**Supporting information**

**Title**: «Late Quaternary history of Siberian stone pine as revealed by genetic and paleoecological data»

**Tree Genetics & Genomes**

**Authors:**

Denis N. Shuvaev\*, Vladimir L. Semerikov, Galina V. Kuznetsova, Yulia A. Putintseva

**\*Correspondence:** Denis N. Shuvaev, Institute of Plant and Animal Ecology, Ural Branch of the Russian Academy of Sciences, 8 Marta Str. 202, Ekaterinburg 620144, Russia. E-mail: denis.shuvaev@gmail.com

**Appendix S1** Data on samples………………………………………………………….3

**Appendix S2** Development of mtDNA markers……………………………….……....9

**Appendix S3** Species distribution modeling………………………………………......13

**Appendix S4** Standard genetic statistics………...……………………………….....…18

**Appendix S5** TheNJ tree...............................................................................................23

**Appendix S6** Structure results and K values..............................................................24

**Appendix S7** Reference list of paleorecords...................................................................25

**Supplementary Tables**

**Table S1.1** Geographical locations of *P. sibirica* samples

**Table S2.2** Genome assembly statistics for two DNA samples from *P. sibirica* (two trees)

**Table S2.3** Mitochondrial contigs, GenBank accession numbers, detected SNPs, nucleotide sequences of primers used for sequencing and for single-strain conformation polymorphism (SSCP) genotyping, and length of the PCR products amplified from *P. sibirica* DNA

**Table S3.4** List of selected bioclimatic variables

**Table S3.5** Performance statistics for the best models selected on the basis of the user’s pre-defined criteria

**Table S4.6** Standard genetic statistics of *P. sibirica* populations

**Table S4.7** The standard genetic statistics of nuclear microsatellite loci

**Supplementary Figures**

**Fig. S2.1** Genotyping of Siberian stone pine mtDNA markers by the SSCP method in a nondenaturing polyacrylamide gel

**Fig. S3.2** Distribution of all models, statistically nonsignificant models, and selected models

**Fig. S3.3** The relation of TSS' (the true skill statistic for pseudo-absence data) with threshold levels of binary transformations

**Fig. S3.4** The training omission rate and predicted area as a function of the cumulative threshold, averaged across the replicate runs

**Fig. S3.5** The receiver operating characteristic (ROC) curve averaged across the replicate runs

**Fig. S5.6** TheNJ tree based on *DA* distances derived from the *P. sibirica* microsatellite data

**Fig. S6.7** Genetic structure of *P. sibirica* populations as inferred by STRUCTURE

**Appendix S1**

**Table S1.1** Geographical locations of *P. sibirica* samples

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| No. | Sample code (locality name) | Lat. N/Long. E (decimal degrees, WGS84) | Sample size | Sample size for mtDNA | Sample size for nuclear DNA |
| Altai | | | | | |
| 1 | Tal (Talda) | 50.80/85.74 | 30 | 24 | 30 |
| 2 | Art (Artybash) | 51.80/87.29 | 30 | 24 | 30 |
| 3 | Kur (Kurai) | 50.12/87.82 | 30 | 24 | 30 |
| 4 | Ong (Ongudai) | 51.04/85.60 | 30 | 24 | 30 |
| 5 | Sov (Sovetskoe) | 52.04/86.63 | 27 | 24 | ‒ |
| 6 | Ber (Berel') | 49.37/86.43 | 14 | 14 | 14 |
| 7 | Man (Manzherok) | 51.82/85.78 | 10 | 10 | ‒ |
| 8 | V-b (Verkh-Biysk) | 52.05/87.08 | 20 | 20 | ‒ |
| 9 | Akt2K (Aktash2K) | 50.31/87.56 | 31 | 24 | 30 |
| 10 | Sug (Sugul) | 51.90/86.21 | 34 | 24 | 30 |
| 11 | Usk (Uskuch) | 52.04/86.81 | 33 | 24 | ‒ |
| 12 | Akt3K (Aktash3K) | 50.47/87.62 | 30 | ‒ | 30 |
| West Sayan | | | | | |
| 13 | A/O (Abazinskaya/  Oninskaya) | 51.84/89.78 | 30 | 24 | ‒ |
| 14 | Shag (Shagonar) | 51.70/92.94 | 30 | 24 | 30 |
| 15 | AON (Bolsheolonskaya) | 52.11/89.70 | 30 | 24 | 30 |
| 16 | Sh (Shushenskoe) | 52.96/91.72 | 30 | 24 | 30 |
| 17 | U/A (Usinskoe/  Aradanskoe) | 52.70/93.40 | 30 | 24 | 30 |
| 18 | RTKCHR (Cherbi) | 51.90/94.63 | 30 | 24 | 30 |
| 19 | UR (Aradanskaya) | 52.63/93.84 | 30 | 24 | ‒ |
| 20 | ABA (Abazinskoe) | 52.60/90.05 | 30 | 24 | 30 |
| 21 | RTBK/z (Balgazyn) | 50.55/95.01 | 30 | 24 | 30 |
| East Sayan | | | | | |
| 22 | IOT/Tg (Georgievskaya) | 55.48/97.43 | 30 | 24 | 30 |
| 23 | KV/S (Verkhne-Sisimskaya) | 54.77/92.94 | 30 | 24 | 30 |
| 24 | Chun (Chunskaya) | 57.98/96.03 | 30 | 24 | 30 |
| 25 | Tai (Tayshet) | 55.96/97.97 | 16 | 16 | 16 |
| 26 | Arsh (Arshan) | 51.91/102.43 | 30 | 24 | 30 |
| 27 | IONK (Nijneudinskaya) | 54.95/98.36 | 30 | 24 | 30 |
| 28 | IOTSH (Dzhogino) | 56.63/98.33 | 30 | 24 | 30 |
| 29 | IOZCH (Hor-Tagnanskaya) | 53.13/101.28 | 30 | 24 | 30 |
| 30 | CHD (Slyudyanka) | 51.65/103.66 | 34 | 24 | 30 |
| 31 | KAN (Kanskaya) | 56.45/94.38 | 30 | 24 | 30 |
| middle Siberia | | | | | |
| 32 | A/IA (Achinsk) | 56.16/90.90 | 29 | 24 | 29 |
| 33 | BMV (Verkh-Kazanka) | 57.01/92.47 | 30 | 24 | 30 |
| 34 | Bir (Vershina Tyoi) | 53.28/89.59 | 24 | 24 | 24 |
| 35 | M/B (Beret') | 55.68/92.94 | 30 | 24 | 30 |
| 36 | Van (Vanavara) | 60.34/102.30 | 25 | 25 | 25 |
| 37 | TUCH (Tiukhtet) | 56.53/89.31 | 29 | 24 | 29 |
| 38 | IK (Kesovskaya) | 55.30/96.71 | 30 | 24 | ‒ |
| 39 | KAR (Karatuzskoe) | 53.56/92.90 | 30 | 24 | 30 |
| 40 | TK (Rassvet) | 57.00/91.51 | 30 | 24 | ‒ |
| 41 | KGj (Kozul'ka) | 55.78/91.65 | 30 | 24 | ‒ |
| 42 | EM (Mostovskoe) | 56.68/92.94 | 30 | 24 | ‒ |
| 42 | PIR (Pirovskoe) | 57.60/92.30 | 30 | 24 | ‒ |
| 44 | D (Cheremushki) | 55.11/92.25 | 30 | 24 | ‒ |
| 45 | KIZ (Artyomovsk) | 54.28/93.73 | 30 | 24 | ‒ |
| 46 | KT (Mikhailovka) | 56.43/91.69 | 30 | 24 | ‒ |
| 47 | Tur (Turukhansk) | 65.79/87.96 | 30 | 24 | 30 |
| 48 | SE (Pit-Gorodok) | 59.37/93.65 | 29 | 24 | 29 |
| 49 | Ti (Belyaki) | 58.85/96.73 | 30 | 24 | 30 |
| Kuznetsk Alatau | | | | | |
| 50 | Maka­­­(Makarikha) | 54.18/87.80 | 32 | 24 | 32 |
| 51 | Arg-ula (Argat-Yul) | 57.88/86.02 | 30 | 24 | 30 |
| 52 | Mis (Myski) | 53.74/87.96 | 31 | 24 | 30 |
| 53 | Tocha (Chainsk) | 58.08/82.33 | 30 | 24 | 30 |
| 54 | TON (Nibega) | 58.24/84.66 | 30 | 24 | 30 |
| 55 | KOIKa (Rudnichnyi) | 55.94/86.52 | 30 | 24 | 30 |
| 56 | KOM/Mr (Mezhdurechensk) | 53.81/88.30 | 30 | 24 | ‒ |
| West Siberian Plain | | | | | |
| 57 | Sura (Surgut) | 61.50/74.25 | 34 | 24 | 30 |
| 58 | NKa(Noyabrsk) | 63.20/75.45 | 24 | 24 | 24 |
| 59 | NUa (Novy Urengoy) | 66.08/76.68 | 26 | 24 | 26 |
| 60 | CHM (Khanty-Mansiysk) | 61.02/69.02 | 30 | 24 | 30 |
| Urals | | | | | |
| 61 | Ivda (Ivdel') | 60.75/60.45 | 20 | 20 | 20 |
| 62 | Aiza (Ayatskoe) | 57.37/60.70 | 18 | 18 | 18 |
| 63 | Iam (Yamgort) | 64.94/64.38 | 17 | 17 | Iamal as one sample  51 |
| 64 | Mujia (Muzhi) | 65.40/64.70 | 17 | 17 |
| 65 | NIM (Nymvozhgort) | 64.89/63.90 | 5 | 5 |
| 66 | StKiev (Staryi Kievat) | 65.14/64.69 | 8 | 8 |
| 67 | Ovg (Ovgort) | 64.82/63.99 | 4 | 4 |
| Total | ‒ | ‒ | 1821 | 1470 | 1387 |

apopulation samples were studied earlier by Shuvaev and Ibe (2021) using microsatellite markers

**Appendix S2**

*1) De novo sequencing of Siberian stone pine genomic DNA, selection of mtDNA contigs, and the search for polymorphism*

We performed paired-end (PE) whole-genome sequencing for two Siberian stone pine trees growing in natural populations of the West Sayan and the Urals using Illumina HiSeq 2000. For this sequencing, we used a PE DNA library (with the insert size of 400 bp) generated using total DNA isolated with the CTAB method from needles. 1) The library preparation was performed following a standard Illumina TruSeq 2 protocol. For the sequencing, we used the 2×100 cycles Illumina Kit. We employed FastQC and Trimmomatic for quality control and adapter trimming, respectively. 2) Assembly of contigs was performed using CLC ASSEMBLY CELL (QIAGEN, Hilden, Germany). The expected genome size of *P. sibirica* is 28.959 Gbp (Zonneveld 2012). Because of low coverage, the genome assembly was very partial and included only 0.1–0.2% of the expected genome size.

**Table S2.2** Genome assembly statistics for two DNA samples from *P. sibirica* (two trees)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sample | Number of contigs | N50, bp | Maximum length, bp | Total length, Mbp | Genome sequencing coverage |
| Urals | 141786 | 292 | 92477 | 46.57 | 0.2× |
| West Sayan | 879595 | 2622 | 375860 | 283.4 | 17.2× |

Nonetheless, because there are multiple copies of the mitogenome per nuclear genome in one plant cell, we were able to identify several mitochondrial contigs. 3) A search for mitochondrial contigs was carried out using BLASTn and all plant mitochondrial sequences available in the NCBI GenBank and other public databases, such as https://plantgenie.org/FTP and https://treegenesdb.org/FTP/Genomes/Pita/mito for Norway spruce (*Picea abies* [L.] Karst.) and loblolly pine (*Pinus taeda* L.), respectively. Matching Siberian stone pine contigs among the BLASTn hits with an alignment length of more than 100 bp and similarity higher than 90% were selected for further analysis. In total, 173 contigs with the total length of 1,283,256 bp were chosen for the Ural sample and seven contigs with the total length of 2,967,893 bp for the West Sayan sample. The assembly for the Ural sample is fragmented as compared with the West Sayan assembly, as explained by a difference in sequencing coverage: the West Sayan sample was sequenced with 17.2× coverage, whereas the other one, the Ural sample, was sequenced with very low coverage, 0.2× only, to search for possible single-nucleotide polymorphisms (SNPs).

The length of the resulting draft assemblies of the Siberian pine mitochondrial genome is consistent with data on other conifer mitogenomes, which vary widely: from 1.25 Mbp in *P. taeda* (Neale et al. 2014), 3.9 Mbp in *Pinus lambertiana* (Stevens et al. 2016), and 4.3 Mbp (Nystedt et al. 2013) or 4.9 Mbp (Sullivan et al. 2020) for *P. abies* to 11.7 Mbp in *Larix sibirica* (Putintseva et al. 2020).

The selected mtDNA contigs were then subjected to a pairwise comparison of two tree mitogenomes (Ural and West Sayan) using the BLASTN Web service (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome>) for targeted polymorphism searching. The mitogenome comparison gave two distinct SNP sites. PCR primers (Table S2.3) were designed by means of the Primer3 software (Rozen and Skaletsky 2000). These SNP markers were used for a preliminary search for interpopulation differentiation among eight trees representing different parts of the Siberian stone pine range including the Altai, Kuznetsky Alatay, East Sayan, West Sayan, and Urals.

**Table S2.3** Mitochondrial contigs, GenBank accession numbers, detected SNPs, nucleotide sequences of primers used for sequencing and for single-strain conformation polymorphism (SSCP) genotyping, and length of the PCR products amplified from *P. sibirica* DNA

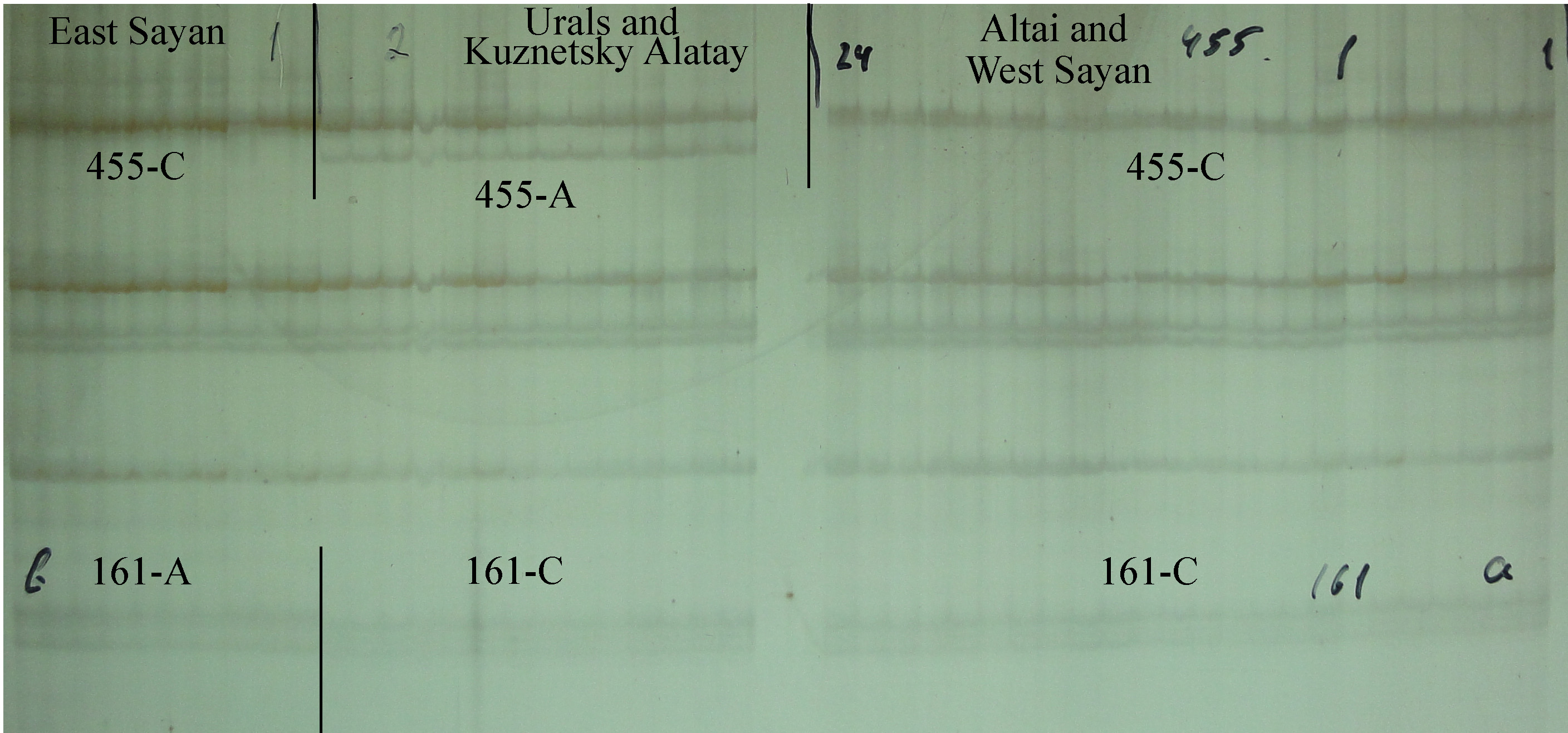
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Contig ID | GenBank accession No. | SNP (position in contig, bp) | Names and nucleotide sequences of primers,  5′→3′ | Length of the PCR product, bp |
| 161 | Contig 161 sequence OL944294 | С/A (10557 bp) | Si161-1Fa: CGAAGCACTCTTCACCTTCC  Si161-1R: TTCATGCTGCCAACTCTTTG | 1471 |
| OL944295 | С | Si161-1R1ab: TTGTGTCGGAACAATCTCCA | 184 |
| OL944296 | A |
| 455 | Contig 455 sequence OL944297 | C/A (4976 bp) | Si455F: CGGGATCCAGTCCAATCTAA  Si455R: TCCCTGCATTCAATGTGTGT | 1020 |
| OL944298 | С | Si455F1a: CGCTCAGGACAGTCCATACA  Si455R2a: CCAGGGAAGCTTTCTCTCCT | 241 |
| OL944299 | A |

aUsed for SSCP genotyping

bUsed in a pair with Si161-1F

2) *Genotyping of SNPs*

The SSCP method for SNP genotyping was described by Semerikov et al. (2019). We present an example of an electropherogram obtained by the SSCP method below.



**Fig. S2.1** Genotyping of Siberian stone pine mtDNA markers by the SSCP method in a nondenaturing polyacrylamide gel. Four variable electrophoretic zones of single-stranded fragments correspond to SNP sites: 455-A, 455-C, 161-A, and 161-C

**Appendix S3**

*1) Bioclimatic variables*

Highly correlating climatic variables (r > |0.8|) were excluded from subsequent SDM (species distribution modeling) analyses. The remaining variables are listed below.

**Table S3.4** List of selected bioclimatic variables

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Bio3 | Bio7 | Bio8 | Bio10 | Bio11 | Bio15 | Bio18 | Bio19 |
| Bio3 | 1.0 | 0.2 | 0.2 | 0.5 | 0.3 | 0.1 | 0.1 | -0.1 |
| Bio7 | 0.2 | 1.0 | 0.2 | 0.1 | -0.7 | 0.2 | 0.0 | -0.6 |
| Bio8 | 0.2 | 0.2 | 1.0 | 0.6 | 0.3 | 0.4 | 0.4 | 0.0 |
| Bio10 | 0.5 | 0.1 | 0.6 | 1.0 | 0.6 | 0.3 | 0.1 | 0.2 |
| Bio11 | 0.3 | -0.7 | 0.3 | 0.6 | 1.0 | 0.0 | 0.1 | 0.6 |
| Bio15 | 0.1 | 0.2 | 0.4 | 0.3 | 0.0 | 1.0 | 0.2 | -0.3 |
| Bio18 | 0.1 | 0.0 | 0.4 | 0.1 | 0.1 | 0.2 | 1.0 | 0.3 |
| Bio19 | -0.1 | -0.6 | 0.0 | 0.2 | 0.6 | -0.3 | 0.3 | 1.0 |

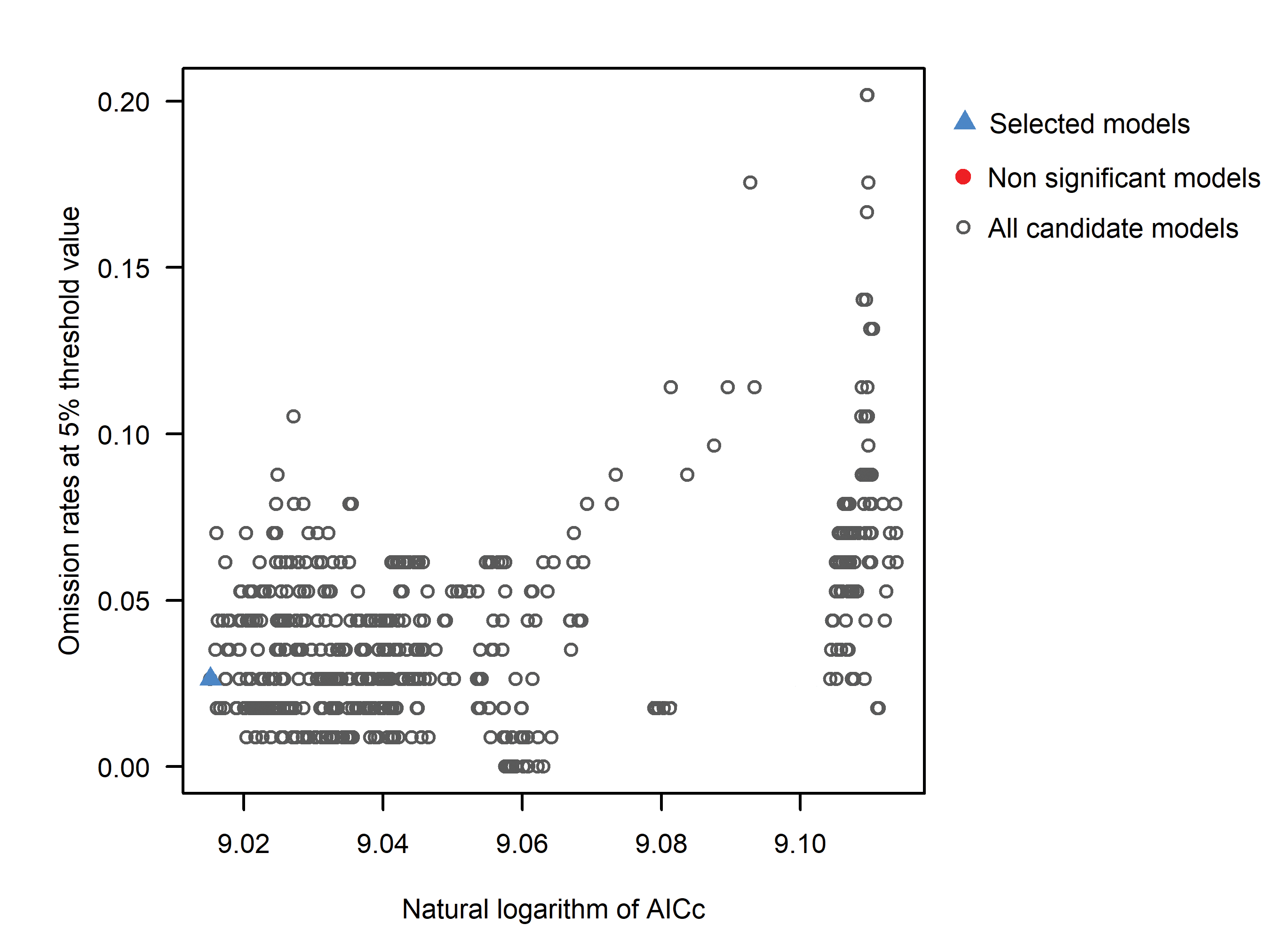
Bio3: isothermality, Bio7: the temperature annual range, Bio8: mean temperature of the wettest quarter, Bio10: mean temperature of the warmest quarter, Bio11: mean temperature of the coldest quarter, Bio15: precipitation seasonality, Bio18: precipitation in the warmest quarter, and Bio19: precipitation in the coldest quarter

*2) Model calibration and the selection process*

Overall, 555 candidate models—with parameters reflecting all combinations of five regularization multiplier settings, three feature class combinations, and 37 distinct sets of environmental variables—were evaluated. Model performance was assessed based on statistical significance (Partial\_ROC: the partial area under the receiver operating characteristic), omission rates (ORs), and the Akaike information criterion corrected for small sample sizes (AICc).

**Table S3.5** Performance statistics for the best models selected on the basis of the user’s pre-defined criteria

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Mean\_AUC\_ratio | Partial\_ROC | Omission\_rate\_at\_5% | AICc | delta\_AICc | W\_AICc | num\_parameters |
| M\_0.5\_F\_lqp\_Set\_2 | 1.691 | 0 | 0.026 | 8227.39 | 0 | 0.92 | 19 |



**Fig. S3.2** Distribution of all models, statistically nonsignificant models, and selected models

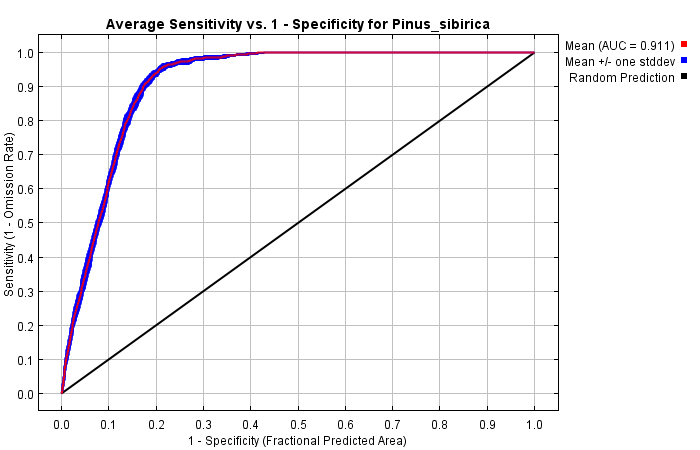
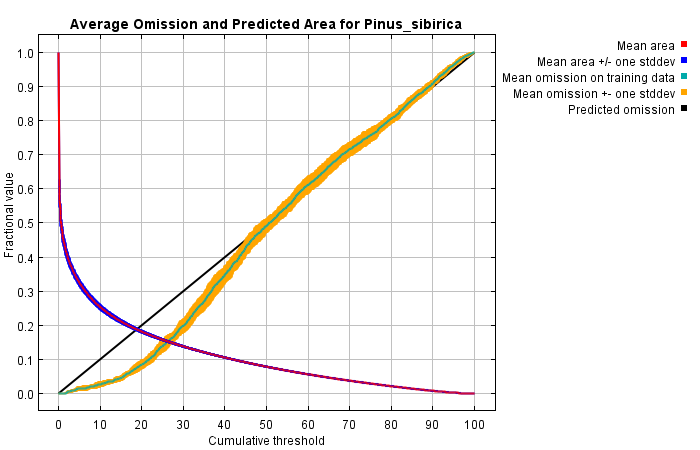


**Fig. S3.3** The relation of TSS' (the true skill statistic for pseudo-absence data) with threshold levels of binary transformations

*3) Results*

All algorithms of modeling were performed with 500 iterations across 10 replicates. The AUC (the area under the curve) value was 0.911 ± 0.003 for the random 50% of test datasets when the MAXENT model for eight variables was employed.

**Fig. S3.4** The training omission rate and predicted area as a function of the cumulative threshold, averaged across the replicate runs



**Fig. S3.5** The receiver operating characteristic (ROC) curve averaged across the replicate runs

The red curve shows how well the model fits the training data (“fit”). The blue curve indicates how well the model fits the test data and is a real test of the model's predictive ability. The black line shows the situation that would be expected if the reliability of the model's predictions were at a random level. The closer the blue line to the upper left corner, the better the model predicts the findings contained in the test set.

**Appendix S4**

**Table S4.6** Standard genetic statistics of *P. sibirica* populations

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sample name | | *NA*±SE | *Ne*±SE | *HO*±SE | *HE*±SE |
| Altai | | | | | |
| Tal | | 3.375±0.565 | 2.116±0.316 | 0.352±0.046 | 0.466±0.066 |
| Art | | 3.375±0.46 | 2.060±0.305 | 0.400±0.065 | 0.443±0.074 |
| Kur | | 3.500±0.5 | 1.952±0.22 | 0.369±0.052 | 0.447±0.056 |
| Ong | | 3.500±0.5 | 2.077±0.3 | 0.429±0.057 | 0.450±0.073 |
| Ber | | 3.125±0.515 | 2.185±0.245 | 0.378±0.08 | 0.490±0.075 |
| Akt2K | | 3.625±0.532 | 1.904±0.142 | 0.458±0.052 | 0.449±0.051 |
| Sug | | 3.250±0.412 | 2.000±0.195 | 0.428±0.072 | 0.460±0.061 |
| Akt3K | | 3.500±0.5 | 1.933±0.219 | 0.404±0.068 | 0.431±0.069 |
| Mean±SE | | 3.406±0.498 | 2.028±0.243 | 0.402±0.061 | 0.455±0.066 |
| West Sayan | | | | | |
| Shag | | 3.750±0.559 | 1.931±0.196 | 0.421±0.057 | 0.446±0.053 |
| AON | | 3.625±0.532 | 1.985±0.176 | 0.440±0.061 | 0.458±0.065 |
| Sh | | 3.625±0.706 | 2.024±0.249 | 0.467±0.058 | 0.461±0.057 |
| U/A | | 3.500±0.681 | 2.306±0.443 | 0.443±0.047 | 0.490±0.065 |
| RTKCHR | | 3.500±0.567 | 2.029±0.245 | 0.424±0.06 | 0.458±0.064 |
| ABA | | 3.250±0.366 | 2.004±0.202 | 0.450±0.078 | 0.456±0.068 |
| RTBK/z | | 3.125±0.398 | 2.067±0.265 | 0.438±0.063 | 0.460±0.071 |
| Mean±SE | | 3.482±0.544 | 2.049±0.254 | 0.440±0.061 | 0.461±0.063 |
| East Sayan | | | | | |
| Tai | | 3.000±0.378 | 2.091±0.328 | 0.422±0.039 | 0.452±0.069 |
| Arsh | | 3.500±0.5 | 2.091±0.34 | 0.412±0.078 | 0.438±0.081 |
| M/B | | 3.625±0.498 | 2.179±0.391 | 0.379±0.049 | 0.454±0.075 |
| IONK | | 3.500±0.5 | 2.015±0.357 | 0.397±0.076 | 0.414±0.08 |
| IOTSH | | 3.500±0.681 | 2.207±0.335 | 0.438±0.074 | 0.470±0.082 |
| IOZCH | | 3.750±0.62 | 2.334±0.471 | 0.398±0.068 | 0.474±0.082 |
| CHD | | 3.250±0.648 | 2.426±0.553 | 0.430±0.06 | 0.449±0.097 |
| IOT/Tg | | 3.375±0.46 | 2.102±0.305 | 0.413±0.07 | 0.453±0.077 |
| KV/S | | 3.875±0.693 | 2.245±0.402 | 0.434±0.058 | 0.471±0.076 |
| Van | | 3.750±0.648 | 2.132±0.392 | 0.450±0.085 | 0.441±0.079 |
| Chun | | 3.375±0.565 | 2.016±0.335 | 0.412±0.066 | 0.427±0.077 |
| KAN | | 3.750±0.62 | 2.121±0.365 | 0.424±0.084 | 0.446±0.076 |
| Mean±SE | | 3.521±0.568 | 2.163±0.381 | 0.417±0.067 | 0.449±0.079 |
| middle Siberia | | | | | |
| KAR | | 3.375±0.498 | 2.161±0.358 | 0.448±0.065 | 0.456±0.076 |
| SE | | 3.625±0.532 | 1.989±0.262 | 0.399±0.047 | 0.442±0.065 |
| Ti | | 3.500±0.567 | 2.123±0.388 | 0.409±0.088 | 0.431±0.083 |
| Tur | | 3.125±0.35 | 1.930±0.212 | 0.423±0.052 | 0.441±0.056 |
| BMV | | 3.125±0.35 | 2.017±0.239 | 0.417±0.064 | 0.451±0.068 |
| A/IA | | 3.375±0.42 | 2.013±0.288 | 0.442±0.09 | 0.436±0.074 |
| Mean±SE | | 3.354±0.453 | 2.039±0.291 | 0.423±0.068 | 0.443±0.07 |
| Kuznetsk Alatau and West Siberian Plain | | | | | |
| NU | | 3.500±0.463 | 2.297±0.324 | 0.356±0.053 | 0.500±0.069 |
| NK | | 3.125±0.227 | 2.142±0.237 | 0.433±0.069 | 0.494±0.055 |
| Mak | | 3.625±0.596 | 2.296±0.359 | 0.461±0.067 | 0.489±0.077 |
| Arg-ul | | 3.500±0.5 | 2.021±0.206 | 0.463±0.053 | 0.466±0.057 |
| Bir | | 3.500±0.5 | 1.875±0.286 | 0.365±0.069 | 0.398±0.072 |
| Mis | | 3.375±0.565 | 1.807±0.168 | 0.371±0.056 | 0.408±0.064 |
| Toch | | 3.375±0.498 | 2.240±0.325 | 0.442±0.071 | 0.484±0.074 |
| TON | | 3.500±0.463 | 2.154±0.265 | 0.481±0.072 | 0.481±0.068 |
| KOIK | | 3.250±0.491 | 2.214±0.313 | 0.421±0.076 | 0.473±0.083 |
| TUCH | | 3.375±0.532 | 1.971±0.256 | 0.411±0.067 | 0.430±0.074 |
| Mean±SE | | 3.413±0.483 | 2.102±0.274 | 0.420±0.065 | 0.462±0.069 |
| Urals and West Siberian Plain | | | | | |
| Aiz | | 3.000±0.378 | 2.267±0.279 | 0.472±0.071 | 0.507±0.064 |
| Ivd | 3.125±0.35 | 2.072±0.226 | 0.419±0.037 | 0.478±0.054 |
| CHM | 3.375±0.42 | 2.331±0.25 | 0.467±0.026 | 0.537±0.048 |
| Sur | 3.375±0.42 | 2.060±0.177 | 0.483±0.052 | 0.487±0.048 |
| Iamal | 3.125±0.398 | 2.162±0.218 | 0.397±0.035 | 0.500±0.055 |
| Mean | 3.200±0.393 | 2.178±0.23 | 0.448±0.044 | 0.502±0.054 |
| Total mean±SE | 3.419±0.07 | 2.097±0.041 | 0.423±0.009 | 0.459±0.01 |

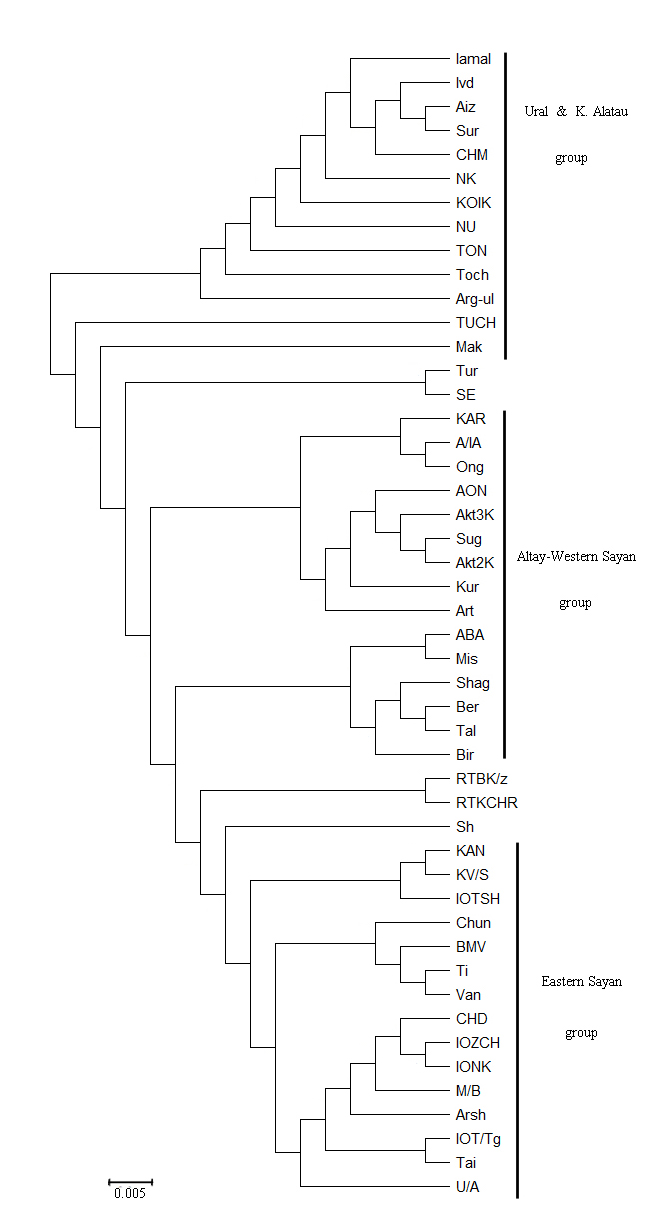
*NA*: the allele number, *Ne*: the effective allele number, *HO*: observed heterozygosity, *HE*: expected heterozygosity, and SE: standard error

**Table S4.7** The standard genetic statistics of nuclear microsatellite loci

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID No. | Locus name | *HO* | *HE* | *NA* | Allelic size, bp |
| 1 | PS\_80612 | 0.553 | 0.585 | 7 | 162, 165, 168, 171, 177, 180, 183 |
| 2 | PS\_364418 | 0.467 | 0.761 | 12 | 154, 157, 160, 163, 166, 169, 172, 175, 178, 181, 184, 187 |
| 3 | PS\_1375177 | 0.389 | 0.402 | 8 | 200, 203, 206, 224, 227, 233, 236, 239 |
| 4 | PS\_1502048 | 0.477 | 0.533 | 6 | 183, 186, 192, 195, 198, 201 |
| 5 | PS\_31489 | 0.471 | 0.490 | 2 | 186, 189 |
| 6 | PS\_25981 | 0.413 | 0.442 | 3 | 170, 174, 178 |
| 7 | PS\_39709 | 0.495 | 0.566 | 4 | 198, 202, 214, 226 |
| 8 | PS\_718958 | 0.095 | 0.097 | 4 | 189, 195, 204, 207 |

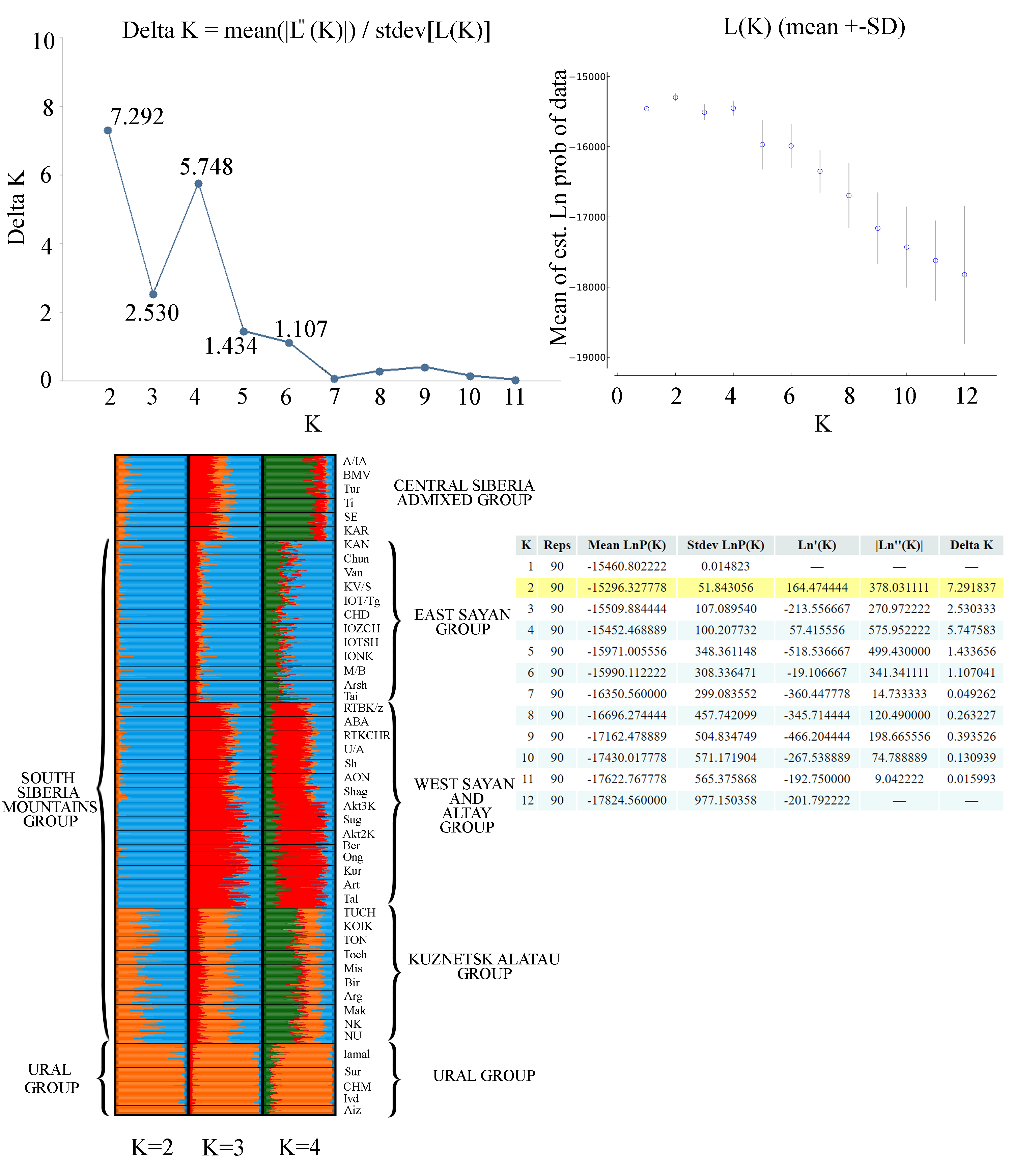
*HO*: observed heterozygosity, *HE*: expected heterozygosity, and *NA*: the allele number

**Appendix S5**



**Fig. S5.6** The NJ tree based on *DA* distances derived from the *P. sibirica* microsatellite data

**Appendix S6**



**Fig. S6.7** Genetic structure of *P. sibirica* populations as inferred by STRUCTURE

**Appendix S7**

The reference list of paleorecords cited in **Fig. 3** (in the main text)

Unlabeled dots refer to Binney et al. (2009, 2017). Unlabeled black dots in the MIS 3 map refer to Krivonogov (1988).

1 – Kobe F, Leipe C, Shchetnikov AA, Hoelzmann P, Gliwa J, Olschewski P, Goslar T, Wagner M, Bezrukova EV, Tarasov PE (2022) Not herbs and forbs alone: Pollen‐based evidence for the presence of boreal trees and shrubs in Cis‐Baikal (Eastern Siberia) derived from the Last Glacial Maximum sediment of Lake Ochaul. Journal of Quaternary Science 37(5):868–883. <https://doi.org/10.1002/jqs.3290>

2-3 – Panova NK, Antipina TG, Zaretskaya NE (2008) New data on palynology, geochronology and stratigraphy of lacustrine-marsh deposits in the Middle Urals. Palynology: Stratigraphy and geoecology. Scientific papers of the XII Russian Palynological conference 2:188–194

4 – Danukalova G, Kurmanov R, Yakovlev A, Osipova E, Zinovyev E, Arslanov K (2016) Palaeoenvironment of the middle and upper Neopleistocene at the Gornovo Upper Palaeolithic site (Southern Ural foreland, Russia). Quaternary International 420:24–46. <https://doi.org/10.1016/j.quaint.2015.08.049>

5 – Baryshnikov GYa (1995) Fossil floras of the transitional zones of the Altai. Altai Republic Flora's Journal 1:111–115

Baryshnikov GYa (1996) Fossil vegetation in terraced complexes of the Altai. Proceedings of the South Siberian Botanical Garden 129–135

6-7 – Gurina AA, Dudko RYu, Tshernyshev SE, Zinovyev EV, Legalov AA (2019) Late Pleistocene insects from the Dubrovino site at Ob River (West Siberia, Russia) and their paleoenvironmental significance. Palaeontologia Electronica 22.1.3A:1–18. <https://doi.org/10.26879/914>

8 – Gurina AA, Dudko RYu, Legalov AA, Zinovyev EV (2021) Late Pleistocene refugium of forest beetles (Coleoptera) in the Ob river region of Novosibirskaya oblast. Paleontology, Stratigraphy and Paleogeography of the Mesozoic and Cenozoic in Boreal Regions: Proceedings of the Online Scientific Session. <https://doi.org/10.18303/B978-5-4262-0104-0-252>

9 – Zinovyev EV, Borodin AV, Kotov AA, Korkin SE (2019) Palaeoenvironment of MIS 5 in the north of Western Siberia, reconstructed on the sub-fossil insect, crustacean and plant macrofossil data. Quaternary International 534:171‒182. <https://doi.org/10.1016/j.quaint.2019.03.030>

Laukhin SA, Shilova GN (2007) Palynological characteristic of Karganian deposits in Surgut Priobye (Late Pleistocene of West-Siberian Plain). Geologija 58:25–33

10 – Zinovyev EV (2020) Insect fauna of the Urals and the Western Siberian Plain in the Quaternary period. Dissertation, Institute of Plant and Animal Ecology, Ural Branch of the Russian Academy of Sciences

11 – Zinoviev EV (2003) Late Karginian entomocomplexes of low reaches of Irtysh River

as exemplified by sites at Skorodum-95 and Kazakovka-95. Euroasian Entomological Journal 2(2):83–93

12 – Dudko RYu, Danukalova GA, Gurina AA, Ivanov AV, Mikhailov YE, Osipova EM, Prosvirov AS, Solodovnikov AYu, Legalov AA, Zinovyev EV (2022) Insects and molluscs of the Late Pleistocene at the Gornovo site (Southern Ural foreland, Russia): New data on palaeoenvironment reconstructions. Quaternary International 632:154–177.

<https://doi.org/10.1016/j.quaint.2021.10.003>

13 – Shchetnikov AA, Bezrukova EV, Maksimov FE, Kuznetsov VYu, Filinov IA (2016) Environmental and climate reconstructions of the Fore-Baikal area during MIS 5-1: Multiproxy record from terrestrial sediments of the Ust-Oda section (Siberia, Russia). Journal of Asian Earth Sciences 129:220–230. <https://doi.org/10.1016/j.jseaes.2016.08.015>

14 – Laukhin S, Shilova G (2005) Results of palynological analysis of Bedoba section

(Late Pleistocene of Middle Siberia). Geologija 49:40–47

15 – Astakhov VI, Mangerud J (2005) The age of the Karginsky Interglacial strata on the Lower Yenisei. Doklady Earth Sciences 403:673–676. <https://people.uib.no/ngljm/PDF_files/Astakhov%20and%20Mangerud%202005,%20Karginsky.pdf>

16 – Laukhin SA, Arslanov KhA, Shilova GN, Velichkevich FYu, Maksimov FE, Kuznetsov VYu, Chernov SB, Tertychnaya TV (2006) Paleoclimates and chronology of the Middle Würm Megainterstadial on the West Siberian Plain. Geography 411A:1457–1461. <https://doi.org/10.1134/S1028334X06090273>

17 – Laukhin SA, Arslanov KhA, Maksimov FE, Kuznetsov VYu, Shilova GN, Velichkevich FYu, Chernov SB, Nikonorov YaA (2008) New outcrop of buried Kazantsevo Peat at lower reaches of the Irtysh River. Doklady Earth Sciences 419:200–204. <https://doi.org/10.1134/S1028334X08020025>

18 – Zinovyev EV (2012) Entomological characteristics of Kazantsev (Eemian) interglacial deposits of Late Pleistocene from Karymkarsky Sor (low reaches of Ob) location in West Siberia, Russia. Euroasian Entomological Journal 11(5):401–409

**References**

Binney HA, Willis KJ, Edwards ME, Bhagwat SA, Anderson PM, Andreev AA, Blaauw M, Damblon F, Haesaerts P, Kienast F, Kremenetski KV, Krivonogov SK, Lozhkin AV, MacDonald GM, Novenko EY, Oksanen P, Sapelko TV, Valiranta M, Vazhenina L (2009) The distribution of late-Quaternary woody taxa in northern Eurasia: Evidence from a new macrofossil database. Quaternary Science Reviews 28(23–24):2445–2464. <https://doi.org/10.1016/j.quascirev.2009.04.016>

Binney H, Edwards M, Macias-Fauria M, Lozhkin A, Anderson P, Kaplan JO, Andreev A, Bezrukova E, Blyakharchuk T, Jankovska V, Khazina I, Krivonogov S, Kremenetski K, Nield J, Novenko E, Ryabogina N, Solovieva N, Willis K, Zernitskaya V (2017) Vegetation of Eurasia from the last glacial maximum to present: Key biogeographic patterns. Quaternary Science Reviews 157:80–97. <https://doi.org/10.1016/j.quascirev.2016.11.022>

Krivonogov SK (1988) Stratigraphy and paleogeography of the Lower Irtysh river in the last glaciation (by carpological data). Nauka Press, Novosibirsk. <https://www.geokniga.org/bookfiles/geokniga-stratigrafiya-i-paleogeografiya-nizhnego-priirtyshya-v-epohu-poslednego-oledene.pdf>

Neale DB, Wegrzyn JL, Stevens KA, Zimin AV, Puiu D, Crepeau MW et al (2014) Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. Genome Biology 15(R59):1‒13. <https://doi.org/10.1186/gb-2014-15-3-r59>

Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin Y-C, Scofield DG et al (2013) The Norway spruce genome sequence and conifer genome evolution. Nature 497:579–584. <https://doi.org/10.1038/nature12211>

Phillips SJ (2017) A brief tutorial on Maxent. <http://biodiversityinformatics.amnh.org/open_source/maxent/>

Putintseva YA, Bondar EI, Simonov EP et al (2020) Siberian larch (*Larix sibirica* Ledeb.) mitochondrial genome assembled using both short and long nucleotide sequence reads is currently the largest known mitogenome. BMC Genomics 21(654). <https://doi.org/10.1186/s12864-020-07061-4>

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: Misener S and Krawetz SA (Eds.) Bioinformatics Methods and Protocols. Methods in Molecular Biology. Humana Press, Totowa, pp 365‒386. <https://doi.org/10.1385/1-59259-192-2:365>

Stevens KA, Wegrzyn JL, Zimin A, Puiu D, Crepeau M, Cardeno C et al (2016) Sequence of the sugar pine megagenome. Genetics 204(4):1613–1626. <https://doi.org/10.1534/genetics.116.193227>

Sullivan AR, Eldfjell Y, Schiffthaler B, Delhomme N, Asp T, Hebelstrup KH, Keech O, Öberg L, Møller IM, Arvestad L, Street NR, Wang X-R (2020) The mitogenome of Norway spruce and a reappraisal of mitochondrial recombination in plants. Genome Biology and Evolution 12(1):3586–3598. <https://doi.org/10.1093/gbe/evz263>

Zonneveld BJM (2012) Conifer genome sizes of 172 species, covering 64 of 67 genera, range from 8 to 72 picogram. Nordic Journal of Botany 30(4):490‒502. <https://doi.org/10.1111/j.1756-1051.2012.01516.x>