

На правах рукописи

ПАНЬКОВА

Марина Владимировна

**СТРУКТУРА И ЭВОЛЮЦИЯ ГЕНОВ ГОРМОНА РОСТА
ЛОСОСЁВЫХ РЫБ (SALMONIDAE)**

03.02.07 – генетика

АВТОРЕФЕРАТ

диссертации на соискание учёной степени

кандидата биологических наук

Владивосток – 2016

Работа выполнена в Федеральном государственном автономном образовательном учреждении высшего профессионального образования «Дальневосточный федеральный университет» (ДВФУ)

Научный руководитель:

доктор биологических наук, профессор **Брыков Владимир Алексеевич**

Официальные оппоненты:

Челомина Галина Николаевна, доктор биологических наук, профессор, Федеральное государственное бюджетное учреждение науки Биолого-почвенный институт Дальневосточного отделения Российской академии наук, главный научный сотрудник лаборатории паразитологии

Щербаков Дмитрий Юрьевич, доктор биологических наук, профессор, Федеральное государственное бюджетное учреждение науки Лимнологический институт Сибирского отделения Российской академии наук, заведующий лабораторией геносистематики

Ведущая организация:

Федеральное государственное бюджетное учреждение науки Институт общей генетики им. Н.И. Вавилова Российской академии наук

Защита диссертации состоится «27» мая 2016 г. в 13 часов на заседании диссертационного совета Д 005.008.01 при Федеральном государственном бюджетном учреждении науки Институт биологии моря им. А.В. Жирмунского Дальневосточного отделения Российской академии наук по адресу: 690041, г. Владивосток, ул. Пальчевского, д. 17. Факс: (423)2310-900, e-mail: inmarbio@mail.primorye.ru

Отзывы просим присылать на e-mail: mvaschenko@mail.ru

С диссертацией можно ознакомиться в библиотеке и на сайте Федерального государственного бюджетного учреждения науки Институт биологии моря им. А.В. Жирмунского Дальневосточного отделения Российской академии наук: <http://www.imb.dvo.ru/misc/dissertations/index.php/soviet-d-005-008-01/25-pankova-marina-vladimirovna>

Автореферат разослан

« ____ » _____ 2016 г.

Ученый секретарь диссертационного совета, кандидат биологических наук

Ващенко

Ващенко Марина Александровна

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Дубликации генов играют очень важную роль в эволюции. Как правило, копии дублицированных или амплифицированных генов оказываются под меньшим давлением отбора, накапливают изменения с более высокой частотой и, со временем могут приобретать новые функции.

Гормон роста, также известный как соматотропин, является ключевым белком, ответственным за регуляцию соматического роста и участвует в многочисленных физиологических процессах, включая ионный баланс, липидный и белковый обмен, размножение, иммунный ответ, а также различные аспекты поведения (Coker, Arman, 2009). Очевидно из функции, что последовательность гена гормона роста должна быть консервативной. В большинстве групп позвоночных животных, где этот ген представлен единичной копией, это действительно так (Wallis, 1996; Rajesh, Mudjundar, 2007), хотя в некоторых таксонах скорости дивергенции могут отличаться на порядок величины.

Ген гормона роста типично представлен в геноме млекопитающих и птиц одной копией, в то же время есть исключения (Yuri et al., 2008; Arai, Iigi, 2010). Два гена GH найдено у воробьиных птиц (Yuri et al., 2008; Arai, Iigi, 2010). У высших приматов он представлен кластером из пяти гомологичных генов, которые, за исключением собственно гена гормона роста, экспрессируются только в плаценте (Hirt et al., 1987), но у других приматов такого кластера не обнаружено. У большинства исследованных видов рыб ген гормона роста представлен двумя несвязанными функциональными паралогичными генами, GH1 и GH2 (Devlin, 1993). Структура генов у рыб типична для позвоночных и включает 6 экзонов и 5 интронов, при этом показано, что у рыб ген гормона роста обладает более высоким уровнем изменчивости, чем у других позвоночных, и эволюционирует с большей скоростью (Ruupänen, Primmer, 2006), что, вероятно, обусловлено наличием двух функциональных копий гена.

Лососевые рыбы представляют собой уникальную группу, сформировавшуюся после события автотетраплоидизации и последующей дивергенции, и, таким образом, являются естественными и относительно недавними полиплоидами (Allendorf, Thorgaard, 1984). Как следствие, многие гены

в этой таксономической группе видов оказались множественными, в том числе и ген гормона роста.

Целью настоящей работы было сравнение структуры и дивергенции последовательностей двух генов гормона роста (GH1 и GH2) у видов гольцов рода *Salvelinus*. Для этого были поставлены следующие **задачи**:

1. Получить полные нуклеотидные последовательности двух генов гормона роста и определить их структуру у четырёх азиатских видов гольцов (*S. curilus*, *S. malma*, *S. taranetzi*, *S. levanidovi*).
2. Провести сравнительный анализ полученных нуклеотидных последовательностей двух генов GH у четырёх азиатских видов гольцов и лососевых в целом.
3. Провести филогенетический и сравнительный анализ дивергенции интронных и экзонных последовательностей генов GH лососевых рыб.

Степень разработанности. Известно, что многие гены в геномах эукариот являются дублированными или множественными. Значимость дубликаций генов для эволюции видов очевидна, однако сам процесс приобретения новых функций у дублированных копий изучен слабо (Zhang, 2003; Magadum et al., 2013). Гены гормона роста во многих группах позвоночных являются дублированными. Наиболее изученными являются гены гормона роста у человека. Менее изученными являются дублированные гены гормона роста у рыб, хотя использование их в генетической инженерии оказалось очень успешным (Du et al., 1992; Wu et al., 2003). Полные последовательности гена гормона роста известны у немногих видов даже у лососёвых рыб. Практически нет данных по сравнению дивергенции функционально различных участков генов-паралогов внутри таксонов и большая часть публикаций по GH лососёвых посвящена исследованию интронных участков (Driscoll et al., 1998; Oakley, Phillips, 1999; Phillips et al., 2004). Возможно, что субфункционализация генов-паралогов определяется не только кодирующими последовательностями, но и некодирующими участками.

Научная новизна. Впервые получены и детально охарактеризованы полные нуклеотидные последовательности двух генов гормона роста, GH1 и GH2, у четырёх азиатских видов гольцов: *S. curilus*, *S. malma*, *S. taranetzi*, *S. levanidovi*.

Полученные последовательности зарегистрированы в международной базе данных GenBank/NCBI. Впервые проведён сравнительный анализ дивергенции интронных и экзонных последовательностей генов GH гольцов рода *Salvelinus*. Показано, что скорость дивергенции двух паралогичных генов гормона роста лососевых рыб различна. Это обусловлено влиянием разнонаправленного и отличающегося по силе отбора на дублированные в процессе эволюции гены. Филогенетический и сравнительный анализы экзонных и интронных последовательностей свидетельствуют о давней независимой эволюции генов GH1 и GH2.

Теоретическая и практическая значение работы. Полученные результаты дополнили существующие на сегодняшний день представления о структуре гена гормона роста рыб. Настоящая работа вносит существенный вклад в исследования эволюции дублированных генов гормона роста рыб, а также филогении гольцов и семейства лососёвых рыб (*Salmonidae*) в целом. Результаты работы могут быть использованы для разработки практических заданий для студентов университета при анализе длинных фрагментов ДНК. Полученные результаты используются в настоящее время для создания полноразмерной генетической конструкции гормона роста с целью получения трансгенных линий рыб с более высокой скоростью роста в условиях аквакультуры.

Методология и методы диссертационного исследования. В данной работе были применены различные молекулярно-генетические методы получения и анализа фрагментов ДНК. Для амплификации фрагментов GH генов использовали методы полимеразной цепной реакции (ПЦР) и клонирования. Для разделения продуктов ПЦР и определения размеров полученных фрагментов использовали метод электрофореза в агарозном геле. При определении нуклеотидной последовательности фрагментов ДНК использовали метод флуоресцентно-меченых терминаторов (ddNTP). Полученные данные обработаны с помощью современных статистических программ.

Положения, выносимые на защиту:

1. Ген гормона роста представлен у лососёвых рыб двумя паралогичными генами, которые имеют сходную структуру.

2. Гены GH1 и GH2 лососёвых находятся под разным давлением очищающего отбора.

3. Направленность и сила отбора в генах-паралогах может быть различной в разных филетических линиях.

Степень достоверности результатов. Достоверность результатов исследования была обеспечена использованием современных молекулярно-генетических подходов и статистической обработкой материалов. О достоверности экспериментальных результатов также свидетельствует анализ многих клонов при клонировании фрагментов и воспроизводимость полученных последовательностей.

Апробация работы. Результаты работы были представлены на X и XI региональных конференциях студентов, аспирантов вузов и научных организаций Дальнего Востока России «Актуальные проблемы экологии, морской биологии и биотехнологии» (Владивосток, 2011 и 2012 гг.), на VII Международном симпозиуме по гольцам (Южно-Сахалинск, 2012) и на II международной конференции «Современные проблемы биологической эволюции» (Москва, 2014).

Публикации. По теме диссертации опубликовано 7 работ, в том числе три статьи в журналах из списка, рекомендованного ВАК.

Структура и объём диссертации. Диссертация изложена на 140 страницах и состоит из введения, обзора литературы, материалов и методов, результатов, обсуждения, заключения, выводов, списка литературы и приложения. Работа содержит 32 таблицы и 20 рисунков. Список литературы состоит из 174 наименований, из них 167 на английском языке.

Личный вклад автора. Экспериментальная часть работы была выполнена соискателем самостоятельно. Соискатель непосредственно участвовал в анализе и интерпретации полученных результатов, в представлении результатов на конференциях и подготовке публикаций по результатам исследований.

Благодарности. Выражаю огромную благодарность своему научному руководителю д.б.н. Вл.А. Брыкову за руководство, помощь и ценные советы на всех этапах исследования; д.б.н. Э.Я. Костецкому за поддержку и внимание в период аспирантуры. Отдельную благодарность выражаю к.б.н. В.В. Паньковой и

к.б.н. Д.М. Атопкину за помощь в освоении методов молекулярной биологии и разностороннее содействие и поддержку, оказанные в ходе проведения работы.

Работа выполнена при частичной финансовой поддержке РФФ (грант № 14-50-00034), ДВО РАН (грант № 09-I-П22-01) и Правительства РФ (грант № 11.G34.31.0010).

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Материалы и методы

Объектами исследования послужили 4 вида голецов северо-западной части Тихого Океана: северная мальма – *Salvelinus malma*, южная мальма – *S. curilus*, *syn. S. malma krascheninnikovi*, голец Таранца – *S. taranetzi*, голец Леванидова – *S. levanidovi*. В работе использовалась одна особь каждого вида. Для сравнительного и филогенетического анализа из базы данных GenBank/NCBI были взяты полные последовательности ДНК генов GH *Oncorhynchus nerka*, *Oncorhynchus tshawytscha*, *Salmo salar*, *Coregonus lavaretus*, последовательности мРНК GH *Salmo trutta*, *Oncorhynchus mykiss*, *Esox lucius*, а также последовательности интронов С и D генов GH1 и GH2 22 видов лососевых рыб.

Аmplification последовательностей генов GH1 и GH2 проводили с использованием пяти пар специфических праймеров (рис. 1). Продукты амплификации разделяли в 1.2% агарозном геле, вырезали и далее использовали для клонирования и секвенирования.

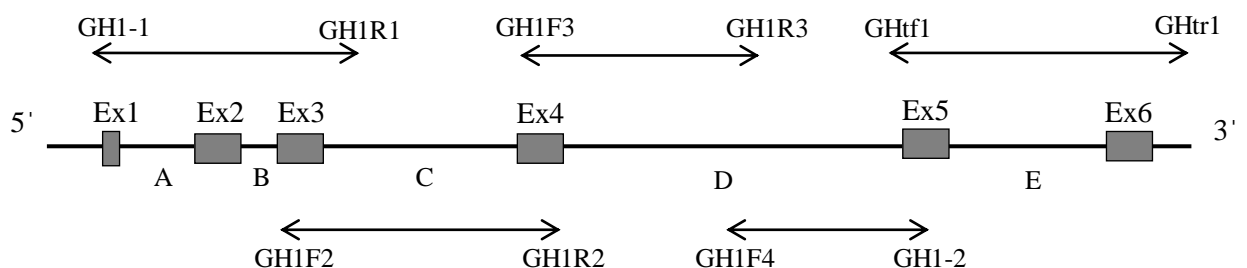


Рис. 1. Стратегия амплификации и секвенирования генов GH1 и GH2 видов рода *Salvelinus*.

Построение филогенетических деревьев на основе нуклеотидных последовательностей выполняли методом объединения ближайшего соседа (Neighbor-Joining, NJ), максимального правдоподобия (maximum likelihood, ML) в программе PAUP 4b10 (Swofford, 2002), и Байесовским методом (Bayesian Inference, BI) в программе MrBayesV.3.1.2 (Huelsenbeck, Ronquist, 2001; Ronquist, Huelsenbeck, 2003). Построение филогенетических деревьев по аминокислотным последовательностям выполняли методами NJ и ML в программе MEGA 5.0 (Tamura et al. 2011). Устойчивость полученных филогенетических деревьев оценивали методом бутстреп-анализа (bootstrap analysis; Felsenstein, 1985) используя 1000 реплик.

Количество синонимичных замен на синонимичный сайт (dS) и количество несинонимичных замен на несинонимичный сайт (dN) были оценены методом Нея-Годжобори (Nei-Gojobori, 1986) в программе MEGA 5.0 (Tamura et al., 2011). Определение вида отбора произведено с помощью дифференции несинонимичной и синонимичной дистанций ($Dd = dN - dS$). Вероятность отклонения от нулевой гипотезы определяли Z-тестом в программе MEGA 5.0 (Tamura et al., 2011). Для определения сайтов, находящихся под действием отбора использовали три метода (SLAC – single-likelihood ancestor counting, FEL – fixed effects likelihood, REL – random effects likelihood), входящие в пакет программ HyPhy, доступного на сервере Datamonkey (www.datamonkey.org, Kosakovsky Pond, Frost, 2005).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Структура генов

Структура генов GH1 и GH2 у всех исследованных видов совпадает и включает в себя 6 экзонов (1–6) и 5 интронов (A, B, C, D, E) (рис. 2). Последовательности гена GH1 для четырёх видов гольцов составили: *S. curilus* (3666 п.н.), *S. malma* (3673 п.н.), *S. taranetzi* (3675 п.н.), *S. levanidovi* (3672 п.н.). Для гена GH2: *S. curilus* (3305 п.н.), *S. malma* (3305 п.н.), *S. taranetzi* (3300 п.н.), *S. levanidovi* (3299 п.н.). Размер последовательности каждой копии гена GH незначительно отличается у разных видов гольцов, но варьирует между GH1 и

GH2. Разница в длине между копиями GH возникает за счёт разной длины интронов.

Размер каждого экзона двух генов GH идентичен у всех исследованных видов гольцов и составляет 74 п.н., 140 п.н., 117 п.н., 156 п.н., 147 п.н. и 63 п.н. соответственно. Первый экзон генов GH включает нетранслируемую область, кодирующую лидерную последовательность от точки начала транскрипции до стартового кодона ATG и транслируемую область длиной 10 п.н., которая включает первые три кодона и первый нуклеотид четвёртого кодона сигнального пептида, отщепляемого после трансляции при созревании функционального белка (рис. 2). Второй экзон включает последние два нуклеотида четвёртого кодона и остальные нуклеотиды для 18 аминокислот сигнального пептида, а также аминокислот зрелого белка с 1 по 28. Экзоны III, IV, V, VI кодируют 29–67, 68–119, 120–168, 169–188 аминокислоты соответственно.

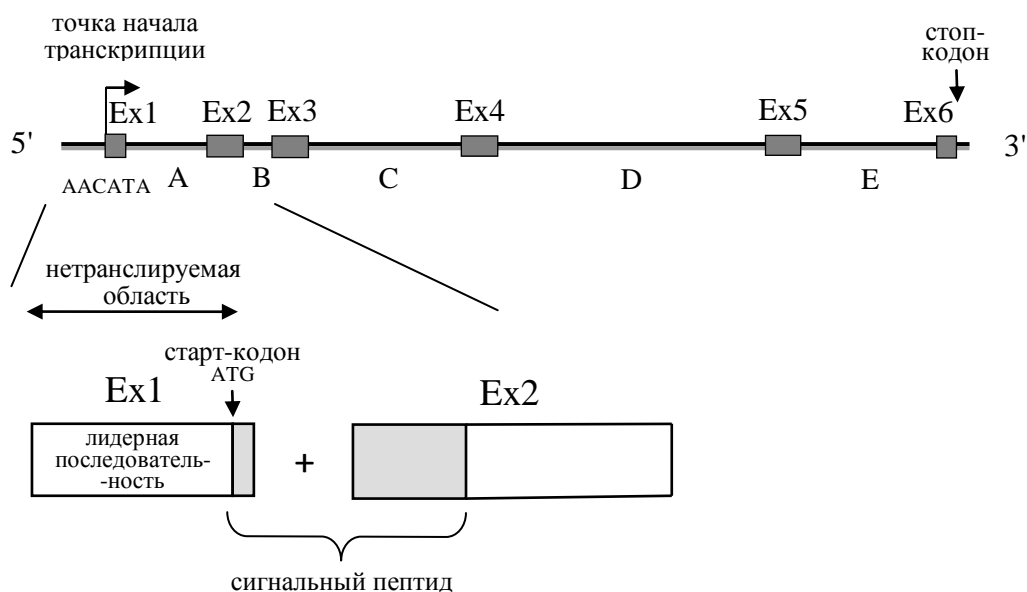


Рис. 2. Схематичное изображение структуры гена гормона роста.

Кодирующая область генов GH1 и GH2 (630 п.н.) соответствует аминокислотной последовательности предшественника длиной 210 остатков, включающих 188 аминокислот зрелого белка и 22 аминокислоты сигнального пептида. Открытая рамка считывания генов GH1 гольцов, начинается ATG кодоном и заканчивается стоп-кодом TAG. Открытая рамка считывания генов GH2 также начинается ATG кодоном, но заканчивается стоп-кодом TAA. Такая

структура генов согласуется с данными других авторов для других видов лососевых рыб.

В зрелой последовательности GH1 и GH2 гольцов найдено четыре функционально важных остатка цистеина, положение которых совпадает в обоих GH у всех четырёх видов (49, 161, 178, 186). Идентичное положение цистеинов характерно для GH других представителей семейства Salmonidae (Rentier-Delrue et al., 1989).

Сравнительный анализ генов GH четырёх видов гольцов рода *Salvelinus* показал высокий уровень сходства кодирующих последовательностей. В гене GH1 гольцов выявлено две нуклеотидные замены, которые не изменяют аминокислотную последовательность GH1. В кодирующей части гена GH2 обнаружено 9 нуклеотидных замен. Две замены приводят к изменению аминокислотной последовательности (табл. 1). Таким образом, ген GH2 гольцов более изменчив, чем ген GH1. При сравнении последовательностей генов GH1 и GH2 гольцов выявлено 34 нуклеотидные замены. Последовательности зрелого белка отличаются 8 аминокислотными заменами, сигнальный пептид – одной аминокислотой (табл. 1).

Анализ аминокислотных последовательностей GH гольцов с последовательностями других видов лососёвых рыб (*O. nerka*, *O. mykiss*, *O. tshawytscha*, *S. salar* и *S. trutta*) показал, что GH1 более изменчив (18 аминокислотных замен), чем GH2 (15 аминокислотных замен). Количество нуклеотидных замен, наоборот, больше в гене GH2 (48 замен), чем в GH1 (36 замен). Однако при сравнении последовательностей GH отдельно для видов каждого рода (*Salvelinus*, *Oncorhynchus*, *Salmo*), показано, что у видов рода *Oncorhynchus*, также как и у видов рода *Salvelinus*, последовательность GH1 более консервативна, чем GH2. Последовательности GH, включая сигнальный пептид, у *O. nerka*, *O. mykiss*, *O. tshawytscha* содержат 5 и 10 аминокислотных замен для GH1 и GH2 соответственно. У видов рода *Salmo* (*S. salar*, *S. trutta*), наоборот, последовательность GH2 более консервативна, чем GH1 (9 и 4 аминокислотных замен для GH1 и GH2 соответственно) (табл. 1).

Таблица 1

Аминокислотные замены в GH лососёвых рыб

	Положение замены	Сигнальный пептид			Зрелый белок																							
		8	13	18	1	12	20	29	30	34	46	50	58	61	73	76	90	96	100	119	123	125	134	135	149	152	181	
GH1	<i>S. curilus</i>	L	V	S	M	N	M	E	V	P	L	N	I	H	H	Y	T	I	S	T	D	V	Q	Q	D	V	S	
	<i>S. malma</i>
	<i>S. taranetzi</i>
	<i>S. levanidovi</i>
	<i>S. salar</i>	M	G	L	.	F	.	M	.	K
	<i>S. trutta</i>	M	.	.	I	S	L	D	G	.	.	.	V	.	.	F	I	M	A
	<i>O. nerka</i>	M	A	.	I	S	L	D	G	.	.	I	V	.	.	F	I	M	A	H	E	.	.	.
	<i>O. tshawytscha</i>	M	.	.	I	S	L	D	G	.	.	.	V	.	.	F	I	M	A	.	.	L
	<i>O. mykiss</i>	M	.	.	I	S	L	D	G	.	.	.	V	.	.	F	I	M	A
GH2	<i>S. curilus</i>	M	L	.	G	S	.	.	.	Q	.	.	A	.	.	K	.	.	.	H	.	.	.	
	<i>S. malma</i>	M	L	.	G	S	.	.	.	Q	.	.	A	.	.	K	.	.	.	H	.	.	.	
	<i>S. taranetzi</i>	M	L	.	G	S	.	.	.	Q	.	.	A	H	.	.	.	
	<i>S. levanidovi</i>	M	L	.	G	S	.	.	.	Q	.	.	A	M	.	K	.	.	.	H	.	.	.	
	<i>S. salar</i>	M	L	.	G	S	.	.	.	Q	.	F	A	M	.	K	.	.	.	H	.	I	.	
	<i>S. trutta</i>	M	.	G	.	.	L	.	G	Q	.	F	.	M	.	K	.	.	.	H	.	.	.	
	<i>O. nerka</i>	M	L	.	G	S	.	.	.	Q	R	F	.	M	.	E	E	I	.	H	.	.	.	
	<i>O. tshawytscha</i>	M	L	.	G	S	V	.	.	Q	.	F	.	M	Y	K	.	.	L	H	.	.	.	
	<i>O. mykiss</i>	M	.	G	.	.	L	.	G	Q	.	F	.	M	.	K	.	.	.	H	.	.	Y	

Дивергенция кодирующих последовательностей генов GH

Средняя величина различий между кодирующими последовательностями паралога у лососевых рыб составила 4,8% нуклеотидных замен, внутри видов рода *Salvelinus* 4,4%. При анализе различными методами были получены сходные топологии филогенетических деревьев. На рис. 3 приведено филогенетическое дерево, построенное на основе нуклеотидных последовательностей экзонов с использованием Байесовского метода. Последовательности генов GH1 и GH2 с высокой поддержкой объединяются в отдельные кластеры, что доказывает давнюю независимую их эволюцию и наличие двух генов гормона роста уже у предковых форм лососевых рыб. Топологии двух клад в целом согласуются друг с другом, хотя по некоторым ветвям обнаруживаются различия. Последовательности генов видов родов *Salvelinus* и *Oncorhynchus* формируют отдельные субкластеры, в то время как последовательности генов видов рода *Salmo* не образуют такого субкластера. В случае филогении по гену GH1, *S. salar* действительно формирует базальный кластер по отношению к гольцам и тихоокеанским лососям. В то же время *S. trutta* с высокой поддержкой кластеризуется с видами рода *Oncorhynchus*, но не с близкородственным видом *S. salar*. В случае филогении по гену GH2, *S. trutta* также объединяется с тихоокеанскими лососями, а *S. salar* кластеризуется с гольцами рода *Salvelinus* (рис. 3). Данные по аминокислотным последовательностям подтверждают неопределенность в положении рода *Salmo* относительно двух других таксонов (рис. 4).

Наиболее вероятно, что это явление может быть обусловлено отбором. Присутствие отбора было проанализировано путём проверки нулевой гипотезы (Но: $dN = dS$) против трёх альтернативных гипотез (H1): 1) тест на наличие отбора ($dN \neq dS$), 2) положительный отбор ($dN > dS$), 3) отрицательный отбор ($dN < dS$), используя Z-тест. Определение вида отбора произведено с помощью дифференции несинонимичной и синонимичной дистанций ($Dd = dN - dS$).

Для генов GH1 и GH2 получены отрицательные значения дифференции дистанций ($dN - dS$) (табл. 2), что позволяет предположить очищающий отбор. Для определения статистической достоверности этого предположения был проведён Z-тест. В тесте на наличие отбора ($dN \neq dS$) и тесте на очищающий отбор ($dN < dS$)

получены значения вероятности (p) меньше 0,05 (табл. 2). Следовательно, нулевая гипотеза ($dN = dS$) отклоняется в пользу альтернативной гипотезы отрицательного отбора. Таким образом, эволюция GH1 и GH2 происходила под воздействием очищающего отбора.

Таблица 2

Средние значения синонимичных и несинонимичных дистанций, дифференций дистанций, вероятностей отклонения нулевой гипотезы в пользу альтернативных (p) по результатам Z-теста для генов GH1 и GH2 лососёвых.

	GH1	GH2
dN	0,0158 ± 0,0036	0,0095 ± 0,0027
dS	0,0434 ± 0,0105	0,0900 ± 0,0161
Dd (dN – dS)	-0,0276 ± 0,0109	-0,0804 ± 0,0161
p (dN ≠ dS)	0,0201	0,0000027
p (dN > dS)	1,0000	1,0000
p (dN < dS)	0,0073	0,0000016

Примечание. Значения $p < 0,05$ считаются значимыми.

Средняя величина ω , оцененная методом SLAC, составила 0,255 для GH1 и 0,149 для GH2. Ни один из использованных методов (SLAC, FEL, REL) не выявил сайтов, находящихся под положительным отбором. В гене GH1 количество кодонов, находящихся под действием отрицательного отбора, составило 4 и 12 для SLAC и FEL соответственно. Метод REL все кодоны распознал как находящиеся под очищающим отбором. Четыре сайта (16-Phe, 18-Ser, 95-His, 121-Asn) идентифицированные SLAC, также выявлены методом FEL. Это говорит о том, что эти сайты находятся под очищающим отбором с очень большой вероятностью.

В гене GH2 количество кодонов, находящихся под действием очищающего отбора, составило 1, 17 и 30 для SLAC, FEL и REL соответственно. Один сайт (179-Glu), обнаруженный методом SLAC, также выявлен двумя другими методами. При сравнении результатов НуPhy анализа двух генов GH обнаружено, что четыре кодона (16-Phe, 95-His, 170-Gly, 178-Tyr), находящиеся под действием отрицательного отбора в гене GH1, также выявлены в гене GH2.

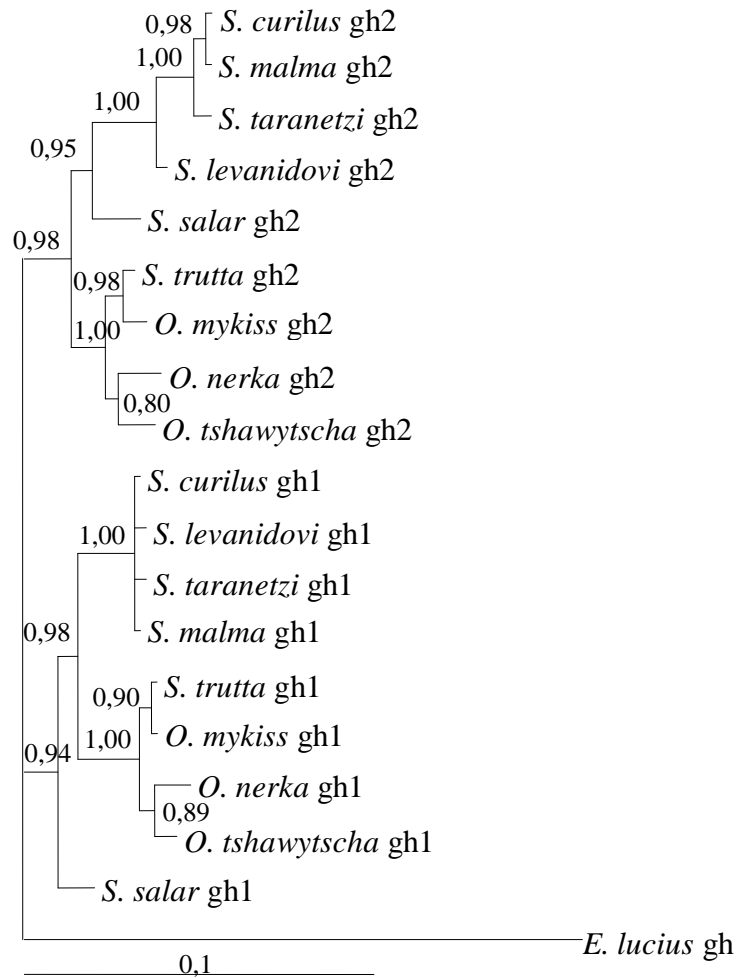


Рис. 3. Филогенетическое дерево, реконструированное по данным экзонов генов GH1 и GH2 с помощью байесовского моделирования.

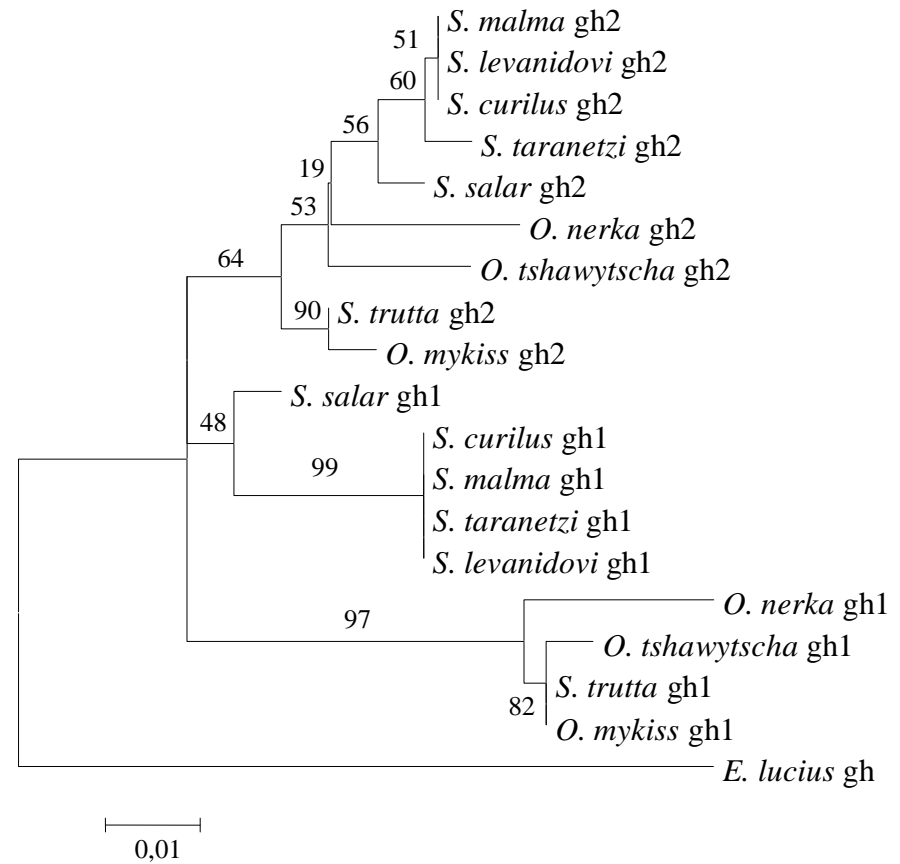


Рис. 4. Филогенетическое дерево, реконструированное на основе аминокислотных последовательностей GH1 и GH2 методом максимального правдоподобия.

Дивергенция некодирующих последовательностей генов GH

Средняя величина различий между интронами паралогичных генов оказалась в 2 раза больше, чем между экзонами и составила в среднем 9,6% нуклеотидных замен для всех видов, а внутри видов рода *Salvelinus* – 9,0%. Филогенетическое дерево, построенное с использованием Байесовского метода, приведено на рис. 5. Аналогично тому, что наблюдается по экзонам (рис. 3), данные подтверждают давнюю независимую эволюцию генов гормона роста. Интроны каждого из генов объединяются в отдельные кластеры, и с высокой поддержкой отличаются друг от друга. Из данных также видно, что интроны видов рода *Salvelinus* и *Oncorhynchus* формируют отдельные субкластеры внутри кластеров генов-паралогов.

Для выяснения возможных закономерностей эволюции генов GH в большинстве филогенетических линиях семейства Salmonidae были использованы последовательности интронов C и D, которые доступны в базе данных GenBank/NCBI для многих видов лососевых рыб. Последовательности интронов C и D объединяли при анализе каждого гена. На рис. 6 приведено филогенетическое древо интронов, полученное с использованием VI-метода построения деревьев. Также как и в предыдущих случаях, интронные участки формируют два независимых кластера, соответствующих двум генам-паралогам. Отчетливо формируются в случае интронов обоих генов субкластеры, объединяющие виды *Oncorhynchus*, *Salvelinus*, *Salmo*. В то же время, очевидно, что положение некоторых видов по разным генам в филогенетической схеме отличается. В частности, наиболее существенные различия касаются положения сахалинского тайменя *Parahucho perryi* и гольца *S. alpinus*. В случае гена GH1, *S. alpinus* объединяется с *S. namaycush*, а гена GH2 – с двумя видами мальмы. *P. perryi* объединяется с *H. hucho* в случае гена GH1, а в случае гена GH2 кластеризуется с гольцами (рис. 6).

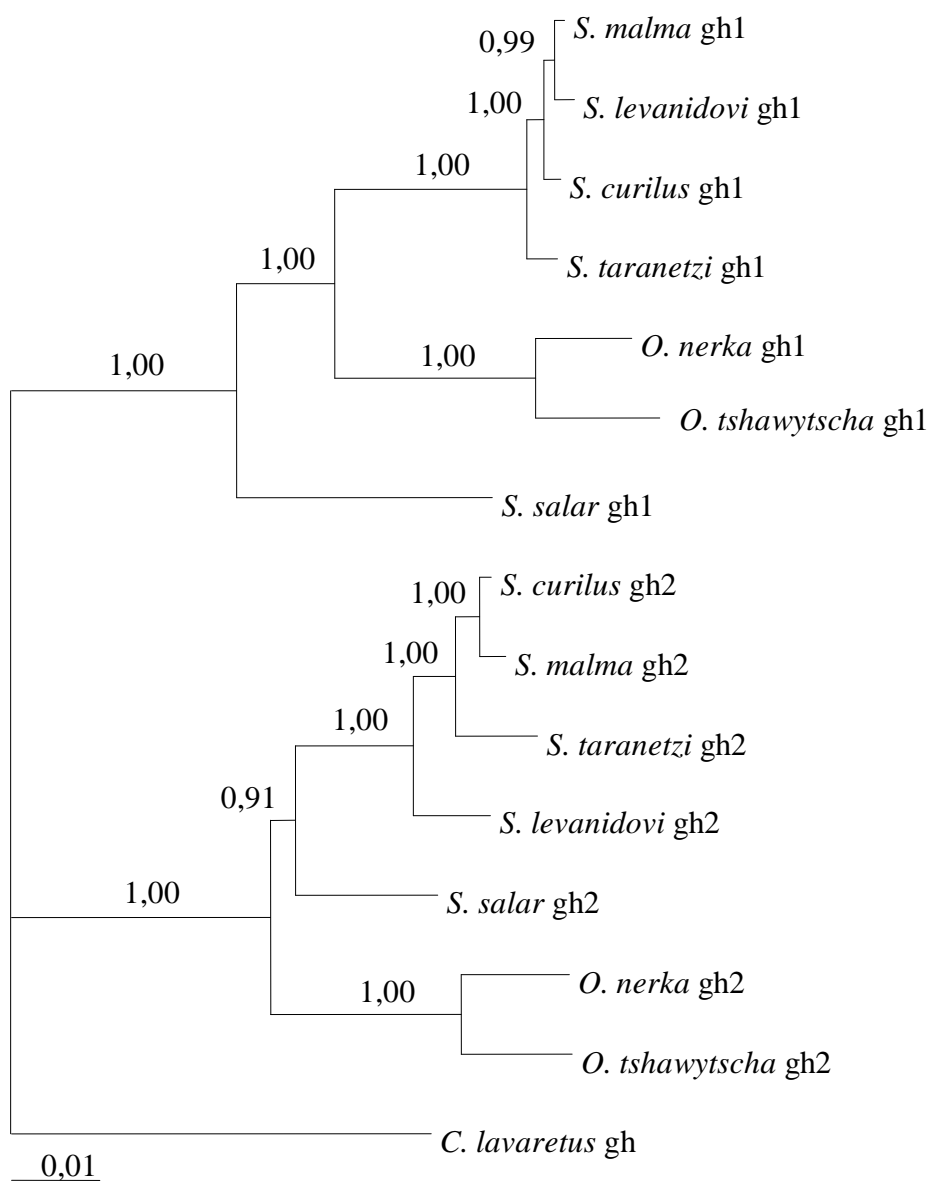


Рис. 5. Филогенетическое дерево, реконструированное по данным интронов генов GH1 и GH2 с помощью байесовского моделирования.

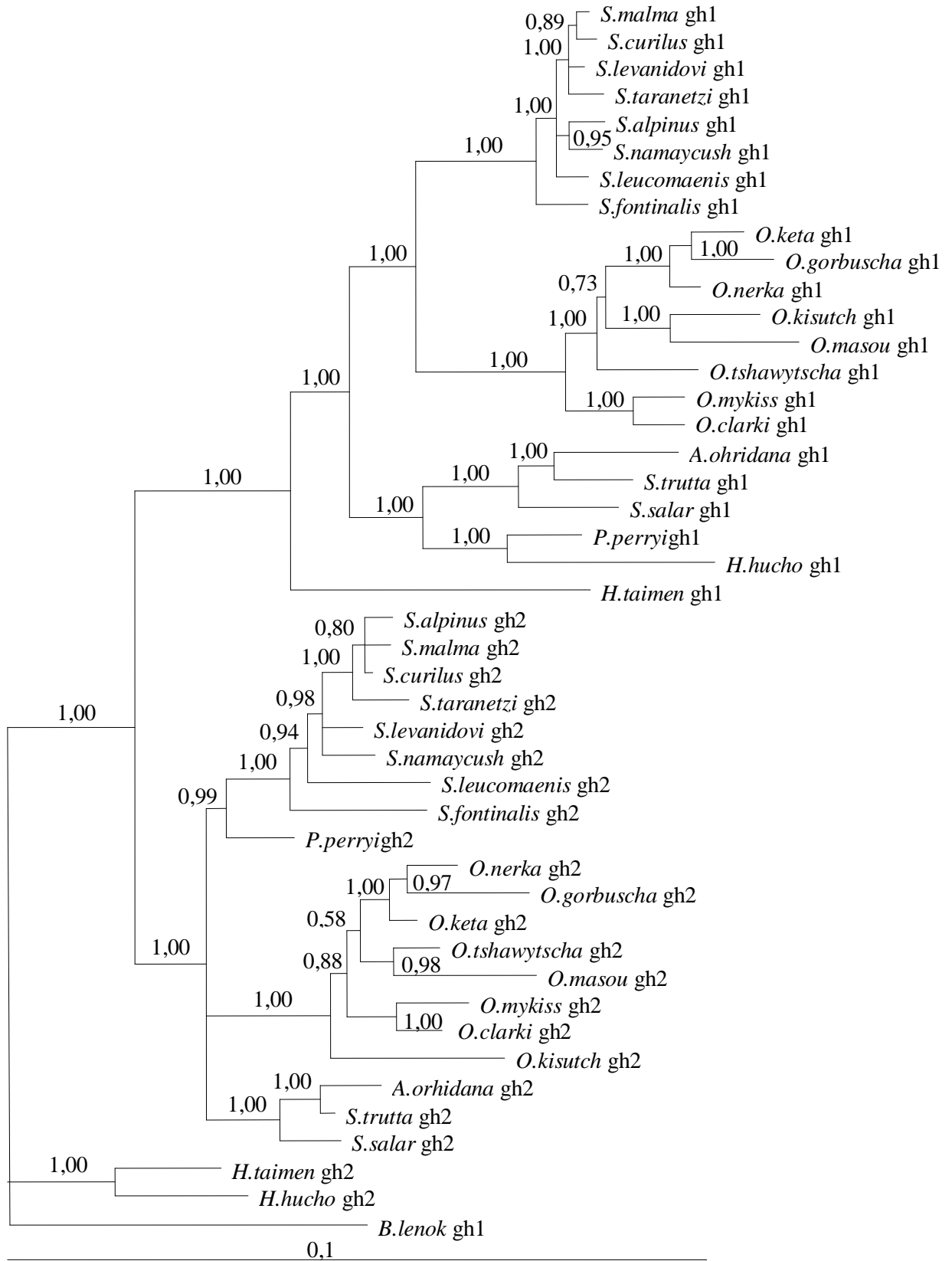


Рис. 6. Филогенетическое дерево, реконструированное по данным интронов С и D двух генов GH лососевых рыб с помощью байесовского моделирования.

ОБСУЖДЕНИЕ

Показано, что все четыре вида рода *Salvelinus* содержат два гена GH, как и другие исследованные виды семейства Salmonidae. Структура генов GH гольцов сходна со структурой генов GH других видов лососёвых рыб. Этого следовало ожидать, исходя из важной функциональной роли гормона роста в функционировании организмов рыб. Полученные в работе данные дают основание считать, что оба гена, по-видимому, функциональны. В пользу этого говорят следующие факты: 1. Оба паралога консервативны в кладе гольцов; 2. Последовательности экзонов паралогов у всех видов содержат открытую рамку считывания общей длиной 630 п.н.; 3. Предсказанная аминокислотная последовательность высоко консервативна, включая положения цистеиновых остатков; 4. Последовательности обоих генов GH сходны с теми, которые выявлены у *S. salar* и *O. tshawytscha*, у которых оба гена функциональны (Von Schalburg et al., 2008); 5. Последовательности обоих генов находятся под влиянием очищающего отбора.

Полученные данные по объединению кодирующих последовательностей генов GH1 и GH2 в отдельные кластеры доказывают давнюю независимую их эволюцию и наличие двух генов гормона роста уже у предковых форм лососевых рыб. В целом филогения родов *Oncorhynchus* и *Salvelinus*, полученная на основе последовательностей обоих генов GH, достаточно хорошо согласуется с филогенией этих таксонов, основанных на других подходах и признаках. В то же время положение видов рода *Salmo* на филограмме с высокой поддержкой не всегда конгруэнтно с имеющимися схемами дивергенции видов в этой группе (рис. 3). Наиболее вероятным следует полагать, что это явление может быть обусловлено отбором. При этом отбор, по-видимому, был разнонаправлен. Поскольку тесты не выявляют каких-либо сайтов, находящихся под влиянием положительного отбора, необходимо предположить, что это обусловлено разным давлением очищающего отбора на гены-паралоги.

Более убедительными и доказательными данными являются различия GH1 и GH2 по скорости аминокислотных замен. Скорость аминокислотных замен для всех видов, исходя из максимального времени начальной дивергенции

паралогичных генов (100 млн. лет), составила $0,182 \times 10^{-9}$ для GH1 и $0,109 \times 10^{-9}$ для GH2 на сайт в год. Если рассчитывать отдельно для каждого таксона, то выявляются различия. В GH1 у гольцов вообще не обнаружено аминокислотных замен, в GH2 хотя и обнаруживаются аминокислотные замены, но скорость низкая – $0,0238 \times 10^{-9}$ на сайт в год. В GH2 у лососей рода *Oncorhynchus* скорость накопления аминокислотных замен ($0,158 \times 10^{-9}$) выше в 2 раза по сравнению с GH1 ($0,079 \times 10^{-9}$). В то же время у представителей рода *Salmo* обнаруживается обратная картина: скорость накопления аминокислотных замен в GH1 ($0,214 \times 10^{-9}$) превышает скорость в GH2 ($0,095 \times 10^{-9}$) более чем в 2 раза. Очевидно, что в одном таксоне очищающий отбор более интенсивный по одному гену-паралогу, а в остальных таксонах – по другому. Таким образом, наши данные дают основания полагать, что скорость эволюции паралогичных генов GH у лососевых рыб может сильно отличаться как внутри одного таксона, так и между таксонами.

Это также следует при сравнении изменчивости в последовательностях экзонов и интронов генов GH. Известно, что интронные участки генов дивергируют с большей скоростью по сравнению с экзонными, поскольку находятся под меньшим давлением отбора. В нашем случае средняя величина различий между экзонами гена GH1 у всех видов гольцов составила 0,16% нуклеотидных замен, а между интронами этого же гена 0,49%, и различия значимы ($p < 0,05$). В гене GH2 среднее различие по экзонам составило 0,74%, а по интронам – 1,32% ($p < 0,01$).

Из полученных данных следует, что, как и в случае с экзонами, в интронах генов GH2 гольцов скорость накопления нуклеотидных замен выше, чем в интронах генов GH1. Это может означать, что интроны паралогов, по-видимому, тоже находятся под разной степенью очищающего отбора. Однако при сравнении генетических дистанций, рассчитанных на основе последовательностей интронов C и D для большинства видов лососёвых рыб (исключив виды рода *Salvelinus*), более консервативными оказались последовательности интронов гена GH2, а не GH1, как в случае с гольцами. Средняя генетическая дистанция по интронам GH1 составила 6,77%, а по GH2 – 5,14%, и различия высоко значимы ($p \ll 0,01$). Таким образом,

можно полагать, что у большинства видов лососевых рыб интроны гена GH2 находятся под большим влиянием отбора.

Наличие отбора по нуклеотидным последовательностям предполагает, что интронные последовательности несут функциональную нагрузку. То, что интронные участки в генах гормона роста могут нести функциональную нагрузку и находится под действием отбора, следует из наличия в интронах С и D генов GH лососёвых регуляторных элементов (рис. 7). В четвёртом интроне (D) лососёвых рыб обнаружен CRE-элемент (Von Schalburg et al., 2008), ответственный за связывание с цАМФ-зависимым транскрипционным фактором (CREB) (Montminy, 1990). В третьем интроне (C) обнаружены консенсусные последовательности сайтов связывания гипофиз-специфичного транскрипционного фактора Pit-1, входящие в состав двух АТ-богатых областей (Bernardini et al., 1999), а также эстроген-чувствительный элемент (ERE) (Driscoll et al., 1998, Phillips et al., 2004) (рис. 7).

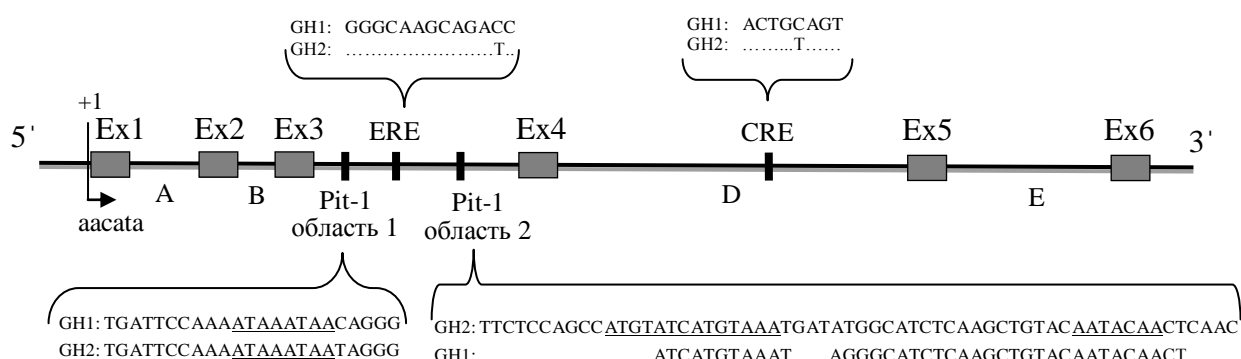


Рис. 7. Схема расположения регуляторных элементов в интронах генов GH Salmonidae. Последовательности элементов указаны для *S. levanidovi*. Сайты связывания Pit-1 подчеркнуты.

Дополнительными факторами, способными оказывать влияние на консервативность и различия в скорости дивергенции интронных последовательностей в паралогичных генах, являются эффекты хичхайкинга и/или эффект Хилла-Робертсона (Fay, Wu, 2000; Comeron, Kreitman, 2002). Представляется, что отрицательный отбор по экзонам не только ограничивает

возможность изменений даже по синонимичным положениям в самих экзонах, но отбор интерферирует на прилежащие некодирующие интронные участки (Comeron, Kreitman, 2002). В нашем случае при сравнении дивергенции в генах-паралогах видно, что высокая консервативность экзонных участков в GN1 определяет и большую консервативность интронных участков в этом гене. Снижение давления отбора по экзонным последовательностям GN2 снижает давление отбора и по интронам.

ЗАКЛЮЧЕНИЕ

Проведенные исследования позволили выявить закономерности организации и эволюции паралогичных генов гормона роста у лососевых рыб. Сравнительный анализ структуры паралогичных генов гормона роста у всех видов лососевых рыб, включая гольцов, показал высокое сходство и консервативность. Экзонные последовательности генов оказались весьма консервативными и находятся под влиянием очищающего отбора. Эти факты дают основание считать, что оба гена-паралога функциональны и, возможно, характеризуются тканеспецифичной экспрессией.

Филогенетический анализ экзонных и интронных последовательностей показал, что скорость дивергенции паралогичных генов гормона роста различается в разных филетических линиях у лососевых рыб и это определяется разным давлением отрицательного отбора.

Сравнительный анализ последовательностей генов GN показал, что интронные участки дивергируют с более высокой скоростью, примерно в 2 раза, по сравнению с экзонами. В то же время, очевидно, что и интронные последовательности находятся под влиянием отрицательного отбора. Это может определяться как эффектом хичхайкинга и Хилла-Робертсона, так и тем, что в интронах обнаруживаются потенциальные регуляторные последовательности.

ВЫВОДЫ

1. Ген гормона роста гольцов рода *Salvelinus*, как у других видов лососёвых рыб, представлен двумя несвязанными паралогичными генами, GH1 и GH2.
2. Структура генов GH1 и GH2 гольцов сходна и включает 6 экзонов и 5 интронов. Размер каждого экзона двух генов идентичен у всех видов гольцов. Длина генов варьирует за счёт длины интронов.
3. Филогенетический анализ экзонных и интронных последовательностей генов GH1 и GH2 свидетельствует о давней дупликации гена гормона роста и парафилетичности двух генов в семействе Salmonidae.
4. Высокая гомология последовательностей экзонов и тест на отбор подтверждают эффект отрицательного отбора по кодирующим последовательностям генов GH у всех видов гольцов. Последовательности экзонов гена GH2 гольцов накапливает нуклеотидные замещения с большей скоростью, чем экзоны гена GH1 и, следовательно, находятся под меньшим влиянием отбора.
5. Открытая рамка считывания, высокая консервативность предсказанной аминокислотной последовательности, идентичное положение цистеиновых остатков и наличие очищающего отбора дают основание считать, что оба гена-паралога функциональны.
6. Интронные последовательности в генах гормона роста также находятся под действием отрицательного отбора, вследствие наличия в них потенциальных регуляторных элементов, а также, вероятно, эффектов хичхайкинга и Хилла-Робертсона.
7. Неконгруэнтность топологий филогенетических отношений и разная скорость дивергенции как экзонов, так и интронных последовательностей генов-паралогов свидетельствуют о влиянии на них отрицательного и отличающегося по силе отбора в разных филетических линиях.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в журналах из списка, рекомендованного ВАК:

1. **Панькова М.В.**, Брыков В.А., Панькова В.В., Атопкин Д.М. Гены гормона роста. Дивергенция последовательностей интронов у гольцов рода *Salvelinus* // Генетика. 2013. Т. 49, № 6. С. 743–750.

2. **Панькова М.В.**, Брыков В.А. Дивергенция интронов в паралогичных генах гормона роста у лососёвых рыб выявляет эффект отбора // Доклады Академии наук. 2013. Т. 451, № 3. С. 351–354.

3. Каменская Д.Н., **Панькова М.В.**, Атопкин Д.М., Брыков В.А. Гены гормона роста у рыб: доказательства функциональности паралогичных генов у гольца *Salvelinus levanidovi* // Молекулярная биология. 2015. Т. 49, № 5. С. 770–776.

Работы в сборниках трудов и материалах конференций:

4. **Панькова М.В.** Сравнение дивергенции интронов и экзонов гена гормона роста у лососевидных рыб // X Региональная конференция студентов, аспирантов вузов и научных организаций Дальнего Востока России, 4–6 мая 2011 г., Владивосток. – Владивосток: Изд-во Дальневосточного университета, 2011. С. 214–215.

5. **Панькова М.В.** Эволюция генов гормона роста у гольцов рода *Salvelinus* (Pisces) // XI Региональная конференция студентов, аспирантов вузов и научных организаций Дальнего Востока России, 3–4 мая 2012 г., Владивосток. – Владивосток: Изд-во Дальневосточного университета, 2012. С. 222–224.

6. Brykov V.A., **Pankova M.V.** Structure and evolution of the growth hormone genes of the chars (genus *Salvelinus*) // 7th International Charr Symposium, 3–6 September 2012, Yuzhno-Sakhalinsk, Russia: Program and Abstracts of Presentations. – Yuzhno-Sakhalinsk, 2012. P. 51.

7. Брыков В.А., **Панькова М.В.** Эволюция кодирующих и не кодирующих последовательностей в генах гомона роста лососевых рыб // Современные проблемы биологической эволюции: материалы II Международной конференции, 11–14 марта 2014 г., Москва. – Москва: ГДМ, 2014. С. 117–119.

Марина Владимировна Панькова

Структура и эволюция генов гормона роста лососёвых рыб (Salmonidae)

03.02.07 – генетика

Автореферат

диссертации на соискание ученой степени

кандидата биологических наук