**СЕКВЕНИРОВАНИЕ И СБОРКА ГЕНОМОВ ЛИСТВЕННИЦЫ СИБИРСКОЙ (*LARIX SIBIRICA* LEDEB.) И СОСНЫ СИБИРСКОЙ КЕДРОВОЙ (*PINUS SIBIRICA DU TOUR*) И ПРЕДВАРИТЕЛЬНЫЕ ДАННЫЕ АНАЛИЗА ТРАНСКРИПТОМА**

***Орешкова Н.В.\*1,2, Путинцева Ю.А.1, Кузмин Д.А.1, Шаров В.В..1, Бирюков В.В.1,***

***Дейч К.О.1,6, Ибе A.A.1,6, Шилкина E.A.6***,***Крутовский К.В.1,3,4,5***

1Сибирский Федеральный Университет, Красноярск, Россия

2Институт леса им. В.Н. Сукачева СО РАН, Красноярск, Россия

3Институт общей генетики им. Н.И. Вавилова РАН, Москва, Россия

4Геттингенский университет, Геттинген, Германия

5Техасский агро-механический университет, Колледж Стейшин, Техас, США

6Филиал Российского центра защиты леса «Центр защиты леса Красноярского края», Красноярск, Россия

\*e-mail: oreshkova@ksc.krasn.ru

Лиственница сибирская (*Larix sibirica* Ledeb.) и сосна кедровая сибирская (*Pinus sibirica* Du Tour.) являются важными видами бореальных лесов Сибири и имеют огромное экономическое, экологическое и эстетическое значение. В связи с тем, что данные виды не являются модельными объектами геномных исследований, на текущий момент накоплено крайне мало знаний об их генах и геномах.Сложность исследований геномов этих видов связана также с их огромными размерами. Так, геном лиственницы сибирской составляет 12,03 Gbp, сосны сибирской кедровой – 23,6 Gbp, что в 4 и 7 раз больше размера генома человека.

Основной целью наших исследований является проведение полногеномного секвенирования, сборки, сравнительного анализа и аннотирования геномов лиственницы сибирской и кедра.

Секвенирование осуществляется с использованием секвенаторов Illumina HiSeq 2000 и MiSeq. Для *L. sibirica* произведено секвенирование парноконцевых (PE) библиотек длиной 250-500 bp и сцепленноконцевых (MP) библиотек для фрагментов 3000 bp, полученных из хвои, мегагаметофитов и каллусной культуры. Для *P.sibirica* секвенированы аналогичные библиотеки, полученные из мегагаметофитов. На текущий момент покрытие геномов составляет 74X и 25X для лиственницы и кедра соответственно. Также, произведено секвенирование транскриптома почек лиственницы сибирской.

Полученные данные секвенирвоания позволили получить черновые сборки геномов *L. sibirica* и *P. sibirica*, которые использовались для поиска микросателлитных локусов и дизайна праймеров к ним с целью дальнейшего выявления высокополиморфных и высокоинформативных SSR-маркеров для популяционно-генетических исследований. Для лиственницы сибирской впервые был собран и аннотирован хлоропластный геном, для сосны сибирской удалось улучшить сборку опубликованного в Genbank хлоропластного генома (FJ899558.1), закрыв пропуски суммарной длиной в 16085 bp. По результатам секвенирования образцов, обогащенных мтДНК проведены черновая сборка митохондриального генома и его предварительный анализ.

Сборка транскриптома позволила получить 43717 юнигенов общей длиной более 26 Mbp, длина наиболее длинного юнигена составила 8512 bp, N50 = 1330 bp, число юнигенов длиной более 1000 п.н.о. составило 6919. Сравнительный анализ полученной сборки с опубликованными сборками транскриптомов некоторых хвойных показал достаточно высокую степень сходства полученных контигов и контигов предыдущих сборок. Таким образом, на данный момент получена сборка, составляющая около 70% от предполагаемой общей длины транскриптома лиственницы сибирской.

*Работа выполнена в рамках проекта «Геномные исследования основных бореальных лесообразующих хвойных видов и их наиболее опасных патогенов в Российской Федерации», финансируемого Правительством РФ (договор № 14.Y26.31.0004).*

**GENOME SEQUENCING AND ASSEMBLY OF SIBERIAN LARCH (LARIX SIBIRICA LEDEB.) AND SIBERIAN PINE (PINUS SIBIRICA DU TOUR.)**

**AND PRELIMENARY TRANSCRIPTOME DATA**

***Oreshkova N.V.\*1,2, Putintseva Yu.A.1, Kuzmin D.A.1, Sharov V.V.1, Biryukov V.V.1, Makolov S.V.1, Deych K.O.1,6, Ibe A.A.1,6, Shilkina E.A.6***,***Krutovsky K.V.1,3,4,5***

1Siberian Federal University, Krasnoyarsk, Russia

2V.N. Sukachev Institute of Forest SB RAS, Krasnoyarsk, Russia

3N.I. Vavilov Institute of General Genetics RAS, Moscow, Russia

4University of Göttingen, Göttingen, Germany

5Texas A&M University, College Station, Texas, USA

6Branch of the Russian Centre for Forest Protection Centre for Forest Protection of Krasnoyarsk Territory, Krasnoyarsk, Russia

\*e-mail: oreshkova@ksc.krasn.ru

The whole genomes of Siberian larch (*Larix sibirica* Ledeb.) and Siberian pine (*Pinus sibirica* Du Tour.) were sequenced using Illumina HiSeq 2000 and MiSeq, and their first draft genome assemblies were generated*.* We searched for microsatellite loci in the assemblies and designed PCR primers to identify and develop highly polymorphic and informative SSR-markers for population genetic studies. For the first time the chloroplast genome of Siberian larch has been assembled and annotated. For Siberian pine we improved the chloroplast genome assembly published in Genbank (FJ899558.1) by closing gaps with the total gap length of 16085 bp. The draft assembly of mitochondrial genomes for these species has been done. The transcriptome assembly consisted of 43717 unigenes with a total length of ~26 Mbp. The longest unigene was 8512 bp; N50 = 1330 bp, and the number of unigenes longer than 1 Kbp was 6919. The obtained transcriptome assembly represented ~70% of the estimated total transcriptome in Siberian larch and was similar to other published conifer transcriptomes.

*This study was supported by Research Grant No. 14.Y26.31.0004 from the Government of the Russian Federation.*