**разработкА маркеров митохондриальной ДНК ключевых ХВОЙНЫХ видов сибирских бореальных лесов НА ОСНОВЕ геномного секвенирования и их использование в филогеографии**

Семериков В.Л.1, Путинцева Ю.А.2, Орешкова Н.В. 2,3, Крутовский К.В.2,4,5

1Институт экологии растений и животных УрО РАН, Екатеринбург, Россия, semerikov@ipae.uran.ru;

2Научно-образовательный центр геномных исследований Сибирского федерального университета, Красноярск, Россия; 3Института леса им. В.Н.Сукачева СО РАН, Красноярск, Россия; 4Гёттингенский университет, Гёттинген, Германия; 5Институт общей генетики им. Н. И. Вавилова РАН, Москва, Россия

Изучение истории биологического вида на основе анализа географического распределения генетической изменчивости составляет предмет филогеографии - отрасли популяционной генетики. Для видов семейства сосновые, ключевых хвойных видов сибирских бореальных лесов, одними из наиболее информативных генетических маркеров для филогеографических исследований являются маркеры митохондриальной ДНК. В частности, филогеография сосны обыкновенной в пределах большей части ареала на основе изменчивости митохондриальной ДНК ранее не была изучена в силу отсутствия полиморфных маркеров. Для их разработки мы использовали два подхода. Во-первых, с помощью метода обращенной ПЦР мы исследовали некодирующие фланкирующие области митохондриальных генов сосны обыкновенной и далее, на основе полученной последовательности разрабатывали новые пары ПЦР праймеров для поиска полиморфизма в амплифицированных участках посредством ресеквенирования в небольшой выборке особей разного географического происхождения. Этот способ позволил выявить изменчивость в генах *cox*I и *atp*A в европейской России. К сожалению, изменчивость *atp*A оказалась жестко сцепленной с изменчивостью уже изученного маркера *nad*7 и, поэтому, не является информативной. Изменчивость *cox*I позволила разделить выявленный ранее гаплотип ***c*** на два: ***сa*** и ***cc***. К западу от Твери встречены три гаплотипа – ***a***, ***ca*** и ***сс***. К востоку от Владимира – только ***a***, между этими пунктами – ***a*** и ***ca***. Второй подход основан на секвенировании общей ДНК, которая содержит ядерный и цитоплазматические (митохондриальный и хлоропластный) геномы, и включал следующие этапы: частичное секвенирование геномов сосны,кедра, сибирской лиственницы и пихты с помощью секвенатора Illumina HiSeq 2000, картирование ридов на уже доступные сборки митохондриального геномов, отбор ридов представляющих митохондриальный геном и сборку контигов, а также использование всех контигов для поиска гомологичных нуклеотидных сиквенсов в базе данных GenBank с помощью программы BLAST и отбор контигов, имеющих гомологию с митохондриальными генами растений. В результате было отобрано около 618 контигов общей длиной 1414000 п.н. На их основе были разработаны ПЦР праймеры для выборочного ресеквенирования отдельных участков с помощью метода Сангера. Преимущество при выборе отдавалось участкам, содержащим тандемные повторы, длиной более 3 повторов и мотивом длиной более 10 нуклеотидов. Участки общей длиной около 70000 п.н. были ресеквенированы у 8 деревьев, представляющих Восточную Европу, Кавказ и Монголию. Было выявлено 3 однонуклеотидных полиморфных маркёров – т.н. «снипов» (SNPs) по которым сосны с Кавказа отличались от остальной части ареала. Хотя в основной части ареала они не были изменчивы, но по двум из них обнаружена изменчивость в пределах Кавказа. Кроме того, в одном из контигов был найден минисателлит, имеющий мотив длиной 30 нуклеотидов, повторяющийся от 1 до 42 раз. Его генотипирование у 192 сосен сибирского, монгольского и восточноевропейского происхождения выявило 20 аллелей.

Работа поддержана РФФИ, проект 13-04-01028 и Правительством РФ (договор № 14.Y26.31.0004).

Development of mitochondrial DNA markers for conifers – key species of Siberian boreal forests from genomic sequencing and their use in phylogeography.

Semerikov V.L.1, Putintseva Y.A.2, Oreshkova N.V. 2,3, Krutovsky K.V.2,4,5

1Institute of Plant and Animal Ecology, Ural Branch of Russian Academy of Sciences,

Ykaterinburg, Russia, semerikov@ipae.uran.ru;

2 Genome Research and Education Center, Siberian Federal University, Krasnojarsk, Russia; 3V.N.Sukachev Institute of Forest, Siberian Branch of Russian Academy of Sciences, Krasnoyarsk, Russia;

4 University of Göttingen, Göttingen, Germany;

5N.I.Vavilov Institute of General Genetics Russian Academy of Science, Moscow, Russia

Studying the history of the species on the basis of analysis of the geographical distribution of genetic variation is the subject of phylogeography - section of population genetics. For conifer species, key elements of Siberian boreal forest, one of the most informative genetic markers for phylogeographic studies are markers of mitochondrial DNA. The phylogeography of Scots pine based on the variability of mitochondrial DNA within the great part of the area has not been previously studied due to the lack of polymorphic markers. For their development, we used two approaches. First, using the method of reverse-PCR, we investigated the non-coding flanking regions of the mitochondrial genes of Scots pine and further, based on the obtained sequences developed new pairs of PCR primers to find polymorphism in the amplified fragments by resequencing in a small sample of individuals of different geographic origin. This method revealed variability in genes *cox*I and *atp*A in European Russia. Unfortunately, polymorphism in atpA was tightly linked to the variability of the previously investigated marker *nad*7 and, therefore, is not informative. Variability of *cox*I allowed to split the previously identified haplotype ***c*** into two: ***ca*** and ***cc***. Three haplotype - ***a, ca*** and ***cc*** were met to the west of Tver. To the east of Vladimir - just ***a***, between these points - ***a*** and ***ca***. The second approach is based on the sequencing of the total DNA including nuclear and cytoplasmic (mitochondrial and chloroplast) genomes, and included the following steps: a partial sequencing of the genomes of Scots pine, Siberian stone pine, larch and fir via sequencer Illumina HiSeq 2000, mapping of the obtained rides on the already available assemblies of mitochondrial genomes, selection of the rides representing mitochondrial genome, assembling of contigs, and also the use of contigs for BLAST search in the GenBank database and selection of contigs having homology to mitochondrial genes of plants. As a result, 618 contigs were selected of a total length about 1414000 bp. Then the PCR primers were designed for partial amplification and resequencing of them using the Sanger method. Regions containing tandem repeats longer than 3 repeats with motif longer than 10 nucleotides had an advantage in the choice of the target for re-sequencing. The selected fragments with a total length of about 70,000 bp were re-sequenced in eight trees representing Eastern Europe, Caucasus and Mongolia. 3 SNPs markers were revealed which segregated the pines in Caucasus from the rest of the range. Although they were not variable within the main part of the range, two of them were polymorphic within Caucasian. Furthermore, one minisatellite marker was found having a 30 nucleotide motif and repeat number ranged between 1 to 42. Its genotyping in 192 trees of Siberian, Mongolian and East European origin revealed 20 alleles.

The research was supported by Russian Fund for Basic Research, project 13-04-01028 and by Government of Russian Federation (contract 14.Y26.31.0004).