0501-poplar

* [Computational analysis:](#computational-analysis)

## Computational analysis:

DAP-Seq: Mapping, quantification and peak calling

Adapters were trimmed off from raw reads with “trim\_galore” (!CITE!TRIM\_GALORE). Raw reads were mapped to the genome "sPta717-v1-1" (downloaded from "http://aspendb.uga.edu/index.php/databases/spta-717-genome") with Bowtie2 ((*1*)) under argument:"--no-mixed --no-discordant --no-unal -k2". Duplicate reads were removed with Picard using default setting (!CITE!PICARD). The resultant alignment is used for downstream quantification and visualisation

Genomic binding profile was quantified in RPKM (Reads Per Kilobase per Million mapped reads) using a bin-size of 10bp. For each bin, \text{RPKM}_{\text{bin}}=\frac{\# \text{Reads covering bin}} { \text{bin-size} } \cdot \frac{10^6}{ \#\text{Mapped reads}}

For each treated DAP-Seq library, peaks were called against a control library using MACS2 (!CITE!MACS2) with argument "--keep-dup 1 -p 0.1". Peaks from each DAP-Seq were further fitlered for FC>3.0.

Any gene with a peak within +/- 6000bp of its start codon is considered a target gene. Best arabidopsis homologue is reported for each target gene, using annotation file "Ptrichocarpa\_444\_v3.1.annotation\_info.txt" (bulk downloaded from Phytozome 12 https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org\_Ptrichocarpa\_er).

Availability and external packages

Post-processing code is available as tarball as attached in the supplementary. It depends on python2 packages [pymisca](https://github.com/shouldsee/pymisca) and [synotil](https://github.com/shouldsee/synotil).

DAP-Seq data are available from: [GEO\_ACCESSION]

GNU-parallel (!CITE!GNUPARA) was used in paralleling the computational analysis.

Bedtools (!CITE!BEDTOOLS) was used for intersection of peaks, making genomic windows.

1. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nature methods*. **9**, 357–359 (2012).