



Секвенирование мегагеномов хвойных. Геномные и эпигеномные механизмы адаптации лесных древесных видов



Константин Валерьевич Крутовский



*Внс лаборатории популяционной генетики
Института общей генетики им. Н. И. Вавилова РАН*



*Научный руководитель Научно-образовательного центра
геномных исследований и лаборатории лесной геномики
Сибирского федерального университета*



*Профессор генетики отделения лесной генетики и селекции
Гёттингенского университета, Германия*



*Адъюнкт профессор отделения по изучению и управлению
экосистем Техасского агро-механический университета, США*



**Проект: «Геномные исследования
основных бореальных лесообразующих
хвойных видов и их наиболее опасных
патогенов в Российской Федерации»
грант Правительства РФ №14.У26.31.0004**

**Крутовский К.В.^{1,2,3}, Орешкова Н.В.^{2,4}, Путинцева Ю.В.²,
Павлов И.Н.^{2,4}, Кузьмин Д.А.², Шилкина Е.А.^{2,5},
Ибе А.А.^{2,5}, Дейч К.О.^{2,5}, Третьякова И.Н.⁴, Ваганов Е.А.²**



¹*Институт общей генетики им. Н.И. Вавилова РАН, Москва*



²*Сибирский Федеральный Университет, Красноярск*

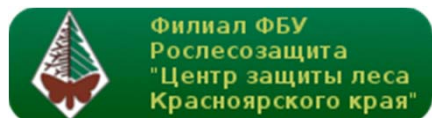


³*Гёттингенский университет, Германия*



⁴*Институт леса им. В.Н. Сукачева СО РАН, Красноярск*

⁵*Филиал ФБУ «Рослесозащита» «Центр защиты леса Красноярского*



края», Красноярск

Основные задачи проекта

1. **Создание новой, конкурентоспособной, современной, динамичной и хорошо оснащённой лаборатории геномных исследований**, интегрированной как в научную, так и в образовательную инфраструктуру СФУ, способной проводить современные научные исследования и готовить специалистов в области геномики и биоинформатики.
2. **Получение эталонных, полностью просеквенированных и аннотированных геномов основных лесообразующих пород Российской Федерации** лиственницы сибирской (*Larix sibirica* Ledeb.), сосны обыкновенной (*Pinus sylvestris* L.) и сосны кедровой сибирской (*Pinus sibirica* Du Tour.) и их наиболее опасных патогенов, в первую очередь грибов, входящих в комплексы *Armillaria mellea* s. l. и *Heterobasidion annosum* s. l., которые вызывают катастрофическое усыхание российских бореальных лесов.
3. **Разработка высокоинформативные молекулярно-генетические маркёры**, которые могут быть эффективно использованы для определения происхождения древесины, изучения и мониторинга генетической изменчивости хвойных лесов, их адаптации к изменению климата и для создания селекционных и природоохранных программ. Эти же задачи были заявлены как наиболее приоритетные для лесного хозяйства в Комплексной программе развития биотехнологий в Российской Федерации на период до 2020 года, утверждённой Правительством Российской Федерации 24 апреля 2012 г.

Создана новая лаборатория – единственная такого рода в стране!



СГААГАГАСГТТТАГССССГТТГАА
ТСТТАГССССГТТГАААААААААААА
АААААААААААААААААААААААААА
АААААААААААААААААААААААААА

ТТГАААТТТТТГАААААААААААААА
АТТААААААААААААААААААААААА
ТТТАААААААААААААААААААААА
АААААААААААААААААААААААААА

Центр геномных исследований
Научно-образовательный центр (ФУ)



Новости

Главная

Лаборатория лесной геномики

Руководитель

Сотрудники

Оборудование

Учебная лаборатория

Нам помогают

Партнеры

Контакты

СМИ о нас

Лаборатория лесной геномики

Лаборатория лесной геномики создана в Сибирском федеральном университете в апреле 2014 года в рамках выполнения проекта «Геномные исследования основных бореальных лесообразующих хвойных видов и их наиболее опасных патогенов в Российской Федерации» Правительством РФ в рамках постановления № 220 от 9 апреля 2010 г. «О мерах по привлечению ведущих ученых в российские образовательные учреждения высшего профессионального образования»

[Приказ о создании лаборатории](#)

[Положение о лаборатории](#)

<http://genome.sfu-kras.ru>

Руководитель лаборатории



Крутовский Константин Валерьевич

Координатор проекта



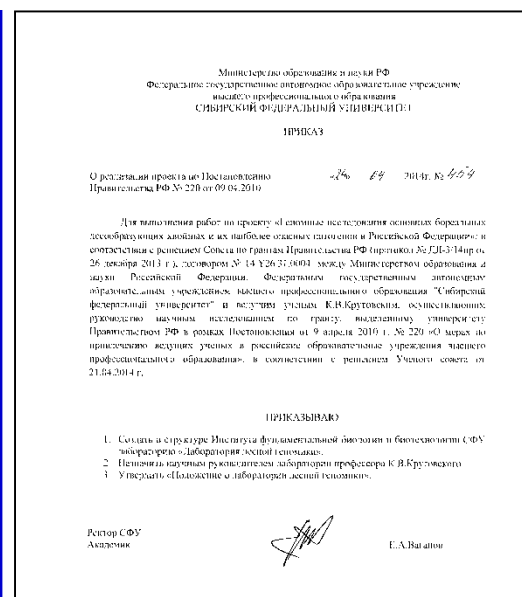
Шулепина Светлана Петровна

Сотрудники лаборатории



© 2012 НОЦ «Центр геномных исследований» СФУ

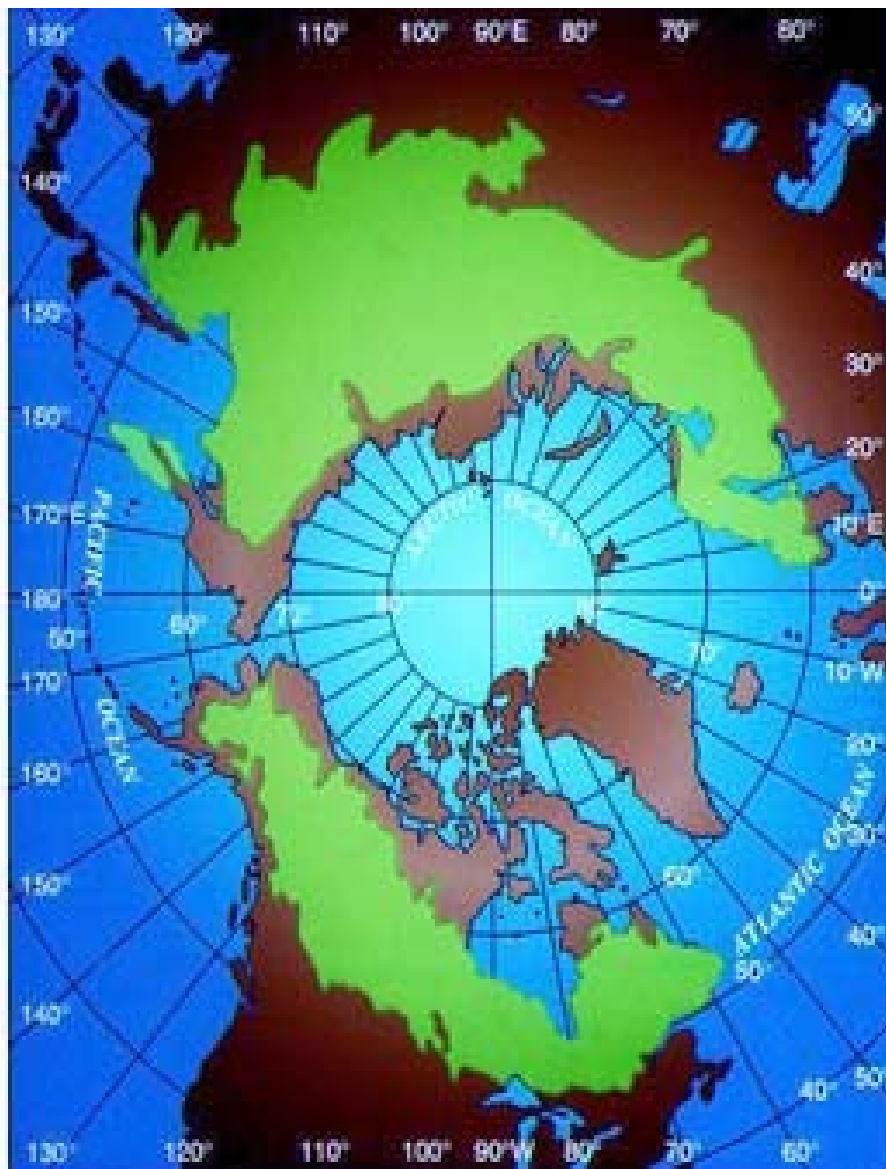
По вопросам редактирования сайта обращаться yuliya-putintseva@rambler.ru



Более 30 человек, включая ВТК, в основном молодые специалисты и студенты участвуют в проекте!

Цикл лекций «Современные методы фитопатологии, геномики и биоинформатики в изучении и решении проблем сибирских лесов», 3-4 июня 2015 г., СФУ, Красноярск

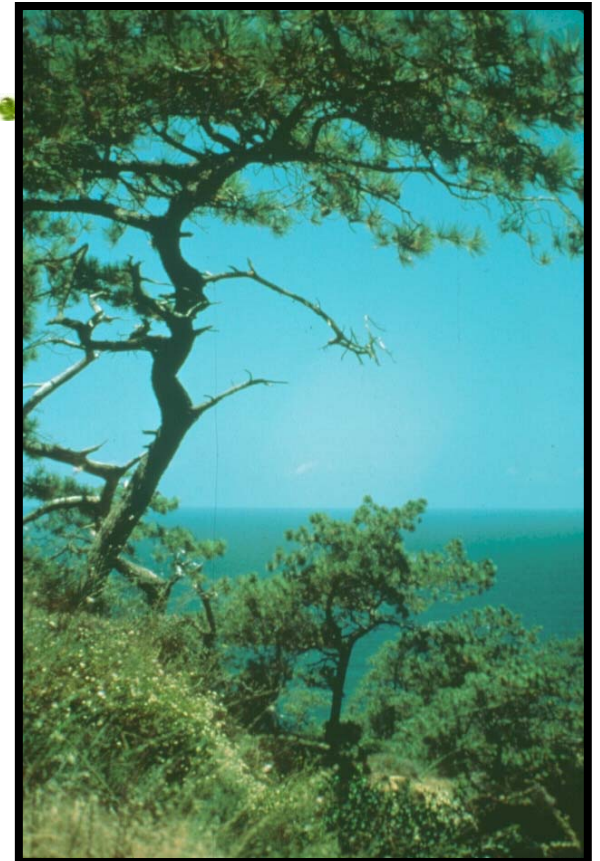
Бореальные леса - самые большие биогеоценозы, которые требуют защиты и сохранения



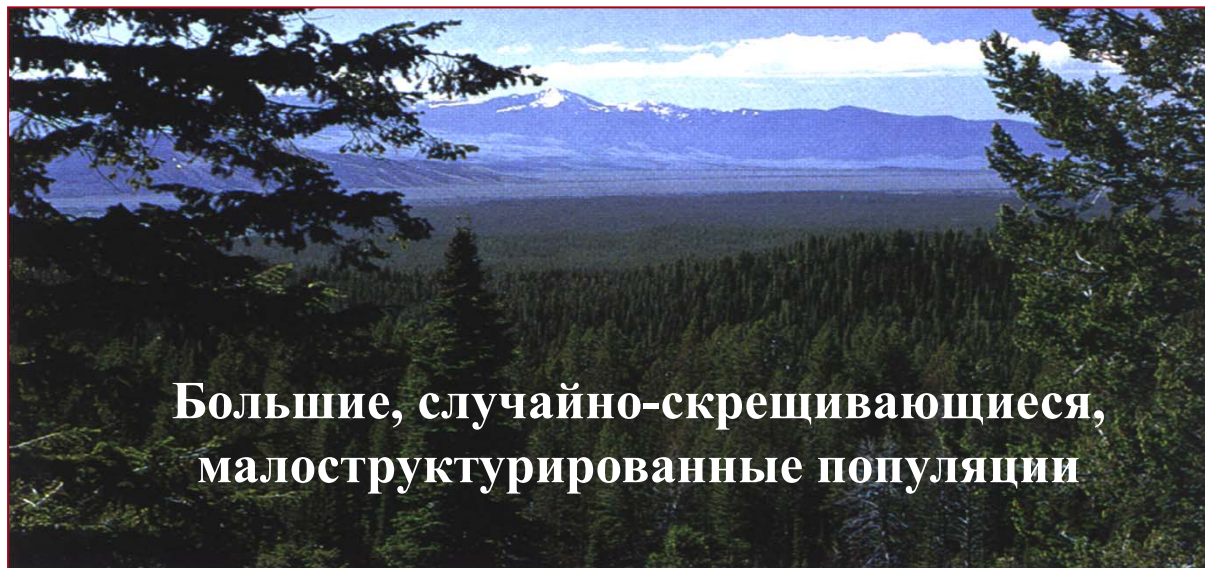
- ~35% всех лесов мира
- главным образом хвойные

Хвойные

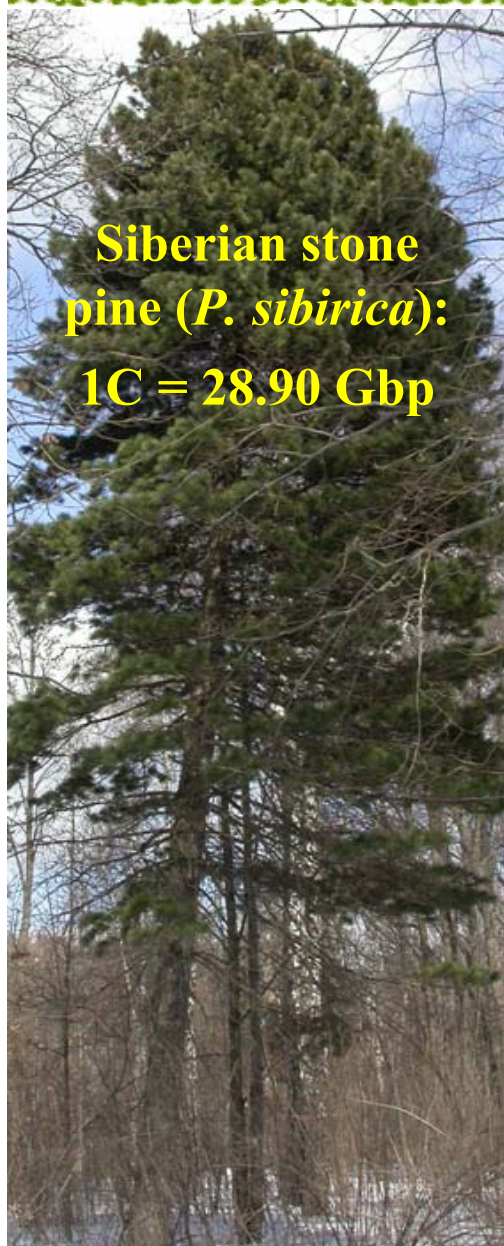
- экстремалы: самые долгоживущие, самые крупные, и т.д.
- крайне разнообразные
- самые эволюционно древние
- самые экономически важные
- эстетичные
- самое доступное и реальное средство для борьбы с парниковым эффектом (через связывание углерода)
- биогорючее (единственный реальный возобновляемый альтернативный источник энергии)



Самые эволюционно древние среди растений



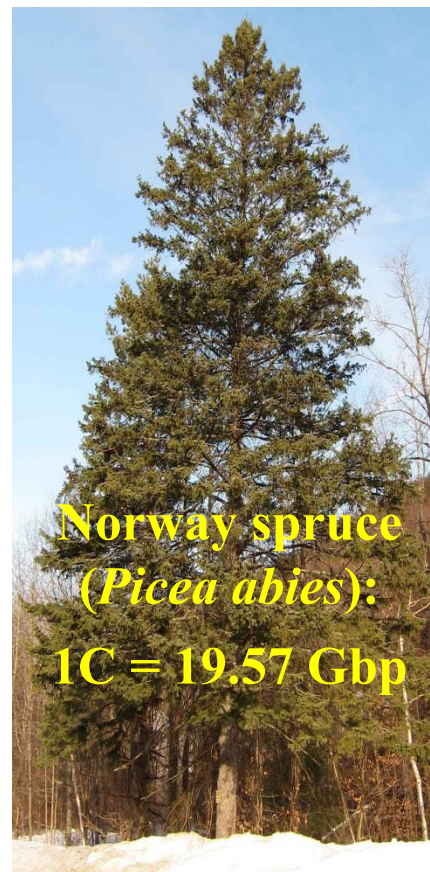
Космический масштаб задач: Гигантский размер геномов хвойных – в 4-9 раз больше генома человека!



Siberian stone
pine (*P. sibirica*):
1C = 28.90 Gbp



Scots pine (*Pinus
sylvestris*):
1C = 22.47 Gbp



Norway spruce
(*Picea abies*):
1C = 19.57 Gbp



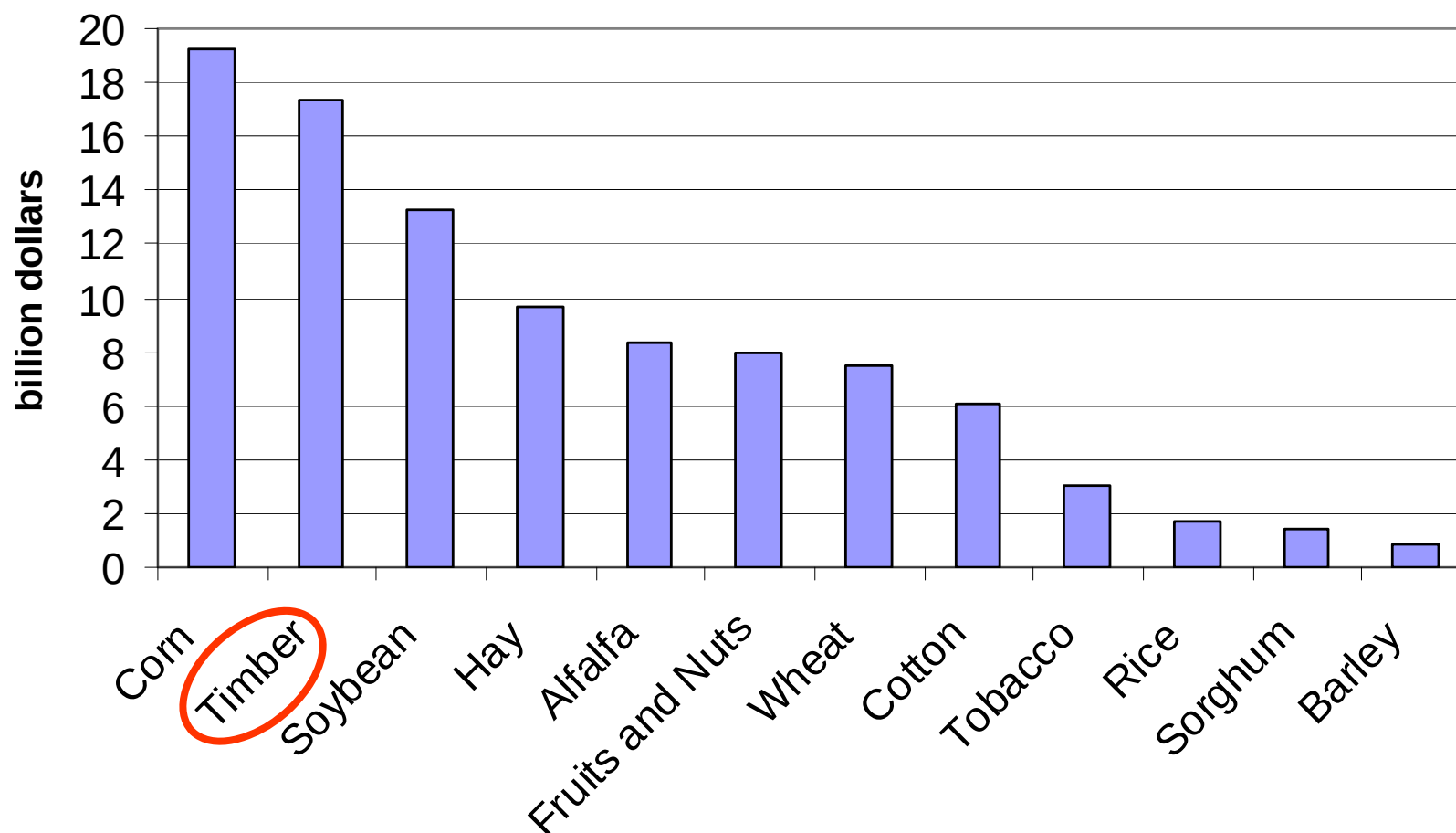
Siberian larch
(*Larix sibirica*):
1C = 12.03 Gbp

Human
genome,
*Homo
sapiens*: 1C
= 3.20 Gbp



Самые экономически и хозяйственно важные среди с.-х. растений

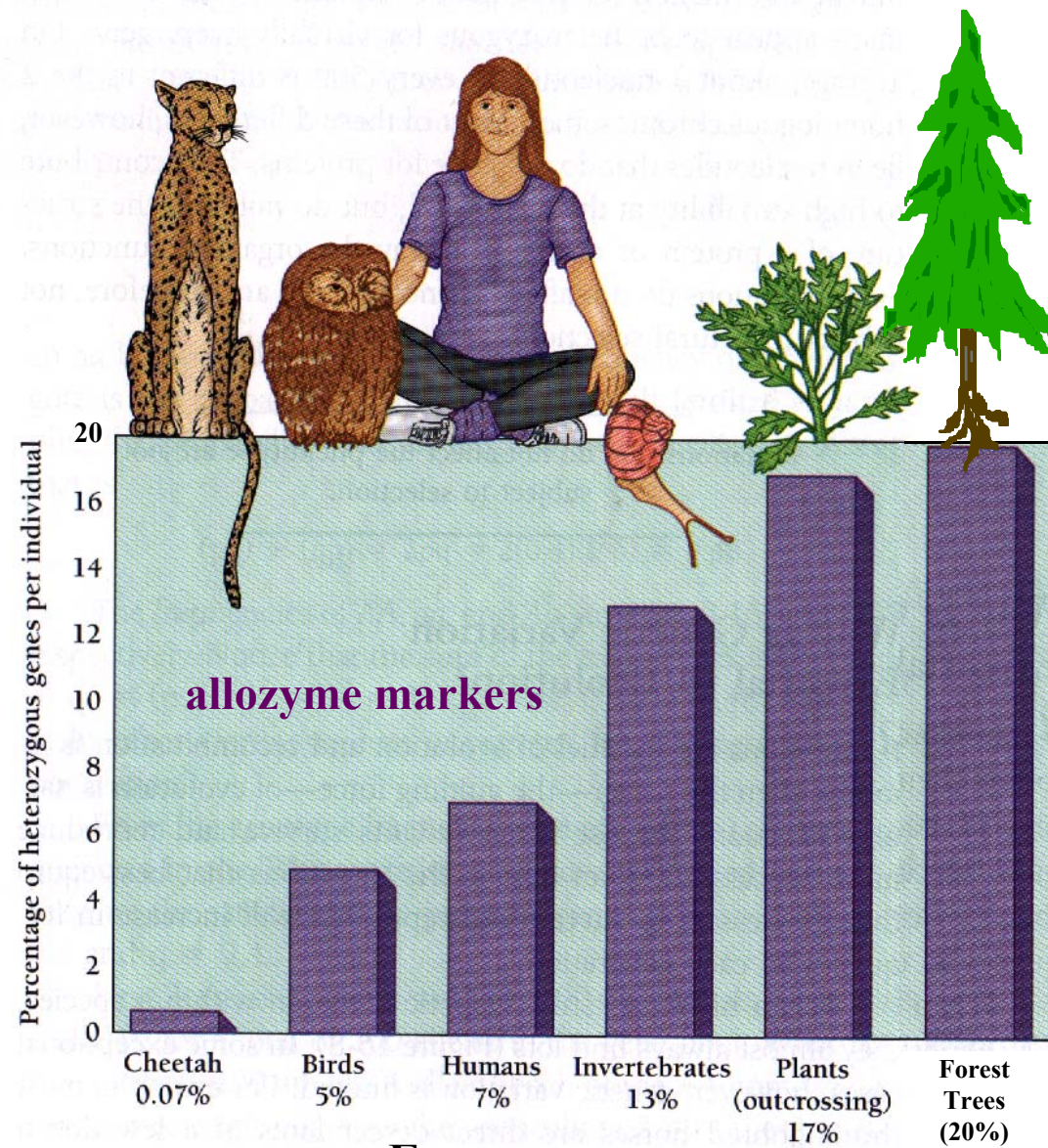
**Стоимость основных сельско-хозяйственных культур и
древесины в США (USDA Economics Research Service 2005)**



Источник уникальных продуктов для:

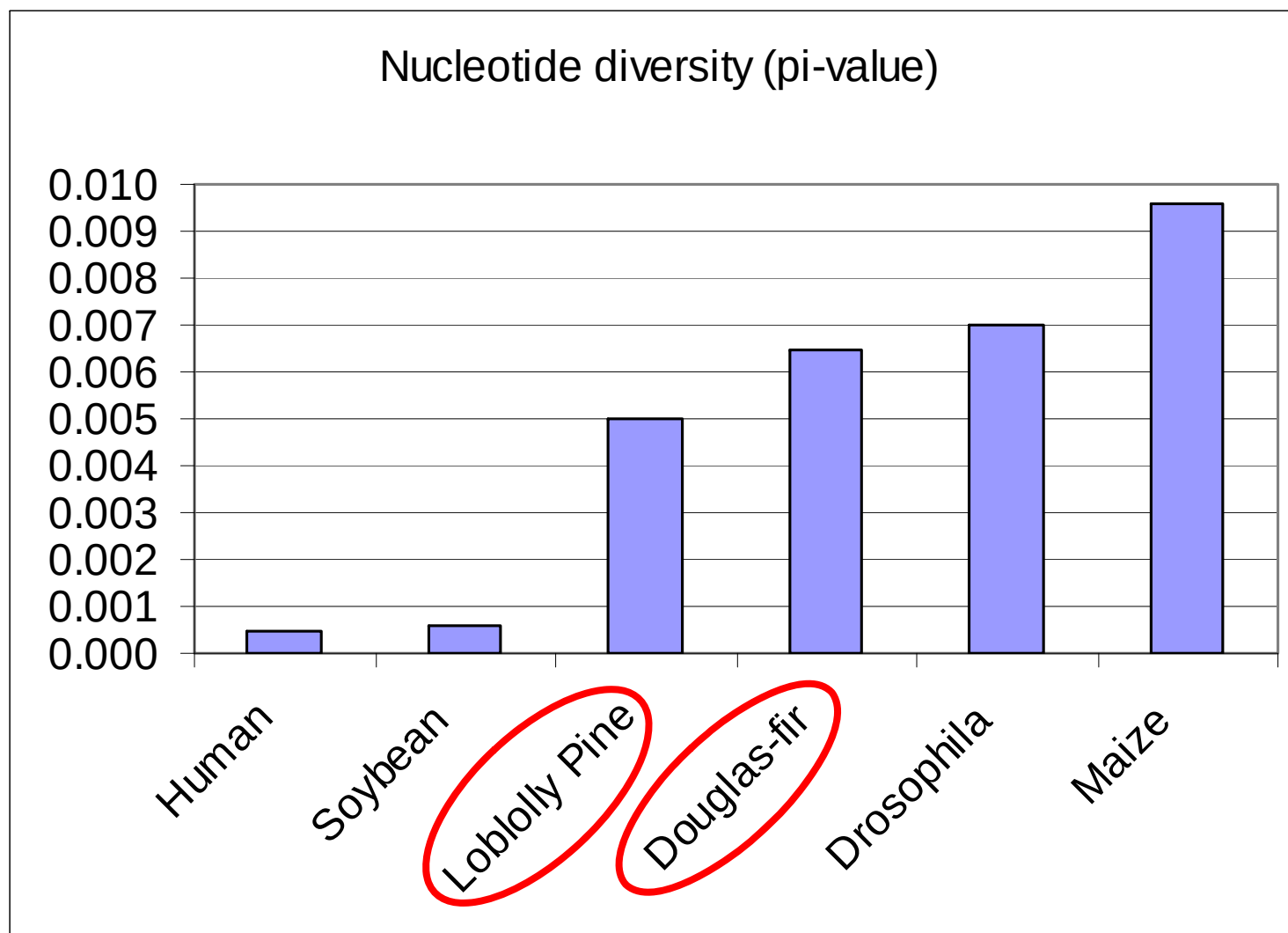
- строительной, бумажной и мебельной пром-сти
- получения альтернативной энергии и биотоплива (спирт на основе целлюлозы)
- медицины (например, противоопухолевые препараты на основе таксола)
- авиационной и ракетной техники (высококачественное реактивное топливо на основе альфа-пиренов)
- борьбы с вредителями (экологически безопасные пестициды и репелленты на основе терпенов и др. вторичных метаболитов)

Самые генетически изменчивые



Гетерозиготность

Уровень нуклеотидной изменчивости

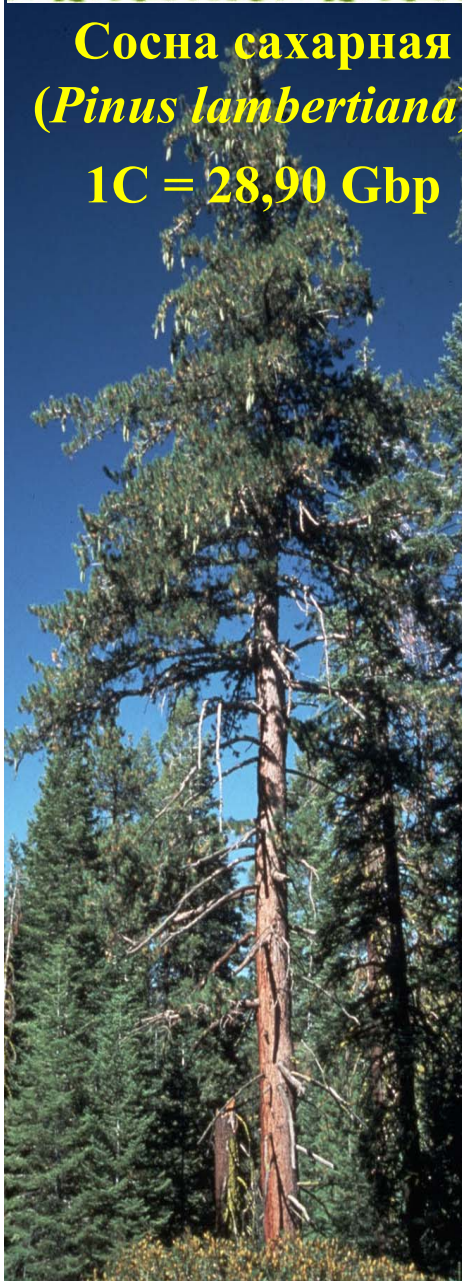


Krutovsky K.V. & D.B. Neale. 2005. Nucleotide diversity and linkage disequilibrium in cold-hardiness and wood quality-related candidate genes in Douglas-fir. *Genetics* 171(4): 2029-2041.

Самый большой размер генома

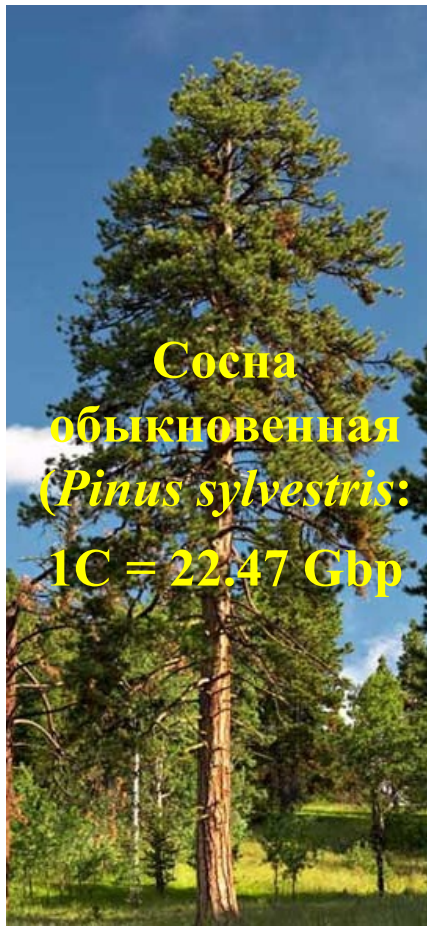
Сосна сахарная
(*Pinus lambertiana*):

1C = 28,90 Gbp



Сосна обыкновенная
(*Pinus sylvestris*):

1C = 22.47 Gbp



Сосна ладанная
(*Pinus taeda*):

1C = 21,61 Gbp



Ель обыкновенная
(*Picea abies*):

1C = 19,57 Gbp



Геном
человека,
Homo sapiens: 1C
= 3,20 Gbp



Почему и зачем нужны геномные исследований в лесном хозяйстве



Какая выгода лесной генетике и защите леса от расшифровки генома основных видов хвойных?

- идентификация и аннотация всех функциональных генов и регуляторных элементов (включая короткие РНК, транскрипционные факторы и т.д.) и определение метаболических сетей генов, контролирующих адаптацию и устойчивость к болезням
- разработка высоко информативные генетические маркеры (прежде всего микросателлитных локусов и однонуклеотидных полиморфизмов - SNP – т.н. «снипов»), которые могут быть использованы в генетических исследований популяций и для создания генетической базы данных (наподобие молекулярно-генетических штрих-кодов для отдельных популяций) для борьбы с нелегальной заготовкой и торговлей древесины
- разработка полногеномных генетических маркеров для обнаружения связи между генетической изменчивостью (SNP, аллели, гаплотипы и генотипы) с изменчивостью адаптивных и селекционно-ценных признаков и фенотипов, и с факторами окружающей среды для лучшего понимания генетического контроля адаптивных, селекционных и экономически важных признаков
- разработка полногеномных генетических маркеров для геномной селекции быстрорастущих и более устойчивых пород с ценными признаками
- системная биология: интеграция протеомики, транскриптомики и метаболомики
- референсный геном для картирования при повторном секвенировании (ресеквенировании)

Проекты полного *de novo* секвенирования геномов хвойных

Вид	Ведущие учреждение (руководитель проекта, бюджет, год начала)
<u>Ель обыкновенная</u> Norway spruce (<i>Picea abies</i>)	Umeå Plant Science Centre, Sweden (Dr. Pär Ingvarsson, \$12M, 2010)
<u>Ель белая, ситка и чёрная</u> White, Sitka and Black spruce (<i>Picea glauca</i> , <i>P. sitchensis</i> & <i>P. mariana</i>)	Université Laval, Canada (Dr. John MacKay, \$10M, 2010)
<u>Ладанная сосна</u> Loblolly pine (<i>Pinus taeda</i>), <u>Лжетсуга</u> Douglas-fir (<i>Pseudotsuga menziesii</i>), <u>Сахарная сосна</u> Sugar pine (<i>Pinus lambertiana</i>)	University of California, Davis, USA (Dr. David Neale, \$15M, 2011)
<u>Средиземноморская сосна</u> Maritime pine (<i>Pinus pinaster</i>), <u>сосна обыкновенная</u> Scots pine (<i>Pinus sylvestris</i>)	European Union (Dr. María-Teresa Cervera, INIA CIFOR, Spain, \$10M, 2012)
<u>Сибирская лиственница</u> Siberian larch (<i>Larix sibirica</i>), <u>сибирская кедровая сосна</u> Siberian pine (<i>Pinus sibirica</i>), <u>сосна обыкновенная</u>	Siberian Federal University, Russia (Dr. Konstantin Krutovsky, \$2M, 2012, \$3M, 2014)

Колоссальная сложность задачи – гигантский размер геномов хвойных!

Сравнительный размер геномов хвойных в проектах полногеномного секвенирования

Вид ¹	Кол-во ДНК, pg (1С)	Размер ДНК, Gbp (1С)	Соотношение к геному человека
Геном человека, <i>Homo sapiens</i>²	3,47	3,20	1
Лиственница сибирская (<i>Larix sibirica</i>)³	12,30	12,03	4
Лиственница Гмелина (<i>Larix gmelinii</i>) ³	14,00	13,69	4
Псевдотсуга (<i>Pseudotsuga menziesii</i>) ⁴	19,05	18,63	6
Ель обыкновенная (<i>Picea abies</i>)⁵	20.01	19.57	6
Ель белая (<i>Picea glauca</i>)³	20.20	19,76	6
Сосна ладанная (<i>Pinus taeda</i>)⁴	22.10	21.61	7
Сосна обыкновенная (<i>Pinus sylvestris</i>)⁵	22,98	22,47	7
Сосна приморская (<i>Pinus pinaster</i>) ³	24.35	23.81	7
Сосна сибирская (<i>Pinus sibirica</i>)⁶	29,55	28,90	9
Сосна сахарная (<i>Pinus lambertiana</i>) ⁷	29,55	28,90	9

¹Все виды семейства Сосновые (Pinaceae), включая данные виды имеют 12 пар хромосом, за исключением псевдотсуги, у которой их 13 (<http://data.kew.org/cvalues>) ; ²IHGSC 2004; ³Ohri & Khoshoo 1986; ⁴O'Brien *et al.* 1996; ⁵Fuchs *et al.* 2008; ⁶Zonneveld BJM. 2012; ⁷Wakamiya *et al.* 1993.

Современные методы секвенирования ДНК

- Выпуск с 2005 г. коммерческих высокопроизводительных секвенаторов ДНК на основе новых технологий массивного параллельного секвенирования компаниями Roche и Illumina
- Создание в 2012 г. Научно-образовательного центра геномных исследований и в 2014 г. Лаборатории лесной геномики Сибирского федерального университета при поддержке Отдела генетики и селекции Центра защиты леса Красноярского края и Лаборатории лесной генетики и селекции Института леса им. В. Н. Сукачева СОРАН, оборудованной самым производительным секвенатором Illumina HiSeq 2000 и суперкомпьютером (<http://genome.sfu-kras.ru>)

СФУ | Поступление | Наука | Обучение | Мой университет | Спорт | Сайты СФУ

Поиск

Центр геномных исследований
Научно-образовательный центр (СФУ)

Главная » Лаборатория лесной геномики

Оборудование

Основное оборудование для секвенирования

- Секвенатор
- Кластерная станция
- Вычислительный кластер (сервер)
- Специализированный комплекс IBM для высокопроизводительных вычислений

Вспомогательное оборудование

- Амплификатор
- Амплификатор Real Time

Специализированный комплекс IBM для высокопроизводительных вычислений

Комплекс обеспечит эффективную обработку и хранение данных, производимых Центром геномных исследований СФУ и станет основой сервиса для решения широкого круга задач в области геномных исследований и клинической геномики. Сервис предполагает включение в него специализированного биохимического оборудования, высокопроизводительного оборудования с уникальными характеристиками, специализированного программного ПО и ПО с открытым кодом.

В состав комплекса входят:

- 96 ядерный SMP сервер IBM x3950 X6 с объемом ОЗУ 3 ТБ;
- Гибридный сетевой сервер IBM 6130 M6 с двумя GPU NVIDIA Tesla K20 обеспечивающий суммарную линейную производительность для вычислений одинарной точности с плавающей точкой до 7 TFlops;
- Подсистема хранения данных IBM Storwize V3700 объемом 48 ТБ.

Комплекс работает под управлением ОС SuSE Linux Enterprise Server 11, установлена параллельная файловая система IBM GPFS, система мониторинга Ganglia, система пакетной обработки Torque.



Генетическая и языковая передача информации



Нуклеотиды (А, Т, С и G) **Буквы алфавита**



Ген

**Строчка в тексте / Предложение
(состоящее из 3-х буквенных
«слов»-триплетов или кодонов у
протеин кодирующих генов)**



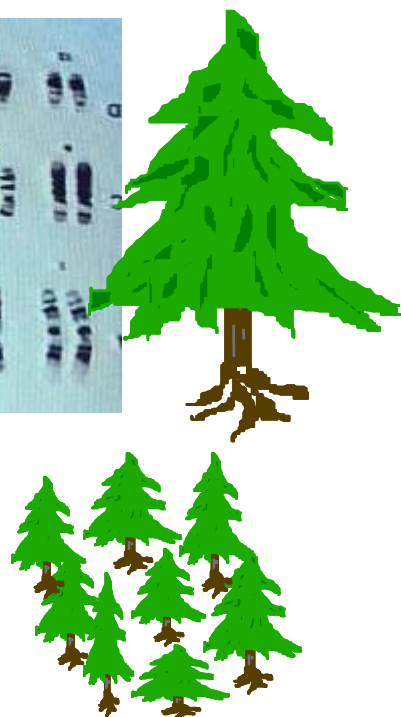
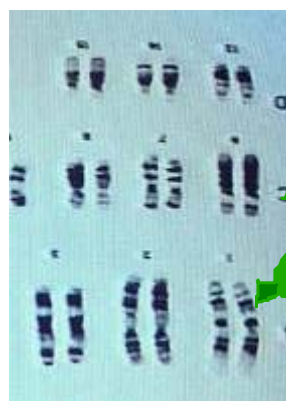
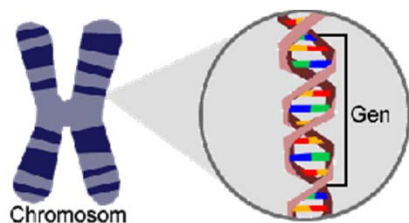
Хромосома



Геном



Генофонд

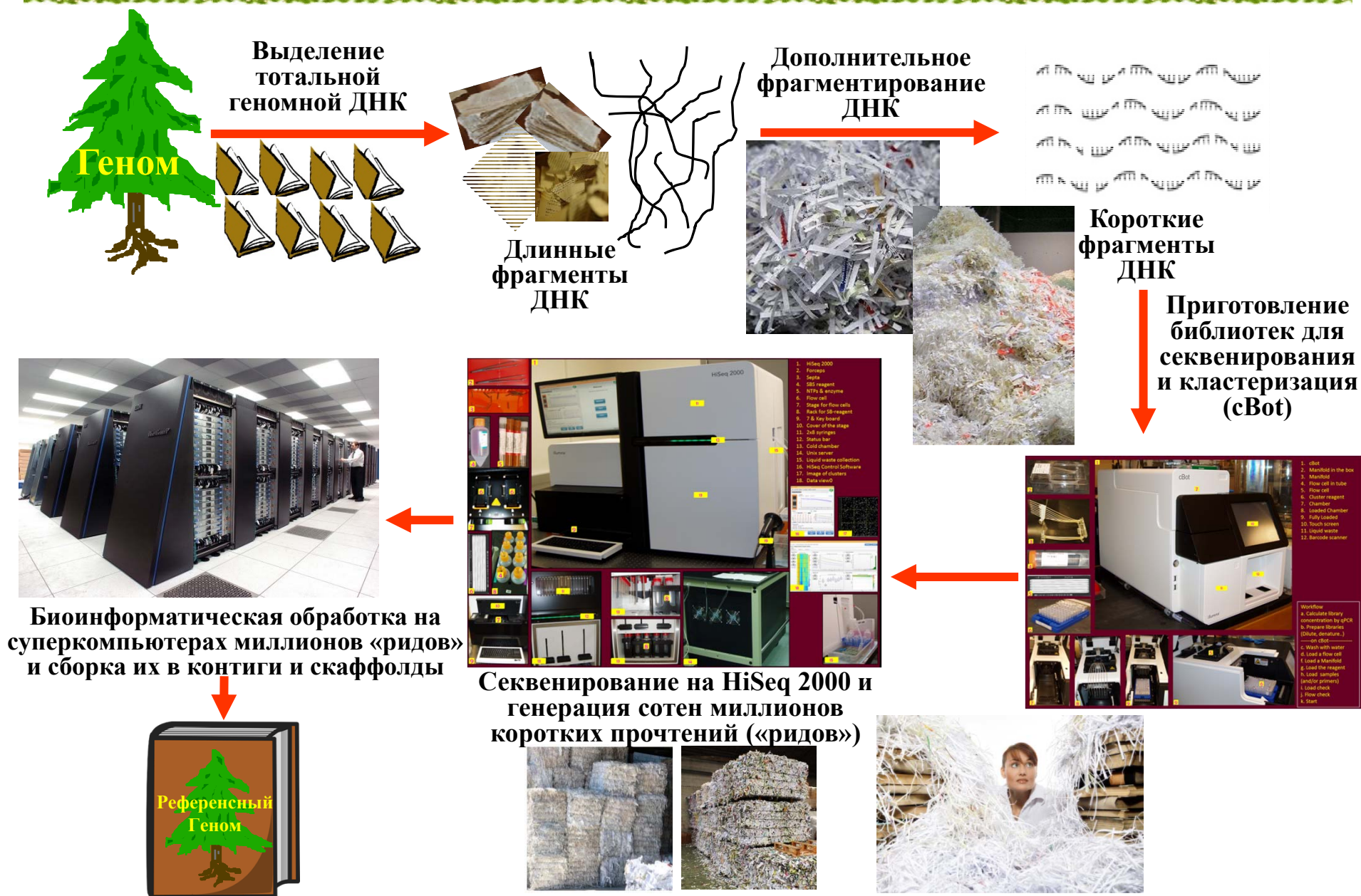


Глава



Основная задача геномики - полное секвенирование и расшифровка генома

Проблема полногеномного секвенирования с помощью новых секвенаторов: возможность секвенировать только короткими фрагментами 100-150 нм



Полученные данные полногеномного секвенирования

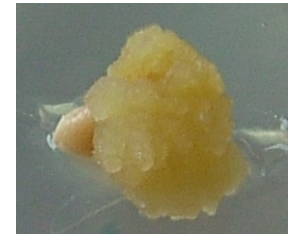
- *Larix sibirica* (12,03 Gbp):

хвоя и гаплоидная клеточная линия

парно-концевые (PE) библиотеки: 180, 250, 400 и 500 bp

сцепленно-концевая (MPE) библиотека: 5 Kbp

общая длина ридов хорошего качества ~891 Gbp (~74X)



- *Pinus sibirica* (28,90 Gbp):

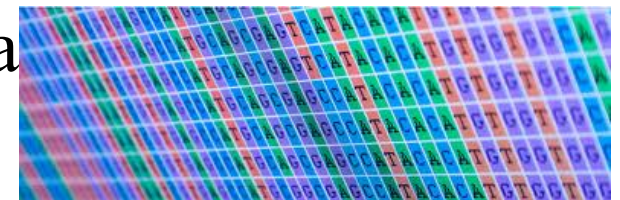
отдельный эндосперм

PE библиотеки: 250 и 500 bp

MPE библиотеки: 3 и 5 Kbp

общая длина ридов хорошего качества

~595 Gbp (~21X)



Сравнительные данные полногеномного секвенирования

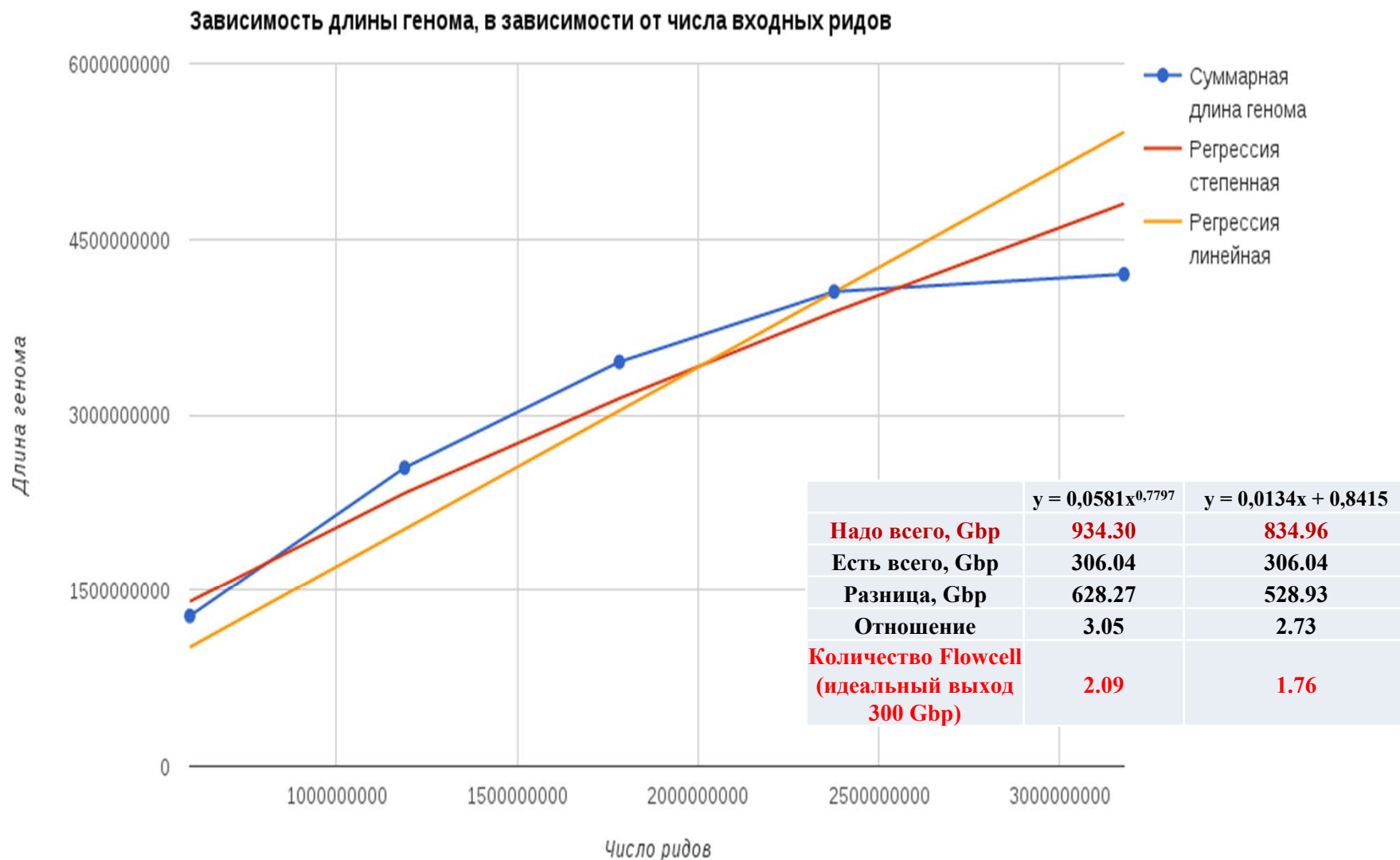
Metrics	<i>Larix sibirica</i>	<i>Pinus sibirica</i>	<i>Pinus lambertiana</i>	<i>Pinus taeda</i> ¹	<i>Picea abies</i> ²	<i>Populus trichocarpa</i> ³	<i>Eucalyptus grandis</i> ⁴	<i>Quercus robur</i> ⁵	<i>Arabidopsis thaliana</i> ⁶	<i>Vitis vinifera</i> ⁶	<i>Amborella trichopoda</i> ⁷
Chromosomes (1n)	12	12	12	12	12	19	11	12	5	19	13
Estimated genome size, Gbp	12.029	28.900	28.900	21.614	19.570	0.484	0.611	0.905	0.156	0.416	0.868
Assembled genome size (sequence in contigs), Gbp	4.949	4.957	25.167	20.148 ⁸	12.019*	0.423	0.640	1.34*	0.135*	0.487*	0.706*
%	41	17	87	93	61	87	105	148	86	117	81
Total span of scaffolds, Gbp	N/A	N/A	34.076	23.180	12.019	0.410	0.691	1.34	0.135	0.487	0.706
Contig N50, bp	1,754	429	8,112	8,206	N/A	129,000	2,261	N/A	N/A	N/A	N/A
Scaffold N50, bp	N/A	N/A	209,180*	66,920	4,869	3,100,000	5,000	260,000	N/A	N/A	4,900,000
G+C content, %	36.6	36.7	37.2	38.2	37.9	33.3	N/A	N/A	35.0	36.2	35.5
TE content, %	N/A	N/A	N/A	79	70	42	50	N/A	15.3	41.4	N/A
Number of genes ⁸	N/A	N/A	N/A	50,172	58,587 ⁹	36,393	46,315	N/A	27,160	25,663	25,347
Average CDS length, bps	N/A	N/A	N/A	965	723	1,143	1,568	N/A	1,102	1,095	969
Average intron length, bps	N/A	N/A	N/A	2,741	1,020	366	1,936	N/A	182	933	1,538
Maximum intron length, bps	N/A	N/A	N/A	318,524	68,269	4,698	N/A	N/A	10,234	38,166	175,748

*Scaffold statistics

Зависимость размера собираемого генома от числа входных ридов

Уравнение степенной регрессии : $Y = 505,93782 * \text{pow}(X; 0,7343)$

Уравнение линейной регрессии : $Y = 1,7036 * X + 16,6397$



Микросателлитные локусы лиственницы сибирской (*Larix sibirica*)

Motif, bp	Loci selected	Mean locus length, bp	Per 1 Gbp
2	563	20.6	28,570
3	140	31.8	7,100
4	7	53.9	360
5	3	60.0	150
6	10	59.9	510

Genome-wide Microsatellite Analyzing Tool (*GMATo*)

```

1121   GGGAGTTCCA CCGATCCTAA GATCATAGTG GTTTTCTTTA GAGACTTCAC GCCCCTACC ACGGATGCCT
1191   TCGCCCACCC ATGCGCACCT ACATCAATAA CACAAACAGG AAAATGCAGA CCCCTTCCA TACCTGTGCG
1261   GCCATCATAA ATACATGCTC GCCCCTCCTT TCTTCCCTTA ACGCACACGA AAGAAGGAAT GGCCAAGTGG
1331   TGGAGCGTTA CTCTTCCAAT ATCCCGACAT AGTGCCTCCT ATAATACTAT ATATATAACG TCTTCCGCTT
1401   AATTGCATTA TTAATTATAT CACAGCATTA TAACTTATAT AACTAECTTA GATTTAACCA TCCAAAGTAA
1471   TAATAATAAT AATAATAATA TTATTATTAT TATTATTATT ACTTTGGATG GTTAAATCTA AGTTAGTTAT
1541   ATAAGTTATA ATGCTGTGAT ATAATTAATA ATGCAATTAA GCGGAAGACG TTATATATAT AGTATTATAG
1611   GAGGCACTAT GTCGGGATAT TGGAAGAGTA ACGCTCCACC ACTTGCCCAT TCCTTCTTTC GTGTGCGTTA
1681   AGGGAAGAAA GGAGGGGCGA GCATGTATTT ATGATGGCCG CACAGGTATG GAAGGGGGTC TGCATTTTCC
1751   TGTTTGTGTT ATTGATGTAG GTGCGCATGG GTGGGCGAAG GCATCCGTGG TAGGGGGCGT GAAGTCTCTA
    
```

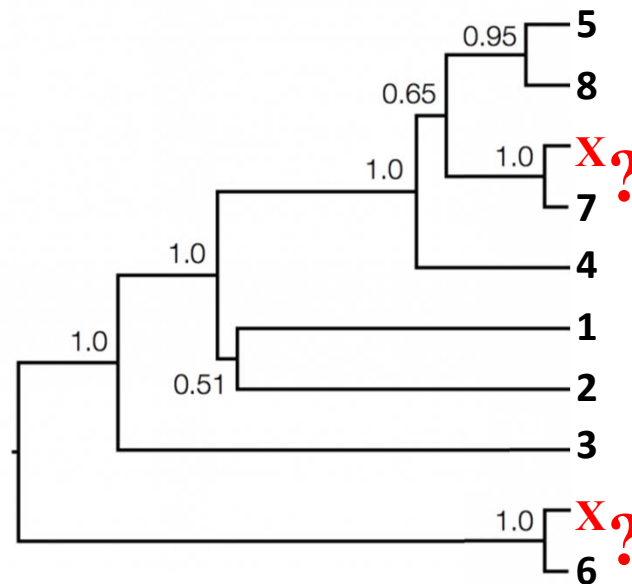
- для лиственницы синтезированы и протестированы ПЦР праймеры для 87 локусов;
- выявлено 2 полиморфных праймера информативных по гетерозиготности;
- в дальнейшем планируется расширить число праймеров и использовать сиквенсы трёх деревьев для более информативного поиска
- для сосны кедровой сибирской *Pinus sibirica* Du Tour протестированы ПЦР праймеры для 70 локусов и 20 оказались полиморфными

Использование молекулярно-генетических маркёров для борьбы с нелегальной заготовкой и торговлей древесиной

Условная лесная карта, на которой генетически различающиеся популяции выделены разным номерами



На дендрограмме эти популяции расположены на основе генетических различий между ними



↑
Построение
новой
дендрограммы

Выделение
ДНК и гено-
типирование

Сверка с компьютерной базой данных по частотам аллелей маркёров (микросателлитные локусы и «снипы») для всех основных популяций данной породы в данном регионе

Анализ повторяющихся последовательностей в геноме сосны сибирской кедровой

- Участки ДНК, включённые в геном, последовательность которых состоит из повторяющихся фрагментов.

~50% генома человека

~32% *Arabidopsis thaliana*

~70% кукурузы

~35% риса,

~42% *Populus trichocarpa*

~82% хвойных!

Повторяющиеся последовательности в геноме

Часто повторяющиеся последовательности

- **Сателлиты**
- **Микросателлиты**
- **Минисателлиты**

Умеренно повторяющиеся последовательности

- **Ретроэлементы**
обратная транскрипция
синтез ДНК на матрице РНК
обратной транскриптазой
- **ДНК-транспозоны** (*cut and paste*) с участием
транспозазы

Ретроэлементы:

- 1) LTR-содержащие ретроэлементы, включающие ретротранспозоны и эндогенные ретровирусы. Ty1/copia и Ty3/gypsy
- 2) ретроэлементы, не содержащие LTR, составленные из длинных (LINE) и коротких (SINE) диспергированных элементов.

Анализ повторяющихся последовательностей в геноме сосны сибирской кедровой

Данные полногеномного секвенирования сосны сибирской кедровой:

- контиги длиной > 500 bp
- общее число таких контигов = 2,389,750
- суммарная длина = 2 149 954 272 bp
- максимальная длина = 14,020 bp
- N50 = 919 bp
- N90 = 554 bp

27

Подходы к поиску повторяющихся последовательностей

- *Поиск повторяющихся последовательностей с использованием библиотек повторов.*
- *Сигнатурный подход (The signature-based approaches).*
- *De Novo подходы:*
 - *Self-comparison approaches*
 - *K-mer and spaced seed approaches*

RepeatModeler pipeline

- <http://www.repeatmasker.org/RepeatModeler.html>
- Включает в себя программы:
 - RECON
 - RepeatScout (с использованием k -меров
<http://repeatscout.bioprotects.org>)
 - TRF
 - RepeatMasker
 - Использует выходные данные программ RECON и RepeatScout для создания, уточнения и классификации консенсусных моделей предполагаемых повторов.

Классификация найденных повторов

Тип повтора	Количество
Unknown	1623
LTR/Gypsy	145
LTR/Copia	100
LINE/L1	65
DNA/hAT-Tag1	15
LINE/RTE-X	8
Simple repeat	8
DNA/CMC-EnSpm	8
RC/Helitron	7
LTR/Pao	6
LINE/Penelope	5
Satellite	5
DNA	4
LTR	3

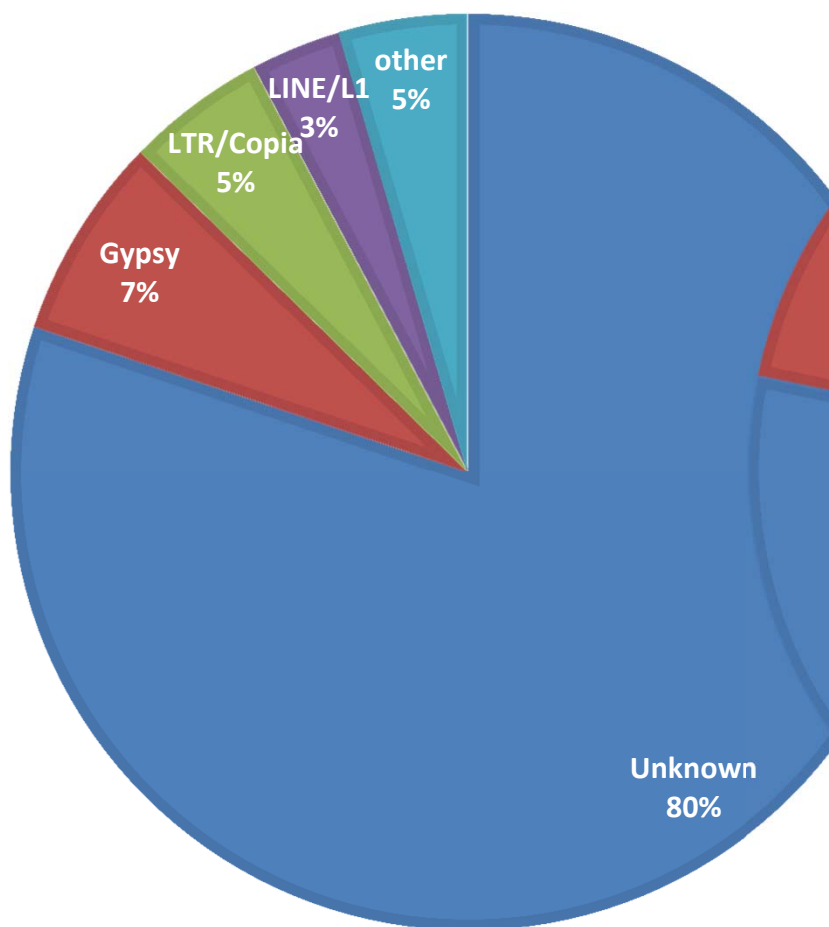
Тип повтора	Количество
DNA/MULE-MuDR	3
LINE/L2	3
buffer	2
DNA/hAT-Charlie	2
SINE	2
LINE/CR1	2
DNA/hAT-hATw	1
LINE/RTE-BovB	1
LINE/R1	1
DNA/hAT-Ac	1
LTR/ERV1	1
LINE/Tad1	1
LINE/I	1
DNA/Ginger	1
LINE/L1-Tx1	1

Повторяющиеся элементы, найденные в геноме лиственницы сибирской

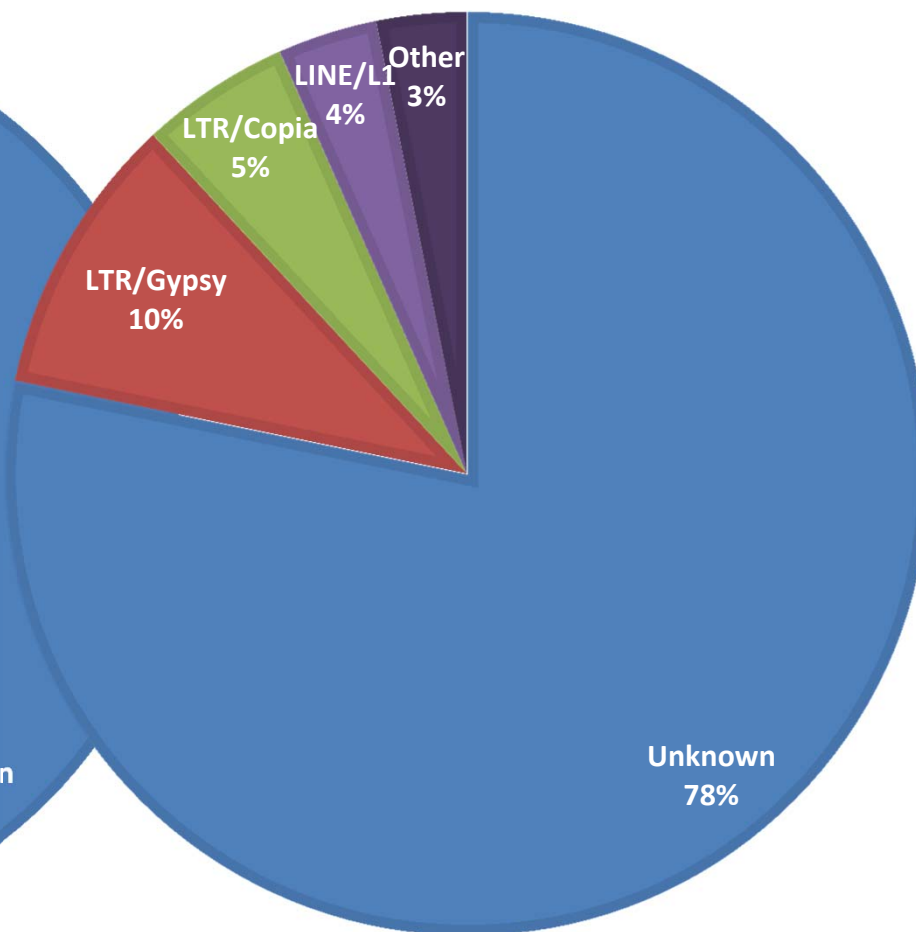
Тип повтора	Количество
Unknown	1815
LTR/Gypsy	229
LTR/Copia	122
LINE/L1	81
DNA/hAT-Tag1	18
DNA/CMC-EnSpm	12
RC/Helitron	12
LINE/RTE-X	7
LINE/Penelope	6
Simple_repeat	6
LINE/L1-Tx1	5
LTR/Caulimovirus	5
LTR	4
DNA	4
DNA/Sola	4
Satellite	3
SINE?	3
DNA/hAT	3
buffer	3
LTR/Pao	2
LINE/L2	2

Тип повтора	Количество
LINE/R1	2
LTR/Ngaro	2
LTR/ERVK	2
DNA/hAT-Tip100	2
DNA/hAT-Ac	2
DNA/MULE-MuDR	2
LTR/DIRS	2
LTR/ERV1	2
DNA/Dada	1
LTR/ERVL-MaLR	1
LINE/R2	1
LINE	1
SINE/ID	1
DNA/PIF-Harbinger	1
LINE/RTE-BovB	1
DNA/hAT-Charlie	1
DNA/TcMar-Fot1	1
LTR/Lenti	1
LINE/I	1
LINE/CR1	1
DNA/P	1

Сосна сибирская кедровая



Лиственница сибирская

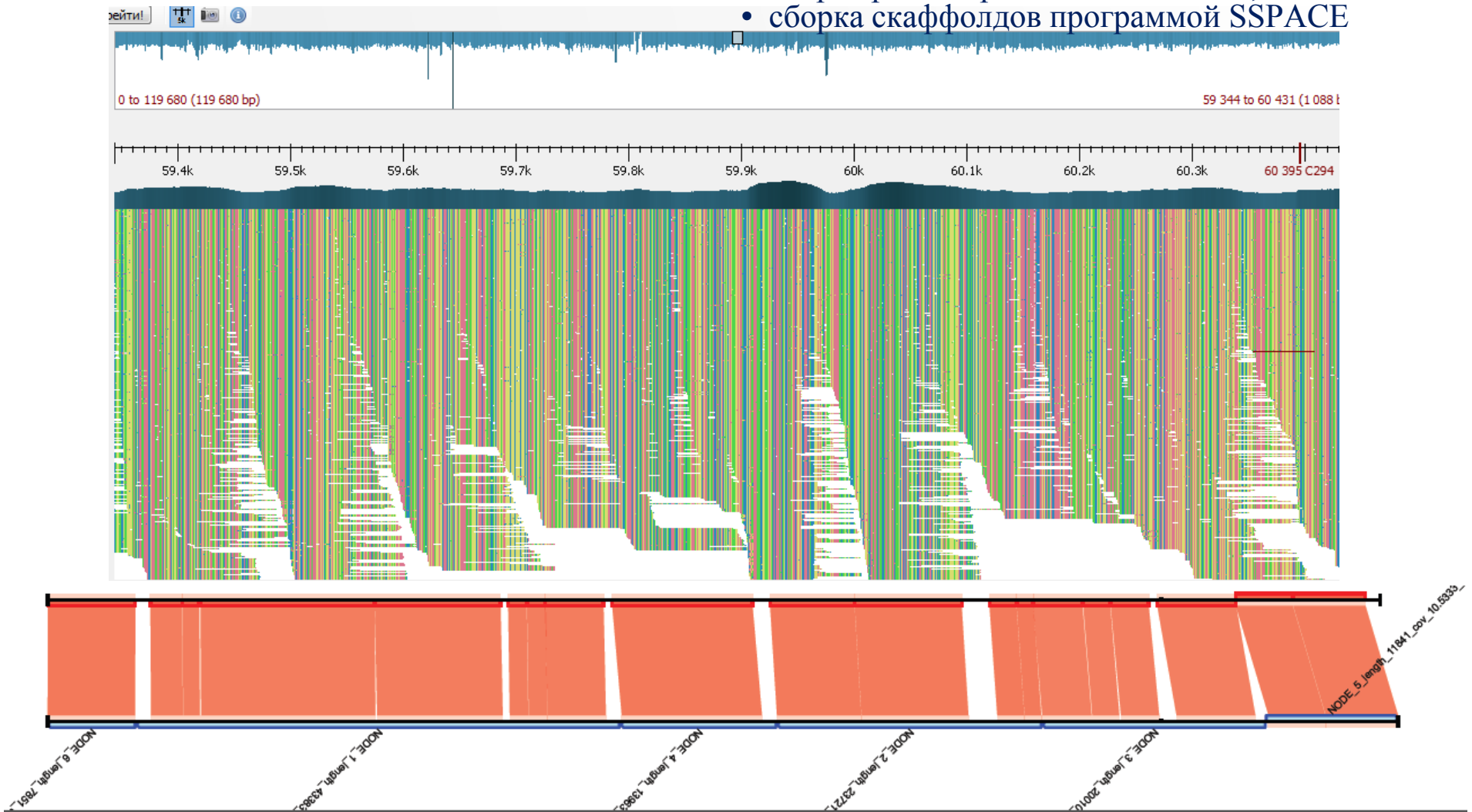


**большинство найденных повторов не
классифицированы!**

Картирование коротких сиквенсов (ридов) сибирской лиственницы на хлоропластный геном европейской лиственницы из Genbank и сборка

L. decidua (NC_016058, 122,474 bp, complete)

- картирование с помощью Bowtie2 и BWA программ
- сборка ридов при помощи SPAdes;
- сборка скаффолдов программой SSPACE



Аннотация хлоропластных геномов лиственницы



<i>Вид</i>	<i>Длина, bp</i>	<i>Сборка</i>	<i>Общее число генов</i>	<i>РНК гены</i>	<i>Протеи новые гены</i>	<i>Genbank</i>
<i>L. decidua</i>	122,474	полная	110	38	72	NC_016058
<i>L. occidentalis</i>	119,680	неполная	105	31	68	FJ899578
<i>L. sibirica</i>	122,561	полная	121	34	87	

- SNPs искали при помощи программы UGENE (Call Variants with SAMtools).
- Обнаружено 13 SNPs, 2 из них – в кодирующих районах

Результаты картирования коротких сиквенсов (ридов) сибирской лиственницы и кедра на митохондриальные геномы ели обыкновенной и ладанной сосны из соответствующих баз данных

P. taeda mitochondrial genome assembly (loblolly.ucdavis.edu/bipod/ftp/Genome_Data/genome/pinerefseq/Pita):

Scaffolds	Contigs	Total length, bp	Max length, bp	Min length, bp	N50, bp	N90, bp
4	31	1,263,957	256,879	124	193,087	58,893

Picea abies mitochondrial genome assembly (dl.dropboxusercontent.com/u/8576950/putative_mitochondrial_scaffold.fasta.zip):

Scaffolds	Contigs	Total length, bp	Max length, bp	Min length, bp	N50, bp	N90, bp
?	105	2,400,824	78,675	?	30,157	12,361

Species	Reads mapped to <i>Pinus taeda</i> mitochondrial genome, %	Reads mapped to <i>Picea abies</i> mitochondrial genome, %
<i>P. sibirica</i>	0.93	0.63
<i>P. sylvestris</i>	3.57	0.84
<i>L. sibirica</i>		0.12
<i>A. sibirica</i>	0.22	0.33

Mitochondrial genome assemblies for 4 conifer species in our study:

Species	Contigs	Total length, bp	Max length, bp	N50, bp	N90, bp
<i>P. sibirica</i>	431	482,064	9,882	2,274	398
<i>P. sylvestris</i>	2,586	1,190,450	19,705	2,003	177
<i>A. sibirica</i>	248	147,965	3,432	1,031	240
<i>L. sibirica</i>	209	69,615	1,039	333	227

De novo сборка транскриптома *Larix sibirica* Ledeb.

- 
- мРНК почек лиственницы сибирской (Illumina MiSeq);
 - Предварительная обработка – Trimmomatic и FastQC;
 - Сборка – Trinity;
 - Качество – PRINSEQ и QUAST;
 - Вспомогательные инструменты - Bowtie2 и SAMtools.

De novo сборка транскриптома *Larix sibirica* Ledeb.

- Общая длина – **26,493,048 bp**;
- Картирование ридов на сборку:
 - общий % выравнивания - **~87%**
 - покрытие – **23X** и **8014X** (среднее и максимальное)
- Число контигов (юнигенов) – **43,686** (из них **6,919** >1000 bp);
- Длина максимального контига – **8512 bp**;
- N50 – **878 bp**
- GC,% – **43,57**

Сравнительный анализ транскриптомов хвойных



Параметр	<i>Larix sibirica</i>	<i>Pseudotsuga menziesii</i>	<i>Picea abies</i>	<i>Pinus taeda</i>
Общая длина	26,493,048	32,492,919	22,248,988	73,956,391
Число КОНТИГОВ	43,686	25,041	19,026	48,842
Макс. контиг	8,512	7,562	6,424	9,828
N50	878	1,644	1,443	1,933

Comparative whole genome sequencing of major phytopathogens of boreal forests



Local die offs of pine-fir stand in Siberia and haploid culture of *Armillaria mellea* s.l., isolated from the roots of the affected trees (fresh dead wood and very weak trees)



The local spot of root fungus in pine forests of natural origin in Siberia and haploid culture *Heterobasidion annosum* (Fr.) Bref. s.s., isolated from the roots of the affected wood trees

Кратко полученные научные результаты:

- Все поставленные на 2014 г. задачи выполнены в полном объёме, а по некоторым показателям даже перевыполнены. Кроме того, дополнительно к запланированным задачам были получены чистые грибные культуры *Armillaria mellea* и *Heterobasidion annosum* с контрастной патогенностью, что изначально планировалось только в следующем 2015 г.
- Впервые получены предварительные рабочие сборки ядерного генома лиственницы и кедра.
- Заложены основы для полной сборки и аннотации геномов лиственницы и кедра.
- Впервые полностью собраны и проаннотированы хлоропластные геномы лиственницы и сосны.
- Впервые получены предварительные рабочие сборки митохондриального генома лиственницы и кедра.
- Полученные геномные нуклеотидные сборки могут быть и были использованы для поиска высокополиморфных молекулярно-генетических маркёров, таких как микросателлитные локусы, которые могут быть эффективно использованы для решения практических и природоохранных задач лесного хозяйства.

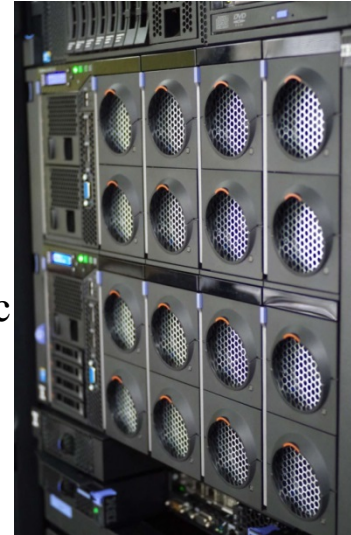
Приобретены и запущены:

Специализированный комплекс IBM для высокопроизводительных вычислений

В состав комплекса входят:

- 96 ядерный SMP сервер IBM x3950 X6 с ОЗУ 3 ТБ;
- Гибридный счетный сервер IBM dx360 M4 с двумя GPU NVIDIA Tesla K20 обеспечивающий суммарную пиковую производительность для вычислений одинарной точности с плавающей точкой до 7 Tflops;
- Подсистема хранения данных IBM Storwize V3700 объемом 48 Тб.
- Комплекс работает под управлением ОС SuSe Linux Enterprise Server 11, установлена параллельная файловая система IBM GPFS, система мониторинга Ganglia, система пакетной обработки Torque.

секвенатор Illumina MiSeq



Образовательная работа

Разработаны и утверждены 3 образовательных магистерских курса:

- «Молекулярная экология»
- «Геномика»
- «Популяционная генетика»

для обучающихся по магистерской программе 06.04.01.06 «Геномика и биоинформатика», направление 06.04.01. «Биология», реализуемой в Институте фундаментальной биологии и биотехнологии СФУ на кафедре «Защита и современные технологий мониторинга лесов»

Работает постоянный семинар по биоинформатике и геномике (рук-ли Юлия Путинцева, Дмитрий Кузмин и Наталья Орешкова)

Проведены многочисленные лекции

Публикации опубликованные (7)

- *Krutovsky K.V., Tretyakova I.N., Oreshkova N.V., Pak M.E., Kvitko O.V., Vaganov E.A.* Somaclonal variation of haploid in vitro tissue culture obtained from Siberian larch (*Larix sibirica* Ledeb.) megagametophytes for whole genome de novo sequencing // In Vitro Cellular and Developmental Biology – Plant. 2014. V. 50. No. 5. P. 655-664.
- *Крутовский К.В., Орешкова Н.В., Путинцева Ю.А., Ибе А.А., Дейч К.О., Шилкина Е.А.* Предварительные результаты полногеномного секвенирования сибирской лиственницы (*Larix sibirica* Ledeb.) и сибирской кедровой сосны (*Pinus sibirica* Du Tour.) // Сибирский лесной журнал. 2014. № 4. С. 77-81.
- *Крутовский К.В.* Перспективы использования геномных исследований в лесном хозяйстве // Сибирский лесной журнал. 2014. № 4. С. 11-15.
- *Шилкина Е.А., Орешкова Н.В., Ибе А.А., Дейч К.О., Крутовский К.В.* Разработка цитоплазматических SSR-маркеров для исследований сосны кедровой сибирской (*Pinus sibirica* Du Tour) // Сибирский лесной журнал. 2014. Т. 1. № 4. С. 21-24.
- *Крутовский К.В.* Исследования гигантских эволюционно древних геномов хвойных: популяционная, ландшафтная и экологическая геномика и геномная селекция // Труды VI Съезда ВОГиС. Ростов на Дону. 2014. С. 126-127.
- *Крутовский К.В., Орешкова Н.В., Путинцева Ю.А., Шилкина Е.А., Павлов И.Н., Ибе А.А., Дейч, К.О., Третьякова И.Н., Ваганов Е.А.* Геномные исследования бореальных лесов // Материалы Всероссийской научной конференции с международным участием, посвященная 70-летию Института леса им. В.Н. Сукачева СО РАН «Лесные биогеоценозы бореальной зоны: география, структура, функции, динамика». Красноярск, 2014. С. 19-22.
- *Tretyakova I., Ivanitskaya A., Park M., Voroshilova E., Oreshkova N., Krutovsky K.* Megagametophyte in vitro tissue culture of *Pinus sibirica* and *Larix sibirica* and somaclonal variation // In Park, Y.S. and Bonga, J.M. (eds.). Proceedings of the 3rd international conference of the IUFRO unit 2.09.02 on “Woody plant production integrating genetic and vegetative propagation technologies”, pp. 45-57. September 8-12, 2014. Vitoria-Gasteiz, Spain.

Публикации в печати (2), поданы (1) и рукописи (7)

- **Павлов И.Н.** Биотические и абиотические факторы усыхания хвойных лесов Сибири и Дальнего Востока // Сибирский экологический журнал (Contemporary Problems of Ecology). 2015 (в печати).
- **Семериков В.Л., Путинцева Ю.А., Орешкова Н.В., Крутовский К.В.** Разработка новых маркеров митохондриальной ДНК сосны обыкновенной (*Pinus sylvestris* L.) для популяционно-генетических и филогеографических исследований // Генетика. 2015 (в печати).
- **Babushkina E.A., Vaganov E.A., Grachev A.M., Oreshkova N.V., Belokopytova L.V., Kostyakova T.V., Krutovsky K.V.** Study of the effects of individual genetic heterozygosity on general homeostasis, heterosis and resilience in Siberian larch (*Larix sibirica* Ledeb.) using dendrochronology and microsatellite loci genotyping // Dendrochronologia. 2015 (submitted).
- **Bondar E.I., Putintseva Yu.A., Oreshkova N.V., Krutovsky K.V.** Siberian larch (*Larix sibirica* Ledeb.) chloroplast genome assembly, annotation and single nucleotide polymorphisms // (рукопись)
- **Biryukov V.V., Putintseva Yu.A., Oreshkova N.V., Krutovsky K.V.** Transcriptome *de novo* assembly and analysis in Siberian larch (*Larix sibirica* Ledeb.) // (рукопись)
- **Новикова С.В., Путинцева Ю.А., Орешкова Н.В., Крутовский К.В.** Анализ повторяющихся последовательностей в геноме сосны сибирской // (рукопись)
- **Ушакова О.А., Путинцева Ю.А., Орешкова Н.В., Дейч К.О., Крутовский К.В.** Разработка SSR-маркеров для *Larix sibirica* Ledeb. // (рукопись)
- **Павлов И.Н., Крутовский К.В., Литовка Ю.А., Рязанова Т.В., Чупрова Н.А.** Патогенные и дереворазрушающие свойства *Porodaedalea niemelaei* M. Fischer в редколесьях *Larix gmelinii* в условиях вечной заморозки // (рукопись)
- **Feranchuk S., Putintseva Y., Titievsky A., Oreshkova N., Kuzmin D., Ioffe Y., Farkov M., Ostretsov B., Lemberg D., Brodsky L., Krutovsky K.** Diversity of genomic repeats across larch, pine, and spruce species // (рукопись)
- **Belokon M., Polyakova T.A., Mudrik E.A., Belokon Y., Putintseva Y., Oreshkova N.V., Krutovsky K.V., Politov D.V.** Тестирование новой группы ядерных микросателлитных маркеров сосны кедровой сибирской, полученных при полногеномном секвенировании // (рукопись)

Благодарность за помощь и поддержку!



д.б.н. А.А. Онучин
д.б.н. И.Н. Третьякова
д.б.н. Л.И. Милютин
д.б.н. Е.Н. Муратова



Акад. РАН, д.б.н. Е.А. Ваганов
к.ф.-м.н. С.В. Первухин
к.ф.-м.н. В.А. Сапожников
к.с.-х.н. С.В. Верховец
к.ф.-м.н. Д.А. Кузьмин



В.В. Солдатов
В. Г. Разнобарский
к.б.н. Е.А. Шилкина

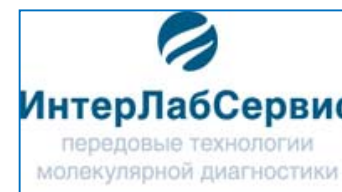


ИВИ СО РАН

д.ф.-м.н. М.Г. Садовский



Чл.-корр. РАН, д.б.н. Н.К. Янковский
д.б.н. Д.В. Политов
д.б.н. Е.И. Рогаев



А. Кирпий
М. Фурцев
А. Килин
В. Мытников
Н. Егоров

Научно-образовательный центр геномных исследований и Лаборатория лесной геномики Сибирского федерального университета



Genome Research and Education Center
Siberian Federal University

Русский

Лаборатория лесной геномики

- Руководитель
- Сотрудники
- Оборудование
- Учебная лаборатория
- Нам помогают
- Партнеры
- Контакты
- СМИ о нас

Сотрудники




в.н.с., к.б.н. Наталья Орешкова,
ответственная за пробоподготовку
и секвенирование

н.с. Юлия Путинцева,
ответственная за обработку
данных секвенирования



д.б.н., в.н.с. Павлов Игорь, ответственный за
направление фитопатологии



к.б.н. Суковатова Ирина



к.т.н. Дмитрий Кузьмин (БТК),
ответственный за вычислительный
кластер



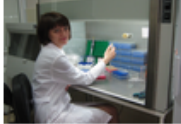
к.б.н. Елена Шиликова (БТК),
ответственная за разработку
образовательных курсов



н.с., к.с.-т.н. Алексей Ибе,
пробоподготовка



н.с. Ксения Дейч,
пробоподготовка



Шеллер Марина (БТК),
разработке образовательных курсов



Юшкова Татьяна (БТК),
фитопатология



Центр геномных исследований
Научно-образовательный центр (ФУ)

English

Новости

Главная

Лаборатория лесной геномики

- Руководитель
- Сотрудники
- Оборудование
- Учебная лаборатория
- Нам помогают
- Партнеры
- Контакты
- СМИ о нас

Лаборатория лесной геномики

Лаборатория лесной геномики создана в Сибирском федеральном университете в апреле 2014 года в рамках выполнения проекта «Геномные исследования основных бореальных лесообразующих хвойных видов и их наиболее опасных патогенов в Российской Федерации» Правительством РФ в рамках постановления № 220 от 9 апреля 2010 г. «О мерах по привлечению ведущих ученых в российские образовательные учреждения высшего профессионального образования».

[Приказ о создании лаборатории](#)

[Положение о лаборатории](#)

Руководитель лаборатории



Крутовский Константин Валерьевич

Координатор проекта



Шулелина Светлана Петровна

Сотрудники лаборатории



© 2012 НОЦ «Центр геномных исследований» СФУ
По вопросам редактирования сайта обращаться yuliya-putintseva@rambler.ru

Our projects

The main current project of the Center is the de novo sequencing and annotation of complete genomes of two conifer species - Siberian larch and Siberian stone pine - the main boreal forest tree species in Siberia and the

http://genome.sfu-kras.ru/en/forest_genomics

Особая благодарность за участие!



Студенты и магистранты СФУ:

Е.И. Бондарь, В.В. Бирюков,
В.В. Шаров, С.В. Маколов, О.
А. Ушакова, С. В. Новикова

Genome Research and Education Center
Siberian Federal University



Благодарность за финансирование

- Сибирский федеральный университет



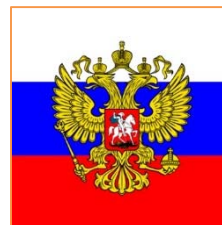
- Министерство образования и науки РФ



- Российский Фонд Фундаментальных Исследований



- Правительство РФ



Благодарность сотрудникам

Научно-образовательного центра геномных исследований Сибирского федерального университета

<http://genome.sfu-kras.ru/main>



Спасибо за внимание!

Желающие участвовать в проекте могут обратиться к Наталье Орешковой (секвенирование, oreshkova@ksc.krasn.ru) или Юлии Путинцевой (биоинформатика, yuliya-putintseva@rambler.ru)
http://genome.sfu-kras.ru/forest_genomics

