнопузырным (Ostariophysii) и четвероногим (Venkatesh, Brenner, 1997). Еще одним аргументом в пользу такой теории стали результаты, полученные Раджешем и Маджумдарой (Rajesh, Majumdar, 2007). Авторы анализировали кДНК гормона роста различных видов рыб и показали (Rajesh, Majumdar, 2007), что виды, у которых количество экзонов и интронов в гене одинаковое, кластеризуются вместе. Так виды, которые относятся к отрядам лососеобразных, окунеобразных и камбалообразных объединялись отдельно от видов, которые принадлежали отрядам карпообразных и сомообразных.

Анализ микросателлитных последовательностей, которые встречаются в интронах гена гормона роста рыб, также может указывать на ряд филогенетических особенностей. Например, в четвертом интроне гена гормона роста щуковых (Esocidae) присутствует минисателлитный повтор длиной 33 нуклеотида, количество его копий варьирует от 7 до 16 среди различных видов щуковых (Barnett et al., 2007). Идентичная последовательность присутствует и в гене гормона роста лососевых, но только в единичной копии. Это может указывать на присутствие такой последовательности у общего предка и увеличение копий этой последовательности у щуковых после дивергенции щуковых и лососевых (Barnett et al., 2007).

У костистых рыб гены гормона роста отличаются не только своими размерами, которые варьируют между различными отрядами, но и количеством генов, представленных в геноме. Среди рыб встречаются представители, у которых ген гормона роста, как и у большинства млекопитающих, представлен одной копией. Одна копия гена гормона роста была найдена у бурой фугу (*F. rubripes*) (Venkatesh, Brenner, 1997), дорады (*S. aurata*) (Almuly et al., 2000), у исследованных видов рода лабео (*Labeo*) (Rajesh, Majumdar, 2007), и у некоторых представителей карпообразных: белого амура (*Stenopharyngodon idellus*) (Zhu et al., 1992) и белого толстолобика (*Hypophthalmichthys molitrix*) (Hong, Schartl, 1993).

В геномах остальных исследованных видов рыб ген гормона роста представлен двумя несвязанными паралогичными генами *gh1* и *gh2*. Два гена обнаружено у тилляпии (*O. niloticus*) (Ber, Daniel, 1992). Две копии гена тилляпии имеют одинаковую структуру (5 интронов, 6 экзонов) и обе копии являются функциональными, поскольку было

показано, что они кодируют одинаковую белковую последовательность. Среди карпообразных встречаются виды с двумя копиями гена гормона роста. По две копии гена имеют все виды семейства чукучановые (Catostomidae) (Bart et al., 2010) и некоторые представители семейства карповых (Choi et al., 1990). У большинства лососеобразных ген гормона роста также представлен двумя копиями (Male et al., 1992; Devlin, 1993; Du et al., 1993; Панькова и др., 2017). Увеличение числа генов связано с полной дупликацией генома, в результате которой, некоторые из видов стали тетраплоидными, и в их геноме присутствует по две копии гена гормона роста, а те виды, которые сохранили диплоидный геном, имеют одну копию гена гормона роста (Zhu et al., 1992; Bart et al., 2010).

У лососевых, которые представляют собой уникальный случай автотетраплоидизации 25–100 миллионов лет назад, оба гена существуют на протяжении всего времени дивергенции видов в этой группе (Allendorf, Thorgaard, 1984). По две копии гена гормона роста были идентифицированы в геноме основных представителей семейства лососевых (Salmonidae): атлантического лосося (*Salmo salar*) (Johansen et al., 1989; Male et al., 1992), нерки (*Oncorhynchus nerka*) (Devlin, 1993), кеты (*Oncorhynchus keta*) (Kawauchi et al., 1986), кижуча (*Oncorhynchus kisutch*) (Forbes et al., 1994), радужной форели или микижи (*Oncorhynchus mykiss*) (Agellon et al., 1988), чавычи (*Oncorhynchus tshawytscha*) (Du et al., 1993), у четырех видов гольцов рода *Salvelinus*: северной мальмы (*S. malma*), южной мальмы (*S. curilus*), гольца Леванидова (*S. levanidovi*), гольца Таранца (*S. taranetzi*) (Панькова и др., 2017). У всех исследованных видов оба гена состоят из шести экзонов и пяти интронов. Границы экзонов с интронами фланкированы такими же консенсусными последовательностями, как и у остальных позвоночных животных: на 5'-концах интронов – GT, а на 3'-концах – AG. Размер экзонов двух генов у всех видов лососёвых идентичен и составляет 74 п.н., 140 п.н., 117 п.н., 156 п.н., 147 п.н. и 63 п.н. для I–VI экзонов соответственно (Agellon, Chen, 1986; Male et al., 1992; Du et al., 1993; Devlin, 1993; Панькова и др., 2017). Интроны в генах гормона роста лососевых отличаются по своему размеру. Отличия эти связаны с присутствием инсерций и делеций внутри последовательностей интронов. Из-за разной длины интронов варьирует размер генов гормона роста как среди разных видов лососевых, так и между двумя копиями гена *gh* одного вида. Интрон С в генах *gh1* тихоокеанского и атлантического лосося больше, чем

интрон С в генах *gh2* у этих видов. Длина интрона С в гене *gh1* составляет примерно 700–800 п.н., а длина интрона С в генах *gh2* составляет 450–500 п.н., соответственно (Male et al., 1992; Yang et al., 1997). У представителей рода *Oncorhynchus* (тихоокеанские лососи) пятый интрон Е в гене *gh1* в несколько раз больше (~ 600 п.н.), чем в гене *gh2* (~200–250 п.н.) (Devlin, 1993).

Делеции разной длины были обнаружены в третьем и четвертом интронах в генах гормона роста некоторых видов лососевых (Панькова и др., 2013). В гене *gh1*, в интроне С у гольцов *Salvelinus curilus* и *Salvelinus fontinalis* обнаружена делеция из 6 нуклеотидов (GTCTAC) у остальных гольцов рода *Salvelinus* присутствуют две тандемные копии этого участка. Еще одна делеция размером 9 нуклеотидов (ATCACAATC) присутствует в интроне С у *Salvelinus namaycush*. Укороченный вариант такой делеции (ACAATC) был обнаружен у *O. nerka* и *O. tshawytscha*. В интроне D гена *gh1* у *Salvelinus leucomaenis* дважды повторяется последовательность из 11 нуклеотидов (CAGTGACATGA), которая у остальных лососевых представлена единственной копией. В третьем интроне (С) гена *gh2* у атлантического лосося (*S. salar*), нерки (*O. nerka*), чавычи (*O. tshawytscha*) и *S. leucomaenis* была обнаружена делеция размером 167 п.н.. Подобные вставки и делеции присутствуют во всех интронах кроме второго, его длина в обеих копиях гена одинаковая (Панькова и др., 2013).

Такие различия, которые встречаются в некодирующих последовательностях между двумя копиями генов гормона роста, позволяют использовать интроны в качестве маркеров для оценки филогенетических межродовых и межвидовых отношений внутри семейства. Анализ последовательностей интронов позволяет оценить темп эволюции и действие отбора на гены, которые возникли в результате дупликации. Филогенетический анализ интронов гена гормона роста использовался для изучения внутривидовых отношений тихоокеанских лососей и гольцов (Devlin, 1993; McKay et al., 1996; Westrich et al., 2002; Панькова и др., 2013). На основании филогенетического анализа интронов С и D генов *gh1* и *gh2* были показаны сестринские отношения между родами *Oncorhynchus* и *Salvelinus*, а не между *Salmo* и *Oncorhynchus*, как ожидалось по аналогии с морфологическими данными (Oakley, Phillips, 1999; Phillips et al., 2004).

Фланкирующие области двух генов гормона роста, также, как и интроны менее консервативны и демонстрируют низкий уровень гомологии. Гены *gh1* и *gh2* радужной форели имеют 78,6% и 89,2% гомологии в 3'- и 5'-нетранслируемых областях соответственно, по сравнению с 96,8% гомологии в кодирующих областях (Chen et al., 1989). Анализ генов гормона роста нерки (Delvin, 1993) показал, что некодирующие области двух генов существенно отличаются, это может быть связано с небольшим генетическим обменом между двумя генами.

Различия во фланкирующих областях двух копий генов могут отражаться на уровне их экспрессии. Янг с соавторами (Yang et al., 1997) показали, что гены *gh1* и *gh2* неравномерно экспрессируются в гипофизе рыб разного пола и рыб, находящихся на разных стадиях развития. У десятидневных мальков и двухлетних самок уровень мРНК *gh1* был заметно выше, чем уровень мРНК *gh2*. Однако у самцов того же возраста таких различий в экспрессии не наблюдалось. Авторы связывают более высокий уровень экспрессии гена *gh1* у самок с присутствием в интроне С гена *gh1* эстроген-чувствительного элемента (ERE), который отсутствует в интроне С гена *gh2* (Yang et al., 1997).

Помимо двух копий генов гормона роста, в геноме кижуча (*Oncorhynchus kisutch*) (Forbes et al., 1994), симы (*O. masou*) (Nakayama et al., 1999) и чавычи (*O. tshawytscha*) (Du et al., 1993) была обнаружена еще одна копия гена, которая оказалась очень похожа на ген *gh2* и находится в половой Y-хромосоме. В дальнейшем авторы установили (Du et al., 1993; Forbes et al., 1994), что данный ген является псевдогеном гена гормона роста (*ghψ*). Результаты показали, что ген не является функциональным и, вероятно, происходит от гена *gh2*. Отсутствие в гене части пятого экзона и пятого интрона, наличие стоп-кодона в пятом экзоне и неправильного сайта сплайсинга на границе первого интрона и второго экзона указывают на то, что ген *ghψ* не является функциональным. Отсутствие какой-либо формы псевдогена у самок дало основание полагать что данный ген встречается только у самцов, сцеплен с полом, передается от отца к сыну, поскольку расположен на Y-хромосоме, и был найден только у представителей тихоокеанских лососей (Du et al., 1993; Nakayama et al., 1999).

1.4. Эволюция генов гормона роста у позвоночных животных

Ген гормона роста принадлежит к семейству генов, которое кодирует гормоны белковой природы. Помимо гена гормона роста к этому семейству относятся: гены пролактина, соматолактина и хорионического соматомаммотропина (плацентарный лактоген). Между генами в семействе была выявлена значительная гомология по нуклеотидным последовательностям, что дает основания некоторым авторам предполагать, что ген гормона роста и другие гены, представленные в семействе, произошли от гена-предшественника путем дупликации и последующей дивергенции (Miller, Eberhardt, 1983; Chen et al, 1989; Walker, 1991; Wallis 1996; Daza et al., 2009).

Ген гормона роста, в отличие от остальных генов семейства, встречается в геномах всех позвоночных животных. Полную нуклеотидную последовательность гена гормона роста удалось обнаружить в одном из наиболее архаичных таксонов, в надклассе бесчелюстных типа хордовых, у морской миноги (*Petromyzon marinus*) (Moriyama et al., 2006). На основании полученных данных авторы (Kawauchi et al., 2002; Moriyama et al., 2006) делают вывод, что, именно, ген гормона роста является тем самым геном-предшественником, в результате, дупликации которого образуются другие гены, входящие в семейство. Авторы связывают это с тем, что гормон роста играет важную роль в развитии организма и на протяжении эволюции всех позвоночных сохраняет и выполняет очень важную функцию – стимуляцию роста, что особенно было необходимо для выживания предковых форм, вымерших остракодерм, древних представителей бесчелюстных.

Более позднее появление в геномах позвоночных нуклеотидных последовательностей новых гормонов, пролактина и соматолактина, могло сыграть важную роль в адаптации животных к меняющимся условиям среды и способствовать освоению позвоночными ареалов, которые раньше они не заселяли, за счет выполнения этими гормонами новых функций, не свойственных гормону роста (Kawauchi et al., 2002).

Появление новой функции (неофункционализация) у гена, который возникает в результате дупликации – это один из возможных вариантов эволюции копии гена. Помимо появления новой функции, те функции, которые были у гена изначально, могут быть

разделены между дублированными копиями (субфункционализация); также дублированный ген может стать псевдогеном в результате мутаций, которые способствуют появлению стоп-кодона в кодирующей последовательности (Zhang, 2003).

Дубликация гена-предшественника, которая послужила началом эволюции отдельных генов гормона роста и пролактина, должна была произойти примерно 350 миллионов лет назад. Присутствие гормона роста и пролактина у рыб и земноводных, которые разошлись около 400 миллионов лет назад подтверждает появление этих генов на раннем этапе эволюции позвоночных (Miller, Eberhardt, 1983). В дальнейшем гены-предшественники современных генов гормона роста и пролактина распределяются по разным хромосомам, например, у человека ген гормона роста находится в 17 хромосоме, а ген пролактина расположен на хромосоме 6 (Walker, 1991). Считается, что еще один ген, который относится к этому семейству и встречается только у млекопитающих – плацентарный лактоген, возник уже после расхождения грызунов и приматов 80 миллионов лет назад.

Сходство между генами плацентарного лактогена и геном гормона роста у приматов намного больше, чем, например, сходство гена гормона роста приматов с геном гормона роста грызунов. В других группах млекопитающих, где встречаются плацентарные лактогены, это скорее производные пролактина, а не гена гормона роста (Walker, 1991). Подтверждением этого служит тот факт, что геном-предшественником для плацентарного лактогена у грызунов является ген пролактина, а у приматов в качестве гена-предшественника выступает ген гормона роста. В своей работе Даквос с соавторами (Duckworth et al., 1986) показали это при сравнении последовательностей ДНК плацентарного лактогена крысы (rPL) с последовательностями гена пролактина крысы (rPRL) и гена гормона роста крысы (rGH). Из полученных данных очевидно, что нуклеотидная последовательность плацентарного лактогена крысы (rPL) более гомологична последовательностям пролактина крысы (rPRL) и пролактина человека hPRL 52% и 51% соответственно, чем последовательностям генов rGH, hGH или hPL, для которых выявлено по 34% гомологи для каждого гена (Duckworth et al., 1986). Следовательно, плацентарный лактоген возник в результате дубликации предкового гена hGH после отделения основных порядков млекопитающих, во время эволюции приматов.

Это подтверждается при сравнении их нуклеотидных последовательностей (Hirt et al., 1987; Walker, 1991). Например, у человека последовательности мРНК гена гормона роста и плацентарного лактогена являются гомологичными на 93,5%, в свою очередь сравнение последовательностей мРНК пролактина с мРНК гена гормона роста и с мРНК плацентарного лактогена имеет идентичность 42% и 41% соответственно. Вместе с геном гормона роста плацентарный лактоген формирует целый кластер близкорасположенных генов, который встречается в геноме человека и других высших приматов. Считается, что гены гормона роста и плацентарного лактогена, формирующие этот кластер, возникли из гена-предшественника в результате дупликации в интервале последних 10–15 миллионов лет (Chen et al., 1989).

Эволюция локуса протекала в три этапа. Авторы (Chen et al., 1989; Walker, 1991) предполагают, что в результате первого этапа дупликации гена-предшественника (Pre-GH/PL) образовался локус, который состоял уже из двух предковых генов: гена Pre-GH и гена Pre-PL. Дупликация этих предковых генов на втором этапе привела к образованию четырех генов: hGH-N, hPL-1/hPL-2, hGH-V и hPL-3. Последний этап дупликации гена hPL-1/hPL-2, который состоялся около 5 миллионов лет назад, привел к появлению современных генов hPL-1 и hPL-2.

Стоит отметить, что данный локус у высших приматов, который состоит из генов гормона роста и плацентарного лактогена, помимо последовательностей генов гормона роста и плацентарного лактогена содержит еще и 48 последовательностей Alu повторов, которые располагаются в локусе в обеих ориентациях; некоторые последовательности представлены в усеченной форме (Barsh et al., 1983). Положения этих последовательностей можно рассматривать как ориентиры, отражающие эволюционные события, происходившие в этом локусе. Баш с соавторами читают (Barsh et al., 1983), что ни один приобретенный Alu элемент не был утрачен во время эволюции локуса. Такая высокая плотность Alu элементов может быть связана с их непосредственным участием в дупликации генов, которые как раз и могли происходить за счет неравного кроссинговера между различными неаллельными Alu элементами. Усеченный Alu элемент на 3'-конце гена-предшественника Pre-GH/PL мог стать той самой «горячей точкой» начала

дупликации, которая привела к образованию генов Pre-GH и Pre-PL (Barsh et al., 1983; Hirt et al., 1987).

На происхождение генов семейства из одного общего предкового гена указывает не только высокий уровень гомологии нуклеотидных последовательностей, но и высокая степень сходства их аминокислотных последовательностей. Внутри каждого белка имеются области внутренней гомологии, которые указывают на то, что предшественник этого семейства генов был образован путем дупликации предкового гена, который состоял из меньшего количества структурных элементов и в результате дупликации образовался ген-предшественник, содержащие пять экзонов и четыре интрона (Miller, Eberhardt, 1983; Hirt et al., 1987)

В структуре генов, представленных в данном семействе, прослеживается ряд общих черт, которые подтверждают гипотезу о происхождении гена гормона роста, пролактина и плацентарного лактогена от общего гена-предшественника. Количество экзонов и интронов совпадает во всех трех генах (пять экзонов и четыре интрона); идентичны и границы сайтов сплайсинга, даже несмотря на то, что интроны гена пролактина больше, чем интроны гена гормона роста и плацентарного лактогена. При выравнивании консервативных областей границы соединения экзонов с интронами совпадают у трех генов. Присутствие четырех высоко гомологичных областей во всех трех генах, служит доказательством дупликации участков предкового гена, в результате которой сформировался ген-предшественник Pre-GH/PL, состоящий уже из пяти экзонов и четырех интронов, из которого и эволюционировали современные ген гормона роста, пролактин и плацентарный лактоген. Некоторые авторы (Barta et al., 1981; Miller, Eberhardt, 1983) считают, что предковый, ген был меньше по размерам и имел примитивную структуру. Он состоял из регуляторной области, которая включала в себя элементы необходимые для экспрессии, и кодирующей последовательности. Кодирующая часть в последующем и дублировалась.

В результате дупликации возникли четыре одинаковых копии кодирующей последовательности, сформировавших один ген, что и объясняет наличие дублированных областей и гомологии аминокислот генов гормона роста, пролактина и плацентарного лактогена. Эти области внутренней гомологии расположены в пределах II,

IV и V экзонов. На пятый экзон приходится две гомологичные области. Возможно, между третьей и четвертой копиями предкового гена находился интрон, и он был удален. Появление двух гомологичных областей в экзоне пять объясняется еще одной версией: сначала возникли экзоны II и IV, а уже их транскрипт матричной РНК был повторно встроен за ними и таким образом сформировался пятый экзон. В результате этой вставки и мог возникнуть пятый экзон. Экзон III не имеет областей гомологии и авторы (Miller, Eberhardt, 1983) отмечают, что он возник отдельно с появлением новых функций, связанных с регуляцией углеводного обмена. И в результате из гена предшественника с такой структурой эволюционировали современные гены гормона роста, пролактина и плацентарного лактогена. На основании того, что внутренняя гомология наблюдается в нуклеотидных последовательностях II, IV и V экзонов считается (Barta et al., 1981; Miller, Eberhardt, 1983), что именно эти части гена возникли за счет дублирования меньшего сегмента предкового гена.

В эволюции гена гормона роста прослеживается нестандартная закономерность: медленная базальная скорость изменений последовательности прерывается периодами быстрой эволюции (Wallis, 1996; Wallis, 2014). У млекопитающих, включая сумчатых (Saunders et al., 1998), базальная скорость очень медленная ($0,21-0,28 \times 10^{-9}$ аминокислотных замен на сайт в год), а сама последовательность гена гормона роста очень консервативна и похожа на последовательность предкового гена. Например, у свиньи и собаки, которые относятся к разным отрядам млекопитающих, белковые последовательности гормона роста идентичны (Wallis, 1994). Уоллис отмечает (Wallis, 1994), что последовательность гена гормона роста свиньи, вероятно, имеет максимальное сходство с последовательностью предкового гена. С последовательностью гена гормона роста свиньи, также сходны последовательности альпаки, лошади, кита и слона (Manioui et al., 2004). Периоды быстрой эволюции гена гормона роста сопровождались увеличением скорости в 25–50 раз (Lioupiis et al., 1997; Wallis, Wallis, 2001). Первый период приходится на момент после расхождения полуобезьян и высших приматов, но до дивергенции узконосых и широконосых обезьян (последовательности гормона роста человека и макаки-резус очень похожи). Такие быстрые изменения гормона роста в линии ведущей к высшим приматам оказались особенно заметными, поскольку замены

составили около 35% всех аминокислотных остатков. Период быстрой эволюции у приматов происходил еще до дублирования генов, образовавших в дальнейшем кластер (Wallis, 1994; Wallis et al., 2001; Wallis, Wallis, 2001).

Второй период увеличения скорости эволюции отмечался после дивергенции мозолоногих и жвачных парнокопытных, но до того момента, как разошлись козлиные (Caprinae) и бычьи (Bovinae) (Lioupis et al., 1997). Последовательности гормона роста в этой группе животных имеют большое сходство между собой и отличаются от предковой последовательности гормона роста. Например, белковые последовательности гормона роста быка и овцы похожи между собой, но отличаются от последовательности гормона роста свиньи на 18 аминокислотных остатков (Lioupis et al., 1997).

Стоит отметить, что такое резкое увеличение скорости эволюции в значительной степени характерно только для экзонов. Для других структурных элементов: сигнального пептида, интронов, 5'- и 3'-регуляторных областей не отмечается всплеска быстрой эволюции (Lioupis et al., 1997; Wallis, Wallis, 2001). Изменения скорости, влияющие на последовательность аминокислот и не затрагивающие остальные части гена, могут указывать как на потерю функции, которую выполняет белок (ослабление отрицательного (очищающего) отбора), так и на расширение функции и действие положительного отбора (Wallis, Wallis, 2001). Поскольку основная функция гормона роста – стимуляция роста и развития – сохраняется на всех этапах эволюции позвоночных животных, аминокислоты, которые образуют основное гидрофобное ядро, остаются консервативными. А такая быстрая эволюция, в результате которой изменения коснулись небольших участков, например, ответственных за взаимодействие гормона роста с рецепторами, скорее указывает на коэволюцию гормона роста вместе с его рецептором (GHR), во время эволюции приматов или на адаптивный характер изменений и расширение функций гормона, но не на их потерю (Wallis et al., 2001).

Расширение функций гормона роста у приматов может быть связано с большей лактогенной активностью, которая не характерна для других млекопитающих, с изменением специфичности рецепторов, с появлением новых физиологических механизмов, связанных с плацентарными белками (Wallis 1997; Wallis et al., 2001). После появления у приматов дополнительных копий гена гормона роста, которые возникли в

результате дупликации и образовали целый кластер (Chen et al., 1989), произошло перераспределение функции, а за геном гормона роста сохранилась его основная функция, связанная с регуляцией роста. В результате появления дополнительных генов давление на ген гормона роста снизилось, а молекулярная скорость эволюции вернулась к базальному уровню. После периода быстрых изменений возвращение скорости эволюции к базальному уровню происходит и у парнокопытных (Lioupis et al., 1997; Wallis et al., 2001).

Еще один отряд млекопитающих, у которых было показано увеличение скорости эволюции гормона роста, это – грызуны (Wallis, Wallis, 1995). У грызунов и зайцеобразных скорость эволюции увеличивается незначительно, всего в два-три раза. Например, отличие последовательности кролика от последовательности свиньи всего на 7 аминокислотных остатков, говорит в пользу небольшого увеличения скорости (Wallis, Wallis, 1995).

У птиц скорость эволюции гена гормона роста оказалась немного быстрее ($0,56-0,80 \times 10^{-9}$ аминокислотных замен на сайт), чем скорость эволюции у млекопитающих в период, так называемой, «медленной фазы» ($0,21-0,28 \times 10^{-9}$ аминокислотных замен на сайт), но значительно ниже той скорости, которая характерна для «быстрых фаз» эволюции гена гормона роста млекопитающих (Buggiotti, Primmer, 2006). Стоит отметить, что ген гормона роста у птиц характеризуется более постоянной и относительно низкой скоростью молекулярной эволюции, которая увеличивается не больше, чем в 1,4 раза (Buggiotti, Primmer, 2006).

Однако и среди птиц встречаются группы, у которых эволюция нуклеотидных и аминокислотных последовательностей гормона роста протекает быстрее. Такие существенные изменения скорости можно наблюдать в отряде воробьинообразных (Yuri et al., 2008), который является одним из самых многочисленных отрядов птиц. В результате дупликации в геноме воробьинообразных представлены две копии гена гормона роста, которые сохраняются в процессе эволюции и обе являются функциональными. Юри с соавторами (Yuri et al., 2008) считают, что обе копии быстро эволюционировали с момента дупликации и подвергались давлению отбора разной интенсивности. Скорость эволюции нуклеотидных последовательностей у этих паралогичных генов больше чем в 2 раза, а темпы эволюции аминокислотных

последовательностей в 10 раз выше, чем в ближайших родственных группах птиц, не относящихся к воробьинообразным. Резкое увеличение скорости может иметь адаптивное значение, поскольку, находясь под разным давлением отбора и эволюционируя настолько быстро, гены-паралоги могут приобретать функциональные различия (Yuri et al., 2008).

У остальных представителей птиц (Buggiotti, Primmer, 2006) быстрые изменения характерны только для отдельных частей гена. И если у млекопитающих изменения затрагивали только экзоны, а другие структурные элементы гена, такие как сигнальный пептид, интроны, регуляторные последовательности не показывали быстрых изменений, то у птиц скорость замещения аминокислот в области сигнального пептида от 1 до 19 раз выше, чем показано для зрелого белка (Buggiotti, Primmer, 2006). А скорость эволюции сигнального пептида гормона роста превышала скорость эволюции зрелых белков в 2,0 – 4,4 раза. В результате высокой изменчивости сигнального пептида возникают различные посттрансляционные модификации. Они могут приводить к увеличению структурного разнообразия, а также оказывать влияние на скорость созревания белка. От скорости созревания гормона роста зависит и уровень его секреции, который потенциально может быть предметом естественного отбора. Контроль темпа роста или поддержание определенного темпа роста в пределах одного вида указывает на адаптивный характер изменений (Buggiotti, Primmer, 2006).

Медленная эволюция гормона роста характерна и для большинства низших позвоночных животных – $0,19-0,39 \times 10^{-9}$ аминокислотных замен в год (Wallis, 1996). Это сопоставимо с базальной скоростью эволюции млекопитающих ($0,21-0,28 \times 10^{-9}$). Однако и среди представителей низших позвоночных могли происходить вспышки быстрой эволюции. Увеличение скорости отмечается во время эволюции амфибий ($0,86 \pm 0,13 \times 10^{-9}$ аминокислотных замен на сайт в год), хрящевых ($0,54 \pm 0,10 \times 10^{-9}$ аминокислотных замен на сайт в год) и костистых рыб ($1,52 \pm 0,26 \times 10^{-9}$ аминокислотных замен на сайт в год). Такие периоды высокой скорости эволюции носят скорее случайный характер, потому что общая картина эволюции низших позвоночных животных похожа на ту, которая характерна для млекопитающих, где преобладает медленная базальная скорость молекулярной эволюции, со случайными длительными всплесками (Wallis, 1996).

1.5. Структура регуляторных участков гена гормона роста позвоночных

Ген гормона роста позвоночных экспрессируется только в специализированных клетках передней доли гипофиза – соматотрофах. Тканеспецифический характер экспрессии обеспечивается различными механизмами взаимодействия белковых факторов транскрипции с регуляторными последовательностями нуклеиновых кислот. В регуляторной (промоторной) области гена гормона роста расположены ТАТА-боксы, сайты связывания с различными факторами транскрипции, элементы ответа на гормоны и другие соединения. Располагаются эти последовательности вплотную к гену на участке длиной около 500 п.н. перед сайтом начала транскрипции. В результате взаимодействий транскрипционных факторов с регуляторными последовательностями формируется транскрипционный комплекс, который активирует транскрипцию и контролирует интенсивность, время и уровень экспрессии гена.

ТАТА-бокс, или бокс Хогнеса, представлен консервативной последовательностью (5'-ТАТААА-3'), располагается на расстоянии до –30 п.н. левее точки инициации транскрипции и встречается в генах гормона роста всех позвоночных животных (Barret et al., 2013). В гене гормона роста человека, крупного рогатого скота и грызунов ТАТА-бокс встречается в положении –26/–33 п.н. от точки инициации транскрипции и представлен последовательностью 5'-ТАТАААА-3' (Barta et al., 1981; DeNoto et al., 1981; Gordon et al., 1983). Среди парнокопытных встречаются виды, у которых ТАТА-бокс располагается дальше от стартового кодона (Wallis, Wallis, 2001). У представителей семейства оленьковые (Chevrotain) ТАТА-бокс находится на расстоянии –69/–75 п.н., а у благородного оленя (*Cervus elaphus*) ТАТА-бокс расположен на расстоянии от точки начала транскрипции в положении –93/–99 п.н. (Lioupiis et al., 1997). Несмотря на важную роль, которую играет ТАТА-бокс в инициации транскрипции, известны случаи, когда происходят изменения в консенсусной последовательности ТАТААА (Breathnach, Chambon, 1981; Bucher, 1990). У одногорбого верблюда (*Camelus dromedarius*) и обыкновенного бегемота (*Hippopotamus amphibius*) последовательность ТАТА-бокса не идентична той последовательности ТАТА-бокса, которая встречается в гене гормона роста большинства млекопитающих (Maniou et al., 2004). Несмотря на замену первого

нуклеотида в ТАТА-боксе одногорбого верблюда (*Camelus dromedarius*), ген гормона роста продолжает экспрессироваться, о чем свидетельствует выделение гормона роста из гипофиза одногорбого верблюда (Martinat et al., 1990).

Количество ТАТА-боксов в промоторах гена гормона роста позвоночных может отличаться у разных видов. В 5'-фланкирующей области гена гормона роста курицы было обнаружено два ТАТА-боксов (Tanaka et al., 1992). Первый ТАТА-бокс (ТАТАТАА) располагается на расстоянии –24 п.н. от точки инициации транскрипции, а второй ТАТА-бокс (ТАТАААТ) был найден в положении –145 п.н.. Авторы считают (Tanaka et al., 1992), что только первый ТАТА-бокс отвечает за связывание с транскрипционным фактором, поскольку занимает более проксимальное положение относительно точки начала транскрипции. Пока неясно, принимает ли второй ТАТА-бокс участие в сборке транскрипционного комплекса. С последовательностью ТАТА-боксов связываются транскрипционные факторы, привлекающие РНК-полимеразу к сайту начала транскрипции. Вначале ТАТА-бокс связывается с ТАТА-связывающимся белком (ТВР), который расплетает ДНК, а затем с ним уже связываются транскрипционные факторы TFIIID, а к его верхней и нижней части уже присоединяются транскрипционные факторы TFIIA и TFIIВ (Barret et al., 2013).

Участки промотора, которые отвечают за экспрессию гена гормона роста непосредственно в клетках гипофиза, представлены сайтами связывания с гипофиз-специфичным фактором транскрипции Pit-1 или GHF-1 (growth hormone factor 1), известным также как POU домен (POU1F1) (Ingraham et al., 1988; Mangalam et al., 1989; Karin et al., 1990). Обычно сайты связывания с транскрипционным фактором Pit-1 насыщены нуклеотидами А и Т и содержат последовательность ТАТНСАТ. Было показано, что для успешной активации транскрипции гена гормона роста необходимо взаимодействие транскрипционного фактора Pit-1 с несколькими сайтами связывания. У большинства позвоночных животных в 5'-фланкирующей области гена гормона роста имеется два сайта связывания с фактором Pit-1. В промоторе гена гормона роста млекопитающих положение и нуклеотидные последовательности этих сайтов достаточно консервативны (Theill, Karin, 1993; Krawczak et al., 1999). Например, в промоторной области гена гормона роста человека, первый сайт расположен от точки начала

транскрипции на расстоянии от -55 до -90 п.н., а второй от -106 до -131 п.н., соответственно (Walker, 1991). У оленьковых наблюдается смещение сайтов связывания с фактором Pit-1 ближе к точке начала транскрипции. Первый сайт связывания занимает положение от -5 до -13 п.н., второй сайт расположен на расстоянии от -44 до -51 п.н. (Lioupis et al., 1997).

Было показано (Jones et al., 1995), что двух сайтов связывания с транскрипционным фактором Pit-1 не всегда бывает достаточно для эффективной экспрессии гена гормона роста. Например, при трансфекции в оплодотворенные ооциты мыши гена гормона роста человека hGH-N с прилегающим промотором, содержащим только два сайта связывания с Pit-1, не удалось обнаружить следов экспрессии гена hGH-N у взрослых особей (Jones et al., 1995). В ряде работ было показано, что успешная экспрессия гена гормона роста человека hGH-N в системе *in vivo* наблюдается только в присутствии сверхчувствительных к ДНКазе I сайтов (Jones et al., 1995; Shewchuk et al., 2006).

Сверхчувствительные сайты (HS) представляют открытые, не упакованные в нуклеосомы участки ДНК в хроматине, которые необходимы для активации транскрипции и последующей высокоуровневой тканеспецифичной экспрессии. Поскольку гены гормона роста человека вместе с плацентарными лактогенами в процессе эволюции сформировали кластер, расположенных друг за другом генов, то по аналогии с кластером генов бета-глобина, участки, содержащие сверхчувствительные сайты, получили название – область, контролирующая локус (locus control region), сокращенно LCR (Jones et al., 1995). LCR можно рассматривать как один из классов цис-элементов, которые функционируют для изменения структуры хроматина. Подобное изменение структуры хроматина может приводить к активации или подавлению экспрессии генов. Область, контролирующая локус, расположена перед кластером генов гормона роста и охватывает участок от -15 т.п.н. до -32 т.п.н. (Jones et al., 1995; Shewchuk et al., 2006).

В состав LCR входят пять сверхчувствительных сайтов. Сверхчувствительные сайты HS I и HS II активны в хроматине клеток гипофиза, сайт HS IV активен в клетках плаценты, а сайты HS III и HS V могут быть активны как в клетках гипофиза, так и в клетках плаценты (Su et al., 2000; Shewchuk et al., 2006). Было показано, что сверхчувствительный сайт HS I, расположенный в районе -15 т.п.н., играет не только

важную роль в активации транскрипции гена гормона роста человека hGH-N, но и обеспечивает процесс ацетилирования гистонов H3/H4 на участке –32 т.п.н.. Подобная модуляция структуры хроматина способствует продолжительной высокоуровневой экспрессии гена hGH-N, которую не могут обеспечить сайты связывания с Pit-1, расположенные в прилегающем промоторе. В сверхчувствительном сайте HS I присутствует три сайта связывания с транскрипционным фактором Pit-1: Pit-1-A, Pit-1-B, Pit-1-C (Shewchuk et al., 2006). Нуклеотидная последовательность сайтов Pit-1-A–Pit-1-C отличается от последовательности сайтов Pit-1, ассоциированных с промотором.

Отличия нуклеотидных последовательностей не влияют на способность транскрипционного фактора Pit-1 взаимодействовать с сайтами расположенными в LCR и в промоторной области. Отсутствие строгой консервативной последовательности позволяет транскрипционному фактору Pit-1 менять конформацию для того, чтобы установить контакты между ДНК-связывающими доменами и нуклеотидными последовательностями. Сигналом к изменению конформации фактора Pit-1 может служить присутствие разных нуклеотидов в одном и том же положении: сайты связывания с транскрипционным фактором Pit-1, расположенные в промоторе, содержат тимин, а в сайтах Pit-1 сверхчувствительного сайта HS I в этом же положении находится аденин. Взаимодействие факторов с остальными сверхчувствительными сайтами, вероятно, подавляет экспрессию гена гормона роста человека hGH-N в клетках плаценты – лактотрофах (Shewchuk et al., 2006).

Для генов гормона роста и плацентарных лактогенов человека и высших приматов характерна высокая степень гомологии нуклеотидных последовательностей (более 94%) (Su et al., 2000). Несмотря на такое сходство экспрессируются гены в разных тканях. Высокоуровневую тканеспецифичную экспрессию обеспечивают различные регуляторные комплексы. Экспрессия плацентарных лактогенов регулируется так называемыми P-последовательностями (Norquay et al., 2003; Norquay et al., 2006). P-последовательности подобно сверхчувствительным сайтам LCR (Jones et al., 1995) обеспечивают оптимальный уровень экспрессии плацентарных лактогенов в клетках плаценты и блокируют их экспрессию в клетках гипофиза – соматотрофах. Располагаются P-последовательности в регуляторных областях альтернативного гена гормона роста

(hGH-V) и плацентарных лактогенов (hPL-1, hPL-2, hPL-3) на расстоянии 2000 п.н. от точки начала транскрипции (Norquay et al., 2003). Норквей с соавторами (Norquay et al., 2006) показали, что Р-последовательность содержит три сайта связывания с транскрипционными факторами: PSE-A, PSE-B и PSE-C. С сайтом PSE-B связывается транскрипционный ядерный фактор-1 (nuclear factor-1, NF-1). NF-1 относится к универсальным факторам и принимает участие в регуляции экспрессии большинства генов эукариот (Norquay et al., 2003). В случае с плацентарными лактогенами и альтернативным геном гормона роста hGH-V, фактор NF-1 выступает в роли фактора, обеспечивающего тканеспецифическую экспрессию генов в клетках плаценты. Сайты PSE-A и PSE-C взаимодействуют с факторами, которые формируют комплекс, подавляющий экспрессию плацентарных лактогенов в клетках гипофиза. Транскрипционный ядерный фактор гепатоцитов-3 α (HNF-3 α), связываясь с сайтом PSE-C, инициирует взаимодействие регуляторного фактора RFX-1 (regulatory factor X1) с фактором NF-1 и с сайтом PSE-A. Сформированный комплекс блокирует процесс ацетилирования гистонов на участках хроматина, и экспрессия плацентарных лактогенов в соматотрофах не происходит.

Обеспечивать тканеспецифичную экспрессию гена гормона роста человека и плацентарных лактогенов могут не только отдельные регуляторные элементы, но и изменения нуклеотидной последовательности прилегающего промотора. Например, отсутствие экспрессии плацентарных лактогенов в гипофизе может отражать изменения в последовательности, с которой связывается транскрипционный фактор Pit-1. В плаценте из-за отсутствия транскрипционного фактора Pit-1 ген гормона роста может не экспрессироваться. В гене hGH-V, который кодирует альтернативный гормон роста, наблюдается пять нуклеотидных замен в проксимальном сайте Pit-1 по сравнению с геном hGH-N. Возможно, именно с этими заменами и с участием других транскрипционных факторов связано то, что уровень экспрессии гена hGH-V в плаценте ниже, чем уровень экспрессии гена hGH-N в клетках гипофиза (Perez-Maya et al., 2016).

В промоторах всех пяти генов, входящих в кластер, был найден сайт связывания с транскрипционным фактором Sp1 (Jones et al., 1995). Транскрипционный фактор Sp1 также известный как белок, связывающийся с GC-боксом и аналогичными

последовательностями, относится к семейству транскрипционных факторов Sp/KLF (specificity protein/kruppel-like factor), представлен в больших количествах, участвует в регуляции транскрипции множества генов (Briggs et al., 1986; Kadonaga et al., 1987), в том числе и генов, экспрессия которых происходит в гипофизе и плаценте. Джонс с соавторами отмечают (Jones et al., 1995), что последовательность сайта связывания с транскрипционным фактором Sp1 частично перекрывается с последовательностью дистального сайта связывания с транскрипционным фактором Pit-1. Возможно, что факторы Pit-1 и Sp1 взаимодействуют с промоторами плацентарных лактогенов и генов гормона роста взаимноисключающим образом. Например, в клетках гипофиза поддерживается высокая концентрация фактора Pit-1, который связывается с двумя сайтами в промоторе гена гормона роста и ингибирует действие фактора Sp1. В условиях нехватки транскрипционного фактора Pit-1 фактор Sp1 связывается с дистальным сайтом фактора Pit-1, что позволяет поддерживать транскрипцию плацентарных лактогенов PL и гена гормона роста GH-V в клетках плаценты на оптимальном уровне (Lemaigre et al., 1989; Barrera-Saldana, 1998).

Сайт связывания с транскрипционным фактором Sp1 является далеко не единственным сайтом в промоторе гена гормона роста, расположенным вблизи сайтов Pit-1. Между проксимальным и дистальным сайтами связывания с Pit-1 находится сайт связывания с транскрипционным фактором Zn15, получивший название Z-бокс. Фактор Zn15 относится к семейству белков, со структурным доменом типа «цинковые пальцы». Сравнение нуклеотидных последовательностей промотора человека, крупного рогатого скота и грызунов показало, что Z-бокс встречается у всех исследованных видов, и находится на расстоянии -95/-110 п.н. от точки начала транскрипции. Положение и нуклеотидная последовательность Z-бокса достаточно консервативны (Lipkin et al., 1993). Такое расположение Z-бокса в промоторе гена гормона роста обеспечивает взаимодействие факторов транскрипции, которые связываются с сайтами Pit-1 и Z-боксом, и способствует эффективной экспрессии гена гормона роста.

Транскрипционные факторы, взаимодействуя с последовательностью промотора, могут как активировать, так и подавлять процесс транскрипции. Участки промотора, при связывании с которыми транскрипционных факторов экспрессия гена прекращается,

получили название отрицательных регуляторных элементов (NRE). Отрицательные регуляторные элементы (NRE) сохраняются в промоторе гена гормона роста большинства млекопитающих (Wallis et al., 2001), исключение составляет ген гормона роста собаки (Maniou et al., 2004). Стоит отметить, что связывание транскрипционного фактора с NRE не подавляет процесс транскрипции полностью, а поддерживает его на оптимальном уровне. В промоторе гена гормона роста NRE выполняет функцию репрессора избыточной транскрипции. В экспериментах с промотором гена гормона роста крупного рогатого скота (bGH), который был соединен с репортерным геном хлорамфеникол-ацетилтрансферазы (CAT), было показано, что после удаления участка промотора, содержащего NRE, экспрессия репортерного гена в клетках HeLa увеличивалась в 5 раз, а в клетках GH₃ крысы в 2,2 раза (Park, Roe, 1996). В промоторе гена гормона роста быка было найдено три отрицательных регуляторных элемента (Park, Roe, 1996). NRE1 занимает положение -316/-308 п.н., NRE2 находится на расстоянии -259/-251 п.н., NRE3 расположен ближе всего к точке инициации транскрипции и занимает положение -117/-108 п.н.. Консенсусная последовательность для трех элементов NRE состоит из девяти нуклеотидов (5'-AANATGGNG/T-3'), и имеет консервативную область из четырех нуклеотидов ближе к 5'-концу. У других представителей млекопитающих, которые принадлежат к разным отрядам, был описан только NRE3 и было показано, что он частично перекрывается с дистальным сайтом связывания Pit-1 (dPit-1) (Lioupis et al., 1997; Wallis et al., 2001; Maniou et al., 2004). С последовательностями NRE, вероятно, связывается многофункциональный транскрипционный фактор YY1 (Park, Roe, 1996).

Еще одним элементом, который присутствует в большинстве промоторов гена гормона роста позвоночных животных, является элемент ответа на циклический АМФ (цАМФ, CRE). Элемент CRE млекопитающих содержит либо асимметричный TGACG-мотив, либо симметричный (палиндромный) TGACGTCA-мотив (Montminy et al., 1990). У некоторых приматов (малый лори) (Wallis et al., 2001), парнокопытных (Maniou et al., 2004) и грызунообразных (кролики) (Wallis O., Wallis M., 1995) в ключевом мотиве CRE (CGTCA) имеется ряд замен, из-за которых CRE не может функционировать и участвовать в регуляции транскрипции. С элементом ответа на цАМФ связывается транскрипционный фактор CREB. CREB принадлежит к семейству белков, связывающихся с CRE

((CREB)/активаторы фактора 1 (ATF1)), которые взаимодействуют с CRE и после фосфорилирования протеинкиназой A (PKA) активируют транскрипцию, взаимодействуя с белками, связывающимися с CREB (CBP) (Montminy et al., 1990; Arany et al., 1994; Kwok et al., 1994). Регуляция транскрипции с участием CREB может проходить и в отсутствие элементов ответа на цАМФ. У крыс это связано с появлением CREB-независимого от цАМФ механизма стимуляции генов, содержащих сайты Pit-1 (Xu et al., 1998). Белки, связывающиеся с CREB (CBP) взаимодействуют непосредственно с сайтами Pit-1 и являются кофакторами транскрипционного фактора Pit-1 (Zanger et al., 1999; Argenton et al., 2002).

Универсальные факторы тоже участвуют в регуляции транскрипции гена гормона роста. В промоторе гена гормона роста человека на расстоянии $-267/-289$ п.н. от точки начала транскрипции был обнаружен сайт связывания сразу с двумя факторами: AP-2 (activating protein 2) и NF-1 (nuclear factor-1) (Walker, 1991). Факторы AP-2 и NF-1 представляют целые семейства факторов, которые участвуют в регуляции экспрессии большинства генов эукариот. Стоит отметить, что взаимодействуют факторы с сайтом связывания в промоторе гена гормона роста взаимоисключающим образом. Это значит, что фактор AP-2 связывается с областью промотора, только когда в клетках отсутствует фактор NF-1. Несмотря на то, что домены транскрипционных факторов AP-2 и NF-1, ответственные за связывание с ДНК, не имеют общих структурных особенностей, наличие областей богатых пролином позволяет предположить (Courtois et al., 1990), что AP-2 и NF-1 взаимодействуют с определенным компонентом транскрипционного механизма, который является общим для этих факторов.

Помимо факторов транскрипции в регуляции экспрессии генов гормона роста позвоночных могут принимать участие гормоны (гормоны щитовидной железы и глюкокортикоиды) (Truss, Veato 1993). Регуляция транскрипции осуществляется за счет связывания рецепторов гормонов с соответствующими сайтами – элементами ответа на гормоны (HRE), которые могут быть рассеяны по всему гену, и встречаться не только в промоторной области, но и в интронах (Yang et al., 1997).

Элемент ответа на тиреоидные гормоны (TRE) может представлять собой прямой, а иногда и инвертированный повтор двух гексамерных полусайтов (half-sites) (AGGTCA),

разделённых спейсером размером от трех до пяти нуклеотидов. (Leidig et al., 1992; Wu et al., 2001). В промоторе гена гормона роста крыс последовательность TRE представлена двумя копиями несовершенного прямого повтора, которые отделены друг от друга четырьмя нуклеотидами (5'-gaaAGGTAAgatcAGGGACg-3') (Koenig et al., 1987). Располагается элемент ответа на тиреоидные гормоны немного дальше дистального сайта Pit-1: у крыс в положении -173/-192 п.н. от точки начала транскрипции (Koenig et al., 1987), у мышей в положении -164/-179 п.н. от точки начала транскрипции (Das et al., 1996).

Рецептор тиреоидного гормона (TR) может связываться с TRE как мономер, или как гетеродимер совместно с рецептором ретиноевой кислоты (RXR). Стоит отметить, что участие в активации транскрипции гетеродимера TR/RXR повышает эффективность экспрессии генов. У человека тиреоидный гормон оказывает разное влияние на экспрессию генов гормона роста и плацентарных лактогенов (Walker, 1991). Поскольку транскрипционный фактор Pit-1 не участвует в активации транскрипции плацентарных лактогенов, его функции могут выполнять другие факторы. Одним из таких факторов и становится TR. При взаимодействии с рецептором тиреоидного гормона экспрессия плацентарных лактогенов увеличивается в четыре раза, а экспрессия гена гормона роста только в два (Leidig et al., 1992). Отличаются и последовательности TRE между плацентарными лактогенами и геном гормона роста. Последовательность TRE среди плацентарных лактогенов остается более консервативной, чем в гене гормона роста GH-N (Perez-Maya et al., 2016).

Консенсусная последовательность для глюкокортикоидов (GRE) представляет собой несовершенный инвертированный повтор GGTACAnnnTGTTCT, где правая часть повтора – гексануклеотид TGTTCT – более консервативна, чем левая (Beato et al., 1989). Рецептор глюкокортикоидов связывается с GRE в виде гомодимера. При этом с каждым из гексануклеотидов инвертированного повтора взаимодействует одна молекула рецептора (Truss, Beato, 1993). Элемент ответа на глюкокортикоиды (GRE) встречается в первом интроне гена гормона роста большинства млекопитающих, парнокопытных, но никаких следов последовательности не было найдено в гене гормона роста грызунообразных (крыса, кролик) (Wallis O., Wallis M., 1995). Глюкокортикоиды и

тиреоидные гормоны являются важными физиологическими регуляторами транскрипции, при их совместном действии уровень экспрессии гена гормона роста повышается в несколько раз (Martial et al., 1977).

1.5.1. Структура регуляторных участков гена гормона роста у рыб

Несмотря на то, что нуклеотидные последовательности гена гормона роста получены для большого числа рыб (Ber, Daniel, 1992; Male et al., 1992; Zhu et al., 1992; Du et al., 1993; Hong, Scharl, 1993; Venkatesh, Brenner, 1997; Rajesh, Majumdar, 2007; Sekar et al., 2014), исследования механизмов, участвующих в инициации и регуляции транскрипции, ограничены отдельными представителями из нескольких отрядов (Ohkubo et al., 1996; Bernardini et al., 1999; Sekkali et al., 1999; Almuly et al., 2005; Von Schalburg et al., 2008). В промоторе гена гормона роста рыб представлены сайты связывания с такими же транскрипционными факторами, гормонами и другими регуляторными элементами, как и у большинства позвоночных животных. Однако встречаются сайты связывания с транскрипционными факторами, которые у других позвоночных не были описаны. Объясняется это тем, что гормон роста помимо основной функции стимуляции роста выполняет функции, связанные с метаболизмом углеводов и липидов, регуляцией водно-солевого баланса и процессом размножения, а также принимает участие в адаптации к различным водным средам при миграции из морского водоема в пресный и наоборот (Almuly et al., 2005).

ТАТА-бокс присутствует во всех известных промоторных областях гена гормона роста рыб (Ber, Daniel, 1992; Devlin, 1993; Ohkubo et al., 1996; Almuly et al., 2005). Занимает консервативное положение в промоторе, располагается на расстоянии $-23/-25$ п.н. от точки начала транскрипции; нуклеотидная последовательность (ТАТААА) идентична среди представителей разных отрядов рыб (Sekkali et al., 1999; Almuly et al., 2005). В промоторах генов гормона роста радужной форели (*Oncorhynchus mykiss*) rtGH1 и rtGH2 было обнаружено еще два ТАТА-боксов, расположенных гораздо дальше от точки начала транскрипции (Yang et al., 1997). Один из этих ТАТА-боксов (ТАТААТА) находится в положении $-399/-393$ п.н. в гене rtGH1 и $-402/-398$ п.н. в гене rtGH2 соответственно. Второй ТАТА-бокс (ТАТААСА) в гене rtGH1 находится на расстоянии $-657/-651$ п.н. от

точки начала транскрипции, а в гене *rtGH2* на расстоянии –615/–611 п.н.. Функция этих двух консенсусных последовательностей ТАТА еще пока не ясна (Yang et al., 1997).

Как и у большинства позвоночных животных, в промоторе гена гормона роста рыб, присутствуют сайты связывания с транскрипционным фактором Pit-1, ответственным за тканеспецифичную экспрессию гена гормона роста. По два сайта связывания с Pit-1 было обнаружено в промоторе у желтохвоста (*Seriola quinqueradiata*) (Almuly et al., 2000) и азиатского паралихта (*Paralichthys olivaceus*) (Tanaka et al., 1995), в остальных известных промоторах гена гормона роста рыб количество сайтов связывания с Pit-1 варьирует. В промоторе желтохвоста (*S. quinqueradiata*) первая последовательность (ТТСААСАТ) располагается в положении –119/–112 п.н., вторая последовательность (АССТССАТ) расположена на антисмысловой цепи в положении –112/–105 п.н. (Ohkubo et al., 1996). Последовательности двух сайтов связывания с Pit-1 у азиатского паралихта (*P. olivaceus*) обладают достаточно высокой гомологией (78%) с аналогичными сайтами промотора радужной форели и находятся в положении от –70 до –53 п.н. и от –133 до –144 п.н., соответственно (Tanaka et al., 1995).

У мозамбикской тиляпии (*Oreochromis mossambicus*) было обнаружено несколько предполагаемых последовательностей, связывающихся с Pit-1 (Sekkali et al., 1999). Последовательности отличались не больше чем на 1 или 2 нуклеотида от консенсусной последовательности 5'-А(А/Т)(А/Т)ТАТНСАТ-3', и были фланкированы с 5'-конца областью богатой нуклеотидами А/Т. Рядом с ТАТА-боксом была обнаружена высококонсервативная последовательность (5'-GATGAАТТТАААСАТ-3') в положении –56/–42 п.н., которая встречается в большинстве промоторов других рыб и содержит потенциальный сайт для связывания с Pit-1, ориентированный в обоих направлениях (Sekkali et al., 1999). У золотистого спара (*Sparus aurata*) и латеса (*Lates calcarifer*) помимо основных двух сайтов связывания с Pit-1 (5'-АТGAАТТТААТСАТ-3'), которые являются частью палиндромного повтора из 20 нуклеотидов, в дистальной части промотора встречается и несколько дополнительных сайтов связывания (Almuly et al., 2005). У всех исследованных представителей отряда окунеобразных расстояние между ТАТА-боксом и сайтом связывания с транскрипционным фактором Pit-1 составляет 12 нуклеотидов. При

сравнении с промоторами представителей из других отрядов такой закономерности не выявлено (Yowe, Epping 1995; Almuly et al., 2005).

Наибольшее число сайтов связывания с транскрипционным фактором Pit-1 было обнаружено в 5'-фланкирующей области гена гормона роста лососевых рыб (Yamada et al., 1993; Argenton et al., 1996; Yang et al., 1997; Bernardini et al., 1999; Von Schalburg et al., 2008). При исследовании транскрипционного фактора Pit-1 радужной форели (*O. mykiss*) Ямада с соавторами показал (Yamada et al., 1993), что в промоторе гена гормона роста форели с фактором Pit-1 связываются не два сайта, как у млекопитающих, а четыре. Все четыре сайта (G1 -238/-220 п.н.; G2 -182/-149 п.н.; G3 -145/-113 п.н.; G4 -66/-15 п.н.), расположены достаточно компактно перед точкой начала транскрипции, на участке длиной 300 п.н.. На этом участке встречаются сайты связывания и с другими факторами транскрипции, которые вместе с сайтами Pit-1 составляют активный центр промотора (Yamada et al., 1993). Несколько сайтов связывания с Pit-1, было обнаружено у атлантического лосося (*Salmo salar*) и чавычи (*Oncorhynchus tshawytscha*) (Von Schalburg et al., 2008). Стоит отметить, что консенсусная последовательность сайта Pit-1 у рыб отличается от консенсусной последовательности, которая встречается у млекопитающих (T/A)(T/A)TATNCAT. Отличия в консенсусной последовательности для сайта Pit-1 встречаются и у птиц. На основании исследований (Tanaka et al., 1992; Yamada et al., 1993) было предложено (Ohkubo et al., 1996) ввести новую консенсусную последовательность ((T/A)NCTNCAT), для сайтов связывания Pit-1 у костистых рыб и птиц.

У костистых рыб также, как и у млекопитающих, цАМФ принимает непосредственное участие в регуляции транскрипции. В промоторе рыб элемент ответа на цАМФ (CRE) представлен асимметричной последовательностью TGACG. В промоторе гена гормона роста латеса (*L. calcarifer*) (Almuly et al., 2000), также, как и в промоторе форели (*O. mykiss*) (Argenton et al., 1996), этот элемент расположен между двумя функциональными сайтами связывания с транскрипционным фактором Pit-1, но относительно точки начала транскрипции элемент ответа на цАМФ у латеса и форели занимает разное положение. Положение проксимального сайта связывания с цАМФ полностью совпадает у рыб, относящихся к одному отряду окунеобразные (Almuly et al., 2005): золотистый спар (*S. aurata*), латес (*L. calcarifer*) и желтохвост (*S. quinquerediata*). В

промоторе у золотистого спара (*S. aurata*) и желтохвоста (*S. quinquerediata*) были обнаружены два дополнительных CRE элемента, которые не сопряжены с сайтами Pit-1. Вероятно, что у этих видов взаимодействие с цАМФ происходит не напрямую, как было показано для форели (*O. mykiss*) (Argenton et al., 1996), а с использованием других механизмов. Так, в промоторе мозамбикской тиляпии (*O. mossambicus*) не было обнаружено ни одной последовательности CRE, а белки, связывающиеся с CREB, принимают участие в транскрипции, взаимодействуя непосредственно с сайтами Pit-1 (Sekkali et al., 1999). (Согласно современной систематике (Betancur-R et al., 2017), перечисленные виды относятся к самостоятельным отрядам: золотистый спар (*S. aurata*) – отряд спарообразные; желтохвост (*S. quinquerediata*) – отряд ставридообразные; мозамбикская тиляпия (*O. mossambicus*) – отряд циклообразные).

Вместе с тканеспецифичным фактором Pit-1, непосредственное участие в регуляции транскрипции гена гормона роста у рыб принимают глюкокортикоиды, рецепторы ретиноевой кислоты и тиреоидного гормона. Регуляция транскрипции осуществляется за счет связывания комплекса рецепторов гормонов с элементами ответа на гормоны (HRE), которые встречаются не только в промоторе, но и в других частях гена (Beato et al., 1989). В промоторе гена гормона роста радужной форели (*O. mykiss*) Бернардини с соавторами (Bernardini et al., 1999) была описана последовательность элемента ответа на глюкокортикоиды (GRE) (TATACTnnnTGTTCC). Она отличается дистальным гексануклеотидом от последовательности GRE (GCTACAnnnTGTTCT), описанной для человека и некоторых приматов, но имеет частичное сходство с последовательностями которые взаимодействуют с глюкокортикоидами в промоторе гена гормона роста крысы и других представителей млекопитающих. Проксимальная часть последовательности GRE (TGTTCT) в промоторе радужной форели (*O. mykiss*) частично перекрывается с сайтом F2 транскрипционного фактора Pit-1, а дистальная часть GRE отделена всего тремя нуклеотидами от элемента ответа на цАМФ. Таким образом, данная область может представлять собой составной энхансер, который обеспечивает полноценную Pit-1-зависимую базальную транскрипцию гена гормона роста радужной форели и регулирует его экспрессию в зависимости от физиологических потребностей (Bernardini et al., 1999; Argenton et al., 2002). У золотистого спара (*S. aurata*) обнаружены

только две проксимальные части элемента GRE (Almuly et al., 2005). Расстояние между ними составляет 20 нуклеотидов, а относительно точки начала транскрипции они находятся гораздо дальше, чем в промоторе радужной форели (*O. mykiss*). Сайтов связывания с цАМФ рядом с элементами GRE обнаружено не было, вероятно, у золотистого спара (*S. aurata*) цАМФ и глюкокортикоиды не оказывают взаимного эффекта на регуляцию транскрипции (Almuly et al., 2005).

Элементы ответа на ретиноевую кислоту (RARE и RXRE) состоят обычно из двух прямых повторов (AGGTCA), разделенных пятью или двумя нуклеотидами, с которыми связывается рецептор ретиноевой кислоты. Рецептор ретиноевой кислоты может быть представлен как в форме гомодимера (RXR/RXR), так и в форме гетеродимера (RAR/RXR) (Blomhoff, Blomhoff, 2006). Сайт связывания с гетеродимерной формой рецептора ретиноевой кислоты (RAR/RXR) (5'-TCACATacaaaaacAGGTCC-3') встречается в промоторе генов *gh1* и *gh2* атлантического лосося (*Salmo salar*), чавычи (*Oncorhynchus tshawytscha*) (Von Shalburg et al., 2008) и радужной форели (*O. mykiss*) (Yang et al., 1997) в положении -106/-125 п.н. от точки инициации транскрипции.

Последовательность TRE, с которой взаимодействует рецептор тиреоидного гормона, не была описана для промоторов гена гормона роста рыб. Однако в экспериментах по оценке гормон-зависимой транскрипции в первичной культуре клеток гипофиза карпа при одновременной трансфекции рецепторов тиреоидного гормона, ретиноевой кислоты и гена люциферазы, находящегося под контролем промотора гена гормона роста атлантического лосося, удалось зафиксировать достаточно высокий уровень экспрессии репортерного гена (Farchi-Pisanty et al., 1997). В присутствии двух форм рецептора тиреоидного гормона, TR α и TR β , экспрессия гена люциферазы увеличивалась в два и десять раз, соответственно. А при воздействии на клетки трийодтиронином активность экспрессии увеличивалась в пять раз в присутствии рецептора TR α и в тридцать пять раз в присутствии рецептора TR β (Farchi-Pisanty et al., 1997). При совместном действии рецепторов тиреоидного гормона и ретиноевой кислоты (RAR α и RAR γ) экспрессия гена люциферазы увеличивалась в тридцать пять раз при взаимодействии рецепторов TR β /RAR α и TR β /RAR γ и в шестнадцать раз при взаимодействии промотора с рецепторами TR α /RAR α и TR α /RAR γ , соответственно

(Sternberg, Moav, 1999). Взаимодействие с рецепторами тиреоидных гормонов (TR) указывает на присутствие элемента ответа TRE в промоторе рыб, а увеличение уровня экспрессии при совместном воздействии рецепторов TR и RAR подтверждают, что гетеродимеры TR/RXR являются наиболее активной формой при регуляции транскрипции (Kliwer et al., 1992; Sternberg, Moav, 1999).

В промоторе гена гормона роста золотистого спара (*S. aurata*) (Almuly et al., 2005) встречаются предполагаемые сайты связывания с факторами N-Oct-3, HNF-3 и MEF-2, которые не являются тканеспецифичными и участвуют в регуляции транскрипции большинства генов. Авторы объясняют присутствие в промоторе сайтов связывания с этими факторами тем, что экспрессия гена гормона роста у рыб не ограничивается только гипофизом, а может протекать и в других тканях: в жабрах, сердце, кишечнике и в лейкоцитах периферических кровеносных сосудов. Экспрессия гена гормона роста в этих тканях может отличаться от экспрессии в гипофизе и не зависеть от Pit-1, а присутствие в промоторе сайтов связывания с другими транскрипционными факторами позволяет регулировать транскрипцию в этих тканях. Встречаются в промоторе золотистого спара (*S. aurata*) микросателлитные и инвертированные повторы (Almuly et al., 2005). Динуклеотидный повтор (CA)₁₄, расположен на расстоянии -90 п.н. от начала транскрипции, рядом с предполагаемым сайтом связывания с Pit-1. Самая дистальная часть промотора содержит две инвертированные последовательности длиной 60 п.н., расположенные на расстоянии -1275 п.н. и -1645 п.н.. Этот инвертированный повтор имеет потенциал для формирования стабильной структуры стержня из 60 п.н. и петли из 311 п.н. (Almuly et al., 2005).

Сайты связывания с транскрипционными факторами встречаются у рыб не только в промоторной области, но и в интронах гена гормона роста. В третьем интроне (интрон С) у радужной форели (*O. mykiss*) в составе двух АТ-богатых областей были обнаружены консенсусные последовательности (A/T₃NCAT), характерные для сайтов связывания с Pit-1 (Yang et al., 1997; Bernardini et al., 1999). У разных видов рода лабео (*Labeo*) количество сайтов связывания с Pit-1, найденное в третьем интроне, варьирует (Rajesh, Majumdar, 2007). В четвертом интроне (интрон D) у атлантического лосося (*S. salar*) и чавычи (*O. tshawytscha*) был обнаружен элемент ответа на цАМФ, который отличается от

элемента CRE в промоторе гена гормона роста, и представлен палиндромной последовательностью (ACTGCAGT) (Von Schalburg et al., 2008). Элемент ответа на эстрогены (ERE) был идентифицирован только в одном гене гормона роста (rtGH1) радужной форели (*O. mykiss*), в интроне С (Yang et al., 1997). Однако у атлантического лосося (*S. salar*) и чавычи (*O. tshawytscha*) элемент ответа на эстрогены встречается в промоторной области гена *gh2* (Von Schalburg et al., 2008). В своей работе Янг (Yang et al., 1997) приводит последовательности, ответственные за связывание с рецепторами тиреоидного гормона, ретиноевой кислоты, прогестерона, андрогена, минералокортикоидов, которые встречаются в интронах гена гормона роста. Но в тоже время положение этих сайтов в интронах, функция, которую они выполняют, пути их взаимодействия между собой и влияние друг на друга автором описаны не были (Yang et al., 1997).

Таким образом, гены гормона роста исследованы во многих группах позвоночных животных. Структура генов в целом сходна, хотя имеются различия. Наиболее интересным нам представляется тот факт, что у лососевых рыб имеется два гена. Учитывая, что эта группа эволюционировала в течении длительного времени (25–100 млн. лет назад), можно было предполагать, что эти гены могли субфункционализироваться. Приобретение новых функций должно было сопровождаться изменением механизмов регуляции экспрессии генов, что наиболее вероятно обусловлено изменениями в регуляторных элементах (Каменская, Брыков, 2020). Сравнительному анализу регуляторных участков генов и посвящена эта работа.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ

2.1. Объекты исследования

В качестве объектов исследования были выбраны четыре вида гольцов рода *Salvelinus* из северо-западной части Тихого Океана (табл. 2.1). Материал для исследования был взят из коллекции препаратов фиксированных тканей лаборатории генетики Национального научного центра морской биологии им. А. В. Жирмунского ДВО РАН. В работе было использовано по одной особи от каждого вида. Ниже приводиться полное систематическое положение рода исследуемых видов гольцов рода *Salvelinus*.

Тип **Chordata** – Хордовые

Подтип **Vertebrata** – Позвоночные

Инфратип **Gnathostomata** – Челюстноротые

Группа **Osteichthyes** – Костные рыбы

Класс **Actinopterygii** – Лучеперые рыбы

Подкласс **Neopterygii** – Новоперые рыбы

Инфракласс **Teleostei** – Костистые рыбы

Надотряд **Protacanthopterygii**

Отряд **Salmoniformes** – Лососеобразные

Семейство **Salmonidae** – Лососевые

Подсемейство **Salmoninae** – Лососевые

Род *Salvelinus* Richardson, 1836 – Гольцы

S. curilus (синоним *S. malma krascheninnikovi*) Pallas, 1814 – южная азиатская мальма

S. levanidovi Chereshev, Skopets, Gudkov, 1989 – голец Леванидова

S. malma Walbaum, 1792 – северная мальма

S. taranetzi Kaganowsky, 1955 – голец Таранца

Таблица 2.1 – Места сбора материала

| Вид | Место сбора | Географическое положение | Бассейн | Широта/Долгота |
|------------------------------|-------------------|-------------------------------------|----------------------------------------------------------------|-----------------------------------------------------|
| <i>Salvelinus malma</i> | Оз. Крокур | П-ов Камчатка, Азия | Зал. Кроноцкий, бассейн Тихого океана | 54°41'59.8"N 160°22'00.2"E 54.699950, 160.366727 |
| <i>Salvelinus curilus</i> | Руч. Анива | О. Сахалин, юго-восточное побережье | Бассейн Охотского моря | 47°10'18.0"N 143°01'52.2"E 47.171652, 143.031156 |
| <i>Salvelinus levanidovi</i> | Р. Яма | Материковое побережье, Азия | Ямская губа, бассейн Охотского моря | 59°36'54.8"N 154°12'11.2"E 59.615215, 154.203101 |
| <i>Salvelinus taranetzi</i> | Оз. Пекульнейское | Материковое побережье, Азия | Мейныпильгинская озерно-речная система, бассейн Берингова моря | 62°36'41.0"N 177°12'50.7"E 62.611393, 177.214086 |

Нуклеотидные последовательности регуляторных областей, четырех видов гольцов рода *Salvelinus*, были добавлены в GeneBank NCBI к уже размещенным там ранее нуклеотидным последовательностям структурной части гена гормона роста (экзоны и интроны) (табл. 2.2).

Для сравнительного анализа были использованы нуклеотидные последовательности генов гормона роста других видов лососевых рыб, представленных в базе данных GenBank (NCBI) для которых известны 5'- и 3'-фланкирующие области (табл. 2.3).

Таблица 2.2 – Номера последовательностей генов *gh* исследованных видов гольцов в GenBank

| Вид | Ген | |
|------------------------------|------------|------------|
| | <i>gh1</i> | <i>gh2</i> |
| <i>Salvelinus malma</i> | KF772972 | KF772976 |
| <i>Salvelinus curilus</i> | KF772971 | KF772975 |
| <i>Salvelinus levanidovi</i> | KF772974 | KF772978 |
| <i>Salvelinus taranetzi</i> | KF772973 | KF772977 |

Таблица 2.3 – Номера последовательностей генов *gh* лососевых из GeneBank

| Вид | Ген | |
|---------------------------------|------------|------------|
| | <i>gh1</i> | <i>gh2</i> |
| <i>Oncorhynchus keta</i> | | L04688 |
| <i>Oncorhynchus mykiss</i> | | J03797 |
| <i>Oncorhynchus nerka</i> | U14551 | U14535 |
| <i>Oncorhynchus tshawytscha</i> | EU621900 | EU621901 |
| <i>Salmo salar</i> | EU621898 | EU621899 |

2.2. Выделение ДНК

Тотальную ДНК выделяли методом хлороформной экстракции из тканей сердца, фиксированных в 96%-ном этаноле (Заславская и др., 2009). Кусочки ткани гомогенизировали в 500 мкл лизирующего буфера, содержащего 100 мМ трис-НСl рН = 7.5, 100 мМ NaCl, 50 мМ ЭДТА рН = 8.0, 50 мкл 10% додецилсульфата натрия и 1,5 мкл протеиназы К (в концентрации 10 мг/мл) и инкубировали при 54°C 2–3 часа до полного просветления лизата. Затем добавляли 1,5 мкл рибонуклеазы А (в концентрации 10 мг/мл) и инкубировали при 37°C в течение 40 мин. Дважды депротенизировали образцы равным объемом хлороформа 10–15 мин. Центрифугировали 8 мин при 10 000 об/мин. Верхнюю фракцию отбирали, затем осаждали двумя–тремя объемами этилового спирта с добавлением 10 мкл 5М хлорида натрия. Осадок ДНК центрифугировали 8 мин при 10 000 об/мин, сушили и растворяли в 1×TE-буфере, содержащем 10 мМ трис-НСl, рН = 7.5 и 1 мМ ЭДТА, рН = 8.0. После полного растворения ДНК проверяли ее качество и оценивали количество полученного материала с помощью электрофореза в агарозном геле.

2.3. Амплификация

Для амплификации фланкирующих участков генов гормона роста *gh1* и *gh2* гольцов рода *Salvelinus*, с помощью программ Primer Premier 5 (Lalitha, 2000) и Primer 3 (<https://bioinfo.ut.ee/primer3-0.4.0>), были разработаны пять пар праймеров, представленных в таблице 2.4.

Таблица 2.4 – Пары праймеров, использованные для амплификации фланкирующих участков генов *gh1* и *gh2* (Каменская и др., 2015)

| № | Пары праймеров | Последовательность 5'→3' |
|---|----------------|-----------------------------|
| 1 | ghpf2 | TGATGATAACAAGACCCTGTC |
| | ghpr | TGGTACACTGACTCCCCACT |
| 2 | gh2pf1 | TGTCATCTACAGTACCACAGCG |
| | gh2pr1 | TGGCACTAAATTA CT TGTTCAC TC |

Продолжение таблицы 2.4.

| | | |
|---|---------|--------------------------|
| 3 | gh1pfw | AACCGTTGTATAGTGCTCTTTT |
| | gh1prw | ACACATACAACACAAACCTCTG |
| 4 | gh1tf2 | GTTCATTGATCAAGACTGTTCC |
| | gh1tr22 | GTAGAGGTCCTTGTGAATTTTAGG |
| 5 | gh2tf | CAAGACTGGTCTCGAGAAAGTC |
| | gh2tr | TCCAAGGGACACAATAGGAATGTT |

Амплификацию фланкирующих участков проводили с использованием классической полимеразной цепной реакции (ПЦР) в термоциклере Eppendorf Mastercycler (Eppendorf, Германия). В состав реакционной смеси входили следующие компоненты: тотальная ДНК – 20–50 нг; 10×буфер DreamTaq (содержащий KCl, $(\text{NH}_4)_2\text{SO}_4$ и 20мМ MgCl_2) – 2,5 мкл (Thermo Scientific, США), смесь дезоксиинуклеотидтрифосфатов (dNTP) (2,5 мМ каждого) – 2,5 мкл (СибЭнзим, Россия); праймеры (2,5 мкМ) по 2,0 мкл каждого, DreamTaq-полимераза – 1,5 ед. (Thermo Scientific, США) и 15 мкл деионизованной воды до общего объема реакционной смеси 25 мкл. С помощью градиента температур отжига были подобраны следующие условия амплификации. Для амплификации области промотора генов *gh1* и *gh2* использовали три пары праймеров: ghpf2/ghpr, gh2pfl/gh2pr1 и gh1pfw/gh1prw. ПЦР проводили при следующих условиях: начальная денатурация (94°C – 5 мин), далее для 35 циклов: денатурация 94°C – 1 мин; отжиг 55°C – 1 мин; синтез новых цепей 72°C – 1 мин и окончательная достройка цепей 72°C – 5 мин. Область терминатора амплифицировали с использованием пар праймеров gh1tf2/gh1tr22 и gh2tf/gh2tr для *gh1* и *gh2* генов, соответственно, при следующих условиях: начальная денатурация 94°C – 5 мин, для 35 циклов денатурация 94°C – 1 мин; отжиг 60°C – 1 мин; синтез новых цепей 72°C – 1 мин; и окончательная достройка цепей 72°C – 5 мин.

Продукты амплификации были разделены с использованием электрофореза в 1,2%-ном агарозном геле в 0,5×трис-боратном (ТВЕ) буфере (89мМ трис, 89мМ борная кислота, 2,5мМ ЭДТА, pH=8.0), окрашенном бромистым этидием (EtBr) (3,8-диамино-6-этил-5-фенилфенантридиумбромид). По 2 мкл ПЦР-продукта смешивали с 2 мкл 10× буфера для нанесения образцов (0,025 г бромфенолового синего БФС, 4 мл глицерина, 5 мл 0.5М ЭДТА, pH=8.0) и вносили в лунки в геле. Для определения длины амплифицированного

фрагмента использовался стандартный ДНК-маркер (100+bp+1,5kb DNA ladder (СибЭнзим, Россия)). Результаты электрофореза просматривали и регистрировали в геле-документирующей системе GelDoc XR (BioRad, США) и в проходящем ультрафиолетовом свете с длиной волны (365 нм). Далее полученные фрагменты использовали для клонирования и секвенирования.

2.4. Молекулярное клонирование

Молекулярное клонирование – это совокупность экспериментальных процедур, позволяющая осуществлять перенос генетического материала из одного организма в другой. Клонирование даёт возможность размножить нужные фрагменты ДНК *in vivo*. Клонирование включает следующие этапы: лигирование, трансформацию, посев клеток на чашки Петри со средой, идентификацию и отбор колоний, несущих рекомбинантную ДНК. Важную роль в клонировании играет способность клеток поглощать плазмиду с исследуемым фрагментом ДНК. Это становится возможным после того, как клетки подвергают специальной обработке раствором CaCl_2 или высоковольтным разрядом (электропорация); в результате такого воздействия на клетки происходит образование пор в клеточной стенке, что облегчает проникновение молекул в клетку извне. Такие клетки, полученные в результате обработки и способные поглощать чужеродную ДНК называются компетентными.

2.4.1. Получение компетентных клеток *Escherichia coli* для молекулярного клонирования

Для получения компетентных клеток необходимы: культура клеток *Escherichia coli* (штамм *DH5a*), бактериальная среда Лурия-Бертани (LB) (табл. 2.5), буферные растворы RF1 и RF2 (табл. 2.6).

Таблица 2.5 – Ингредиенты для приготовления

1 литра среды LB (Лурия-Бертани)

| Ингредиенты | Жидкая | Твердая |
|-----------------------------|--------|---------|
| Бакто-триптон, г | 10 | 10 |
| Бакто-дрожжевой экстракт, г | 5 | 5 |
| NaCl, г | 10 | 10 |
| Агар, г | – | 12 |

Примечание. 2М NaOH использовали, чтобы довести pH до 7.5. В среде с агаром pH выравнивали после введения и расплавления агара.

1. Среду LB (табл. 2.5) с агаром без антибиотика плавил на водяной бане, остужали до 55°C и заливали чашку Петри. После того как среда в чашке застывала, каплю (~10 мкл) культуры клеток *E. coli* добавляли в чашку, растирали шпателем и инкубировали ночь при 37°C в Environmental Shaker-Incubator ES-20 (BioSan, Латвия).

2. Отбирали материал с одной отдельной колонии с помощью микробиологической петли и переносили в стеклянную пробирку с 2 мл жидкой LB среды без антибиотика и ставили инкубироваться на 200 об/мин при 37°C в шейкер-инкубатор Environmental Shaker-Incubator ES-20 (BioSan, Латвия) на ночь.

3. Готовили буферные растворы RF1 и RF2 (табл. 2.6), содержащие CaCl₂.

Таблица 2.6 – Ингредиенты для приготовления буферных растворов RF1 и RF2

| Ингредиенты | RF1 | RF2 |
|---------------------------|-----|-----|
| 1М MnCl ₂ , мл | 2 | – |
| 1М AcK pH 5.8, мл | 1,2 | – |
| 1М KCl, мл | 4 | 0,1 |

Продолжение таблицы 2.6

| | | |
|-----------------------------|-----|-----|
| 0.5M CaCl ₂ , мл | 0,8 | 1,5 |
| Глицерин 50% | 12 | 5 |
| ddH ₂ O, мл | 20 | 3,4 |
| 50mM MOPS pH 6,8, мл | – | 2 |
| Общий объем | 40 | 10 |

Примечание. Для стерилизации растворов RF1 и RF2 использовали 0.22 микронный фильтр.

4. В стеклянную колбу с круглым дном, в 50 мл жидкой среды LB без антибиотика добавляли 500 мкл ночной затравки (п.1) и ставили инкубироваться при 37°C на 200 об/мин в шейкер-инкубатор (Environmental Shaker-Incubator ES-20 (BioSan, Латвия)) до оптической плотности OD=0,6 (~2 ч). Оптическую плотность оценивали на спектрофотометре Eppendorf (Eppendorf, Германия) против контроля (чистая LB среда).

5. По 25 мл культуры переносили в две полипропиленовые пробирки с коническим дном объемом 50 мл, охлаждали на ледяной бане в течении 10–15 мин. Центрифугировали клетки при 4°C 3 000 об/мин 15 мин, затем удаляли среду.

6. Ресуспендировали осадок в 1 мл буфера RF1 (табл. 2.6) и перед тем, как поместить пробирки на ледяную баню на 15 мин, добавляли по 6,5 мл буфера RF1 (табл. 2.6) в каждую пробирку.

7. Центрифугировали клетки при 4°C 3 000 об/мин 15 мин, затем удаляли буфер RF1, добавляли по 2 мл буфера RF2 (табл. 2.6) и ресуспендировали осадок.

8. Разливали по 100 мкл клеток в пробирки на 1,5 мл и замораживали в жидком азоте, хранили клетки при –80°C.

2.4.2. Лигирование и трансформация

Процедуру клонирования осуществляли с использованием специального набора реагентов CloneJET PCR Cloning Kit (Thermo Scientific, США), согласно инструкции производителя. В состав набора входит специальный термостойкий фермент (DNA Blunting Enzyme), который обрезает концы у исследуемого фрагмента ДНК таким

образом, чтобы нуклеотидные основания на конце фрагмента были комплементарны основаниям на концах вектора, куда будет происходить вставка.

Формирование тупых концов с помощью фермента DNA Blunting Enzyme (Thermo Scientific, США) происходит перед лигированием. Смесь, которая содержит 1,75 мкл ПЦР-продукта, 2,5 мкл 2×буфера, 0,75 мкл ddH₂O и 0,25 мкл DNA blunting enzyme (Thermo Scientific, США) сначала нагревали до 70°C в течение 5 мин, а потом охлаждали на ледяной бане около 10 секунд. Далее проводили лигирование: в полученную смесь добавляли плазмидный вектор pJET/blunt cloning vector – 0,25 мкл (Thermo Scientific, США) и T4 ДНК лигазу – 0,25 мкл (Thermo Scientific, США). Смесь хорошо перемешивали и оставляли на ночь при 22°C.

Трансформация методом теплового шока включала в себя следующие этапы:

1. 100 мкл компетентных клеток *E. coli*, смешивали с лигирующей смесью и оставляли на ледяной бане при 4°C, на 30 мин.
2. Помещали пробирки на 1–1,5 мин в термостат на 42°C, затем снова переносили в лёд на 2 минуты.
3. Добавляли к клеткам 1 мл жидкой LB среды (табл. 2.5), перемешивали и помещали пробирки на 1 час в термостат на 37°C. Плавил среду LB, содержащую агар, на водяной бане, остужали до 55°C, добавляли 90 мкл антибиотика (ампициллин 50 мг/мл) на 100 мл среды и заливали в чашки Петри по 30 мл среды; оставляли чашки до полного застывания среды.
4. После термостата пробирки с трансформированными клетками центрифугировали 3 мин при 5 000 об/мин, сливали надосадочную жидкость.
5. Оставляли в пробирках по 100–150 мкл супернатанта, аккуратно ресуспендировали осадок.
6. Переносили смесь на чашки, распределяли смесь по поверхности среды микробиологическим шпателем, инкубировали чашки ночь в термостате при 37°C.

Стоит отметить, что входящий в набор вектор pJET/blunt cloning vector (Thermo Scientific, США) содержит ген эндонуклеазы рестрикции, которая является летальной для всех штаммов *E. coli*. В результате процедуры лигирования исследуемый фрагмент ДНК

встраивается именно в этот ген, и только клетки, содержащие рекомбинантную плазмиду, выживают и образуют колонии.

2.5. Идентификация рекомбинантных клонов

Отбор искомым рекомбинантных молекул ДНК проводили на основании оценки размера встроённых в векторную молекулу фрагментов. Для анализа клеточных колоний с праймерами, комплементарными к нуклеотидной последовательности вектора, использовали ПЦР. Амплификацию проводили в 20 мкл реакционной смеси, включающей 2 мкл 10×буфера DreamTaq (Thermo Scientific, США); 1 мкл смеси дезоксинуклеотидтрифосфатов (dNTP) (2,5 мМ каждого) (СибЭнзим, Россия); по 1 мкл каждого праймера (2,0 мкМ); 0,8 мкл MgCl₂ (Thermo Scientific, США); 0,2 мкл DreamTaq-полимеразы (Thermo Scientific, США) и 14 мкл деионизированной воды. Реакцию проводили при следующих условиях: начальная денатурация 98°C – 2 мин, 25 циклов (денатурация 95°C – 15 сек, отжиг праймеров 55°C – 15 сек, достройка цепей 72°C – 1 мин) и финальная достройка цепей (72°C – 5 мин), охлаждение до 4 °С. Продукты ПЦР разделяли в 1,2% агарозном геле с маркером молекулярного веса, кратным 100 п.н. (100+bp+1,5kb DNA ladder (СибЭнзим, Россия)). Препараты, в которых была обнаружена вставка исследуемого фрагмента ДНК, очищали осаждением этанолом.

На первом этапе очистки к содержимому пробирок добавляли двойной объем (50 мкл) 96% этанола и оставляли на ночь при –20°C. Затем центрифугировали 20 мин при 4°C 10 000 об/мин, спирт сливали. Следующий этап очистки 70% этанолом: добавляли 50 мкл 70% этанола, центрифугировали 20 мин при 4°C 10 000 об/мин, спирт сливали. Очистку 70% спиртом повторяли дважды. После удаления спирта из пробирок, пробы ставили сушиться на 30 мин в термостате при 37°C. К высушенным пробам добавляли по 15–20 мкл ddH₂O, хорошо перемешивали на Vortex (BioSan, Латвия), до растворения осадка. Концентрацию очищенной вставки также анализировали при помощи метода электрофореза в 1,2% агарозном геле. При необходимости добавляли ddH₂O для выравнивания концентрации.

2.6. Секвенирование

Определение нуклеотидных последовательностей фланкирующих участков генов гормона роста проводили с помощью набора реактивов BigDye Terminator Cycle Sequencing Kit V. 3.1 (Applied Biosystems, США) по методике производителя. Для секвенирования использовали праймеры, входящие в состав набора CloneJET PCR Cloning Kit (Thermo Scientific, США). Реакционная смесь объемом 10 мкл содержала 2 мкл ДНК; 0,8 мкл (2.0 мкМ) прямого или обратного праймера; 1,5 мкл BigDye Terminator Ready Reaction Kit V 3.1 (Applied Biosystems, США); 1,25 мкл 5×буфера для секвенирования и 4,45 мкл деионизированной воды. Температура и время начальной денатурации в реакции 96°C – 1 мин, далее 26 циклов проходили при следующих условиях: денатурация 96°C – 10 сек, отжиг праймеров 55°C – 10 сек, элонгация 60°C – 4 мин.

После амплификации продукты ПЦР пересаждали и очищали этанолом. В пробирки добавляли 2,5 мкл 0,125М ЭДТА и 30 мкл 96% этанола, перемешивали и оставляли при комнатной температуре на 15 мин. Затем центрифугировали 20 мин при 10 000 об/мин с охлаждением 4°C; спирт сливали и дважды добавляли по 30 мкл 70% этанола, перемешивали, центрифугировали, спирт удаляли. Пробы высушивали при 60°C 15 мин, добавляли по 13 мкл формамида (CH₃NO) перемешивали и откручивали. Перед тем как заносить образцы в планшет, проводили денатурацию при 95°C 3 мин. Последовательности нуклеотидов определяли в автоматическом генетическом анализаторе ABI Prism GA3130xl (Applied Biosystems, США) на базе Кафедры клеточной биологии и генетики, Института Мирового океана ДВФУ и в генетическом анализаторе ABI Prism 3500 (Applied Biosystems, США) на базе Национального научного центра морской биологии им. А. В. Жирмунского ДВО РАН.

2.7. Сборка и выравнивание последовательностей

Полученные хроматограммы визуально проверяли на наличие ошибок с помощью программы FinchTV 1.4.0 (<http://finchtv.software.informer.com/1.4>) Сборку нуклеотидных последовательностей осуществляли в пакете программ SeqScape V 2.6 (Applied Biosystems, США), BioEdit 7.2.5 (Hall, 1999). Анализ нуклеотидных последовательностей

и множественное выравнивание проводили с помощью пакета программ MEGA 7.0 (Kumar et al., 2016). При выравнивании использовали алгоритмы ClustalW (Thompson et al., 1994) и Muscle (Edgar, 2004). Полученные последовательности с помощью алгоритма BLASTn (Altschul et al., 1990) сопоставляли с уже имеющимися последовательностями в GenBank NCBI (<https://www.ncbi.nlm.nih.gov/nucleotide/>), чтобы установить гомологию полученных нуклеотидных последовательностей.

2.8. Анализ данных

Филогенетическую реконструкцию, основанную на полных нуклеотидных последовательностях гена гормона роста, включающих: экзоны, интроны и фланкирующие области генов выполняли Байесовским методом (Bayesian Inference, BI) с помощью программы MrBayes v. 3.2.6 (Ronquist et al., 2012) и методом ближайшего соседства (Neighbor-Joining) в программе MEGA 7.0 (Kumar et al., 2016). Оптимальная модель нуклеотидных замен (T92+G) была выбрана в программе jModelTest 2.1.1 (Darriba et al., 2012) согласно информационному критерию Акаике. Нуклеотидное разнообразие (π) и дивергенцию (Dxy) последовательностей гена гормона роста оценивали в программе DnaSP 5.10.01 (Librado, Rozas, 2009) с использованием метода скользящего окна (sliding window). Данный метод применяется для выявления закономерностей изменчивости нуклеотидных последовательностей за счет определения средних значений в окне заданной длины, которое перемещается вдоль последовательности (Proutski, Holmes, 1998).

ГЛАВА 3. РЕЗУЛЬТАТЫ

3.1. Характеристика регуляторных участков гена гормона роста у гольцов рода *Salvelinus*

Были получены нуклеотидные последовательности 5'-фланкирующего промоторного участка паралогичных генов гормона роста *gh1* и *gh2* для четырех видов гольцов рода *Salvelinus*. Длина промоторного участка гена *gh1* составила: *S. malma* – 1202 п.н., *S. curilus* – 1267 п.н., *S. levanidovi* – 1197 п.н. и *S. taranetzi* – 1207 п.н.. Длина промотора гена *gh2* составила: *S. malma* – 1183 п.н., *S. curilus* – 1193 п.н., *S. levanidovi* – 1172 п.н. и *S. taranetzi* – 1193 п.н..

Промоторы обоих генов-паралогов содержат характерный для большинства генов эукариот высококонсервативный мотив – ТАТА-бокс, в положении –22 п.н. от точки инициации транскрипции. Последовательность ТАТА-бокса консервативна у всех четырех видов гольцов рода *Salvelinus* в обоих генах. Помимо ТАТА-бокса в промоторной области были обнаружены другие регуляторные элементы. К ним относятся А/Т-богатые регионы, содержащие сайты связывания с гипофиз-специфичным транскрипционным фактором Pit-1, а также области промотора, ответственные за взаимодействие с другими активаторами и инициаторами транскрипции, в частности – с рецепторами гормонов.

У гольцов с транскрипционным фактором Pit-1 взаимодействуют четыре сайта связывания: F1 – F4. Последовательность сайта F1, расположенного в положении –39/–61 п.н., консервативна в каждом гене у всех четырех видов гольцов и отличается лишь тремя нуклеотидными заменами между генами *gh1* и *gh2* (рис. 3.1). Изменения локализованы во фланкирующих областях сайта и не затрагивают его центральную часть. Сайт связывания F2, расположенный в положениях –118/–140 п.н. в гене *gh1* и –117/–139 п.н. в гене *gh2*, консервативен в каждом гене и имеет две нуклеотидные замены между двумя генами. Третий сайт (F3) находится в положении –155/–183 п.н. в гене *gh1*, и в положении –153/–181 п.н. в гене *gh2*. Последовательность сайта F3 гена *gh1* консервативна у всех четырех исследованных видов гольцов. В гене *gh2* последовательность сайта F3 у *S. curilus* и *S. malma* идентична и имеет по одной нуклеотидной замене, по сравнению с сайтом F3 *S. taranetzi* и *S. levanidovi*. Положение сайта F4 в гене *gh1* совпадает у всех гольцов –264/–

295 п.н.. В гене *gh2* границы сайта F4 у *S. malma*, *S. curilus* и у *S. taranetzi* располагаются на расстоянии –263/–294 п.н., у *S. levanidovi* границы сайта F4 смещены на 2 нуклеотида –265/–296 п.н.. Сайты F3 и F4 в генах *gh1* и *gh2* имеют большее число нуклеотидных замен по сравнению с сайтами F1 и F2 (рис. 3.1).

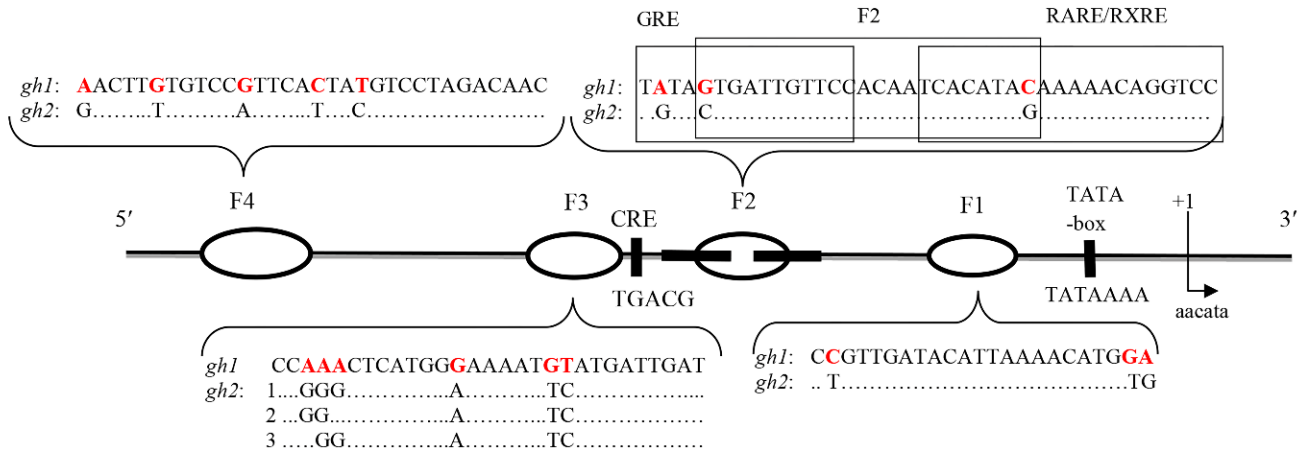


Рисунок 3.1 – Схема расположения сайтов связывания (F1–F4) с гипофиз-специфичным транскрипционным фактором Pit-1; элементов ответа на цАМФ (CRE) и глюкокортикоиды (GRE); и рецептора ретиноевой кислоты (RARE/RXRE) в промоторной области генов гормона роста гольцов. Цифрами обозначены виды гольцов, имеющие различия в сайте F3 гена *gh2*: 1 – *S. malma* и *S. curilus*; 2 – *S. levanidovi*; 3 – *S. taranetzi* (Каменская и др., 2015, с модификациями).

В промоторе обоих генов гормона роста гольцов были найдены элементы ответа на циклический АМФ (CRE), глюкокортикоиды (GRE) и сайты связывания с рецептором ретиноевой кислоты (RARE/RXRE) (рис. 3.1). Последовательность элемента CRE представлена в промоторах гена гормона роста *gh1* и *gh2* ассиметричным мотивом (TGACG), идентична в двух генах и располагается между сайтами F2 и F3 в положениях –148/–152 п.н. и –147/–151 п.н. в генах *gh1* и *gh2*, соответственно.

Элемент ответа на глюкокортикоиды (GRE) представлен в промоторах гена гормона роста гольцов несовершенным инвертированным повтором (рис. 3.1). В дистальной части элемента ответа на глюкокортикоиды имеются две нуклеотидные замены, проксимальная часть элемента ответа консервативна в обоих генах у всех четырех видов. Положение элемента ответа на глюкокортикоиды (GRE): –130/–144 п.н. в гене *gh1* и –129/–143 п.н. в гене *gh2*.

У гольцов сайт связывания с рецептором ретиноевой кислоты (RARE/RXRE) присутствует в промоторной области обоих генов и занимает положение –105/–124 п.н. в

гене *gh1* и –106/–125 п.н., в гене *gh2*. Элемент ответа на глюкокортикоиды (GRE) и сайт связывания с рецептором ретиноевой кислоты (RARE/RXRE) частично перекрываются с сайтом F2, ответственным за связывание с транскрипционным фактором Pit-1; элемент GRE перекрывается с дистальной частью, а элемент RARE/RXRE с проксимальной частью сайта F2 (рис. 3.1). Промоторы обоих генов гормона роста гольцов на всем своем протяжении не содержат часто встречающийся в промоторах других генов эукариот СААТ-бокс и сайт связывания с рецептором тиреоидного гормона (TRE), описанный в промоторе гена гормона роста некоторых видов позвоночных животных.

В процессе исследования промоторных областей паралогичных генов гормона роста гольцов рода *Salvelinus* было установлено, что большинство функциональных сайтов связывания, ответственных за взаимодействие с основными факторами транскрипции, сосредоточено вблизи стартового кодона, поэтому в дальнейший анализ по оценке нуклеотидного разнообразия и дивергенции вошли не все полученные 5'-фланкирующие последовательности, а только участки длиной 300 п.н..

Длина 3'-терминирующей последовательности генов *gh1* и *gh2* была получена только у *S. levanidovi* и составила 1196 п.н. и 606 п.н., соответственно. 3' участок обоих генов включает поли-А последовательность (ААТAAA) в положениях +4137 п.н. и +3794 п.н. для *gh1* и *gh2* генов, соответственно. Других потенциальных цис-регуляторных последовательностей и шпилечных участков в терминирующей части не обнаружено, поэтому анализ не проводился.

3.2. Сравнительный анализ паралогичных генов и их промоторов у лососевых рыб

3.2.1. Изменчивость в 5'-цис-прилежащем промоторном участке

Распределения изменчивых сайтов и величин дивергенции оценивались в 5'-фланкирующих участках длиной 300 п.н. как между видами внутри рода *Salvelinus*, (рис. 3.2 – 3.5) так и при сравнении промоторных последовательностей гольцов с промоторами генов гормона роста: нерки (*O. nerka*), чавычи (*O. tshawytscha*) и атлантического лосося (*S. salar*) (рис. 3.6 – 3.8 и рис. 3.10).

На рисунке 3.2 представлено сравнение промоторных последовательностей гена *gh1* между видами гольцов рода *Salvelinus*. Видно, что последовательности достаточно консервативны и встречаются только единичные участки изменчивости. Последовательности промотора южной азиатской мальмы (*S. curilus*) и гольца Леванидова (*S. levanidovi*) полностью совпадают и не имеют различий. В результате сравнения промотора северной мальмы (*S. malma*) с промотором южной азиатской мальмы (*S. curilus*) и с промотором гольца Леванидова (*S. levanidovi*), наблюдается один участок изменчивости, локализованный между сайтом F1 и ТАТА-боксом. Один участок изменчивости встречается при сравнении южной азиатской мальмы (*S. curilus*) и гольца Леванидова (*S. levanidovi*) с гольцом Таранца (*S. taranetzi*). Однако расположен этот участок изменчивости между сайтами F1 и F2, а не как в парах *S. malma* – *S. levanidovi* и *S. malma* – *S. curilus* (рис. 3.2). В паре видов северная мальма (*S. malma*) и голец Таранца (*S. taranetzi*) промоторные последовательности имеют два участка изменчивости, один из которых расположен между сайтом F1 и ТАТА-боксом, а второй между сайтами F1 и F2.

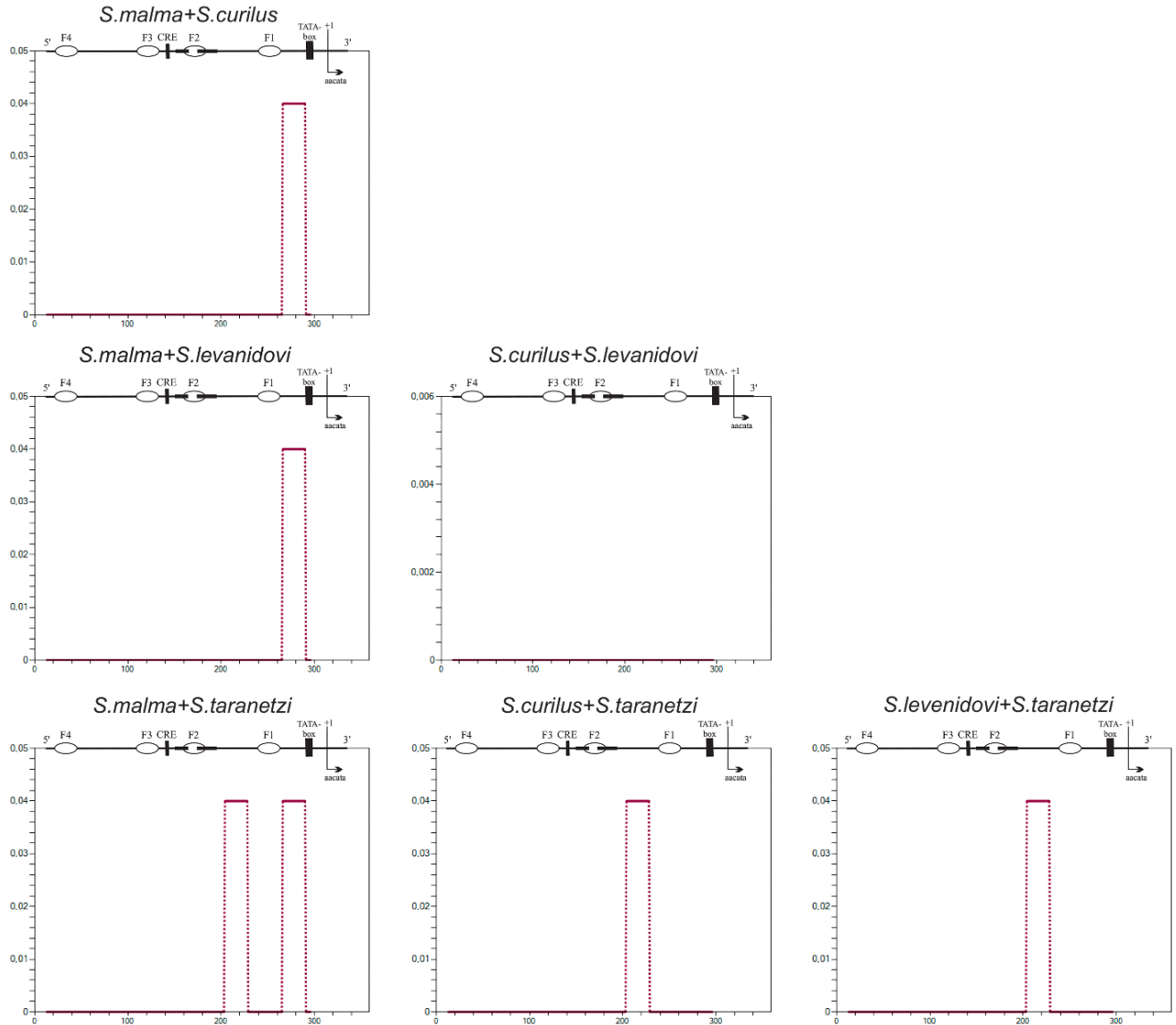


Рисунок 3.2 – Анализ уровня дивергенции промоторных участков гена *gh1* методом скользящего окна между видами гольцов рода *Salvelinus*. Размер рамки – 10 п.н., длина шага – 1 п.н. По оси *x* отмечены положения нуклеотидов, по оси *y* – уровень дивергенции.

Промоторные участки гена гормона роста *gh2* у четырех видов гольцов рода *Salvelinus* так же, как и последовательности промотора гена *gh1*, имеют низкий уровень дивергенции (рис. 3.3). Полностью совпадают последовательности промотора у северной мальмы (*S. malma*) и южной азиатской мальмы (*S. curilus*). Два участка дивергенции наблюдается при сравнении северной мальмы (*S. malma*) и южной азиатской мальмы (*S. curilus*) с гольцом Леванидова (*S. levanidovi*) и гольцом Таранца (*S. taranetzi*).

Расположены данные участки изменчивости между сайтами связывания с транскрипционным фактором Pit-1 F3 и F4, ближе к сайту F3. При сравнении последовательности промотора гольца Леванидова (*S. levanidovi*) с гольцом Таранца (*S. taranetzi*) встречается только один участок изменчивости и расположен он между сайтами F3 и F4.

Из рисунков 3.4 и 3.5 видно, что промоторные участки в обоих генах представителей рода *Salvelinus* очень консервативны по всей длине. Нуклеотидное разнообразие в промоторах обоих генов составляет 0,00327 в гене *gh1* и 0,00545 в гене *gh2* (таблица 3.1). Значения дивергенции 0,00218 и 0,00654 для промоторов генов *gh1* и *gh2*, соответственно. В гене *gh1* наблюдаются два участка изменчивости: один непосредственно прилегает к ТАТА-боксу, второй локализован между сайтами F1 и F2 (рис. 3.4). Два участка с повышенным уровнем изменчивости встречаются и в гене *gh2*, но они отличаются по своей локализации. Один находится в непосредственной близости к сайту F3, второй участок располагается между сайтами F3 и F4 (рис. 3.5). Графики, построенные методом скользящего окна, для нуклеотидной изменчивости и дивергенции 5'-фланкирующих участков генов гормона роста *gh1* и *gh2* гольцов рода *Salvelinus* в значительной степени совпадают.

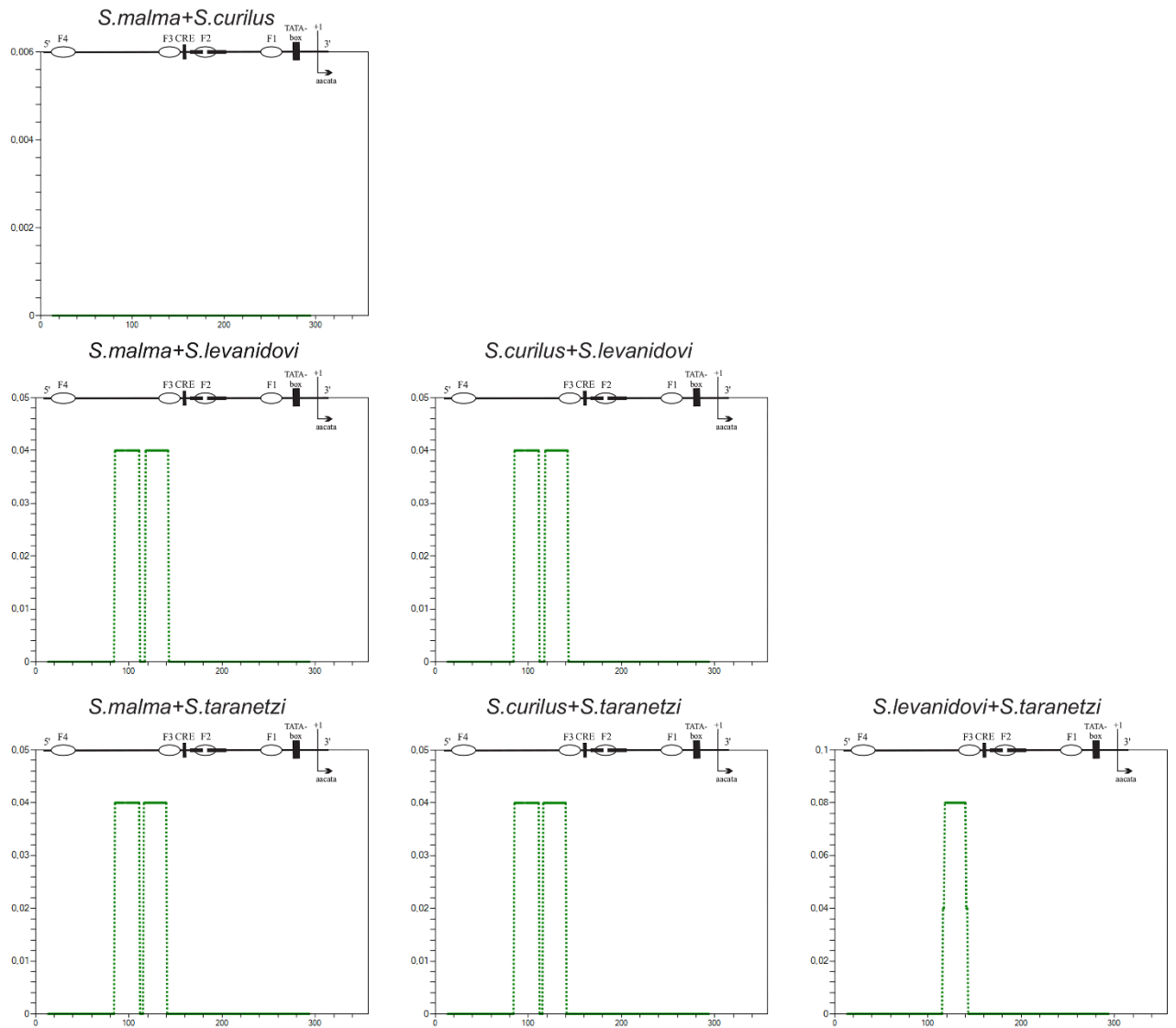


Рисунок 3.3 – Анализ уровня дивергенции промоторных участков гена *gh2* методом скользящего окна между видами гольцов рода *Salvelinus*. Размер рамки – 10 п.н., длина шага – 1 п.н. По оси *x* отмечены положения нуклеотидов, по оси *y* – уровень дивергенции.

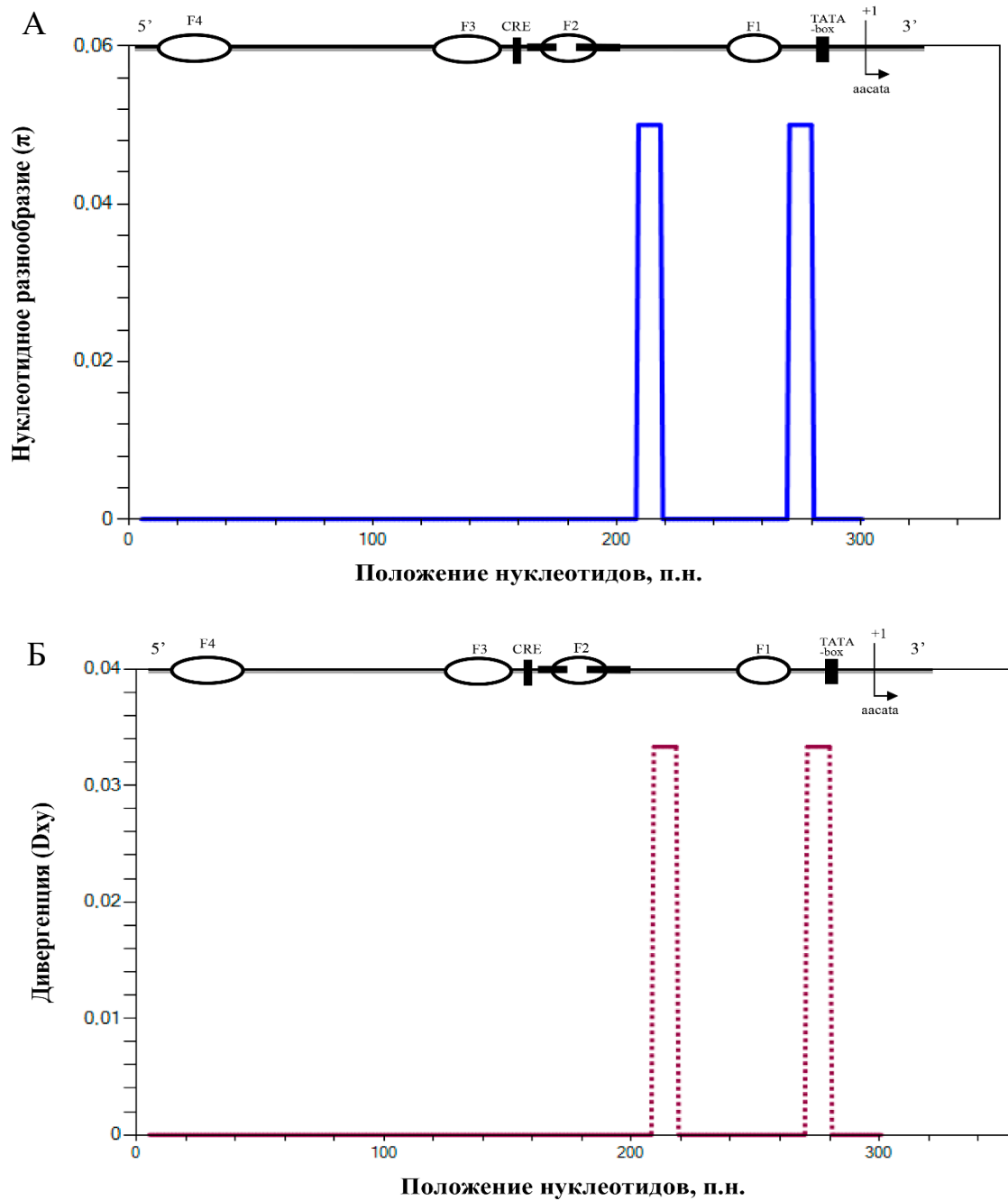


Рисунок 3.4 – Анализ нуклеотидной изменчивости и дивергенции промоторных участков гена *gh1* голцов рода *Salvelinus* методом скользящего окна. Размер рамки – 10 п.н., длина шага – 1 п.н. По оси x отмечены положения нуклеотидов, по оси y – нуклеотидное разнообразие (π) (А) и дивергенция (Б). А– уровень нуклеотидной изменчивости промоторных участков гена *gh1*. Б – уровень дивергенции промоторных участков гена *gh1*.

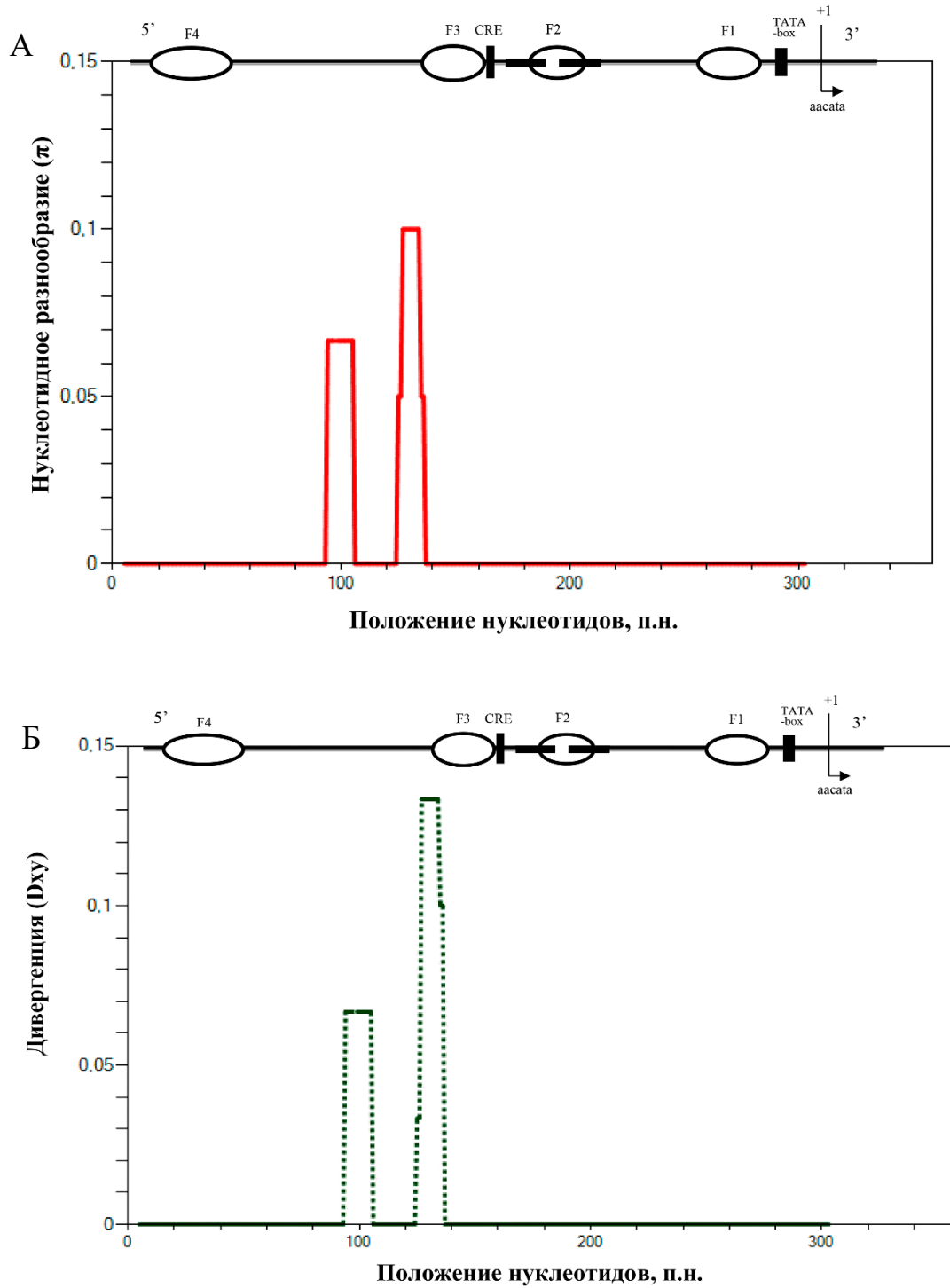


Рисунок 3.5 – Анализ нуклеотидной изменчивости и дивергенции промоторных участков гена *gh2* гольцов рода *Salvelinus* методом скользящего окна. Размер рамки – 10 п.н., длина шага – 1 п.н. По оси x отмечены положения нуклеотидов, по оси y – нуклеотидное разнообразие (π) (А) и дивергенция (В). А– уровень нуклеотидной изменчивости промоторных участков гена *gh2*. Б – уровень дивергенции промоторных участков гена *gh2*.

Таблица 3.1 – Нуклеотидное разнообразие промоторной области генов гормона роста *gh1* и *gh2*

| Наименование таксонов | Нуклеотидное разнообразие (π) \pm Стандартное отклонение | |
|---------------------------------------------------------------------------------|------------------------------------------------------------------|-----------------------|
| | Промотор <i>gh1</i> | Промотор <i>gh2</i> |
| Род <i>Salvelinus</i> | 0,00327 \pm 0,00111 | 0,00545 \pm 0,00145 |
| Семейство Salmonidae (<i>Salvelinus</i> , <i>Oncorhynchus</i> , <i>Salmo</i>) | 0,04922 \pm 0,01143 | 0,05721 \pm 0,01548 |

В результате сравнительного анализа промоторных областей генов *gh1* и *gh2* гольцов рода *Salvelinus* с другими представителями семейства Salmonidae, были показаны множественные участки дивергенции (рис. 3.6, 3.7). В промоторе гена *gh1* наиболее изменчивые участки можно наблюдать в сайте F2 и вблизи сайта F4, при сравнении четырех видов гольцов рода *Salvelinus* с чавычей (*O. tshawytscha*) (рис. 3.6). Однако, сравнивая последовательности промотора гольцов рода *Salvelinus* с неркой (*O. nerka*) участок с наибольшим уровнем дивергенции находится рядом с сайтом F1, а участки, ассоциированные с сайтами F2 и F4, имеют более низкий уровень изменчивости. Последовательности промотора гена *gh1* северной мальмы (*S. malma*), гольца Леванидова (*S. levanidovi*) и гольца Таранца (*S. taranetzi*) при сравнении с атлантическим лососем (*S. salar*) имеют два наиболее выраженных участка дивергенции: один прилегает к сайту F1, другой к сайту F4. В паре видов *S. curilus* – *S. salar* участок с наибольшим уровнем дивергенции приходится на центральную часть последовательности промотора (рис. 3.6).

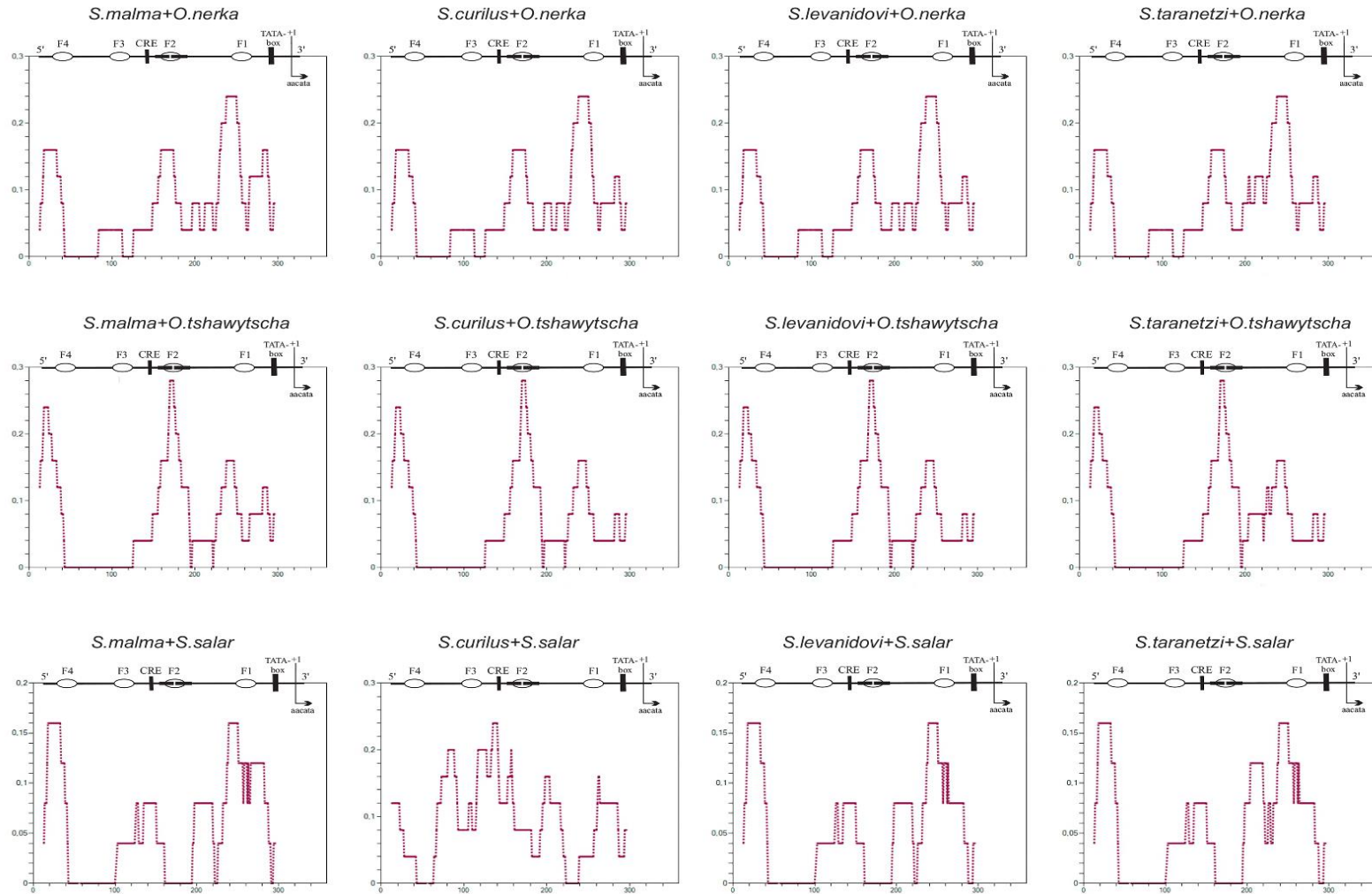


Рисунок 3.6 – Анализ уровня дивергенции промоторных участков гена *ghl* методом скользящего окна среди представителей трех родов семейства Salmonidae: *Oncorhynchus*, *Salmo*, *Salvelinus*. Размер рамки – 10 п.н., длина шага – 1 п.н. По оси *x* отмечены положения нуклеотидов, по оси *y* – уровень дивергенции.

Как видно из рисунка 3.7, промотор гена *gh2* более вариабельный, чем промотор гена *gh1*. Участки дивергенции представлены не единичными пиками, а протяженными областями. Сравнение четырех видов гольцов рода *Salvelinus* с неркой (*O. nerka*) и чавычей (*O. tshawytscha*) показывает, что один участок дивергенции располагается между сайтами F3 и F4, а другой включает в себя сайт F2 и часть прилегающей к нему последовательности со стороны сайта F1. При сравнении четырех видов гольцов рода *Salvelinus* с атлантическим лососем (*S. salar*) на участке между сайтами F3 и F4 наблюдается один четко выраженный пик, а часть промотора, которая включает сайт F2 и прилегающую к нему последовательность можно отнести к участкам со средним уровнем дивергенции. В паре видов *S. taranetzi* – *S. salar* можно выделить еще один пик, который приходится на сайт F3. Несмотря на повышенный уровень дивергенции, сайты связывания с транскрипционными факторами остаются достаточно консервативными (рис. 3.7). В результате анализа, объединяющего промоторные последовательности гена *gh1* всех исследованных видов лососевых рыб, была выявлена консервативность ТАТА-бокса, сайта F1 и CRE последовательности. Сайты F2 и F3 могут быть охарактеризованы как участки со средним уровнем изменчивости (рис. 3.8). Однако, стоит отметить, что при сравнении гольцов рода *Salvelinus* с неркой (*O. nerka*) и атлантическим лососем (*S. salar*) сайты F2 и F3 остаются консервативными (рис. 3.6). Самым вариабельным из сайтов связывания с транскрипционным фактором Pit-1 оказался сайт F4 (рис. 3.8). Исключение составляет сравнение пары видов *S. curilus* – *S. salar*, где сайт F4 имеет незначительную дивергенцию (рис. 3.6).

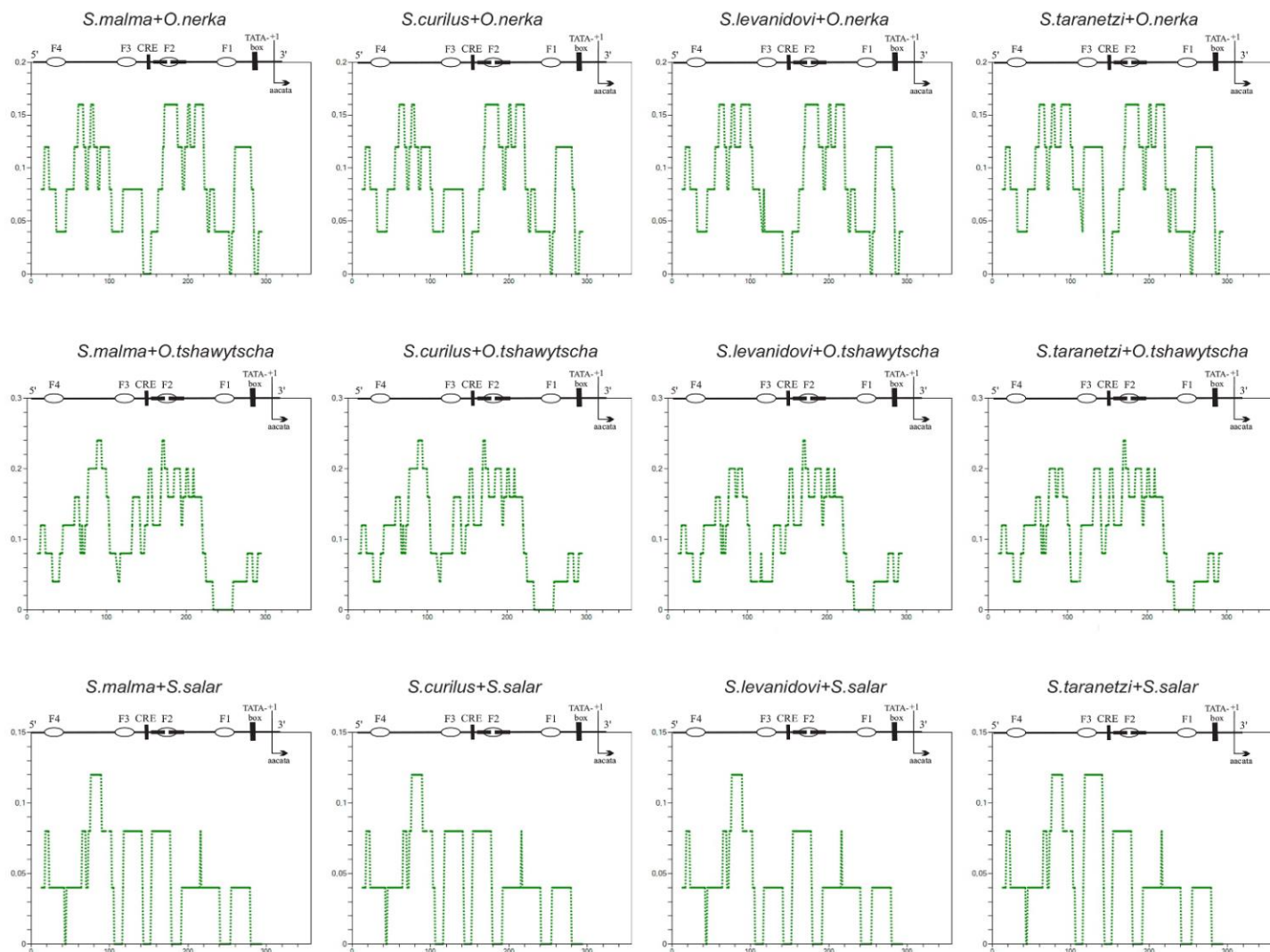


Рисунок 3.7 – Анализ уровня дивергенции промоторных участков гена *gh2* методом скользящего окна среди представителей трех родов семейства Salmonidae: *Oncorhynchus*, *Salmo*, *Salvelinus*. Размер рамки – 10 п.н., длина шага – 1 п.н. По оси *x* отмечены положения нуклеотидов, по оси *y* – уровень дивергенции.

В результате анализа, объединяющего промоторные последовательности гена *gh1* всех исследованных видов лососевых рыб, была выявлена консервативность ТАТА-бокса, сайта F1 и CRE последовательности. Сайты F2 и F3 могут быть охарактеризованы как участки со средним уровнем изменчивости (рис. 3.8). Однако, стоит отметить, что при сравнении гольцов рода *Salvelinus* с неркой (*O. nerka*) и атлантическим лососем (*S. salar*) сайты F2 и F3 остаются консервативными (рис. 3.6). Самым вариабельным из сайтов связывания с транскрипционным фактором Pit-1 оказался сайт F4 (рис. 3.8). Исключение составляет сравнение пары видов *S. curilus* – *S. salar*, где сайт F4 имеет незначительную дивергенцию (рис. 3.6).

Элемент RARE/RXRE консервативен, лишь одна нуклеотидная замена была найдена у *O. nerka*. В элементе GRE выявлено две нуклеотидные замены. И встречаются эти замены так же, как и в элементе RARE/RXRE, только у представителей рода *Oncorhynchus*. Интересно отметить, что у *O. tshawytscha* часть сайта F2, которая не перекрывается элементами GRE и RARE/RXRE, в трех из четырех положений нуклеотидов содержит замены: две трансверсии и одну транзицию (рис. 3.9). Наличие замен на этом участке приводит к тому, что при сравнении последовательностей промотора гольцов рода *Salvelinus* с чавычей (*O. tshawytscha*) сайт F2 оказывается самым изменчивым (рис. 3.6). Еще один высококонсервативный участок (≈ 50 п.н.) расположен вблизи сайта F4. Последовательности с высоким уровнем изменчивости локализованы около сайтов F1 и F2 (рис. 3.8).

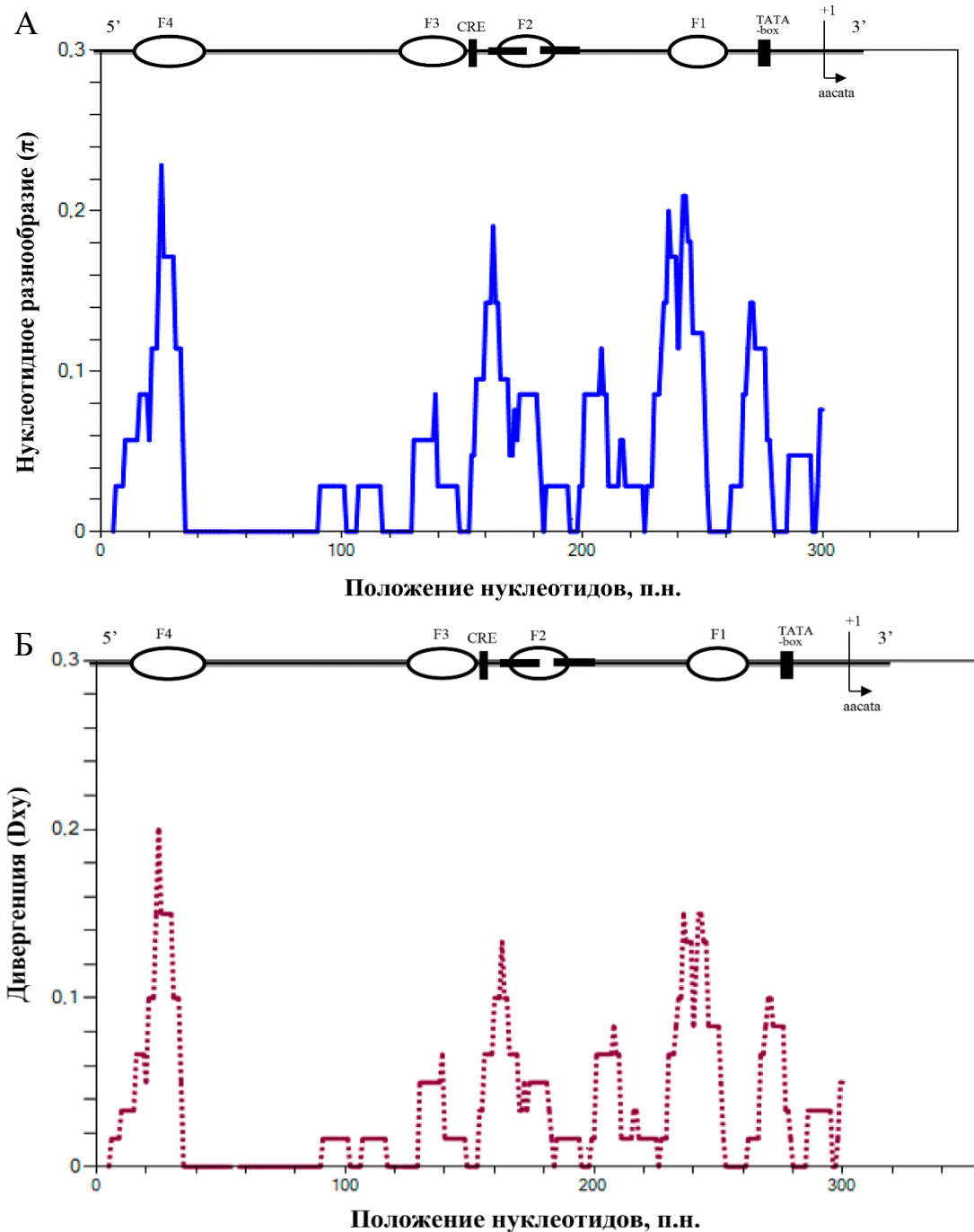


Рисунок 3.8 – Анализ нуклеотидной изменчивости и дивергенции промоторных участков гена *gh1* представителей семейства Salmonidae методом скользящего окна. Размер рамки – 10 п.н., длина шага – 1 п.н. По оси *x* отмечены положения нуклеотидов, по оси *y* – нуклеотидное разнообразие (π) (А) и дивергенция (В). А – уровень нуклеотидной изменчивости промоторных участков гена *gh1* среди представителей трех родов семейства Salmonidae: *Oncorhynchus*, *Salmo*, *Salvelinus*. Б – уровень дивергенции промоторных участков гена *gh1* среди представителей трех родов семейства Salmonidae: *Oncorhynchus*, *Salmo*, *Salvelinus*.

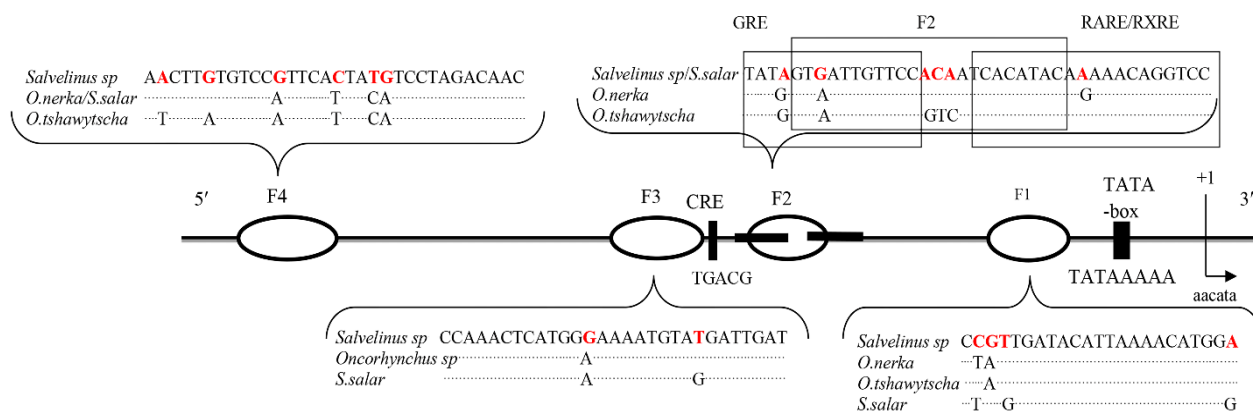


Рисунок 3.9 – Схема расположения сайтов связывания с транскрипционными факторами в промоторной области гена гормона роста *gh1* у представителей трех родов семейства Salmonidae. В сайтах связывания с гипофиз-специфичным транскрипционным фактором Pit-1 (F1–F4); элемента ответа на глюкокортикоиды (GRE) и рецептора ретиноевой кислоты (RARE/RXRE) отмечены различия в нуклеотидной последовательности между видами.

Последовательность промотора гена *gh2* у всех видов лососевых рыб (рис. 3.10) оказалась более изменчива по сравнению с промоторной областью гена *gh1*. Консервативными остаются ТАТА-боксы, сайт F1, элементы RARE/RXRE и CRE. Часть сайта F3 и сайт F4 можно отнести к участкам со средним уровнем изменчивости. Наиболее изменчивыми оказались участки расположенные между сайтами связывания с транскрипционным фактором Pit-1: F1 и F2, а также между сайтами F3 и F4 (рис. 3.10). Как и в промоторе гена *gh1*, часть сайта F2 у представителей рода *Oncorhynchus*, расположенная между GRE и RARE/RXRE, содержит замены. Замены у *O. nerka* встречаются в двух положениях из четырех; у *O. tshawytscha* отличаются три нуклеотида, при этом среди замен присутствуют как транзиции, так и трансверсии (рис. 3.11).

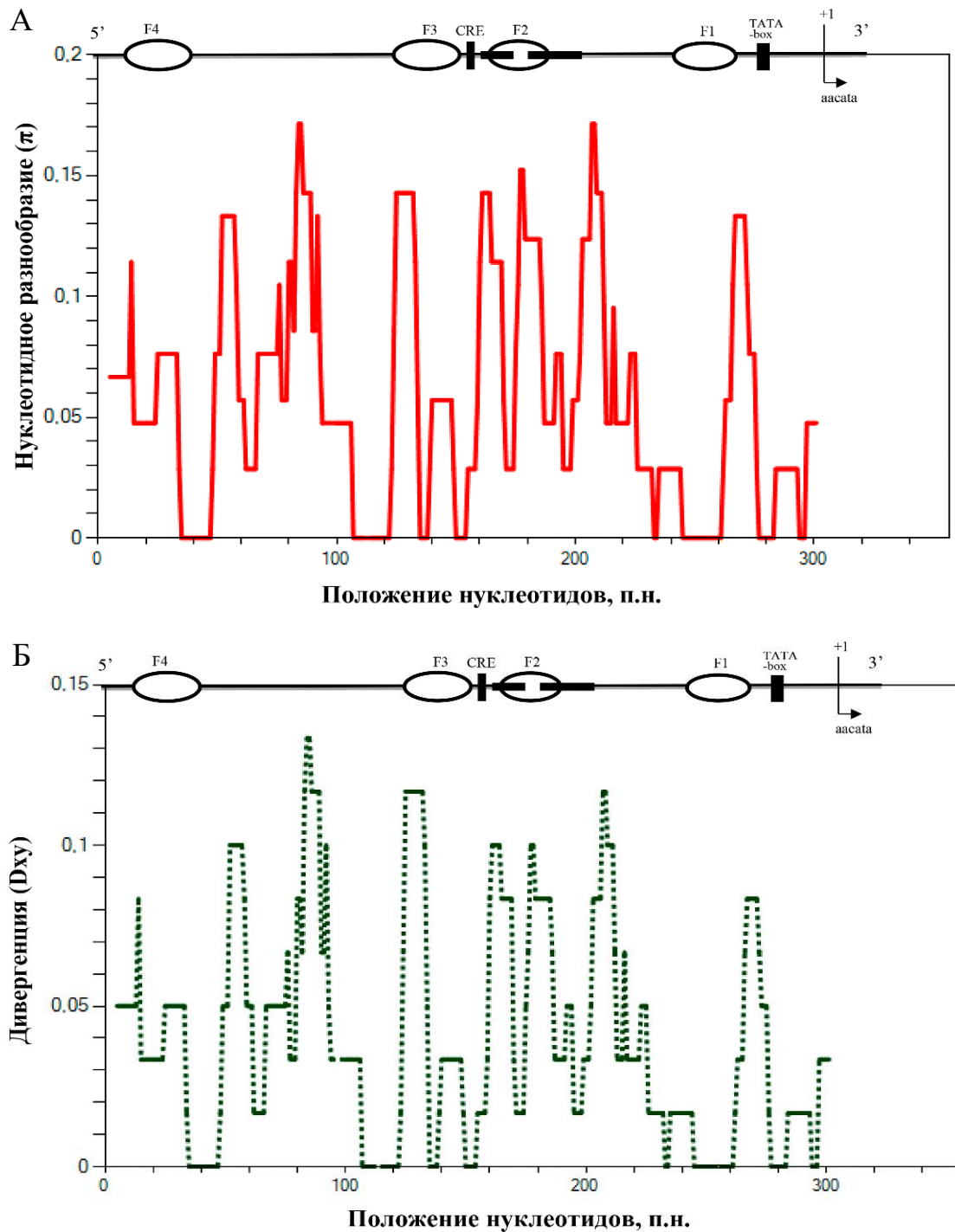


Рисунок 3.10 – Анализ нуклеотидной изменчивости и дивергенции промоторных участков гена *gh2* представителей семейства Salmonidae методом скользящего окна. Размер рамки – 10 п.н., длина шага – 1 п.н. По оси *x* отмечены положения нуклеотидов, по оси *y* – нуклеотидное разнообразие (π) (А) и дивергенция (В). А– уровень нуклеотидной изменчивости промоторных участков гена *gh2* среди представителей трех родов семейства Salmonidae: *Oncorhynchus*, *Salmo*, *Salvelinus*. Б – уровень дивергенции промоторных участков гена *gh1* среди представителей трех родов семейства Salmonidae: *Oncorhynchus*, *Salmo*, *Salvelinus*.

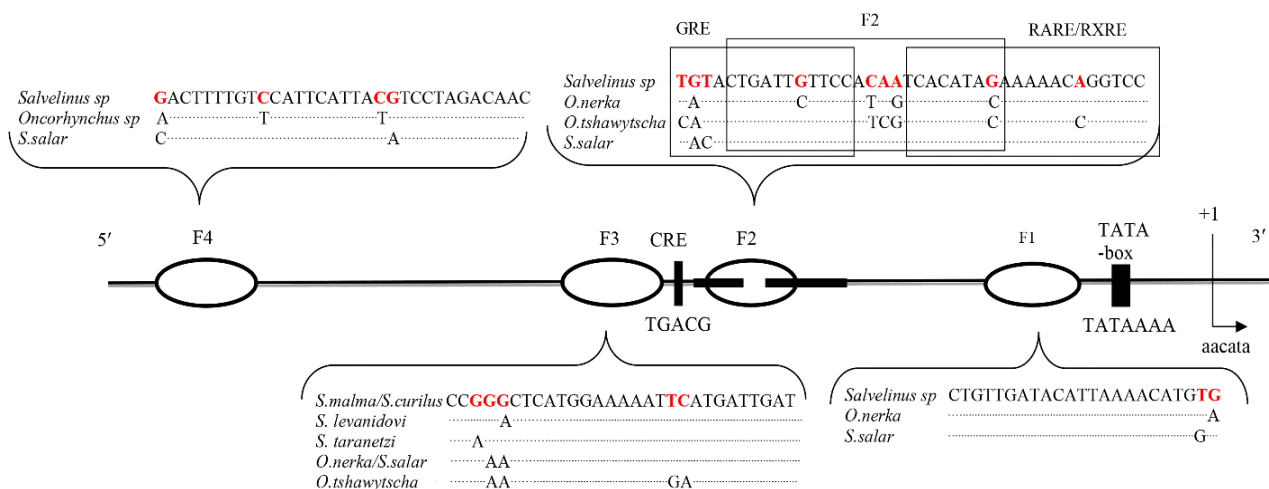


Рисунок 3.11 – Схема расположения сайтов связывания с транскрипционными факторами в промоторной области гена гормона роста *gh2* у представителей трех родов семейства Salmonidae. В сайтах связывания с гипофиз-специфичным транскрипционным фактором Pit-1 (F1–F4); элементе ответа на глюкокортикоиды (GRE) и рецепторе ретиноевой кислоты (RARE/RXRE) отмечены различия в нуклеотидной последовательности между видами.

Несмотря на то, что при сравнении промоторных областей гольцов рода *Salvelinus* с промоторами нерки (*O. nerka*), чавычи (*O. tshawytscha*) и атлантического лосося (*S. salar*) не все участки сохраняют консервативность, характерную для промоторов гольцов, уровень нуклеотидного разнообразия данного участка промотора у всех исследованных видов между генами не сильно различается: в гене *gh1* он составил $\pi=0,04922$, а в гене *gh2* $\pi=0,05721$ (табл. 3.1). Кривые нуклеотидной изменчивости и дивергенции, как и в случае сравнения гольцов рода *Salvelinus*, в значительной степени совпадают по каждому гену.

3.2.2. Изменчивость структурных последовательностей генов

Результаты анализа методом скользящего окна транскрибируемой части паралогичных генов гормона роста у гольцов рода *Salvelinus* и их сравнение с генами гормона роста других лососевых рыб представлены на рисунках 3.12 – 3.19. Как видно из рисунка 3.12, кодирующая часть гена *gh1* у гольцов рода *Salvelinus* имеет очень низкий уровень дивергенции, значения D_{xy} для пар видов *S. malma* – *S. curilus*, *S. malma* – *S. levanidovi* и *S. curilus* – *S. levanidovi* составляют 0,00246, 0,00245 и 0,00382, соответственно (Приложение III). Самый высокий уровень дивергенции среди последовательностей гена *gh1* гольцов можно наблюдать при сравнении гольца Леванидова (*S. levanidovi*) с гольцом Таранца (*S. taranetzi*), $D_{xy}=0,00626$. Уровень дивергенции в парах видов *S. malma* – *S. taranetzi* и *S. curilus* – *S. taranetzi* отличается незначительно.

Вариабельные участки, в большинстве случаев, приходятся на последовательности интронов, однако некоторые интроны имеют очень низкий уровень дивергенции (рис. 3.12). Например, при сравнении северной мальмы (*S. malma*) с южной азиатской мальмой (*S. curilus*) интроны А и В сохраняют консервативность. В паре видов *S. malma* – *S. levanidovi* на достаточно протяженном участке последовательности, которая включает часть интрона D, экзон 5, интрон Е целиком и экзон 6, полностью отсутствуют различия. Консервативные участки небольшой длины встречаются и в интронах С и D в паре видов *S. malma* – *S. curilus*, и в интронах А и D при сравнении гольца Леванидова (*S. levanidovi*) с северной мальмой (*S. malma*) и южной азиатской мальмой (*S. curilus*). Сравнивая гольца Таранца (*S. taranetzi*) с тремя другими видами гольцов рода *Salvelinus* можно заметить, что все интроны имеют достаточно высокий уровень дивергенции, особенно выделяется интрон А.

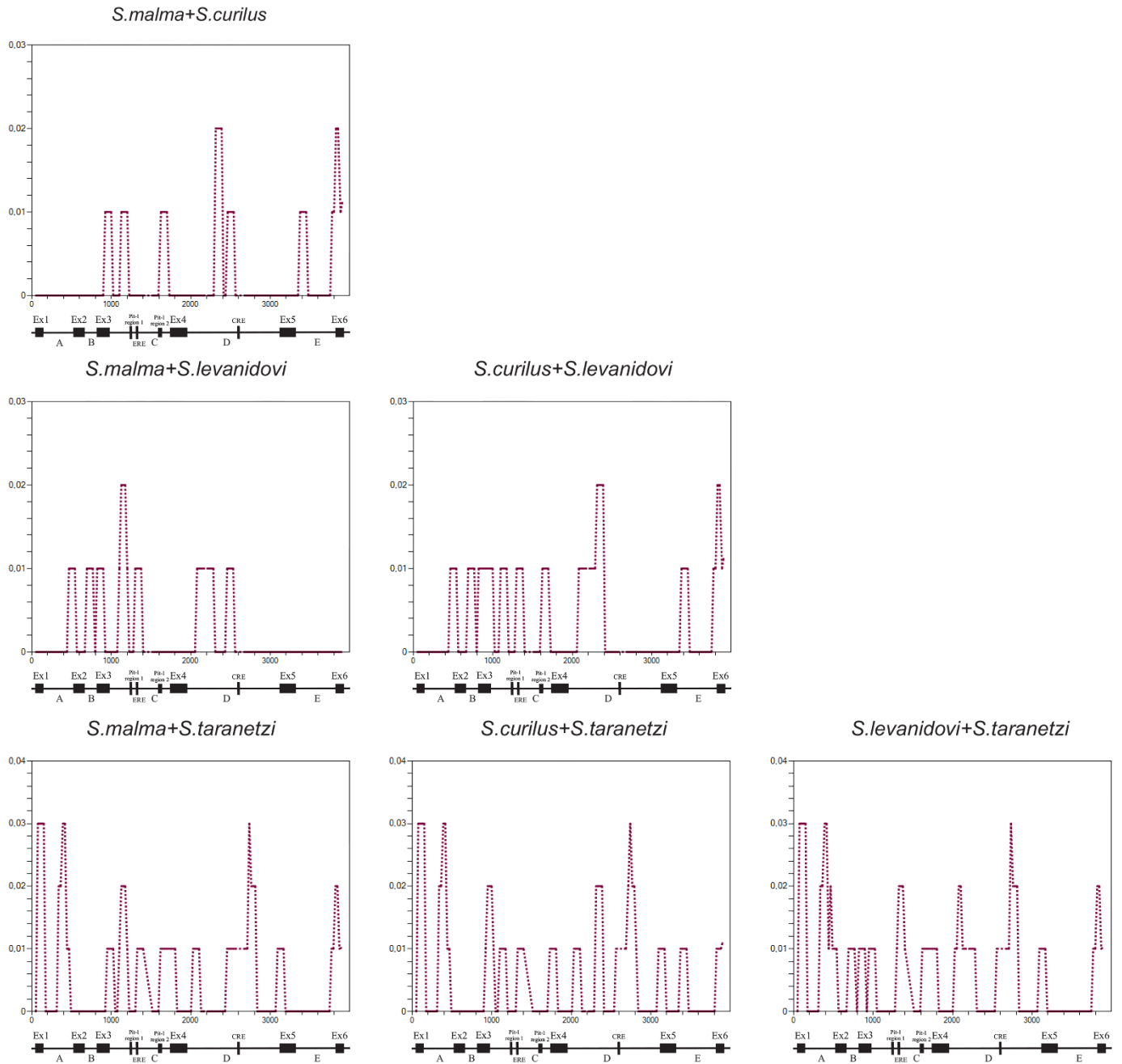


Рисунок 3.12 – Анализ уровня дивергенции гена *ghl* гольцов рода *Salvelinus* методом скользящего окна. Размер рамки – 100 п.н., длина шага – 25 п.н. По оси *x* отмечены положения нуклеотидов, по оси *y* – уровень дивергенции. Черными прямоугольниками обозначены экзоны и потенциальные регуляторные элементы. Тонкой линией обозначены интроны.

Уровень дивергенции транскрибируемой части гена *gh2* у гольцов рода *Salvelinus* значительно выше, чем в гене *gh1* (рис. 3.13). Самыми консервативными оказались последовательности видов *S. malma* – *S. curilus*, значение Dху между ними составляет 0,00272. В паре видов *S. levanidovi* – *S. taranetzi* наблюдается самое большое значение Dху (0,02309). Для всех остальных пар видов значения Dху отличаются незначительно (Приложение IV).

В гене *gh2* у гольцов рода *Salvelinus* консервативные интроны практически не встречаются. Исключение составляют северная мальма (*S. malma*) и южная азиатская мальма (*S. curilus*): при сравнении этих видов интроны А и Е совсем не имеют различий, а в интронах С и D встречаются небольшие консервативные участки (рис. 3.13). У остальных видов гольцов рода *Salvelinus* последовательности интронов изменчивы по всей длине. При сравнении последовательности гена *gh2* гольца Леванидова (*S. levanidovi*) с северной мальмой (*S. malma*) и южной азиатской мальмой (*S. curilus*) самый высокий уровень дивергенции наблюдается в интронах А и С. Сравнивая гольца Таранца (*S. taranetzi*) с тремя другими видами гольцов рода *Salvelinus* видно, что все интроны имеют достаточно низкий уровень дивергенции, только в интроне D у всех видов наблюдается один высокий пик.

На рисунках 3.14 и 3.15, объединяющих в анализе все последовательности гольцов рода *Salvelinus*, видно, что ассоциация консервативных участков с экзонами сохраняется. Пики варибельных участков приходятся на интронные последовательности участки с низким уровнем дивергенции в интронах при сравнении одновременно всех видов гольцов рода *Salvelinus* практически не встречаются. У гольцов в гене *gh1* наиболее высокие пики можно наблюдать в интронах А, D и Е, а в гене *gh2* только в интронах А и D. Интрон Е в гене *gh2* имеет средние по высоте пики. Ген *gh1* у гольцов более консервативен по сравнению с геном *gh2*.

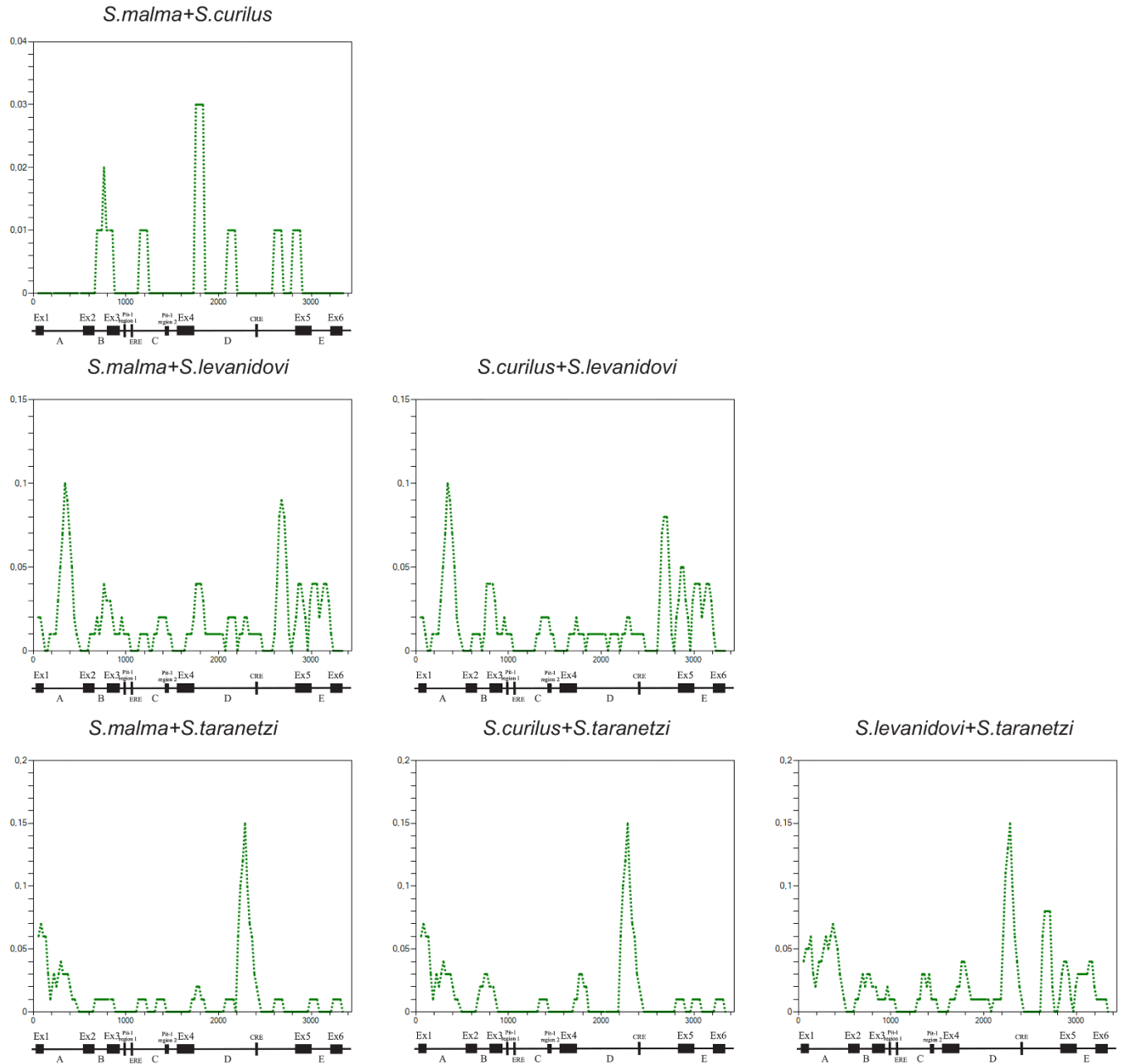


Рисунок 3.13 – Анализ уровня дивергенции гена *gh2* гольцов рода *Salvelinus* методом скользящего окна. Размер рамки – 100 п.н., длина шага – 25 п.н. По оси *x* отмечены положения нуклеотидов, по оси *y* – уровень дивергенции. Черными прямоугольниками обозначены экзоны и потенциальные регуляторные элементы. Тонкой линией обозначены интроны.

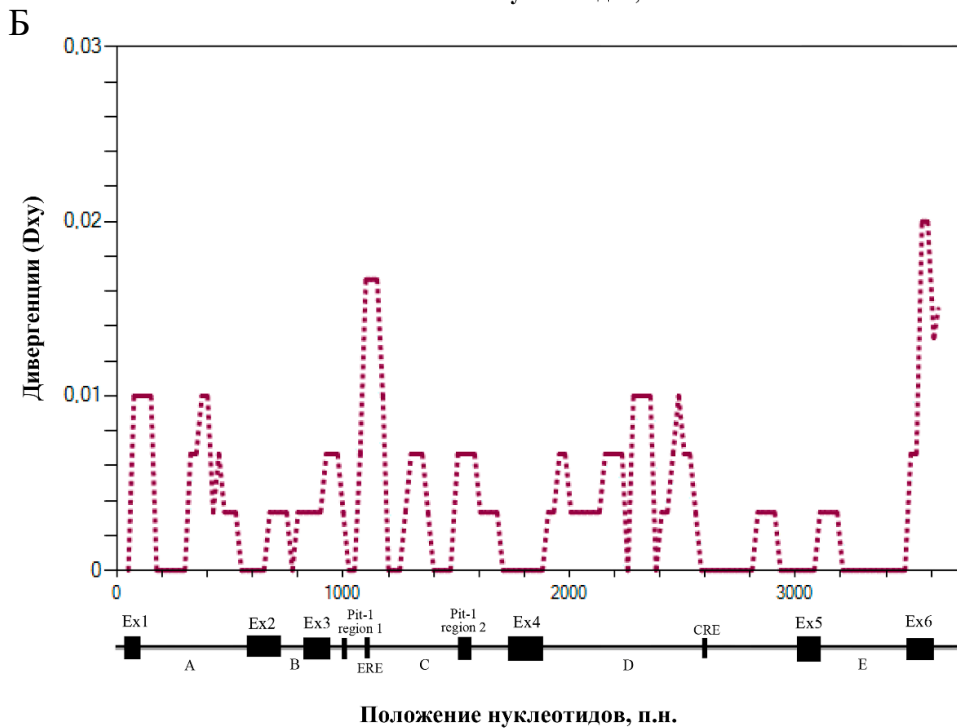
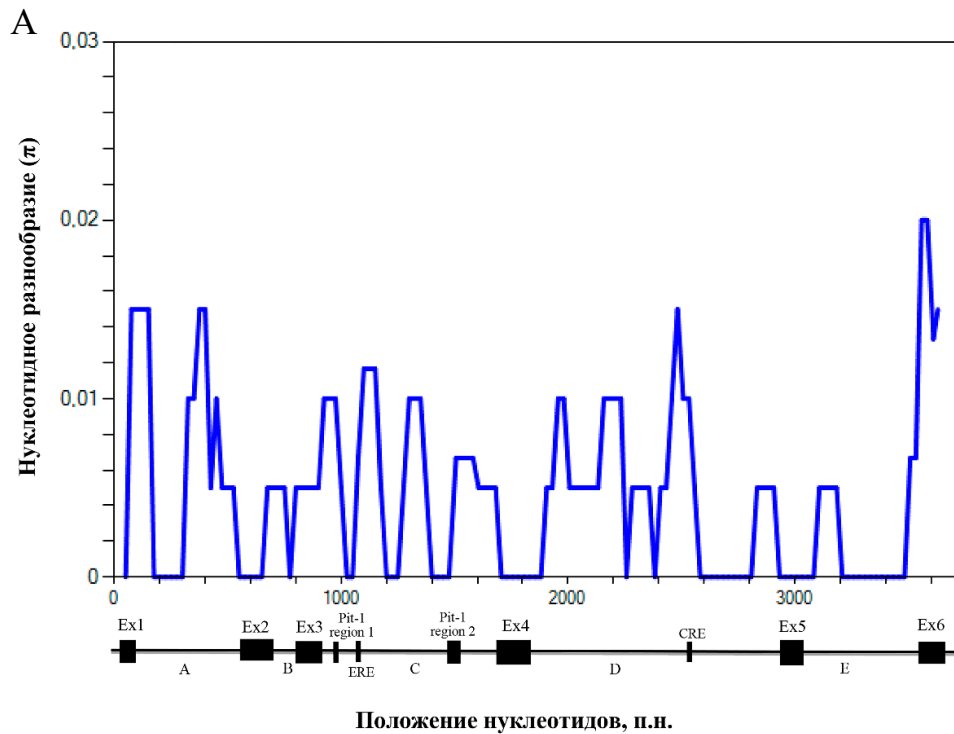


Рисунок 3.14 – Анализ нуклеотидной изменчивости и дивергенции гена *gh1* гольцов рода *Salvelinus* методом скользящего окна. Размер рамки – 100 п.н., длина шага – 25 п.н. По оси x отмечены положения нуклеотидов, по оси y – нуклеотидное разнообразие (π) (А) и дивергенция (Б). А – уровень нуклеотидной изменчивости кодирующей части гена *gh1*. Б – уровень дивергенции кодирующей части гена *gh1*. Черными прямоугольниками обозначены экзоны и потенциальные регуляторные элементы. Тонкой линией обозначены интроны

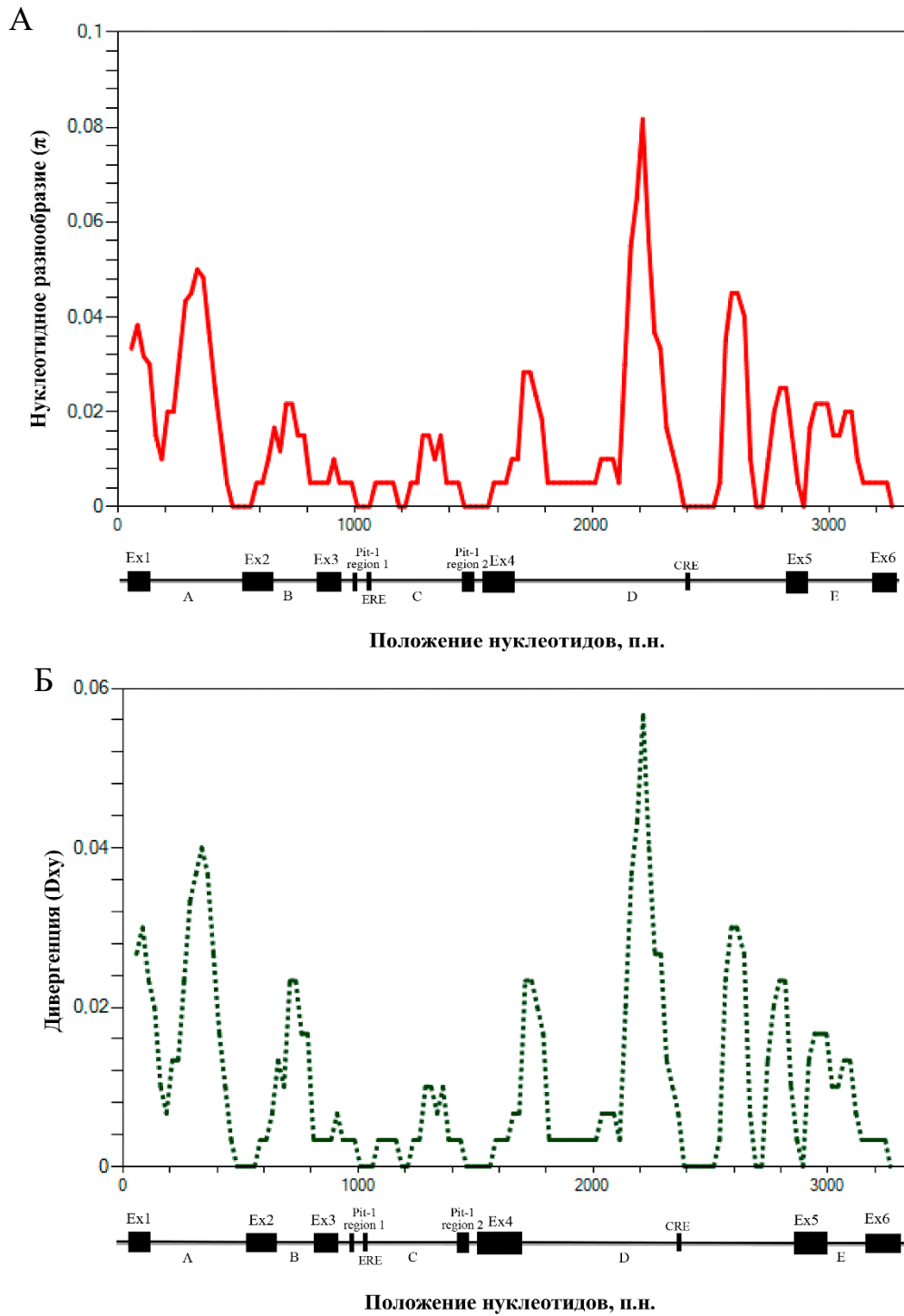


Рисунок 3.15 – Анализ нуклеотидной изменчивости и дивергенции гена *gh2* гольцов рода *Salvelinus* методом скользящего окна. Размер рамки – 100 п.н., длина шага – 25 п.н. По оси x отмечены положения нуклеотидов, по оси y – нуклеотидное разнообразие (π) (А) и уровень дивергенция (Б). А – уровень нуклеотидной изменчивости кодирующей части гена *gh2*. Б – уровень дивергенции кодирующей части гена *gh2*. Черными прямоугольниками обозначены экзоны и потенциальные регуляторные элементы. Тонкой линией обозначены интроны.

Уровень нуклеотидного разнообразия в гене *gh1* составляет $\pi=0,00446$, в гене *gh2* $\pi=0,01423$ (таблица 3.2), величина пиков изменчивости в гене *gh2* в 2–3 раза больше, чем в гене *gh1* (рис. 3.14, 3.15). Значения дивергенции для транскрибируемой части генов *gh1* и *gh2* составил 0,00364 и 0,01064, соответственно. Кривые графиков нуклеотидного разнообразия и дивергенции в значительной степени совпадают. На графиках дивергенции (рис. 3.14 Б, 3.15 Б) также видно, что экзоны имеют меньше различий, интроны более вариабельны.

Таблица 3.2 – Нуклеотидное разнообразие транскрибируемой части генов гормона роста *gh1* и *gh2*

| Наименование таксонов | Нуклеотидное разнообразие(π) \pm Стандартное отклонение | |
|---------------------------------------------------------------------------------------|-----------------------------------------------------------------|----------------------------------------|
| | Транскрибируемая часть гена <i>gh1</i> | Транскрибируемая часть гена <i>gh2</i> |
| Род <i>Salvelinus</i> | 0,00446 \pm 0,00108 | 0,01423 \pm 0,00397 |
| Семейство Salmonidae (<i>Salvelinus</i> , <i>Oncorhynchus</i> , <i>Salmo</i>) | 0,03737 \pm 0,00836 | 0,03683 \pm 0,00604 |

Сравнивая транскрибируемые участки генов *gh1* и *gh2* гольцов рода *Salvelinus* отдельно с неркой (*O. nerka*), чавычей (*O. tshawytscha*) и атлантическим лососем (*S. salar*) (рис 3.16, 3.17) можно увидеть, что частота пиков увеличивается, участков с высоким уровнем дивергенции становится больше приходятся такие участки в основном на последовательности интронов, при этом экзоны сохраняют консервативность. При сравнении гена *gh1* четырех видов гольцов рода *Salvelinus* с неркой (*O. nerka*) (рис. 3.16) участки наибольшей дивергенции встречаются в интронах С и D. Интрон А имеет самый высокий пик изменчивости при сравнении последовательностей гена *gh1* гольцов рода *Salvelinus* с чавычей (*O. tshawytscha*). Сравнивая ген *gh1* атлантического лосося (*S. salar*) и гольцов рода *Salvelinus*, видно, что общий уровень дивергенции немного меньше, чем при сравнении последовательностей гольцов с неркой (*O. nerka*) и чавычей (*O. tshawytscha*), это отражается и в значениях Dху (Приложение III). достаточно большой делецией нуклеотидов у атлантического лосося (*S. salar*).

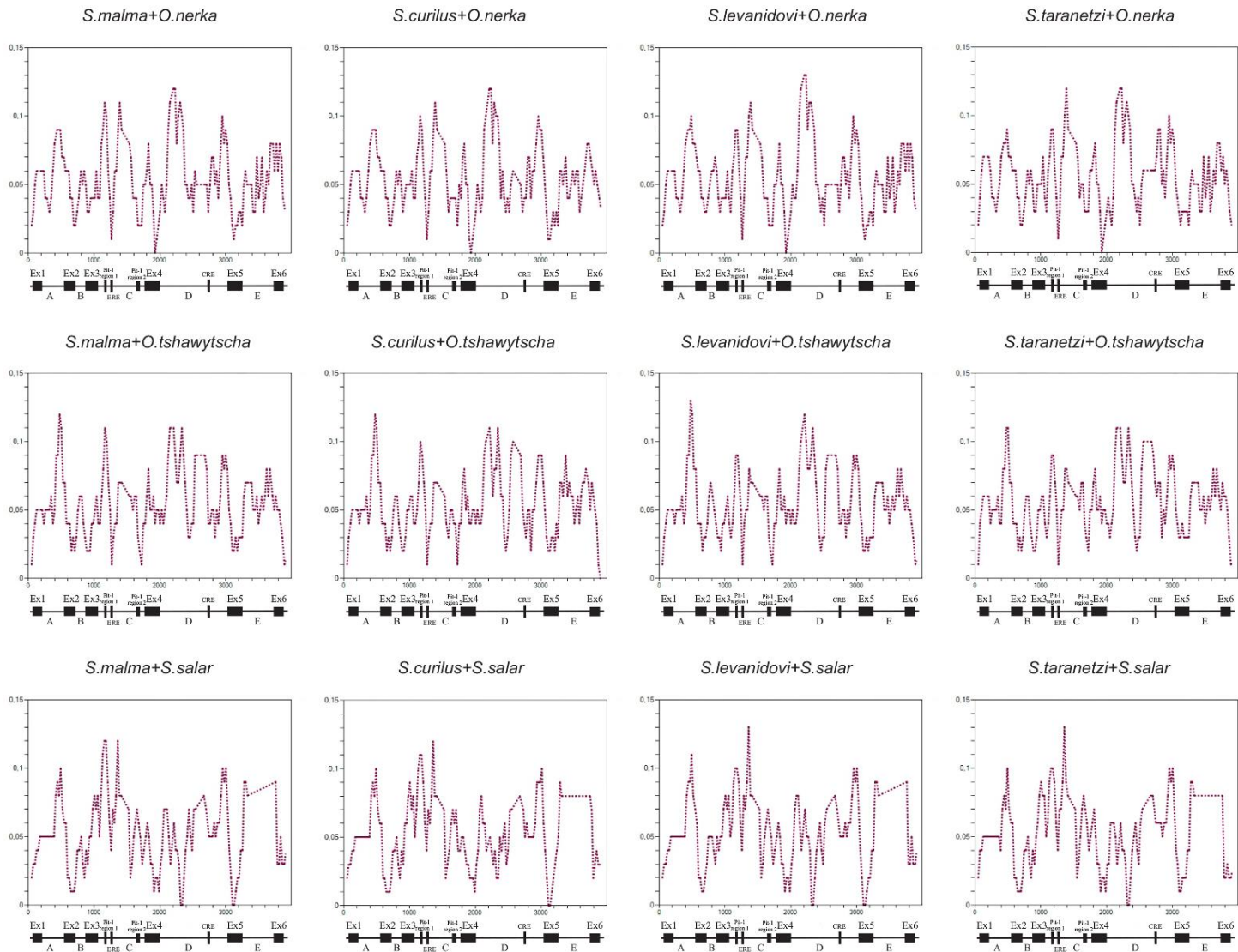


Рисунок 3.16 – Анализ уровня дивергенции кодирующей части гена *ghl* представителей семейства Salmonidae методом скользящего окна. Размер рамки – 100 п.н., длина шага – 25 п.н. По оси *x* отмечены положения нуклеотидов, по оси *y* – уровень дивергенции. Черными прямоугольниками обозначены экзоны и потенциальные регуляторные элементы. Тонкой линией обозначены интроны.

Самые высокие пики можно наблюдать в интроне С, а необычная форма пика интрона Е обусловлена достаточно большой делецией нуклеотидов у атлантического лосося (*S. salar*). В транскрибируемой части гена *gh2*, как и в транскрибируемой части гена *gh1*, наблюдается достаточно высокий уровень дивергенции (рис. 3.17). При сравнении четырех видов гольцов рода *Salvelinus* с представителями рода *Oncorhynchus* самый высокий пик находится в интроне А. У атлантического лосося (*S. salar*) и гольцов рода *Salvelinus* при сравнении последовательностей гена *gh2* самым изменчивым оказался интрон D. Стоит отметить, что при сравнении гольцов рода *Salvelinus* с остальными представителями семейства Salmonidae, имеется высокий пик, ассоциированный с участком последовательности интрона D, прилегающий непосредственно к пятому экзону.

На рисунках 3.18 и 3.19, где транскрибируемые участки генов *gh1* и *gh2* гольцов рода *Salvelinus* объединены для сравнения с последовательностями других представителей семейства Salmonidae, больших различий не выявлено. Уровень нуклеотидного разнообразия между генами отличается незначительно (таблица 3.2). Значения дивергенции также имеют небольшие различия: 0,02741 для гена *gh1* и 0,02980 для гена *gh2*. Наибольшие пики, соответствующие нуклеотидным заменам в случае гена *gh1*, приходятся на интроны А, С, D и Е, но не на интрон В (рис. 3.18), в то время как, у гена *gh2* интроны А, В и D накапливают большее количество мутаций, а относительно консервативным является интрон С (рис. 3.19). Последовательности интронов переменчивы не на всем протяжении, встречаются и консервативные области. На рисунках 3.18 и 3.19 видно, что пики изменчивости приходятся преимущественно на центральные части интронов, а участки, прилежащие к концам экзонов, более консервативны. Также области интронов, содержащие потенциальные регуляторные элементы, менее изменчивы. Например, консенсусные участки сайтов связывания с транскрипционным фактором Pit-1 (Pit-1 участок 1 и Pit-1 участок 2), которые обнаружены в интроне С обоих генов, остаются консервативными в генах *gh1* и *gh2* только при сравнении четырех видов гольцов рода *Salvelinus* (рис. 3.14, 3.15). При сравнении генов *gh1* четырех видов гольцов рода *Salvelinus* с генами гормона роста других лососевых рыб консервативность этих участков не сохраняется (рис. 3.18). В свою очередь элемент ERE в интроне С и элемент

ответа на цАМФ (CRE) в интроне D гена *gh1* не являются высококонсервативными при сравнении генов *gh1* всех видов лососевых. В гене *gh2* участки интронов, содержащие данные регуляторные элементы, менее изменчивы, чем в гене *gh1* (рис. 3.19).

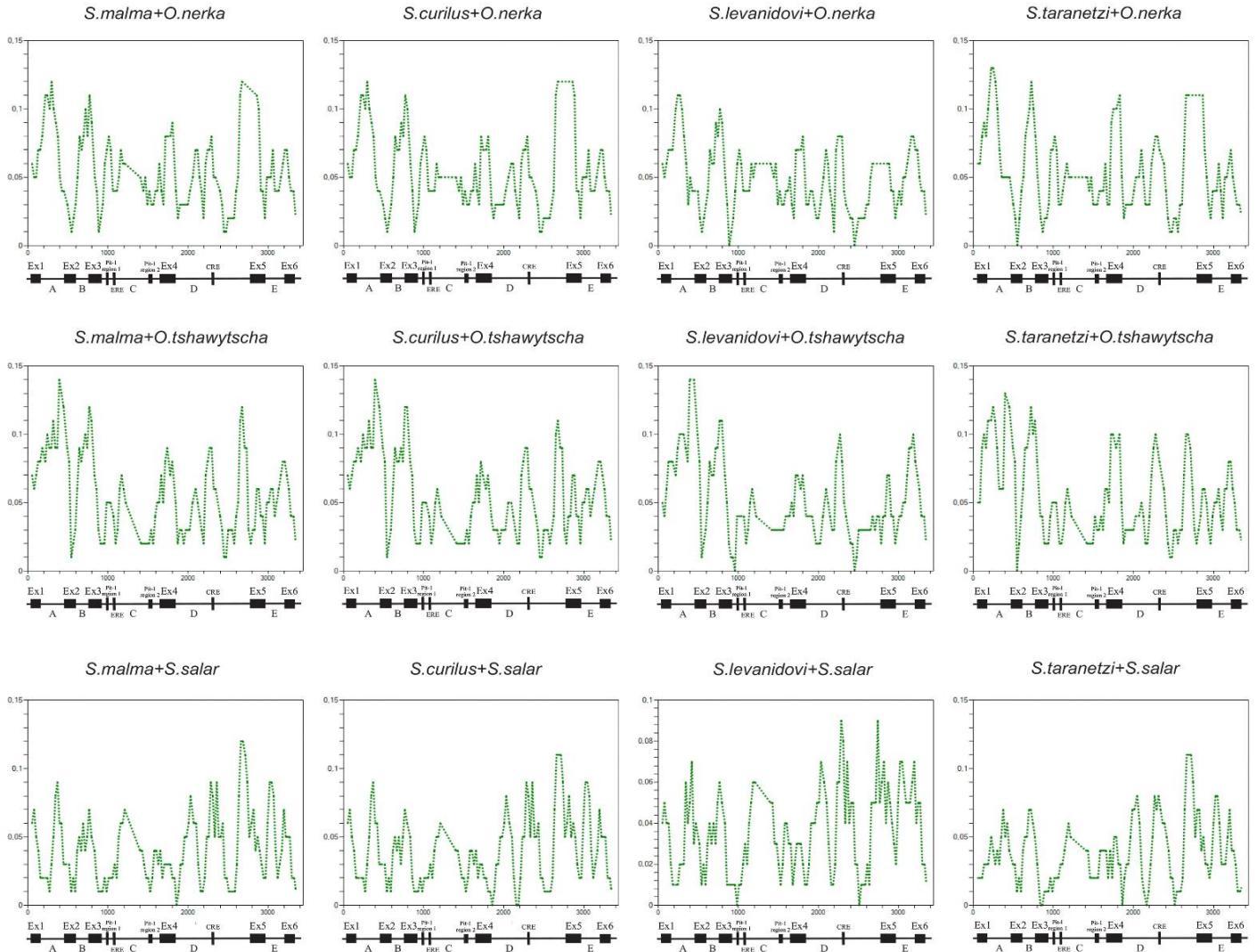


Рисунок 3.17 – Анализ уровня дивергенции кодирующей части гена *gh2* представителей семейства Salmonidae методом скользящего окна. Размер рамки – 100 п.н., длина шага – 25 п.н. По оси *x* отмечены положения нуклеотидов, по оси *y* – уровень дивергенции. Черными прямоугольниками обозначены экзоны и потенциальные регуляторные элементы. Тонкой линией обозначены интроны.

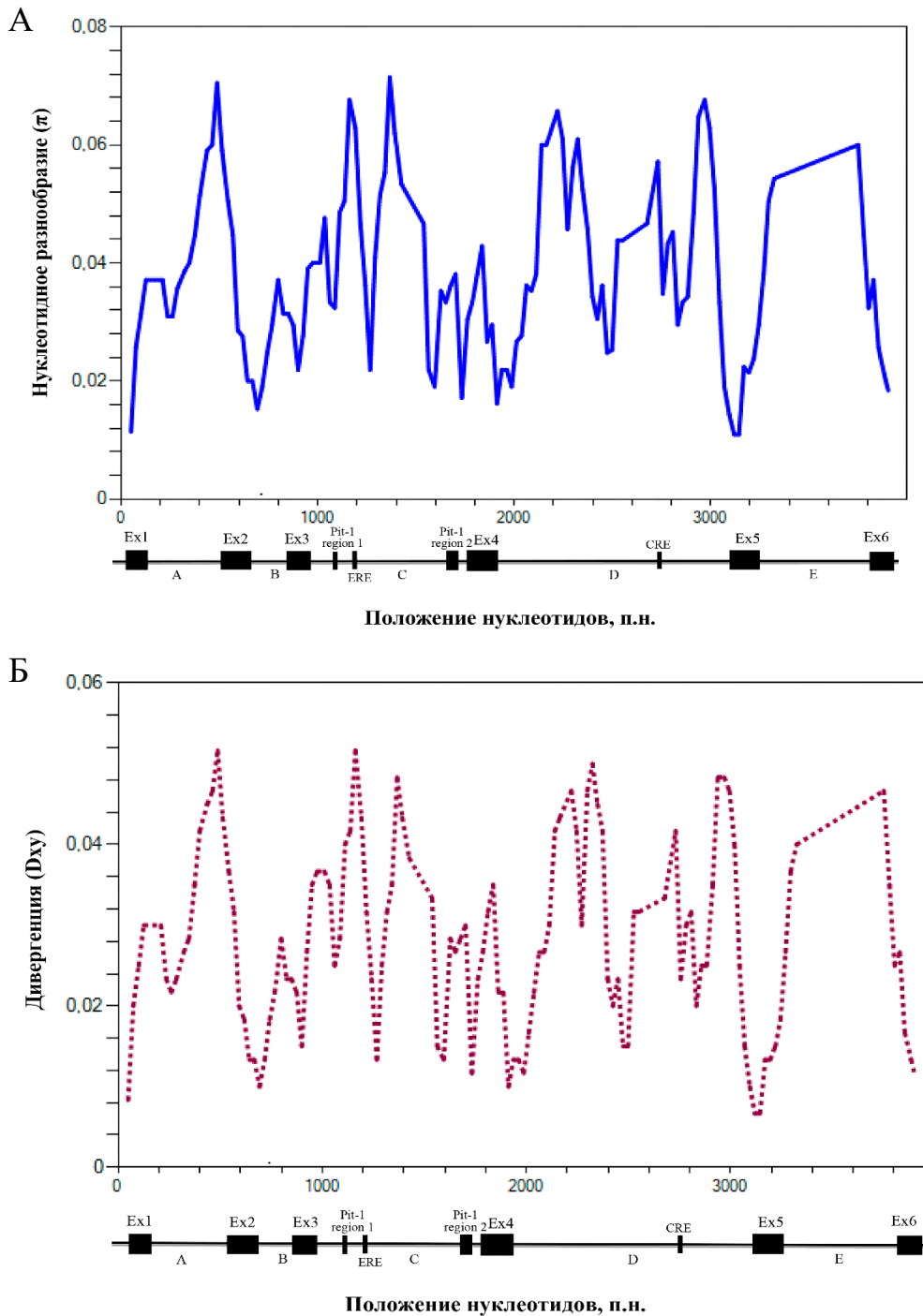


Рисунок 3.18 – Анализ нуклеотидной изменчивости и дивергенции гена *gh1* представителей семейства Salmonidae методом скользящего окна. Размер рамки – 100 п.н., длина шага – 25 п.н. По оси x отмечены положения нуклеотидов, по оси y – нуклеотидное разнообразие (А) и дивергенция (Б). А – уровень нуклеотидной изменчивости кодирующей части гена *gh1* среди представителей трех родов семейства Salmonidae: *Oncorhynchus*, *Salmo*, *Salvelinus*. Б – уровень дивергенции кодирующей части гена *gh1* среди представителей трех родов семейства Salmonidae: *Oncorhynchus*, *Salmo*, *Salvelinus*. Черными прямоугольниками обозначены экзоны и потенциальные регуляторные элементы. Тонкой линией обозначены интроны.

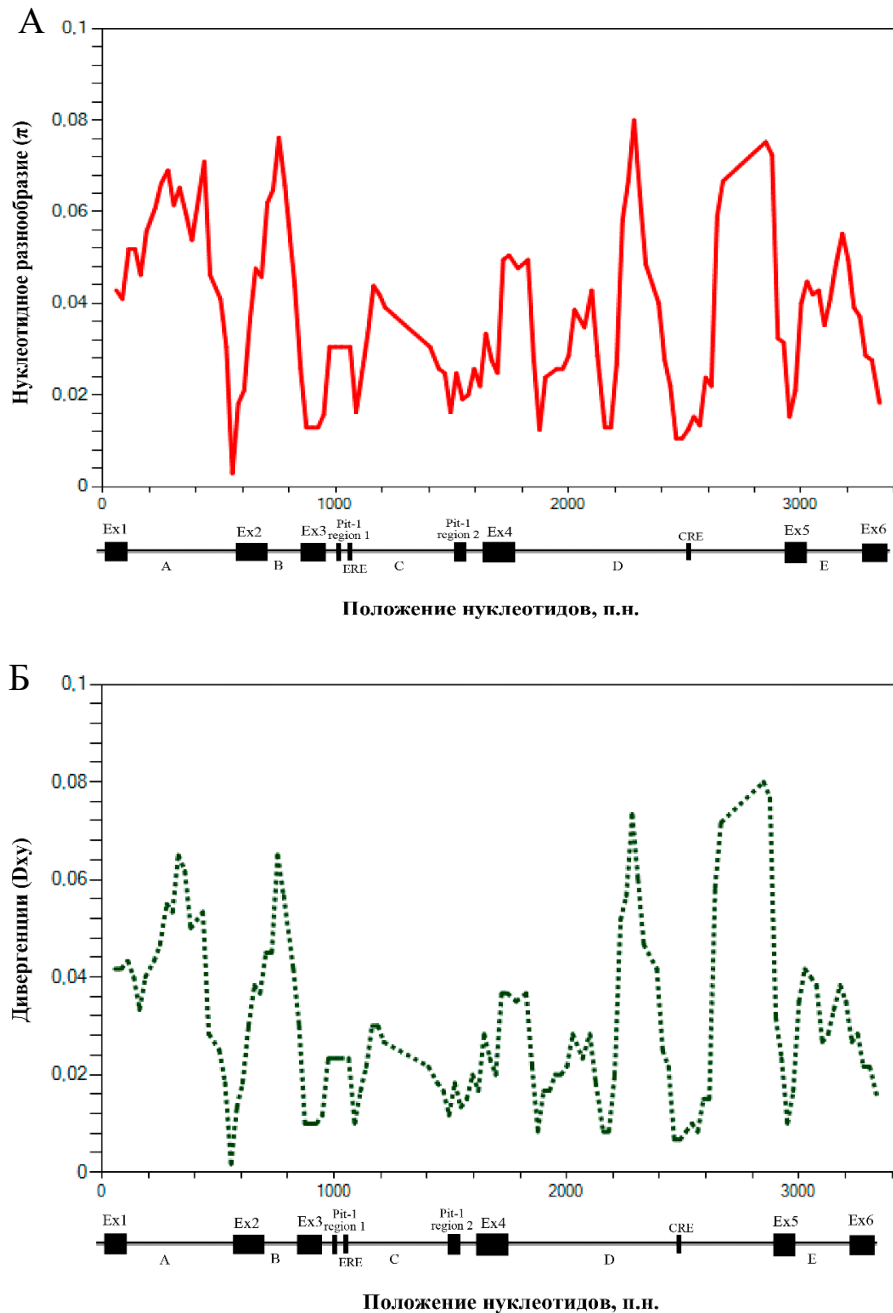


Рисунок 3.19 – Анализ нуклеотидной изменчивости и дивергенции гена *gh2* представителей семейства Salmonidae методом скользящего окна. Размер рамки – 100 п.н., длина шага – 25 п.н. По оси x отмечены положения нуклеотидов, по оси y – нуклеотидное разнообразие (А) и дивергенция (Б). А – уровень нуклеотидной изменчивости кодирующей части гена *gh2* среди представителей трех родов семейства Salmonidae: *Oncorhynchus*, *Salmo*, *Salvelinus*. Б – уровень дивергенции кодирующей части гена *gh2* среди представителей трех родов семейства Salmonidae: *Oncorhynchus*, *Salmo*, *Salvelinus*. Черными прямоугольниками обозначены экзоны и потенциальные регуляторные элементы. Тонкой линией обозначены интроны.

Для того чтобы показать действие отрицательного отбора на разные участки гена, был проведен сравнительный анализ уровня изменчивости внутри каждого экзона и интрона среди представителей трех родов семейства Salmonidae. В гене *gh1* у всех исследованных видов лососевых короткие экзоны оказались менее вариабельными. Самым консервативным оказался шестой экзон (63 п.н.), и также незначительный уровень нуклеотидной изменчивости ($\pi=0,00386$) выявлен в первом экзоне (74 п.н.). В остальных экзонах гена *gh1*: втором (140 п.н.), третьем (117 п.н.), четвертом (156 п.н.) и пятом (147 п.н.) – уровень нуклеотидной изменчивости колеблется от 0,018 до 0,029. Наибольшее нуклеотидное разнообразие в гене *gh1* характерно для четвертого экзона $\pi=0,0293$ (табл. 3.3).

Таблица 3.3 – Нуклеотидное разнообразие в экзонах генов гормона роста *gh1* и *gh2*

| Экзоны | <i>gh1</i> | <i>gh2</i> |
|------------|----------------------------------------------------------------|----------------------------------------------------------------|
| | Нуклеотидное разнообразие (π) ±Стандартное отклонение | Нуклеотидное разнообразие (π) ±Стандартное отклонение |
| Эхон 1 | 0,00386±0,00265 | 0,03501±0,00633 |
| Эхон 2 | 0,02585±0,00745 | 0,03946±0,00851 |
| Эхон 3 | 0,02768±0,00599 | 0,01099±0,00236 |
| Эхон 4 | 0,0293±0,00642 | 0,02869±0,00563 |
| Эхон 5 | 0,01846±0,00507 | 0,03693±0,00654 |
| Эхон 6 | 0 | 0,00756±0,00272 |
| Все экзоны | 0,02070±0,00480 | 0,02832±0,00482 |

В гене *gh2*, как и в гене *gh1*, наименьшую изменчивость можно отметить для шестого экзона (табл. 3.3). Однако уровень изменчивости первого экзона в гене *gh2* гораздо выше, чем в гене *gh1* и составляет $\pi=0,03501$. Несмотря на то, что длина третьего экзона составляет 117 п.н., уровень нуклеотидного разнообразия не самый высокий $\pi=0,01099$ и по сравнению с третьим экзоном гена *gh1*, где $\pi=0,02768$, третий экзон в гене *gh2* более консервативен. Наибольший уровень нуклеотидного разнообразия характерен

для второго экзона гена *gh2* $\pi=0,03946$. Уровень нуклеотидного разнообразия пятого экзона составляет $\pi=0,03693$. Первый и пятый экзон в гене *gh2* более изменчивы, чем первый и пятый экзон в гене *gh1*. Уровень изменчивости четвертого экзона в гене *gh2* оказался такой же, как и в четвертом экзоне гена *gh1* $\pi=0,02869$, однако изменчивость у этого экзона одна из самых низких по сравнению с другими экзонами гена *gh2*.

Уровень нуклеотидного разнообразия в интронах, как правило, выше, чем в экзонах, поскольку интроны – это некодирующие последовательности и находятся под меньшим давлением отрицательного отбора. В среднем уровень изменчивости по интронам не превышает значения 0,05, за исключением интрона В гена *gh2*. Несмотря на то, что интрон В самый короткий из всех интронов генов гормона роста, именно в интроне В гена *gh2* наблюдается наибольший уровень нуклеотидного разнообразия $\pi=0,05982$. В свою очередь интрон В гена *gh1* является самым консервативным. Уровень изменчивости интрона В гена *gh1* составляет $\pi=0,02963$, что соответствует уровню изменчивости некоторых экзонов (табл. 3.3). Наибольший уровень изменчивости в гене *gh1* наблюдается в интроне Е – $\pi=0,05007$. Изменчивость интрона А меньше изменчивости интрона Е и составляет $\pi=0,04699$. Уровень изменчивости самых длинных интронов гена *gh1* С и D составляет всего $\pi=0,04244$ и $0,04190$, соответственно (табл. 3.4).

В гене *gh2* уровень изменчивости в длинных интронах С и D один из самых низких. Несмотря на то, что у четырех видов гольцов рода *Salvelinus* интрон С второй по величине, а у представителей рода *Oncorhynchus* и *Salmo* он немного короче и сопоставим по длине с интроном А, уровень изменчивости в интроне С всего $\pi=0,02835$. В свою очередь, в интроне А наблюдается одно из самых высоких значений уровня изменчивости $\pi=0,05153$. В интроне D уровень изменчивости составляет $\pi=0,03607$. Интрон Е в гене *gh2* имеет небольшие размеры, однако уровень нуклеотидного разнообразия составляет $\pi=0,04321$ (табл. 3.4).

Таблица 3.4 – Нуклеотидное разнообразие в интронах

генов *gh1* и *gh2* лососевых рыб

| Интрон | Salmonidae, вид | <i>gh1</i> | | <i>gh2</i> | |
|--------|-----------------------|-----------------|---------------------------------------------------------------------------|-----------------|---------------------------------------------------------------------------|
| | | Длина (п.н.) | Нуклеотидное разнообразие (π) \pm Стандартное отклонение | Длина (п.н.) | Нуклеотидное разнообразие (π) \pm Стандартное отклонение |
| A | <i>O. nerka</i> | 392 | 0,04699 \pm 0,00934 | 470 | 0,05153 \pm 0,01034 |
| | <i>O. tshawytscha</i> | 403 | | 457 | |
| | <i>S. salar</i> | 464 | | 461 | |
| | <i>S. curilus</i> | 455 | | 463 | |
| | <i>S. levanidovi</i> | 455 | | 463 | |
| | <i>S. malma</i> | 455 | | 463 | |
| | <i>S. taranetzi</i> | 455 | | 465 | |
| B | <i>O. nerka</i> | 136 | 0,02963 \pm 0,00646 | 138 | 0,05982 \pm 0,01198 |
| | <i>O. tshawytscha</i> | 136 | | 138 | |
| | <i>S. salar</i> | 142 | | 135 | |
| | <i>S. curilus</i> | 136 | | 124 | |
| | <i>S. levanidovi</i> | 136 | | 123 | |
| | <i>S. malma</i> | 136 | | 124 | |
| | <i>S. taranetzi</i> | 136 | | 123 | |
| C | <i>O. nerka</i> | 710 | 0,04244 \pm 0,01050 | 445 | 0,02835 \pm 0,00666 |
| | <i>O. tshawytscha</i> | 723 | | 445 | |
| | <i>S. salar</i> | 821 | | 455 | |
| | <i>S. curilus</i> | 715 | | 624 | |
| | <i>S. levanidovi</i> | 721 | | 623 | |
| | <i>S. malma</i> | 721 | | 624 | |
| | <i>S. taranetzi</i> | 721 | | 623 | |
| D | <i>O. nerka</i> | 1145 | 0,04190 \pm 0,00976 | 964 | 0,03607 \pm 0,00466 |
| | <i>O. tshawytscha</i> | 1170 | | 1108 | |
| | <i>S. salar</i> | 1117 | | 1173 | |
| | <i>S. curilus</i> | 1043 | | 1171 | |
| | <i>S. levanidovi</i> | 1043 | | 1168 | |
| | <i>S. malma</i> | 1044 | | 1171 | |
| | <i>S. taranetzi</i> | 1044 | | 1171 | |
| E | <i>O. nerka</i> | 586 | 0,05007 \pm 0,01030 | 220 | 0,04321 \pm 0,00923 |
| | <i>O. tshawytscha</i> | 592 | | 217 | |
| | <i>S. salar</i> | 218 | | 221 | |
| | <i>S. curilus</i> | 620 | | 226 | |
| | <i>S. levanidovi</i> | 620 | | 225 | |
| | <i>S. malma</i> | 620 | | 226 | |
| | <i>S. taranetzi</i> | 622 | | 226 | |
| | Все интроны | | 0,04218 \pm 0,00944 | | 0,03958 \pm 0,00646 |

Если сравнивать значения нуклеотидного разнообразия в экзонах, интронах и промоторных областях, то промоторы, несмотря на большое количество сайтов связывания с транскрипционными факторами, которые достаточно консервативны внутри одного семейства у разных видов, оказываются ближе по уровню изменчивости к некодирующим участкам – интронам, чем к кодирующим – экзонам.

Различия в нуклеотидном составе также могут указывать на разное действие отбора на отдельные участки гена. Содержание GC в интронах генов гормона роста *gh1* и *gh2* достаточно однородное и не превышает 40%. Интроны, имеющие экстремально высокие или наоборот очень низкие значения по содержанию GC в генах гормона роста не обнаружены. Наименьшее содержание GC в гене *gh1* приходится на интрон В (36,9%), а наибольшее на интрон С (39,6%). В гене *gh2* на интрон В наоборот приходится самое большое содержание GC-пар 39,2%, а самое низкое значение у интрона Е (35,5%). Взаимосвязь между длиной интронов и содержанием GC-пар также не прослеживается. Интрон D, как самый длинный, является не самым насыщенным GC: 37,2% в гене *gh1* и 36,3% в гене *gh2*. А самый короткий интрон В только в гене *gh2* имеет высокий уровень содержания GC. В том же гене *gh2* интрон Е характеризуется небольшой длиной, но уровень содержания GC самый низкий по всем интронам (табл. 3.5).

Таблица 3.5 – Содержание GC-пар в интронах генов *gh1* и *gh2* лососевых рыб

| Участки гена | G+C (%) | |
|--------------|------------|------------|
| | <i>gh1</i> | <i>gh2</i> |
| Интрон А | 38,4 | 37,6 |
| Интрон В | 36,9 | 39,2 |
| Интрон С | 39,6 | 38,5 |
| Интрон D | 37,2 | 36,3 |
| Интрон Е | 38,1 | 35,5 |

3.3. Филогения генов гормона роста

Филогенетическая схема дивергенции транскрибируемой части генов-паралогов на основе Байесовского подхода (BI) приведена на рисунке 3.20. Филогенетическое дерево, построенное на основании NJ метода, имеет сходную топологию. Видно, что каждый из генов-паралогов всех видов объединяются в отдельный кластер. Внутри каждого из кластеров виды родов *Salvelinus*, *Oncorhynchus* и *Salmo* формируют отдельные субкластеры. Исключением является ген *gh2 Salmo salar*, который выступает в положении сестринского таксона по отношению к гольцам рода *Salvelinus*.

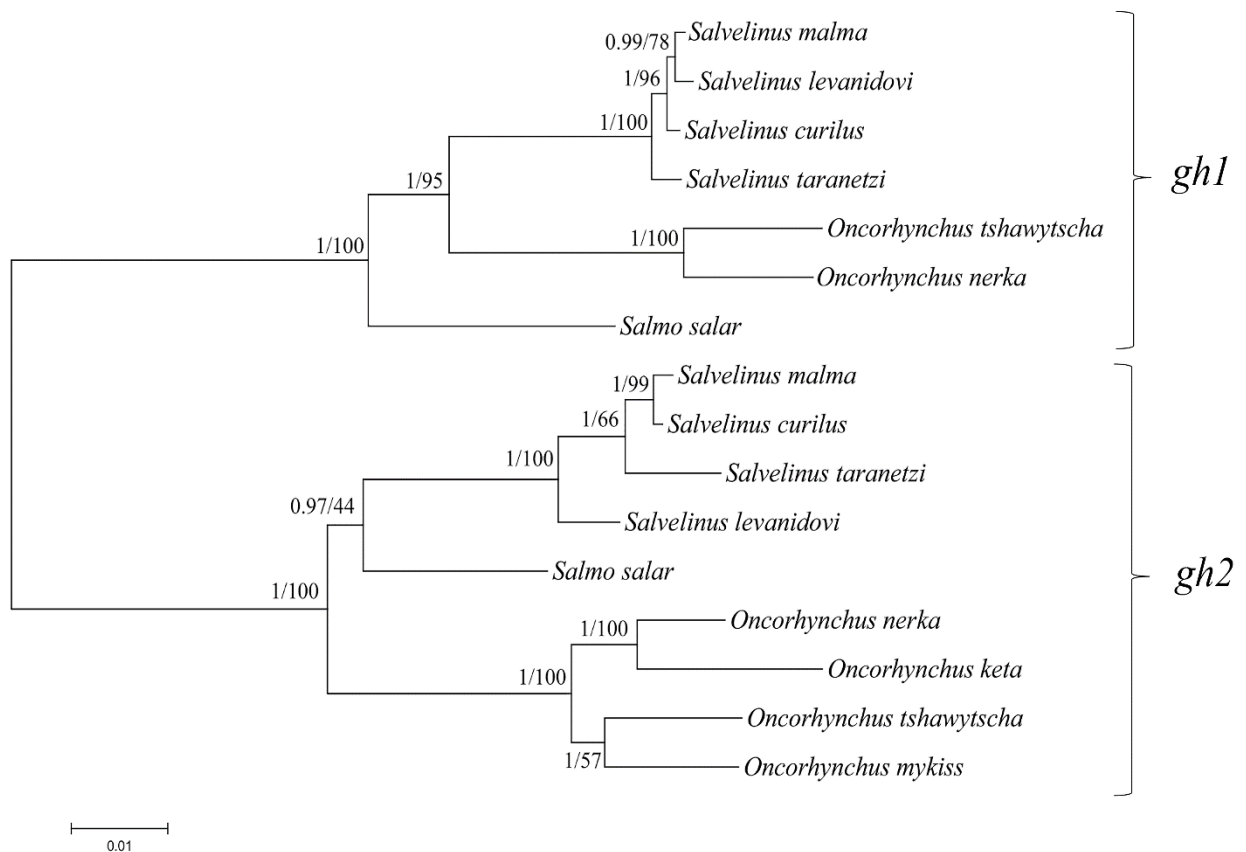


Рисунок 3.20 – Филогенетическая реконструкция, выполненная при помощи байесовского метода. В узлах ветвей указаны апостериорные вероятности/бутстреп-поддержки для деревьев BI/NJ (Каменская и др., 2017, с модификациями).

Схема филогении объединенных последовательностей генов-паралогов, включающих регуляторную часть и транскрибируемую части, приведена на рисунке 3.21. Топология NJ дерева идентична топологии дерева, построенному Байесовским методом. Кластеры сформированы согласно родовой принадлежности видов, *Salmo salar* формирует неразрешенную ветвь на этом дереве.

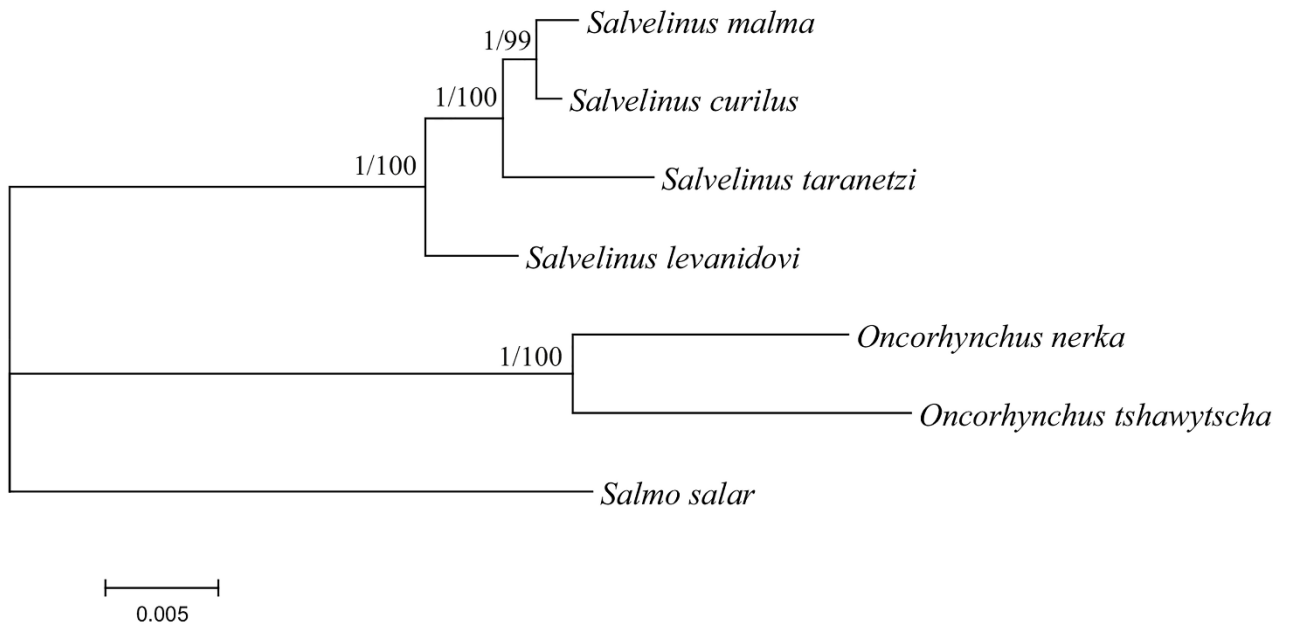


Рисунок 3.21 – Филогенетическая реконструкция, объединенных последовательностей генов гормона роста *gh1* и *gh2* выполненная при помощи байесовского метода. В узлах ветвей указаны апостериорные вероятности/бутстреп-поддержки для деревьев BI/NJ.

ГЛАВА 4. ОБСУЖДЕНИЕ

4.1. Структура промоторных участков гена гормона роста лососевых

Исследование промоторных областей паралогичных генов гормона роста *gh1* и *gh2* у четырех видов гольцов рода *Salvelinus* (*S. malma*, *S. curilus*, *S. levanidovi* и *S. taranetzi*) и сравнение их с промоторами генов гормона роста нерки (*O. nerka*), чавычи (*O. tshawytscha*) и атлантического лосося (*S. salar*) показали, что на участке длиной 300 п.н. перед точкой инициации транскрипции, сосредоточены ключевые сайты связывания с транскрипционными факторами и другими лигандами, которые отвечают за формирование полноценного транскрипционного комплекса и эффективную тканеспецифичную экспрессию гена гормона роста в соматотрофах. В промоторах обоих генов всех исследованных видов гольцов встречаются последовательность ТАТА-бокса, сайты связывания с гипофиз-специфичным транскрипционным фактором Pit-1 (F1–F4), элементы ответа на цАМФ (CRE), глюкокортикоиды (GRE) и сайты связывания с рецептором ретиноевой кислоты (RARE/RXRE).

ТАТА-бокс является обязательным элементом промотора во всех описанных генах гормона роста позвоночных животных. В большинстве промоторов ТАТА-бокс находится на расстоянии около 25–30 п.н. от точки начала транскрипции, но у некоторых представителей парнокопытных ТАТА-бокс располагается дальше (Lioupis et al., 1997; Wallis, Wallis, 2001). У четырех исследованных нами видов гольцов последовательность ТАТА-бокса, состоящая из шести нуклеотидов, была найдена в обоих генах-паралогах на расстоянии –22 п.н. от точки начала транскрипции. При сравнении последовательностей промотора гольцов с промоторами атлантического лосося (*S. salar*), нерки (*O. nerka*) и чавычи (*O. tshawytscha*) было показано, что у этих видов ТАТА-бокс также состоит из шести нуклеотидов и находится на расстоянии –22 п.н. от точки начала транскрипции в генах *gh1* и *gh2*, соответственно. Поскольку у лососевых, как и у представителей семейства спаровые (Almuly et al., 2005) и цихловые (Sekkali et al., 1999), положение и последовательность нуклеотидов, составляющих ТАТА-бокс, совпадают, можно говорить о высокой консервативности ТАТА-бокса в генах гормона роста у рыб из разных отрядов.

В промоторах гена гормона роста могут встречаться дополнительные ТАТА-боксы. Два дополнительных ТАТА-боксов было описано в промоторной области генов гормона роста радужной форели (*Oncorhynchus mykiss*) (Yang et al., 1997). У курицы в гене гормона роста помимо основного ТАТА-боксов был описан еще один (Tanaka et al., 1992). Однако о взаимодействии транскрипционного фактора ТВР с дополнительными ТАТА-боксами и непосредственного их участия в образовании транскрипционного комплекса ничего не известно. У лососевых на участке промотора длиной в 300 п.н. дополнительных ТАТА-боксов нам обнаружить не удалось.

Для формирования полноценного транскрипционного комплекса обязательных элементов промотора, таких как ТАТА-боксов, как правило, бывает недостаточно. В промоторной области должны присутствовать последовательности, с которыми будут взаимодействовать определенные тканеспецифические факторы транскрипции. За транскрипцию генов гормона роста непосредственно в соматотрофах отвечает гипофиз-специфичный транскрипционный фактор Pit-1. Для успешной транскрипции гена гормона роста необходимо взаимодействие транскрипционного фактора Pit-1 с несколькими сайтами связывания. У большинства позвоночных животных (Walker, 1991; Ohkubo et al., 1996; Liouris et al., 1997; Maniou et al., 2004), в том числе и у представителей разных отрядов рыб (Tanaka et al., 1995), в промоторе гена гормона роста присутствуют по два сайта связывания с транскрипционным фактором Pit-1. Однако при исследовании взаимодействия фактора транскрипции Pit-1 с промотором одного из генов гормона роста радужной форели Ямада с соавторами (Yamada et al., 1993) показали, что фактор Pit-1 в промоторе связывается не с двумя сайтами, а с четырьмя. При сравнении промоторных областей генов гормона роста четырех видов гольцов рода *Salvelinus*, атлантического лосося (*S. salar*), нерки (*O. nerka*) и чавычи (*O. tshawytscha*) были найдены четыре сайта связывания с транскрипционным фактором Pit-1 в гене *gh1* и в гене *gh2* у каждого вида. Нуклеотидная последовательность и положение сайта F1 достаточно консервативны между двумя генами *gh* у всех исследуемых видов лососевых. В гене *gh1* сайт F1 имеет немного больше различий между видами, чем в гене *gh2*. При сравнении с сайтом F1 гольцов три замены встречается у атлантического лосося (*S. salar*), две замены у нерки

(*O. nerka*), и одна замена у чавычи (*O. tshawytscha*). Замены сосредоточены на концах последовательности и не затрагивают центральную часть сайта F1 (рис. 3.9).

Сайт F1 в гене *gh2* гольцов совпадает с данным сайтом у *O. tshawytscha* и отличается одним нуклеотидом от сайта F1 *O. nerka* и *S. salar* (рис. 3.11). В сайтах F2 и F3 у лососевых найдено больше различий по сравнению с сайтом F1. Положение сайта F2 консервативно внутри каждого гена. Границы сайта F3 в гене *gh2* совпадают у всех видов лососевых. Положение сайта F3 в гене *gh1* смещено на один нуклеотид у гольцов (–154/–182 п.н.) по сравнению с остальными видами лососевых (–155/–183 п.н.). Сайт F4 оказался самым переменчивым по сравнению с другими сайтами. Его положение идентично в обоих генах, за исключением гена *gh1* *S. salar* (–258/–289 п.н.). Каждый F-сайт богат нуклеотидами А и Т и содержит консенсусный мотив (Т/А)NCTNCAT, предложенный Окубо с соавторами, как универсальный для костистых рыб и птиц (Okubo et al., 1996). На основании полученных данных, лососевых рыб можно называть уникальной группой позвоночных животных, у которых на прилегающем к транскрибируемой части гена промоторном участке было обнаружено не два сайта связывания с Pit-1, а четыре.

В некоторых случаях, в частности при оценке уровня экспрессии гена гормона роста человека hGH-N в системе *in vivo* было показано, что двух сайтов связывания с фактором Pit-1 находящихся в промоторе для обеспечения активной транскрипции не достаточно, и необходимо присутствие сверхчувствительных к ДНКазе I сайтов (HS), один из которых, HSI, содержит три дополнительных сайта связывания с транскрипционным фактором Pit-1. У человека эти сверхчувствительные сайты расположены на расстоянии начиная от –15 т.п.н. до –32 т.п.н. от основного промотора (Jones et al., 1995; Shewchuk et al., 2006). У лососевых настолько протяженная последовательность никогда не исследовалась и однозначно нельзя сказать будут ли найдены у лососевых рыб дополнительные сайты связывания с Pit-1 на таком большом расстоянии от точки начала транскрипции. Консенсусный мотив одного из дополнительных сайтов Pit1-C, был найден в составе сайта F3 гена *gh2* лососевых, но на основании сравнения только нуклеотидной последовательности нельзя сделать вывод, что участок сайта F3 гена гормона роста лососевых является предшественником сайта связывания с Pit-1 (Pit1-C), локализованного в HSI человека. Однако, согласно работам по оценке эффективности экспрессии гена

гормона роста лососевых, присутствие в промоторе сайтов F2 и F3, с которыми связывается транскрипционный фактор Pit-1, увеличивает уровень экспрессии гена гормона роста лососевых до оптимального, по аналогии с сайтами HS для гена гормона роста человека (Argenton et al., 2002; Shewchuk et al., 2006).

Увеличение уровня экспрессии генов гормона роста может происходить не только за счет взаимодействия транскрипционного фактора Pit-1 с сайтами F2 и F3, но и из-за присутствия на этом участке сайтов связывания с рецепторами гормонов и другими факторами, которые вместе с Pit-1 участвуют в процессе активации транскрипции (Argenton et al., 1996; Almuly et al., 2005). Элемент ответа на глюкокортикоиды (GRE) и сайт связывания с рецептором ретиноевой кислоты (RARE/RXRE) перекрывают сайт F2 в промоторе у всех исследованных видов лососевых. С проксимальной частью сайта F2 совпадает участок последовательности RARE/RXRE, а большая часть элемента ответа на глюкокортикоиды GRE перекрывает дистальную часть сайта F2. Свободной от перекрытия в сайте F2 остается область из четырех нуклеотидов, которая у представителей рода *Oncorhynchus* в обоих генах содержит разные типы мутаций (транзиции и трансверсии). Возможно, именно мутации обеспечивают пластичность данного участка при взаимодействии с транскрипционными факторами (рис. 3.9 и 3.11).

Последовательность сайта связывания с рецептором ретиноевой кислоты RARE/RXRE, описанная ранее для чавычи и атлантического лосося (Von Schalburg et al., 2008), также была найдена нами в промоторе нерки и четырех видов гольцов рода *Salvelinus* (Каменская и др., 2017). С рецептором связываются две области, разделенные восьмью парами нуклеотидов. Участки, ответственные за взаимодействие с рецептором ретиноевой кислоты консервативны в обоих генах-паралогах у всех исследованных видов лососевых, за исключением гена *gh2* чавычи: в проксимальной части элемента первый аденин заменен на цитозин. В последовательности, разделяющей участки связывания с рецептором ретиноевой кислоты, имеются единичные замены и в гене *gh1*, и в гене *gh2*. В гене *gh1* только у нерки наблюдается транзиция: один из аденинов заменен на гуанин. А в гене *gh2* происходит трансверсия: за первым аденином у гольцов и атлантического лосося следует гуанин, а у нерки и чавычи цитозин.

Элемент ответа на глюкокортикоиды (GRE), как и сайт связывания с рецептором ретиноевой кислоты, состоит из двух участков. Но, в отличие от RARE/RXRE, участки связывания с рецептором глюкокортикоидов расположены ближе к друг другу, поскольку их разделяет только три нуклеотида, а не восемь как в случае RARE/RXRE. Проксимальная часть элемента GRE (TGTTCC), которая совмещена с сайтом F2, в гене *gh1* идентична у всех видов лососевых, а в гене *gh2* у нерки имеется одна замена: вместо гуанина – цитозин. От последовательности GRE, описанной для млекопитающих (Beato et al., 1989), проксимальная часть элемента GRE лососевых отличается только на один нуклеотид: на 3'-конце у лососевых – цитозин, а у млекопитающих – тимин. Высокая консервативность данного участка при сравнении среди разных групп позвоночных указывает на его важную роль при взаимодействии с рецепторами глюкокортикоидов.

Дистальная часть элемента ответа на глюкокортикоиды в гене *gh1* содержит только одну замену: у представителей рода *Oncorhynchus* в последовательности GGT в первом положении гуанин, а не аденин, как у гольцов и атлантического лосося. Еще одна замена у нерки и чавычи в элементе GRE гена *gh1* присутствует в последовательности, разделяющей сайты связывания с рецептором гормона, у них на этом участке вместо гуанина – аденин. В гене *gh2* нет замен в последовательности, разделяющей сайты GRE, но встречается несколько транзиций в дистальной области. У гольцов в дистальной части GRE между двумя тимидами стоит гуанин (TGT), у нерки гуанин заменяется на аденин (TAT). По две замены в этом участке обнаружено у атлантического лосося и чавычи: одна замена также во втором положении, когда гуанин меняется на аденин, а вторая транзиция – это замена тимина на цитозин, только у чавычи эта замена происходит в первом положении (CAT), а у атлантического лосося в третьем (TAC). Дистальная часть элемента GRE более вариабельная не только у лососевых, она отличается и при сравнении внутри разных отрядов млекопитающих (Beato et al., 1989).

Через три нуклеотида от элемента GRE расположен элемент ответа на цАМФ (CRE). Элемент ответа на цАМФ присутствует в промоторах генов гормона роста *gh1* и *gh2* всех исследованных видов лососевых. У нерки, чавычи и атлантического лосося элемент CRE, так же, как и у четырех видов гольцов, представлен в промоторе ассиметричной последовательностью (TGACG), с которой взаимодействует

транскрипционный фактор CREB. Ассиметричной последовательности может быть недостаточно для эффективного связывания с CREB и в таком случае, вероятно, необходимо участие транскрипционного фактора Pit-1. Известно, что в результате мутаций или полного отсутствия в промоторе элементов ответа на цАМФ транскрипционный фактор CREB может напрямую взаимодействовать с сайтами транскрипционного фактора Pit-1 или выступать в качестве его кофактора (Argenton et al., 1996; Bernardini et al., 1999; Argenton et al., 2002; Almuly et al., 2005). Помимо основного мотива CRE в промоторе могут встречаться и дополнительные элементы ответа на цАМФ. Два дополнительных сайта связывания с CREB описаны у золотистого спара (*S. aurata*) и желтохвоста (*S. quinquerediata*) (Almuly et al., 2005), расположены они в промоторе гораздо дальше относительно проксимального CRE и не ассоциированы с сайтами связывания Pit-1. На исследуемых в данной работе участках промотора длиной 300 п.н. генов *gh1* и *gh2* лососевых рыб представлена только одна последовательность элемента ответа на цАМФ. Расположена она между сайтами связывания с транскрипционным фактором Pit-1 F2 и F3. Такое положение элемента ответа на цАМФ указывает на то, что транскрипционные факторы Pit-1 и CREB взаимодействуют с данным участком. Только остается пока неясным – связываются ли они одновременно или транскрипционный фактор CREB является участником альтернативного пути активации транскрипции генов гормона роста *gh1* и *gh2* у лососевых рыб в отсутствие транскрипционного фактора Pit-1.

В результате сравнительного анализа было показано, что большинство сайтов связывания с факторами транскрипции, которые принимают участие в формировании транскрипционного комплекса, в генах гормона роста *gh1* и *gh2* у лососевых рыб сосредоточено на участке промотора длиной 300 п.н. Сайты ответственные за взаимодействие с транскрипционным фактором CREB, с рецепторами глюкокортикоидов (GRE) и с рецепторами ретиноевой кислоты (RARE/RXRE) состоят из одинакового количества нуклеотидов в обоих генах-паралогах. У всех исследованных видов семейства Salmonidae: CRE представлен последовательностью из пяти нуклеотидов; сайты GRE и RARE/RXRE сформированы двумя мотивами из пяти (GRE) и шести (RARE/RXRE) нуклеотидов, разделенных тремя и восьмью нуклеотидами соответственно. В последовательностях не найдено никаких инсерций или делеций, которые бы нарушали

целостность функциональных мотивов. В гене *gh1* последовательности CRE, GRE и RARE/RXRE полностью консервативны и не имеют различий в функционально значимых консенсусных последовательностях, а в гене *gh2* встречаются только несколько точечных мутации.

Положение элементов ответа на цАМФ, глюкокортикоиды и ретинойды совпадает как в гене *gh1*, так и в гене *gh2*. У четырех видов гольцов, атлантического лосося, нерки и чавычи в гене *gh1* CRE, GRE и RARE/RXRE расположены на расстоянии –148/–152 п.н.; –130/–144 п.н. и –105/–124 п.н. от точки начала транскрипции; в гене *gh2* положение данных сайтов смещается на один нуклеотид: CRE – –147/–151 п.н., GRE – –129/–143 п.н., RARE/RXRE –106/–125 п.н.. Сайты находятся настолько близко к друг другу, что элемент ответа на глюкокортикоиды и сайт связывания с рецептором ретиноевой кислоты частично перекрывают сайт F2 с обеих сторон, а элемент ответа на цАМФ отделен всего тремя нуклеотидами от GRE и двумя нуклеотидами от сайта F3 транскрипционного фактора Pit-1.

В экспериментах по влиянию отдельных областей промотора на эффективность транскрипции промоторы гена гормона роста радужной форели с разным количеством сайтов связывания с транскрипционным фактором Pit-1 соединяли с геном люциферазы (Argenton et al., 1996; Argenton et al., 2002). В результате авторами было показано (Argenton et al., 1996), что инициация транскрипции может происходить в отсутствие участка промотора, содержащего сайты F2, F3 и элементов ответа на другие лиганды. Транскрипционный фактор Pit-1 взаимодействует с сайтом F1, но достигнуть оптимального уровня транскрипции в присутствии только одного сайта связывания с Pit-1 не удастся. Напротив, транскрипционная активность репортерного гена, соединенного с промотором, который включает в себя участок последовательности с сайтами F2 и F3 увеличивалась в 3,5 раза. Однако при наличии мутаций в сайте F1 два дистальных сайта связывания с Pit-1 (F2 и F3) не могут обеспечить оптимальную Pit-1-зависимую транскрипцию. А низкий уровень такой активности может указывать на то, что для успешного взаимодействия транскрипционного фактора Pit-1 с сайтами F2 и F3 необходимо присутствие сайтов связывания с дополнительными факторами транскрипции. Возможно, что именно в результате взаимодействия транскрипционного

фактора CREB, рецепторов глюкокортикоидов и ретиноевой кислоты со своими сайтами связывания, заключенными между последовательностями F2 и F3, формируется сложный многокомпонентный белковый комплекс, под действием которого промотор гена гормона роста изменяет свою конформацию таким образом, что транскрипционный фактор Pit-1 может связаться с сайтами F2 и F3 в более выгодном положении, что и приводит к увеличению транскрипционной активности.

Как правило, рецептор гормона, связываясь со своим сайтом, не может активировать транскрипцию специфичного гена, но может облегчить взаимодействие основного фактора транскрипции, изменяя структуру хроматина прилежащих участков таким образом, что сайты, расположенные на расстоянии нескольких сотен нуклеотидов, становятся доступными для связывания с транскрипционным фактором. Возможно, что в результате связывания GRE и RARE/RXRE со своими рецепторами последовательность гена гормона роста изменяется таким образом, что элемент ответа на цАМФ и участки связывания с транскрипционным фактором Pit-1, которые были найдены в интронах С и D в обоих генах гормона роста лососевых, оказываются в непосредственной близости к промоторной области и взаимодействуют с транскрипционными факторами совместно с сайтами промотора. Например, расстояние между элементами ответа на цАМФ в промоторе и интроне С составляет 2616 и 2617 п.н. для генов *gh1* и *gh2*, соответственно. Это кратно примерно 10 нуклеосомным повторам ДНК. Если предположить, что благодаря GRE и RARE/RXRE эти сайты оказываются сближены, то в зависимости от того, как работают эти участки (кооперативно или же, наоборот, по принципу антагонистов при связывании с лигандами), могут определяться и различия в функциональности генов-паралогов (Каменская и др., 2015).

Таким образом, в результате сравнения нуклеотидных последовательностей промоторных областей генов гормона роста *gh1* и *gh2* гольцов рода *Salvelinus* с промоторами атлантического лосося, нерки и чавычи было показано, что оба гена-паралога содержат одинаковый набор сайтов связывания с факторами транскрипции. У всех исследованных видов лососевых в обоих генах на участке промотора длиной 300 п.н. были найдены ТАТА-боксы, четыре сайта связывания с гипофиз-специфичным транскрипционным фактором Pit-1 (F1–F4), сайты связывания с транскрипционным

фактором CREB, с рецепторами глюкокортикоидов и ретиноевой кислотой. Расположение данных сайтов на небольшом расстоянии друг от друга дают основания предполагать, что они формируют ядро промотора гена гормона роста и играют ключевую роль в инициации транскрипции. Не стоит исключать, что при исследовании более протяженных участков, прилегающих к гену, могут быть найдены дополнительные сайты связывания с факторами транскрипции, как это было показано у других видов позвоночных. Но согласно экспериментам по оценке экспрессии репортерных генов, находящихся под промотором, содержащим только эти сайты, данного участка достаточно для эффективной транскрипции. А область длиной 73 нуклеотида, содержащая сайты, которые отвечают за взаимодействие сразу с несколькими факторами транскрипции, может выступать в качестве активного центра промотора.

Взаимодействие с элементами ответа на цАМФ, глюкокортикоиды и с сайтом связывания с рецептором ретиноевой кислоты не только может облегчать связывание фактора Pit-1 со своими сайтами для формирования сложного транскрипционного комплекса, но и в отсутствие Pit-1 привлекать альтернативные факторы транскрипции для сохранения транскрипционной активности, тем самым поддерживая адекватный уровень экспрессии в ответ на внешние или внутренние факторы. Несмотря на то, что последовательности сайтов связывания с транскрипционными факторами достаточно консервативны в обоих генах-паралогах, анализ нуклеотидного разнообразия и дивергенции могут показать, насколько сильно изменчивы последовательности промотора между генами *gh1* и *gh2*, и какого рода эволюционные силы действуют на данные последовательности.

4.2. Сравнительный анализ изменчивости генов

4.2.1. Изменчивость транскрибируемой части гена

В случае большинства дублированных генов одна копия, как правило, сохраняет первоначальную функцию, а другая последовательность, накапливая изменения с более высокой частотой, либо перестает функционировать, превращаясь в псевдоген, либо приобретает новую функцию. Две копии гена гормона роста появились еще у предковой формы лососевых рыб в результате процесса тетраплоидизации и сохранились на

протяжении всего времени дивергенции видов в этой группе (McKay et al., 2004). Для описанных ранее последовательностей гена гормона роста известно, что обе копии не несут дополнительных стоп-кодонов, имеют открытые рамки считывания и кодируют одинаковое количество аминокислотных остатков (Agellon et al., 1988; Male et al., 1992; Du et al., 1993; Devlin, 1993). Все эти особенности указывают на то, что оба гена могли сохранить способность функционировать, несмотря на длительное существование в геноме.

В результате анализа методом скользящего окна двух копий гена гормона роста представителей трех родов семейства Salmonidae (*Oncorhynchus*, *Salmo*, *Salvelinus*) показано, что между разными видами внутри одного семейства последовательности генов *gh1* и *gh2* слабо изменчивы. Уровень нуклеотидного разнообразия транскрибируемой части гена *gh1* при сравнении четырех видов гольцов оказался низким и составил 0,00446. Уровень изменчивости в гене *gh2* при сравнении последовательностей четырех видов гольцов оказался в несколько раз выше $\pi=0,01423$. Аналогичная закономерность наблюдается и при сравнении значений дивергенции: для гена *gh1* дивергенция составляет 0,00364, а для гена *gh2* 0,01064. Связано это с тем, что последовательности транскрибируемой части гена гольцов оказались более варибельными между видами. В свою очередь эти данные подтверждают, что одна из копий гена находится под меньшим давлением отбора и накапливает мутации с высокой частотой. При сравнении транскрибируемых последовательностей генов гормона роста у гольцов можно говорить о том, что ген *gh1* более консервативен по сравнению с геном *gh2*. (рис. 3.14, 3.15) В меньшей степени эта закономерность прослеживается при сравнении соответствующих генов с другими видами лососевых. После добавления в анализ последовательностей генов гормона роста *O. nerka*, *O. tshawytscha* и *S. Salar* больших различий в значениях нуклеотидного разнообразия не наблюдается: для гена *gh1* $\pi=0,03737$, для гена *gh2* $\pi=0,03683$. Значения дивергенции тоже имеют небольшие различия 0,02741 для гена *gh1* и 0,02980 для гена *gh2*. Как видно, при этом уровень нуклеотидного разнообразия и дивергенции увеличился в несколько раз, но общая закономерность сохранилась: минимальные значения на графике соответствуют экзонам, а максимальные значения оказались ассоциированы с интронами (рис. 3.18, 3.19)

4.2.2. Изменчивость экзонов

Рассчитав уровень нуклеотидного разнообразия для каждого экзона, удалось показать, что не все экзоны одинаково консервативны. Несмотря на то, что размер экзонов как между генами *gh1* и *gh2*, так и между исследованными видами лососевых совпадает (табл. 3.3), строгой корреляция между длиной экзонов и уровнем нуклеотидного разнообразия не прослеживается (Каменская и др., 2020). Наличие изменчивости или ее отсутствие, как, например, в шестом экзоне гена *gh1*, может быть обусловлено функциональными особенностями экзонов, а не их длиной. Шестой экзон гена *gh1* оказался самым консервативным ($\pi=0$) и самым коротким (63 п.н.). Отсутствие изменчивости в этом экзоне может быть связано с тем, что все кодируемые им аминокислоты, входят в состав одного из пяти GD доменов (GD5). В отличие от доменов GD1–GD4, которые отвечают за специфическое взаимодействие с рецепторами гормона роста, домен GD5, согласно рентгеноструктурному анализу, расположен в центре молекулы и участвует в формировании определенной конфигурации гормона и его стабилизации (Watahiki et al., 1989). В этот же домен у всех костистых рыб входит консервативный участок С-терминальной области, состоящий из тринадцати аминокислот, пять из которых приходятся на шестой экзон. Два из четырех аминокислотных остатка цистеина, которые участвуют в формировании вторичной структуры белка, за счет образования дисульфидных связей, также кодируются шестым экзоном. В шестом экзоне гена *gh2*, по сравнению с геном *gh1*, уровень нуклеотидного разнообразия выше ($\pi=0,00756$), но это связано с наличием у нерки и чавычи одной замены, которая не влияет на аминокислотный состав. По сравнению с остальными экзонами гена *gh2* нуклеотидное разнообразие шестого экзона самое низкое.

Низкий уровень нуклеотидного разнообразия в первом экзоне можно было бы объяснить его небольшой длиной (74 п.н.), но это справедливо только для гена *gh1* $\pi=0,00386$. В гене *gh2* уровень нуклеотидного разнообразия составил $\pi=0,03501$, что в несколько раз больше значений полученных для третьего ($\pi=0,01099$) и четвертого ($\pi=0,02869$) экзонов, которые длиннее первого экзона в 1.5 и 2 раза, соответственно. У всех исследованных видов первый экзон кодирует транскрибируемую, но не

транслируемую лидерную последовательность. Лидерная последовательность может содержать важные функциональные элементы, необходимые для эффективной инициации трансляции: участки внутренней посадки рибосомы, внутренние открытые рамки считывания, железозависимые и другие элементы (Barrett et al., 2013). Стоит отметить, что в гене *gh1* лидерная последовательность консервативна у всех исследованных видов и имеет только одну нуклеотидную замену у атлантического лосося. В гене *gh2* лидерная последовательность отличается не только количеством замен между видами, но и размерами. Так у гольца Таранца из-за делеции в шесть нуклеотидов длина лидерной последовательности составляет всего 59 п.н., тогда как у остальных видов 64 п.н..

Для остальных экзонов также не наблюдается строгой корреляции между увеличением уровня нуклеотидного разнообразия и их длиной. Несмотря на то, что второй, третий, четвертый и частично пятый экзон кодируют аминокислоты входящие в состав консервативных доменов GD, в отличие от шестого экзона, строгой взаимосвязи между уровнем нуклеотидного разнообразия и функциональными особенностями для этих экзонов показать не удалось. Например, четвертый экзон кодирует аминокислоты, которые входят в состав сразу двух доменов: GD3 и GD4. А значение нуклеотидного разнообразия в четвертом экзоне гена *gh1* по сравнению с другими экзонами гена самое высокое. Четвертый экзон – единственный, где значения нуклеотидного разнообразия между генами *gh1* и *gh2* практически совпадают (табл. 3.3). Пятый экзон кодирует 49 аминокислот, из которых 34 не входят ни в один из консервативных доменов GD и можно было бы предполагать, что это будет в равной степени отражаться на изменчивости экзона пять в обоих генах. Но, как видно из таблицы 3.3 высокий уровень нуклеотидного разнообразия наблюдается только в пятом экзоне гена *gh2*, в пятом экзоне гена *gh1* уровень нуклеотидного разнообразия один из самых низких. Возможно, в зависимости от функции данная последовательность аминокислот занимает различные положения в генах-паралогах на этапах формирования вторичных и третичных белковых структур. Второй экзон кодирует 18 аминокислот сигнального пептида и первые 28 аминокислот зрелого белка. Несмотря на то, что в нуклеотидной последовательности сигнального пептида встречается всего четыре замены в каждом гене и почти все аминокислоты зрелого белка входят в состав домена GD1, уровень нуклеотидного разнообразия во

втором экзоне гена *gh2* в 1,5 раза выше, чем во втором экзоне гена *gh1*. Связаны ли различия в значениях нуклеотидного разнообразия между генами *gh1* и *gh2* в экзоне три с функциональными или конфигурационными особенностями белковой молекулы пока не известно.

Анализ нуклеотидного разнообразия показал, насколько сильно в дублированных копиях генов могут быть изменчивы отдельные участки относительно друг друга, даже такие консервативные как экзоны. Несмотря на то, что уровень изменчивости экзонов гена *gh2* на порядок выше, чем экзонов гена *gh1*, оба гена-паралога лососевых сохраняют свой функциональный потенциал. Это обусловлено тем, что в нуклеотидной последовательности экзонов представлены в основном синонимичные замены, не приводящие к заменам аминокислот (Панькова и др., 2017). Однако, различия между формами гормона роста могут проявляться на этапе созревания и формирования белковых молекул, в зависимости от выполняемых ими функций.

4.2.3. Изменчивость интронов

Уровень изменчивости в разных участках гена зависит от скорости накопления мутаций. Особенно это касается некодирующих последовательностей, таких как интроны и регуляторные области. Как правило, такие участки находятся под ослабленным давлением отбора. На основании оценки нуклеотидного разнообразия каждого интрона в генах *gh1* и *gh2* было показано, что уровень изменчивости интронов, так же, как и уровень изменчивости экзонов, не зависит от их длины (табл. 3.4). В гене *gh1* наименьший уровень нуклеотидного разнообразия ($\pi=0,02963$) действительно приходится на короткий интрон В. Однако в гене *gh2* уровень нуклеотидного разнообразия в этом интроне самый высокий по сравнению с остальными интронами в обоих генах. Самый низкий уровень нуклеотидного разнообразия в гене *gh2* получен для второго по длине интрона С $\pi=0,02835$. В гене *gh1* интрон С длиннее и значение нуклеотидного разнообразия по сравнению с интроном С гена *gh2* больше в 1.5 раза $\pi=0,04244$. Самый высокий уровень нуклеотидного разнообразия в гене *gh1* приходится на интрон Е $\pi=0,05007$, который короче интрона С на 100 нуклеотидов. Несмотря на то, что в гене *gh2* последовательность интрона Е немного длиннее интрона В, уровень нуклеотидного разнообразия в 1,2 раз

выше, чем в самом длинном интроне D. В гене *gh2* интрон D имеет наибольшую длину последовательности, но по сравнению с остальными интронами достаточно низкое нуклеотидное разнообразие $\pi=0,03607$. Интрон D в гене *gh1* короче интрона D гена *gh2* на 100 п.н., но значение нуклеотидного разнообразия немного выше $\pi=0,04190$. В интроне A достаточно высокий уровень изменчивости по сравнению с остальными интронами: в гене *gh1* нуклеотидное разнообразие интрона A составило $\pi=0,04699$, а в гене *gh2* $\pi=0,05153$.

Несмотря на то, что интроны C и D имеют большие размеры, по уровню нуклеотидного разнообразия они оказались менее изменчивыми, чем более короткие интроны. Невысокое значение уровня нуклеотидного разнообразия может быть связано с наличием потенциальных регуляторных элементов, которые могут принимать участие в изменении нуклеотидной последовательности таким образом, чтобы сформировать полноценный транскрипционный комплекс. Так в интроне C у обоих генов обнаружен элемент ответа на эстрогены (ERE), который взаимодействует с рецептором эстрогена (ER) (Панькова и др., 2013). Элемент ответа на эстрогены представлен в обоих генах гормона роста лососевых и состоит из 13 пар нуклеотидов. В гене *gh2* последовательность ERE полностью консервативна у всех исследованных видов, а в гене *gh1* содержит всего по одной замене у *S. malma*, *S. curilus* и *S. salar*. Помимо элемента ответа на эстрогены в интроне C имеется еще один потенциальный регуляторный участок, содержащий две АТ-богатые области, в которых можно выделить сайты связывания с транскрипционным фактором Pit-1. При сравнении данных сайтов внутри каждого гена между разными видами лососевых эти сайты оказались консервативны (Каменская и др., 2017).

В интроне D имеется последовательность CRE-элемента. Последовательности CRE внутри каждого гена-паралога у всех исследованных видов лососевых идентичны, а между генами отличаются только одним нуклеотидом. В отличие от промоторной последовательности в интроне D CRE-элемент представлен палиндромом, что обеспечивает эффективное взаимодействие с фактором CREB без участия дополнительных лигандов (Montminy et al., 1990; Zanger et al., 1999; Argenton et al., 2002). В интроне D имеется еще один консервативный участок, состоящий из прямых повторов GATT (McKay et al., 2004). В гене *gh1* у всех исследованных видов лососевых встречается два таких повтора, в гене *gh2* у нерки и чавычи два повтора, а остальные виды имеют по

три повтора. Известно, что первые интроны большинства генов млекопитающих могут содержать регуляторные элементы и наравне с промотором принимать участие в регуляции транскрипции, при этом уровень изменчивости относительно остальных интронов у них остается низкий (Almuly et al., 2000; Kalari et al., 2006). Однако ни у кого из представителей лососевых в генах гормона роста не были описаны регуляторные элементы в интроне A. По уровню нуклеотидного разнообразия интрон A является одним из самых изменчивых в обоих генах-паралогах.

Интроны уже давно не относят к бесполезной или, по-другому, к «мусорной» ДНК. В последнее время появляется все больше доказательств, указывающих на то, что интроны включают в себя нетранслируемые РНК (микроРНК, малые ядершковые РНК), элементы транскрипции и элементы контроля сплайсинга (Gazave et al., 2007; Zhu et al., 2009). В зависимости от выполняемых функций и наличия в них регуляторных элементов интроны также могут находиться под разным давлением отбора. Отбор оказывает влияние на уровень изменчивости, нуклеотидный состав и даже на длину последовательности интрона. Для некоторых млекопитающих и насекомых установлен ряд закономерностей, которые оказались справедливы не для всех видов. Например, при сравнении интронных участков многих ортологичных генов человека и шимпанзе Газаве с соавторами (Gazave et al., 2007) показали строгую положительную корреляцию между длиной интронов и их дивергенцией, а также строгую отрицательную корреляцию между длиной интронов и содержанием GC-пар. Кроме того, авторы (Gazave et al., 2007) утверждают, что скорость дивергенции зависит от порядкового номера интрона в гене. И на примере первого интрона показывают, что, в отличие от остальных интронов, они богаче GC, длиннее и быстрее дивергируют. Хеддрилл с соавторами (Haddrill et al., 2005), в результате сравнения 225 интронных участков в геномах *Drosophila melanogaster* и *D. simulans*, напротив, выявили строгую отрицательную корреляцию между длиной интронов и их дивергенцией; положительную корреляцию между длиной интронов и содержанием GC-пар и также показали отрицательную корреляцию между дивергенцией интронов и содержанием GC-пар. Как считают авторы (Haddrill et al., 2005), такая взаимосвязь может просто отражать локальные вариации в скорости мутаций, поскольку не учитывает корреляцию между длиной интронов и их дивергенцией (Haddrill et al., 2005). Сравнивая

последовательности интронов из полных геномов пяти видов рыб, авторам (Chaurasia et al., 2014) не удалось установить какие-либо строгие закономерности между длиной интронов, дивергенцией и содержанием GC.

В нашем случае для интронов генов гормона роста также не установлено общей закономерности между длиной интрона и содержанием GC. Если рассматривать каждый интрон в отдельности, то можно обнаружить как положительную, так и отрицательную корреляцию между этими параметрами (табл. 3.5). Например, в интронах В, Е и С гена *gh1* можно наблюдать положительную корреляцию длины интрона с содержанием GC поскольку с увеличением длины увеличивается и содержание GC. Отрицательная корреляция между длиной и содержанием GC наблюдается в интроне D, который имеет самый большой размер, но низкое содержание GC. Содержание GC в интроне А немного больше, чем в интроне Е, который длиннее интрона А на 200 нуклеотидов. В гене *gh2* содержание GC положительно коррелирует с длиной интронов Е, А и С, а отрицательно с интроном D и интроном В. Для интрона В гена *gh2* прослеживается еще и положительная корреляция между содержанием GC и уровнем изменчивости. Но для остальных интронов строгой связи между содержанием GC и уровнем изменчивости не наблюдается. Это может быть связано со сходным содержанием GC-пар во всех интронах (Каменская и др., 2020). Закономерности, показанные для первых интронов генов человека и шимпанзе, в гене гормона роста лососевых тоже не наблюдаются. Первый интрон в генах *gh1* и *gh2* имеет среднюю длину последовательности и не самый богатый по содержанию GC относительно других интронов. На примере интронов одного гена, изолированного от остального генома, сложно установить какие именно закономерности, описанные для млекопитающих или насекомых, будут наблюдаться для всех интронов лососевых. Возможно, для лососевых будут прослеживаться совсем другие закономерности, так как в отличие от теплокровных животных, холоднокровные позвоночные характеризуется низким уровнем содержания GC и неоднородным нуклеотидным составом (Bernardi, 2007).

4.2.4. Изменчивость регуляторной области

Уровень нуклеотидного разнообразия, полученный для промоторной области у паралогичных генов гормона роста, соответствует некоторым закономерностям, установленным для транскрибируемой части гена. При сравнении регуляторных участков среди четырех видов гольцов рода *Salvelinus* промотор гена *gh1* ($\pi=0,00327$) оказался более консервативен, чем промотор гена *gh2* ($\pi=0,00545$). Согласно рисункам 3.4 и 3.5 в промоторах обоих генов наблюдается всего два пика с максимальными значениями, но их положение между генами не совпадает. В промоторе гена *gh1* один пик находится между ТАТА-боксом и сайтом F1, а другой между сайтом F1 и сайтом F2. В гене *gh2* пики расположены дальше относительно точки начала транскрипции в непосредственной близости к сайту F3.

Если в анализ добавить последовательности нерки, чавычи и атлантического лосося, то промоторная область у гена *gh1* также имеет более низкий уровень изменчивости, чем у гена *gh2* $\pi=0,04922$ и $\pi=0,05721$, соответственно. Несмотря на то, что промотор гена *gh2* оказался более изменчив, чем промоторная область гена *gh1*, количество сайтов связывания с транскрипционными факторами, необходимое для успешной инициации транскрипции, сохраняется в обоих генах у всех исследованных видов. Уровень изменчивости в промоторных областях оказался выше, чем уровень изменчивости транскрибируемой части обоих генов-паралогов.

Согласно полученным данным, можно отметить, что в отличие от промотора, по уровню нуклеотидного разнообразия транскрибируемая часть генов гормона роста *gh1* и *gh2*, достаточно консервативна и не отличается между генами. Но если сгруппировать экзоны и интроны и оценить изменчивость кодирующих и некодирующих участков, то более заметные отличия в генах *gh1* и *gh2* окажутся не только в промоторной области. Уровень нуклеотидного разнообразия в экзонах гена *gh2* ($\pi=0,02832$) в 1,6 раз больше чем в экзонах гена *gh1* ($\pi=0,02070$). Повышенный уровень нуклеотидного разнообразия в экзонах гена *gh2* относительно гена *gh1* указывает, что давление отбора на кодирующую часть гена *gh2* снижено. Возможно, на уровне нуклеотидной последовательности различия не критичны, поскольку оба гена у всех лососевых имеют открытую рамку

считывания и могут кодировать полноценный белок (Male et al., 1992; Du et al., 1993; Devlin, 1993; Панькова и др., 2017). Однако если оценивать уровень нуклеотидного разнообразия по каждому экзону между генами, то значения могут сильно отличаться, как например, между первыми и пятыми экзонами генов *gh1* и *gh2*. Возможно, существенные различия между генами-паралогами проявляются пока не на уровне нуклеотидных последовательностей, а на уровне взаимодействий с разного рода лигандами и выполняемых белком функций.

Последовательности интронов, как и ожидалось, оказались более изменчивы, чем экзоны. При этом пики максимальной изменчивости приходятся во многих случаях на среднюю часть последовательности интронов. Такое распределение может быть обусловлено тем, что все интроны фланкированы одинаковыми сайтами сплайсинга (на 5'-концах интронов – GT, на 3'-концах – AG), находящихся под действием отбора.

В отличии от экзонов, интроны в гене *gh1* оказались более изменчивы, чем интроны гена *gh2* $\pi=0,04218$ и $\pi=0,03958$ соответственно. Можно было бы предполагать, что в паралогичных генах гормона роста под ослабленным действием отбора в каждом гене находятся разные участки. Однако если исключить из анализа интроны атлантического лосося, потому что интрон E в гене *gh1* короче относительно остальных видов на 400 п.н., и сравнивать только последовательности *Salvelinus* и *Oncorhynchus*, то и по уровню изменчивости интронов ген *gh1* ($\pi=0,03422$) оказывается более консервативен, чем ген *gh2* ($\pi=0,03867$). Что тоже подтверждает ослабленное давление отбора на ген *gh2*. Несмотря на то, что самые длинные интроны C и D содержат потенциальные регуляторные элементы, на общем уровне изменчивости интронов это не отразилось.

Изменчивость как экзонов, так и интронов может зависеть от изменчивости соседних участков (Zhu et al., 2009). Например, в гене гормона роста *gh2* за первым экзоном с высоким уровнем изменчивости следует интрон A, у которого тоже достаточно высокий уровень изменчивости, далее второй экзон, как и следующий за ним интрон B, имеют самые высокие значения нуклеотидного разнообразия среди экзонов и интронов гена *gh2* соответственно. Однако следующий за интроном B третий экзон, как и четвертый имеют относительно низкие значения нуклеотидного разнообразия, как и интроны C и D,

которые их фланкируют. В гене *gh1* уровень изменчивости экзонов и следующих за ним интронов изменяется более равномерно.

Согласно данным, полученным Паньковой с соавторами (Панькова и др., 2017), отбор на гены гормона роста лососевых действует с разной силой, в результате в одном из паралогах мутаций возникают с более высокой частотой, что приводит к увеличению его уровня изменчивости. В свою очередь, данные, полученные по оценке нуклеотидного разнообразия промоторного участка и транскрибируемой части паралогичных генов *gh1* и *gh2* методом скользящего окна показывают, что оба гена у лососевых достаточно консервативны, имеют низкий уровень изменчивости даже при сравнении видов внутри одного семейства. Консервативное строение промоторной области наблюдается на протяжении всего времени дивергенции лососевых и может служить еще одним свидетельством функциональности обоих генов-паралогов и их общего происхождения. На функциональность обоих генов-паралогов указывает и анализ аминокислотной последовательности генов гормона роста лососевых. Предсказанная аминокислотная последовательность высококонсервативна, включая положения цистеиновых остатков, а открытая рамка считывания имеет общую длину 630 п.н. (Панькова и др., 2017).

Схема дивергенции транскрибируемых участков генов, включая прилегающий цис-регуляторный участок, согласуется с ранее опубликованными схемами, полученными на основании интронных последовательностей (Панькова и др., 2013) и подтверждает давнюю и независимую эволюцию паралогичных генов гормона роста у лососевых рыб. На рисунках 3.20 и 3.21 показано, что накопление нуклеотидных замен (мутаций) в каждом из генов-паралогов определяется временем независимой дивергенции. Объединенные последовательности двух генов-паралогов более точно отражают последовательность дивергенции видов в таксоне, подтверждая представление о том, что чем длиннее используемые в анализе последовательности ДНК, тем лучше они отражают филогению видов внутри таксонов (Lynch, Katju, 2004).

Лососевые прошли через четыре этапа дупликации, и многие гены у них оказались множественными, в том числе и представленные двумя копиями гены гормона роста. За все время дивергенции видов в этой группе (25–100 млн. лет) ни один из этих двух генов не стал псевдогеном и не накопил такое количество замен, которое сопровождалось бы

кодированием белка с новыми функциями (неофункционализация), отличными от тех, которые выполняет гормона роста. Действие очищающего отбора на оба гена-паралога дает основания считать, что такая избыточность (присутствие продуктов обоих генов гормона роста) имеет важное значение в жизненном цикле лососевых рыб. Возможно, что гены включаются на разных стадиях жизненного цикла. Либо, поскольку отбор действует на паралоги с разной силой, именно сейчас и происходит процесс субфункционализации, когда один из генов адаптируется для выполнения части функций гена гормона роста или готовится взять на себя какую-то новую функцию. Или, когда параметры окружающей среды начнут меняться, второй ген сможет быстрее подстроиться к новым условиям и такая важная функция, как регуляция роста и развития сохранится.

ЗАКЛЮЧЕНИЕ

Полногеномные дубликации способствовали появлению нового генетического материала на ранних этапах эволюции позвоночных животных (Lynch, Conery, 2000; Zhang, 2003). В результате вторичной диплоидизации многие гены оказались утрачены, а копии генов, которые сохранились, объединились в генные семейства. Как правило, дублицированные гены, находясь под ослабленным давлением отбора, накапливают изменения с более высокой частотой и имеют возможность модифицировать свою структуру и функции. У лососевых рыб, как естественных и относительно недавних полиплоидов, многие гены оказались множественными, в том числе и ген гормона роста, представленный двумя копиями (Male et al., 1992; Du et al., 1993; Devlin, 1993). Гены гормона роста лососевых не стали исключением и, как большая часть дублицированных генов, испытывают давление отбора разной интенсивности, что может приводить к появлению большего числа замен в одном из паралогов и увеличению его изменчивости. Полученные в настоящей работе данные по оценке нуклеотидного разнообразия и дивергенции промоторного участка и транскрибируемой части паралогичных генов *gh1* и *gh2* методом скользящего окна показывают, что оба гена у лососевых достаточно консервативны, имеют низкий уровень изменчивости даже при сравнении видов внутри одного семейства. Консервативное строение промоторной области наблюдается на протяжении всего времени дивергенции лососевых и вместе с открытой рамкой считывания, высокой консервативностью аминокислотной последовательности и идентичным положением цистеиновых остатков, может указывать на то, что оба гена сохранили способность функционировать, несмотря на длительное совместное существование в геноме.

Исследование структуры промоторных областей паралогичных генов гормона роста *gh1* и *gh2* у четырех видов гольцов рода *Salvelinus* (*S. malma*, *S. curilus*, *S. levanidovi* и *S. taranetzi*) и сравнение их с промоторами генов гормона роста нерки (*O. nerka*), чавычи (*O. tshawytscha*) и атлантического лосося (*S. salar*) показали, что на участке длиной 300 п.н. перед точкой инициации транскрипции, сосредоточены ключевые сайты связывания с транскрипционными факторами и другими лигандами, которые отвечают за формирование полноценного транскрипционного комплекса и эффективную

тканеспецифичную экспрессию гена гормона роста в соматотрофах. В промоторах обоих генов всех исследованных видов лососевых встречаются последовательность ТАТА-бокса, сайты связывания с гипофиз-специфичным транскрипционным фактором Pit-1 (F1–F4), сайты связывания с транскрипционным фактором CREB (CRE), с рецепторами глюкокортикоидов (GRE) и ретиноевой кислотой (RARE/RXRE). Расположение данных сайтов на небольшом расстоянии друг от друга дает основания предполагать, что они формируют ядро промотора гена гормона роста и играют важную роль в инициации транскрипции. Отсутствие существенных различий в структурных элементах промотора указывает на потенциальную возможность сохранять функцию синтеза активного белка с обоих генов паралогов. Это может иметь важное значение для лососевых рыб, поскольку копии гена гормона роста могут включаться на разных стадиях жизненного цикла лососевых или при изменении условий окружающей среды, например, при колебаниях осмотического давления в процессе миграций.

ВЫВОДЫ

1. Промоторные участки в паралогичных генах *gh1* и *gh2* у четырех видов гольцов рода *Salvelinus*: северная мальма – *Salvelinus malma*, южная азиатская мальма – *S. curilus*, голец Леванидова – *S. levanidovi*, голец Таранца – *S. taranetzi*, представлены высококонсервативными последовательностями и включают сайты связывания необходимые для взаимодействия с тканеспецифичными транскрипционными факторами и другими лигандами.

2. Сравнительный анализ 5'-фланкирующих участков паралогичных генов гормона роста *gh1* и *gh2* четырех видов гольцов рода *Salvelinus* с последовательностями нерки (*O. nerka*), чавычи (*O. tshawytscha*) и атлантического лосося (*S. salar*) показал, что у всех исследованных видов в промоторе содержится одинаковый набор структурных элементов, необходимых для сборки транскрипционного комплекса.

3. У всех семи исследованных видов лососевых в обоих паралогичных генах на участке промотора длиной 300 п.н. встречаются ТАТА-боксы, сайты связывания с гипофиз-специфичным транскрипционным фактором Pit-1 (F1–F4), сайты связывания с транскрипционным фактором CREB (CRE), с рецепторами глюкокортикоидов (GRE) и с рецепторами ретиноевой кислоты (RARE/RXRE). Расположение данных сайтов на небольшом расстоянии друг от друга дает основания предполагать, что они формируют ядро промотора гена гормона роста и играют ключевую роль в инициации транскрипции.

4. В результате сравнительного анализа было показано, что последовательность и положение ТАТА-боксов, GRE и CRE-мотивов, RARE/RXRE и F1 сайтов, консервативны в двух генах-паралогах у всех исследованных видов лососевых. Положение сайтов F2 – F4 между генами различно.

5. Паралогичные гены гормона роста *gh1* и *gh2* четырех видов гольцов рода *Salvelinus*, нерки (*O. nerka*), чавычи (*O. tshawytscha*) и атлантического лосося (*S. salar*) содержат в своем промоторе четыре сайта связывания с гипофиз-специфичным транскрипционным фактором Pit-1. Такое количество сайтов связывания с Pit-1 на коротком участке промотора, прилегающем непосредственно к транскрибируемой части гена, встречается только у лососевых рыб.

6. Некоторые регуляторные элементы (CRE и Pit-1) встречаются не только в промоторных участках, но и в структурной части гена, в частности – в интронах. Возможно, что они кооперативно взаимодействуют друг с другом в составе хроматина.

7. Низкий уровень дивергенции указывает на общее происхождение паралогичных генов гормона роста и сохранение функционального потенциала обоих генов.

СПИСОК ЛИТЕРАТУРЫ

1. Жимулев И.Ф. Хромомерная организация политенных хромосом – Новосибирск: Наука. 1994. – 564 с.
2. Жимулев И.Ф. Современные представления о структуре гена // Соросовский образовательный журнал. 2000. Т. 6, № 7. С. 17–24.
3. Задесенец К.С, Рубцов Н.Б Дупликация генома в эволюции животных // Генетика. 2018. Т. 54, № 10. С. 1107–1119.
4. Заславская Н.И., Скурихина Л.А., Панькова В.В., Рязанова И.Н. Методы генетических исследований морских организмов: учеб. пособие. – Владивосток: изд. ДВГУ. 2009. – 160 с.
5. Каменская Д.Н., Панькова М.В., Атопкин Д.М., Брыков Вл.А. Гены гормона роста у рыб: доказательства функциональности паралогичных генов гольца *Salvelinus levanidovi* // Молекулярная биология. 2015. Т. 49, № 5. С. 770–776.
6. Каменская Д.Н., Панькова М.В., Атопкин Д.М., Брыков Вл.А. Дивергенция паралогичных генов гормона роста и цис-регуляторных участков у лососевых рыб // Молекулярная биология. 2017. Т. 51, № 2. С. 314–323.
7. Каменская Д.Н., Брыков Вл.А. Гены гормона роста у рыб: структура и дивергенция // Биология моря. 2020. Т. 46, № 4. С. 1–12.
8. Каменская Д.Н., Панькова М.В., Брыков Вл.А. Изменчивость экзонов и интронов в генах гормона роста у лососевых рыб // Молекулярная биология. 2020. Т. 54, № 6. С. 975–979.
9. Льюин Б. Гены. – М: БИНОМ: Лаборатория знаний. 2011. – 896 с.
10. Панькова М.В., Брыков Вл.А., Панькова В.В., Атопкин Д.М. Гены гормона роста рыб. Дивергенция последовательностей интронов у гольцов рода *Salvelinus* // Генетика. 2013. Т. 49, № 6. С. 1–8.
11. Панькова М. В., Кухлевский А. Д., Брыков Вл.А. Гены гормона роста: дивергенция кодирующих последовательностей у лососевых рыб // Генетика. 2017. Т. 53, №2. С. 201–213.

12. Разин С. В., Гаврилов А. А., Ульянов С. В. Регуляторные элементы эукариотического генома, контролирующие транскрипцию // Молекулярная биология. Т. 49, №. 2. С. 212–223.
13. Agellon L.B., Chen T.T. Rainbow trout growth hormone: molecular cloning of cDNA and expression in *Escherichia coli* // DNA. 1986. Vol. 5, № 6. P. 463–471.
14. Agellon L.B., Davies S.L., Chen T.T., Powers D.A. Structure of a fish (rainbow trout) growth hormone gene and its evolutionary implications // Proceedings of the National Academy of Sciences of the USA. 1988. Vol. 85, № 14. P. 5136–5140.
15. Allendorf F.W., Thorgaard G.H. Tetraploidy and the evolution of salmonid fishes // In: Evolutionary Biology of Fishes. – N. Y.: Plenum Press. 1984. PP. 1–53.
16. Almuly R., Cavari B., Ferstman H., Kolodny O., Funkenstein B. Genomic structure and sequence of the gilthead seabream (*Sparus aurata*) growth hormone-encoding gene: identification of minisatellite polymorphism in intron I // Genome. 2000. Vol. 43, № 5. P. 836–845.
17. Almuly R., Poleg-Danin Y., Gorshkov S., Gorshkova G., Rapoport B., Soller M., Kashi Y., Funkenstein B. Characterization of the 5'-flanking region of the growth hormone gene of the marine teleost, gilthead sea bream *Sparus aurata*: analysis of a polymorphic microsatellite in the proximal promoter // Fisheries Science. 2005. Vol. 71, № 3. P. 479–490.
18. Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J.J. Basic local alignment search tool // Journal of Molecular Biology. 1990. Vol. 215, № 3. P. 403–410.
19. Antequera F., Bird A. Number of CpG islands and genes in human and mouse // Proceedings of the National Academy of Sciences of the USA. 1993. Vol. 90, № 24. P. 11995–11999.
20. Arany Z., Sellers W.R., Livingston D.M., Eckner R. E1A-associated p300 and CREB-associated CBP belong to a conserved family of coactivators // Cell. 1994. Vol. 77, № 6. P. 799–800.
21. Arai N., Iigo M. Duplicated growth hormone genes in a passerine bird, the jungle crow (*Corvus macrorhynchos*) // Biochemical and Biophysical Research Communications. 2010. Vol. 397, № 3. P. 553–558.

22. Argenton F., Bernardini S., Puttini S., Colombo L., Bortolussi M. A TGACG motif mediates growth-hormone-factor-1/pituitary-transcriptional-activator-1-dependent cAMP regulation of the rainbow trout growth-hormone promoter // *European Journal of Biochemistry*. 1996. Vol. 238, № 3. P. 591–598.
23. Argenton F., Vianello S., Bernardini S., Lopreiato R., Colombo L., Bortolussi M. Trout GH promoter analysis reveals a modular pattern of regulation consistent with the diversification of GH gene control and function in vertebrates // *Molecular and Cellular Endocrinology*. 2002. Vol. 189, № 1–2. P. 11–23.
24. Ashburner M. A Biologist's view of the Drosophila genome annotation assessment project // *Genome Research*. 2000. Vol. 10, № 4. P. 391–393.
25. Ávila-Mendoza J., Carranza M., Pérez-Rueda E., Luna M., Arámburo C. Characterization of pituitary growth hormone and its receptor in the green iguana (*Iguana iguana*) // *General and Comparative Endocrinology*. 2014. Vol. 203, № 1. P. 281–295.
26. Balakirev E. S., Ayala F. J. Pseudogenes: Are they “junk” or functional DNA? // *Annual Review of Genetics*. 2003. Vol. 37. P. 123–151.
27. Barnett K.R., Hopkins R.L., Peyton D.K. A minisatellite in the growth hormone gene of Esocidae is derived from a single copy element in the salmonid genome // *Copeia*. 2007. Vol. 2007, № 1. P. 205–211.
28. Baron M.H. Developmental regulation of the vertebrate globin multigene family // *Gene Expression*. 1996. Vol. 6, № 3. P. 129–137.
29. Barrera-Saldana H.A. Growth hormone and placental lactogen: biology, medicine and biotechnology // *Gene*. 1998. Vol. 211, № 1. P. 11–18.
30. Barrett L. W. Fletcher S. Wilton S. D. Untranslated gene regions and other non-coding elements regulation of eukaryotic gene – N.Y.: Springer. 2013. – 56 p.
31. Barsh G. S., Seeburg P. H., Gelinas R. E. The human growth hormone gene family: structure and evolution of the chromosomal locus // *Nucleic Acids Research*. 1983. Vol. 11, № 12. P. 3939–3958.
32. Bart H.L., Reneau P.C., Doosey M.H., Bell C.B. Evolutionary divergence of duplicate copies of the growth hormone gene in suckers (Actinopterygii: Catostomidae) // *International Journal of Molecular Sciences*. 2010. Vol. 11, № 3. P. 1090–1102.

33. Barta A., Richards R.I., Baxter J.D., Shine J. Primary structure and evolution of rat growth hormone gene // Proceedings of the National Academy of Sciences of the USA. 1981. Vol. 78, № 8. P. 4867–4871.
34. Beadle, G.W., Tatum, E.L. Genetic control of biochemical reactions in *Neurospora* // Proceedings of the National Academy of Science of the USA. 1941. Vol. 27, № 11. P. 409–506.
35. Beato M., Chalepakis G., Schauer M., Slater E.P. DNA regulatory elements for steroid hormones // Journal of Steroid Biochemistry. 1989. Vol. 32, № 5. P. 737–748.
36. Ber R., Daniel V. Structure and sequence of the growth hormone-encoding gene from *Tilapia nilotica* // Gene. 1992. Vol. 113, № 2. P. 245–250.
37. Berget S.M., Moore C., Sharp P.A. Spliced segments at the 5' terminus of adenovirus 2 late mRNA // Proceedings of the National Academy of Sciences of the USA. 1977 Vol.74, № 8. P. 3171-3175.
38. Bernardi G. The neoselectionist theory of genome evolution // Proceedings of the National Academy of Sciences of the USA. 2007. Vol. 104, № 20. P. 8385–8390.
39. Bernardini S., Argenton F., Vianello S., Colombo L., Bortolussi M. Regulatory regions in the promoter and third intron of the growth hormone gene in rainbow trout, *Oncorhynchus mykiss* Walbaum // General and Comparative Endocrinology. 1999. Vol. 116, № 2. P. 261–271.
40. Berthelot C., Brunet F., Chalopin D., Juanchich A., Bernard M., Noe B., Bento P., Da Silva C., Labadie K., Alberti A., Aury J.-M., Louis A., Dehais P., Bardou P., Montfort J., Klopp C., Cabau C., Gaspin C., Thorgaard G. H., Boussaha M., Quillet E., Guyomard R., Delphine G., Bobe J., Volff J.-N., Genet C., Wincker P., Jaillon O., Crollius H. R., Guiguen Y. The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates // Nature Communications. 2014. Vol. 5, № 3657. P. 1–10.
41. Betancur-R R., Wiley E. O., Arratia G., Acero A., Bailly N., Masaki M., Lecointre G., Ortí G. Phylogenetic classification of bony fishes // BMC Evolutionary Biology. 2017. Vol. 7. Article No. 162. Doi: 10.1186/s12862-017-0958-3.
42. Blomhoff R., Blomhoff H.K. Overview of retinoid metabolism and function // Journal of Neurobiology. 2006. Vol. 66, № 7. P. 606–630.

43. Breathnach R., Chambon P. Organization and expression of eucaryotic split genes coding for proteins // *Annual Review of Biochemistry*. 1981. Vol. 50. P. 349–383.
44. Briggs M.R., Kadonaga J.T., Bell S.P., Tjian R. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1 // *Science*. 1986. Vol. 234 № 4772. P. 47–52.
45. Brown D.T. Histone variants: are they functionally heterogeneous? // *Genome Biology*. 2001. Vol. 2, № 7. P. 1–6.
46. Bucher P. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences // *Journal of Molecular Biology*. 1990. Vol. 212, № 4. P. 563–578.
47. Buggiotti L., Hellstrom M.A., Primmer C.R. Characterization of the first growth hormone gene sequence for a passerine bird – the pied flycatcher (*Ficedula hypoleuca*) // *DNA Sequence*. 2006. Vol. 17, № 6. P. 401–406.
48. Buggiotti L., Primmer C.R. Molecular evolution of the avian growth hormone gene and comparison with its mammalian counterpart // *Journal of Evolutionary Biology*. 2006. Vol. 19, № 3. P. 844–854.
49. Buratowski S. Mechanisms of gene activation // *Science*. 1995. Vol. 270, № 5243. P. 1773–1774.
50. Byrne C.R., Wilson B.W., Ward K.A. The isolation and characterisation of the ovine growth hormone gene // *Australian Journal of Biological Sciences*. 1987. Vol. 40, № 4. P. 459–468.
51. Chaurasia A., Tarallo A., Berna L., Yagi M., Agnisola C., D’Onofrio G. Length and GC content variability of introns among teleostean genomes in the light of the metabolic rate hypothesis // *PloS One*. 2014. Vol. 9, № 8. Article No. e103889. Doi: 10.1371/journal.pone.0103889.
52. Chen E.Y., Liao Y.C, Smith D.H., Barrera-Saldana H.A., Gelinas R.E., Seeburg P.H. The human growth hormone locus: nucleotide sequence, biology, and evolution // *Genomics*. 1989. Vol. 4, № 4. P. 479–497.

53. Chen T.T., Agellon L.B., Lin C.M., Tsai H.J., Zhang P., Gonzalez-Villasenor L.I., Powers D.A. Evolutionary implications of two rainbow trout growth hormone genes // *Fish Physiology and Biochemistry*. 1989. Vol. 7, № 1–6. P. 381–385.
54. Chiou C.S., Chen H.T., Chang W.C. The complete nucleotide sequence of the growth-hormone gene from the common carp (*Cyprinus carpio*) // *Biochimica et Biophysica Acta – Gene Structure and Expression*. 1990. Vol. 1087, № 1. P. 91–94.
55. Chow L.T., Gelinas R.E., Broker T.R., Roberts R.J. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA // *Cell*. 1977. Vol. 12, № 1. P. 1–8.
56. Courtois S.J., Lafontaine D.A., Lemaigre F.P., Durviaux S.M., Rousseau G.G. Nuclear factor-I and activator protein-2 bind in a mutually exclusive way to overlapping promoter sequences and trans-activate the human growth hormone gene // *Nucleic Acids Research*. 1990. Vol. 18, № 1. P. 57–64.
57. Darriba D., Taboada G.L., Doallo R., Posada D. jModelTest 2: more models, new heuristics and parallel computing // *Nature Methods*. 2012. Vol. 9 № 8. P. 772.
58. Das P., Meyer L., Seyfert H.M., Brockmann G., Schwerin M. Structure of the growth hormone encoding gene and its promoter in mice // *Gene*. 1996. Vol. 169, № 2. P. 209–213.
59. Daza D.O., Sundström G., Larsson T.A., Larhammar D. Evolution of the growth hormone–prolactin–somatolactin system in relation to vertebrate tetraploidizations // *Trends in Comparative Endocrinology and Neurobiology*. 2009. Vol. 1163, № 1. P. 491–493.
60. DeNoto F.M., Moore D.D., Goodman H.M. Human growth hormone DNA sequence and mRNA structure: possible alternative splicing // *Nucleic Acids Research*. 1981. Vol. 9, № 15. P. 3719–3730.
61. Deutsch M., Long M. Intron-exon structures of eukaryotic model organisms // *Nucleic Acids Research*. 1999. Vol. 27, № 15. P. 3219–3228.
62. Devlin R.H. Sequence of sockeye salmon type 1 and 2 growth hormone genes and the relationship of rainbow trout with Atlantic and Pacific salmon // *Canadian Journal of Fisheries and Aquatic Sciences*. 1993. Vol. 50, № 8. P. 1738–1748.
63. Devlin R.H., Biagi C.A., Yesaki T.Y. Growth, viability and genetic characteristics of GH transgenic coho salmon strains // *Aquaculture*. 2004. Vol. 236, № 1–4. P. 607–632.

64. Du S.J., Gong Z., Fletcher G.L., Shears M.A., King M.J., Idler D.R., Hew C.L. Growth enhancement in transgenic Atlantic salmon by the use of an “all fish” chimeric growth hormone gene construct // *Nature Biotechnology*. 1992. Vol 10, № 2. P. 176–181.
65. Du S.J., Devlin R.H., Hew C.L. Genomic structure of growth hormone genes in chinook salmon (*Oncorhynchus tshawytscha*): presence of two functional genes, GH-I and GH-II, and a male-specific pseudogene, GH- ψ // *DNA and Cell Biology*. 1993. Vol. 12, № 8. P. 739–751.
66. Duckworth M. L., Kirk K., Friesen H.G. Isolation and identification of a cDNA clone of rat placental lactogen II // *The Journal of Biological Chemistry*. 1986. Vol. 261, № 23. P. 10871–10878.
67. Edgar R.C. Muscle: multiple sequence alignment with high accuracy and high throughput // *Nucleic Acids Research*. 2004. Vol. 32, № 5. P. 1792–1797.
68. Elder J.F., Turner B.J. Concerted evolution of repetitive DNA sequences in eukaryotes // *The Quarterly Review of Biology*. 1995. Vol. 70, № 3. P. 297–320.
69. Farchi-Pisanty O., Sternberg H., Moav B. Transcriptional regulation of fish growth hormone gene // *Fish Physiology and Biochemistry*. 1997. Vol. 17, № 1–6. P. 237–246.
70. Forbes S.H., Knudsen K.L., North T.W., Allendorf F.W. One of two growth hormone genes in coho salmon is sex-linked // *Proceedings of the National Academy of Sciences of the USA*. 1994. Vol. 91, № 5. P. 1628–1631.
71. Gazave E., Marqués-Bonet T., Fernando O., Charlesworth B., Navarro A. Patterns and rates of intron divergence between humans and chimpanzees // *Genome Biology*. 2007. Vol. 8, № 2. Article No. R21. Doi:10.1186/gb-2007-8-2-r21.
72. George D.L., Phillips J.A. 3rd, Francke U., Seeburg P.H. The genes for growth hormone and chorionic somatomammotropin are on the long arm of human chromosome 17 in region q21 to qter // *Human Genetics*. 1981. Vol. 57, № 2. P. 138–141.
73. Gonzalez-Alvarez R., Revol de Mendoza A., Esquivel Escobedo D., Corrales Felix G., Rodriguez-Sanchez I., Gonzalez V., Davila G., Cao Q., de Jong P., Fu Y.X., Barrera Saldana H.A. Growth hormone locus expands and diverges after the separation of New and Old World monkeys // *Gene*. 2006. Vol. 380, № 1. P. 38–45.

74. Gordon D.F., Quick D.P., Erwin C.R., Donelson J.E., Maurer R.A. Nucleotide sequence of the bovine growth hormone chromosomal gene // *Molecular and Cellular Endocrinology*. 1983. Vol. 33, № 1. P. 81–95.
75. Green M.M., Green K.C. Crossing-over between alleles at the lozenge locus in *Drosophila melanogaster* // *Proceedings of the National Academy of Sciences of the USA* 1949. Vol. 35, № 10. P. 586–591.
76. Haddrill P. R., Charlesworth B., Halligan D. L., Andolfatto P. Patterns of intron sequence evolution in *Drosophila* are dependent upon length and GC content // *Genome Biology*. 2005. Vol. 6. R67.
77. Hall T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT // *Nucleic Acids Symposium Series*. 1999. Vol. 41. P. 95–98.
78. Harper M.E., Barrera-Saldana H.A., Saunders G.F. Chromosomal localization of the human placental lactogen-growth hormone gene cluster to 17q22-24 // *American Journal of Human Genetics*. 1982. Vol. 34, № 2. P. 227–234.
79. Harvey S., Hull K.L. Growth hormone // *Endocrine*. 1997. Vol. 7, № 3. P. 267–279.
80. Hirt H., Kimelman J., Birnbaum M.J., Chen E.Y., Seeburg P.H., Eberhardt N.L., Barta A. The human growth hormone locus: structure, evolution and allelic variation // *DNA*. 1987. Vol. 6, № 1. P. 59–70.
81. Holland L. Z., Albalat R., Azumi K. et al. The amphioxus genome illuminates vertebrate origins and cephalochordate biology // *Genome Research*. 2008. Vol. 18, № 7. P. 1100–1111.
82. Hong Y., Scharl M. Sequence of the growth hormone (GH) gene from the silver carp (*Hypophthalmichthys molitrix*) and evolution of the GH genes in vertebrates // *Biochimica et Biophysica Acta – Gene Structure and Expression*. 1993. Vol. 1174 № 3. P. 285–288.
83. Huang H., Brown D.D. Overexpression of *Xenopus laevis* growth hormone stimulates growth of tadpoles and frogs // *Proceedings of the National Academy of Sciences of the USA*. 2000. Vol. 97, № 1. P. 190–194.

84. Ingraham H.A., Chen R.P., Mangalam H.J., et al. A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype // *Cell*. 1988. Vol. 55, № 3. P. 519–529.
85. Johansen B., Johnsen O.C., Valla S. The complete nucleotide sequence of the growth-hormone gene from Atlantic salmon (*Salmo salar*) // *Gene*. 1989. Vol. 77, № 2. P. 317–324.
86. Jones B.K., Monks B.R., Liebhaber S.A., Cooke N.E. The human growth hormone gene is regulated by a multicomponent locus control region // *Molecular and Cellular Biology*. 1995. Vol. 15, № 12. P. 7010–7021.
87. Kadonaga J.T., Carner K.R., Masiarz F.R., Tjian R. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain // *Cell*. 1987. Vol. 51, № 6. P. 1079–1090.
88. Kalari K. R., Casavant M., Bair T. B., Keen H. L., Comeron J. M., Casavant T. L., Scheetz T. E. First Exons and Introns – A survey of GC content and gene structure in the human genome // *In Silico Biology*. 2006. Vol. 6, № 3. P.237–242.
89. Kansaku N., Soma A., Furukawa S., Hiyama G., Okabayashii H., Guemene D., Kuhnlein U., Zadworny D. Sequence of the domestic duck (*Anas platyrhynchos*) growth hormone-encoding gene and genetic variation in the promoter region // *Animal Science Journal*. 2008. Vol. 79, № 2. P.163–170.
90. Karin M., Theill L., Castrillo J. L., McCormick A., Brady H. Tissue-specific expression of the growth hormone gene and its control by growth hormone factor-1 // *Recent Progress in Hormone Research*. 1990. Vol. 46. P. 43–57.
91. Kawauchi H., Moriyama S., Yasuda A., Yamaguchi K., Shirahata K., Kubota J., Hirano T. Isolation and characterization of chum salmon growth hormone // *Archives of Biochemistry and Biophysics*. 1986. Vol. 244, № 2. P. 542–552.
92. Kawauchi H., Suzuki K., Yamazaki T., Moriyama S., Nozaki M., Yamaguchi K., Takahashi A., Youson J., Sower S.A. Identification of growth hormone in the sea lamprey, an extant representative of a group of the most ancient vertebrates // *Endocrinology*. 2002. Vol. 143, № 12. P. 4916–4921.

93. Kliewer S.A., Umesono K., Heyman R., Mangelsdorf D., Dyck Y., Evans R. Retinoid X receptor-COUP-TF interactions modulate retinoic acid signaling // Proceedings of the National Academy of Sciences of the USA. 1992. Vol. 89, № 4. P. 1448–1452.
94. Kobayashi T., Yasuda A., Yamaguchi K., Kawavchi H. Kikuyama S. The complete amino acid sequence of growth hormone of the bullfrog (*Rana catesbeiana*) // Biochimica et Biophysica Acta – Protein Structure and Molecular Enzymology. 1991. Vol. 1078, № 3. P. 383–387.
95. Koenig R.J., Brent G.A., Warne R.L., Larsen P.R., Moore D.D. Thyroid hormone receptor binds to a site in the rat growth hormone promoter required for induction by thyroid hormone // Proceedings of the National Academy of Sciences of the USA. 1987. Vol. 84, № 16. P. 5670–5674.
96. Kondrashov F.A., Kondrashov A.S. Role of selection in fixation of gene duplications // Journal of Theoretical Biology. 2006. Vol. 239, № 2. P. 141–151.
97. Krakauer D.C., Nowak M.A. Evolutionary preservation of redundant duplicated genes // Seminars in Cell and Developmental Biology. 1999. Vol. 10, № 5. P. 555–559.
98. Krawczak M., Chuzhanova N.A., Cooper D.N. Evolution of the proximal promoter region of the mammalian growth hormone gene // Gene. 1999. Vol. 237, № 1. P. 143–151.
99. Kumar S., Stecher G., Tamura K. MEGA 7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets // Molecular Biology and Evolution. 2016. Vol. 33, № 7. P. 1870–1874.
100. Kwok R.P., Lundblad J.R., Chrivia J.C., Richards J.P., Bächinger H.P., Brennan R.G., Roberts S.G., Green M.R., Goodman R.H. Nuclear protein CBP is a coactivator for the transcription factor CREB // Nature. 1994. Vol. 370, № 6486. P. 223–226.
101. Lalitha S. Primer Premier 5 // Biotechnology Software and Internet Report. 2000. Vol. 1, № 6. P. 270–272.
102. Leggatt R.A., Biagi C.A., Smith J.L., Devlin R.H. Growth of growth hormone transgenic coho salmon *Oncorhynchus kisutch* is influenced by construct promoter type and family line // Aquaculture. 2012. Vol. 356–357. P. 193–199.
103. Leidig F., Shepard A.R., Zhang W.G., Stelter A., Cattini P.A., Baxter J.D., Eberhardt N.L. Thyroid hormone responsiveness in human growth hormone-related genes.

Possible correlation with receptor-induced DNA conformational changes // *Journal of Biological Chemistry*. 1992. Vol. 267, № 2. P. 913–921.

104. Lemaigre F.P., Peers B., Lafontaine D.A., Mathy-Hartert M., Rousseau G.G., Belayew A., Martial J.A. Pituitary-specific factor binding to the human prolactin, growth hormone, and placental lactogen genes // *DNA*. 1989. Vol. 8, № 3. P. 149–159.

105. Li M., Gao Z., Ji D., Zhang S. Functional characterization of GH-like homolog in amphioxus reveals an ancient origin of GH/GH receptor system // *Endocrinology*. 2014. Vol. 155, № 12. P. 4818–4830.

106. Librado P., Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data // *Bioinformatics*. 2009. Vol. 25, № 11. P. 1451–1452.

107. Lioupis A., Wallis O.C., Wallis M. Cloning and characterisation of the gene encoding red deer (*Cervus elaphus*) growth hormone: implications for the molecular evolution of growth hormone in artiodactyls // *Journal of Molecular Endocrinology*. 1997. Vol. 19, № 3. P. 259–266.

108. Lioupis A., Nevo E., Wallis M. Cloning and characterisation of the gene encoding mole rat (*Spalax ehrenbergi*) growth hormone // *Journal of Molecular Endocrinology*. 1999. Vol. 22, № 1. P. 29–36.

109. Lipkin S.M., Naar A.M., Kalla K.A., Sack R.A., Rosenfeld M.G. Identification of a novel zinc finger protein binding a conserved element critical for Pit-1-dependent growth hormone gene expression // *Genes and Development*. 1993. Vol. 7, № 9. P. 1674–1687.

110. Liu W.S., Ma J.E., Li W.X., Zhang J.G., Wang J., Nie Q.H., Qiu F.F., Fang M.X., Zeng F., Wang X., Lin X.R., Zhang L., Chen S.H., Zhang X.Q. The long intron 1 of growth hormone gene from Reeves' turtle (*Chinemys reevesii*) correlates with negatively regulated gh expression in four cell lines // *International Journal of Molecular Sciences*. 2016. Vol. 17, № 4. P. 1–17.

111. Lynch M., Conery J. S. The evolutionary fate and consequences of duplicate genes // *Science*. 2000. Vol. 290, № 5494. P. 1151–1155.

112. Lynch M., Katju V. The altered evolutionary trajectories of gene duplicates // *Trends in Genetics*. 2004. Vol. 20, № 11. P. 544–549.

113. Magadum S., Banerjee U., Murugan P., Gangapur D., Ravikesavan R. Gene duplication as a major force in evolution // *Journal of Genetics*. 2013. Vol. 92. № 1. P. 155–161.
114. Male R., Nerland A.N., Lorens J.B., Telle W., Lossius I., Totland G.K. The complete nucleotide sequence of the Atlantic salmon growth hormone I gene // *Biochimica et Biophysica Acta (BBA) – Gene Structure and Expression*. 1992. Vol. 1130, № 3. P. 345–348.
115. Mangalam H.J. Albert V.R. Ingraham, H.A., Kapiloff M., Wilson L., Nelson C., Elsholtz H., Rosenfeld M.G. A pituitary POU domain protein, Pit-1, activates both growth hormone and prolactin promoters transcriptionally // *Genes and Development*. 1989. Vol. 3, № 7. P. 946–958.
116. Maniou Z., Wallis O.C., Wallis M. Episodic molecular evolution of pituitary growth hormone in Cetartiodactyla // *Journal of Molecular Evolution*. 2004. Vol. 58, № 6. P. 743–753.
117. Martial J.A., Seeburg P.H., Guenzi D., Goodman H.M., Baxter J.D. Regulation of growth hormone gene expression: synergistic effects of thyroid and glucocorticoid hormones // *Proceedings of the National Academy of Sciences of the USA*. 1977. Vol. 74, № 10. P. 4293–4295.
118. Martinat N., Anouassi A., Huet J. C., Pernollet J. C., Segard V., Combarous Y. Purification and partial characterization of growth hormone from the dromedary (*Camelus dromedarius*) // *Domestic Animal Endocrinology*. 1990. Vol. 7, № 4. P. 527–536.
119. Marzluff W. F., Gongidi P., Woods K. R., Jin J., Maltais L. J. The human and mouse replication-dependent histone genes // *Genomics*. 2002. Vol. 80, № 5. P. 487–498.
120. McKay S.J., Devlin R.H., Smith M.J. Phylogeny of Pacific salmon and trout based on growth hormone type-2 and mitochondrial NADH dehydrogenase subunit 3 DNA sequences // *Canadian Journal of Fisheries and Aquatic Sciences*. 1996. Vol. 53, № 5. P. 1165–1176.
121. McKay S.J., Trautner J., Smith M.J., Koop B.F., Devlin R.H., Evolution of duplicated growth hormone genes in autotetraploid salmonid fishes // *Genome*. 2004. Vol. 47, № 4. P. 714–723.
122. Miller W.L., Eberhardt N.L. Structure and evolution of the growth hormone gene family // *Endocrine Reviews*. 1983. Vol. 4, № 2. P. 97–130.

123. Montminy M.R., Gonzalez G.A., Yamamoto K.K. Regulation of cAMP-inducible genes by CREB // Trends in Neurosciences. 1990. Vol. 13, № 5. P. 184–188.
124. Moriyama S., Oda M., Takahashi A., Sower S. A., Kawauchi H. Genomic structure of the sea lamprey growth hormone-encoding gene // General and Comparative Endocrinology. 2006. Vol. 148, № 1. P. 33–40.
125. Moriyama S., Oda M., Yamazaki T. et al. Gene structure and functional characterization of growth hormone in dogfish, *Squalus acanthias* // Zoological Science. 2008. Vol. 25, № 6. P. 604–613.
126. Nakayama I., Biagi C.A., Koide N., Devlin R.H. Identification of a sex-linked GH pseudogene in one of two species of Japanese salmon (*Oncorhynchus masou* and *O. rhodurus*) // Aquaculture. 1999. Vol. 173, № 1–4. P. 65–72.
127. Norquay L.D., Yang X., Sheppard P., Gregoire S., Dodd J.G., Reith W., Cattini P.A. RFX1 and NF-1 associate with P sequences of the human growth hormone locus in pituitary chromatin // Molecular Endocrinology. 2003. Vol. 17, № 6. P. 1027–1038.
128. Norquay L.D., Yang X., Jin Y., Detillieux K.A., Cattini P.A. Hepatocyte nuclear factor-3 α binding at P sequences of the human growth hormone locus is associated with pituitary repressor function // Molecular Endocrinology. 2006. Vol. 20, № 3. P. 598–607.
129. Noso T., Lance V.A., Kawauchi H. Complete amino acid sequence of crocodile growth hormone // General and Comparative Endocrinology. 1995. Vol. 98, №3. P. 244–252.
130. Oakley T.H., Phillips R.B. Phylogeny of Salmonine fishes based on growth hormone introns: Atlantic (*Salmo*) and Pacific (*Oncorhynchus*) salmon are not sister taxa // Molecular Phylogenetics and Evolution. 1999. Vol. 11, № 3. P. 381–393.
131. Ohkubo T., Araki M., Tanaka M., Sudo S., Nakashima K. Molecular cloning and characterization of the yellowtail GH gene and its promoter: a consensus sequence for teleost and avian Pit-1/GHF-1 binding sites // Journal of Molecular Endocrinology. 1996. Vol. 16, № 1. P. 63–72.
132. Ohno S. Evolution by gene duplication. – N.Y.: Springer-Verlag. 1970. – 160 p.
133. Orkin S. H. Regulation of globin gene expression in erythroid cells // European Journal of Biochemistry. 1995. Vol. 231, № 2. P. 271–281.

134. Owerbach D., Martial J.A., Baxter J.D., Rutter W.J., Shows T.B. Genes for growth hormone, chorionic somatomammotropin and a growth hormone-like gene are located on chromosome 17 in humans // *Science*. 1980. Vol. 209, № 4453. P. 289–292.
135. Panicz R., Sadowski J., Drozd R. Genetic and structural characterization of the growth hormone gene and protein from tench, *Tinca tinca* // *Fish Physiology and Biochemistry*. 2012. Vol. 38, № 6. P. 1645–1653.
136. Park K-Y., Roe J-H. Identification of a negative regulatory site in the upstream region of bovine growth hormone gene // *Biochemical and Biophysical Research Communications*. 1996. Vol. 219, № 2. P. 354–358.
137. Portin P. The concept of the gene: short history and present status // *The Quarterly Review of Biology*. 1993. Vol. 68, № 2. P. 173–223.
138. Portin P. Historical development of the concept of the gene // *The Journal of Medicine and Philosophy*. 2002. Vol. 27, № 3. P. 257–286.
139. Perez-Maya A.A., Wallis M., Barrera-Saldana H.A. Structure and evolution of the gorilla and orangutan growth hormone loci // *Mammalian Genome*. 2016. Vol. 27, № 9–10. P. 511–523.
140. Phillips R.B., Matsuoka M.P., Konkel N.R., McKay S. Molecular systematic and evolution of the growth hormone introns in the Salmoninae // *Environmental Biology of Fishes*. 2004. Vol. 69, № 1–4. P. 433–440.
141. Proutski V., Holmes E. SWAN: sliding window analysis of nucleotide sequence variability // *Bioinformatics*. 1998. Vol. 14, № 5. P. 467–468.
142. Rajesh R., Majumdar K.C. A comparative account of the structure of the growth hormone encoding gene and genetic interrelationship in six species of the genus *Labeo* // *Fish Physiology and Biochemistry*. 2007. Vol. 33, № 4. P. 311–333.
143. Rodriguez-Sanchez I.P., Tejero M.E., Cole S.A., Comuzzie A.G., Nathanielsz P.W., Wallis M, Barrera-Saldana H.A. Growth hormone-related genes from baboon (*Papio hamadryas*): characterization, placental expression and evolutionary aspects // *Gene*. 2010. Vol. 450, № 1–2. P. 1–7.
144. Ronquist F., Teslenko M., Van Der Mark P., Ayres D. L., Darling A., Höhna S., Larget B., Liu L., Suchard M. A., Huelsenbeck J. P. MrBayes 3.2: efficient bayesian phylogenetic

inference and model choice across a large model space // *Systematic Biology*. 2012. Vol. 61, № 3. P. 539–542.

145. Ryyänänen H.J., Primmer C.R. Varying signals of the effects of natural selection during teleost growth hormone gene evolution // *Genome*. 2006. Vol. 49, № 1. P. 42–53.

146. Saunders M.C., Deakin J., Harrison G. A., Curlewis J. D. cDNA cloning of growth hormone from the brushtail possum (*Trichosurus vulpecula*) // *General and Comparative Endocrinology*. 1998. Vol. 111, № 1. P. 68–75.

147. Sekar M., Singh S.D., Gupta S. Cloning and characterization of *Pangasianodon hypophthalmus* growth hormone gene and its heterologous expression // *Applied Biochemistry and Biotechnology*. 2014. Vol. 173, № 6. P. 1446–1468.

148. Sekkali B., Brim H., Muller M., Argenton F., Bortolussi M., Colombo L., Belayew A., Martial J.A. Structure and functional analysis of a tilapia (*Oreochromis mossambicus*) growth hormone gene: activation and repression by pituitary transcription factor Pit-1 // *DNA and Cell Biology*. 1999. Vol. 18, № 6. P. 489–502.

149. Sekkali B., Belayew A., Bortolussi M., Martial J.A., Muller M. Pit-1 mediates cell-specific and cAMP-induced transcription of the tilapia GH gene // *Molecular and Cellular Endocrinology*. 1999. Vol. 152, № 1–2. P. 111–123.

150. Shewchuk B.M., Ho Y., Liebhaber S.A., Cooke N.E. A single base difference between Pit-1 binding sites at the hGH promoter and locus control region specifies distinct Pit-1 conformations and functions // *Molecular and Cellular Biology*. 2006. Vol. 26, № 17. P. 6535–6546.

151. Sternberg H., Moav B. Regulation of the growth hormone gene by fish thyroid/retinoid receptors // *Fish Physiology and Biochemistry*. 1999. Vol. 20, № 4. P. 331–339.

152. Su Y., Liebhaber S.A., Cooke N. The Human growth hormone gene cluster locus control region supports position-independent pituitary- and placenta-specific expression in the transgenic mouse // *Journal of Biological Chemistry*. 2000. Vol. 275, № 11. P. 7902–7909.

153. Tanaka M., Hosokawa Y., Watahiki M., Nakashima K. Structure of the chicken growth hormone-encoding gene and its promoter region // *Gene*. 1992. Vol. 112, № 2. P. 235–239.

154. Tanaka M., Toma Y., Ohkubo T., Sudo S., Nakashima K. Sequence of the flounder (*Paralichthys olivaceus*) growth hormone-encoding gene and its promoter region // *Gene*. 1995. Vol. 165, № 2. P. 321–322.
155. Theill L.E., Karin M. Transcriptional control of GH expression and anterior pituitary development // *Endocrine Reviews*. 1993. Vol. 14, № 6. P. 670–689.
156. Thompson J.D., Higgins D.G., Gibson T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions specific gap penalties and weight matrix choice // *Nucleic Acids Research*. 1994. Vol. 22, № 22. P. 4673–4680.
157. Truss M., Beato M. Steroid hormone receptors: interaction with deoxyribonucleic acid and transcription factors // *Endocrine Reviews*. 1993. Vol. 14, № 4. P. 459–478.
158. Valinsky A., Shani M., Gootwine E. Restriction fragment length polymorphism in sheep at the growth hormone locus is the result of variation in gene number // *Animal Biotechnology*. 1990. Vol. 1, № 2. P. 135–144.
159. Vanin E. F Processed pseudogenes: characteristics and evolution // *Annual Review of Genetics*. 1985. Vol. 19. P. 253–272.
160. Venkatesh B., Brenner S. Genomic structure and sequence of the pufferfish (*Fugu rubripes*) growth hormone-encoding gene: a comparative analysis of teleost growth hormone genes // *Gene*. 1997. Vol. 187, № 2. P. 211–215.
161. Von Schalburg K.R., Yazawa R., de Boer J, Lubieniecki K.P., Goh B., Straub C.A., Beetz-Sargent M.R., Robb A., Davidson W.S., Devlin R.H., Koop B.F. Isolation, characterization and comparison of Atlantic and Chinook salmon growth hormone 1 and 2 // *BMC Genomics*. 2008. Vol. 9. Article No. 522. Doi:10.1186/1471-2164-9-522.
162. Wagner A. The fate of duplicated genes: loss or new function? // *Bioessays*. 1998. Vol. 20, № 10. P. 785–788.
163. Walker W.H., Fitzpatrick S.L., Barrera-Saldana H.A., Resendez-Perez D., Saunders G.F. The human placental lactogen genes: structure, function, evolution and transcriptional regulation // *Endocrine Reviews*. 1991. Vol. 12, № 4. P. 316–328.
164. Wallis M. Variable evolutionary rates in the molecular evolution of mammalian growth hormones // *Journal of Molecular Evolution*. 1994. Vol. 38, № 6. P. 619–627.

165. Wallis M. The molecular evolution of vertebrate growth hormones: a pattern of near-stasis interrupted by sustained bursts of rapid change // *Journal of Molecular Evolution*. 1996. Vol. 43, № 2. P. 93–100.
166. Wallis M. Function switching as a basis for bursts of rapid change during the evolution of pituitary growth hormone // *Journal of Molecular Evolution*. 1997. Vol. 44, № 3. P. 348–350.
167. Wallis M. Molecular evolution of growth hormone // *The Biochemist*. 2014. Vol. 36, № 1. P. 4–8.
168. Wallis O.C., Wallis M. Cloning and characterization of the rabbit growth hormone-encoding gene // *Gene*. 1995. Vol. 163, № 2. P. 253–256.
169. Wallis O.C., Wallis M. Molecular evolution of growth hormone (GH) in Cetartiodactyla: cloning and characterization of the gene encoding GH from a primitive ruminant, the chevrotain (*Tragulus javanicus*) // *General and Comparative Endocrinology*. 2001. Vol. 123, № 1. P. 62–72.
170. Wallis O.C., Zhang Y.P., Wallis M. Molecular evolution of GH in primates: characterization of the GH genes from slow loris and marmoset defines an episode of rapid evolutionary change // *Journal of Molecular Endocrinology*. 2001. Vol. 26, № 3. P. 249–258.
171. Walsh J.B. How often do duplicated genes evolve new functions? // *Genetics*. 1995. Vol. 139, № 1. P. 421–428.
172. Watahiki M., Yamamoto M., Yamakawa M., Tanaka M., Nakashima K. Conserved and unique amino acid residues in the domains of the growth hormones. Flounder growth hormone deduced from the cDNA sequence has the minimal size in the growth hormone prolactin gene family // *Journal of Biological Chemistry*. 1989. Vol. 264, № 1. P. 312–316.
173. Westrich K.M., Konkol N.R., Matsuoka M.P., Phillips R.B. Interspecific relationships among charrs based on phylogenetic analysis of nuclear growth hormone intron sequences // *Environmental Biology of Fishes*. 2002. Vol. 64, № 1–3. P. 217–222.
174. Woychik R.P., Camper S.A., Lyons R.H., Horowitz S., Goodwin E.C., Rottman F.M. Cloning and nucleotide sequencing of the bovine growth hormone gene // *Nucleic Acids Research*. 1982. Vol. 10, № 22. P. 7197–7210.

175. Wu Y., Xu B., Koenig R.J. Thyroid hormone response element sequence and the recruitment of retinoid X receptors for thyroid hormone responsiveness // *Journal of Biological Chemistry*. 2001 Vol. 276, №6. P.3929–3936.
176. Xu, L., Lavinsky, R.M., Dasen, J.S., et al. Signal-specific co-activator domain requirements for Pit-1 activation // *Nature*. 1998. Vol. 395, № 6699. P. 301–306.
177. Yamada S., Hata J., Yamashita S. Molecular cloning of fish Pit-1 cDNA and its functional binding to promoter of gene expressed in the pituitary // *Journal of Biological Chemistry*. 1993. Vol. 268, №32. P. 24361–24366.
178. Yamaguchi K., Yasuda A., Lewis U. J., Yokoo Y., Kawauchi H. The complete amino acid sequence of growth hormone of an elasmobranch, the blue shark (*Prionace glauca*) // *General and Comparative Endocrinology*. 1989. Vol. 73, №2. P. 252–259.
179. Yamano Y., Abe M., Mikawa S., Kioka N., Manabe E., Sakai H., Komano T., Utsumi K., Iritani A. Structural analysis of repetitive DNA sequences the goat growth hormone gene region // *Agricultural and Biological Chemistry*. 1991. Vol. 55, №3. P. 633–639.
180. Yang B.Y., Chan K.M., Lin C.M., Chen T.T. Characterization of rainbow trout (*Oncorhynchus mykiss*) growth hormone 1 gene and the promoter region of growth hormone 2 gene // *Archives of Biochemistry and Biophysics*. 1997. Vol. 340, № 2. P. 359–368.
181. Yang M., Lin Y., Fan J., Yin Y., Yu P., Meng F., Du X., Han X., Cao X., Kong F., Huang A., Huang L., Zeng X., Bu G.. Characterization of growth hormone (GH) in Chinese soft-shelled turtle: Molecular identification, capability in activating GH receptor and tissue distribution // *Aquaculture Reports*. 2020. Vol. 18, P. 100407.
182. Yanofsky C., Carlton B.C., Guest J.R., Helinski D.R., Henning U.. On the colinearity of gene structure and protein structure // *Proceedings of the National Academy of Science of the USA*. 1964. Vol. 51, № 2. P. 266–272.
183. Yasuda A., Yamaguchi K., Papkoff H., Yokoo Y., Kawauchi H. The complete amino acid sequence of growth hormone from the Sea Turtle (*Chelonia mydas*) // *General and Comparative Endocrinology*. 1989. Vol. 13, № 2. P. 242–251.
184. Yowe D.L., Epping R.J. Cloning of the barramundi growth hormone encoding gene: a comparative analysis of higher and lower vertebrate growth hormone genes // *Gene*. 1995. Vol. 162, № 2. P. 255–259.

185. Yuri T., Kimball R.T., Braun E.L., Braun M.J. Duplication of accelerated evolution and growth hormone gene in passerine birds // *Molecular Biology and Evolution*. 2008. Vol. 25, № 2. P. 352–361.
186. Zanger K., Cohen L.E., Hashimoto K., Radovick S., Wondisford E. A novel mechanism for cyclic adenosine 3',5'-monophosphate regulation of gene expression by CREB-binding protein // *Molecular Endocrinology*. 1999. Vol. 13, № 2. P. 268–275.
187. Zhang J. Evolution by gene duplication: an update // *Trends in Ecology and Evolution*. 2003. Vol.18, № 6. P. 292–298.
188. Zhu Z., He L., Chen T.T. Primary-structural and evolutionary analyses of the growth-hormone gene from grass carp (*Ctenopharyngodon idellus*) // *European Journal of Biochemistry*. 1992. Vol. 207, № 2. P. 643–648.

ПРИЛОЖЕНИЕ I

Таблица

Значения дивергенции промоторной области гена гормона роста *gh1* для представителей трех родов семейства Salmonidae (*Oncorhynchus*, *Salmo*, *Salvelinus*) ± стандартное отклонение.

| Виды | <i>S. malma</i> | <i>S. curilus</i> | <i>S. levanidovi</i> | <i>S. taranetzi</i> | <i>O. nerka</i> | <i>O. tshawytscha</i> |
|-----------------------|-----------------|-------------------|----------------------|---------------------|-----------------|-----------------------|
| <i>S. malma</i> | | | | | | |
| <i>S. curilus</i> | 0,00329±0,00164 | | | | | |
| <i>S. levanidovi</i> | 0,00329±0,00164 | 0,00000 | | | | |
| <i>S. taranetzi</i> | 0,00658±0,00329 | 0,00329±0,00164 | 0,00329±0,00164 | | | |
| <i>O. nerka</i> | 0,07921±0,03960 | 0,07591±0,03795 | 0,07591±0,03795 | 0,07921±0,03960 | | |
| <i>O. tshawytscha</i> | 0,07591±0,03795 | 0,07261±0,03630 | 0,07261±0,03630 | 0,07591±0,03795 | 0,04276±0,02138 | |
| <i>S. salar</i> | 0,05705±0,02852 | 0,05369±0,02685 | 0,05369±0,02685 | 0,05705±0,02852 | 0,06376±0,03188 | 0,06711±0,03356 |

Примечание. Оранжевым цветом выделены значения дивергенции, полученные при сравнении четырех видов гольцов рода *Salvelinus*. Синим цветом значения дивергенции, полученные при сравнении гольцов с другими видами семейства Salmonidae.

ПРИЛОЖЕНИЕ II

Таблица

Значения дивергенции промоторной области гена гормона роста *gh2* для представителей трех родов семейства Salmonidae (*Oncorhynchus*, *Salmo*, *Salvelinus*) ± стандартное отклонение.

| Виды | <i>S. malma</i> | <i>S. curilus</i> | <i>S. levanidovi</i> | <i>S. taranetzi</i> | <i>O. nerka</i> | <i>O. tshawytscha</i> |
|-----------------------|-----------------|-------------------|----------------------|---------------------|-----------------|-----------------------|
| <i>S. malma</i> | | | | | | |
| <i>S. curilus</i> | 0,00000 | | | | | |
| <i>S. levanidovi</i> | 0,00658±0,00329 | 0,00658±0,00329 | | | | |
| <i>S. taranetzi</i> | 0,00658±0,00329 | 0,00658±0,00329 | 0,00658±0,00329 | | | |
| <i>O. nerka</i> | 0,08638±0,04319 | 0,08638±0,04319 | 0,08638±0,04319 | 0,09302±0,04651 | | |
| <i>O. tshawytscha</i> | 0,10631±0,05316 | 0,10631±0,05316 | 0,09967±0,04983 | 0,10631±0,05316 | 0,05316±0,02658 | |
| <i>S. salar</i> | 0,04319±0,02159 | 0,04319±0,02159 | 0,03987±0,01993 | 0,04651±0,02326 | 0,08725±0,04362 | 0,10403±0,05201 |

Примечание. Оранжевым цветом выделены значения дивергенции, полученные при сравнении четырех видов гольцов рода *Salvelinus*. Синим цветом значения дивергенции, полученные при сравнении гольцов с другими видами семейства Salmonidae.

ПРИЛОЖЕНИЕ III

Таблица

Значения дивергенции транскрибируемой части гена гормона роста *gh1* для представителей трех родов семейства Salmonidae (*Oncorhynchus*, *Salmo*, *Salvelinus*) ± стандартное отклонение.

| Виды | <i>S. malma</i> | <i>S. curilus</i> | <i>S. levanidovi</i> | <i>S. taranetzi</i> | <i>O. nerka</i> | <i>O. tshawytscha</i> |
|-----------------------|-----------------|-------------------|----------------------|---------------------|-----------------|-----------------------|
| <i>S. malma</i> | | | | | | |
| <i>S. curilus</i> | 0,00246±0,00123 | | | | | |
| <i>S. levanidovi</i> | 0,00245±0,00123 | 0,00382± 0,00191 | | | | |
| <i>S. taranetzi</i> | 0,00545±0,00272 | 0,00546±0,00273 | 0,00626±0,00313 | | | |
| <i>O. nerka</i> | 0,05311±0,02655 | 0,05265±0,02632 | 0,05396±0,02698 | 0,05365±0,02682 | | |
| <i>O. tshawytscha</i> | 0,05288±0,02644 | 0,05242±0,02621 | 0,05316±0,02658 | 0,05426±0,02713 | 0,02216±0,01108 | |
| <i>S. salar</i> | 0,05014±0,02507 | 0,05023± 0,02512 | 0,05108±0,02554 | 0,05012±0,02506 | 0,06185±0,03092 | 0,06101±0,03051 |

Примечание. Оранжевым цветом выделены значения дивергенции, полученные при сравнении четырех видов гольцов рода *Salvelinus*. Синим цветом значения дивергенции, полученные при сравнении гольцов с другими видами семейства Salmonidae.

ПРИЛОЖЕНИЕ IV

Таблица

Значения дивергенции транскрибируемой части гена гормона роста *gh2* для представителей трех родов семейства Salmonidae (*Oncorhynchus*, *Salmo*, *Salvelinus*) ± стандартное отклонение.

| Виды | <i>S. malma</i> | <i>S. curilus</i> | <i>S. levanidovi</i> | <i>S. taranetzi</i> | <i>O. nerka</i> | <i>O. tshawytscha</i> |
|-----------------------|-----------------|-------------------|----------------------|---------------------|-----------------|-----------------------|
| <i>S. malma</i> | | | | | | |
| <i>S. curilus</i> | 0,00272±0,00136 | | | | | |
| <i>S. levanidovi</i> | 0,01790±0,00895 | 0,01638±0,00819 | | | | |
| <i>S. taranetzi</i> | 0,01304±0,00652 | 0,01335±0,00667 | 0,02309±0,01155 | | | |
| <i>O. nerka</i> | 0,05323±0,02662 | 0,05219±0,02610 | 0,04773±0,02387 | 0,05232±0,02616 | | |
| <i>O. tshawytscha</i> | 0,05533±0,02767 | 0,05434±0,02717 | 0,05073±0,02536 | 0,05378±0,02689 | 0,02440±0,01220 | |
| <i>S. salar</i> | 0,04050±0,02025 | 0,03954±0,01977 | 0,03537±0,01768 | 0,03833±0,01916 | 0,04525±0,02263 | 0,04801±0,02401 |

Примечание. Оранжевым цветом выделены значения дивергенции, полученные при сравнении четырех видов гольцов рода *Salvelinus*. Синим цветом значения дивергенции, полученные при сравнении гольцов с другими видами семейства Salmonidae.