

Hybridization capture of Larix candidate adaptive genes from sedimentary ancient DNA

Stefano Meucci¹, Luise Schulte¹, Kathleen R. Stoof-Leichsenring¹, Stefan Kruse¹, Konstantin Krutovsky², and Ulrike Herzschuh¹

¹Alfred Wegener Institute, Polar Terrestrial Environmental Systems, Potsdam, Germany

²Department of Forest Genetics and Forest Tree Breeding, Georg-August-University of Göttingen, Büsgenweg 2, 37077 Göttingen, Germany

Introduction

Study area: Siberia



Aim

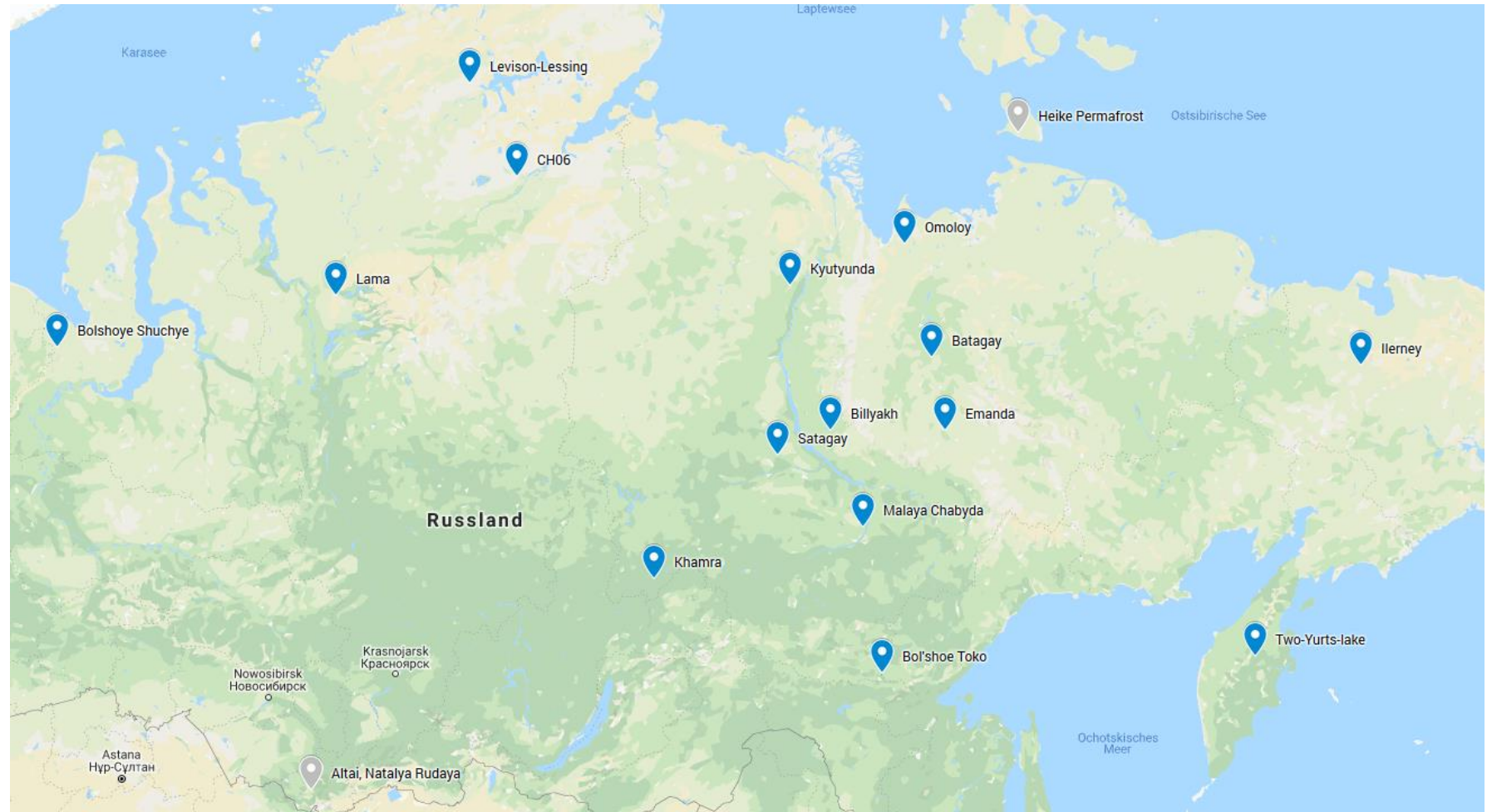
To retrieve sedimentary ancient DNA sequences of *Larix* candidate adaptive genes in order to identify possible genetic variation in space and time.

Data source

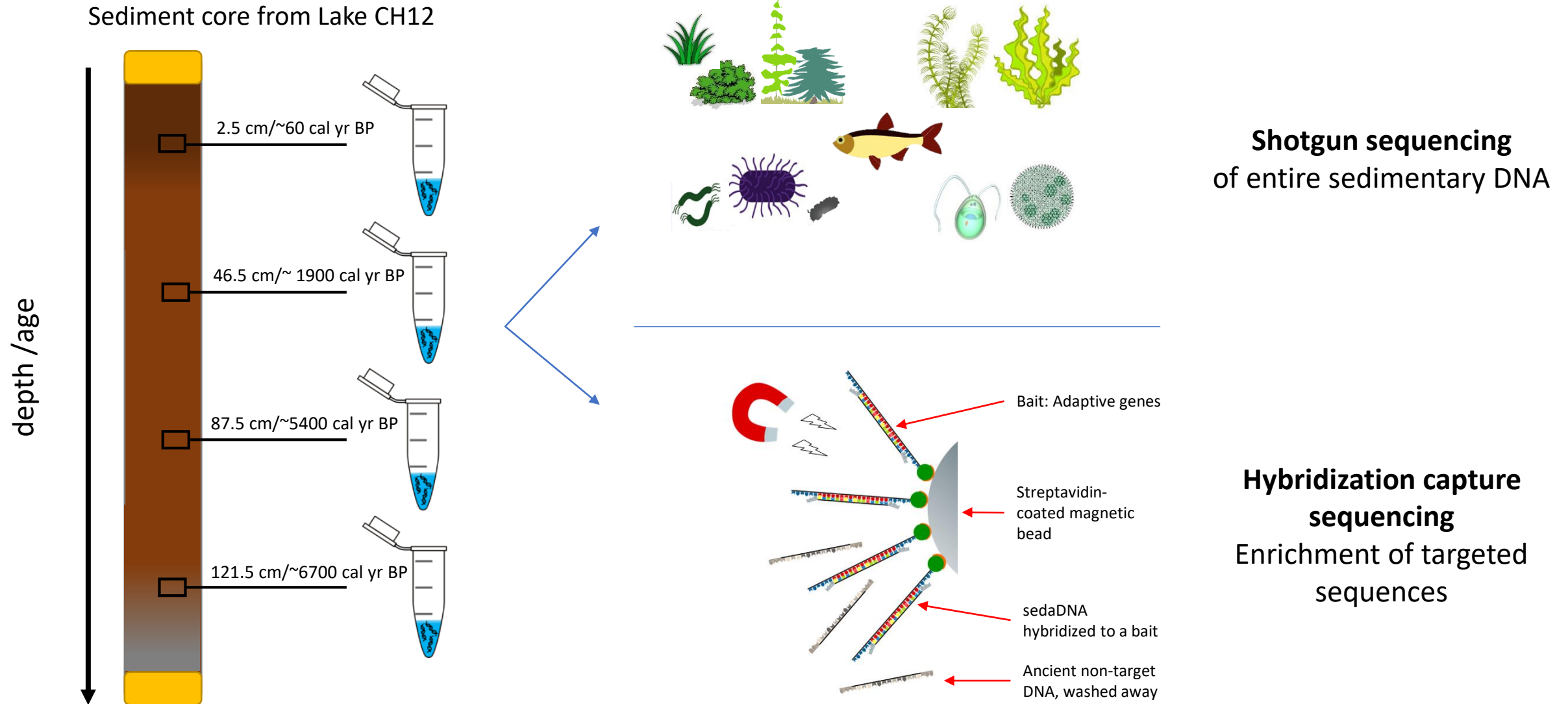
Sedimentary ancient DNA (sedaDNA) retrieved from 8 lake sediment cores across Siberia

Geographic position of the Lakes

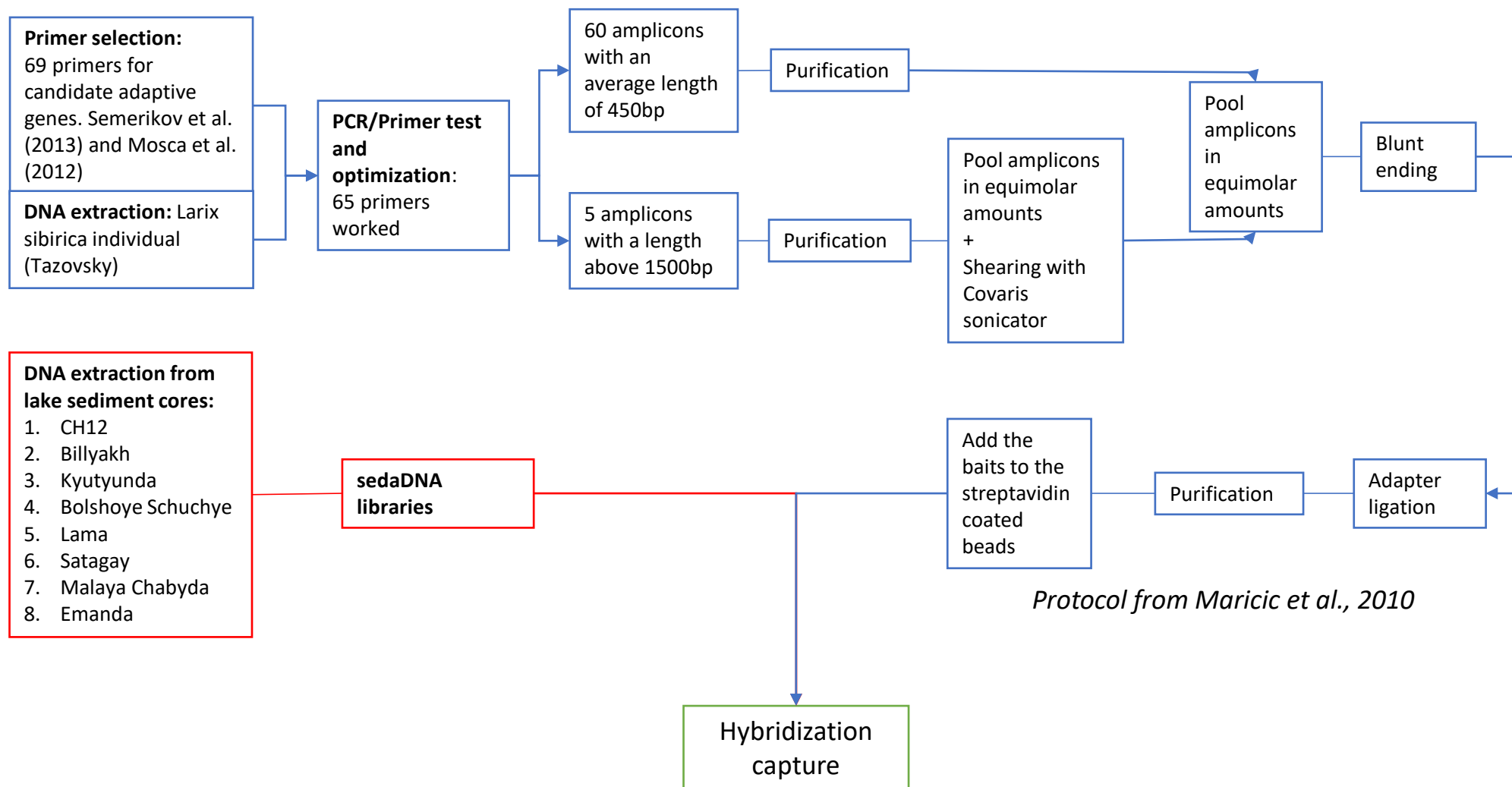
- CH12
- Billyakh
- Kyutyunda
- Bolshoye Schuchye
- Lama
- Satagay
- Malaya Chabyda
- Emanda



Method: Hybridization capture vs Shotgun

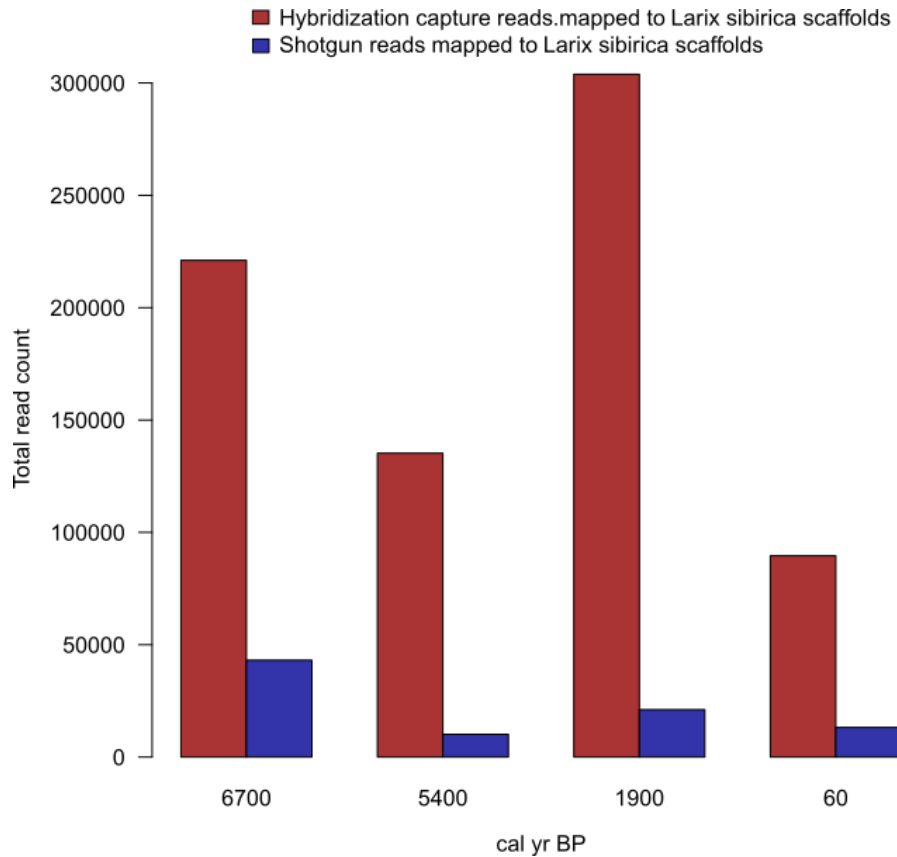
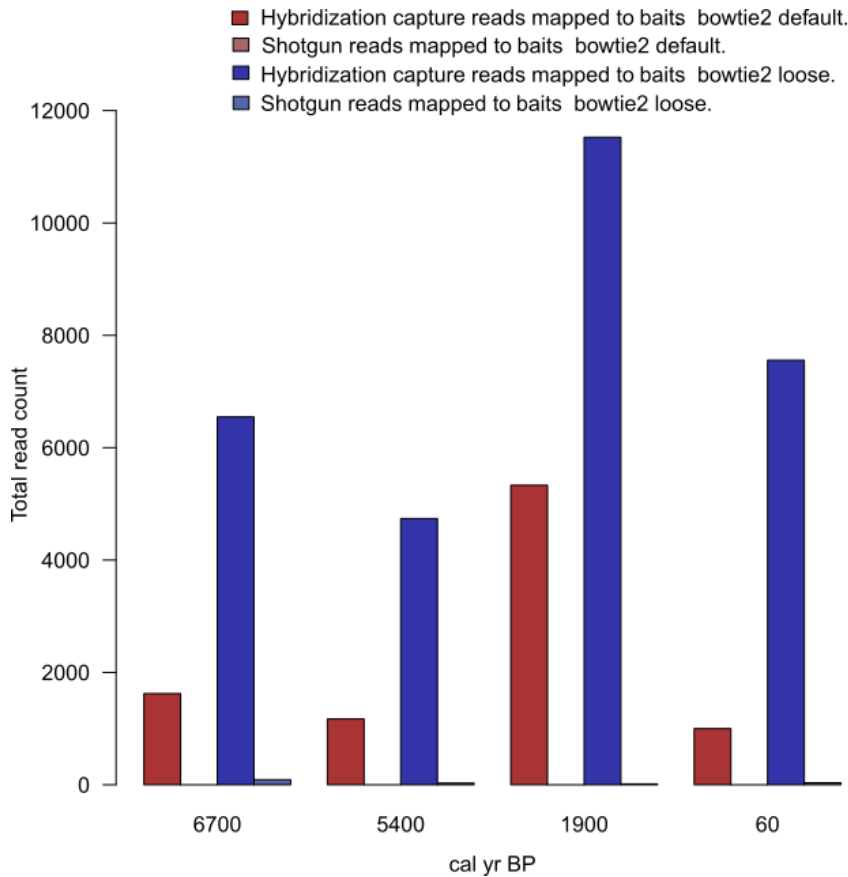


Workflow

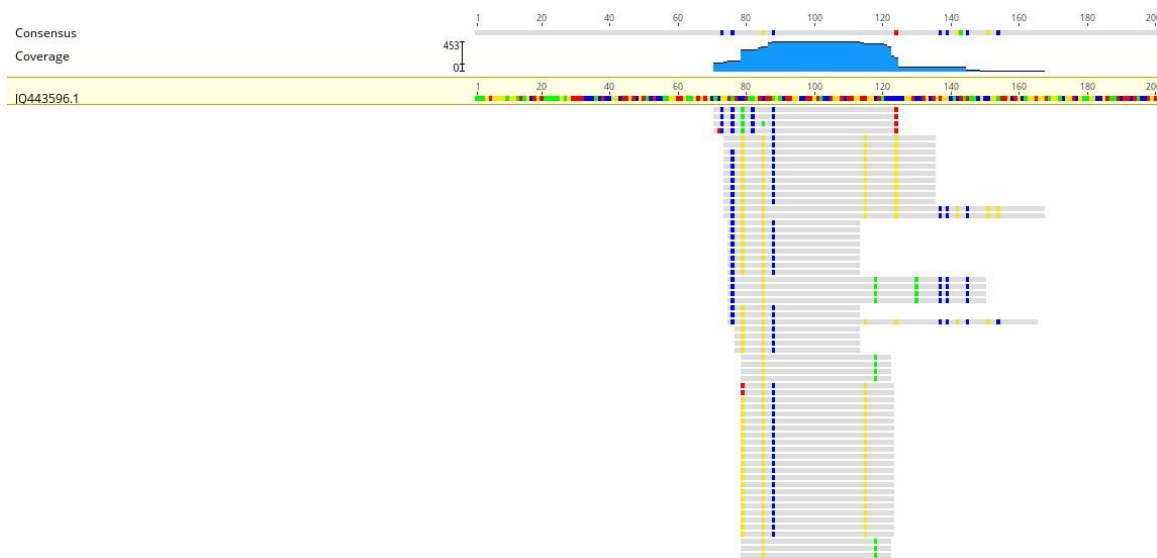


Hybridization capture vs Shotgun

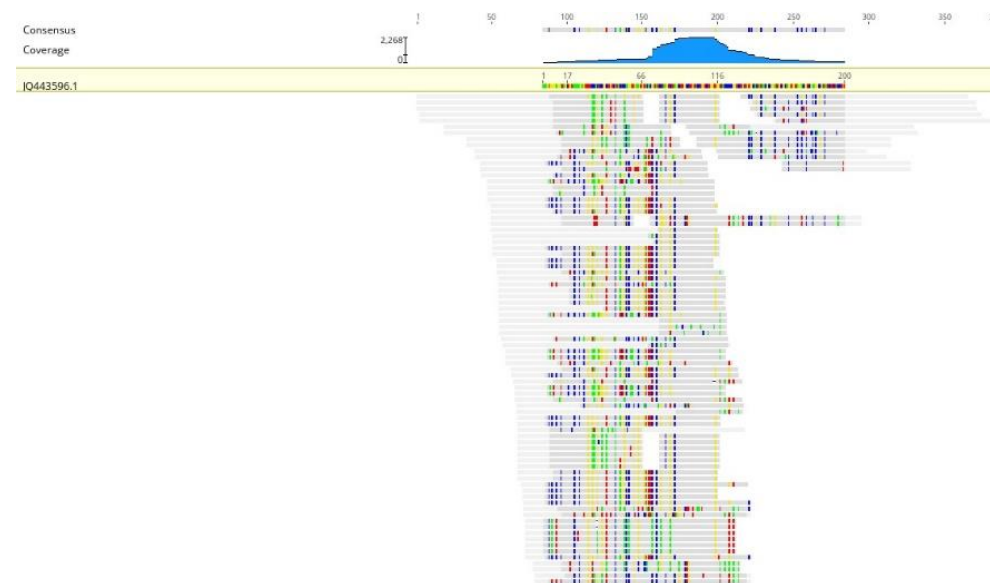
The total count of mapped reads to our targeted reference sequences show that it was possible to retrieve our extremely rare targets with a very efficient enrichment in the capture samples.



Mapping of captured reads to 1 of the 65 target sequences



Default mapping

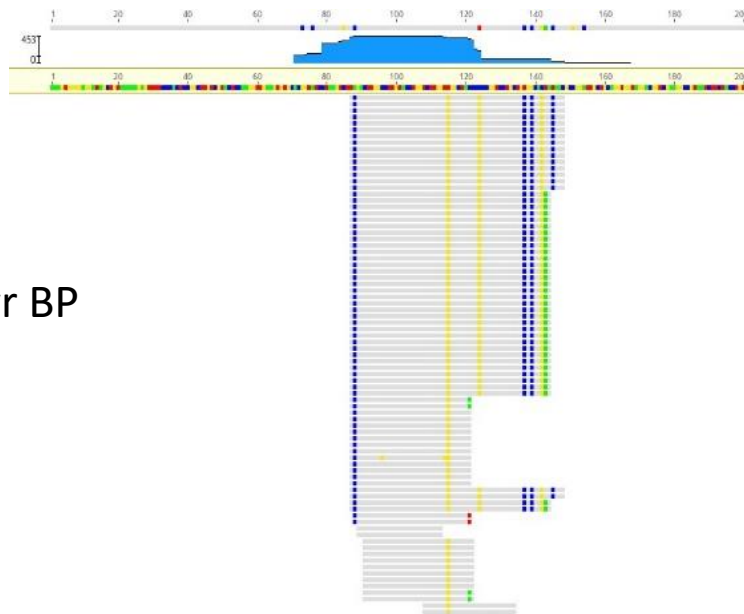


Loose mapping

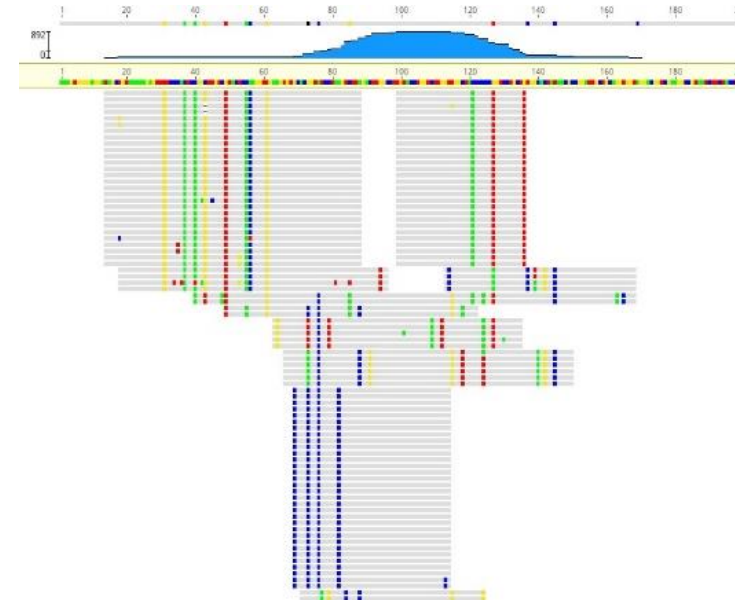
Increase of unspecific sequences

Results

Captured reads mapped to the reference of one targeted gene in the oldest and youngest core samples from lake CH12



6700 cal yr BP



60 cal yr BP

Next step: Evaluation of genetic variation in space and time

Conclusions and next steps



- For the first time, it was possible to retrieve extremely rare *Larix* nuclear genome sequences (candidate adaptive genes) from sedaDNA.
- This was possible thanks to the very efficient enrichment obtained by our hybridization capture system. Indeed, the shotgun method performed on the same samples was not able to retrieve our targeted sequences.
- About 60% of the baits captured only 1 to 5 *Larix* specific sedaDNA fragments -> extremely rare in targeted sequences in sedaDNA.
- There are conserved regions between our baits and other regions of the *Larix* genome as well as in other genomes -> unspecific capture
- The high amount of reads mapping on *Larix sibirica* scaffolds might be due to unspecific fragments from *Larix* nc which become baits for further unspecific capture.
- A high duplication rate (~30%) might occur due to the high number of amplification cycles used for a double hybridization capture.