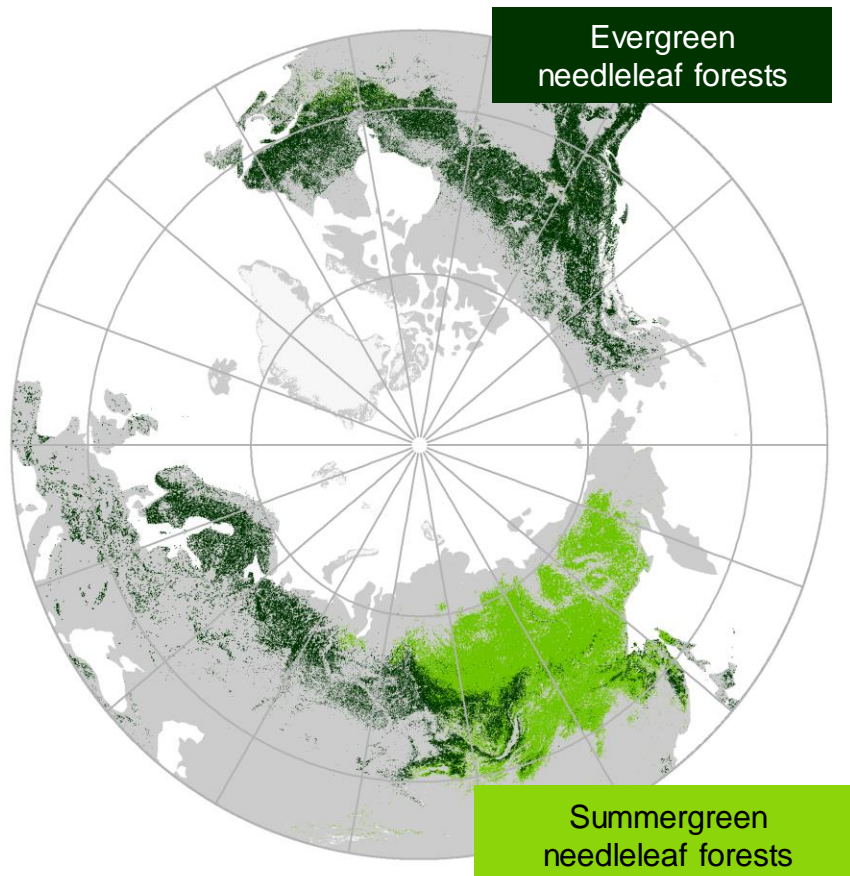S. Kruse

Dahurian Larch (Larix gmelinii)

Hybridization capture of larch (*Larix* Mill.) chloroplast genomes from sedimentary ancient DNA reveals past changes of Siberian forest

Luise Schulte[1,2], Nadine Bernhardt[1]*, Kathleen Stoof-Leichsenring[1], Heike H. Zimmermann[1], Luidmila A. Pestryakova[3], Laura S. Epp[1], Ulrike Herzschuh[1,2,4]

1: Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar und Meeresforschung, Forschungsstelle Potsdam, Potsdam, Germany, 2: Institut für Biochemie and Biologie, Universität Potsdam, Potsdam, Germany, 3: Institute of Natural Sciences, North-Eastern Federal University of Yakutsk, Yakutsk, Russia, 4: Institut für Geowissenschaften, Universität Potsdam, Potsdam, Germany

HELMHOLTZ

Contact: luise.schulte@awi.de

Background: Asian Russia is dominated by Larch Forest

Evergreen needleleaf forests

Summergreen needleleaf forests

Fig. 1 Borel forests on northern hemisphere, adapted from: *Herzschuh et al., Nature Communications, 2016*

- Lach forest cover vast areas and provide critical ecosystem services
- Understanding past population dynamics is important for predictions of larch to a changing climate

Aim:
Learn from genetic information of ancient *Larix* population out of DNA isolated from lake sediment cores

erc   Project: Glacial*Legacy*

HELMHOLTZ

Two sequencing approaches

Shotgun data set

depth /age

algae

terrestrial plants

fungi

animals
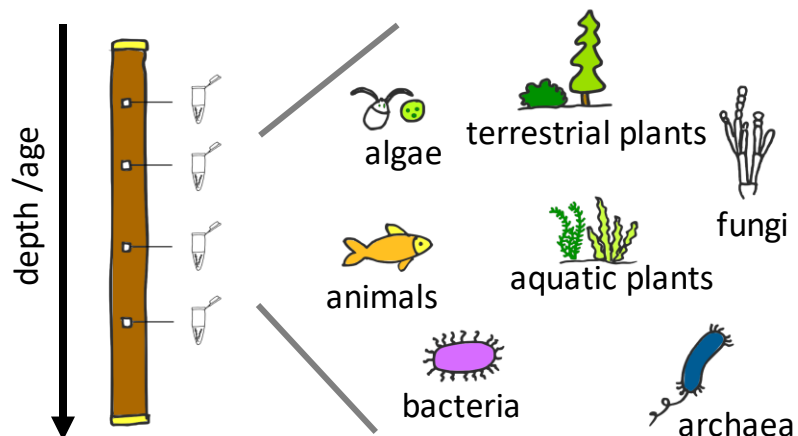
aquatic plants

bacteria

archaea

Fig. 2

Shotgun sequencing of entire sedimentary DNA
- Single stranded library preparation protocol following Gansauge et al. 2017
- DNA from all organism groups living in and around the lake
- Majority of DNA can not be assigned to species level
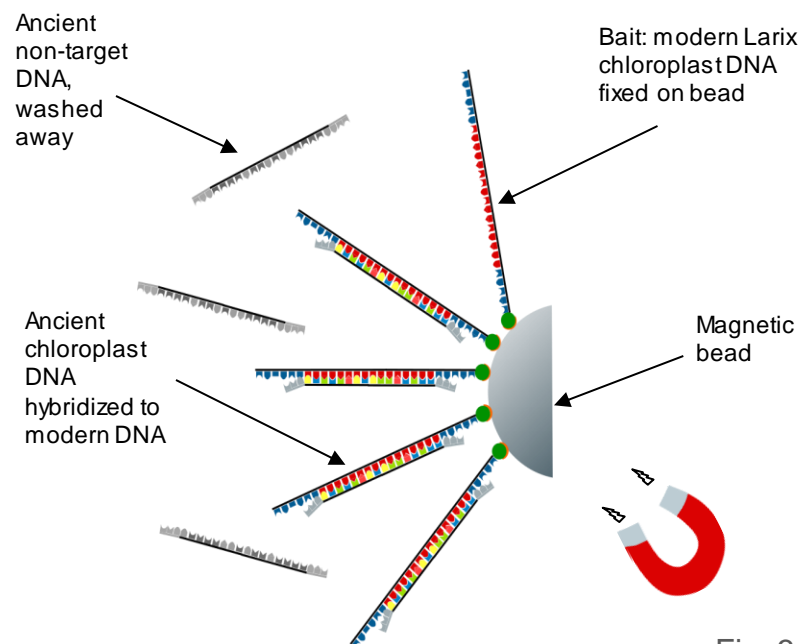
Capture data set

Ancient non-target DNA, washed away

Bait: modern Larix chloroplast DNA fixed on bead

Magnetic bead

Ancient chloroplast DNA hybridized to modern DNA

Fig. 3

Enrichment for larch (*Larix*) chloroplast DNA via hybridization **capture**
- Cost effective baits using self-made long-range PCR products spanning the complete chloroplast genome
- Less sequencing effort with more DNA of target species

HELMHOLTZ

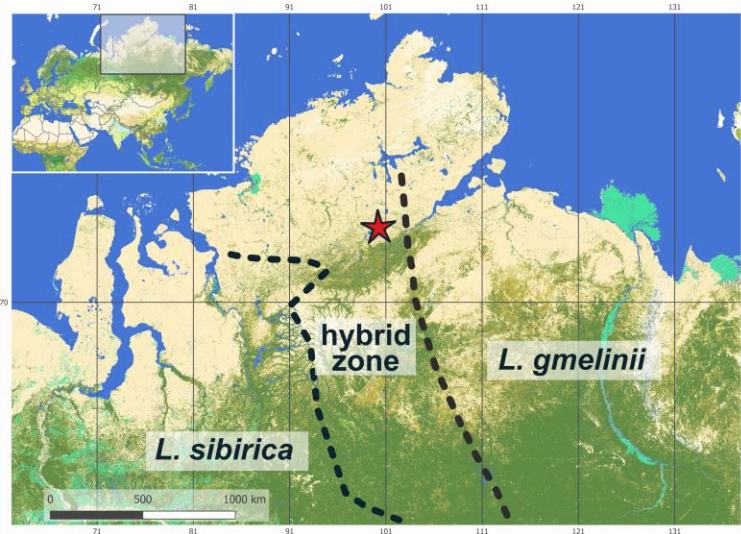# Study Site & Previous Studies

Fig. 4 Study site and species delineations (acc. to Abaimov et al 2010), Map done with © ESA 2010 and UCLouvain

Mitochondrial marker results:

○ type "*Larix sibirica*"

● type "*Larix gmelinii*"
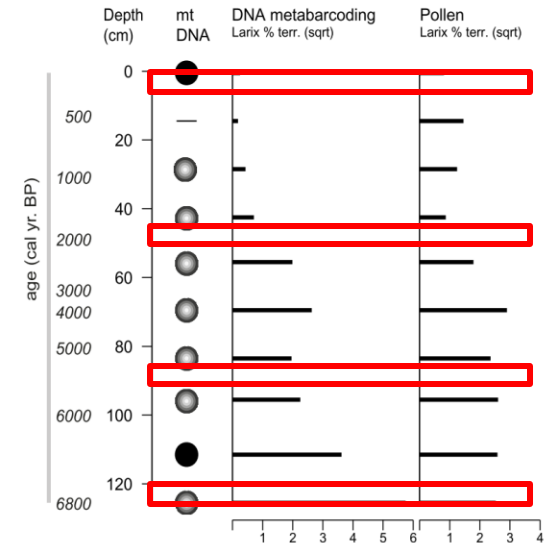
◉ both types

New samples for current study:

Fig. 5 Previous results adapted from Epp et al. 2018, Scientific Reports

- Pilot study on 4 samples from short core form Taymyr peninsula dating back 6,800 years BP
- Core previously studied using pollen, DNA metabarcoding and one mitochondrial marker
- Known facts:
  - vegetation turn-over from forest to single-tree tundra in the last 7000y
  - core lies in hybrid zone of two *Larix* species *L. gmelinii* and *L. sibirica*
  - mixed signal of both *Larix* species with mitochondrial marker

Proof of concept:
- Can we capture very sparse target DNA out of complex mixtures like ancient sedimentary DNA?
- How much of the chloroplast genome can we recover?

Results: Hybridization capture enriches eukaryotic DNA
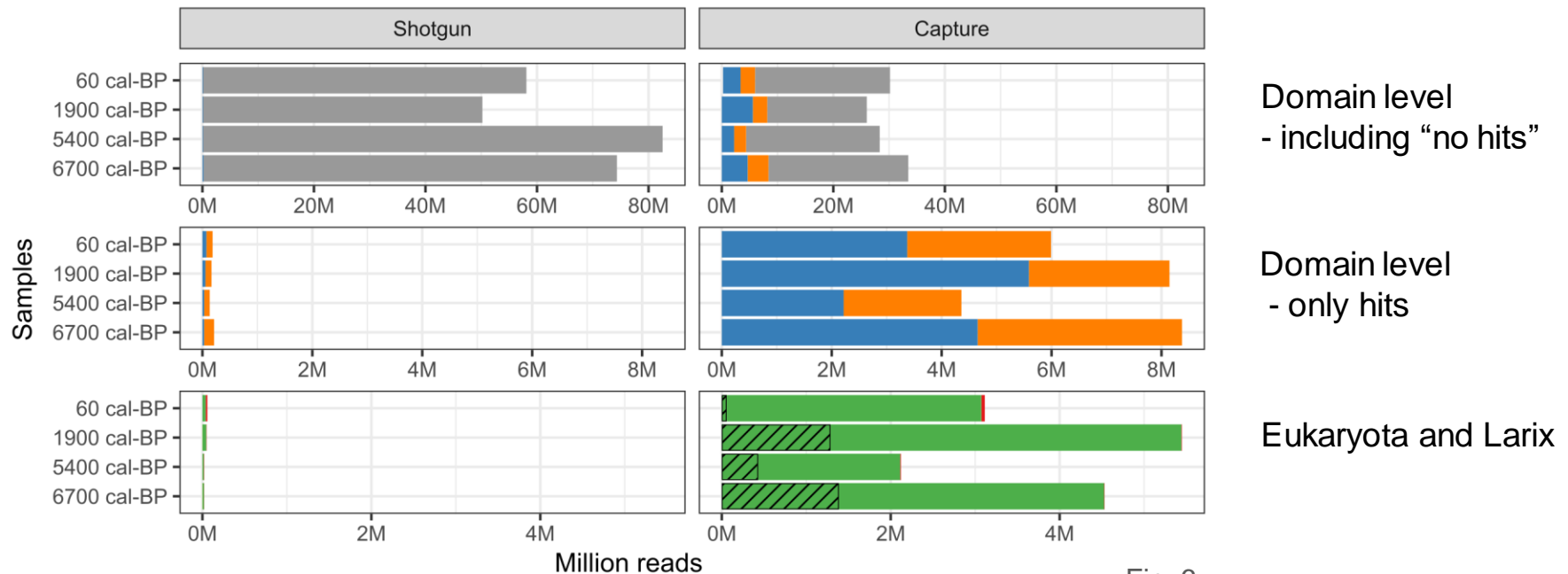

Fig. 6

Sequence reads classified with kraken2 against NCBI non-redundant nucleotide database (nt) at high confidence (--confidence 0.9)

- Viridiplantae enrichment ranged from 77.8 to 236.9-fold enrichment of captured data in respect to shogun data.
- The number of Larix-classified reads per sample corresponds to an increase of around 800 to 1160-fold compared to the shotgun data

Results: SNP analysis resolves species level

Fig. 7
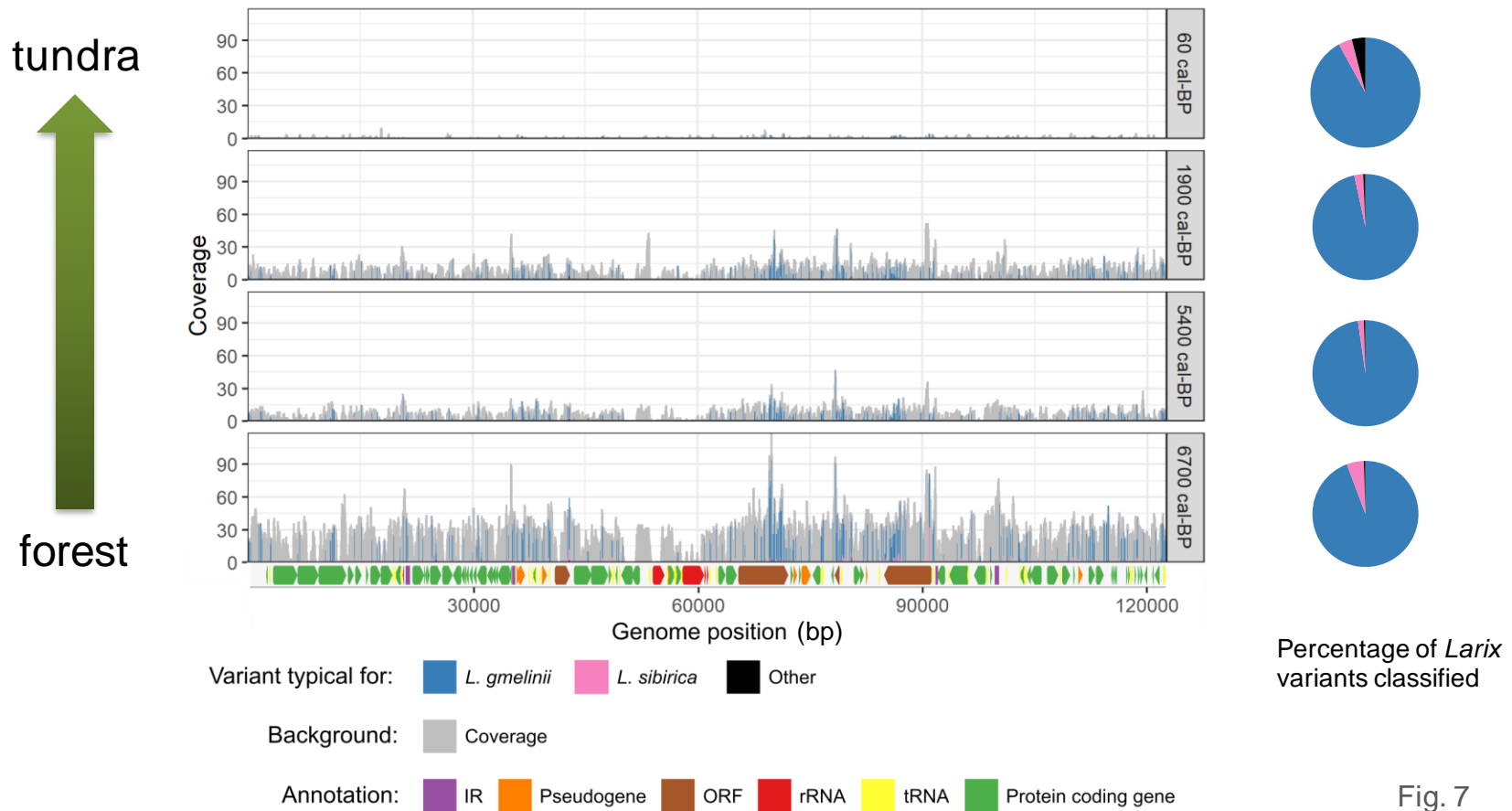
- Alignment of *Larix* classified reads against *L.* chloroplast (cp) genome (capture dataset)
- Coverage per position in grey --> declining coverage corresponds to vegetation turn-over
- 294 positions on cp-genome differing between two species *L. gmelinii* and *L. sibirica*
- Colour indicates which variant was found in ancient reads --> more *L. gmelinii* than *L. sibirica*

HELMHOLTZ

Results: Recovery of near-complete chloroplast genomes

Gaps in alignments
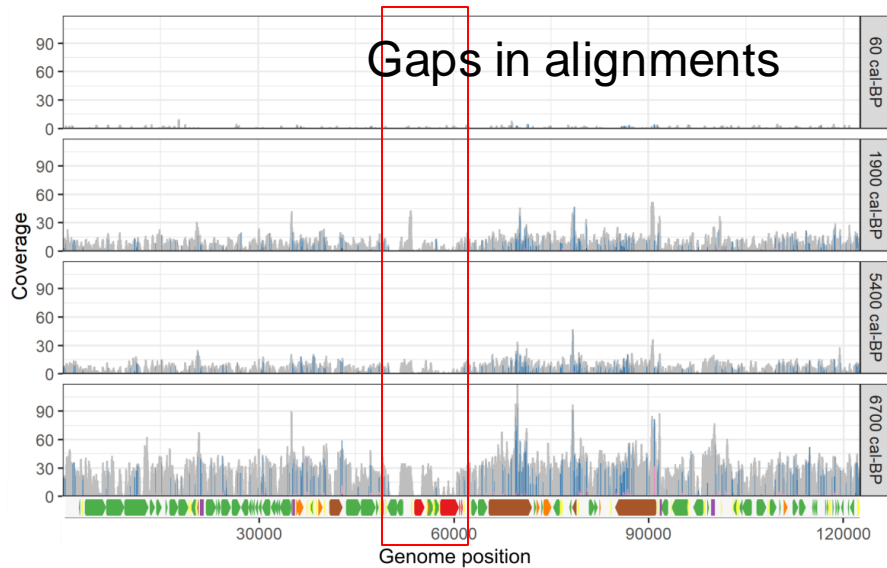
Fig. 8

Variant typical for: ▮ *L. gmelinii* ▮ *L. sibirica* ▮ Other

Background: ▮ Coverage

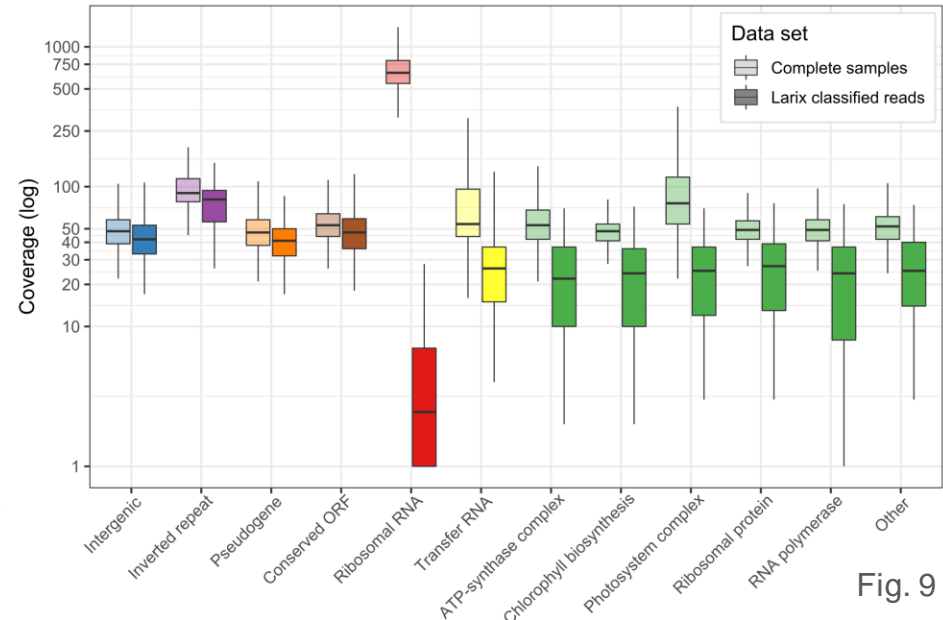Annotation: ▮ IR ▮ Pseudogene ▮ ORF ▮ rRNA ▮ tRNA ▮ Protein coding gene

Fig. 9

Coverage of alignment according to functional annotation: Comparison between full capture data set and only the reads assigned exclusively to *Larix*

- Alignment has gaps even in well-covered oldest "forest" sample

- Reads covering well conserved regions as ribosomal RNA and coding genes cannot be exclusively assigend to *Larix*
- Complete *Larix* chloroplast information most probably conserved in ancient samples

HELMHOLTZ

Summary and outlook

- Hybridization capture with long-range PCR products efficient method to target enrich ancient sedimentary DNA

- Near-complete *Larix* chloroplast genomes can be recovered

- Whole chloroplast genome analysis allows high resolution of species

Next steps

- Apply method on cores spanning last glacial maximum

- Apply method on spatially distributed cores

- Capture mitochondrial sequences along with chloroplast

Contact: luise.schulte@awi.de

HELMHOLTZ