DE GRUYTER
OPEN

Whole-genome draft assembly of *Populus tremula* x *P. alba* clone INRA 717-1B4

Malte Mader[1], Marie-Christine Le Paslier[2], Rémi Bounon[2], Aurélie Bérard[2], Patricia Faivre Rampant[2], Matthias Fladung[1], Jean-Charles Leplé[3,4], and Birgit Kersten[1]

[1] Thünen Institute of Forest Genetics, Grosshansdorf, Germany
[2] EPGV US 1279, INRA, CEA, IG-CNG, Université Paris-Saclay, 91000 Evry, France
[3] AGPF, INRA, 45075, Orléans, France
[4] present address: BIOGECO, INRA, Université Bordeaux, 33610 Cestas, France

Corresponding author: Dr. Birgit Kersten, E-mail: birgit.kersten@thuenen.de

Abstract

Populus trichocarpa and *P. deltoides* are the only *Populus* species known to date to have a publicly available nuclear genome sequence that has been assembled to chromosomes and annotated (https://phytozome.jgi.doe.gov/). Here we focus on the clone INRA 717-1B4, a female *P. tremula x P. alba* (*P. x canescens*) interspecific hybrid that is universally used by scientists worldwide as a tree model in transgenic experiments. The already available INRA 717-1B4 nuclear genomic resource (v1.1 of sPta717 at http://aspendb.uga.edu/index.php/databases/spta-717-genome) presents only INRA 717-1B4 genomic regions with high similarity to the *P. trichocarpa* genomic reference sequences. We assembled draft genomic scaffolds by a combination of *de novo* assembly with reference-based assembly using 30x resequencing NGS data (Illumina MiSeq® and Ion Torrent Ion PGM™) of INRA 717-1B4. In total, 419,969 scaffolds of length larger than 500 bp were generated. The mean length of the scaffolds is 2,166 bp and the size of the largest scaffold 84,573 bp. The N50 contig length is 3,850 bp when considering contigs larger than 1,000 bp. Probably due to the high level of heterozygosity of this interspecific hybrid, the accumulated scaffold length is with 0.9 GB about twice the expected size of the haploid nuclear genome. DNA sequences of the genomic scaffolds of INRA 717-1B4 are publicly available for Blast analyses and download via the new INRA web portal at https://urgi.versailles.inra.fr/Species/Forest-trees/Populus/Clone-INRA-717-1B4/. This new genomic sequence resource will complement the already available INRA 717-1B4 resources and will facilitate the future optimization of genetic transformation experiments to discover gene function.

Keywords: *INRA 717-1B4, P. x canescens, poplars, next generation sequencing, genome assembly, scaffolds, model tree*
Running title: Draft genome sequence of Populus × canescens

Introduction

The genus *Populus* consists of about 29 different species, classified in six different sections (*Populus* (formerly Leuce), *Tacamahaca, Aigeiros, Abaso, Turanga* and *Leucoides*; Eckenwalder, 1996). Most *Populus* species are dioecious, although reports of hermaphroditism have been published (reviewed in Slavov and Zhelev, 2010). Due to its small genome size, rapid growth, easy vegetative propagation, huge genomic resources and easy biotechnological handling ability, this genus has become a model species for forest tree genomics (Wullschleger et al., 2002), resulting in the publication of the full nuclear genome sequence of western black cottonwood, *P. trichocarpa* (section *Tacamahaca*; Tuskan et al., 2006). Very recently the sequence of the nuclear genome of *P. deltoides* WV94 (v2.1), a member of the section *Aigeiros*, was published at Phytozome v12 (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_PdeltoidesWV94er).

Sequences of genomic scaffolds from whole genome *de novo* assemblies for two aspen species (*P. tremula, P. tremuloides*; section *Populus*) and a *P. tremula x tremuloides* hybrid are publicly available at PopGenIE (v3; http://popgenie.org/; Sjödin et al., 2009). One *de novo* draft assembly for *P. nigra* (104,431 contigs with N50 of 6,130 bp) was produced by Pinosio et al. (2016) and corresponds to an assembly length of 339.5 Mbp with a mean contig length of 3,251 bp.
One *P. euphratica* genome assembly (Populus euphratica_1.1) has a size of 496.5 Mb with the contig-N50 of the assembled sequence being 40.4 Kb and scaffold-N50 being 482 Kb (Ma et al., 2013; sequence available at the The Salinity Tolerant Poplar Database: http://lzume.iok.la/stpd.

Only few of the *Populus* species are of high scientific and economic importance, and serve as source material for breeding purposes (Licht and Isebrands, 2005; Walle et al., 2007). Important characteristics of some interspecific hybrids are high growth rates and broad applicability ranging from wood and paper to energy production (Dickmann and Stuart, 1993). Nowadays, breeding of *Populus* is intensified worldwide due to

DOI:10.1515/sg-2016-0019
edited by the Thünen Institute of Forest Genetics

the application of biotechnological methods (Fladung et al., 2012) and the availability of huge genomic resources (Stettler et al., 1996; Stanton et al., 2010). Superior clones of various *Populus* species have been developed and are commonly used for biomass production in short rotation plantations. Few poplar clones are easily accessible for genetic transformation by *Agrobacterium tumefaciens* infection (Jansson and Douglas, 2007).

The poplar clone largely used as model and transformed for the first time by Leple et al. in 1992 is the interspecific hybrid clone INRA 717-1B4, *P. tremula x P. alba* (*P. x canescens*). Many transgenic poplar lines based on this clone have been transferred to the field and tested for commercial application (Walter et al., 2010; Pilate et al., 2012). A search of publications that used the clone INRA 717-1B4 as a model experimental tree species for physiological and genomic studies returned 245 publications (Additional files 1 and 2).

A "substituted genome sequence of *P. tremula x alba* INRA 717-1B4" (sPta717) was created "by introducing genomic variants of INRA 717-1B4 into the *P. trichocarpa* reference genome" using 20x resequencing NGS (next generation sequencing) data (v1.1 of sPta717 at http://aspendb.uga.edu/index.php/databases/spta-717-genome; Zhou et al., 2015). This resource has been applied for screening of 717 gRNAs for sequence variants to optimize gRNA design in CRISPR/Cas genome editing experiments (Zhou et al., 2015).

Unfortunately, the sPta717 sequence represents only genomic regions of INRA 717-1B4 that show a high similarity to the *P. trichocarpa* reference genome. To provide a basis for the future optimization of genetic transformation experiments also in target regions that are very specific for INRA 717-1B4 and not displayed in the sPta717 sequence, we decided to generate and provide draft genomic scaffolds of INRA 717-1B4 by assembly (*de novo* and reference-based) of 30x resequencing NGS data.

Material and Methods

Plant material

The INRA clone 717-1B4 is a female clone, originating from a cross between a female *Populus tremula* tree (tree # 5903) harvested on March 3, 1959 in the Parroy forest (Meurthe Moselle, France) and a male *Populus alba* tree (tree # 6072) harvested on February 22, 1960 close to the Pont du Gard (France). This hybrid was part of a breeding program initiated by Dr. Michel Lemoine in the 1960's (Lemoine, 1973). Trees 5903 and 6072, installed in the nurseries of Nancy, were lost during storms in 1999, and unfortunately vegetative copies of the parents are not available.

Whole genome shot gun sequencing of 717-1B4

Ramets of INRA 717-1B4 were grown in the greenhouse. The youngest leaves were harvested, shock-frozen in liquid nitrogen and stored at -80°C. Frozen leaves were then ground with a pestle and mortar to a fine powder in liquid nitrogen. Total genomic DNA was extracted using the Dolezel-MATAB (mixed alkyltrimethylammonium bromide) method (Ky et al., 2000). Genomic DNA was then prepared for the Illumina MiSeq and Ion Torrent PGM™ sequencing systems. The whole genome sequencing (WGS) DNA sequencing library for Illumina MiSeq sequencing was created using the TruSeq® DNA PCR-Free LT kit (Illumina®). Briefly, sample preparation was performed with 2 µg of DNA using the low sample protocol. The mean fragment size was 550 bp. All enzymatic steps and cleaning steps, including fragmentation using AFA (Adaptive Focused Acoustics™) technology on focused-ultrasonicator E210 (Covaris), were performed according to the manufacturer's instructions. On board clusters generation and 2 × 250 bp paired-end sequencing by synthesis cycles (MiSeq Reagent Kit v2) were performed on a MiSeq (Illumina®) according to the manufacturer's instructions.

A WGS DNA sequencing library of 717-1B4 was prepared with the Ion Xpress™ Plus Fragment Library Kit for NGS sequencing using the Ion Torrent PGM™ sequencing platform (Personal Genome Machine® (PGM™) Sequencer, Life Technologies, USA). Briefly, total genomic DNA (100 ng) was sheared using the Ion Shear™ Plus Reagents and used for preparing the sequencing library according to the Ion Xpress™ Plus gDNA Fragment Library kit (catalog number 4471252) following Ion Torrent PGM™ protocol (Life Technologies, USA). The resulting individual DNA library was quality checked and quantified using the Qubit® 2.0 Fluorometer and the Qubit® dsDNA HS Assay Kit following the manufacturer's specification (Life Technologies, USA). Following template amplification and enrichment on the Ion OneTouch™ 2 System (Ion OneTouch™ 2 Instrument for amplification and Ion OneTouch™ ES enrichment system, Life Technologies, USA) using the Ion PGM™ Template OT2 400 Kit (catalog number 4479878), the sample was loaded onto one PGM Ion 318™ Chip v2 and sequenced using the Ion PGM™ Hi-Q™ Sequencing Kit (catalog number A25592, Life Technologies, USA) according to manufacturer's protocol.

Bioinformatic analyses

If not stated otherwise, all bioinformatic analyses steps were performed using CLC Genomics Workbench (CLC GWB; v7.0.4; CLC bio, a QIAGEN Company, Aarhus, Denmark). We pursued a hybrid assembly strategy integrating MiSeq and Ion Torrent reads (Figure 1).

Read trimming

Initial quality control of the NGS reads (MiSeq or Ion Torrent) was performed with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Based on the FastQC reports all reads were trimmed using the trimming tool of the CLC GWB. All MiSeq short reads containing adapter sequences (*i.e.* Illumina TrueSeq Universal Adapter, Index adapter I and TruSeq control sequences) were trimmed. Reads containing bases with a quality Phred score less than 20 were trimmed by setting the quality limit of the trimming tool to a value of 0.01. Ten nucleotides 3' and 5' of every read were removed and finally every read with a length of less than 100 bp was discarded. To reduce the amount of redundant data all overlapping read pairs were merged and duplicate reads were removed from the data set.

All Ion Torrent reads containing the adapter sequence (3'-ATCA CCGACTGCCCATAGAGAGGCTGAGAC-5') were trimmed. Reads containing bases with a quality Phred score less than 15 were trimmed by setting the quality limit of the trimming tool to a value of 0.03. Furthermore reads with a length less than 75 bp and duplicate reads were discarded.

De novo-assembly of 717-1B4

1- De novo-assembly of MiSeq and Ion Torrent reads

De novo sequence assembly of the MiSeq reads was performed with a preset word size of 29, length fraction of 0.9, similarity fraction of 0.95 and without the scaffolding option (Figure 1). These assembly parameters provided the least misassemblies according to the quality assessment generated by the software QUAST (Gurevich et al., 2013). The minimum and maximum read pair distances were set to 379 bp and 579 bp, respectively (based on the mean fragment size of the 717-1B4 genomic library). All remaining parameters were set to default (with back-mapping option per default).

De novo sequence assembly of the trimmed Ion Torrent reads was performed with a length fraction of 0.9 and a similarity fraction of 0.95. All remaining parameters were set to default.

2- Generation of additional 717-1B4 consensus contigs after mapping MiSeq reads to the P. trichocarpa reference genome

Trimmed MiSeq reads were mapped to all *P. trichocarpa* scaffolds (v3.0; downloaded from Phytozome) using default parameters, but a length fraction of 0.94 and a similarity fraction of 0.98. Non-mapped reads were kept. Based on the mappings, consensus contig sequences were derived using default parameters.

3- Superassembly of MiSeq contigs, Ion Torrent contigs and consensus contigs from mapping

All contigs including MiSeq contigs, Ion Torrent contigs and consensus contigs from mapping were subjected to a superassembly, in combination with the related unmapped reads of the original assemblies and the mapping (default parameters with scaffolding option, but with a length fraction of 0.9 and a similarity fraction of 0.95; Figure 1). All scaffolds and contigs from the superassembly larger than 500 bp were collected together with all unmapped contigs larger than 500 bp from the original assemblies/mapping in one file as final scaffolds of the superassembly (419,969 scaffolds; Figure 1). The scaffolds were sorted by size (starting with largest), successively numbered and stored in a FASTA file.

Results

NGS of total DNA of *P. tremula x P. alba* INRA 717-1B4 resulted in the generation of about 45.124 million Illumina (MiSeq®) raw reads in pairs (2x250 bp; 20x coverage) and about 61.797 million Ion Torrent (Ion PGM™) raw reads (10x coverage). In total, 419,969 scaffolds of length larger than 500 bp were generated from INRA 717-1B4 by a combination of *de novo* assembly and reference-based assembly followed by superassembly (Figure 1).

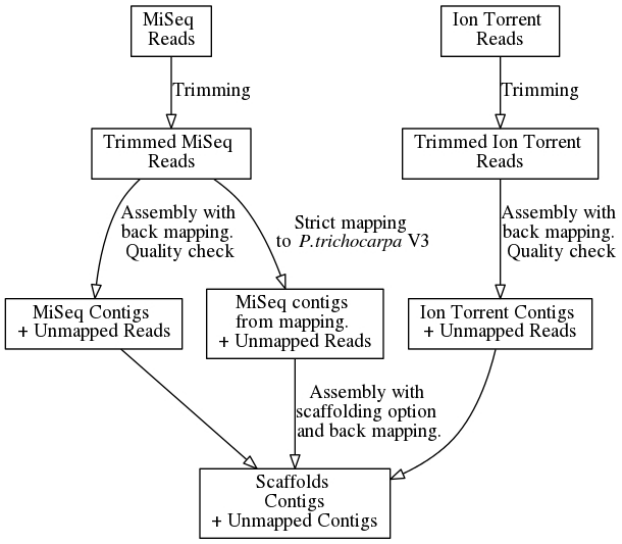

Figure 1
Bioinformatic workflow of *de novo* assembly and scaffolding of INRA 717-1B4 NGS reads. Mapping options were chosen relatively strict. The length fraction and similarity fraction for the mapping against *P. trichocarpa* were specified with 0.94 and 0.98, respectively. All remaining mapping options were set to default.

The accumulated scaffold length is with about 0.9 GB larger than the expected size of the haploid nuclear genome of 0.440 GB for *P. tremula* and 0.509 GB for *P. alba* (estimation based on the C-value of the haploid genomes at the Plant DNA C-values Database; http://data.kew.org/cvalues/). The mean length of the scaffolds is 2,166 bp and the size of the largest scaffold 84,573 bp. The N50 contig length is 3,850 bp (when considering contigs > 1,000 bp). The six largest scaffolds represent the largest chloroplast or mitochondrial scaffolds as determined by BlastN analysis (data not shown) versus the genomic sequences of both organelles (Kersten et al., 2016).

The sequences of the 419,969 scaffolds of INRA 717-1B4 are publicly available for different BLAST analyses at the URGI blast home page (https://urgi.versailles.inra.fr/blast/; "Whole-genome draft assembly of *Populus tremula x P. alba* clone INRA 717-1B4") and for download at OSF (Open Sciences Framework; https://osf.io/) in the following project: https://osf.io/xhjh6/ (doi: 10.17605/OSF.IO/XHJH6; ARK: c7605/osf.io/xhjh6). Access to the BLAST and download pages together with detailed information on INRA 717-1B4 is also provided at the new INRA web portal (https://urgi.versailles.inra.fr/Species/Forest-trees/Populus/Clone-INRA-717-1B4/).

It has to be considered that heterozygote SNP positions are not indicated by nucleotide ambiguity codes (IUPAC) in the DNA sequences of the scaffolds because the assembler is not using these codes. Only one of the two alleles will be displayed

in the scaffold sequence at heterozygote SNP or InDel positions.

Table 1 provides a summary of the chromosomal distribution of the INRA 717-1B4 scaffolds when mapped to the reference sequence of *P. trichocarpa* (v3 downloaded from Phytozome). In total, only 36.6 % of the scaffolds mapped to the 19 chromosomes using a length fraction of 80% and a similarity fraction of 90 % in the mapping. The average coverage of the chromosomes is 0.54. Chromosome 9 shows the highest coverage (0.72) and chromosome 19 the lowest (0.39).

Table 1
Summary of the results for the mapping of the INRA 717-1B4 scaffold sequences against the *P. trichocarpa* reference sequence (chromosome 1-19; v3 at Phytozome; Tuskan et al., 2006).

Reference sequence	Reference length	Average coverage	Total scaffold count
Chromosome 1	50495391	0.53	20143
Chromosome 2	25263035	0.58	10875
Chromosome 3	21816808	0.59	9720
Chromosome 4	24267051	0.50	9367
Chromosome 5	25890704	0.55	10724
Chromosome 6	27912125	0.57	11627
Chromosome 7	15610913	0.52	6080
Chromosome 8	19465461	0.69	10054
Chromosome 9	12948742	0.72	6893
Chromosome 10	22580532	0.64	10725
Chromosome 11	18501271	0.44	6607
Chromosome 12	15760346	0.48	5762
Chromosome 13	16320717	0.54	6458
Chromosome 14	18920894	0.57	7903
Chromosome 15	15278577	0.55	6366
Chromosome 16	14494361	0.50	5563
Chromosome 17	16080358	0.45	5708
Chromosome 18	16958300	0.49	6381
Chromosome 19	15942145	0.39	5055

A small section of the mapping of the INRA 717-1B4 scaffold sequences versus *P. trichocarpa* chromosome 9 is shown in Figure 2. One subregion is covered by two scaffolds which may represent the two different haplotypes of that region. As INRA 717-1B4 is an interspecific hybrid, the heterozygosity level is expected to be high. The generation of more than one scaffold is especially expected in regions with high heterozygosity. In general, there are many regions that are covered by more than one scaffold in the mapping. Scaffolds that cover a given region may include the same sequence or slightly different sequences representing this region. Different sequences may represent the two different haplotypes or two (or more) mixed sequence representations of the different haplotypes in the given region.

Discussion

The draft genomic scaffolds of INRA 717-1B4 created in this study supplement the existing genomic sequence resources for this clone, namely the "sPta717 Genome" described in the Introduction and the complete DNA sequences of the chloroplast and the mitochondrial genomes which we published recently (Kersten et al., 2016). In contrast to the "sPta717 Genome", our new resource provides also access to sequences of INRA 717-1B4 in genomic regions that show no or only a low similarity to the *P. trichocarpa* reference genome.

The relatively low mean coverage detected when mapping the INRA 717-1B4 scaffold sequences to the *P. trichocarpa* reference chromosomes reflects the sequence differences between INRA 717-1B4 (*P. tremula x P. alba*) and the *P. trichocarpa* reference, which are members of the two different *Populus* sections (*Populus* and *Tacamahaca*; see Introduction). Interestingly, the lowest coverage was reported for chromosome 19 that includes the sex-linked region in different *Populus* species (summarized, *e.g.*, in Kersten et al., 2017). The location of this region at chromosome 19 differs in members of the section *Populus* and *Tacamahaca* (summarized, *e.g.*, Kersten et al., 2017) and a different genomic architecture of sex determination was postulated (Geraldes et al., 2015); thus, larger sequence differences are not unexpected for this chromosome when comparing members of these sections.

The further combination of the INRA 717-1B4 scaffolds to chromosomal sequences is highly desired but very complicated due to the high level of heterozygosity in this interspecific hybrid. As shown in Christe et al. (2016), chromosomes of perfect F1 individuals of an interspecific hybrid are (almost) completely of heterospecific ancestry. Sequence data of the parents of INRA 717-1B4 would therefore be very helpful to fulfil this task and to resolve the two haplotypes; however, unfortunately, the parents of INRA 717-1B4 are not available anymore. New next generation sequencing methods capable of delivering long read lengths with high accuracy are forthcoming and will undoubtedly facilitate resolution of highly heterozygous regions as well as large insertions/deletions, copy number variations etc. The future inclusion of heterozygote SNPs and InDels in the INRA 717-1B4 nuclear genomic sequence would add additional value to this genomic resource.

Acknowledgements

We acknowledge the support by CEA-IG/CNG performing the DNA QC in its DNA and Cell Bank service (Anne Boland's team) and providing access to their Illumina Sequencing Platform. We thank Isabelle Le Clainche and Aurélie Chauveau for performing DNA extraction and sequencing. We are grateful to Katrin Groppe, Manuela Will, and Susanne Bein for technical assistance. We thank Karine Robineau for search of publications. We are grateful to Françoise Alfama-Depauw to release

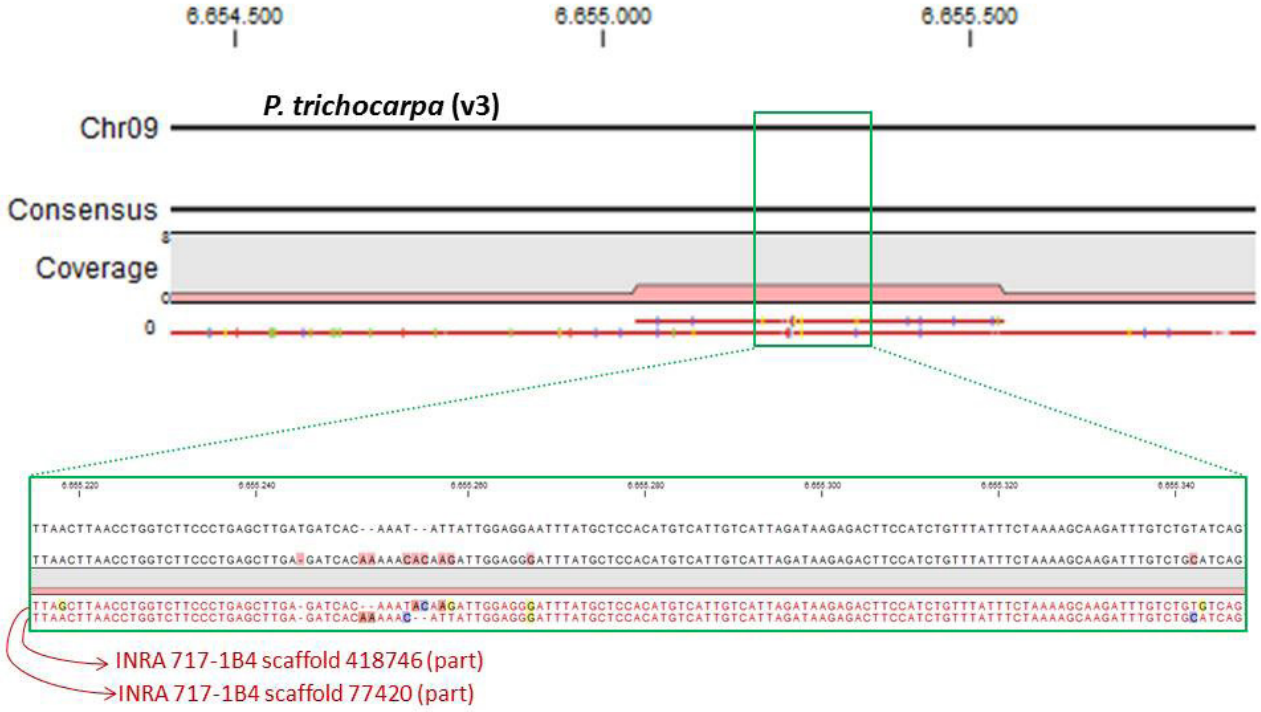

Figure 2
Part of the mapping of INRA 717-1B4 scaffolds to the DNA sequence of *P. trichocarpa* chromosome 9 (at the top) and enlargement of one subregion where two scaffolds mapped (at the bottom)

the draft genome within the URGI Blast web page and the URGI to maintaining the infrastructure.

References

Christe C, Stölting KN, Bresadola L, Fussi B, Heinze B, Wegmann D, Lexer C (2016) Selection against recombinant hybrids maintains reproductive isolation in hybridizing *Populus* species despite F1 fertility and recurrent gene flow. Mol Ecol 25 (11):2482-2498. http://dx.doi.org/10.1111/mec.13587

Dickmann D, Stuart KW (1993) The culture of poplars in eastern North America. Michigan: Michigan State University, 1st edition (1983), 168 p

Eckenwalder JE (1996) Systematics and evolution of *Populus*. In: Stettler RF, Bradshaw Jr. HD, Heilman PE, Hinckley TM (eds) Biology of *Populus* and its implications for management and conservation. Ottawa, Ontario, Canada: NRC Research Press, National Research Council of Canada, pp 7–32, ISBN 0660165066, 9780660165066

Fladung M, Altosaar I, Bartsch D, Baucher M, Boscaleri F, Gallardo F, Häggman H, Hoenicka H, Nielsen K, Paffetti D, Séguin A, Stotzky G, Vettori C (2012) European discussion forum on transgenic tree biosafety. Nature Biotechnology 30 (1): 37-38. http://dx.doi.org/10.1038/nbt.2078

Geraldes A, Hefer CA, Capron A, Kolosova N, Martinez-Nuñez F, Soolanayakanahally RY, Stanton B, Guy RD, Mansfield SD, Douglas CJ, Cronk QC (2015) Recent Y chromosome divergence despite ancient origin of dioecy in poplars (*Populus*). Mol Ecol 24 (13):3243-3256. http://dx.doi.org/10.1111/mec.13126

Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUAST: quality assessment tool for genome assemblies. Bioinformatics 29 (8):1072-1107. http://dx.doi.org/10.1093/bioinformatics/btt086

Jansson S, Douglas CJ (2007) *Populus*: A model system for plant biology. Annu Rev Plant Biol 58:435-458. http://dx.doi.org/10.1146/annurev.arplant.58.032806.103956

Kersten B, Faivre Rampant P, Mader M, Le Paslier MC, Bounon R, Berard A, Vettori C, Schroeder H, Leple JC, Fladung M (2016) Genome sequences of *Populus tremula* chloroplast and mitochondrion: Implications for holistic poplar breeding. PLoS One 11 (1):e0147209. http://dx.doi.org/10.1371/journal.pone.0147209

Kersten B, Pakull B, Fladung M (2017) Genomics of sex determination in dioecious trees and woody plants. Trees – Structure and Function 31 (4):1113–1125. http://dx.doi.org/10.1007/s00468-017-1525-7

Ky CL, Barre P, Lorieux M, Trouslot P, Akaffou S, Louarn J, Charrier A, Hamon S, Noirot M (2000) Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.). Theor Appl Genet 101 (4):669-676. http://dx.doi.org/10.1007/s001220051529

Lemoine M (1973) Amélioration des peupliers de la section Leuce sur sols hydromorphes [Thesis]: University of Nancy, France

Leplé JC, Brasileiro ACM, Michel MF, Delmotte F, Jouanin L (1992) Transgenic poplars: expression of chimeric genes using four different constructs. Plant Cell Rep 11:137-141. http://dx.doi.org/10.1007/BF00232166

Licht LA, Isebrands JG (2005) Linking phytoremediated pollutant removal to biomass economic opportunities. Biomass Bioenergy 28 (2):203-218. http://dx.doi.org/10.1016/j.biombioe.2004.08.015

Ma T, Wang J, Zhou G, Yue Z, Hu Q, Chen Y, Liu B, et al. (2013) Genomic insights into salt adaptation in a desert poplar. Nat Commun 4:2797. http://dx.doi.org/10.1038/ncomms3797

Pilate G, Dejardin A, Leple JC (2012) Chapter 1 - Field trials with lignin-modified transgenic trees. In: Lise J, Catherine L (eds) Advances in Botanical Research. Volume 61 (Lignins). London, Oxford, Boston, New York, San Diego: Academic Press, pp 1-36. ISBN 0124160239

Pinosio S, Giacomello S, Faivre-Rampant P, Taylor G, Jorge V, Le Paslier MC, Zaina G, Bastien C, Cattonaro F, Marroni F, Morgante M (2016) Characterization of the poplar pan-genome by genome-wide identification of structural variation. Mol Biol Evol 33 (10):2706-2719. http://dx.doi.org/10.1093/molbev/msw161

Sjödin A, Street NR, Sandberg G, Gustafsson P, Jansson S (2009) The *Populus* Genome Integrative Explorer (PopGenIE): a new resource for exploring the *Populus* genome New Phytol 182 (4):1013-1025. http://dx.doi.org/10.1111/j.1469-8137.2009.02807.x

Slavov GT, Zhelev P (2010) Salient biological features, systematics, and genetic variation of *Populus*. In: Jansson S, Bhalerao R, Groover A (eds) Genetics and Genomics of *Populus*, Plant Genetics and Genomics: Crops and Models. Ber-

lin, Heidelberg, New York: Springer, pp 15-38, ISBN 978-1-4419-1540-5 (Print) 978-1-4419-1541-2 (Online)

Stanton BJ, Neale DB, Li S (2010) *Populus* breeding: From the classical to the genomic approach. In: Jansson S, Bhalerao R, Groover A (eds) Genetics and Genomics of *Populus*. Dordrecht, Netherlands: Springer Science+Business Media, pp 309-348

Stettler RF, Bradshaw Jr. HD, Heilman PE, Hinckley TM (1996) Biology of *Populus* and its implications for management and conservation. Ottawa, Ontario, Canada: NRC Research Press, National Research Council of Canada, 539 p, ISBN 0660165066, 9780660165066

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, et al. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313 (5793):1596-1604. http://dx.doi.org/10.1126/science.1128691

Walle IV, Van Camp N, Van de Casteele L, Verheyen K, Lemeur R (2007) Short-rotation forestry of birch, maple, poplar and willow in Flanders (Belgium) II. Energy production and CO2 emission reduction potential. Biomass Bioenergy 31 (5):276-283. http://dx.doi.org/10.1016/j.biombioe.2007.01.019

Walter C, Fladung M, Boerjan W (2010) The 20-year environmental safety record of GM trees. Nat Biotechnol 28 (7):656-658. http://dx.doi.org/10.1038/nbt0710-656

Wullschleger SD, Jansson S, Taylor G (2002) Genomics and forest biology: *Populus* emerges as the perennial favorite. Plant Cell 14 (11):2651-2655. http://dx.doi.org/10.1105/tpc.141120

Zhou X, Jacobs TB, Xue LJ, Harding SA, Tsai CJ (2015) Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate:CoA ligase specificity and redundancy. New Phytol 208 (2):298-301. http://dx.doi.org/10.1111/nph.13470