

УДК 575.113+582.475.4

## РАЗРАБОТКА МИКРОСАТЕЛЛИТНЫХ МАРКЕРОВ СОСНЫ ОБЫКНОВЕННОЙ С ДЛИННЫМИ МОТИВАМИ НА ОСНОВЕ ПОЛНОГЕНОМНОГО *de novo* СЕКВЕНИРОВАНИЯ

Н. В. Орешкова<sup>1, 2, 3</sup>, Е. И. Бондар<sup>1, 3</sup>, В. В. Шаров<sup>1, 3</sup>, К. В. Крутовский<sup>3, 4, 5, 6</sup>

<sup>1</sup> Федеральный исследовательский центр «Красноярский научный центр СО РАН»  
660036, Красноярск, Академгородок, 50

<sup>2</sup> Институт леса им. В. Н. Сукачева СО РАН – обособленное подразделение ФИЦ КНЦ СО РАН  
660036, Красноярск, Академгородок, 50/28

<sup>3</sup> Сибирский федеральный университет  
660041, Красноярск, пр. Свободный, 79

<sup>4</sup> Институт общей генетики им. Н. И. Вавилова РАН  
119991, Москва, ул. Губкина, 3

<sup>5</sup> Воронежский государственный лесотехнический университет им. Г. Ф. Морозова  
394087, Воронеж, ул. Тимирязева, 8

<sup>6</sup> Гёттингенский университет им. Георга-Августа  
Германия, 37077, Гёттинген

E-mail: oreshkova@ksc.krasn.ru, bone-post@yandex.ru, vsharov@sfu-kras.ru,  
konstantin.krutovsky@forst.uni-goettingen.de

Поступила в редакцию 03.12.2024 г.

Благодаря современным методам полногеномного секвенирования стала возможной разработка большого числа ДНК-маркеров, как селективно-нейтральных, так и функциональных. Наиболее информативными, воспроизводимыми, относительно недорогими и высокополиморфными являются микросателлитные локусы, или SSRs (Simple Sequence Repeats). Полногеномное секвенирование значительно облегчает их поиск и разработку. Статья посвящена разработке новых микросателлитных маркеров для сосны обыкновенной (*Pinus sylvestris* L.). Из полученной черновой геномной сборки сосны обыкновенной были отобраны несколько тысяч контигов, содержащих микросателлитные локусы с нуклеотидными мотивами. Поиск был специально сфокусирован на локусах с длиной мотива более двух нуклеотидов как наиболее надежных для генотипирования даже в простом гель-электрофорезе. Всего протестировано 39 пар праймеров, из них отобраны 6 локусов с три-, тетра- и пентануклеотидными повторами, показавших высокий уровень полиморфизма, надежное генотипирование. Они были дополнительно протестированы в двух популяциях из Северо-Енисейского и Курагинского районов Красноярского края и сравнены с другими популяциями и наборами маркеров по опубликованным данным. Разработанные маркеры могут быть использованы в различных популяционно-генетических исследованиях и для идентификации происхождения древесины и растительного материала.

**Ключевые слова:** *Pinus sylvestris* L., генетическое разнообразие, полногеномное секвенирование, гетерозиготность, хвойные, популяционно-генетический анализ, микросателлитные маркеры, SSRs.

DOI: 10.15372/SJFS20250103

### ВВЕДЕНИЕ

Сохранение генетического разнообразия основных лесообразующих пород, в том числе сосны обыкновенной (*Pinus sylvestris* L.), широко распространенного древесного вида бореаль-

ных лесов, является острой проблемой современности. В результате стрессового воздействия природных и антропогенных факторов среды, фрагментации ареала и нарушения естественно-исторической структуры популяций снижаются генетическое разнообразие и устойчивость

популяций. Для успешного решения проблемы сохранения их генетического разнообразия необходимы постоянный мониторинг и изучение популяционно-генетической структуры основных лесобразующих видов (Fussi et al., 2016; Bajc et al., 2020). Одним из наиболее информативных способов получения данных о генетической изменчивости являются ДНК-маркеры, позволяющие определить внутривидовую аллельную и генную изменчивость, дифференциацию на различных иерархических уровнях, а также степень инбридинга (Porth, El-Kassaby, 2014). К наиболее полиморфным ДНК-маркерам относятся микросателлиты, или SSRs (Simple Sequence Repeats). Данные генетические маркеры обладают рядом преимуществ: высоким уровнем полиморфизма, относительно равномерным распределением в геноме и широкой представленностью, наличием большого числа аллелей, кодоманантностью (Merritt et al., 2015; Nowakowska, 2016; Portela et al., 2024).

Разработка микросателлитных маркеров для различных хвойных пород с использованием транскриптомных и полногеномных данных активно осуществляется. Так, получены SSR-маркеры для тиса остроконечного (*Taxus cuspidata* Siebold & Zucc. ex Endl.) (Ueno et al., 2015), т. западно-гималайского (*T. contorta* Griff.) (Majeed et al., 2019), секвойи вечнозеленой (*Sequoia sempervirens* (D. Don) Endl.) (Breidenbach et al., 2019), лиственницы сибирской (*Larix sibirica* Ledeb.) (Орешкова и др., 2017, 2019; Oreshkova et al., 2017, 2019), сосны сибирской (*Pinus sibirica* Du Tour) (Белоконь и др., 2016; Belokon' et al., 2016), пихты сибирской (*Abies sibirica* Ledeb.) и п. белой (*A. alba* Mill.) (Oreshkova et al., 2023).

Сосна обыкновенная относится к основным лесобразующим породам Евразии. Известно, что ее древесина обладает хорошими физико-механическими свойствами, также она используется как источник живицы, из сосновых опилок делают гидролизный спирт и т. д. Значительное негативное влияние на рост, развитие и жизнеспособность популяций этого вида оказывают антропогенные факторы (загрязнения, урбанизация, фрагментация ареала, сплошные вырубki и т. д.), что при достижении определенного предела может привести к их гибели (Richardson, 1998; Farjon, 2005; Tóth et al., 2017). Экологический стресс в условиях техногенеза значительно влияет на репродуктивные способности сосны обыкновенной, в частности приводит к образованию пыльцы низкого ка-

чества и снижению семенной продуктивности (Носкова, Третьякова, 2006). В связи с высоким экономическим значением сосны обыкновенной и ее подверженностью негативным техногенным влияниям необходимы регулярные мониторинговые наблюдения популяций данной хвойной породы. Одним из эффективных способов определения влияния промышленных выбросов представляется исследование генетического разнообразия и популяционно-генетической структуры вида. Как было отмечено выше, высокополиморфные микросателлитные локусы могут использоваться для мониторинга генетической изменчивости хвойных. Однако имеется объективная необходимость в разработке новых удобных, надежных и простых в анализе генетических маркеров, применимых к исследованиям популяционно-генетической структуры вида даже в просто оборудованных лабораториях. Большинство существующих микросателлитных маркеров для сосны обыкновенной – это локусы с динуклеотидными мотивами, генотипирование с помощью которых путем проведения простого агарозного или полиакриламидного гель-электрофореза крайне затруднительно и требуется капиллярный электрофорез с использованием относительно дорогостоящего оборудования для фрагментного анализа. Целью данного исследования стала разработка простых для анализа новых микросателлитных маркеров с три-, тетра и пентануклеотидными мотивами для сосны обыкновенной на основе данных полногеномного *de novo* секвенирования в дополнение к используемым (Sebastiani et al., 2012; Fang et al., 2014; Żukowska et al., 2023).

## МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЙ

**Популяции.** Биологическим материалом для отработки методов послужили образцы хвои сосны обыкновенной, предоставленные сотрудниками Центра защиты леса Красноярского края и собранные из двух популяций: южной (выборка Кизирская) и северной (выборка Северо-Енисейская), расположенных соответственно в Северо-Енисейском и Курагинском районах Красноярского края. Названия популяций и выборок даны по названию местных лесничеств, в которых они отобраны (табл. 1).

**Поиск микросателлитных маркеров.** Для поиска *in silico* микросателлитных локусов и дизайна праймеров использовался геном сосны обыкновенной, асемблированный *de novo*

**Таблица 1.** Местоположение и географические координаты двух изученных популяций сосны обыкновенной в Красноярском крае

Популяции (лесничество)	Район	Квартал	Выдел	Географические координаты, с. ш. / в. д.
Северо-Енисейская	Северо-Енисейский	73	33	59°22' / 93°39'
Кизирская	Курагинский	297	24	54°16' / 93°42'

**Таблица 2.** Условия отбора микросателлитных локусов сосны обыкновенной

Мотив	Минимальное пороговое число повторов мотива
3-нуклеотидный	22
4-нуклеотидный	13
5-нуклеотидный	10

с помощью ПО CLC Assembly Cell в лаборатории лесной геномики Сибирского федерального университета. Суммарная длина сборки равна 8.957 млрд нуклеотидных оснований (н. о.), число скаффолдов – 21.44 млн,  $N_{50} = 510$  н. о. Поиск *in silico* микросателлитных локусов выполнялся программой MISA (Beier et al., 2017), дизайн праймеров – программой Primer 3 (Rozen, Skaletsky, 1999; Untergasser et al., 2012). Микросателлитные локусы были отфильтрованы по условиям, показанным в табл. 2.

Микросателлитные локусы отбирались по расположению не ближе, чем 20 нуклеотидов от начала и конца контига, для того, чтобы имелась возможность подобрать ПЦР-праймеры. Расположение сайтов отжига для ПЦР праймеров выбрано так, чтобы размер амплифицированных фрагментов, включающих микросателлитный локус, был в диапазоне 140–280 н. о. В итоге для дальнейшего тестирования выбрано 39 пар праймеров.

**Тестирование и отбор полиморфных микросателлитных локусов.** Образцы тотальной ДНК выделены модифицированным методом с применением цетилтриметиламмонийбромида (СТАВ) из образцов хвои, собранной с индивидуальных деревьев и высушенной с помощью силикагеля (Doyle J. J., Doyle J. L., 1990). Для ПЦР использовали готовые реагенты GenePak PCR Core производства ООО «Лаборатория Изоген» (Москва, Россия). Продукты амплификации разделяли путем электрофореза в 6%-м полиакриламидном геле с использованием трис EDTA-боратного электродного буфера в камерах для вертикального электрофореза. После окончания электрофореза гели окрашивали в растворе бромистого этидия с последующей

визуализацией в ультрафиолетовом свете. В качестве маркера стандартных длин использовали ДНК плазмиды pBR322 кишечной палочки (*Escherichia coli* (Migula)), обработанную рестриктазой Hpa II. Размер фрагментов определяли путем сопоставления с размером стандартных маркеров в программе Photo-Capt (Vilber Lourmat, Marnes-la-Valle, France).

Праймеры тестировали путем подбора и оптимизации параметров программы ПЦР. После оптимизации условий амплификации каждая пара праймеров тестировалась на четырех образцах, принадлежащих к одной из выборок сосны обыкновенной с целью отбора наиболее успешных праймеров и локусов для дальнейших исследований на большей популяционной выборке. Амплификацию ядерных микросателлитных локусов сосны обыкновенной проводили в режиме ПЦР с *touchdown* – последовательным снижением температуры в циклах, что должно уменьшать вероятность неспецифической амплификации (Korbie, Mattick, 2008) (табл. 3).

По результатам тестирования отобраны шесть полиморфных, надежно генотипируемых локусов, демонстрирующих хорошо интерпретируемые электрофореграммы. Отобранные SSR-маркеры в дальнейшем использовались для исследования полиморфизма выборок из двух популяций – Северо-Енисейской и Кизирской. Для них рассчитаны основные параметры оценки уровня генетического разнообразия, популяционной структуры и степени генетической подразделенности популяций.

**Параметры генетического разнообразия.** Для оценки уровня генетического разнообразия рассчитаны следующие параметры: число аллелей на локус ( $N_A$ ), наблюдаемая ( $H_o$ ) и ожидаемая ( $H_e$ ) гетерозиготности, эффективное число аллелей ( $N_e$ ) и индексы фиксации ( $F_{IS}$ ,  $F_{IT}$  и  $F_{ST}$ ) с помощью программы GenAlEx 6.5 (Peakall, Smouse, 2006). Степень дифференциации и индексы фиксации вычислены с использованием AMOVA (*analysis of molecular variance*) – подхода в программе GenAlEx 6.5. В дополнение к обычному индексу дифференциации  $F_{ST}$  рас-

**Таблица 3.** Программа ПЦР-амплификации образцов ДНК сосны обыкновенной

Этап амплификации	Температура, °С	Время	Число циклов
Первичная денатурация	94	1 мин	1
Денатурация	94	30 с	9
Отжиг праймеров	60 с понижением на 1 °С каждый цикл (до 50)	30 с	
Элонгация	72	1 мин	
Денатурация	94	30 с	
Отжиг праймеров	50	30 с	24
Элонгация	72	30 с	
Финальная элонгация	72	10 мин	1
Охлаждение	4	–	–

считан стандартизированный индекс (Meirmans, 2006), более точно оценивающий дифференциацию (Meirmans, Hedrick, 2011). Проверка на наличие *null* аллелей и соответствующая корректировка параметров осуществлены с помощью программы Micro-Checker (Van Oosterhout et al., 2004, 2006).

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ И ИХ ОБСУЖДЕНИЕ

**Поиск, тестирование и отбор полиморфных микросателлитных локусов.** Первичный тест показал, что из 39 первоначально отобранных пар праймеров для 16 отсутствовала амплификация, у 7 наблюдалась нестабильная амплификация, 9 давали мономорфный продукт, одна пара – сложный спектр, возможно вызванный амплификацией более чем одного локуса, и только 6 пар праймеров демонстрировали стабильные и надежно генотипируемые спектры (табл. 4).

Для примера фотографии гелей после электрофореза продуктов ПЦР четырех локусов показаны на рисунке.

**Генетическая изменчивость двух популяций сосны обыкновенной по шести новым ядерным микросателлитным локусам.** С помощью отобранных 6 локусов (табл. 4) изучена генетическая изменчивость двух выборок сосны обыкновенной из Северо-Енисейской и Кизинской популяций, различающихся по составу и частотам встречаемости выявленных микросателлитных аллелей, достоверных для четырех локусов по значению теста  $\chi^2$  на гетерогенность (табл. 5).

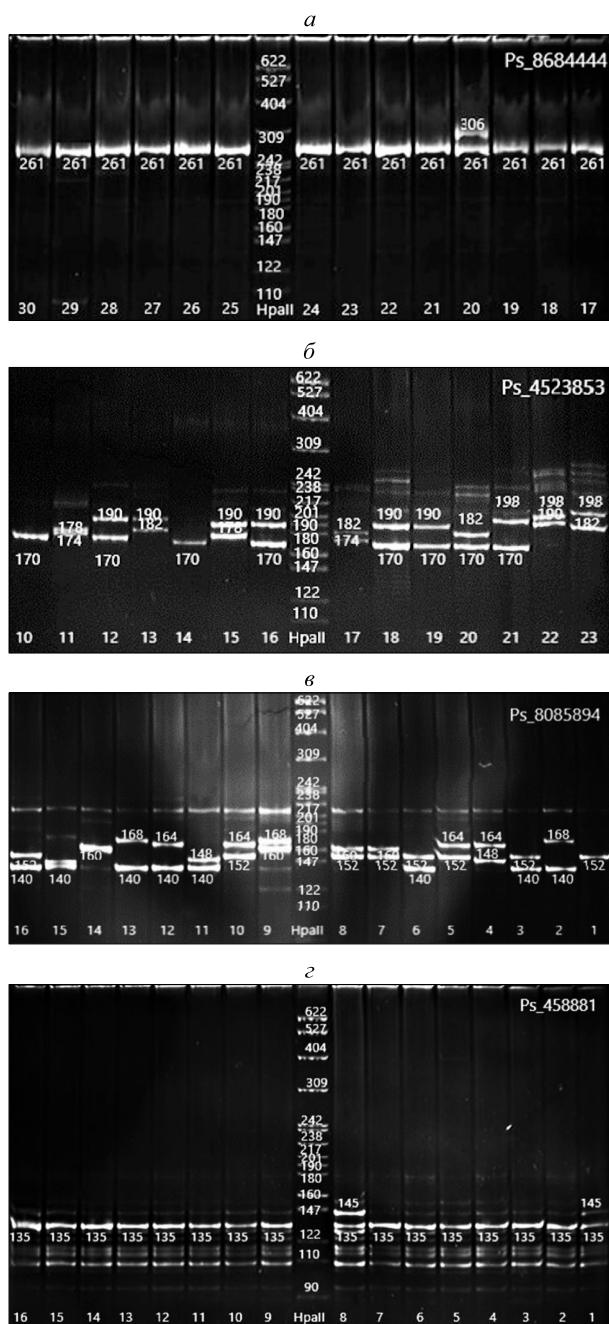
Всего выявлено 34 аллельных варианта (табл. 5). В выборке Северо-Енисейская было обнаружено 29 аллелей (из них 4 редких, частота встречаемости которых менее 5%), а в выборке Кизинская – 30 (из них 10 – редкие).

В локусе *Ps\_3015784* в обеих выборках выявлены скрытые *null*-аллели. В локусе

**Таблица 4.** Отобранные микросателлитные локусы сосны обыкновенной

Локус	Мотив*	Нуклеотидные последовательности первого (F) и второго (R) праймеров в паре	Число аллелей	Длина амплифицированных фрагментов, н. о.
<i>Ps_8684444</i>	(ATG)22	F: TGAAGATGTCATGTTGCACTTTCT R: TAACGTCGCCATCGACAACA	3	261–306
<i>Ps_13653311</i>	(ACAT)17	F: ACCAGGAACCTCTTCATGTCA R: GGTCACCACGGATCACGATT	3	184–208
<i>Ps_4523853</i>	(TATG)17	F: CGCATGTAAAGGTTATGGCCTC R: TGGTAGATTGTAGTGCATTGCC	9	170–206
<i>Ps_8085894</i>	(TATG)15	F: ACCAACTCCTAGGTATCTGGTAA R: TCCAGCACCGTCAACACAAT	6	140–168
<i>Ps_3015784</i>	(CATA)15	F: AGCATAGAACCAACAAGCAACT R: TGTTGGAGGGGCTAGTCGAT	10	244–308
<i>Ps_458881</i>	(AAAGG)11	F: AGGAGAAATCACCAAGACGAGA R: GCTGAAGCTACTTCATCCTTGC	3	115–145

\* Число повторов (после скобок) для каждого локуса соответствует числу повторов каждого мотива в исходном контиге, содержащем этот локус.



Фотография гелей после электрофореза продуктов ПЦР ядерных микросателлитных локусов сосны обыкновенной: *Ps\_8684444*(а), *Ps\_4523853* (б), *Ps\_8085894* (в), *Ps\_458881* (г).

Трехзначные цифры на фото обозначают длину амплифицируемых ДНК-фрагментов, соответствующих аллелям данных локусов. Нижний ряд одно- и двухзначных цифр обозначает номера образцов индивидуальных деревьев. Hра II – маркер стандартных длин фрагментов с цифрами, обозначающими их размер.

*Ps\_4523853* они обнаружены только в выборке Северо-Енисейская. Частоты скрытых *null*-аллелей рассчитаны с помощью программы Micro-Checker (Van Oosterhout et al., 2004, 2006), исходя из предположения о том, что популяции

находятся в равновесном состоянии согласно правилу Харди-Вайндберга (Van Oosterhout et al., 2004, 2006).

В исследованных выборках сосны эти локусы имели самый высокий уровень аллельного разнообразия – по 10 и 9 аллелей соответственно (табл. 5).

Для выявления уровня генетического разнообразия в каждой из двух выборок сосны обыкновенной рассчитаны основные показатели генетической изменчивости на основании аллельных частот 6 локусов (табл. 6).

Расчет основных параметров генетической изменчивости показал, что включенные в исследование выборки сосны обыкновенной схожи по уровню генетического разнообразия.

Наиболее высокое значение среднего числа аллелей на локус выявлено в выборке Кизирская из более южного района ( $N_a = 5.0 \pm 1.2$ ), а эффективного числа аллелей на локус – в выборке Северо-Енисейская ( $N_e = 3.1 \pm 0.9$ ). Самые высокие значения наблюдаемой и ожидаемой в гетерозиготности отмечены в выборке Северо-Енисейская ( $H_o = 0.472 \pm 0.132$ ;  $H_e = 0.504 \pm 0.134$ ).

Различия по средним значениям данных генетических показателей для этих выборок не достоверны согласно рассчитанному *t*-критерию Стьюдента.

В выборке сосны обыкновенной Северо-Енисейская выявлен небольшой дефицит гетерозиготных генотипов ( $F_{IS} = 0.077 \pm 0.045$ ). Выборка Кизирская находится близко к равновесному состоянию ( $F_{IS} = -0.018 \pm 0.053$ ).

Согласно *t*-критерию Стьюдента, различие между выборками по индексу фиксации является достоверными с уровнем значимости  $p < 0.001$ .

С помощью *F*-статистик на основе анализа молекулярной дисперсии (AMOVA) оценена степень генетической подразделенности двух исследованных природных популяций сосны обыкновенной (табл. 7).

Для каждого из 6 полиморфных локусов рассчитаны индексы фиксации внутри популяции ( $F_{IS}$ ), внутри общей (тотальной) выборки ( $F_{IT}$ ) и между популяциями ( $F_{ST}$ ). Обычный индекс дифференциации очевиден почти вдвое и более занижен по сравнению со стандартизированным индексом  $F'_{ST}$ .

Анализ популяционной структуры сосны обыкновенной (табл. 7) показал, что в изучаемых выборках наблюдается небольшой, но достоверный 7 %-й дефицит гетерозиготных генотипов ( $F_{IS} = 0.072$ ,  $p < 0.023$ ). Полокусные значения

**Таблица 5.** Частота аллелей в исследованных выборках двух популяций сосны обыкновенной и значение теста  $\chi^2$  на гетерогенность и его достоверность ( $p$ )

Локус	Аллель	Популяция		$\chi^2$	$p$
		Северо-Енисейская	Кизирская		
<i>Ps_8684444</i>	261	0.950	0.935	0.191	0.979
	279	–	0.022		
	306	0.050	0.043		
<i>Ps_1365311</i>	184	0.083	0.087	1.727	0.631
	196	0.900	0.826		
	208	0.017	0.087		
<i>Ps_4523853</i>	170	0.150	0.392	87.916	$3.11 \times 10^{-6}$
	174	–	0.043		
	178	0.083	0.043		
	182	0.217	0.196		
	186	0.017	–		
	190	0.233	0.239		
	198	0.150	0.087		
	206	0.033	–		
<i>null</i>	0.117	–			
<i>Ps_8085894</i>	140	0.184	0.087	32.084	0.006
	148	0.083	0.087		
	152	0.450	0.652		
	160	0.083	0.109		
	164	0.133	0.043		
	168	0.067	0.022		
<i>Ps_3015784</i>	244	0.117	0.217	121.573	$5.7 \times 10^{-9}$
	248	–	0.043		
	256	0.283	0.109		
	264	0.217	0.152		
	268	–	0.022		
	276	0.183	0.109		
	284	0.050	0.239		
	296	0.033	0.022		
	308	–	0.022		
<i>null</i>	0.117	0.065			
<i>Ps_458881</i>	115	0.050	–	22.948	$4.14 \times 10^{-5}$
	135	0.767	0.913		
	145	0.183	0.087		
Число аллелей	34	29 (4)*	30 (10)*		

\* В скобках показано число редких аллелей с частотой встречаемости менее 5 %.

показывают, что наиболее значимый дефицит гетерозигот наблюдается в локусах *Ps\_3015784* и *Ps\_458881*. Уровень генетической дифференцированности между популяциями ( $F_{ST}$ ) относительно небольшой, но достоверный, варьирует от  $-0.015$  (*Ps\_8684444*) до  $0.036$  (*Ps\_458881*), составляя в среднем  $0.026$  ( $p < 0.001$ ). Это свидетельствует о том, что только 2.6 % выявленной генетической изменчивости изученных выборок сосны объясняется различием между популяци-

ями. Внутри популяций сосредоточено 97.4 % всего генетического разнообразия. Наибольший вклад в дифференциацию изученных популяций вносят локусы *Ps\_4523853*, *Ps\_3015784* и *Ps\_458881* (табл. 7).

Анализ данных по генетическому полиморфизму выборок из популяций сосны обыкновенной показал относительно высокий уровень генетического разнообразия, но несколько меньший при сопоставлении с аналогичными пока-

**Таблица 6.** Генетическая изменчивость выборок из популяций сосны обыкновенной по шести микросателлитным локусам

Популяция	Локус	$N_a$	$N_e$	$H_o$	$H_e$	$F_{IS}$
Северо-Енисейская	<i>Ps_8684444</i>	2	1.1	0.100	0.095	-0.053
	<i>Ps_1365311</i>	3	1.2	0.133	0.183	0.271
	<i>Ps_4523853*</i>	8	5.9	0.800	0.832	0.038***
	<i>Ps_8085894</i>	6	3.7	0.667	0.728	0.084*
	<i>Ps_3015784*</i>	7	5.2	0.800	0.808	0.010
	<i>Ps_458881</i>	3	1.6	0.333	0.376	0.114**
В среднем...		$4.8 \pm 1.0$	$3.1 \pm 0.9$	$0.472 \pm 0.132$	$0.504 \pm 0.134$	$0.077 \pm 0.045$
Кизирская	<i>Ps_8684444</i>	3	1.1	0.130	0.124	-0.053
	<i>Ps_1365311</i>	3	1.4	0.348	0.302	-0.150
	<i>Ps_4523853</i>	6	3.9	0.739	0.740	0.001
	<i>Ps_8085894</i>	6	2.2	0.565	0.545	-0.036
	<i>Ps_3015784*</i>	10	6.3	0.652	0.841	0.225***
	<i>Ps_458881</i>	2	1.2	0.174	0.159	-0.095
В среднем...		$5.0 \pm 1.2$	$2.7 \pm 0.8$	$0.435 \pm 0.104$	$0.452 \pm 0.124$	$-0.018 \pm 0.053$
В среднем для обеих популяций...		$4.9 \pm 0.8$	$2.9 \pm 0.6$	$0.454 \pm 0.080$	$0.478 \pm 0.087$	$0.030 \pm 0.036$

Примечание.  $N_a$  – число аллелей на локус;  $N_e$  – эффективное число аллелей на локус;  $H_o$  – наблюдаемая гетерозиготность;  $H_e$  – ожидаемая гетерозиготность;  $F_{IS}$  – индекс фиксации;  $\pm$  – стандартная ошибка. \* Предполагается наличие *null*-аллелей.

**Таблица 7.** Показатели  $F$ -статистик для сосны обыкновенной

Локус	$n$	$F_{IS}$	$F_{IT}$	$F_{ST}$	$F'_{ST}$
<i>Ps_8684444</i>	3	-0.034	-0.050	-0.015	-0.017
<i>Ps_1365311</i>	3	0.055	0.056	0.001	0.002
<i>Ps_4523853</i>	9	0.042	0.070	0.029*	0.149
<i>Ps_8085894</i>	6	0.059	0.083	0.025	0.072
<i>Ps_3015784</i>	10	0.124*	0.152**	0.032*	0.203
<i>Ps_458881</i>	3	0.082	0.114	0.036	0.050
В среднем	5.7	0.072*	0.096**	0.026***	0.051

Примечание.  $N$  – число аллелей;  $F_{IS}$  – индивидуальный коэффициент инбридинга особи внутри популяции;  $F_{IT}$  – общий коэффициент инбридинга особи относительно тотальной выборки обеих популяций;  $F_{ST}$  – коэффициент межпопуляционной дифференциации (степень генетической дифференциации между популяциями);  $F'_{ST}$  – стандартизированный индекс. \*  $p < 0.05$ . \*\*  $p < 0.01$ . \*\*\*  $p < 0.001$ .

зателями, полученными другими авторами для сосны обыкновенной в основном на основе микросателлитных маркеров с динуклеотидными мотивами (Ильинов, Раевский, 2015, 2016; Шуваев и др., 2020, 2022; Żukowska et al., 2023).

Так, при исследовании полиморфизма в популяциях сосны обыкновенной в средней подзоне тайги Карелии и в Южно-Карельском лесосеменном районе количество аллелей на локус варьировало от 7.50 до 10.25, ожидаемая гетерозиготность ( $H_o$ ) – от 0.51 до 0.68 на локус, а наблюдаемая ( $H_e$ ) – от 0.28 до 0.50. Анализ подразделенности карельских популяций на основе  $F$ -статистик Райта (Wright, 1965) выявил дефицит гетерозигот как на популяционном уровне ( $F_{IS} = 0.25$ ), так относительно вида в целом ( $F_{IT} = 0.32$ ) (Ильинов, Раевский, 2015), что от-

личает эти популяции сосны обыкновенной от исследованных нами популяций Курагинского и Северо-Енисейского районов Красноярского края, находящихся ближе к равновесному состоянию. При исследовании авторами малонарушенных естественных сосняков северной и средней подзон тайги Карелии в Карельском и Южно-Карельском лесосеменных районах получены сравнительно близкие уровни аллельного разнообразия в популяциях и дефицита гетерозиготных генотипов как относительно популяции, так и вида в целом ( $F_{IS} = 0.055$ ;  $F_{IT} = 0.114$ ) (Ильинов, Раевский, 2016).

Популяции сосны обыкновенной с Карпатских гор и Паннонского бассейна характеризуются чуть более высокими по сравнению с исследованными нами популяциями пока-

зателями уровня ожидаемой гетерозиготности  $H_e$ , которые варьируют от 0.511 до 0.615 (Liewlaksaneeyanawin et al., 2004).

В работах Д. Н. Шуваева с соавт. (2020, 2022) также отмечается немного повышенный уровень гетерозиготности у сосны обыкновенной из Красноярского края ( $H_o = 0.519-0.610$ ;  $H_e = 0.550-0.612$ ). Однако уровень подразделенности популяций в исследуемой части ареала оказался такой же относительно низкий ( $F_{ST} = 0.026$ ), что свидетельствует о низкой степени генетической дифференциации между популяциями.

Повышенные значения основных показателей генетической изменчивости, полученные другими авторами, в первую очередь связаны с разными наборами микросателлитных локусов, используемых в исследованиях. Для анализа популяционно-генетической изменчивости сосны обыкновенной авторами использовались микросателлитные локусы, разработанные как для нее, так и для других видов этого рода – с. ладанной (*P. taeda* L.), с. приморской (*P. pinaster* Aiton) и в основном с динуклеотидными мотивами как значительно более полиморфными, чем локусы с более длинными мотивами (Ильинов, Раевский, 2015, 2016; Liewlaksaneeyanawin et al., 2004; Шуваев и др., 2020, 2022; Zukowska et al., 2023). В своей работе мы специально использовали локусы с длинными мотивами, разработанные именно для вида сосны обыкновенной и пригодные для анализа на простом геле-электрофорезе.

Выявленный дефицит гетерозиготных генотипов в выборке Северо-Енисейская ( $F_{IS} = 0.077 \pm 0.045$ ) вряд ли можно объяснить только инбридингом и самоопылением, поскольку половина локусов не имела существенного дефицита гетерозигот. Хотя в определенной степени близкородственное скрещивание может быть вызвано в Северо-Енисейском районе наличием локальных вырубок, инициированных местным населением, и пожарами как природного, так и антропогенного характера с последующим восстановлением популяций сосны обыкновенной из ограниченного числа сохранившихся деревьев.

## ЗАКЛЮЧЕНИЕ

Из 39 протестированных пар праймеров, разработанных для амплификации микросателлитных локусов и основанных на контигах черновой сборки сосны обыкновенной, содержащих tandemные простые последовательности нукле-

отидных повторов, для амплификации были отобраны 6 наиболее перспективных пар средне- и высокополиморфных микросателлитных локусов с три- (*Ps\_8684444*), тетра- (*Ps\_13653311*, *Ps\_4523853*, *Ps\_8085894*, *Ps\_3015784*) и пентануклеотидными (*Ps\_458881*) мотивами. Полученные уровни изменчивости были достаточно высокими, но несколько меньшими по сравнению с показателями, полученными для сходных сибирских популяций сосны обыкновенной на основе микросателлитных локусов с моно- и динуклеотидными мотивами. Это было ожидаемо, но данный недостаток компенсируется более надежным генотипированием и возможностью использования более простого геле-электрофореза с недорогим оборудованием, что важно для быстрого тест-анализа в полевых условиях. Эти локусы могут быть использованы в дальнейшем в качестве генетических маркеров для индивидуальной генетической идентификации и в популяционно-генетических исследованиях сосны обыкновенной. Нуклеотидные последовательности контигов, использованных для проектирования праймеров, могут быть предоставлены авторами по запросу.

*Работа выполнена в рамках государственного задания ФИЦ КНЦ СО РАН по теме «Геномные, популяционные и метагеномные исследования лесных биоценозов. Таксономический и филогенетический анализ растительных сообществ лесных экосистем» № FWES 2025 0015.*

*Авторы выражают благодарность магистранту кафедры геномики и биоинформатики ИФБиТ СФУ Яну Викторовну Стенину за участие в лабораторных исследованиях и помощь в обработке полученных данных, а также признательность сотрудникам Центра защиты леса Красноярского края Алексею Александровичу Ибе и Елене Алексеевне Шилкиной за предоставленный биологический материал (хвоя) двух выборок сосны обыкновенной.*

## СПИСОК ЛИТЕРАТУРЫ

- Белоконь, М. М., Политов Д. В., Мудрик Е. А., Полякова Т. А., Шатохина А. В., Белоконь Ю. С., Орешкова Н. В., Путинцева Ю. А., Шаров В. В., Кузьмин Д. А., Крутовский К. В. Разработка микросателлитных маркеров сосны кедровой сибирской (*Pinus sibirica* Du Tour) по результатам полногеномного *de novo* секвенирования // Генетика. 2016. Т. 52. № 12. С. 1418–1427.
- Ильинов А. А., Раевский Б. В. Сравнительная оценка генетического разнообразия естественных популяций и клоновых плантаций сосны обыкновенной и ели



- финской в Карелии // Экол. генет. 2015. Т. 13. № 4. С. 55–67.
- Ильинов А. А., Раевский Б. В. Состояние генофонда сосны обыкновенной *Pinus sylvestris* L. в Карелии // Сиб. лесн. журн. 2016. № 5. С. 45–54.
- Носкова Н. Е., Третьякова И. Н. Влияние стресса на репродуктивные способности сосны обыкновенной // Хвойные бореал. зоны. 2006. Т. 23. № 3. С. 54–63.
- Орешкова Н. В., Бондар Е. И., Путинцева Ю. А., Шаров В. В., Кузьмин Д. А., Крутовский К. В. Разработка ядерных микросателлитных маркеров с длинными (трех-, четырех-, пяти- и шестинуклеотидными) мотивами для трех видов лиственницы на основе полногеномного *de novo* секвенирования лиственницы сибирской (*Larix sibirica* Ledeb.) // Генетика. 2019. Т. 55. № 4. С. 418–425.
- Орешкова Н. В., Путинцева Ю. А., Шаров В. В., Кузьмин Д. А., Крутовский К. В. Разработка микросателлитных маркеров лиственницы сибирской (*Larix sibirica* Ledeb.) на основе полногеномного *de novo* секвенирования // Генетика. 2017. Т. 53. № 11. С. 1278–1284.
- Шуваев Д. Н., Ибе А. А., Щерба Ю. Е., Сухих Т. В., Шилкина Е. А., Усова Е. А., Лисотова Е. В., Ренях М. В., Ступакова О. М. Разработка панели ядерных микросателлитных локусов для оценки легальности происхождения древесины сосны обыкновенной в Красноярском крае // Хвойные бореал. зоны. 2020. Т. 38. № 5–6. С. 297–304.
- Шуваев Д. Н., Ибе А. А., Щерба Ю. Е., Сухих Т. В., Шилкина Е. А., Шеллер М. А., Усова Е. А., Лисотова Е. В., Ренях М. В., Ступакова О. М. Молекулярно-генетическая характеристика популяций сосны обыкновенной в Красноярском крае по панели 15 ядерных микросателлитных локусов // Лесоведение. 2022. № 5. С. 530–539.
- Bajc M., Aravanopoulos F. A., Westergren M., Fussi B., Kavaliauskas D., Alizoti P., Kiourtsis F., Kraigher H. (Eds.) Manual for forest genetic monitoring. 1<sup>st</sup> ed. Ljubljana: Silva Slovenica Publ. Centre, 2020. 326 p.
- Beier S., Thiel T., Münch T., Scholz U., Mascher M. MISA-web: a web server for microsatellite prediction // Bioinformatics. 2017. V. 33. N. 16. P. 2583–2585.
- Belokon' M. M., Politov D. V., Mudrik E. A., Polyakova T. A., Shatokhina A. V., Belokon' Yu. S., Oreshkova N. V., Putintseva Yu. A., Sharov V. V., Kuz'min D. A., Krutovsky K. V. Development of microsatellite genetic markers in Siberian stone pine (*Pinus sibirica* Du Tour) based on the *de novo* whole genome sequencing // Rus. J. Genet. 2016. V. 52. Iss. 12. P. 1263–1271 (Original Rus. Text © M. M. Belokon', D. V. Politov, E. A. Mudrik, T. A. Polyakova, A. V. Shatokhina, Yu. S. Belokon', N. V. Oreshkova, Yu. A. Putintseva, V. V. Sharov, D. A. Kuz'min, K. V. Krutovsky, 2016, publ. in Genetika. 2016. V. 52. N. 12. P. 1418–1427.).
- Breidenbach N., Gailing O., Krutovsky K. V. Development of novel polymorphic nuclear and chloroplast microsatellite markers in coast redwood (*Sequoia sempervirens*) // Plant Genet. Res.: Characterization and Utilization. 2019. V. 17. N. 3. P. 293–297.
- Doyle J. J., Doyle J. L. Isolation of plant DNA from fresh tissue // Focus. 1990. V. 12. N. 1. P. 13–15.
- Fang P., Niu Sh., Yuan H., Li Zh., Zhang Y., Yuan L., Li W. Development and characterization of 25 EST-SSR markers in *Pinus sylvestris* var. *mongolica* (Pinaceae) // Appl. Plant Sci. 2014. V. 2. Iss. 1. Article number 1300057.
- Farjon A. Pines: drawings and descriptions of the genus *Pinus*. 2<sup>nd</sup> ed. Leiden, Boston: Brill, 2005. 235 p.
- Fussi B., Westergren M., Aravanopoulos F., Baier R., Kavaliauskas D., Finzgar D., Alizoti P., Bozic G., Avramidou E., Konnert M., Kraigher H. Forest genetic monitoring: an overview of concepts and definitions // Environ. Monitor. Assess. 2016. V. 188. Article number 493. 12 p.
- Korbie D. J., Mattick J. S. Touchdown PCR for increased specificity and sensitivity in PCR amplification // Nat. Protocols. 2008. V. 3. N. 9. P. 1452–1456.
- Liewlaksaneeyanawin C., Ritland C. E., El-Kassaby Y. A., Ritland K. Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs // Theor. Appl. Genet. 2004. V. 109. N. 2. P. 361–369.
- Majeed A., Singh A., Choudhary Sh., Bhardwaj P. Transcriptome characterization and development of functional polymorphic SSR marker resource for Himalayan endangered species, *Taxus contorta* (Griff) // Industr. Crops Products. 2019. V. 140. Article number 111600.
- Meirmans P. G. Using the AMOVA framework to estimate a standardized genetic differentiation measure // Evolution. 2006. V. 60. N. 11. P. 2399–2402.
- Meirmans P. G., Hedrick P. W. Assessing population structure: FST and related measures // Molecul. Ecol. Res. 2011. V. 11. N. 1. P. 5–18.
- Merritt B. J., Culley T. M., Avanesyan A., Stokes R., Brzyski J. An empirical review: characteristics of plant microsatellite markers that confer higher levels of genetic variation // Appl. Plant Sci. 2015. V. 3. N. 8. Article number 1500025.
- Nowakowska J. Microsatellite markers in analysis of forest-tree populations // Microsatellite markers / I. Y. Abdurakhmonov (Ed.). Open Sci. INTECH, 2016. P. 95–116.
- Oreshkova N. V., Bondar E. I., Putintseva Yu. A., Sharov V. V., Kuzmin D. A., Krutovsky K. V. Development of nuclear microsatellite markers with long (tri-, tetra-, penta- and hexanucleotide) motifs for three larch species based on the *de novo* whole genome sequencing of Siberian larch (*Larix sibirica* Ledeb.) // Rus. J. Genet. 2019. V. 55. N. 4. P. 444–450 (Original Rus. Text © N. V. Oreshkova, E. I. Bondar, Yu. A. Putintseva, V. V. Sharov, D. A. Kuzmin, K. V. Krutovsky, 2019, publ. in Genetika. 2019. V. 55. N. 4. P. 418–425).
- Oreshkova N. V., Bondar E. I., Sharov V. V., Dhungana S. P., Gailing O., Krutovsky K. V. Population genetic variation of microsatellite markers developed for Siberian fir (*Abies sibirica* Ledeb.) and European silver fir (*Abies alba* Mill.) using whole genome sequencing data // Plant Genet. Res.: Characterization and Utilization. 2023. V. 21. N. 2. P. 149–158.
- Oreshkova N. V., Putintseva Yu. A., Sharov V. V., Kuzmin D. A., Krutovsky K. V. Development of microsatellite genetic markers in Siberian larch (*Larix sibirica* Ledeb.) based on the *de novo* whole genome sequencing // Rus. J. Genet. 2017. V. 53. N. 11. P. 1194–1199 (Original Rus. Text © N. V. Oreshkova, Yu. A. Putintseva, V. V. Sharov, D. A. Kuzmin, K. V. Krutovsky, 2017, publ. in Genetika. 2017. V. 53. N. 11. P. 1278–1284).

- Peakall R. O. D., Smouse P. E. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research // *Molecul. Ecol. Not.* 2006. V. 6. N. 1. P. 288–295.
- Portela R. M., Medalha L. M. D. S., Quadros J. F. D. S., De Lara J. H., Feitosa Q. A., Costa J. G., Paludeto J. G. Z., Gomes C. A. F. C. Microsatellite markers in the sustainable management of forest genetic resources // *IOSR J. Business and Manag.* 2024. V. 26. Iss. 2. Ser. 7. P. 37–45.
- Porth I., El-Kassaby Y. A. Assessment of the genetic diversity in forest tree populations using molecular markers // *Diversity*. 2014. V. 6. N. 2. P. 283.
- Richardson D. M. (Ed.). *Ecology and biogeography of Pinus*. Cambridge, UK: Cambridge Univ. Press, 1998. 527 p.
- Rozen S., Skaletsky H. Primer3 on the WWW for general users and for biologist programmers // *Bioinformatics methods and protocols. Methods in molecular biology* / S. Misener, S. Krawetz (Eds.). Totowa: Humana Press, 1999. V. 132 P. 365–386.
- Sebastiani F., Pinzauti F., Kujala S. T., Gonzalez-Martinez S. C., Vendramin G. G. Novel polymorphic nuclear microsatellite markers for *Pinus sylvestris* L. // *Conserv. Genet. Res.* 2012. V. 4. N. 2. P. 231–234.
- Tóth E. G., Vendramin G. G., Bagnoli F., Cseke K., Höhn M. High genetic diversity and distinct origin of recently fragmented Scots pine (*Pinus sylvestris* L.) populations along the Carpathians and the Pannonian Basin // *Tree Genet. Genom.* 2017. V. 13. Article number 47.
- Ueno S., Wen Y., Tsumura Y. Development of EST-SSR markers for *Taxus cuspidata* from publicly available transcriptome sequences // *Biochem. Syst. Ecol.* 2015. V. 63. P. 20–26.
- Untergasser A., Cutcutache I., Koressaar T., Ye J., Faircloth B. C., Remm M., Rozen S. G. Primer3 – new capabilities and interfaces // *Nucleic Acids Res.* 2012. V. 40. Article number 115.
- Van Oosterhout C., Hutchinson W. F., Wills D. P. M., Shipley P. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data // *Molecul. Ecol. Not.* 2004. V. 4. N. 3. P. 535–538.
- Van Oosterhout C., Weetman D., Hutchinson W. F. Estimation and adjustment of microsatellite null alleles in nonequilibrium populations // *Molecul. Ecol. Not.* 2006. V. 6. N. 1. P. 255–256.
- Wright S. The interpretation of population structure by F-statistics with special regard to systems of mating // *Evolution*. 1965. V. 19. N. 3. P. 395–420.
- Żukowska W. B., Wójcikiewicz B., Lewandowski A., Łászló R., Wachowiak W. Genetic variation of Scots pine (*Pinus sylvestris* L.) in Eurasia: impact of postglacial recolonization and human-mediated gene transfer // *Ann. For. Sci.* 2023. V. 80. Iss. 1. Article number 42.

## DEVELOPMENT OF MICROSATELLITE MARKERS WITH LONG MOTIFS FOR SCOTS PINE BASED ON WHOLE-GENOME *de novo* SEQUENCING

N. V. Oreshkova<sup>1, 2, 3</sup>, E. I. Bondar<sup>1, 3</sup>, V. V. Sharov<sup>1, 3</sup>, K. V. Krutovsky<sup>3, 4, 5, 6</sup>

<sup>1</sup> Federal Research Center Krasnoyarsk Scientific Center, Russian Academy of Sciences, Siberian Branch Akademgorodok, 50, Krasnoyarsk, 660036 Russian Federation

<sup>2</sup> V. N. Sukachev Institute of Forest, Russian Academy of Sciences, Siberian Branch Federal Research Center Krasnoyarsk Scientific Center, Russian Academy of Sciences, Siberian Branch Akademgorodok, 50/28, Krasnoyarsk, 660036 Russian Federation

<sup>3</sup> Siberian Federal University Prospekt Svobodny, 79, Krasnoyarsk, 660041 Russian Federation

<sup>4</sup> N.I. Vavilov Institute of General Genetics, Russian Academy of Sciences Gubkin str., 3, Moscow, 119991 Russian Federation

<sup>5</sup> G.F. Morozov Voronezh State Forest Engineering University Timiryazev str., 8, Voronezh, 394087 Russian Federation

<sup>6</sup> Georg-August University of Göttingen Gottingen, 37077 Germany

E-mail: oreshkova@ksc.krasn.ru, bone-post@yandex.ru, vsharov@sfu-kras.ru, konstantin.krutovsky@forst.uni-goettingen.de

Modern methods of whole-genome sequencing have made it possible to develop a large number of DNA markers, both selectively neutral and functional. Microsatellite loci are the most informative, reproducible, relatively inexpensive and highly polymorphic among other genetic markers. Whole-genome sequencing significantly simplifies their search and development. The article is devoted to the development of new microsatellite markers for Scots pine (*Pinus sylvestris* L.). Several thousand contigs containing microsatellite loci with tri-, tetra- and pentanucleotide motifs were selected in the draft genome assembly of Scots pine obtained by the authors. The search was specifically focused on loci with a repetitive motif length longer than two nucleotides, as the most reliable for genotyping even in a simple gel electrophoresis. A total of 39 primer pairs were tested. Of these, six loci with tri-, tetra- and pentanucleotide repeats were ultimately selected, which showed a high level of polymorphism, reliable genotyping, and were additionally tested in two populations from the Severo-Yeniseisky and Kuraginsky districts in Krasnoyarsk Region (East Siberia, Russia) and compared with other populations and marker sets according to published data. The developed markers can be used in the future in various population genetic studies and for identifying the origin of wood and plant material.

**Keywords:** *Pinus sylvestris* L., genetic diversity, whole genome sequencing, heterozygosity, conifers, population genetic analysis, microsatellite markers, SSRs.

**How to cite:** Oreshkova N. V., Bondar E. I., Sharov V. V., Krutovsky K. V. Development of microsatellite markers with long motifs for Scots pine based on whole-genome *de novo* sequencing // *Sibirskij Lesnoj Zhurnal* (Sib. J. For. Sci.). 2025. N. 1. P. 23–33 (in Russian with English abstract and references).