

ГЕНОМИКА.
ТРАНСКРИПТОМИКА

УДК 575.832:597.553.2

ИЗМЕНЧИВОСТЬ ЭКЗОНОВ И ИНТРОНОВ
В ГЕНАХ ГОРМОНА РОСТА У ЛОСОСЕВЫХ РЫБ

© 2020 г. Д. Н. Каменская^{а, *}, М. В. Панькова^б, В. А. Брыков^{а, б}

^аНациональный научный центр морской биологии им. А. В. Жирмунского Дальневосточного отделения
Российской академии наук, Владивосток, 690041 Россия

^бДальневосточный федеральный университет, Школа естественных наук, Владивосток, 690012 Россия

*e-mail: kamenskaya.daria@mail.ru

Поступила в редакцию 16.03.2020 г.

После доработки 22.05.2020 г.

Принята к публикации 26.05.2020 г.

Дубликации генов — один из основных механизмов формирования нового генетического материала в процессе эволюции. Как правило, одна из копий дублированного гена оказывается под меньшим давлением очищающего отбора, накапливает мутации с более высокой скоростью, что со временем может привести к появлению у дублированного гена новой функции. В результате нескольких раундов полиплоидизации многие гены у лососевых рыб оказались множественными, в том числе и ген гормона роста. В работе проведен анализ нуклеотидного разнообразия в паралогичных генах гормона роста — *gh1* и *gh2*. Показано, что уровень изменчивости в интронах выше, чем в экзонах. Изменчивость в каждом из экзонов слабо связана с их длиной и может определяться функциональной значимостью кодируемой ими белковой части. Выявлен относительно высокий уровень изменчивости в экзонах гена *gh2* по сравнению с *gh1*, что, вероятно, связано с текущим процессом субфункционализации генов.

Ключевые слова: гены гормона роста, полиморфизм генов, дубликация генов, отрицательный отбор, *Salvelinus*, Salmonidae

DOI: 10.31857/S0026898420060051

ВВЕДЕНИЕ

Накопление мутаций в различных участках генома эукариот в процессе эволюции неравномерно и зависит от многих факторов. Белоккодирующие последовательности и регуляторные элементы находятся, как правило, под жестким давлением отрицательного отбора. Некодирующие последовательности накапливают мутации с более высокой частотой.

Гормон роста относится к наиболее изученным полипептидным гормонам. Он выполняет ряд жизненно важных функций: отвечает за регуляцию соматического роста и участвует в многочисленных физиологических процессах, включая ионный баланс, липидный и белковый обмен, иммунный ответ, процесс размножения, а также различные аспекты поведения [1, 2]. Белковая последовательность гормона роста позвоночных достаточно консервативна, а кодирующие его гены хорошо изучены во многих таксономических группах и обычно включают 5 экзонов и 4 интрона [3–5]. У большинства позвоночных, в том числе и у некоторых видов рыб (Cypriniformes, Siluriformes), ген гормона роста представлен единствен-

ной копией [6, 7]. Гены гормона роста других рыб, в том числе и лососевых, состоят из 6 экзонов и 5 интронов [8–11]. У лососевых рыб ген гормона роста представлен двумя копиями: *gh1* и *gh2*, — что, вероятно, свидетельствует о его дубликации у предковой формы [12]. Структура двух копий гена гормона роста, *gh1* и *gh2*, лососевых сходна. Различия в размере этого гена, как у разных видов, так и между двумя копиями гена одного вида, возникает за счет разной длины интронов [13, 14].

Мы сравнили нуклеотидное разнообразие в каждом экзоне и интроне генов *gh1* и *gh2* лососевых для выяснения возможных закономерностей эволюции как самого гена, так и особенностей (потенциальных путей) дивергенции в случае появления дублированных копий.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Полноразмерные нуклеотидные последовательности генов *gh1* и *gh2* четырех видов тихоокеанских гольцов рода *Salvelinus*, использованные в анализе, получены и подробно описаны нами ранее [15–17]. Размещены последовательности в базе данных GeneBank/NCBI под следующими

номерами: *Salvelinus malma gh1* – KF772972.2, *gh2* – KF772976.2; *Salvelinus curilus gh1* – KF772971.2, *gh2* – KF772975.2; *Salvelinus levanidovi gh1* – KF772974.2, *gh2* – KF772978.2; *Salvelinus taranetzki gh1* – KF772973.2, *gh2* – KF772977.2. Последовательности генов гормона роста других видов из семейства Salmonidae взяты из базы данных GeneBank/NCBI: *Salmo salar* (атлантический лосось) *gh1* – EU621898.1, *gh2* – EU621899.1; *Oncorhynchus nerka* (нерка) *gh1* – U14551.2, *gh2* – U14535.2; *Oncorhynchus tshawytscha* (чавыча) *gh1* – EU621900.1, *gh2* – EU621901.1. Содержание GC-пар в интронах оценивали с помощью программы MEGA 6.0 [18]. Анализ нуклеотидного разнообразия был выполнен в программе DnaSP 5.10.01 [19].

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Полученные данные по нуклеотидному разнообразию для каждого экзона и интрона паралогичных генов гормона роста и их длины приведены в табл. 1 и 2. Как мы и ожидали, короткие экзоны оказались менее вариабельными. Так, в случае гена *gh1* первый (74 п.н.) и шестой (63 п.н.) экзоны практически не содержали нуклеотидных замен. В остальных экзонах гена *gh1*: втором (140 п.н.), третьем (117 п.н.), четвертом (156 п.н.) и пятом (147 п.н.) – нуклеотидное разнообразие было выше, но при этом примерно одинаковое – от 0.018 до 0.029 (табл. 1).

Иная картина наблюдается в случае паралогичного гена *gh2*. Наименьшая изменчивость отмечается, как и в случае гена *gh1*, для шестого экзона, немного более изменчив третий экзон. Остальные экзоны гена *gh2*, в том числе и первый, характеризуются в 3–4 раза более высоким уровнем изменчивости.

На основании анализа нуклеотидных последовательностей интронов паралогов *gh1* и *gh2* можно сделать вывод, что уровень их изменчивости не зависит от длины (табл. 2).

Консервативность первого и шестого экзонов в генах гормона роста лососевых рыб может определяться не длиной, а функциональной значимостью. Первый экзон этого гена кодирует нетранслируемую 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 четвертого кодона сигнального пептида, отщепляемого после трансляции при созревании функционального белка. Высокая консервативность этого экзона очевидна, так как лидерная последовательность включает в себя важные функциональные элементы, необходимые для эффективной инициации трансляции: участки внутренней посадки рибосомы, внутренние открытые рамки считывания, железозависимые элементы и др. [20]. Консервативность шестого экзона определяется, вероятно, участием кодируемого им фрагмента в правильной укладке полипептидной цепи после трансляции. Считается, что для этого процесса важно наличие двух остатков цистеина в этой области белка [21].

Относительная консервативность остальных экзонов в генах гормона роста лососевых определяется тем, что значительная часть, кодируемой ими аминокислотной последовательности участвует в формировании вторичной структуры – образовании α -спиралей. В тоже время допускаются синонимичные замены нуклеотидов, не приводящие к заменам аминокислот.

Относительно высокий уровень изменчивости экзонов гена *gh2* у лососевых рыб по сравнению с геном *gh1* может быть обусловлен снижением давления очищающего отбора, допускающего субфункционализацию генов в эволюции.

Уровень изменчивости в том или ином участке гена напрямую связан со скоростью накопления мутаций, а именно со скоростью дивергенции в процессе эволюции. Особенно это касается некодирующих последовательностей, таких как интроны, где давление очищающего отбора снижено.

Таблица 1. Нуклеотидное разнообразие в экзонах генов *gh1* и *gh2* лососевых рыб

Экзон (номер)	Длина, п.н.	Нуклеотидное разнообразие (π) ^a	
		<i>gh1</i>	<i>gh2</i>
1	74	0.00386 ± 0.00265	0.03501 ± 0.00633
2	140	0.02585 ± 0.00745	0.03946 ± 0.00851
3	117	0.02768 ± 0.00599	0.01099 ± 0.00236
4	156	0.02930 ± 0.00642	0.02869 ± 0.00563
5	147	0.01846 ± 0.00507	0.03693 ± 0.00654
6	63	0	0.00756 ± 0.00272

^a Результаты представлены как среднее значение ± стандартное отклонение на основании независимого сравнения семи нуклеотидных последовательностей по каждому экзону.

Таблица 2. Нуклеотидное разнообразие в интронах генов *gh1* и *gh2* лососевых рыб^a

Интрон	Salmonidae, вид	<i>gh1</i>		<i>gh2</i>	
		Длина, п.н.	Нуклеотидное разнообразие (π)	Длина, п.н.	Нуклеотидное разнообразие (π)
A	<i>O. nerka</i>	392	0.04699 ± 0.00934	470	0.05153 ± 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. taranetz</i>	455		465	
B	<i>O. nerka</i>	136	0.02963 ± 0.00646	138	0.05982 ± 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. taranetz</i>	136		123	
C	<i>O. nerka</i>	710	0.04244 ± 0.01050	445	0.02835 ± 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. taranetz</i>	721		623	
D	<i>O. nerka</i>	1145	0.04190 ± 0.00976	964	0.03607 ± 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. taranetz</i>	1044		1171	
E	<i>O. nerka</i>	586	0.05007 ± 0.01030	220	0.04321 ± 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. taranetz</i>	622		226	

^a Результаты представлены как среднее значение ± стандартное отклонение на основании независимого сравнения семи нуклеотидных последовательностей по каждому интрону.

Имеющиеся в литературе данные по изменчивости и дивергенции интронов противоречивы. Gazave с соавт. [22] сравнили интронные участки многих ортологичных генов человека и шимпанзе и нашли строгую положительную корреляцию между длиной интронов и их дивергенцией, а также строгую отрицательную корреляцию между длиной интронов и содержанием GC-пар. Кроме того, авторы утверждают, что скорость дивергенции зависит от порядкового номера интрона в гене.

Haddrill с соавт. [23] сравнили 225 интронных участков в геномах *Drosophila melanogaster* и *D. simulans* и выявили строгую отрицательную корреляцию между длиной интронов и их дивергенцией и между дивергенцией интронов и содержанием GC-пар.

В нашем случае последовательности интронов, как и ожидалось, оказались более изменчивы, чем экзоны. Однако в генах гормона роста лососевых практически не наблюдается связи между уровнем изменчивости интронов и их дли-

Таблица 3. Содержание GC-пар в интронах генов *gh1* и *gh2* лососевых рыб

Интрон	G + C, %	
	<i>gh1</i>	<i>gh2</i>
A	38.4	37.6
B	36.9	39.2
C	39.6	38.5
D	37.2	36.3
E	38.1	35.5

ной. Лишь в гене *gh1*, в самом коротком интроне В, эта величина примерно в 1.5 раза ниже, чем в других интронах (табл. 2). В гене *gh2* слабо изменчивым оказался интрон С – не самый короткий из интронов. Ранее мы обнаружили вероятные регуляторные последовательности в интронах С и D, и эти элементы действительно консервативны [18]. Тем не менее консервативность этих коротких участков не сказывается на общем уровне изменчивости. Также мы не обнаружили связи между содержанием GC-пар и уровнем изменчивости, но это может быть связано со сходным содержанием GC-пар во всех интронах (табл. 3).

Ранее Comerop и Kreitman предполагали [24], что степень очищающего отбора по экзонам коррелирует с изменчивостью прилежащих к ним интронных участков по механизму интерферирующего отбора. Полученные нами данные не подтверждают этой гипотезы (табл. 1, 2).

Таким образом, нами установлено, что изменчивость экзонов генов гормона роста, *gh1* и *gh2*, лососевых рыб не зависит от длины последовательности и в основном определяется функциональной значимостью кодируемого ими фрагмента белка. Уровень изменчивости экзонов невысокий и сходен в этих генах-паралогах, что свидетельствует об их функциональной значимости. Изменчивость в интронах этих генов оказалась выше, чем в экзонах, но, как и для экзонов, не связана с длиной. Не выявлено корреляции между нуклеотидным разнообразием в интронах и прилегающих экзонах.

Написание статьи не потребовало специального финансирования.

Статья не содержит каких-либо исследований с участием людей или животных в качестве объектов исследований.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Bolton J.P., Collie N.L., Kawachi H., Hirano T. (1987) Osmoregulatory actions of growth hormone in rainbow trout (*Salmo gairdneri*). *J. Endocrinol.* **112**, 63–68.

2. Coker A., Arman A. (2009) Human growth hormone. *Adv. Mol. Biol.* **3**, 9–16.
3. Buggiotti L., Hellstrom M.A., Primmer C.R. (2006) Characterization of the first growth hormone gene sequence for a passerine bird – the pied flycatcher (*Ficedula hypoleuca*). *DNA Seq.* **17**, 401–406.
4. Das P., Meyer L., Seyfert H.M., Brockmann G., Schwerin M. (1996) Structure of the growth hormone encoding gene and its promoter in mice. *Gene.* **169**, 209–213.
5. Gordon D.F., Quick D.P., Erwin C.R., Donelson J.E., Maurer R.A. (1983) Nucleotide sequence of the bovine growth hormone chromosomal gene. *Mol. Cell. Endocrinol.* **33**, 81–95.
6. Panicz R., Sadowski J., Drozd R. (2012) Genetic and structural characterization of the growth hormone gene and protein from tench, *Tinca tinca*. *Fish Physiol. Biochem.* **38**, 1645–1653.
7. Sekar M., Singh S.D., Gupta S. (2014) Cloning and characterization of *Pangasianodon hypophthalmus* growth hormone gene and its heterologous expression. *Appl. Biochem. Biotechnol.* **173**, 1446–1468.
8. Chen Y., Wang Y., He S., Zhu Z. (2004) Cloning and sequencing of the growth hormone gene of large yellow croaker and its phylogenetic significance. *Biochem. Genet.* **42**, 365–375.
9. Tanaka M., Toma Y., Ohkubo T., Sudo S., Nakashima K. (1995) Sequence of the flounder (*Paralichthys olivaceus*) growth hormone-encoding gene and its promoter region. *Gene.* **165**, 321–322.
10. Venkatesh B., Brenner S. (1997) Genomic structure and sequence of the pufferfish (*Fugu rubripes*) growth hormone-encoding gene: a comparative analysis of teleost growth hormone genes. *Gene.* **187**, 211–215.
11. Male R., Nerland A.N., Lorens J.B., Telle W., Lossius I., Totland G.K. (1992) The complete nucleotide sequence of the Atlantic salmon growth hormone I gene. *Biochim. Biophys. Acta.* **1130**, 345–348.
12. McKay S.J., Trautner J., Smith M.J., Koop B.F., Delvin R.H. (2004) Evolution of duplicated growth hormone genes in autotetraploid salmonid fishes. *Genome.* **47**, 714–723.
13. Devlin R.H. (1993) Sequence of sockeye salmon type 1 and 2 growth hormone genes and the relationship of rainbow trout with Atlantic and Pacific salmon. *Can. J. Fish. Aquat. Sci.* **50**, 1738–1748.
14. Du S.J., Devlin R.H., Hew C.L. (1993) 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 Cell Biol.* **12**, 739–751.
15. Каменская Д.Н., Панькова М.В., Атопкин Д.М., Брыков В.А. (2015) Гены гормона роста у рыб: доказательства функциональности паралогичных генов у гольца *Salvelinus levanidovi*. *Молекуляр. биология.* **49**, 770–776.
16. Каменская Д.Н., Панькова М.В., Атопкин Д.М., Брыков В.А. (2017) Дивергенция паралогичных генов гормона роста и *цис*-регуляторных участков у лососевых рыб. *Молекуляр. биология.* **51**, 314–323.
17. Панькова М.В., Кухлевский А.Д., Брыков В.А. (2017). Гены гормона роста: дивергенция кодирую-

- ших последовательностей у лососевых рыб. *Генетика*. **53**, 201–213.
18. Tamura K., Stecher G., Peterson D., Filipski A., Kumar S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729.
 19. Librado P., Rozas J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. **25**, 1451–1452.
 20. Barrett L. W., Fletcher S., Wilton S. D. (2013) *Untranslated Gene Regions and Other Non-coding Elements. Regulation of Eukaryotic Gene Expression*. N.Y.: Springer.
 21. Vestling M., Murphy C., Fenselau C., Chen T.T. (1991) Disulfide bonds in native and recombinant fish growth hormones. *Mol. Mar. Biol. Biotechnol.* **1**, 73–77.
 22. Gazave E., Marqués-Bonet T., Fernando O., Charlesworth B., Navarro A. (2007) Patterns and rates of intron divergence between humans and chimpanzees. *Genome Biol.* **8**, 21.
 23. Haddrill P. R., Charlesworth B., Halligan D. L., Andolfatto P. (2005) Patterns of intron sequence evolution in *Drosophila* are dependent upon length and GC content. *Genome Biol.* **6**, 67.
 24. Comeron J.M., Kreitman M. (2002) Population, evolutionary and genomic consequences of interference selection. *Genetics*. **161**, 389–410.

EXON AND INTRON VARIABILITY IN SALMONIDAE GROWTH HORMONE GENES

D. N. Kamenskaya^{1,*}, M. V. Pankova², and V. A. Brykov^{1,2}

¹*Zhirmunsky National Scientific Center of Marine Biology, Far Eastern Branch, Russian Academy of Sciences, Vladivostok, 690041 Russia*

²*Far Eastern Federal University, School of Natural Sciences, Vladivostok, 690012 Russia*

*e-mail: kamenskaya.daria@mail.ru

A gene duplication is one of the main mechanisms of the formation of new genetic material in the process of evolution. The occurrence of a gene duplication is believed to relax selection pressure on one of the copies. Consequently, this gene accumulates mutations at a higher rate, and overtime it acquires a new function. As a result of several rounds of polyploidization, many genes in salmon fish turned out to be multiple, including the growth hormone gene. The analysis of nucleotide diversity in the paralogous genes of growth hormone, *gh1* and *gh2*, demonstrated that a level of variability in their introns was higher than in the exons. In addition, variability of each exon weakly correlated with its length, and seems to be determined by the functional significance of protein region encoded. The level of variability in the exons of *gh2* gene was higher than that in *gh1* one, which was probably due to the current process of gene subfunctionalization.

Keywords: growth hormone genes, nucleotide polymorphism, gene duplication, negative selection, *Salvelinus*, Salmonidae