

Полногеномные данные подтверждают присутствие кроссинговера у гибридов соболя (*M. zibellina*) и лесной куницы (*M. martes*)

Томаровский А.^{1,2*}, Тотиков А.^{1,2}, Беклемишева В.¹, Перельман П.¹, Сердюкова Н.¹, Бульонкова Т.³, Сидоров М.⁴, Мамаев Н.⁴, Охлопков И.⁴, Мухачева А.⁵, Коняева К.⁶, Абрамов А.⁷, Графодатский А.¹, Кливер С.⁸

¹ Институт молекулярной и клеточной биологии СО РАН, Новосибирск, Россия

² Новосибирский государственный университет, Новосибирск, Россия

³ Институт систем информатики им. А.П. Еришова СО РАН, Новосибирск, Россия

⁴ Институт биологических проблем криолитозоны СО РАН, Якутск, Россия

⁵ Сихотэ-Алинский государственный природный биосферный заповедник им. К.Г. Абрамова, п. Терней, Россия

⁶ Зоопарк «Лесная сказка», Барнаул, Россия

⁷ Зоологический институт РАН, Санкт-Петербург, Россия

⁸ Независимый исследователь, Копенгаген, Дания

* andrey.tomarovsky@gmail.com

Ключевые слова: гибридизация; анализ генетического смешения

Мотивация и цель: Соболя (*M. zibellina*) и лесная куница (*M. martes*) являются близкородственными представителями рода *Martes* (сем. Mustelidae). В зоне перекрывания ареалов этих пушных видов происходит образование гибридов – кидасов, которые характеризуются промежуточными фенотипическими признаками родительских видов [1]. Внутрипопуляционный и межвидовой анализ генетической структуры этих симпатрических видов ранее проводился с использованием подходов низкого разрешения, основанных на анализе митохондриальных и отдельных ядерных маркеров [2, 3]. Используя собственные полногеномные данные 7 образцов *M. martes*, 13 образцов *M. zibellina* и 12 образцов их предполагаемых гибридов, а также полногеномные данные 1 опубликованного образца китайского подвида соболя, мы выполнили ADMIXTURE-анализ исследуемых особей.

Методы и алгоритмы: Поиск, маскирование и фильтрация SNP аутосом и псевдоаутосомного района выполнены при помощи BCFTOOLS, BEDTOOLS, BEDOPS и PLINK [4–7]. В качестве референса использована общедоступная геномная сборка соболя хромосомного уровня [8]. Анализ главных компонент (PCA) проведен с помощью PLINK. Визуализация полученного результата и вычисление объясненной дисперсии каждой компоненты выполнены средствами языка программирования R. Анализ генетического смешения образцов выполнен с использованием ADMIXTURE [8] для 1–6 предполагаемых генетических кластеров исходных популяций (K). Анализ проводился в трехкратной повторности с вычислением ошибок перекрестной валидации и значений среднего попарного сходства. Локальный ADMIXTURE анализ выполнен путем отдельных запусков ADMIXTURE (K = 2) с использованием неперекрывающихся скользящих окон

размером 1 млн п. н. Визуализация выполнена средствами языка программирования Python.

Результаты: На графике PCA (рис. 1, А) исследуемые образцы сформировали две группы относительно первой главной компоненты. Левая группа сформирована лесными кунцами. В свою очередь правая группа, особи из которой больше рассредоточены по вертикали, состоит из соболей. Образцы T87 и T84, а также в меньшей степени T18 расположились практически между двумя группами, что является признаком наиболее выраженного гибридного статуса этих особей.

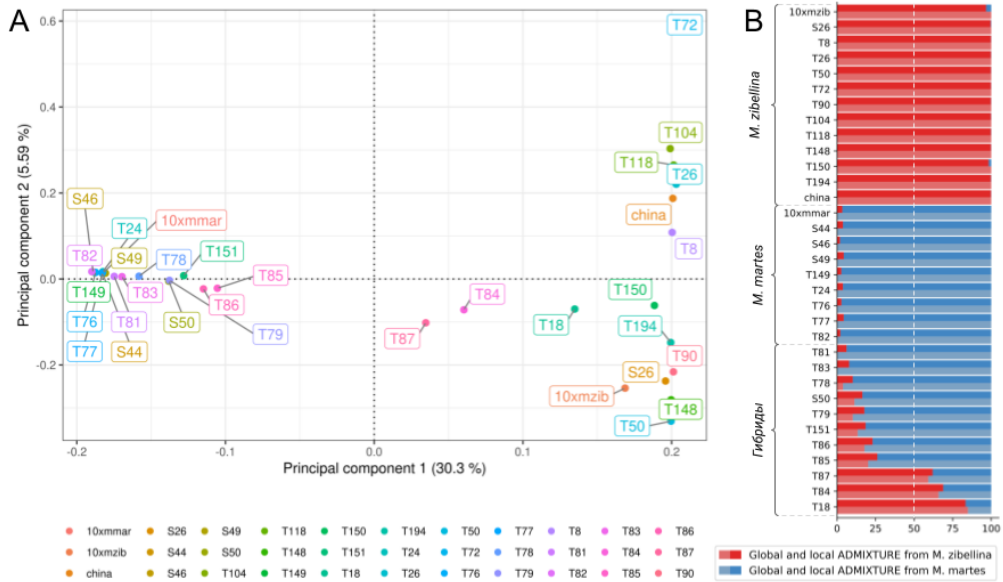


Рис. 1. А – график PCA исследуемых образцов соболя, лесной кунцы и их предполагаемых гибридов; В – результаты глобального и локального ADMIXTURE-анализа (K = 2)

Результаты глобального и локального ADMIXTURE-анализа (K = 2) визуализированы на рис. 1, В. Глобальный ADMIXTURE-анализ выявил наличие интрогрессии в 9 образцах, доля которой колеблется в диапазоне от 3.9 % (T78) до 40.8 % (T87). Следов интрогрессии среди остальных образцов обнаружено не было. Локальный ADMIXTURE-анализ позволил дополнительно выявить интрогрессию в образцах T81 и T83, равную 6.17 и 7.86 % соответственно. В общей сложности мы обнаружили следы интрогрессии от 1 до 5 % в 11 образцах (включая референсные образцы родительских видов).

Локальный ADMIXTURE-анализ на уровне каждой хромосомы дополнительно позволил нам установить мозаичность хромосом у гибридных особей. В качестве примера на рис. 2, А показан гибридный образец T18, на 10-й хромосоме которого выявлены как участки от каждого из родительских видов, так и гибридные участки от обоих из них. Мозаичность представленного и других гибридных образцов проявляется в разных пропорциях и может проявляться на разных хромосомных участках, что свидетельствует о том, что кроссинговер в гибридных образцах не подавлен. Дополнительно показаны две подтвержденных лесных кунцы из Свердловской (рис. 2, В) и Архангельской (рис. 2, С) областей с интрогрессией со

стороны соболя в 3.4 и 4.3 % соответственно. Подобные следы интрогрессии у типичных особей лесной куницы говорят о расширении западной части ареала соболя, а также, вероятно, могут быть следствием побегов соболей со звероферм.

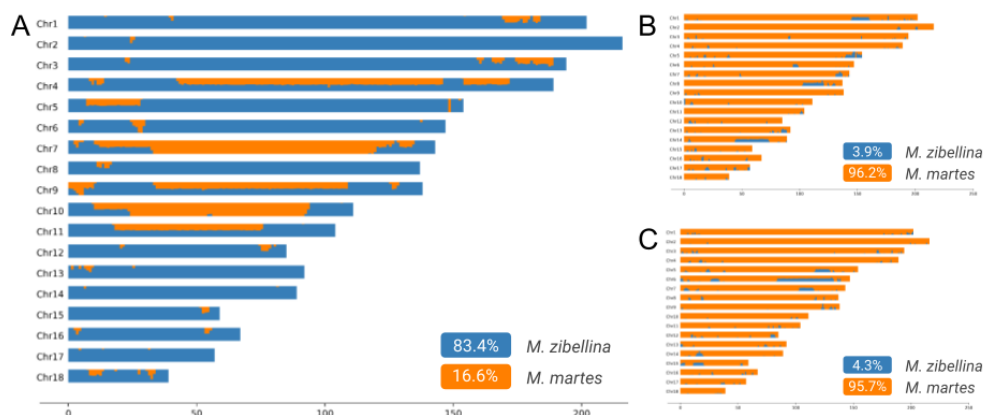


Рис. 2. Результаты локального ADMIXTURE-анализа каждой хромосомы: А – для гибридного образца T18; В и С – для подтвержденных образцов лесной куницы

Выводы: Нам удалось выявить взаимные интрогрессии на уровне каждой хромосомы между соболем и лесной куницей. Активный двунаправленный поток генов наблюдается в геномах особей как из зоны симпатрии, так и в удаленных от этой области частях ареалов родительских видов. Среди образцов лесных куниц не обнаружено ни одной особи без интрогрессии от соболя, что является свидетельством расширения западной части ареала соболя. Из рассмотренных 12 образцов предполагаемых кидасов подтвержденными оказались только 8, тогда как оставшиеся 4 являются лесными куницами с интрогрессией со стороны соболя менее условного порога в 5 %. Полученные результаты подтверждают, что кроссинговер у гибридных особей не подавлен.

Благодарности: Выражаем благодарность Александру Мигуру, Валерию Зелепухину, Виктору Хазагаеву, Валерию Середкину, Валентину Слободенюку, Сергею Крехнову, Ирине Шуваловой, Светлане Зубакиной, Алексею Лахтину, Сергею Цыпленкову, Сергею Писареву, Павлу Резниченко и Валерию Жданюк за помощь в сборе образцов.

Финансирование: Исследование поддержано грантом РФФ № 19-14-00034-П.

Whole-genome data confirm the presence of crossover in hybrids of sable (*M. zibellina*) and pine marten (*M. martes*)

Tomarovsky A.^{1, 2*}, Totikov A.^{1, 2}, Beklemisheva V.¹, Perelman P.¹, Serdyukova N.¹, Bulyonkova T.³, Sidorov M.⁴, Mamaev N.⁴, Okhlopov I.⁴, Mukhacheva A.⁵, Koniaeva K.⁶, Abramov A.⁷, Graphodatsky A.¹, Kliver S.⁸

¹Institute of Molecular and Cellular Biology, SB RAS, Novosibirsk, Russia

²Novosibirsk State University, Novosibirsk, Russia

³A.P. Ershov Institute of Informatics Systems, SB RAS, Novosibirsk, Russia

⁴Institute for Biological Problems of Cryolithozone, SB RAS, Yakutsk, Russia

⁵ Sikhote-Alin State Nature Biosphere Reserve named after K.G. Abramov, Terney, Russia

⁶ “Lesnaya skazka” Zoo, Barnaul, Russia

⁷ Zoological Institute RAS, St. Petersburg, Russia

⁸ Independent scientist, Copenhagen, Denmark

* andrey.tomarovsky@gmail.com

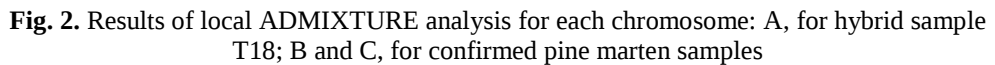
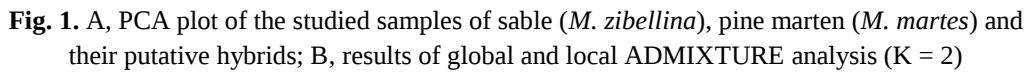
Key words: hybridization; genetic admixture analysis

Motivation and Aim: Sable (*M. zibellina*) and pine marten (*M. martes*) are closely related species of the genus *Martes* (family Mustelidae). In the area of overlapping ranges of these fur species, hybrids – kidases – are formed, which are characterized by intermediate phenotypic traits of the parental species [1]. Intrapopulation and interspecific analysis of the genetic structure of these sympatric species was previously performed using low-resolution approaches based on the analysis of mitochondrial and individual nuclear markers [2, 3]. Using the obtained whole-genome data of 7 samples of *M. martes*, 13 samples of *M. zibellina* and 12 samples of their putative hybrids, as well as the whole-genome data of 1 published sample of the Chinese subspecies of sable, we performed ADMIXTURE analysis of the studied individuals.

Methods and Algorithms: SNPs of autosomes and pseudoautosomal region were identified, masked, and filtered using BCFTOOLS, BEDTOOLS, BEDOPS, and PLINK [4–7]. The publicly available chromosome-level genome assembly of sable was used as a reference [8]. Principal Component Analysis (PCA) was conducted using PLINK. Visualization of the results and calculation of the explained variance for each component were performed using R. Genetic admixture analysis of the samples was performed using ADMIXTURE [8] for 1–6 assumed genetic clusters of ancestral populations (K). The analysis was performed in 3-fold repetition with calculation of cross validation errors and average pairwise similarity values. Local ADMIXTURE analysis was performed by separate runs of ADMIXTURE (K = 2) using non-overlapping sliding windows of 1 Mbp. Visualization was performed using Python.

Results: In the PCA plot (Fig. 1A), the studied samples formed two groups relative to the 1st principal component. The left group is formed by pine martens, while the right group, whose individuals are more dispersed vertically, consists of sables. Samples T87 and T84, and to a lesser extent T18, are located almost between the two groups, indicating a more pronounced hybrid status of these individuals.

The results of global and local ADMIXTURE analysis (K = 2) are visualized in Fig. 1B. Global ADMIXTURE analysis revealed the presence of introgression in 9 samples, ranging from 3.9 % (T78) to 40.8 % (T87). No traces of introgression were found among the remaining samples. Local ADMIXTURE analysis additionally identified introgression in samples T81 and T83 equal to 6.17 and 7.86 % respectively. In total, we detected traces of introgression from 1 to 5 % in 11 samples (including reference samples of the parental species).



895

pine martens indicate the expansion of the western part of the sable's range and may also be a result of sable escapes from fur farms.

Conclusion: We succeeded in identifying mutual introgressions at the level of each chromosome between sable and pine marten. Active bidirectional gene flow was observed in the genomes of individuals both from the sympatric zone and in remote parts of the ranges of the parental species. Among the pine marten samples, not a single individual without sable introgression was found, indicating the expansion of the western part of the sable's range. Out of the 12 examined suspected hybrids, only 8 were confirmed, while the remaining 4 are pine martens with introgression below the arbitrary threshold of 5 %. The results confirm that crossover is not suppressed in hybrid individuals.

Acknowledgements: We acknowledge Alexander M. Migura, Valerii A. Zelepukhin, Victor D. Hazagaev, Valeriy V. Seredkin, Valentin S. Slobodenuk, Sergei D. Krekhnov, Irina A. Shuvalova, Svetlana V. Zubakina, Aleksey A. Lakhtin, Sergei M. Tsiplenkov, Sergei Pisarev, Pavel Reznichenko and Valeriy T. Zhdanuk for helping with sample gathering.

Funding: The study is supported by the RSF grant No. 19-14-00034-P.

Список литературы/References

1. Юргенсон П. Кидас – гибрид соболя и куницы. *Труды Печоро-Илычского заповедника*. 1947;5:145-178
[Yurgenson P. Kidas – a hybrid of sable and marten. *Trudy Pechoro-Ilychskogo zapovednika*. 1947;5:145-178 (in Russian)]
2. Li B. et al. Phylogeography of subspecies of the sable (*Martes zibellina* L.) based on mitochondrial genomes: implications for evolutionary history. *Mamm Biol.* 2021;101(1):105-120. doi 10.1007/s42991-020-00092-0
3. Пищулина С.Л. Взаимодействие популяций лесной куницы и соболя в зоне симпатрии: генетический аспект. М.: ИПЭЭ РАН, 2013
[Pishulina S.L. Interaction of pine marten and sable populations in the sympatry zone: genetic aspect. М.: IPEE RAS, 2013 (in Russian)]
4. Danecek P. et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021;10(2):giab008. doi 10.1093/gigascience/giab008
5. Quinlan A.R., Hall I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841-842. doi 10.1093/bioinformatics/btq033
6. Neph S. et al. BEDOPS: high-performance genomic feature operations. *Bioinformatics*. 2012;28(14):1919-1920. doi 10.1093/bioinformatics/bts277
7. Chang C.C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*. 2015;4(1):7. doi 10.1186/s13742-015-0047-8
8. *Martes_zibellina*. DNA Zoo. https://www.dnazoo.org/assemblies/Martes_zibellina (accessed March 20, 2024) [Electronic resource]
9. Alexander D.H., Novembre J., Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19(9):1655-1664. doi 10.1101/gr.094052.109