

## ИДЕНТИФИКАЦИЯ ПОПУЛЯЦИЙ СОБОЛЯ (*Martes zibellina* L.) ЮЖНОЙ ЧАСТИ АРЕАЛА ВИДА

© 2021 г. С. Н. Каштанов<sup>1, \*</sup>, М. В. Шитова<sup>1</sup>, М. М. Сомова<sup>4</sup>,  
П. А. Филимонов<sup>1</sup>, О. Ю. Тютеньков<sup>5</sup>, А. А. Ходырева<sup>1</sup>, Г. Р. Свищева<sup>1, 3</sup>,  
И. Г. Мещерский<sup>2</sup>, Ю. А. Столповский<sup>1</sup>, В. В. Рожнов<sup>2</sup>

<sup>1</sup>Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, 119991 Россия

<sup>2</sup>Институт проблем экологии и эволюции им. А.Н. Северцова Российской академии наук, Москва, 119071 Россия

<sup>3</sup>Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения  
Российской академии наук, Новосибирск, 630090 Россия

<sup>4</sup>Федеральный исследовательский центр Всероссийский институт животноводства им. академика Л.К. Эрнста,  
Московская область, Подольск, п. Дубровицы, 142132 Россия

<sup>5</sup>Национальный исследовательский Томский государственный университет, Томск, 634050 Россия

\*e-mail: snkashtanov@mail.ru

Поступила в редакцию 25.03.2021 г.

После доработки 15.04.2021 г.

Принята к публикации 20.04.2021 г.

Для исследования современной структуры генофонда соболя южной части ареала использована панель из 15 микросателлитных локусов. Определен высокий уровень дифференциации популяций. Выявлено существование нескольких популяционных кластеров соболя на обширной территории от Западной Сибири до Забайкалья. Эта часть ареала вида отличается высокой численностью соболя, другая особенность — разрешенные к промыслу территории чередуются со значительным числом заповедных. Определены возможности идентификации особей по популяционной и региональной принадлежности. В исследовании дана оценка надежности прогноза популяционной принадлежности (SCOR-коэффициент) особей к объединенной выборке, показавшая, что при увеличении числа локусов с 8 до 15 значимо возрастает точность (количество правильно идентифицированных особей) и надежность идентификации в среднем для каждой отдельной особи. Результаты исследования дополняют референтную базу генетических данных популяций географических регионов вида, позволяют решить ряд проблем, связанных с промыслом на территориях, граничащих с заповедными, что способствует сохранению внутривидового разнообразия соболя.

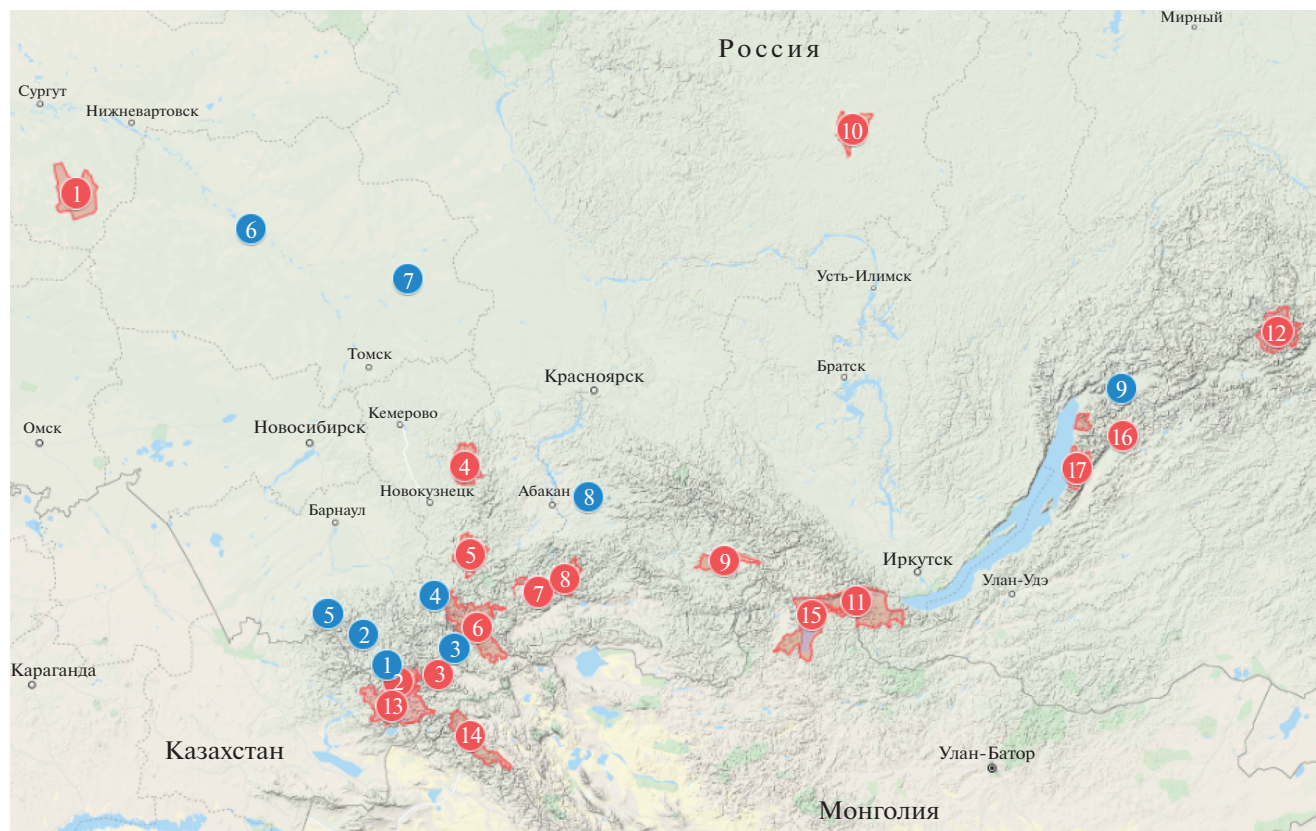
**Ключевые слова:** географические популяции соболя, интродукция, промысел, микросателлитные локусы, филогенетический анализ, генетическая структура, ДНК-идентификация.

**DOI:** 10.31857/S0016675821100052

Ареал соболя Евразии занимает обширную территорию от западных склонов гор Урала до побережья Дальнего Востока, включая и некоторые острова Охотского моря. Промысел соболя — это традиционное занятие населения таежных поселков. Значительное снижение численности вида в первой половине XX в. привело к запрету на промысел соболя, однако после восстановления численности популяций запрет был снят. Сравнительный анализ промысловой пушнины из различных регионов показал, что соболь относится к видам с хорошо выраженной географической изменчивостью. Данные накапливались с XVI в. и касаются признаков, имеющих товарную ценность, таких как окрас и качество меха, размер животных. Наиболее ценные по окрасу меха географические типы соболя обитают в районе оз. Байкал. Наибо-

лее крупный по размеру соболь обитает в равнинной тайге Западной Сибири и на Дальнем Востоке (п-ов Камчатка). Соболь южной части ареала (Алтае-Саянский регион) значительно отличается от соболей соседних географических регионов Среднесибирского плоскогорья и Западной Сибири [1, 2]. Таксономический статус соболя этого региона был определен по ряду количественных признаков как подвид *Martes zibellina averini* Vazhanov, 1943 [3–5].

Выявленные различия на основе морфологических признаков между популяциями географических регионов сохраняются до настоящего времени [2, 6]. Популяции сохранили присущие им признаки, несмотря на фрагментацию ареала в период снижения численности и последующие



**Рис. 1.** Карта выборок географических популяций соболя, охваченных промыслом, и заповедных территорий южной части ареала вида. Синий маркер: 1 – Саяны, 2 – Алтай-3, 3 – Алтай-4, 4 – Алтай-1, 5 – Алтай-2, 6 – Алтай-5, 7 – Томск-2, 8 – Томск-1, 9 – Баргузин; красный маркер: 1 – гос. природный заповедник “Юганский”, 2 – Катунский биосферный резерват, 3 – Сайлогемский национальный парк, 4 – гос. природный заповедник “Кузнецкий Алатау”, 5 – Шорский национальный парк, 6 – Алтайский гос. природный биосферный заповедник, 7 – гос. природный заповедник “Убсунурская котловина”, 8 – Саяно-Шушенский гос. природный заповедник, 9 – гос. природный заповедник Азас, 10 – Тунгусский гос. природный заповедник, 11 – Тункинский национальный парк, 12 – Витимский гос. природный заповедник, 13 – биосферный резерват “Катон-Карагай”, 14 – национальный парк “Алтай-Таван-Богд”, 15 – национальный парк оз. Хувсгул, 16 – Джергинский гос. природный заповедник, 17 – Баргузинский гос. природный биосферный заповедник им. К.А. Забелина.

мероприятия по интродукции соболя [7]. Дифференциация популяций географических регионов подтверждается исследованиями с применением генетических маркеров. Исследования генетической структуры выборок из смежных регионов позволили выявить особенности популяций, миграционные потоки и уровень изоляции отдельных популяций [8, 9].

Южную часть ареала вида населяют несколько географических популяций: соболь северо-запада Алтая, баргузинский соболь и другие [10]. Особенностью региона является наличие как разрешенных к промыслу территорий, так и значительного количества заповедных, всякая деятельность в пределах которых запрещена (рис. 1). Создание надежной референтной базы генетических данных южной части ареала вида позволит определить специфический набор маркеров для выявления популяционной принадлежности пушнины, добываемой промыслом. В дальнейшем это позволит обос-

новать размер промысла и поддерживать баланс численности уникальных географических популяций.

Исследования генетической идентификации природных популяций выполнены для многих видов хищных млекопитающих, например для русского волка [11], норки [12], хорька [13], а также европейской и американской куниц [14, 15]. Анализ популяционной принадлежности и генетическая индивидуализация особей соболя проведены впервые, но исследования генетической структуры популяций даны ранее [8, 9]. Эти результаты легли в основу создания референтной базы данных вида.

## МАТЕРИАЛЫ И МЕТОДЫ

### *Сбор биологического материала*

Сбор биологического материала южной части ареала вида проведен в период 2014–2018 гг. (табл. 1,

**Таблица 1.** Районы сбора биологического материала из популяций соболя южной части ареала

Часть ареала	Выборка	Географическое положение	Административный округ	Координаты (ш., д.)
Южная	Саяны	Западные Саяны, Саянская ландшафтная область	Курагинский район, Красноярский край	53.89°, 92.67°
	Алтай-3	Алтайская ландшафтная область	Улаганский район, Республика Алтай	50.63°, 87.96°
	Алтай-4	»	Турочакский район, Республика Алтай	51.79°, 87.26°
	Алтай-1	»	Усть-Коксинский район, Республика Алтай	50.27°, 85.62°
	Алтай-2	»	Поселок Усть-Кан, р. Чарыш в Республике Алтай	50.93°, 84.76°
	Алтай-5	»	Чарышский район, Алтайский край	51.39°, 83.56°
Юго-западная	Томск-2	Чулымская равнина	Первомайский район, Томская область	57.71°, 86.30°
	Томск-1	Васюганская равнина	Каргасокский район, Томская область	58.88°, 80.67°
Юго-восточная	Баргузин	Забайкалье, Становое нагорье	Северо-Байкальский район, Республика Бурятия	56.07°, 111.45°

рис. 1). Выборки Алтай-1, Алтай-2 и Алтай-5 собраны на территориях северо-западной провинции Алтайской горной области. Выборка Алтай-4 представляет популяции северо-восточной части Горного Алтая, выборка Алтай-3 – восточной части Горного Алтая. Саянская ландшафтная область представлена выборкой из популяции центральной части этого региона Саяны. Западная Сибирь представлена двумя выборками: одна выборка Томск-2 – из района среднего течения р. Чичка-Юл (правый приток Чулыма). В период депрессии численности здесь сохранялись жизнеспособные очаги местного “чулымского” соболя [1], в дальнейшем именно они стали основой современной нативной популяции соболя. Другая выборка Томск-1 – из района р. Васюган (левый приток р. Обь), в период восстановления численности вида сюда было завезено значительное количество восточносибирского соболя [7]. Выборка Баргузин формировалась в районе Станового нагорья северо-восточнее оз. Байкал.

Образцы депонированы в коллекции Института общей генетики им. Н.И. Вавилова РАН.

#### *Выделение ДНК и фрагментарный анализ*

В настоящей работе проанализированы девять выборок соболя, генотипирование проведено по

15 микросателлитным локусам (*Ma1, Ma3, Ma8, Ma15, Ma18, Ma19, Mer041, Mvi072, Mar08, Mar21, Mar36, Mar43, Mar53, Mar58, Mar64*). Описание всех используемых маркеров было дано в [16–18]. Методы выделения ДНК, проведения ПЦР, процедуры фрагментного анализа и идентификации аллелей микросателлитных локусов описаны ранее [8, 19].

#### *Статистический анализ*

Обработка данных фрагментного анализа микросателлитных локусов была выполнена с помощью программных пакетов, разработанных в программной среде R [20]. Основные популяционно-генетические показатели (число аллелей на локус, число аллелей на выборку, расчет попарных значений индекса  $F_{st}$  для оценки генетической дифференциации между популяциями) были получены с использованием пакета Adegnet [21]. В пакете PopGenReport были рассчитаны показатели аллельного богатства (allelic richness), а также частоты встречаемости аллелей [22]. Для оценки распределения частот аллелей использовали пакет DiveRsity [23]. Также с его помощью рассчитывали показатели ожидаемой и наблюдаемой гетерозиготности для каждой выборки, коэффициент инбридинга и  $p$ -значение для равновесного теста

**Таблица 2.** Уровень изменчивости используемых локусов южной части ареала вида

Локус	<i>A</i>	$H_o$	$H_e$	<i>PIC</i>	<i>HWE</i>
<i>Ma1</i>	12	0.849	0.874	0.859	NS
<i>Ma3</i>	3	0.431	0.548	0.460	**
<i>Ma8</i>	14	0.784	0.818	0.795	NS
<i>Ma15</i>	8	0.638	0.670	0.629	NS
<i>Ma18</i>	22	0.826	0.902	0.893	NS
<i>Ma19</i>	12	0.697	0.785	0.758	NS
<i>Mer041</i>	9	0.771	0.817	0.791	NS
<i>Mvi072</i>	11	0.702	0.816	0.792	**
<i>Mar08</i>	10	0.679	0.765	0.729	NS
<i>Mar21</i>	5	0.592	0.657	0.605	NS
<i>Mar36</i>	13	0.716	0.827	0.802	NS
<i>Mar43</i>	11	0.748	0.846	0.826	NS
<i>Mar53</i>	12	0.729	0.867	0.851	***
<i>Mar58</i>	14	0.706	0.766	0.747	NS
<i>Mar64</i>	10	0.665	0.785	0.751	***

Примечание. *A* – число выявленных аллелей,  $H_o$  – наблюдаемая гетерозиготность,  $H_e$  – ожидаемая гетерозиготность, *PIC* – показатель полиморфного информационного содержания, *HWE* – отклонение от равновесия Харди–Вайнберга (NS – незначимо, \*\* – значимо при  $p < 0.001$ , \*\*\* – значимо при  $p < 0.0001$ ).

Харди–Вайнберга. Для количественной оценки микросателлитных локусов использовали показатель уровня полиморфизма (polymorphism information content, *PIC*), рассчитанного в программном приложении Cervus 3.0.7 [24], значения которого основаны не только на числе аллелей в локусе, но и на распределении аллелей в популяциях [25].

Кластеризацию популяций строили методом UPGMA с оценкой надежности узлов путем бутстреп-анализа (100 000 итераций) на базе генетических расстояний Нея [26] в пакете PopprR [27]. На основе байесовского подхода была оценена вероятность принадлежности каждой из проанализированных особей к той или иной географической выборке. Расчет выполняли по алгоритму “Rannala & Mountain” в программе GeneClass2 [28, 29]. На основе коэффициента Мантеля в программе Arlequin 3.1 [30] оценивали уровень корреляции межпопуляционных генетических дистанций и географических расстояний (тест изоляции расстоянием, Isolation by distance).

Значимость коэффициента линейной регрессии ( $R^2$ ) определяли с помощью уровня значимости ( $\alpha$ ) по распределению Фишера.

SCOR-коэффициент, выступающий в качестве показателя надежности прогноза популяционной принадлежности особи, рассчитывали по формуле:

$$SCOR_{i,l} = \frac{L_{i,l}}{\sum_{j=1}^k L_{i,j}},$$

где  $k$  – количество выборок,  $L_{i,l}$  – вероятность того, что  $i$ -тая особь принадлежит выборке  $l$ .

## РЕЗУЛЬТАТЫ

### Генетическая изменчивость

Всего было генотипировано 218 образцов из девяти выборок южной части ареала соболя (рис. 1, табл. 1). Число аллелей в 15 локусах варьировало от 3 до 22 (среднее  $A = 11.067$ ), наблюдаемая гетерозиготность – от 0.431 до 0.893 (среднее  $H_o = 0.703$ , среднее  $H_e = 0.783$ ) (табл. 2). Исходя из значений *PIC*, локус *Ma18* был наиболее информативным (0.893), а *Ma3* – наименее информативным (0.460). В среднем по всем локусам показатель полиморфного информационного содержания составил 0.753, что свидетельствует о высокой информативности используемой микросателлитной панели (табл. 2). Отклонение от равновесия Харди–Вайнберга для всей совокупности исследуемых особей зафиксировано в четырех локусах *Mar53*, *Mar64* ( $p < 0.0001$ ) и *Ma3*, *Mvi072* ( $p < 0.001$ ). Отклонение от равновесия Харди–Вайнберга может проявляться по разным причинам, но практически по каждому локусу выявлен дефицит гетерозигот (табл. 2), что предполагает наличие субпопуляционной структуры и инбридинга, проявление эффекта Валунда [31].

Анализ изменчивости выборок (табл. 3) показал, что аллельное богатство (*Ar*), рассчитанное для 15 локусов, изменялось от 5.19 (Алтай-2) до 5.77 (Баргузин). В популяциях Алтай-1 и Саяны выявлено максимальное количество общих аллелей (74.95 и 74.65%). В четырех популяциях (Саяны, Алтай-1, Алтай-5, Томск-1) выявлено отклонение от равновесия Харди–Вайнберга.

В выборке из Забайкалья Баргузин обнаружено четыре собственных аллеля (private alleles): *Ma8* – 114, 130, *Mar36* – 206, *Mar64* – 193, что может быть связано с возможным влиянием миграций из рефугиума юго-восточной части ареала [8].

Наличие трех приват-аллелей в выборке Томск-2 (*Ma8* – 106, *Mar53* – 235, *Mar64* – 179) вероятно является результатом частичной изоляции популяции соболя этого региона. Три собственных аллеля выявлены в северо-западной части ареала соболя Алтая (Алтай-1, *Ma1* – 228, *Ma18* – 164, *Ma18* – 174), что связано с высокой степенью изоляции популяций соболя этой части ареала. В выборках Саяны и Алтай-5 также обнаружены три

**Таблица 3.** Популяционно-генетическая характеристика выборок соболя южной части ареала вида

Популяция	<i>N</i>	<i>A</i>	%	<i>Ar</i>	<i>H<sub>o</sub></i>	<i>H<sub>e</sub></i>	<i>HWE</i>
Саяны	23	117	74.65	5.67	0.73	0.76	***
Алтай-3	14	104	64.32	5.63	0.71	0.74	NS
Алтай-4	24	113	70.42	5.68	0.74	0.75	NS
Алтай-1	33	116	74.95	5.58	0.67	0.72	**
Алтай-2	10	92	59.79	5.19	0.7	0.72	NS
Алтай-5	28	119	74.48	5.75	0.7	0.75	***
Томск-2	30	111	71.32	5.32	0.7	0.72	NS
Томск-1	24	118	72.32	5.67	0.69	0.71	**
Баргузин	32	118	72.96	5.77	0.7	0.74	NS

Примечание. *N* – число образцов в выборке, *A* – число аллелей в выборке, % – процент общих аллелей, *Ar* – аллельное богатство, *H<sub>o</sub>* – наблюдаемая гетерозиготность, *H<sub>e</sub>* – ожидаемая гетерозиготность, *HWE* – отклонение от равновесия Харди–Вайнберга (NS – незначимо, \*\* – значимо при  $p < 0.01$ , \*\*\* – значимо при  $p < 0.0001$ ).

собственных аллеля (Саяны: *Ma15* – 212, *Mar43* – 151; Алтай-5: *Ma18* – 172), что в данном случае может быть результатом влияния миграционных потоков северных популяций Среднесибирского плоскогорья, не задействованных в настоящем исследовании.

#### Генетическая дифференциация

Анализ парной дифференциации выборок (табл. 4) показал низкую дифференциацию между выборками внутри географического региона Северо-Западного Алтая: Алтай-2 и Алтай-5 ( $F_{st} = 0.015$ ,  $p = 0.172$ ). Между выборками восточной части региона Алтай-3 и Алтай-4 также значения коэффициента низкие ( $F_{st} = 0.016$ ,  $p = 0.047$ ), что может свидетельствовать о существовании в этом регионе только двух популяционных группировок соболя. Низкий коэффициент дифференциации

выявлен между группами Саяны и Баргузин ( $F_{st} = 0.022$ ,  $p = 0.040$ ), хотя географическое расстояние между ними составляет около 1240 км. Это обусловлено тем, что популяция Саяны находится под влиянием миграционных потоков, в том числе и из восточной части ареала [8]. Сравнительно высоко дифференцированы популяции западной части ареала соболя Томск-1 и Томск-2: коэффициент дифференциации  $F_{st} = 0.033$  ( $p = 0.003$ ). Объяснение значительной дифференциации популяций соболя одного географического региона можно найти в истории восстановления популяций Западной Сибири. В данном случае популяция Томск-1 создавалась при значительных интродукциях соболя из восточной части ареала [7], в то время как популяция Томск-2 восстановлена на основе местных популяций.

Несмотря на относительную географическую близость выборок Алтай-1 и Алтай-3 (около 160 км), коэффициент  $F_{st} = 0.038$  ( $p = 0.001$ ), что связано с изолированностью популяций Северо-Западного Алтая.

Глобальная оценка дифференциации между исследуемыми выборками составила  $F_{st} = 0.0463$ . Наибольший вклад в подразделенность вносят локусы *Ma3* ( $F_{st} = 0.111$ ) и *Mar64* ( $F_{st} = 0.102$ ), наименьший – *Ma19* ( $F_{st} = 0.021$ ), что соответствует различию в 5 раз (табл. 5).

Помимо коэффициента фиксации Райта ( $F_{st}$ ) [32, 33], нами также был произведен расчет глобальных значений подразделенности и с помощью других тестов:  $G_{st} = 0.178$ ,  $D = 0.140$ . В данном случае  $G_{st}$  – адаптированный Неем [26] индекс фиксации ( $F_{st}$ ) для мультиаллельных маркеров, рассчитанный в качестве подтверждающей оценки и показавший сопоставимые с ним результаты. Индекс  $D$ , рассчитанный по Йосту [34], был применен в качестве уточняющей меры, поскольку по мнению ряда авторов индексы  $F_{st}$  и  $G_{st}$  могут вводить в заблуждение при расчете для мультиал-

**Таблица 4.** Парные значения  $F_{st}$  и соответствующие  $p$ -значения между выборками соболя южной части ареала

Выборки	Саяны	Алтай-3	Алтай-4	Алтай-1	Алтай-2	Алтай-5	Томск-2	Томск-1	Баргузин
Саяны		0.001	0.010	0.001	0.001	0.001	0.002	0.002	0.040
Алтай-3	0.030		0.047	0.001	0.001	0.002	0.001	0.005	0.006
Алтай-4	0.022	0.016		0.002	0.002	0.004	0.001	0.001	0.035
Алтай-1	0.036	0.038	0.030		0.008	0.011	0.001	0.001	0.001
Алтай-2	0.031	0.037	0.026	0.025		0.172	0.001	0.001	0.009
Алтай-5	0.034	0.028	0.027	0.023	0.015		0.001	0.001	0.001
Томск-2	0.027	0.034	0.040	0.044	0.036	0.046		0.003	0.002
Томск-1	0.031	0.027	0.032	0.046	0.043	0.038	0.033		0.005
Баргузин	0.022	0.028	0.022	0.049	0.030	0.040	0.036	0.035	

Примечание. Под диагональю – значения  $F_{st}$ , над диагональю – соответствующие им  $p$ -value (различия значимы при  $p < 0.05$ ).

**Таблица 5.** Значения  $F$ -статистик Райта используемых 15 микросателлитных локусов

Локус	$F_{is}$	$F_{st}$	$F_{it}$
<i>Ma1</i>	0.0058	0.0270	0.0326
<i>Ma3</i>	0.1270	0.1106	0.2235
<i>Ma8</i>	0.0126	0.0321	0.0443
<i>Ma15</i>	0.0040	0.0584	0.0547
<i>Ma18</i>	0.0668	0.0219	0.0872
<i>Ma19</i>	0.0945	0.0213	0.1138
<i>Mer041</i>	0.0302	0.0305	0.0597
<i>Mvi072</i>	0.1025	0.0470	0.1447
<i>Mar08</i>	0.0694	0.0519	0.1177
<i>Mar21</i>	0.0676	0.0392	0.1042
<i>Mar36</i>	0.0861	0.0607	0.1416
<i>Mar43</i>	0.0892	0.0339	0.1201
<i>Mar53</i>	0.1384	0.0273	0.1619
<i>Mar58</i>	0.0256	0.0600	0.0840
<i>Mar64</i>	0.0687	0.1015	0.1632
Среднее	0.0648	0.0463	0.1081

лельных случаев [35–37], в частности при наличии более чем двух аллелей коэффициенты Райта ( $F_{st}$ ) и Нея ( $G_{st}$ ) не будут достигать единицы даже при отсутствии между популяциями общих аллелей, поскольку в популяциях всегда будет присутствовать некоторая гетерозиготность. В отличие от этого показатель дифференциации  $D$ , который измеряет долю аллельных вариаций среди популяций, будет равен 1 при полной дифференциации и нулю при ее отсутствии даже при высокой вариации внутри популяций.

Однако в то время как  $G_{st}$  ограничивается низкими значениями при высокой гетерозиготности,  $D$  чувствителен к высокополиморфным маркерам (в частности к таким, скорость мутаций которых выше  $10^{-8}$ ) и в случае микросателлитов может быть смещен вверх при наличии большого количества популяций. Одним из дополнительных решений, часто рекомендуемых для оценки структуры популяции с использованием маркеров высокой мутации, является расчет как  $G_{st}$ , так и  $D$ . Маркеры, где  $G_{st}$  недооценивает дифференциацию, должны иметь значительно повышенные значения  $D$ .

В нашем случае противоречия между значениями  $F_{st}$ ,  $G_{st}$  и  $D$  не наблюдается: попарные значения  $G_{st}$  и  $D$  для всех исследуемых выборок пропорциональны соответствующим значениям  $F_{st}$  и  $p$ -value, приведенным в табл. 4.

### Идентификация

Наличие базы данных по частотам микросателлитных локусов для ряда популяций вида дает возможность идентифицировать отдельных особей с неизвестным происхождением. Для этого необходимо знать разрешающую способность имеющейся базы данных, т.е. долю правильно идентифицированных особей. Для решения данной задачи требуется протестировать имеющуюся базу данных: провести симуляционный анализ и определить долю правильно и ошибочно идентифицированных особей среди особей с известным происхождением.

При этом необходимо иметь в виду существование двух типов ошибок идентификации. Первый тип ошибки — это доля особей, ошибочно отнесенных к несоответствующим им популяциям, т.е. доля неправильно идентифицированных особей среди всех особей, отнесенных программой к какой-нибудь популяции. Второй тип ошибки идентификации — это доля особей, ошибочно не включенных в соответствующие им популяции, т.е. неправильно идентифицированных особей среди животных с известным происхождением.

На практике при идентификации особей с неизвестным происхождением обе эти ошибки имеют значение: первый тип ошибки дает представление о количестве “как минимум” особей, а второй — о количестве “как максимум” особей конкретной группы среди всех животных с неизвестным происхождением. Так, при идентификации 70 особей с неизвестным происхождением с использованием протестированной базы данных было выявлено 40 особей, отнесенных программой к западно-алтайской группировке. Разрешающая способность базы данных при использовании первого типа ошибки для западно-алтайской группировки — 87.9%. То есть среди 40 особей как минимум 87.9% (35.2 особи) действительно относятся к западно-алтайской группировке. Разрешающая способность базы данных при использовании второго типа ошибки для западно-алтайской группировки — 81.7%, т.е. 35.2 особи составляют 81.7% от максимального количества особей данной популяции. Отсюда максимально возможное количество (100%) особей западно-алтайской группировки среди 70 анализируемых животных будет равно 43.1 особей ( $43.1 = (35.2 \times 100\%) / 81.7\%$ ). Таким образом, среди 70 животных с неизвестным происхождением к западно-алтайской группировке может относиться от 35.2 до 43.1 особей.

Девять выборок соболя были выделены в шесть популяционных групп: Северо-Западный Алтай (Алтай-1, Алтай-2, Алтай-5), Восточный Алтай (Алтай-3, Алтай-4), Саяны, Томск-1, Томск-2, Баргузин. При формировании популяционных групп учитывались как оценки генетической дифференциации выборок, так и результаты кластеризации по расстояниям Нея (рис. 2).



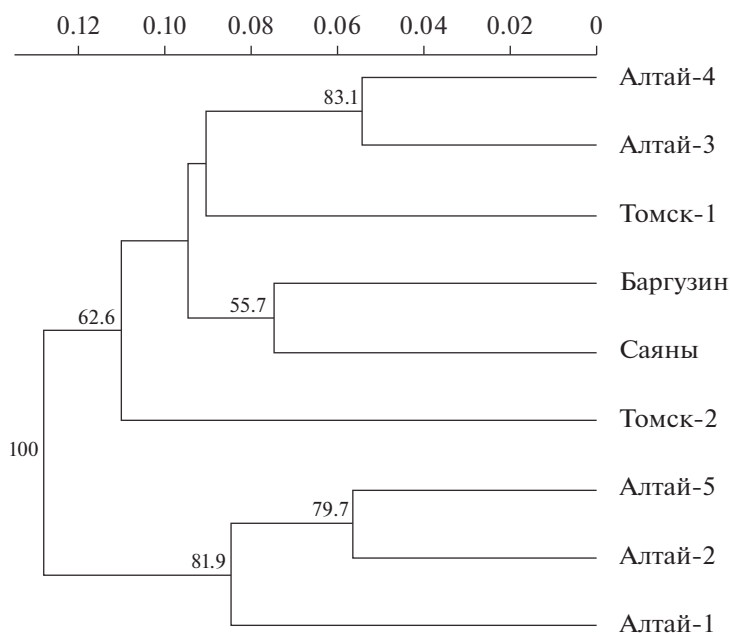


Рис. 2. UPGMA-кластеризация выборок соболя южной части ареала вида. Значения бутстреп-поддержки приведены в узлах ветвления.

Несмотря на то что выборка Алтай-1 значительно отличается от всех других выборок, ее генетический профиль близок к выборкам Алтай-2 и Алтай-5. В связи с этим мы обозначили данную группу выборок как Северо-Западный Алтай. Перечисленные шесть групп были взяты за основу для индивидуальной идентификации особей.

В процессе симуляционного анализа особь, которая в данный момент идентифицировалась, исключалась из базы. Анализ проводился для базы данных по 8, 10, 12 и 15 локусам.

Анализ проведен для выяснения оптимального числа микросателлитных локусов при определении популяционной принадлежности особи. Для 10 и 12 локусов использовались все возможные комбинации (т.е. стандартная панель из восьми локусов (*Ma1*, *Ma3*, *Ma8*, *Ma15*, *Ma18*, *Ma19*, *Mer041*, *Mvi072*) [8, 19] плюс возможные комбинации из оставшихся семи локусов (*Mar08*, *Mar21*, *Mar36*, *Mar43*, *Mar53*, *Mar58*, *Mar64*). Таким образом, для 10 локусов использовались средняя оценка точности и ошибки идентификации из 21 варианта возможных панелей микросателлитных локусов, а для 12 локусов использовались средняя оценка точности и ошибки идентификации по 35 возможным вариантам панели микросателлитных локусов.

Была проведена оценка точности идентификации особей отдельно для каждой из шести выделенных групп соболя: Северо-Западный Алтай (Алтай-1, Алтай-2, Алтай-5), Восточный Алтай (Алтай-3, Алтай-4), Западные Саяны (Саяны),

две региональные группировки Западной Сибири (Томск-2 нативный, Томск-1 интродуцированный) и баргузинский соболь (Баргузин). Анализ показал, что для некоторых группировок точность идентификации незначительно увеличивается при увеличении числа локусов: например, Северо-Западный Алтай и Томск-2. Для Томска-2 (нативный) точность идентификации как по первому (I) типу ошибки, так и по второму (II) типу не изменилась при увеличении числа локусов (рис. 3, 4; табл. 6).

Для других группировок точность идентификации значительно возросла: так, по I типу ошибки для популяции Саяны, баргузинского соболя (Баргузин) и Томск-2 (интродуцированный) на 21, 18.7 и 16.4% соответственно (рис. 3), по II типу ошибки для Томск-2 и Саяны на 29.2 и 17.4% соответственно (рис. 4).

Помимо точности идентификации (количество правильно идентифицированных особей) для характеристики базы данных также важен такой показатель, как надежность идентификации особи. В качестве такого показателя может выступать SCOR-коэффициент. Он отражает надежность прогноза популяционной принадлежности особи и оценивает долю вероятности принадлежности особи к данной популяции, в общей сумме всех возможных вероятностей.

В работе было рассчитано среднее значение индивидуальных SCOR-коэффициентов по всем выборкам для панели микросателлитных локусов из 8, 10, 12 и 15 локусов (средний SCOR –  $77.44 \pm 1.34$ ,

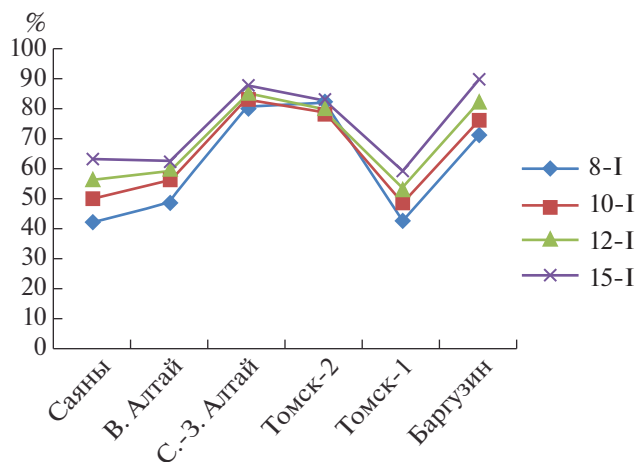


Рис. 3. Оценки точности идентификации особей по базе данных частот 8, 10, 12 и 15 микросателлитных локусов соболя для каждой группировки с учетом I типа ошибки (%).

81.42 ± 0.96, 84.86 ± 0.90, 88.74 ± 1.13 соответственно). Была проведена оценка парной линейной зависимости средней величины SCOR и числа локусов. Коэффициент детерминации  $R^2 = 0.9993$  (оценка значима,  $\alpha = 0.0004$ ), т.е. можно утверждать, что при увеличении числа локусов до 15 надежность прогноза значимо возрастает в среднем для каждой особи (средний SCOR изменяется с  $77.44 \pm 1.34$  (для восьми локусов) до  $88.74 \pm 1.13$  (для 15 локусов)).

Далее мы провели анализ среднего SCOR-коэффициента отдельно для особей, которые были “верно идентифицированы” (т.е. им присвоена та же выборка, из которой особь была взята), и для “неверно идентифицированных” особей (т.е. им присвоена не та выборка, из которой особь была взята) в целом по всем группам. Такой анализ был проведен для восьми и для 15 локусов (табл. 7). Было отмечено, что средний SCOR-коэффициент для “верно идентифицированных” особей досто-

верно выше, чем для “неверно идентифицированных” особей как для восьми ( $\alpha = 5.93 \times 10^{-9}$ ), так и для 15 локусов ( $\alpha = 2.64 \times 10^{-6}$ ).

В тоже время для 15 локусов средний SCOR-коэффициент достоверно выше по обоим категориям особей, чем для панели из восьми локусов (для “неверно идентифицированных” особей  $\alpha = 0.0004$ ; для “верно идентифицированных” особей  $\alpha = 4.82 \times 10^{-6}$ ).

## ОБСУЖДЕНИЕ

К середине XX в. численность соболя в России была восстановлена, на всей территории обитания вида был организован лицензионный промысел. Анализ результатов многолетнего промысла показал, что численность вида в целом стабильна. Международный союз охраны природы (МСОП) присвоил соболю статус “Не вызывающий беспокойства” (Least concern). Но вид на территории страны представлен рядом географических популяций и подвидов, ценность пушнины которых различается между собой в несколько раз. Это приводит к разной интенсивности промысла в регионах. Кроме этого, на исследуемой нами части ареала разрешенные к промыслу территории чередуются с заповедными, всякая деятельность в которых запрещена. Так, промысловый район Северо-Западного Алтая граничит с заповедной территорией Катунского биосферного резервата (Россия) и Биосферного резервата “Катон-Карагай” (Казахстан). Выборки из Восточного Алтая граничат с Алтайским государственным природным биосферным заповедником. Рядом с забайкальским промысловым районом (восточнее оз. Байкал) располагаются три заповедные территории (Джержинский государственный природный заповедник; Витимский государственный природный заповедник; Баргузинский государственный природный биосферный заповедник им. К.А. Забелина) (рис. 1).

Несмотря на миграционную активность соболя, между популяциями существует генетическая

Таблица 6. Коэффициенты детерминации ( $R^2$ ) и уровни значимости ( $\alpha$ ) парной линейной регрессии (зависимость точности идентификации от числа локусов) для каждой выборки для I и II типов ошибок

Популяции	I тип ошибки		II тип ошибки	
	$R^2$	$\alpha$	$R^2$	$\alpha$
Саяны	0.9959	0.0021	0.9607	0.0198
Восточный Алтай (Алтай-3, -4)	0.9979	0.0011	0.8383*	0.0844
Северо-Западный Алтай (Алтай-1, -2, -5)	0.9957	0.0022	0.9480	0.0263
Томск-2	0.0385*	0.8038	0.8043*	0.1032
Томск-1	0.9877	0.0062	0.9677	0.0163
Баргузинский соболь (Баргузин)	0.9531	0.0237	0.9946	0.0027

\* – незначимые коэффициенты детерминации ( $\alpha > 0.05$ ), т.е. статистически подтвержденной зависимости нет.



**Таблица 7.** Средний SCOR-коэффициент для объединенной популяции южной части ареала соболя

Число локусов	“Неверно идентифицированные” особи	“Верно идентифицированные” особи
8	67.18 ± 2.16	82.94 ± 1.51
15	79.46 ± 2.63	91.65 ± 1.15

дифференциация, обусловленная в том числе изоляцией расстоянием – основным фактором формирования межпопуляционных различий на большей части ареала. Тест Мантеля показывает значимую корреляцию между межпопуляционными генетическими дистанциями и географическим расстоянием, разделяющим места обитания соболя ( $cor = 0.528$ ,  $p = 4.18 \times 10^{-7}$ ). Анализ генетической структуры популяций южной части ареала показал значительный уровень дифференциации между популяциями (табл. 4). Максимальные значения  $F_{st}$  (0.044–0.049) выявлены между изолированной популяцией Северо-Западного Алтая (выборки Алтай-1, Алтай-2, Алтай-5) как с популяциями Западной Сибири (Томск-1, Томск-2), так и восточной части исследуемого ареала (популяцией Забайкалья Баргузин). Результаты дифференциации подтверждаются кластерным анализом по методу невзвешенных парногрупповых средних (UPGMA), показано разделение выборок на несколько кластеров (рис. 2). Основываясь на этих данных и предыдущих работах [9], были выделены шесть групп соболя, оптимально отражающих популяционно-генетическую структуру вида на изученной части ареала (табл. 1).

Оценка идентификационной способности базы данных с разным числом локусов показала, что точность идентификации для одних групп может не измениться, для других сильно возрасти. Так, для Томска-2 точность идентификации не изменилась при увеличении числа локусов. Аналогичные данные получены при идентификации изолированной группировки Северо-Западного Алтая (объединенные выборки Алтай-1, Алтай-2, Алтай-5). Это может свидетельствовать о том, что полученные оценки точности являются максимально возможными и дальнейшее увеличение числа локусов не даст ожидаемых результатов. Эти две группы находятся в значительной степени изоляции и миграционные потоки здесь ограничены.

Для других группировок точность идентификации значительно возросла при увеличении числа локусов до 15. Например, по I типу ошибки для популяций Саяны, баргузинского соболя и Томск-1 на 21, 18.7 и 16.4% соответственно (рис. 3), по II типу ошибки для Томск-1 и Саяны на 29.2 и 17.4% соответственно (рис. 4).

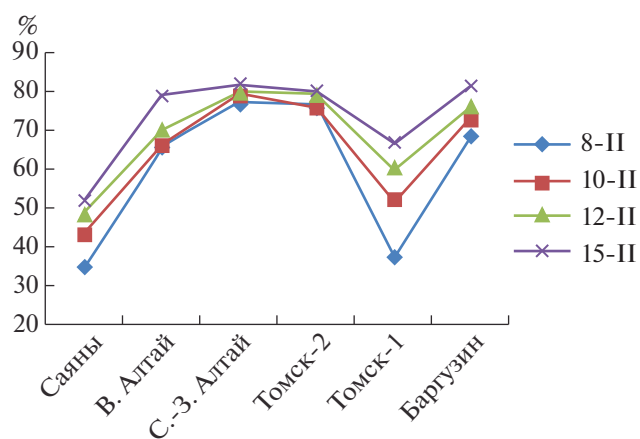
Таким образом, при увеличении числа локусов с восьми до 15 значительно возрастает точность инди-

видуальной идентификации соболя (количество правильно идентифицированных особей).

Анализ среднего значения SCOR-коэффициентов показал, что при увеличении числа локусов с восьми до 15 значительно возрастает надежность индивидуальной идентификации в среднем для каждой отдельной особи (средний SCOR изменяется с  $77.44 \pm 1.34$  – для восьми локусов до  $88.74 \pm 1.13$  – для 15 локусов) (табл. 7).

Средний SCOR-коэффициент “верно идентифицированных” особей по 15 локусам составил 91.65%, что является достаточно хорошей характеристикой базы данных. Более низкий средний SCOR-коэффициент у “неверно идентифицированных” особей может свидетельствовать о том, что данные особи сложно идентифицируются и их генетический профиль трудно однозначно отнести к какой-либо популяции в силу того, что (1) их популяция отсутствует в базе данных, (2) степень дифференциации разных популяций невелика, (3) такие особи могут быть гибридами. В настоящий момент все эти три причины могут иметь место в базе данных соболя. Следует также отметить, что среди “неверно идентифицированных” особей присутствовали особи с высоким значением SCOR-коэффициента, такие особи могут являться реальными мигрантами.

В работе было показано, что при использовании базы данных частот аллелей микросателлитных ло-



**Рис. 4.** Оценки точности идентификации особей по базе данных частот 8, 10, 12 и 15 микросателлитных локусов соболя для каждой группировки с учетом II типа ошибки (%).

кусов для индивидуальной идентификации особей соболя наиболее высокими показателями идентификации обладает база данных из 15 локусов. По результатам анализа был разработан алгоритм идентификации группы особей неизвестного происхождения. Была оценена идентификационная (или разрешающая) способность базы данных для каждой из шести группировок. При использовании панели из 15 микросателлитных локусов достоверно снижается количество неправильно идентифицированных особей и достоверно возрастает надежность идентификации в среднем для каждой отдельной особи по сравнению с панелью из восьми локусов. Точность идентификации изменяется от 65 до 76% в целом для всей базы, а ошибка идентификации уменьшается в 1.5 раза. Также при использовании базы данных из 15 локусов возрастает надежность идентификации в среднем для каждой отдельной особи (SCOR = 88.74).

Одной из задач настоящего исследования являлось определение генетического профиля популяций соболя, обитающих на юге Сибири, что позволило создать референтную базу данных для индивидуальной идентификации особей, обитающих на промысловых и заповедных территориях. В перспективе это позволит сохранить биологическое разнообразие вида и составляющих его географических популяций.

#### СПИСОК ЛИТЕРАТУРЫ

1. Тимофеев В.В., Надеев В.Н. Соболя. М.: Заготиздат, 1955. 403 с.
2. Монахов Г.И., Бакеев Н.Н. Соболя. 2-е изд. М.: Лесная пром-сть, 1981. 240 с.
3. Бажанов В.С. Новый подвид соболя с Алтая // Труды Казахского фил. АН СССР за 1942 год. Алма-Ата: Изд-во АН КазССР, 1943. С. 53–54.
4. Кузнецов Б.А. *Martes zibellina* L., 1758 – Соболя // Млекопитающие Казахстана. М.: МОИП, 1948. С. 77.
5. Лобачев Ю.С., Афанасьев Ю.Г. Соболя *Martes zibellina* Linnaeus, 1758 // Млекопитающие Казахстана. Алма-Ата: Изд-во АН КазССР, 1982. Т. 3. Ч. 2. С. 101–119.
6. Монахов В.Г. Размерная структура соболя в Прибайкалье: подекадный анализ за последние 60 лет // Изв. РАН. Серия биол. 2013. № 1. С. 40–47.
7. Павлов М.П., Корсакова И.Б., Тимофеев В.В. Акклиматизация охотничье-промысловых зверей и птиц в СССР. Киров: Волго-Вят. кн. изд-во, 1973. 536 с.
8. Каштанов С.Н., Свищёва Г.Р., Пищулина С.Л. и др. Географическая структура генофонда соболя (*Martes zibellina* L.): данные анализа микросателлитных локусов // Генетика. 2015. Т. 51. № 1. С. 78–88.
9. Каштанов С.Н., Столповский Ю.А., Мецкерский И.Г. и др. Таксономический статус и генетическая идентификация соболя Алтая (*Martes zibellina averini* Bazhanov, 1943) // Генетика. 2018. Т. 54. № 11. С. 1327–1337.
10. Чернукин Е.М. Экология соболя (*Martes zibellina* Linnaeus, 1758) в Баргузинском заповеднике. Улан-Удэ: Изд-во Бурятского гос. ун-та, 2006. 265 с.
11. Aspi J., Roininen E., Kiiskila J. et al. Genetic structure of the northwestern Russian wolf populations and gene flow between Russia and Finland // Conserv. Genet. 2009. V. 10. P. 815–826.
12. Michaux J.R., Hardy O.J., Justy F. et al. Conservation genetics and population history of the threatened European mink *Mustela lutreola*, with an emphasis on the west European population // Mol. Ecol. 2005. V. 14. P. 2373–2388.
13. Pertoldi C., Breyne P., Cabria M.T. et al. Genetic structure of the European polecat (*Mustela putorius*) and its implication for conservation strategies // J. Zool. 2006. V. 270. P. 102–105.
14. Colli L., Cannas R., Deiana A.M. et al. Microsatellite variability of sardinian pine martens, *martes martes* // Zoolog. Sci. 2011. V. 28. P. 580–586.
15. O'Shea K. Genetic study of recent samples of American marten (*Martes americana*) from Vermont // UVM Honors College Senior Theses. 2014. V. 6. 29 p.
16. Davis C.S., Strobeck C. Isolation, variability, and cross-species amplification of polymorphic microsatellite loci in the family mustelidae // Mol. Ecol. 1998. V. 7. Is. 12. P. 1776–1778.
17. Fleming M.A., Ostrander E.A., Cook J.A. Microsatellite markers for American mink (*Mustela vison*) and ermine (*Mustela erminea*) // Mol. Ecol. 1999. V. 8. № 8. P. 1352–1355.
18. Natali C., Banchi E., Ciofi C. et al. Characterization of 13 polymorphic microsatellite loci in the European pine marten *Martes martes* // Conserv. Genet. Resources. 2010. V. 2. P. 397–399.
19. Каштанов С.Н., Афанасьев К.И., Помапов С.Г., Лазебный О.Е. Микросателлитный анализ двух промысловых популяций соболя (*Martes zibellina* L., 1943) // Генетика. 2011. Т. 47. № 12. С. 1622–1628.
20. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2014.
21. Jombart T. Adegnet: An R package for the multivariate analysis of genetic markers // Bioinformatics. 2008. V. 24(11). P. 1403–1405.
22. Adamack A.T., Gruber B. PopGenReport: Simplifying basic population genetic analyses in R // Methods Ecol. Evol. 2014. V. 5. Is. 4. P. 384–387.
23. Keenan K., McGinnity Ph., Cross T.F. et al. DiveRcity: An R package for the estimation and exploration of population genetics parameters and their associated errors // Methods Ecol. Evol. 2013. V. 4. Is. 8. P. 782–788.
24. Botstein D., White R.L., Skolnick M.H., Davis R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms // Am. J. Hum. Genet. 1980. V. 32. P. 314–331.
25. Gruber B., Adamack A.T. Landgenreport: A new R function to simplify landscape genetic analysis using resistance surface layers // Mol. Ecol. Resour. 2015. V. 15(5). P. 1172–1178.

26. *Nei M.* Genetic distance between populations // *Am. Nat.* 1972. V. 106. P. 283–292.
27. *Kamvar Z.N., Tabima J.F., Grünwald N.J.* Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction // *Peer J.* 2014. V. 2. P. e281. <https://doi.org/10.7717/peerj.281>
28. *Rannala B., Mountain J.L.* Detecting immigration by using multilocus genotypes // *Proc. Natl Acad. Sci. USA.* 1997. V. 94(17). P. 9197–9201.
29. *Piry S., Alapetite A., Cornuet J.-M. et al.* GENECLASS2: A software for genetic assignment and first-generation migrant detection // *J. Hered.* 2004. V. 95. Is. 6. P. 536–539.
30. *Excoffier L., Laval G., Schneider S.* Arlequin (version 3.0): An integrated software package for population genetics data analysis // *Evol. Bioinform.* 2005. № 1. P. 47–50.
31. *Wahlund S.* Zusammensetzung von populationen und korrelationserscheinungen vom standpunkt der vererbungslehre aus betrachtet // *Hereditas.* 1928. V. 11. Is. 1. P. 65–106.
32. *Wright S.* The genetical structure of populations // *Ann. Eugen.* 1951. V. 15. P. 323–354.
33. *Wright S.* The interpretation of population structure by F-statistics with special regard to systems of mating // *Evolution.* 1965. V. 19. P. 395–420.
34. *Jost L.* GST and its relatives do not measure differentiation // *Mol. Ecol.* 2008. V. 17. P. 4015–4026.
35. *Hedrick P.W.* A standardized genetic differentiation measure // *Evolution.* 2005. V. 59. P. 1633–1638.
36. *Jost L.* Reply: D vs G'ST: response to Heller and Siegmund (2009) and Ryman and Leimar (2009) // *Mol. Ecol.* 2009. V. 18. P. 2088–2091.
37. *Kronholm I., Loudet O., de Meaux J.* Influence of mutation rate on estimators of genetic differentiation—lessons from *Arabidopsis thaliana* // *BMC Genetics.* 2010. V. 11. P. 33. <https://doi.org/10.1186/1471-2156-11-33>

## Identification of the Sable (*Martes zibellina* L.) Populations in the Southern Part of the Species' Range

S. N. Kashtanov<sup>a, \*</sup>, M. V. Shitova<sup>a</sup>, M. M. Somova<sup>d</sup>, P. A. Filimonov<sup>a</sup>, O. Yu. Tyutenkov<sup>e</sup>,  
A. A. Khodireva<sup>a</sup>, G. R. Svishcheva<sup>a, c</sup>, I. G. Meshchersky<sup>b</sup>, Yu. A. Stolpovsky<sup>a</sup>, and V. V. Rozhnov<sup>b</sup>

<sup>a</sup>Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 119991 Russia

<sup>b</sup>Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, Moscow, 119071 Russia

<sup>c</sup>Federal Research Center Institute of Cytology and Genetics, Siberian Branch  
of the Russian Academy of Sciences, Novosibirsk, 630090 Russia

<sup>d</sup>Ernst Federal Research Center for Animal Husbandry, Moscow oblast, Podolsk, Dubrovitsy, 142132 Russia

<sup>e</sup>National Research Tomsk State University, Tomsk, 634050 Russia

\*e-mail: snkashtanov@mail.ru

A panel consisting of 15 microsatellite loci was employed to study the modern structure of the sable gene pool. In the southern part of its range, the high level of differentiation among the species' populations was revealed. Several population clusters of the sable samples were found in a vast territory from Western Siberia to Transbaikalia. This part of the species range is distinguished by the sable's high abundance as well as by the fact that controlled hunting areas alternate with numerous reserves. Possibilities of identification of individuals by population and regional affiliations were determined. The study assessed the reliability of the prediction of population belonging (SCOR-coefficient) of individuals in the pooled sample, which showed that with an increase in the number of loci from 8 to 15, the accuracy of identification (the number of correctly identified individuals) and the reliability of identification on average for each individual increase significantly. The results allow one to expand the reference database of genetic data, contribute to the preservation of the intra-specific diversity of the sable and solve a number of problems associated with hunting in the territories bordering with nature conservation areas.

**Keywords:** geographic populations of sable, introduction, hunting, microsatellite analysis, phylogenetic analysis, genetic structure, DNA identification.