ddRADseq library design

J. Ignacio Lucas Lledó

Last update: September 30, 2021

This is based on Salas-Lizana and Oono [2018]. *SphI* recognizes the sequence GCATGC and produces CATG-3' overhangs. And *HindIII* recognizes AAGCTT and produces 5'-AGCT overhangs. Below I show the design for the adapters.

HindIII-P1 adaptor

```
*AGCTCTGTCTCTTATACGAGAACAA
     ||||||||||||
     GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG
```

SphI-P2 adaptor

```
  GTCGGCAGCGTCAGATGTGTATAAGAGACAGCCATG
  ||||||||||||||||||||||||||||||||||
GACGTCTCTTATACACATCTGACGCTGCCGACGA*
```

Ligated fragment

Note that once the adapters are ligated to the genomic fragment (blue), the restriction sites are not available any more, and the fragment would not be digested again.

```
  GTCGGCAGCGTCAGATGTGTATAAGAGACAGCCATGCNNNNNNNNNNAAGCTCTGTCTCTTATACGAGAACAA
  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GACGTCTCTTATACACATCTGACGCTGCCGACGAGCTTNNNNNNNNNNGCATGGCTGTCTCTTATACACATCTGACGCTGCCGACGA
```

Ligated fragment with amplification primers

We use Nextera Index Kit v2 primers to incorporate unique combinations of adapters to every sample. Below, the positions marked as *[i5]* and *[i7]* correspond to the 10-bases long indices. Note these are not Unique Dual Indices, because every individual primer is shared with other samples, but the combination is

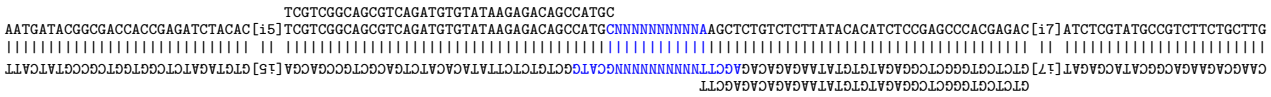
sample-specific. In order to reduce the chance of index hopping, we include a clean-up step before and after the PCR amplification that incorporates the indices. In addition, we include we reserved two pairs of primers to label two samples with unique individual indices, which will allow us to quantify the rate of index hopping.

Amplified fragment

After ligation of adapters and incorporation of indices by amplification PCR, the original DNA fragments will be almost 150 bases longer. In any case, because both enzymes are rare cutters, we will select a wide range of fragment sizes to sequence, say between 300 and 900.

Sequencing primers

The sequencing primers only need to be edited to match the restriction site. This makes the sequence read start at the unknown genomic base, to optimize the yield and increase the complexity of the signal.



References

Rodolfo Salas-Lizana and Ryoko Oono. Double-digest radseq loci using standard illumina indexes improve deep and shallow phylogenetic resolution of lophodermium, a widespread fungal endophyte of pine needles. *Ecology and Evolution*, 8 (13):6638–6651, 2018. doi: 10.1002/ece3.4147.