

Abundant Satellite DNA Differentiates Aspens from Other Closely Related Poplars

Background

- Poplar is an important renewable biofeedstock and a key species in many ecosystems.
- *Populus tremula* × *P. alba* INRA 717-1B4 ('717') is an easily transformed hybrid clone generated at INRA-e, France in the 60's. 717 is widely used in functional genomics and transgenic research.
- Its hybrid genome presents a challenge for sequence-sensitive applications like CRISPR due to high heterozygosity.

Approach

- We sequenced 717 with long-read PacBio HiFi and Hi-C technologies for *de novo* genome assembly through collaboration with JGI and performed a comprehensive analysis of its repetitive genomic landscape.

Results

- We generated a haplotype-resolved, telomere-to-telomere assembly of not one but *two* closely related aspen genomes. Among the 38 chromosomes, 11 are free of gaps. The majority of the chromosomes has only one gap.
- We discovered an abundant megabase satellite DNA, dubbed M147 (monomer of ~147 bp) present in 16 chromosomes from the *P. alba* subgenome and 12 chromosomes of the *P. tremula* subgenome.
- The M147 tandem repeat arrays lie outside of putative centromeres, are much more abundant than known telomeric or centromeric satellite repeats and have the cytogenetic appearance of heterochromatic knobs.
- The M147 satellite DNA is found exclusively in aspens and not in other closely related poplars and cottonwoods.

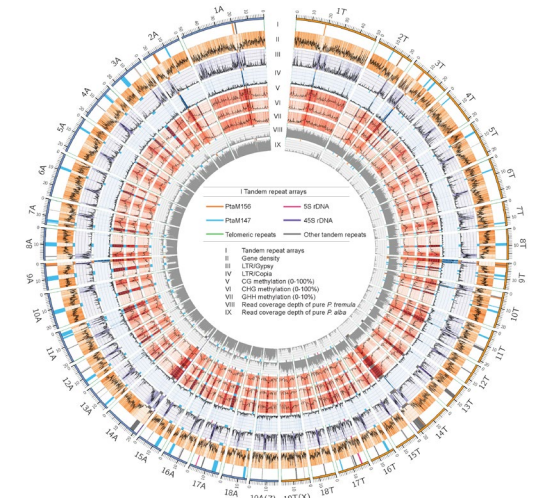
Significance

- The work offers a molecular basis for future hypothesis-testing of PtaM147 heterochromatic knobs as reproductive barriers between aspens and closely related cottonwoods and poplars.
- The haplotype genomic resource generated for this genotype (which is the transformation workhorse) has greatly improved the design and analysis of genome editing experiments that are highly sensitive to sequence polymorphisms.

Zhou et al. *Plant Journal* (2023) 116: 1003-1007. doi: 10.1111/tbj.16454



Plant Journal cover (Nov 2023) featuring transgenic *Populus* (left) and an aspen-specific megabase satellite DNA (top right) that forms heterochromatic knobs (lower right).



The circos show the landscape of tandem repeat arrays, genes, retrotransposons, cytosine methylation, and short read coverage over 19 *P. tremula* (T) and 19 *P. alba* (A) chromosomes.