



УДК 577.218

EDN IXRERA

<https://www.doi.org/10.33910/2687-1270-2022-3-2-185-203>

Организация некодирующих элементов в геномах птиц

А. Ф. Сайфитдинова^{✉1}

¹ Российский государственный педагогический университет им. А. И. Герцена,
191186, Россия, г. Санкт-Петербург, наб. р. Мойки, д. 48

Сведения об авторе

Алсу Фаритовна Сайфитдинова, SPIN-код: [5114-4844](#), Scopus AuthorID: [6603195889](#), ResearcherID: [C-1104-2011](#), ORCID: [0000-0002-1221-479X](#), e-mail: saitfdinova@mail.ru

Для цитирования: Сайфитдинова, А. Ф. (2022) Организация некодирующих элементов в геномах птиц. *Интегративная физиология*, т. 3, № 2, с. 185–203. <https://www.doi.org/10.33910/2687-1270-2022-3-2-185-203> EDN IXRERA

Получена 22 марта 2022; прошла рецензирование 4 мая 2022; принята 12 мая 2022.

Финансирование: Исследование не имело финансовой поддержки.

Права: © А. Ф. Сайфитдинова (2022). Опубликовано Российским государственным педагогическим университетом им. А. И. Герцена. Открытый доступ на условиях [лицензии CC BY-NC 4.0](#).

Аннотация. Обзор посвящен характеристике некодирующих элементов в геномах птиц. Как и большинство живых организмов, птицы в своем геноме несут значительное количество некодирующих последовательностей, которые, по формальным причинам, не могут быть отнесены к категории генов. Традиционно к таким элементам принято относить участки ДНК, выполняющие структурные функции, которые отличаются высоким уровнем консерватизма, а также значительное количество последовательностей, в том числе повторяющихся элементов, функции которых до сих пор не выяснены. В контексте общей характеристики тандемных повторяющихся элементов в геномах птиц особое внимание уделено различиям в количестве копий отдельных повторяющихся элементов. Интересно, что это непосредственно коррелирует с сокращением размеров геномов у птиц. В обзоре подробно описано участие повторяющихся последовательностей в структурной и функциональной организации центромерных и теломерных районов с акцентом на особенности организации перичентромерного и субтеломерного гетерохроматина и их отличия от собственно центромерных и теломерных последовательностей. Данные о распространенности в геномах птиц различных повторов и особенностях их организации соотнесены с уровнем их консерватизма и ролью в поддержании структурной и функциональной организации генома.

Ключевые слова: организация генома, некодирующие ДНК, повторяющиеся последовательности, гетерохроматин, повторы высокого порядка

Organisation of non-coding elements in avian genomes

A. F. Saifitdinova^{✉1}

¹ Herzen State Pedagogical University of Russia, 48 Moika Emb., Saint Petersburg 191186, Russia

Author

Alsu F. Saifitdinova, SPIN: [5114-4844](#), Scopus AuthorID: [6603195889](#), ResearcherID: [C-1104-2011](#), ORCID: [0000-0002-1221-479X](#), e-mail: saitfdinova@mail.ru

For citation: Saifitdinova, A. F. (2022) Organisation of non-coding elements in avian genomes. *Integrative Physiology*, vol. 3, no. 2, pp. 185–203. <https://www.doi.org/10.33910/2687-1270-2022-3-2-185-203> EDN IXRERA

Received 22 March 2022; reviewed 4 May 2022; accepted 12 May 2022.

Funding: The study did not receive any external funding

Copyright: © A. F. Saifitdinova (2022). Published by Herzen State Pedagogical University of Russia. Open access under [CC BY-NC License 4.0](#).

Abstract. The review focuses on non-coding elements in bird genomes. Like most living organisms, birds carry a significant number of non-coding sequences in their genome. Technically, these sequences cannot be classified as genes. They include DNA regions with structural functions and high level of conservatism as well as a considerable number of sequences, including repetitive fractions, whose functions are not yet known. Apart from a general overview of tandem repeats in avian genomes, the paper provides a special focus on differences in copy number of individual repeats. Interestingly, this number directly correlates with a reduction in genome size in birds. Besides, the review provides a detailed description of how repetitive sequences are involved in structural and functional organisation of centromeric and telomeric regions with a special emphasis given to the organisation of pericentromeric and subtelomeric heterochromatin. It also describes what makes these regions different from centromeric and telomeric sequences per se. The data on the distribution of repeats in bird genomes and their specific organisation are compared with the level of their conservatism and their role in maintaining structural and functional organisation of the genome.

Keywords: genome organisation, non-coding DNA, repetitive sequences, heterochromatin, high-order repeats

Введение

Геномы эукариот, помимо кодирующих последовательностей, несущих информацию об аминокислотных последовательностях, включают существенную долю некодирующих белки элементов. Некоторые из них транскрибируются с образованием различных типов РНК, а для части последовательностей транскрипты до сих пор не удалось обнаружить ни в каких типах клеток. Реализация международного исследовательского проекта «Функциональная аннотация геномов животных» (Functional Annotation of Animal Genome, FAANG) позволила подробно охарактеризовать различные компоненты геномов некоторых видов домашних животных, это значительно увеличило объем информации о составе геномов и позволило подробно аннотировать не только гены, но и различные некодирующие элементы (Andersson et al. 2015). Представители класса Птицы характеризуются стабильным кариотипом, который сформирован двумя группами отличающихся по структурным и молекулярным характеристикам хромосом, имеющими разные размеры и получившими название микрохромосом и макрохромосом (Rodionov 1996). Несмотря на то, что размеры геномов у птиц меньше, чем у большинства видов позвоночных (Karusta, Suh 2017), по данным проекта FAANG, число кодирующих последовательностей в геноме курицы не уступает таковому у других видов позвоночных и составляет около 23 тысяч. Сокращение размера генома у представителей этого таксона произошло за счет уменьшения доли повторяющихся элементов, причем псевдогены найдены только у 24% генов, доля повторов в геноме составляет всего 16%, из них на сателлиты и простые повторы приходится около 7%, и только 9% составляют мобильные элементы различной природы, что существенно отличает

птиц от других видов позвоночных. Для объяснения этого феномена высказывались предположения, что наряду с активной утратой повторяющихся элементов и позитивным действием отбора в этом направлении, обусловленным требованиями облегчения веса для обеспечения полета, в клетках зародышевой линии у птиц сформировались препятствующие проникновению и распространению мобильных элементов (Wallis et al. 2004; Wicker et al. 2005). Это может быть связано с более ранним, по сравнению с млекопитающими, определением судьбы клеток зародышевого пути в эмбриогенезе у птиц. В настоящий момент высказанные предположения все еще не имеют экспериментального обоснования, однако данные исследований геномов птиц доказали, что существенному сокращению подверглись сателлиты, а также некоторые виды рассеянных повторов, включая псевдогены и сегментные дубликации, а также ограниченное распространение получили эндогенные вирусы различной природы (Warren et al. 2017; Zhang et al. 2014a). Сохранившиеся некодирующие последовательности отличаются по степени консерватизма, функциональному статусу хроматина и распределению в геноме. К таким элементам принято относить участки ДНК, выполняющие структурные функции, которые отличаются высоким уровнем консерватизма, а также значительное количество последовательностей, функции которых до сих пор не выяснены. К наиболее консервативным некодирующим элементам относятся различные виды структурных и регуляторных элементов, тогда как псевдогены, мобильные элементы, эндогенные вирусы и сателлитные ДНК, составляющие гетерохроматиновые районы, характеризуются низкой степенью консерватизма, как в пределах вида, так и в составе генома одной особи.

Особенности регуляторных элементов у птиц

К регуляторным последовательностям относятся: промоторы и сайты связывания транскрипционных факторов; удаленные регуляторные элементы, такие как повышающие уровень транскрипции гена или группы генов энхансеры и приводящие к понижению или полному подавлению транскрипции гена сайленсеры; различные операторы, в том числе отвечающие за интенсивность транскрипции подконтрольных генов и альтернативный сплайсинг; а также последовательности ДНК, способные блокировать взаимодействие между регуляторными элементами, отвечающие за поддержание топологически ассоциированных доменов хроматина архитектурные элементы. У птиц, как у и большинства видов позвоночных, поддержание пространственной организации хроматина зависит от связывания с белком CTCF (СССТС-связывающий фактор), который впервые был описан именно у курицы (Lobanekov et al. 1990). Его распределение в геномах птиц подробно изучено в рамках реализации международного исследовательского проекта «Энциклопедия элементов ДНК» (ENCODE). Промоторные районы генов у птиц имеют некоторые характерные особенности, отличающие их от большинства очень консервативных регуляторных последовательностей и сайтов связывания транскрипционных факторов (Abe, Gemmell 2014). У курицы средняя длина коровой области промоторов, ответственных за связывание с РНК-полимеразой, короче, чем у млекопитающих, что может объясняться более высокой частотой присутствия в них ТАТА-бокса. Содержание GC пар в их составе варьирует в широком диапазоне от 31,9% до 73,6% (среднее значение составляет 51,5%), причем содержание GC постепенно увеличивается по мере приближения к сайту инициации транскрипции. CpG-островки обнаружены в промоторах 58,3% всех генов, что свидетельствует о высокой роли эпигенетической регуляции. Коровые районы промоторов курицы обогащены полипуриновыми мотивами и G-квадруплексами, в то время как в других районах генома они встречаются редко. Несмотря на общее сокращение числа коротких tandemных повторов (STR), более 10% генов содержат STR в регуляторной области промотора (Abe, Gemmell 2014). Присутствие повторяющихся элементов является характеристикой регулируемых промоторов, способных подстраивать интенсивность транскрипции под потребности клетки в ответ на внешние сигнала

лы (Zopl et al. 1990). Наши исследования также показали высокую частоту локализации STR в интронах; особенно часто такие элементы встречаются в генах, кодирующих трансмембранные белки и транскрипционные факторы (Bolshakova, Saifitdinova 2020), что говорит об их функциональном значении и возможном участии в регуляции альтернативного сплайсинга.

Характеристика повторяющихся элементов у птиц

Сокращение числа копий повторяющихся элементов у птиц создало сложности для их идентификации и клонирования из тотальной ДНК в догеномную эру на основе рутинных методов анализа, т. к. они не образовывали характерных для сателлитов фракций, пригодных для препаративного выделения. Тем не менее некоторые последовательности были охарактеризованы главным образом благодаря рестрикционному анализу и структурным особенностям самих последовательностей. В таблице 1 представлен список депонированных в базу данных GenBank повторяющихся элементов птиц.

Большинство некодирующих элементов представлено в геноме множественными копиями, распределение которых может носить как рассеянный характер, так и tandemный, причем ориентация в пределах блока tandemных повторов может быть как «голова к хвосту», так и «голова к голове», а также их сочетания, что дополнительно может определять свойства как самой последовательности ДНК, так и образующихся транскриптов. Особенности организации таких элементов в геноме создают трудности для сборки и их правильной локализации при анализе данных. В четвертой версии сборки генома курицы Gallus_gallus-4.0 (GCA_000002315.2) их доля составила всего около 12,5%. Улучшение методов анализа позволило увеличить эту цифру до 16,4% в следующей версии сборки Gallus_gallus-5.0 (GCA_000002315.3), однако в основном это коснулось рассеянных повторов. Области tandemно повторенных последовательностей остались непредставленными в сборке до хромосом и обнаруживались фрагментарно в составе отдельных нелокализованных контигов, объединенных под названием chromosome unknown («неопределенная хромосома»).

Шестая версия генома курицы Gallus_gallus-6.0 (GCA_000002315.5) основана на ресеквенированных данных и включает также данные

Табл. 1. Список клонированных и охарактеризованных повторов птиц

Название элемента	Регистрационный номер	Объекты исследования	Ссылка на источник
GGXhoI	M24754	<i>Gallus gallus domesticus</i>	Kodama et al. 1987
CNM	X51431	<i>G. g. domesticus</i>	Matzke et al. 1990
GGSAT (<i>EcoRI</i>)	X57344	<i>G. g. domesticus</i>	Saitoh et al. 1991
TM	X66696	<i>Melleagris gallopavo</i>	Matzke et al. 1992
PR1	X97402	<i>Columba livia</i>	Solovei et al. 1996
CR1	U88211	<i>G. g. domesticus</i>	Haas et al. 1997
CJA- <i>Bgl</i> II	AB035968	<i>Coturnix japonica</i>	Tanaka et al. 2000
FCP	AF160980	<i>Fringilla coelebs</i>	Saifitdinova et al. 2000
MHM	AB046699	<i>G. g. domesticus</i>	Teranishi et al. 2001
GGSpI	AB074190	<i>G. g. domesticus</i>	Itoh, Mizuno 2002
CCH-L3	AB104632	<i>Coturnix chinensis</i>	Yamada et al. 2002
CCH-S7	AB104630	<i>C. chinensis</i>	Yamada et al. 2002
CCH-S9	AB104631	<i>C. chinensis</i>	Yamada et al. 2002
KBL- <i>EcoRI</i>	AB103252	<i>Bubo blakistoni</i>	Yamada et al. 2004
SUH- <i>EcoRI</i>	AB103288	<i>Strix uralensis</i>	Yamada et al. 2004
PO41	CL269118	<i>G. g. domesticus</i>	Wicker et al. 2005
PIR	AF124927	<i>G. g. domesticus</i>	Wicker et al. 2005
NME- <i>Apa</i> I	AB189147	<i>Numida meleagris</i>	Yamada et al. 2006
CCXbaI	AY491001	<i>G. g. domesticus</i>	Li et al. 2007
ZBM1	EU099581	<i>Taeniopygia guttata</i>	Itoh et al. 2008
ZBM2	EU099582	<i>T. guttata</i>	Itoh et al. 2008
LL2R	KT952327	<i>G. g. domesticus</i>	Krasikova et al. 2010
CVI- <i>Msp</i> I	AB872148	<i>Colinus virginianus</i>	Ishishita et al. 2014
CVI- <i>Nsi</i> I	AB872150	<i>C. virginianus</i>	Ishishita et al. 2014
CVI- <i>Hae</i> III	AB872146	<i>C. virginianus</i>	Ishishita et al. 2014
ACH- <i>Sau</i> 3AI	AB872160	<i>Alectoris chukar</i>	Ishishita et al. 2014
CjapSAT	MH475922	<i>C. japonica</i>	Kulak et al. 2018
APL- <i>Hae</i> III	LC416791	<i>Anas platyrhynchos</i>	Uno et al. 2019
AFA- <i>Hinfl</i> -S	LC416770	<i>Anser fabalis</i>	Uno et al. 2019
AFA- <i>Hinfl</i> -L	LC416774	<i>A. fabalis</i>	Uno et al. 2019
CCY- <i>Apa</i> I	LC416776	<i>Cygnus cygnus</i>	Uno et al. 2019

Table 1. List of cloned and described avian repeats

Name of the element	Accession numbers	Species names	References
GGXhoI	M24754	<i>Gallus gallus domesticus</i>	Kodama et al. 1987
CNM	X51431	<i>G. g. domesticus</i>	Matzke et al. 1990
GGSAT (<i>EcoRI</i>)	X57344	<i>G. g. domesticus</i>	Saitoh et al. 1991
TM	X66696	<i>Melleagris gallopavo</i>	Matzke et al. 1992
PR1	X97402	<i>Columba livia</i>	Solovei et al. 1996
CR1	U88211	<i>G. g. domesticus</i>	Haas et al. 1997
CJA- <i>Bgl</i> II	AB035968	<i>Coturnix japonica</i>	Tanaka et al. 2000
FCP	AF160980	<i>Fringilla coelebs</i>	Saifitdinova et al. 2000
MHM	AB046699	<i>G. g. domesticus</i>	Teranishi et al. 2001
GGSpI	AB074190	<i>G. g. domesticus</i>	Itoh, Mizuno 2002
CCH-L3	AB104632	<i>Coturnix chinensis</i>	Yamada et al. 2002
CCH-S7	AB104630	<i>C. chinensis</i>	Yamada et al. 2002
CCH-S9	AB104631	<i>C. chinensis</i>	Yamada et al. 2002
KBL- <i>EcoRI</i>	AB103252	<i>Bubo blakistoni</i>	Yamada et al. 2004
SUH- <i>EcoRI</i>	AB103288	<i>Strix uralensis</i>	Yamada et al. 2004
PO41	CL269118	<i>G. g. domesticus</i>	Wicker et al. 2005
PIR	AF124927	<i>G. g. domesticus</i>	Wicker et al. 2005
NME- <i>Apa</i> I	AB189147	<i>Numida meleagris</i>	Yamada et al. 2006
CCXbaI	AY491001	<i>G. g. domesticus</i>	Li et al. 2007
ZBM1	EU099581	<i>Taeniopygia guttata</i>	Itoh et al. 2008
ZBM2	EU099582	<i>T. guttata</i>	Itoh et al. 2008
LL2R	KT952327	<i>G. g. domesticus</i>	Krasikova et al. 2010
CVI- <i>Msp</i> I	AB872148	<i>Colinus virginianus</i>	Ishishita et al. 2014
CVI- <i>Nsi</i> I	AB872150	<i>C. virginianus</i>	Ishishita et al. 2014
CVI- <i>Hae</i> III	AB872146	<i>C. virginianus</i>	Ishishita et al. 2014
ACH- <i>Sau</i> 3AI	AB872160	<i>Alectoris chukar</i>	Ishishita et al. 2014
CjapSAT	MH475922	<i>C. japonica</i>	Kulak et al. 2018
APL- <i>Hae</i> III	LC416791	<i>Anas platyrhynchos</i>	Uno et al. 2019
AFA- <i>Hinfl</i> -S	LC416770	<i>Anser fabalis</i>	Uno et al. 2019
AFA- <i>Hinfl</i> -L	LC416774	<i>A. fabalis</i>	Uno et al. 2019
CCY- <i>Apa</i> I	LC416776	<i>Cygnus cygnus</i>	Uno et al. 2019

транскриптомов (Lawal et al. 2018), однако и это не позволило восполнить все существующие проблемы в районах локализации tandemных повторов. Существенным преимуществом этой версии является полная доступность для анализа сырых данных, депонированных в GenBank под номером SRP142580. Результаты наших исследований, полученные на основе анализа сырых данных полногеномного секвенирования, дополнили картину разнообразия повторяющихся элементов и особенно информацию о составе tandemных повторов (Komissarov et al. 2018). Данные о наиболее представленных tandemных повторах в геноме курицы подтверждают общую тенденцию к утрате повторяющихся элементов в геномах птиц с преобладанием представителей различных микросателлитов.

Это исследование также позволило дополнить информацию о составе половых хромосом у курицы на основе различий в числе копий отдельных элементов в геномах самцов и самок, а также описать и локализовать на хромосоме W курицы второй по представленности в геноме самок tandemный повтор (GGAAA)n (Komissarov et al. 2018). Дифференцировка нерекombинирующей хромосомы связана с накоплением tandemных повторов, принадлежащих к разным классам, что характерно для эволюции гетероморфных половых хромосом у разных организмов (O'Meally et al. 2010). У птиц с гетерогаметным женским полом доказано монофилетическое происхождение половых хромосом, однако эволюция нерекombинирующей хромосомы W шла с разной скоростью, что привело к существенным структурным отличиям этой половой хромосомы у представителей разных таксонов (Zhou et al. 2014). Несмотря на происхождение от одной и той же пары аутосом, сходство

в организации хромосомы W у разных видов птиц сохраняется главным образом в псевдоаутосомных районах, тогда как спектр и количество tandemных повторов может существенно варьировать (Itoh et al. 2006). Этим объясняется тот факт, что цельнохромосомный зонд к хромосоме W курицы не дает специфичного сигнала на соответствующей хромосоме зебровой амадины, имеющей общее с ней аутосомное происхождение (Itoh, Arnold 2005). У большинства изученных представителей курообразных в состав половой хромосомы W входят протяженные блоки tandemных повторов, относящиеся к следующим семействам: *XhoI*, помимо курицы присутствующие в геномах индейки и обыкновенного фазана; *SspI*, найденные, помимо курицы, в геномах зеленого фазана, японского перепела, павлина, азиатского кеклика и цесарки, а также *ApaI*, обнаруженные в составе половой хромосомы W японского перепела и цесарки (Itoh et al. 2008; Yamada et al. 2006). Повторяющийся элемент PIR входит в состав половых хромосом многих видов, образуя различные варианты минисателлитов с длиной мономера 41 п.н. (Deryusheva et al. 2007; Itoh et al. 2008; Krasikova et al. 2006; Wang et al. 2002). На сегодняшний день половая хромосома W курицы остается наиболее подробно изученной. Для нее определены и картированы основные представители семейств tandemных повторов (табл. 2), составляющих около 82% ее ДНК, тогда как на долю различных рассеянных повторов приходится лишь около 13%. Также для хромосомы W локализованы и подробно охарактеризованы псевдоаутосомные районы с полной сборкой кодирующих элементов (Bellott et al. 2017), которые составляют около 5% от общей длины нуклеотидной последовательности этой хромосомы.

Табл. 2. Tandemные повторы в составе половой хромосомы W курицы

Название элемента	Характеристика	Ссылка на источник
GGSAT (<i>EcoRI</i>)	сателлит с мономером 1,2 т.п. н.	Saitoh, Mizuno 1992
GGXhoI (TTAGGG)n	сателлит с мономером 0,7 т. п. н. теломерный повтор	Saitoh, Mizuno 1992 Solovei et al. 1994
GGSpI	сателлит с мономером 508 п. н.	Itoh, Mizuno 2002
CNM	минисателлит с мономером 41 п. н.	Krasikova et al. 2006
PO41	минисателлит с мономером 41 п. н.	Deryusheva et al. 2007
(GGAAA)n	микросателлит	Komissarov et al. 2018

Table 2. Tandem repeats in the chicken W sex chromosome

Name of the element	Description	References
GGSAT (<i>EcoRI</i>)	satellite with a 1.2 kbp monomer	Saitoh, Mizuno 1992
GGXhoI (TTAGGG)n	satellite with a 0.7 kbp monomer telomere repeat	Saitoh, Mizuno 1992 Solovei et al. 1994
GGSpI	satellite with a 508 bp monomer	Itoh, Mizuno 2002
CNM	minisatellite with a 41 bp monomer	Krasikova et al. 2006
PO41	minisatellite with a 41 bp monomer	Deryusheva et al. 2007
(GGAAA)n	microsatellite	Komissarov et al. 2018

Функциональное значение отдельных повторяющихся элементов в составе хромосомы W курицы все еще остается не до конца понятным. Обнаруженные тандемные повторы отличаются не только длиной мономера, нуклеотидным составом и уровнем консерватизма, они могут по-разному вести себя в соматических клетках. Например, GGSAT и GGXhoI входят в состав конститутивного гетерохроматина, тогда как GGSsp1 и (GGAAA)_n в интерфазных ядрах диспергированы (Itoh, Mizuno 2002; Komissarov et al. 2018). Повторяющийся элемент (GGAAA)_n, помимо формирования крупных блоков на гетероморфной половой хромосоме, присутствует в составе транскрибируемых некодирующих районов различных генов. Некоторые данные свидетельствуют о возможном участии самих повторов или их транскриптов в регуляции активности промоторов или переключении полимеразных комплексов на альтернативные промоторы для получения разных вариантов транскриптов (Bolshakova, Saifitdinova 2020). Повторяющаяся последовательность (GGAAA)_n описана в составе 5' транскрибирующейся некодирующей области гена овотрансферрина у фазана; в геноме курицы описаны два родственных повтора в составе регуляторных областей генов, отвечающих за дифференцировку гонад (Maroteaux et al. 1983). Эта последовательность описана ранее в составе регулируемого промотора гена промежуточного нейрофиламента у курицы (Zopl et al. 1990).

Были обнаружены индивидуальные отличия по числу копий повторяющихся последовательностей у птиц (Wang, Byer 2014). У представителей различных пород кур выявлены различия в количестве некоторых видов повторяющихся элементов, что может объясняться давлением отбора в ходе одомашнивания и приспособления к условиям среды (Lawal et al. 2018; Piégu et al. 2020; Solovei et al. 1994). Имеются данные о развитии фенотипических признаков, связанных с изменением числа копий отдельных последовательностей в некоторых локусах, что приводит к локальному изменению архитектуры хроматина (Griffin et al. 2008; Seol et al. 2019). Хорошо известный фенотип гороховидного гребешка у кур связан с экспансией множественных копий повторяющихся последовательностей в интроне гена транскрипционного фактора SOX5 (Wright et al. 2009). Дупликация 20 т. п. н. (тысяч п. н.) участка в регуляторной области гена эумезодермина EOMES приводит к возникновению раздвоенного гребешка у кур (Dorshorst et al. 2015). Инсерция участка повторов длиной 73 т. п. н., соответствующего фрагменту хромо-

сомы 1, в регуляторную область транскрипционного фактора WNT11 на хромосоме 3 у курицы и вызванное им изменение характера экспрессии гена рецептора к ретиноевой кислоте в ходе развития определяет возникновение признака голошеести у кур (Mou et al. 2011). Инвертированная дупликация участка длиной 400 т. п. н. на хромосоме 20 приводит к комплексным изменениям морфогенеза, вызывающим гиперпигментацию кожи и соединительной ткани, а также потерю бородок высокого порядка в маховых перьях у представителей китайской шелковой породы кур (Dorshorst et al. 2015; Shinomiya et al. 2012). Развитие технологий полногеномного секвенирования и анализа данных постепенно расширяет информацию о роли повторяющихся элементов в формировании индивидуальных морфологических признаков.

Еще большие различия выявляются при сравнении геномов близких видов, о чем свидетельствуют различия в числе копий отдельных повторов у разных видов голубей (Kretschmer et al. 2018; Solovei et al. 1996). Происходит эволюционирование отдельных видов тандемных повторов, в том числе с изменением их локализации. Такие изменения затрагивают и спектр представленных в геноме микросателлитов, которые эволюционируют быстрее всего. Анализ геномов 16 видов птиц из разных систематических групп позволил определить районы, подвергавшиеся дупликациям в эволюции птиц, и эти данные свидетельствуют о том, что такие события происходили у них гораздо реже, чем у млекопитающих. В то же время гораздо чаще они имели место в микрохромосомах, причем решающее значение имел не физический размер хромосомы у данного вида, а ее анцестральное состояние (Skinner et al. 2014). Это свидетельствует о роли конкретного нуклеотидного контекста в создании условий для увеличения числа копий отдельных последовательностей в геноме.

Уникальный пример происходящих изменений на большой эволюционной дистанции демонстрирует сравнение спектра и числа копий повторяющихся элементов в пределах Курообразных на примере геномов домашней курицы и японского перепела. Несмотря на общую тенденцию в эволюции птиц к сокращению числа повторяющихся элементов, здесь можно видеть пример обратного процесса. Геном японского перепела больше при сохранении числа хромосом и основных групп синтении (Nakamura et al. 1990; Schmid et al. 1982). По сравнению с домашней курицей, в геноме японского пере-

пела шире представлены различные повторяющиеся элементы на основе мономера (TATGGGGCAGG). Этот базовый мономер, в свою очередь, участвует в формировании не только наиболее распространенных тандемных повторов с длиной мономера 40–41 п. н., но и образует гетеродимеры и гетеротримеры разной длины, входящие в состав тандемных повторов и повторов высокого порядка (HOR, high order repeat). Анализ геномных данных подтвердил наличие в геноме японского перепела сложных HOR с участием элементов PO41 и CJA-BglII (Deryusheva et al. 2007). Направление отдельных мономеров в составе таких HOR может быть различным, что создает возможности образования транскриптов отдельного элемента в обоих направлениях при сквозной транскрипции и возможности образования фрагментов двунитевых РНК. Разные варианты организации тандемных повторов представлены на рисунке 1.

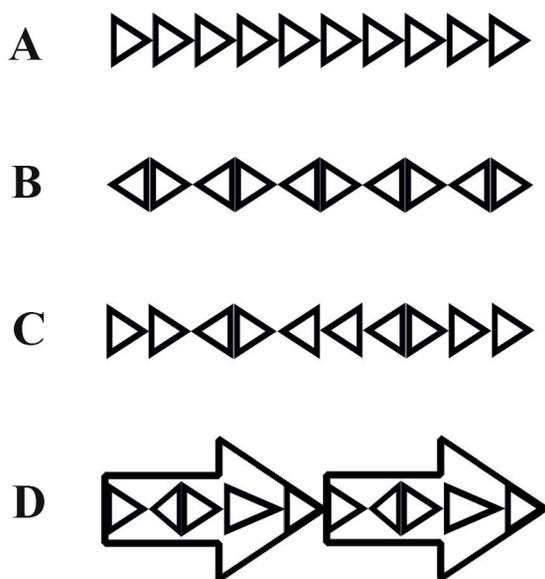


Рис. 1. Схематическое представление организации тандемных повторов в геноме. А. Расположение мономеров друг за другом «голова к хвосту». В. Расположение мономеров во встречных направлениях. С. Расположение мономеров в разном направлении. D. Тандемные повторы в составе HOR могут включать разные мономеры в разной ориентации, образуя мономер более высокого порядка

Fig. 1. A schematic representation of the organisation of tandem repeats in the genome. A. Head-to-tail arrangement of monomers one after another. B. Arrangement of monomers in opposite directions. C. Arrangement of monomers in different directions. D. Tandem repeats within HOR can include different monomers with different orientations, forming a higher order monomer

Представители разных видов птиц характеризуются определенным спектром не только тандемных, но и рассеянных элементов, принадлежащих к разным классам (Wicker et al. 2007). Информация о составе и количестве диспергированных повторяющихся последовательностей в геномах различных видов птиц лучше представлена по сравнению с тандемными повторами в аннотированных полных геномах и базах данных собранных до контигов геномов. Рассеянные повторы представлены разными классами мобильных элементов как полностью функциональными, так и их производными. К ним относятся эндогенные вирусные элементы, включающие сохранившиеся в геноме после инфицирования и проникновения в линию клеток зародышевого пути последовательности однонитевых ДНК и РНК вирусов, а также двунитевых ДНК вирусов, однако у разных видов птиц число копий таких элементов не превышает десятка (Cui et al. 2014). В геномах птиц практически все мобильные элементы представлены ретротранспозонами, относящимися к ретровирусам (ERV, endogenous retroviruses), содержащими длинные терминальные повторы (LTR, long terminal repeats) или их производными (Wicker et al. 2007). В полностью функциональных элементах между двумя идентичными LTR находятся последовательности, кодирующие белок капсида, аспарагиновую протеазу, обратную транскриптазу, РНКазу H, интегразу и белок оболочки. По периферии LTR находятся дублированные таргетные последовательности, по количеству и нуклеотидной последовательности таких повторов, LTR в геномах птиц подразделяют на подсемейства ERV1, ERVK и ERVL (Kapusta, Suh 2017). Особенности организации таргетных последовательностей позволяют идентифицировать и классифицировать даже единичные ERV, оставшиеся после удаления кодирующих последовательностей в результате неаллельной гомологичной рекомбинации (рис. 2А, В). Ретротранспозоны, не имеющие LTR, представлены у птиц длинными рассеянными повторами (LINE, long interspersed nuclear element) и короткими рассеянными повторами (SINE, short interspersed nuclear element), они характеризуются наличием 3'-концевых повторов (рис. 2 С, D, Е). В геномах птиц представлены два основных надсемейства LINE, различающиеся по составу 3'-концевых повторов: AviRTE, характеризующиеся последовательностью (ATG)_n, а также CR1, несущие (SATTCTRT)_n (Weissensteiner, Suh 2019). Кодированная последовательность функциональных LINE несет две открытые рамки

считывания для синтеза эндонуклеазы и обратной транскриптазы, тогда как SINE кодируют малые РНК, а у птиц в основном эти последовательности имеют в своем составе дериваты транспортной РНК (Wicker et al. 2007). То есть SINE относятся к неавтономным элементам и могут перемещаться только за счет активности LINE. Интересно, что если в геноме человека число SINE превосходит 1,5 миллиона копий на геном и они имеют тенденцию к накоплению в обогащенных генами районах, то у представителей разных видов птиц число SINE колеблется в пределах от 6 до 17 тысяч (Karusta, Suh 2017).

ДНК-транспозоны в геномах птиц представлены несущими инвертированные терминальные повторы элементами (рис. 2 F, G). Функциональные полноразмерные мобильные элементы содержат последовательность, кодирующую транспозазу, тогда как дериваты представляют собой миниатюрные инвертированные повторяющиеся элементы, способные к неавтономной транспозиции (Wicker et al. 2007). Независимая эволюция птиц привела к формированию характерных для их геномов особенностей состава и распределения рассеянных повторяющихся элементов, которые сформировались в геноме их общего предка. В отличие от ближайших родственников из числа рептилий, птицы утратили значительную часть ERV, а также существенно сократили долю в геноме коротких рассеянных повторов (Cui et al. 2014; Wallis et al. 2004; Zhang et al. 2014b). В геномах некоторых изученных птиц практически полностью отсутствуют эндогенные вирусные элементы и активные копии SINE, которые в геномах служат источниками вариативности и основой для эволюции благодаря новой функциональности, что ограничивает возможности их распространения даже при снятии давления отбора (Morris et al. 2020; Warren et al. 2017; Wicker et al. 2005).

Наиболее распространенным типом рассеянных повторов у птиц оказались дериваты элемента CR1 (Haas et al. 1997; Morris et al. 2020; Treplin, Tiedemann 2007; Warren et al. 2017; Watanabe et al. 2006). Полноразмерный мономер CR1 из генома курицы (регистрационный номер в GenBank U88211) представляет собой ретротранспозон, относящийся к классу LINE, который содержит две открытые рамки считывания (Haas et al. 1997). В геноме курицы полноразмерные, частично вырожденные и укороченные элементы, относящиеся к семейству CR1 и его производным, составляют до 80% всех рассеянных повторов (Wallis et al. 2004; Warren et al.

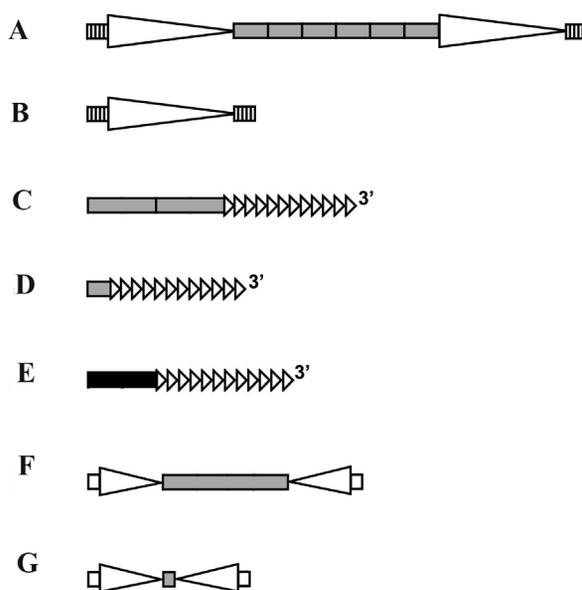


Рис. 2. Схемы организации представленных в геномах птиц рассеянных повторяющихся элементов. Серые прямоугольные блоки показывают белок-кодирующие последовательности мобильных элементов или их дериваты, черные блоки показывают последовательности малых РНК, треугольники обозначают повторяющиеся элементы в составе мобильного элемента, белые прямоугольные блоки обозначают повторы таргетных последовательностей. А. Функциональный LTR ретротранспозон. В. Одиночный LTR. С. Функциональный LINE. D. Неавтономный дериват LINE. E. SINE. F. Функциональный ДНК-транспозон. G. Миниатюрный инвертированный мобильный элемент, представляющий из себя дериват ДНК-транспозона, сохранивший терминальные инвертированные элементы и утративший значительную часть гена транспозазы

Fig. 2. Organisation of scattered repetitive fractions in bird genomes. Gray rectangular blocks show protein-coding sequences of mobile elements or their derivatives; black blocks show small RNA sequences; triangles indicate repetitive fractions in the mobile element; white rectangular blocks indicate repeats of target sequences. A. Functional LTR retrotransposon. B. Single LTR. C. Functional LINE. D. A non-autonomous derivative of LINE. E. SINE. F. Functional DNA transposon. G. A miniature inverted mobile element—a derivative of a DNA transposon that retained terminal inverted elements and lost a significant part of the transposase gene

2017; Wicker et al. 2005). Для встраивания в геном CR1 использует механизм, сходный с теми, которые характерны для других представителей LINE. Он основан на инициации разрыва и праймировании последовательности 3'-конца. Несмотря на то, что подавляющее большинство повторов CR1 представляет собой небольшой

фрагмент производной 3'-концевого повтора мобильного элемента, содержащего в большинстве случаев консенсусную последовательность (CATTC²TRT)(GATTC²TRT)₁₋₃, в геноме курицы были обнаружены и полноразмерные копии длиной 4558 п. н., имеющие в составе функциональные гены эндонуклеазы и обратной транскриптазы (Haas et al. 1997). Это свидетельствует о сохраняющейся способности этого элемента к ретротранспозиции и обеспечению перемещения некоторых укороченных элементов. Подавляющее большинство повторяющихся элементов CR1 утратили способность к транспозиции и дивергировали с образованием различных подсемейств тандемных и рассеянных повторов (Wallis et al. 2004; Wicker et al. 2005). Было показано, что в геномах птиц производные CR1 предпочтительно локализуются в обогащенных генами районах, что нехарактерно для представителей LINE у других групп организмов (Wallen et al. 1996). Это свидетельствует о том, что утрата большинства SINE у предков птиц позволила дериватам CR1 распространиться в районы эухроматина. Подтверждением этому служат результаты физического картирования CR1 у представителей разных видов птиц, относящихся к разным отрядам (Coullin et al. 2005).

Широкое распространение CR1 привело к его эволюционированию даже внутри отдельных геномов, хотя представленность этого повтора у разных видов птиц варьирует в широких пределах (Liu et al. 2009; Watanabe et al. 2006). На основании различий 5'- и 3'-некодирующих районов в геноме курицы были выделены 14 основных подсемейств CR1 (Haas et al. 1997; John, Quinn 2008; Kapusta, Suh 2017), которые, в свою очередь, подразделяются на более мелкие группы на основании различий последовательностей и комбинации 5'- и 3'-концов (Liu et al. 2009). Анализ филогенетических взаимоотношений между различными подсемействами CR1 у курицы и насыщенность их CpG позволила определить их возраст, распределение по хромосомам и степень дивергенции. Наиболее молодые элементы, такие как CR1-B2-2, находятся в обогащенных генами районах микрохромосом, представители самого распространенного семейства CR1-C4 имеют умеренный возраст, и именно они оккупируют обогащенные генами районы макрохромосом, а самое древнее и наименее консервативное семейство CR1-Y4-4 наиболее представлено на хромосоме Z, где степень дивергенции этих элементов достигает почти 30% (Liu et al. 2009).

Другие типы рассеянных повторов у птиц распространены значительно меньше, по срав-

нению с представителями LINE. В геноме курицы содержится приблизительно в пять раз меньше мобильных элементов, относящихся к LTR ретротранспозонам, по сравнению с CR1 (Warren et al. 2017). Более ранние данные указывали на повышенную представленность таких элементов в составе макрохромосом (Wallis et al. 2004), однако с увеличением качества геномных сборок, в особенности в части восполнения пробелов в микрохромосомах, наблюдается выравнивание этого показателя (Warren et al. 2017). Большинство LTR ретротранспозонов в геноме курицы потеряли функциональность и представлены укороченными последовательностями, утратившими некоторые кодирующие элементы, или одиночными LTR, однако около 40% всех найденных элементов оказались интактными и способными к транспозиции. Функциональные LTR ретротранспозоны, вне зависимости от возраста вставки и происхождения, были идентифицированы в составе кластеров, где плотность таких элементов была в пять раз выше, чем в среднем по геному (Wallis et al. 2004). Такие кластеры главным образом локализуются в районах аутосом, характеризующихся низкой скоростью рекомбинации (Warren et al. 2017). Анализ содержания LTR в 48 геномах птиц из разных систематических групп выявил довольно широкий разброс по этому показателю у представителей разных эволюционных ветвей (Zhang et al. 2014b). Активность мобильных элементов стала основной причиной увеличения размеров блоков гетерохроматина у японского перепела (Morris et al. 2020; Saifitdinova et al. 2019). Большая роль экспансии повторяющихся элементов в эволюции гетероморфных половых хромосом уже обсуждалась выше, но здесь важно отметить различия в типах повторяющихся элементов, накапливающихся на разных половых хромосомах, и особенностях их распределения (Komissarov et al. 2018; Saifitdinova et al. 2016). Половая хромосома Z имеет плотность распределения LTR элементов на 64% выше, чем можно было бы ожидать на аутосоме курицы той же длины (Warren et al. 2017). Аналогичные процессы, по-видимому, происходили в эволюции хромосомы линии зародышевого пути (GRC, germ line restricted chromosome) у певчих птиц, в частности у амадины (Torgasheva et al. 2019). Особенности поддержания этой хромосомы в геноме в ряду поколений и отсутствие давления отбора, направленного на облегчение веса организма из-за происходящей в соматических клетках диминуции хроматина, привели к накоплению в ней повторов, относящихся к разным классам. У певчих птиц наблюдается особенно

высокое содержание LTR в геномах, их активность приводит к увеличению разнообразия и независимому эволюционированию этих видов рассеянных повторов (Weissensteiner, Suh 2019).

Анализ геномов разных видов птиц позволил проследить историю внедрения, распространения и эволюции некоторых ДНК-транспозонов (Bertocchi et al. 2017; Morris et al. 2020), хотя их доля по сравнению с описанными выше элементами в геномах птиц существенно ниже. В геномах птиц не обнаружены другие типы ДНК транспозонов, кроме характеризующихся наличием терминальных инвертированных повторов. Фланги терминальных инвертированных повторов несут последовательности, представляющие собой дубликации таргетного сайта. Длина и особенность этих последовательностей помогает определить их принадлежность к конкретной группе элементов, т. к. большинство ДНК транспозонов в геноме потеряли автономность из-за частичной или полной утраты последовательности, кодирующей транспозазу с образованием неавтономных миниатюрных инвертированных мобильных элементов (Wicker et al. 2007). Большая часть ДНК-транспозонов в геномах птиц принадлежит к двум суперсемействам: Tc1/Mariner и hAT (Kapusta, Suh 2017; Warren et al. 2010; Wicker et al. 2005). Элемент Charlie7, принадлежащий к hAT, накапливался в геномах в ходе раннего периода эволюции птиц и присутствует как у курицы, так и у амадины (Kapusta, Suh 2017). У домашней курицы в качестве рассеянных повторяющихся элементов наиболее широко представлены два достаточно древних семейства ДНК-транспозонов Charlie12 и Galluhor (Wallis et al. 2004; Wicker et al. 2005). Последний относится к Mariner-подобным элементам и обнаружен в разном количестве в геномах пяти видов курообразных: домашней курицы, индейки, японского перепела, виргинской американской куропатки и тетерева-косача, а также у одного из видов птиц-носорогов — малайского калао (Bertocchi et al. 2017). Проникновение транспозона Galluhor в геномы представителей обеих групп произошло, по-видимому, путем горизонтального переноса. Количество копий этого транспозона варьирует в широких пределах: от 5 у японского перепела до более чем 10 тысяч у домашней курицы, причем взрывной характер и распространение транспозона Galluhor имело место у общего предка курицы и индейки (Bertocchi et al. 2017).

Организация центромерных районов хромосом

Многие тандемные повторы в геноме выполняют структурные функции, к их числу от-

носятся центромерные и теломерные повторы. Центромеры отвечают за сборку кинетохора, необходимого для прикрепления к микротрубочкам веретена деления и правильной сегрегации хромосом во время клеточных делений. Центромерные районы хромосом имеют особый эпигенетический статус, отличающий их как от эухроматиновых районов генома, так и от гетерохроматина, в том числе непосредственно прилегающего к области формирования кинетохора (Hori et al. 2014; Shang et al. 2016). Ключевым элементом в формировании центромеры является включение в состав нуклеосом центромерного варианта гистона H3 — белка CENP-A (Müller, Almouzni 2017). На культурах клеток курицы было показано, что для его привлечения необходимо ацетилирование гистона H4 по лизиновым остаткам в 5 и 12 положениях (Shang et al. 2016). Эта модификация характерна именно для центромер, тогда как ацетилирование гистона H4 по лизину в 20 положении является характеристикой транскрипционно активного хроматина, однако эта модификация также оказалась необходимой как для формирования функциональных центромер у человека, так и у птиц (Hori et al. 2014; Sullivan, Karpen 2004). Ацетилирование гистона H4, а также невозможность осуществления типичных для гетерохроматина модификаций гистона H3 из-за структурных особенностей его центромерного варианта отличает центромерный хроматин от гетерохроматина, также возникающего в местах локализации тандемных повторов. Дополнительно отличие выражается в отсутствии метилирования цитозина в составе ДНК, участвующей в формировании кинетохора. Эпигенетические особенности являются необходимыми и достаточными для того, чтобы инициировать сборку кинетохора (Bergmann et al. 2011).

Для многих видов было показано наличие в составе центромерной ДНК специфического взаимодействия с другим кинетохорным белком CENP-B, опосредованного высококонсервативным элементом генома, получившим название CENP-B бокса (Masumoto et al. 1989). Собственно центромерные последовательности охарактеризованы у двух видов птиц: последовательность PR1 с длиной мономера 0,9 т. п. н. у скалистого голубя (Solovei et al. 1996), а также GGSAT с длиной мономера 1,2 т. п. н. у домашней курицы, впервые описанный как специфичный для хромосомы W повтор GG*EcoRI* (Saitoh et al. 1991). Обе эти тандемно повторяющиеся последовательности характеризуются наличием CENP-B бокса в составе мономера. У скалисто-

го голубя последовательность PR1 составляет 5% генома, что говорит о ее избыточности. Интересно, что у вяхиря этот повтор выходит за границы собственно центромеры и входит в состав перицентромерного гетерохроматина, тогда как у голубя только на хромосоме 2 он выходит за границы собственно центромеры (Solovei et al. 1996). Присутствующие блоки повтора *GG EcoRI* на хромосоме *W* курицы, составляющие 0,6 % генома, характеризуются наличием вырожденной CENP-B связывающей последовательности и более низким уровнем консерватизма (около 68%) (Saitoh et al. 1991), в то время как центромерные элементы этого повтора (собственно *GGSAT*) составляют не более 0,1% генома (Komissarov et al. 2018). Они характеризуются более высоким консерватизмом и наличием консенсусной последовательности, имеющей сходство с CENP-B боксом млекопитающих: *tTTTCGnnnnAaaCGG GaG*. Считается, что присутствие в составе центромерных последовательностей CENP-B бокса повышает точность связывания кинетохорных белков; в то же время функциональные центромеры формируются и без такого элемента. В этом случае решающую роль играет тандемный характер организации последовательности (Craig et al. 1999).

За исключением почкующихся дрожжей *Saccharomyces cerevisiae* и 25 других видов одноклеточных грибов, принадлежащих к разным родам, центромеры которых формируются исключительно за счет сиквенс-специфического взаимодействия с белками кинетохора (Guin et al. 2020; Ng et al. 1986), у подавляющего большинства изученных на сегодняшний момент эукариот центромерные районы состоят из тандемно повторенных элементов. Отдельные мономеры центромерных сателлитов в них перемежаются другими сателлитами, которые могут различаться между собой, однако последовательности NOR в пределах центромеры отдельной хромосомы имеют очень высокий уровень гомологии, который составляет 95–98%. Протяженность полей NOR может варьировать от 0,5 до 5 миллионов п. н., причем размер центромерного блока отдельной хромосом набора может существенно отличаться у различных индивидов. У курицы аутосомы 5 и 27, а также половая хромосома имеют очень маленькие центромеры (Shang et al. 2010). Индивидуальные особенности центромерных районов, сформированные благодаря участию разных повторяющихся элементов, обеспечивают функциональное сходство и поддерживают индивидуальные видо- и хромосомспецифичные различия.

Определенные последовательности отдельных NOR, сформированные из комбинации различных сателлитов, служат не только для формирования функциональной центромеры, прикрепляющейся к веретену деления, но и участвуют в организации хромосом в ядре, а также могут играть важную роль при конъюгации хромосом в мейозе и поддержании пространственной организации центромерных районов хромосом на разных этапах клеточного цикла (Aldrup-Macdonald et al. 2016; Ivanova et al. 2019; Saifitdinova et al. 2003). Высококонсервативная центромерная повторяющаяся последовательность, ответственная за когезию хроматина и поддержание пространственной организации хромосом, была обнаружена в геноме обыкновенного зяблика (Saifitdinova et al. 2001). Было показано, что она играет важную роль в пространственной организации хромосом не только в митотически делящихся клетках, но и в профазе первого деления мейоза (Saifitdinova et al. 2003). Эта последовательность характеризуется высоким консерватизмом не только в пределах генома одной особи, но и в пределах вида. В то же время у близких видов вьюрковых птиц обнаруживаются повторяющиеся элементы, имеющие сходство с нуклеотидной последовательностью FCP, но не имеющие центромерной локализации и характеризующиеся значительно меньшим уровнем консерватизма. В структуре мономера FCP из генома зяблика были обнаружены регуляторные элементы, характерные для лишенных ТАТА-боксов промоторов РНК-полимеразы II и транскрипционных факторов (Saifitdinova et al. 2001). Имеются данные о том, что транскрипция центромерных районов в интерфазе может играть важную роль для последующей корректной сегрегации хромосом (Sadeghi et al. 2014).

Тандемные повторы распространяются за границы функциональной центромеры и формируют перицентромерные гетерохроматиновые районы, стабилизирующие коровую область центромеры. Присутствующие в них сателлиты обеспечивают взаимодействие между сестринскими хроматидами, способствуя их правильному расхождению во время клеточных делений (Schalch, Steiner 2017; Yi et al. 2018). Увеличение числа копий мономеров в этих районах происходит преимущественно благодаря процессам репарации двунитевых разрывов и рекомбинации (McFarlane, Humphrey 2010; Talbert, Henikoff 2010), причем особенности организации перицентромерных районов разных хромосом у человека свидетельствуют о том, что в формировании гетерохроматиновых

блоков преобладают межхромосомные обмены (Warburton et al. 1996). Хроматин перицентромерных районов имеет типичные черты конститутивного гетерохроматина с высокой степенью метилирования ДНК, компактной упаковкой с участием гетерохроматинового белка HP1 и обогащением коровых нуклеосом гистонами H3, триметилированным по лизину во 2 и 9 положениях (Peng, Kargen 2008). Конститутивный гетерохроматин перицентромерных районов, фланкирующих центромеры, состоит из комбинации различных сателлитных последовательностей и мобильных элементов генома, которые характеризуются снижением уровня гомогенности повторяющихся элементов высокого порядка (Klein, O'Neill 2018). В то же время эти последовательности не являются транскрипционно инертными (Erukashvily, Ponomartsev 2013). Транскрипты этих последовательностей играют важную роль и могут принимать участие в установлении характерного для перицентромерного гетерохроматина статуса на ранних этапах эмбрионального развития (Casanova et al. 2013; Probst et al. 2010). Функциональное значение сателлитных последовательностей и других повторяющихся элементов генома, участвующих в формировании перицентромерного гетерохроматина, на сегодняшний момент не вызывает сомнения (Biscotti et al. 2015; Podgornaya et al. 2018).

Хотя геномы птиц характеризуются значительно более низким содержанием повторяющихся элементов, данные о составе и организации перицентромерного гетерохроматина очень ограничены. В его состав могут входить разные типы сателлитов. Повторяющиеся последовательности с длиной мономера от 42 до 191 п. н. описаны в составе перицентромерного гетерохроматина и относятся к числу мажорных повторов у таких видов, как кряква, зебровая амадина и некоторые виды попугаев (Madsen et al. 1992; Melters et al. 2013; Takki et al. 2022). У многих видов птиц в перицентромерном гетерохроматине обнаружены минисателлиты, такие как CNM у курицы, изначально выделенный из заякоренной на ядерной мембране фракции ДНК (Matzke et al. 1990). В составе перицентромерного гетерохроматина курицы описаны мономеры высокополиморфного инвертированного повтора PIR (Li et al. 2007; Wicker et al. 2005). Повторяющиеся последовательности с длиной мономера около нуклеосомного повтора, имеющие частичную гомологию с 41 п. н. повтором CNM, также описаны в центромерных районах хромосом у некоторых гусеобразных (Uno et al. 2019). Последовательности этих ми-

нисателлитов консервативны для курообразных и часто локализуются в составе перицентромерного гетерохроматина у индейки, виргинской куропатки и разных видов перепелов (Ishishita et al. 2014; Matzke et al. 1992; Tanaka et al. 2000; Wang et al. 2002; Yamada et al. 2002). В ходе эволюции последовательности перицентромерного гетерохроматина подвергаются гомогенизации, которая может происходить хромосомно-специфично (Ishishita et al. 2014), а может формировать сателлиты, характерные для макро- и микрохромосом (Kulak et al. 2018; Takki et al. 2022; Yamada et al. 2002) или, как у азиатского кеклика, приводить к образованию повторяющегося элемента ACH-Sau3AI, характерного для всех хромосом набора (Ishishita et al. 2014). Различные минисателлиты с мономером 41 п. н. также обнаружены в составе центромерных HOR и перицентромерного гетерохроматина всех хромосом у рыбного филина и длиннохвостой неясыти, у которых они формируют не обладающие высоким консерватизмом мономеры более высокого порядка длиной около 167–190 п. н. (Yamada et al. 2004). Мажорные сателлиты с мономерами большей длины входят в состав перицентромерного гетерохроматина у зебровой амадины, китайского и японского перепелов (Kulak et al. 2018; Takki et al. 2022; Yamada et al. 2002). В состав HOR перицентромерного гетерохроматина у японского перепела также могут входить последовательности мобильных элементов, а также фрагменты рибосомных генов (Saifitdinova et al. 2019). Интересно, что GC-состав последовательностей перицентромерного гетерохроматина варьирует в широких пределах (Melters et al. 2013), что также свидетельствует о высокой скорости эволюционирования этих районов хромосом у птиц.

Организация теломерных и субтеломерных районов

Теломеры обеспечивают полную репликацию на концах хромосом, защищают концы линейных хромосом от деградации, препятствуют дестабилизирующей активности незамкнутых нитей ДНК на концах хромосом за счет выступающего 3'-конца, обогащенного гуаниновыми остатками (Wang, Zakian 1990), обеспечивающего работу теломеразного комплекса и способного к формированию внутримолекулярного G-квадруплекса (Kar et al. 2018). У всех изученных видов птиц теломеры образованы каноническим для большинства позвоночных животных тандемным повтором (TTAGGG)_n (Meune et al.

1989), количество копий которого может различаться у представителей разных видов и в клетках разного типа дифференцировки, причем сокращение длины менее чем до 13 повторов, происходящее из-за невозможности репликации 5'-конца отстающей цепи, нарушает функциональность теломеры и ее способность защищать концы хромосом (Carper et al. 2007).

У некоторых видов птиц теломерный повтор относится к числу мажорных, а у домашней курицы он является наиболее представленным в геноме. На его долю приходится до 4% всей ДНК, что составляет более 30 м. п. н. (миллионов п. н.) на гаплоидный геном (Delany et al. 2003; Komissarov et al. 2018). Отчасти такое количество теломерных повторов в геномах птиц объясняется присутствием интерстициальных сайтов (TTAGGG)_n (Galkina et al. 2005; Liangouzov et al. 2002; Nanda, Schmid 1994; Saifitdinova et al. 2003; Solovei et al. 1994). Помимо распространения интерстициальных сайтов, в геноме курицы произошла экспансия теломерного повтора с терминальной локализацией, которая привела к формированию протяженных полей теломерного повтора на концах некоторых хромосом с формированием так называемых «мегателомер» (Delany et al. 2003; 2007). Такие последовательности не могут функционально заместить теломеры на концах хромосом, но они входят в состав субтеломерного гетерохроматина, повышая уровень его хромосомспецифичности за счет усложнения последовательности NOR отдельных субтеломерных районов.

Субтеломерные районы хромосом участвуют в обеспечении связи хромосом с элементами цитоскелета и поддержании определенной пространственной организации в интерфазном ядре благодаря заякориванию хромосом на его оболочке, а также играют важную роль в пространственной организации хромосом в ходе подготовки к клеточным делениям и синапсису гомологов в профазе первого деления мейоза (Calderon et al. 2015; Riethman et al. 2005). В отличие от теломер, которые не подвергаются метилированию из-за особенностей нуклеотидной последовательности, субтеломерные районы характеризуются снижением уровня гомогенности повторов и наличием CpG островков (Blasco 2007), а также присутствием характерных для гетерохроматина модификаций (Gonzalo et al. 2006; Vaquero-Sedas, Vega-Palas 2011). Существенную роль в формировании протяженных полей гетерохроматина субтеломерных районов хромосом играют повторяющиеся элементы разной природы, попадающие в эти районы гены

подвергаются сайленсингу, что хорошо описано как эффект положения гена (Weiler, Wakimoto 1995). Структурно-функциональные особенности субтеломерного гетерохроматина птиц свидетельствуют о том, что он может играть важную роль одновременно в обеспечении уникальности хромосомспецифичного синапсиса гомологов и поддержании структурного единообразия, необходимого для связи с комплексом белков ядерной оболочки.

Заключение

Геномы птиц характеризуются рядом особенностей, связанных с существенным сокращением их размеров, что привело к утрате существенной части повторяющихся последовательностей. Сохранившиеся некодирующие элементы у представителей этой группы имеют ряд характерных особенностей, которые позволяют использовать их в качестве модельных объектов для изучения отдельных классов таких последовательностей. Несмотря на существенное сокращение доли повторяющихся элементов, их распределение и организация свидетельствуют о возможности выполнения важных структурных и регуляторных функций. Исследование их организации может стать ключом к пониманию закономерностей функционирования некодирующих элементов генома, в том числе тех, значение которых до сих пор не удалось установить. Свойства таких последовательностей на молекулярном уровне могут определять их роль в регуляции активности генов и поддержании функциональной архитектуры ядра.

Конфликт интересов

Автор заявляет об отсутствии потенциального или явного конфликта интересов.

Conflict of Interest

The author declares that there is no conflict of interest, either existing or potential.

Соответствие принципам этики

Для подготовки статьи не проводились исследования с участием животных или людей.

Ethics Approval

No research involving animal or human subjects was conducted to prepare this article.

References

- Abe, H., Gemmell, N. J. (2014) Abundance, arrangement, and function of sequence motifs in the chicken promoters. *BMC Genomics*, vol. 15, article 900. <https://doi.org/10.1186/1471-2164-15-900> (In English)
- Aldrup-Macdonald, M. E., Kuo, M. E., Sullivan, L. L. et al. (2016) Genomic variation within alpha satellite DNA influences centromere location on human chromosomes with metastable epialleles. *Genome Research*, vol. 26, no. 10, pp. 1301–1311. <https://doi.org/10.1101/gr.206706.116> (In English)
- Andersson, L., Archibald, A. L., Bottema, C. D. et al. (2015) Coordinated international action to accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes project. *Genome Biology*, vol. 16, no. 1, article 57. <https://doi.org/10.1186/s13059-015-0622-4> (In English)
- Bellott, D. W., Skaletsky, H., Cho, T. J. et al. (2017) Avian W and mammalian Y chromosomes convergently retained dosage-sensitive regulators. *Nature Genetics*, vol. 49, no. 3, pp. 387–394. <https://doi.org/10.1038/ng.3778> (In English)
- Bergmann, J. H., Rodriguez, M. G., Martins, N. M. et al. (2011) Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *EMBO Journal*, vol. 30, no. 2, pp. 328–340. <https://doi.org/10.1038/emboj.2010.329> (In English)
- Bertocchi, N. A., Torres, F. P., del Valle Garnero, A. et al. (2017) Evolutionary history of the mariner element galluhop in avian genomes. *Mobile DNA*, vol. 8, article 11. <https://doi.org/10.1186/s13100-017-0094-z> (In English)
- Biscotti, M. A., Canapa, A., Forconi, M. et al. (2015) Transcription of tandemly repetitive DNA: Functional roles. *Chromosome Research*, vol. 23, no. 3, pp. 463–477. <https://doi.org/10.1007/s10577-015-9494-4> (In English)
- Blasco, M. A. (2007) The epigenetic regulation of mammalian telomeres. *Nature Reviews Genetics*, vol. 8, no. 4, pp. 299–309. <https://doi.org/10.1038/nrg2047> (In English)
- Bolshakova, E. V., Saifitdinova, A. F. (2020) Poisk genov-kandidatov na rol' mishenej regulyatsii polovoj differentsirovki u kuritsy [The search for candidate genes to be targeted for regulation of sex differentiation in chicken]. *Geny i kletki — Genes and Cells*, vol. 15, no. 3, p. 127. (accessed 05.03.2022). (In Russian)
- Calderon, M., Rey, M., Cabrera, A. et al. (2015) The subtelomeric region is important for chromosome recognition and pairing during meiosis. *Scientific Reports*, vol. 4, article 6488. <https://doi.org/10.1038/srep06488> (In English)
- Capper, R., Britt-Compton, B., Tankimanova, M. et al. (2007) The nature of telomere fusion and a definition of the critical telomere length in human cells. *Genes and Development*, vol. 21, no. 19, pp. 2495–2508. <https://doi.org/10.1101/gad.439107> (In English)
- Casanova, M., Pasternak, M., el Marjou, F. et al. (2013) Heterochromatin reorganization during early mouse development requires a single-stranded noncoding transcript. *Cell Report*, vol. 4, no. 6, pp. 1156–1167. <https://doi.org/10.1016/j.celrep.2013.08.015> (In English)
- Coullin, P., Bed'Hom, B., Candelier, J. J. et al. (2005) Cytogenetic repartition of chicken CR₁ sequences evidenced by PRINS in Galliformes and some other birds. *Chromosome Research*, vol. 13, no. 7, pp. 665–673. <https://doi.org/10.1007/s10577-005-1004-7> (In English)
- Craig, J. M., Earnshaw, W. C., Vagnarelli, P. (1999) Mammalian centromeres: DNA sequence, protein composition, and role in cell cycle. *Experimental Cell Research*, vol. 246, no. 2, pp. 249–262. <https://doi.org/10.1006/excr.1998.4278> (In English)
- Cui, J., Zhao, W., Huang, Z. et al. (2014) Low frequency of paleoviral infiltration across the avian phylogeny. *Genome Biology*, vol. 15, no. 12, article 539. <https://doi.org/genomebiology.com/2014/15/12/539> (In English)
- Delany, M. E., Daniels, L. M., Swanberg, S. E. et al. (2003) Telomeres in the chicken: Genome stability and chromosome ends. *Poultry Science*, vol. 82, no. 6, pp. 917–926. <https://doi.org/10.1093/ps/82.6.917> (In English)
- Delany, M. E., Gessaro T. M., Rodrigues, K. L. et al. (2007) Chromosomal mapping of chicken mega-telomere arrays to GGA9, 16, 28 and W using a cytogenomic approach. *Cytogenetic and Genome Research*, vol. 117, no. 1, pp. 54–63. <https://doi.org/10.1159/000103165> (In English)
- Deryusheva, S., Krasikova, A., Kulikova, T. et al. (2007) Tandem 41-bp repeats in chicken and Japanese quail genomes: FISH mapping and transcription analysis on lampbrush chromosomes. *Chromosoma*, vol. 116, no. 6, pp. 519–530. <https://doi.org/10.1007/s00412-007-0117-5> (In English)
- Dorshorst, B., Harun-Or-Rashid, M., Bagherpoor, A. J. et al. (2015) A genomic duplication is associated with ectopic oesodermin expression in the embryonic chicken comb and two duplexcomb phenotypes. *PLoS Genetics*, vol. 11, no. 3, article e1004947. <https://doi.org/10.1371/journal.pgen.1004947> (In English)
- Enukashvily, N. I., Ponomartsev, N. V. (2013) Mammalian satellite DNA: A speaking dumb. *Advances in Protein Chemistry and Structural Biology*, vol. 90, pp. 31–65. <https://doi.org/10.1016/B978-0-12-410523-2.00002-X> (In English)
- Galkina, S., Lukina, N., Zakharova, K. et al. (2005) Interstitial (TTAGGG)_n sequences are not hot spots of recombination in the chicken lampbrush macrochromosomes 1–3. *Chromosome Research*, vol. 13, no. 6, pp. 551–557. <https://doi.org/10.1007/s10577-005-0980-y> (In English)
- Gonzalo, S., Jaco, I., Fraga, M. F. et al. (2006) DNA methyltransferases control telomere length and telomere recombination in mammalian cells. *Nature Cell Biology*, vol. 8, no. 4, pp. 416–424. <https://doi.org/10.1038/ncb1386> (In English)

- Griffin, D. K., Robertson, L. B., Tempest, H. G. et al. (2008) Whole genome comparative studies between chicken and turkey and their implications for avian genome evolution. *BMC Genomics*, vol. 9, article 168. <https://doi.org/10.1186/1471-2164-9-168> (In English)
- Guin, K., Sreekumar, L., Sanyal, K. (2020) Implications of the Evolutionary Trajectory of Centromeres in the Fungal Kingdom. *Annual Review of Microbiology*, vol. 74, pp. 835–853. <https://doi.org/10.1146/annurev-micro-011720-122512> (In English)
- Haas, N. B., Grabowski, J. M., Sivitz, A. B. et al. (1997) Chicken Repeat 1 (CR1) elements, which define an ancient family of vertebrate non-LTR retrotransposons, contain two closely spaced open reading frames. *Gene*, vol. 197, no. 1–2, pp. 305–309. [https://doi.org/10.1016/s0378-1119\(97\)00276-x](https://doi.org/10.1016/s0378-1119(97)00276-x) (In English)
- Hori, T., Shang, W.-H., Toyoda, A. et al. (2014) Histone H4 Lys 20 monomethylation of the CENP-A nucleosome is essential for kinetochore assembly. *Developmental Cell*, vol. 29, no. 6, pp. 740–749. <https://doi.org/10.1016/j.devcel.2014.05.001> (In English)
- Ishishita, S., Tsuruta, Y., Uno, Y. et al. (2014) Chromosome size-correlated and chromosome size-uncorrelated homogenization of centromeric repetitive sequences in New World quails. *Chromosome Research*, vol. 22, no. 1, pp. 15–34. <https://doi.org/10.1007/s10577-014-9402-3> (In English)
- Itoh, Y., Arnold, A. P. (2005) Chromosomal polymorphism and comparative painting analysis in the zebra finch. *Chromosome Research*, vol. 13, no. 1, pp. 47–56. <https://doi.org/10.1007/s10577-005-6602-x> (In English)
- Itoh, Y., Kampf, K., Arnold, A. P. (2006) Comparison of the chicken and zebra finch Z chromosomes shows evolutionary rearrangements. *Chromosome Research*, vol. 14, no. 8, pp. 805–815. <https://doi.org/10.1007/s10577-006-1082-1> (In English)
- Itoh, Y., Kampf, K., Arnold, A. P. (2008) Molecular cloning of zebra finch W chromosome repetitive sequences: Evolution of the avian W chromosome. *Chromosoma*, vol. 117, no. 2, pp. 111–121. <https://doi.org/10.1007/s00412-007-0130-8> (In English)
- Itoh, Y., Mizuno, S. (2002) Molecular and cytological characterization of SspI-family repetitive sequence on the chicken W chromosome. *Chromosome Research*, vol. 10, no. 6, pp. 499–511. <https://doi.org/10.1023/a:1020944414750> (In English)
- Ivanova, N. G., Stefanova, V. N., Ostromyshenskii, D. I. et al. (2019) Tandem repeats in the genome of *Sus scrofa*, their localization on chromosomes and in the spermatogenic cell nuclei. *Russian Journal of Genetics*, vol. 55, no. 7, pp. 835–846. <https://doi.org/10.1134/S0016675819070075> (In English)
- John, St. J., Quinn, T. W. (2008) Identification of novel CR1 subfamilies in an avian order with recently active elements. *Molecular Phylogenetics and Evolution*, vol. 49, no. 3, pp. 1008–1014. <https://doi.org/10.1016/j.ympev.2008.09.020> (In English)
- Kapusta, A., Suh, A. (2017) Evolution of bird genomes—a transposons-eye view. *Annals of the New York Academy of Sciences*, vol. 1389, no. 1, pp. 164–185. <https://doi.org/10.1111/nyas.13295> (In English)
- Kar, A., Jones, N., Arat, N. Ö. et al. (2018) Long repeating (TTAGGG)_n single-stranded DNA self-condenses into compact beaded filaments stabilized by G-quadruplex formation. *Journal of Biological Chemistry*, vol. 293, no. 24, pp. 9473–9485. <https://doi.org/10.1074/jbc.RA118.002158> (In English)
- Klein, S. J., O'Neill, R. J. (2018) Transposable elements: Genome innovation, chromosome diversity, and centromere conflict. *Chromosome Research*, vol. 26, no. 1–2, pp. 5–23. <https://doi.org/10.1007/s10577-017-9569-5> (In English)
- Kodama, H., Saitoh, H., Tone, M. et al. (1987) Nucleotide sequences and unusual electrophoretic behavior of the W chromosome-specific repeating DNA units of the domestic fowl, *Gallus gallus domesticus*. *Chromosoma*, vol. 96, no. 1, pp. 18–25. <https://doi.org/10.1007/BF00285878> (In English)
- Komissarov, A. S., Galkina, S. A., Koshel, E. I. et al. (2018) New high copy tandem repeat in the content of the chicken W chromosome. *Chromosoma*, vol. 127, no. 1, pp. 73–83. <https://doi.org/10.1007/s00412-017-0646-5> (In English)
- Krasikova, A., Deryusheva, S., Galkina, S. et al. (2006) On the positions of centromeres in chicken lampbrush chromosomes. *Chromosome Research*, vol. 14, no. 7, pp. 777–789. <https://doi.org/10.1007/s10577-006-1085-y> (In English)
- Krasikova, A. V., Vasilevskaya, E. V., Gaginskaya, E. R. (2010) Chicken lampbrush chromosomes: Transcription of tandemly repetitive DNA sequences. *Russian Journal of Genetics*, vol. 46, no. 10, pp. 1173–1177. <https://doi.org/10.1134/S1022795410100078> (In English)
- Kretschmer, R., de Oliveira, T. D., de Oliveira Furo, I. et al. (2018) Repetitive DNAs and shrink genomes: A chromosomal analysis in nine Columbidae species (Aves, Columbiformes). *Genetics and Molecular Biology*, vol. 41, no. 1, pp. 98–106. <https://doi.org/10.1590/1678-4685-GMB-2017-0048> (In English)
- Kulak, M., Dyomin, A., Komissarov, A. et al. (2018) New pericentromeric repeat identified in the genome of *Japanese quail*. *Comparative Cytogenetics*, vol. 12, no. 3, pp. 337–338. <https://doi.org/10.3897/CompCytogen.v12i3.27748> (In English)
- Lawal, R. A., Al-Atiyat, R. M., Aljumaah, R. S. et al. (2018) Whole-genome resequencing of red junglefowl and indigenous village chicken reveal new insights on the genome dynamics of the species. *Frontiers in Genetics*, vol. 9, article 264. <https://doi.org/10.3389/fgene.2018.00264> (In English)

- Li, J., Wang, X., Leung, F. C. (2007) The intragenomic polymorphism of a partially inverted repeat (PIR) in *Gallus gallus domesticus*, potential role of inverted repeats in satellite DNAs evolution. *Gene*, vol. 387, no. 1–2, pp. 118–125. <https://doi.org/10.1016/j.gene.2006.08.033> (In English)
- Liangouzov, I. A., Derjusheva, S. E., Saifitdinova, A. F. et al. (2002) Monomers of a satellite DNA sequence of chaffinch (*Fringilla coelebs* L., Aves: Passeriformes) contain short clusters of the TTAGGG repeat. *Russian Journal of Genetics*, vol. 38, no. 12, pp. 1359–1364. <https://doi.org/10.1023/A:1021679520236> (In English)
- Liu, G. E., Jiang, L., Tian, F. et al. (2009) Calibration of mutation rates reveals diverse subfamily structure of galliform CR₁ repeats. *Genome Biology and Evolution*, vol. 1, pp. 119–130. <https://doi.org/10.1093/gbe/evp014> (In English)
- Lobanenkov, V. V., Nicolas, R. H., Adler, V. V. et al. (1990) A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene*, vol. 5, no. 12, pp. 1743–1753. PMID: 2284094 (In English)
- Madsen, C. S., de Kloet, D. H., Brooks, J. E. et al. (1992) Highly Repeated DNA sequences in birds: The structure and evolution of an abundant, tandemly repeated 190-bp DNA fragment in parrots. *Genomics*, vol. 14, no. 2, pp. 462–469. [https://doi.org/10.1016/s0888-7543\(05\)80242-3](https://doi.org/10.1016/s0888-7543(05)80242-3) (In English)
- Maroteaux, L., Heilig, R., Dupret, D. et al. (1983) Repetitive satellite-like sequences are present within or upstream from 3 avian protein-coding genes. *Nucleic Acids Research*, vol. 11, no. 5, pp. 1227–1243. <https://doi.org/10.1093/nar/11.5.1227> (In English)
- Masumoto, H., Masukata, H., Muro, Y. et al. (1989) A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *Journal of Cell Biology*, vol. 109, no. 5, pp. 1963–1973. <https://doi.org/10.1083/jcb.109.5.1963> (In English)
- Matzke, M. A., Varga, F., Berger, H. et al. (1990) A 41–42 bp tandemly repeated sequence isolated from nuclear envelopes of chicken erythrocytes is located predominantly on microchromosomes. *Chromosoma*, vol. 99, no. 2, pp. 131–137. <https://doi.org/10.1007/bf01735329> (In English)
- Matzke, A. J. M., Varga, F., Gruendler, P. et al. (1992) Characterization of a new repetitive sequence that is enriched on microchromosomes of turkey. *Chromosoma*, vol. 102, no. 1, pp. 9–14. <https://doi.org/10.1007/BF00352284> (In English)
- McFarlane, R. J., Humphrey, T. C. (2010) A role for recombination in centromere function. *Trends in Genetics*, vol. 26, no. 5, pp. 209–213. <https://doi.org/10.1016/j.tig.2010.02.005> (In English)
- Melters, D. P., Bradnam, K. R., Young, H. A. et al. (2013) Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biology*, vol. 14, no. 1, article R10. <https://doi.org/10.1186/gb-2013-14-1-r10> (In English)
- Meyne, J., Ratliff, R. L., Moyzis, R. K. (1989) Conservation of the human telomere sequence (TTAGGG)_n among vertebrates. *Proceedings of the National Academy of Sciences of the USA*, vol. 86, no. 18, pp. 7049–7053. <https://doi.org/10.1073/pnas.86.18.7049> (In English)
- Morris, K. M., Hindle, M. M., Boitard, S. et al. (2020) The quail genome: Insights into social behaviour, seasonal biology and infectious disease response. *BMC Biology*, vol. 18, no. 1, article 14. <https://doi.org/10.1186/s12915-020-0743-4> (In English)
- Mou, C., Pitel, F., Gourichon, D. (2011) Cryptic patterning of avian skin confers a developmental facility for loss of neck feathering. *PLOS Biology*, vol. 9, no. 3, article e1001028. <https://doi.org/10.1371/journal.pbio.1001028> (In English)
- Müller, S., Almouzni, G. (2017) Chromatin dynamics during the cell cycle. *Nature Reviews Genetics*, vol. 18, no. 3, pp. 192–208. <https://doi.org/10.1038/nrg.2016.157> (In English)
- Nakamura, D., Tiersch, T. R., Douglass, M. et al. (1990) Rapid identification of sex in birds by flow cytometry. *Cytogenetics and Cell Genetics*, vol. 53, no. 4, pp. 201–205. <https://doi.org/10.1159/000132930> (In English)
- Nanda, I., Schmid, M. (1994) Localization of the telomeric (TTAGGG)_n sequence in chicken (*Gallus Domesticus*) chromosomes. *Cytogenetics and Cell Genetics*, vol. 65, no. 3, pp. 190–193. <https://doi.org/10.1159/000133630> (In English)
- Ng, R., Ness, J., Carbon, J. (1986) Structural studies on centromeres in the yeast *Saccharomyces cerevisiae*. *Basic Life Sciences*, vol. 40, pp. 479–492. https://doi.org/10.1007/978-1-4684-5251-8_36 (In English)
- O'Meally, D., Patel, H. R., Stiglec, R. et al. (2010) Non-homologous sex chromosomes of birds and snakes share repetitive sequences. *Chromosome Research*, vol. 18, no. 7, pp. 787–800. <https://doi.org/10.1007/s10577-010-9152-9> (In English)
- Peng, J. C., Karpen, G. H. (2008) Epigenetic regulation of heterochromatic DNA stability. *Current Opinion in Genetics and Development*, vol. 18, no. 2, pp. 204–211. <https://doi.org/10.1016/j.gde.2008.01.021> (In English)
- Piégu, B., Arensburger, P., Beauclair, L. et al. (2020) Variations in genome size between wild and domesticated lineages of fowls belonging to the *Gallus gallus* species. *Genomics*, vol. 112, no. 2, pp. 1660–1673. <https://doi.org/10.1016/j.ygeno.2019.10.004> (In English)
- Podgornaya, O. I., Ostromyshenskii, D. I., Erukashvily, N. I. (2018) Who needs this junk, or genomic dark matter. *Biochemistry (Moscow)*, vol. 83, no. 4, pp. 450–466. <https://doi.org/10.1134/S0006297918040156> (In English)

- Probst, A. V., Okamoto, I., Casanova, M. et al. (2010) A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Developmental Cell*, vol. 19, no. 4, pp. 625–638. <https://doi.org/10.1016/j.devcel.2010.09.002> (In English)
- Riethman, H., Ambrosini, A., Paul, S. (2005) Human subtelomere structure and variation. *Chromosome Research*, vol. 13, no. 5, pp. 505–515. <https://doi.org/10.1007/s10577-005-0998-1> (In English)
- Rodionov, A. V. (1996) Micro versus macro: A review of structure and functions of avian micro- and macrochromosomes. *Russian Journal of Genetics*, vol. 32, no. 5, pp. 517–527. (In English)
- Sadeghi, L., Siggins, L., Svensson, J. P. et al. (2014) Centromeric histone H2B monoubiquitination promotes noncoding transcription and chromatin integrity. *Nature Structural and Molecular Biology*, vol. 21, no. 3, pp. 236–243. <https://doi.org/10.1038/nsmb.2776> (In English)
- Saifitdinova, A., Derjusheva, S., Krasikova, A. et al. (2003) Lampbrush chromosomes of the chaffinch (*Fringilla coelebs* L.). *Chromosome Research*, vol. 11, no. 2, pp. 99–113. <https://doi.org/10.1023/A:1022859713777> (In English)
- Saifitdinova, A. F., Derjusheva, S. E., Malykh, A. G. et al. (2001) Centromeric tandem repeat from the chaffinch genome: Isolation and molecular characterization. *Genome*, vol. 44, no. 1, pp. 96–103. <https://doi.org/10.1139/gen-44-1-96> (In English)
- Saifitdinova, A. F., Galkina, S. A., Koshel, E. I., Gaginskaya, E. R. (2016) Rol' povtoryayushchikhsya posledovatel'nostej v evolyutsii polovykh khromosom u ptits [The role of repetitive sequences in the evolution of sex chromosomes in birds]. *Tsitologiya*, vol. 58, no. 5, pp. 393–398. (In Russian)
- Saifitdinova, A., Galkina, S., Kulak, M. et al. (2019) Nucleolus organizers transposition in the genome of Japanese quail. *Molecular Cytogenetics*, vol. 12, no. 30, pp. 63. (In English)
- Saifitdinova, A. F., Timofejeva, L. P., Zhurov, V. G., Gaginskaya, E. R. (2000) A highly repeated FCP centromeric sequence from chaffinch (*Fringilla coelebs*) genome is revealed within interchromosomal connectives during. *Tsitologiya*, vol. 42, no. 6, pp. 588–593. PMID: [10953865](https://pubmed.ncbi.nlm.nih.gov/10953865/) (In English)
- Saitoh, Y., Mizuno, S. (1992) Distribution of XhoI and EcoRI family repetitive DNA-sequences into separate domains in the chicken W-chromosome. *Chromosoma*, vol. 101, no. 8, pp. 474–477. <https://doi.org/10.1007/bf00352469> (In English)
- Saitoh, Y., Saitoh, H., Ohtomo, K. et al. (1991) Occupancy of the majority of DNA in the chicken W chromosome by bent-repetitive sequences. *Chromosoma*, vol. 101, no. 1, pp. 32–40. <https://doi.org/10.1007/BF00360684> (In English)
- Schalch, T., Steiner, F. A. (2017) Structure of centromere chromatin: from nucleosome to chromosomal architecture. *Chromosoma*, vol. 126, no. 4, pp. 443–455. <https://doi.org/10.1007/s00412-016-0620-7> (In English)
- Schmid, M., Loser, C., Schmidtke, J. et al. (1982) Evolutionary conservation of a common pattern of activity of nucleolus organizers during spermatogenesis in vertebrates. *Chromosoma*, vol. 86, no. 2, pp. 149–179. <https://doi.org/10.1007/bf00288674> (In English)
- Seol, D., Ko, B. J., Kim, B. et al. (2019) Identification of copy number variation in domestic chicken using whole-genome sequencing reveals evidence of selection in the genome. *Animals*, vol. 9, no. 10, article 809. <https://doi.org/10.3390/ani9100809> (In English)
- Shang, W.-H., Hori, T., Toyoda, A. et al. (2010) Chickens possess centromeres with both extended tandem repeats and short non-tandem-repetitive sequences. *Genome Research*, vol. 20, no. 9, pp. 1219–1228. <https://doi.org/10.1101/gr.106245.110> (In English)
- Shang, W.-H., Hori, T., Westhorpe, F. G. et al. (2016) Acetylation of histone H4 lysine 5 and 12 is required for CENP-A deposition into centromeres. *Nature Communications*, vol. 7, article 13465. <https://doi.org/10.1038/ncomms13465> (In English)
- Shinomiya, A., Kayashima, Y., Kinoshita, K. et al. (2012) Gene duplication of endothelin 3 is closely correlated with the hyperpigmentation of the internal organs (fibromelanosis) in silky chickens. *Genetics*, vol. 190, no. 2, pp. 627–638. <https://doi.org/10.1534/genetics.111.136705> (In English)
- Skinner, B. M., al Mutery, A., Smith, D. et al. (2014) Global patterns of apparent copy number variation in birds revealed by cross-species comparative genomic hybridization. *Chromosome Research*, vol. 22, no. 1, pp. 59–70. <https://doi.org/10.1007/s10577-014-9405-0> (In English)
- Solovei, I., Gaginskaya, E., Macgregor, H. C. (1994) The arrangement and transcription of telomere DNA sequences at the ends of lampbrush chromosomes of birds. *Chromosome Research*, vol. 2, no. 6, pp. 460–470. <https://doi.org/10.1007/bf01552869> (In English)
- Solovei, I., Joffe, B. I., Gaginskaya, E. R. et al. (1996) Transcription of lampbrush chromosomes of a centromerically localized highly repeated DNA in pigeon (*Columba*) relates to sequence arrangement. *Chromosome Research*, vol. 4, no. 8, pp. 588–603. <https://doi.org/10.1007/bf02261722> (In English)
- Sullivan, B. A., Karpen, G. H. (2004) Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nature Structural and Molecular Biology*, vol. 11, no. 11, pp. 1076–1083. <https://doi.org/10.1038/nsmb845> (In English)

- Takki, O., Komissarov, A., Kulak, M. et al. (2022) Identification of centromere-specific repeats in the zebra finch genome. *Cytogenetic and Genome Research*, vol. 162, pp. 55–63. <https://www.doi.org/10.1159/000521716> (In English)
- Talbert, P. B., Henikoff, S. (2010) Centromeres convert but don't cross. *PLOS Biology*, vol. 8, no. 3, article e1000326. <https://doi.org/10.1371/journal.pbio.1000326> (In English)
- Tanaka, K., Suzuki, T., Nojiri, T. et al. (2000) Characterization and chromosomal distribution of a novel satellite DNA sequence of Japanese quail (*Coturnix coturnix japonica*). *Journal of Heredity*, vol. 91, no. 5, pp. 412–415. <https://doi.org/10.1093/jhered/91.5.412> (In English)
- Teranishi, M., Shimada, Y., Hori, T. et al. (2001) Transcripts of the MHM region of the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT locus. *Chromosome Research*, vol. 9, no. 2, pp. 147–165. <https://doi.org/10.1023/A:1009235120741> (In English)
- Torgasheva, A. A., Malinovskaya, L. P., Zadesenets, K. S. et al. (2019) Germline-restricted chromosome (GRC) is widespread among songbirds. *Proceedings of the National Academy of Sciences of the USA*, vol. 116, no. 24, pp. 11845–11850. <https://doi.org/10.1073/pnas.1817373116> (In English)
- Treplin, S., Tiedemann, R. (2007) Specific chicken repeat 1 (CR1) retrotransposon insertion suggests phylogenetic affinity of rockfowls (genus *Picathartes*) to crows and ravens (*Corvidae*). *Molecular Phylogenetics and Evolution*, vol. 43, no. 1, pp. 328–337. <https://doi.org/10.1016/j.ympev.2006.10.020> (In English)
- Uno, Y., Nishida, C., Hata, A. et al. (2019) Molecular cytogenetic characterization of repetitive sequences comprising centromeric heterochromatin in three Anseriformes species. *PLOS One*, vol. 14, no. 3, article e0214028. <https://doi.org/10.1371/journal.pone.0214028> (In English)
- Vaquero-Sedas, M. I., Vega-Palas, M. A. (2011) On the chromatin structure of eukaryotic telomeres. *Epigenetics*, vol. 6, no. 9, pp. 1055–1058. <https://doi.org/10.4161/epi.6.9.16845> (In English)
- Wallen, M. J., Keinainen, R. A., Kulomaa, M. S. (1996) Two chicken repeat one (CR1) elements lacking a silencer-like region upstream of the chicken avidin-related genes *Avr4* and *Avr5*. *Biochimica et Biophysica Acta (BBA)—Gene Structure and Expression*, vol. 1308, no. 3, pp. 193–196. [https://doi.org/10.1016/0167-4781\(96\)00119-4](https://doi.org/10.1016/0167-4781(96)00119-4) (In English)
- Wallis, J. W., Aerts, J., Groenen, A. M. et al. (2004) A physical map of the chicken genome. *Nature*, vol. 432, no. 7018, pp. 793–800. <https://doi.org/10.1038/nature03030> (In English)
- Wang, S., Zakian, V. (1990) Telomere telomere recombination provides an express pathway for telomere acquisition. *Nature*, vol. 345, no. 6274, pp. 456–458. <https://doi.org/10.1038/345456a0> (In English)
- Wang, X., Byers, S. (2014) Copy number variation in chickens: A review and future prospects. *Microarrays*, vol. 3, no. 1, pp. 24–38. <https://doi.org/10.3390/microarrays3010024> (In English)
- Wang, X., Li, J., Leung, F. C. (2002) Partially inverted tandem repeat isolated from pericentric region of chicken chromosome 8. *Chromosome Research*, vol. 10, no. 1, pp. 73–82. <https://doi.org/10.1023/a:1014226412339> (In English)
- Warburton, P. E., Haaf, T., Gosden, J. et al. (1996) Characterization of a chromosome-specific chimpanzee alpha satellite subset: evolutionary relationship to subsets on human chromosomes. *Genomics*, vol. 33, no. 2, pp. 220–228. <https://doi.org/10.1006/geno.1996.0187> (In English)
- Warren, W. C., Hillier, L. W., Tomlinson, C. et al. (2017) A new chicken genome assembly provides insight into avian genome structure. *G3: Genes, Genomes, Genetics*, vol. 7, no. 1, pp. 109–117. <https://doi.org/10.1534/g3.116.035923> (In English)
- Watanabe, M., Nikaido, M., Tsuda, T. T. et al. (2006) The rise and fall of the CR1 subfamily in the lineage leading to penguins. *Gene*, vol. 365, pp. 57–66. <https://doi.org/10.1016/j.gene.2005.09.042> (In English)
- Weiler, K., Wakimoto, B. (1995) Heterochromatin and gene expression in *Drosophila*. *Annual Review of Genetics*, vol. 29, no. 1, pp. 577–605. <https://doi.org/10.1146/annurev.ge.29.120195.003045> (In English)
- Weissensteiner, M. H., Suh, A. (2019) Repetitive DNA: The dark matter of avian genomics. In: R. Kraus (ed.). *Avian genomics in ecology and evolution*. Cham: Springer Publ., pp. 93–150. https://doi.org/10.1007/978-3-030-16477-5_5 (In English)
- Wicker, T., Robertson, J. S., Schulze, S. R. et al. (2005) The repetitive landscape of the chicken genome. *Genome Research*, vol. 15, no. 1, pp. 126–136. <https://doi.org/10.1101/gr.2438005> (In English)
- Wicker, T., Sabot, F., Hua-Van, A. et al. (2007) A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*, vol. 8, no. 12, pp. 973–982. <https://doi.org/10.1038/nrg2165> (In English)
- Wright, D., Boije, H., Meadows, J. R. S. et al. (2009) Copy number variation in intron 1 of *SOX5* causes the pea-comb phenotype in chickens. *PLOS Genetics*, vol. 5, no. 6, article e1000512. <https://doi.org/10.1371/journal.pgen.1000512> (In English)
- Yamada, K., Nishida-Umehara, C., Ishijima, J. et al. (2006) A novel family of repetitive DNA sequences amplified site-specifically on the W chromosomes in Neognathous birds. *Chromosome Research*, vol. 14, no. 6, pp. 613–627. <https://doi.org/10.1007/s10577-006-1071-4> (In English)
- Yamada, K., Nishida-Umehara, C., Matsuda, Y. (2004) A new family of satellite DNA sequences as a major component of centromeric heterochromatin in owls (*Strigiformes*). *Chromosoma*, vol. 112, no. 6, pp. 277–287. <https://doi.org/10.1007/s00412-003-0267-z> (In English)

- Yamada, K., Shibusawa, M., Tsudzuki, M. et al. (2002) Molecular cloning and characterization of novel centromeric repetitive DNA sequences in the blue-breasted quail (*Coturnix chinensis*, Galliformes). *Cytogenetic Genome Research*, vol. 98, no. 4, pp. 255–261. <https://doi.org/10.1159/000071044> (In English)
- Yi, Q., Chen, Q., Liang, C. et al. (2018) HP₁ links centromeric heterochromatin to centromere cohesion in mammals. *EMBO Reports*, vol. 19, no. 4, article e45484. <https://doi.org/10.15252/embr.201745484> (In English)
- Zhang, G., Jarvis, E. D., Gilbert, M. T. P. (2014a) A flock of genomes. *Science*, vol. 346, no. 6215, pp. 1308–1309. <https://doi.org/10.1126/science.346.6215.1308> (In English)
- Zhang, G., Li, C., Li, Q. et al. (2014b) Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, vol. 346, no. 6215, pp. 1311–1320. <https://doi.org/10.1126/science.1251385>
- Zopl, D., Dineva, B., Betz, H. et al. (1990) Isolation of the chicken middle-molecular weight neurofilament (NF-M) gene and characterization of its promoter. *Nucleic Acids Research*, vol. 18, no. 3, pp. 521–529. <https://doi.org/10.1093/nar/18.3.521> (In English)