

Genomic approaches to improvement of *Populus* as a biomass and bioenergy feedstock

Pacific Rim Summit on Industrial
Biotechnology & Bioenergy

Workshop: Genomic Tools for
Optimization of Biomass Production
October 11, 2012

Carl Douglas
Department of Botany
University of British Columbia



Cell wall, poplar genomics, and biomass research

<http://www.botany.ubc.ca/people/carl-douglas>

<http://thedouglaslab.blogspot.ca/>



University of British Columbia



UBC Department of Botany



GenomeCanada



Plant secondary cell walls

Pollen walls

Poplar genomics

Biomass improvement for bioenergy

Molecular basis of phenotypic diversity

POPCAN:

Genetic improvement of poplar trees as a Canadian bioenergy feedstock



Carl Douglas
Shawn Mansfield

Quentin Cronk
Rob Guy
Yousry El-Kasaby
Jürgen Ehling
Geoff Wasteneys

Richard Hamelin
Malcolm Campbell
Tom Maness
Marty Luckert

Poplar Bioenergy: Applied
Genomics Innovation Program
(AGIP Project)

Large Scale Applied Research
Program (POPCAN Project)



GenomeBritishColumbia

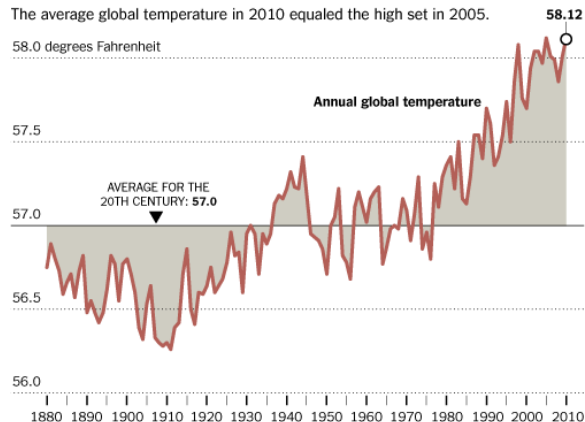


GenomeCanada

Demand for renewable transportation fuels

It's Getting Hot Out There

The average global temperature in 2010 equaled the high set in 2005.



Source: National Oceanic and Atmospheric Administration National Climatic Data Center

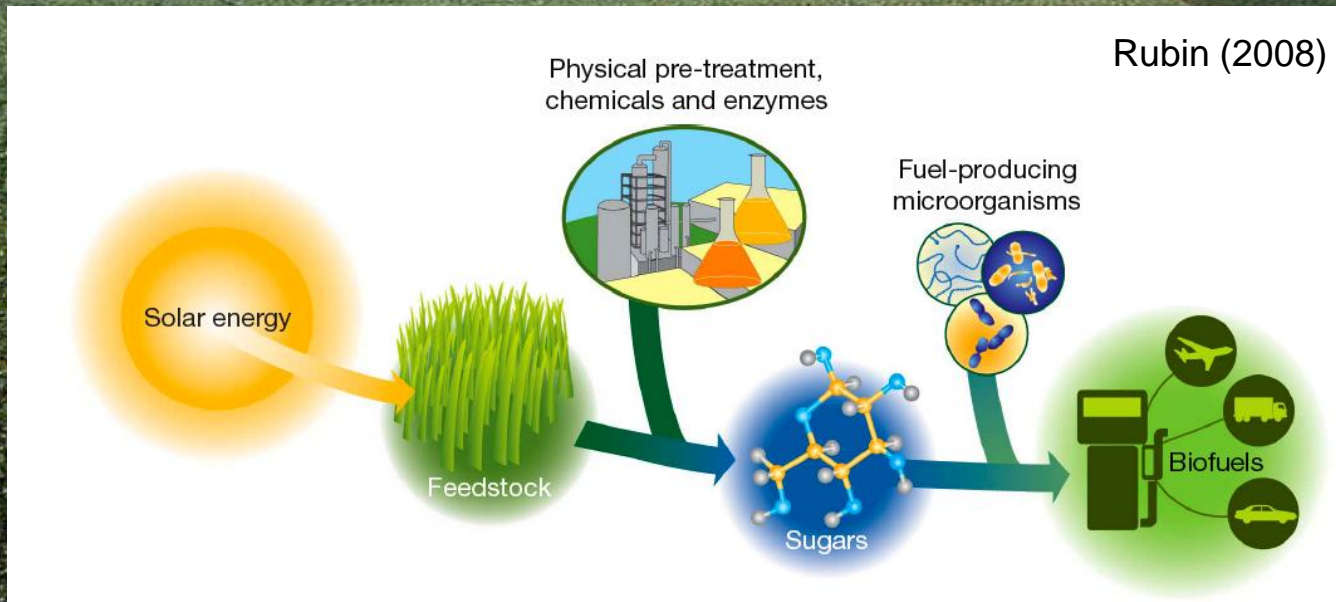
2010 second warmest year on record



36 billion litres/yr light vehicle gasoline consumption in Canada

- **At the current 5% Federal ethanol supplementation target in Canada ~2 billion litres are required annually**
- **Higher targets (10%) likely in the future**
- **BC Bioenergy Strategy (BC production of 50% of BC renewable transportation fuel)**

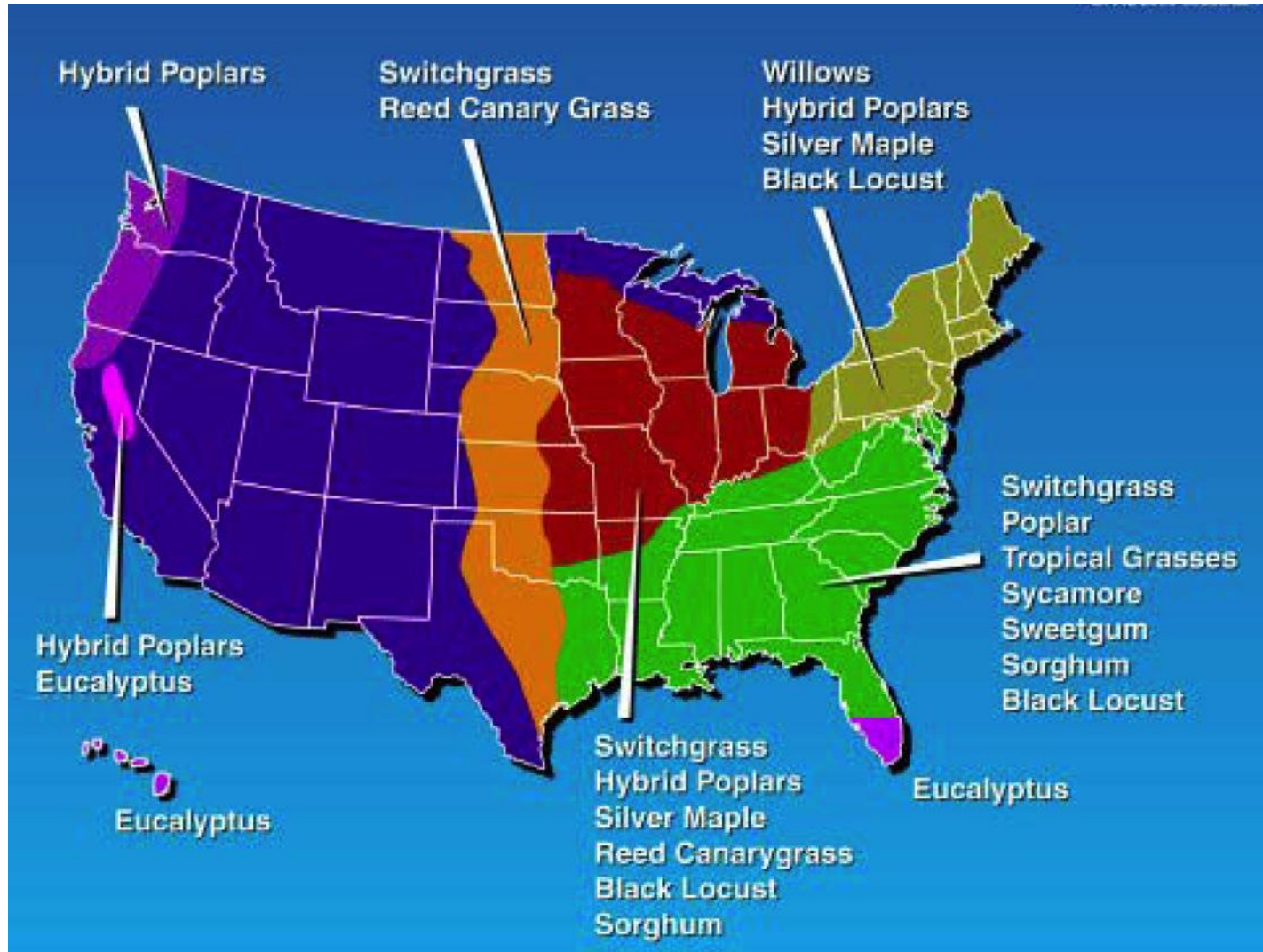
Current Canadian production is ~1.3 billion litres (agricultural feedstocks)



Secondary cell walls in trees: captured solar energy
Targets for lignocellulosic renewable energy



Geographic distribution of biomass crops in the US



Wright et al DOE-ORNL-EERE

Biomass crop productivity

Crop	Total Yield (ton/ha/yr)
Sorghum	40-55
Sugarcane	80-120
Switchgrass	14-18
Miscanthus	30-41
Corn Stover	3-5
Poplar	12-24



J. Eaton

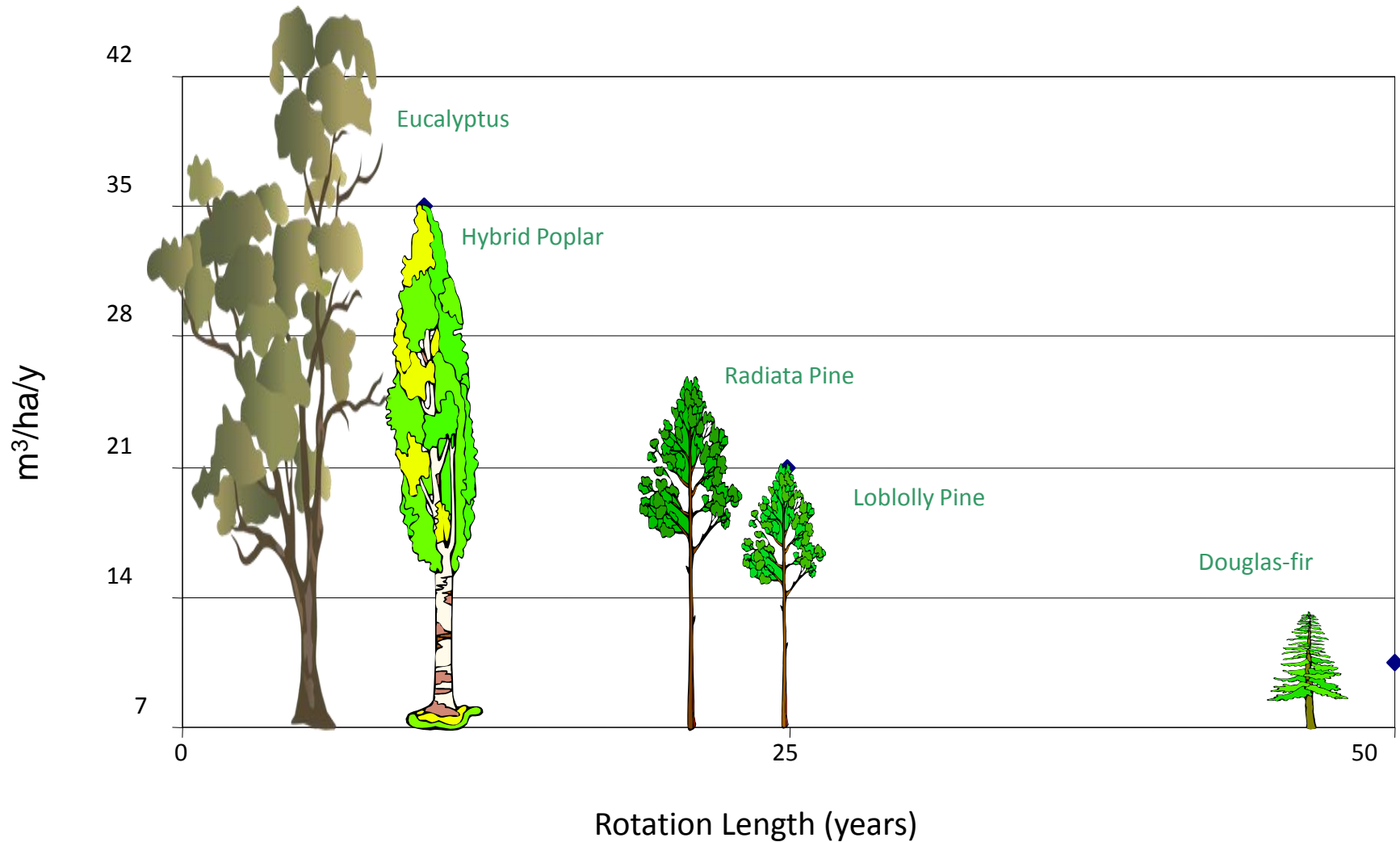
Fast growing trees: poplar



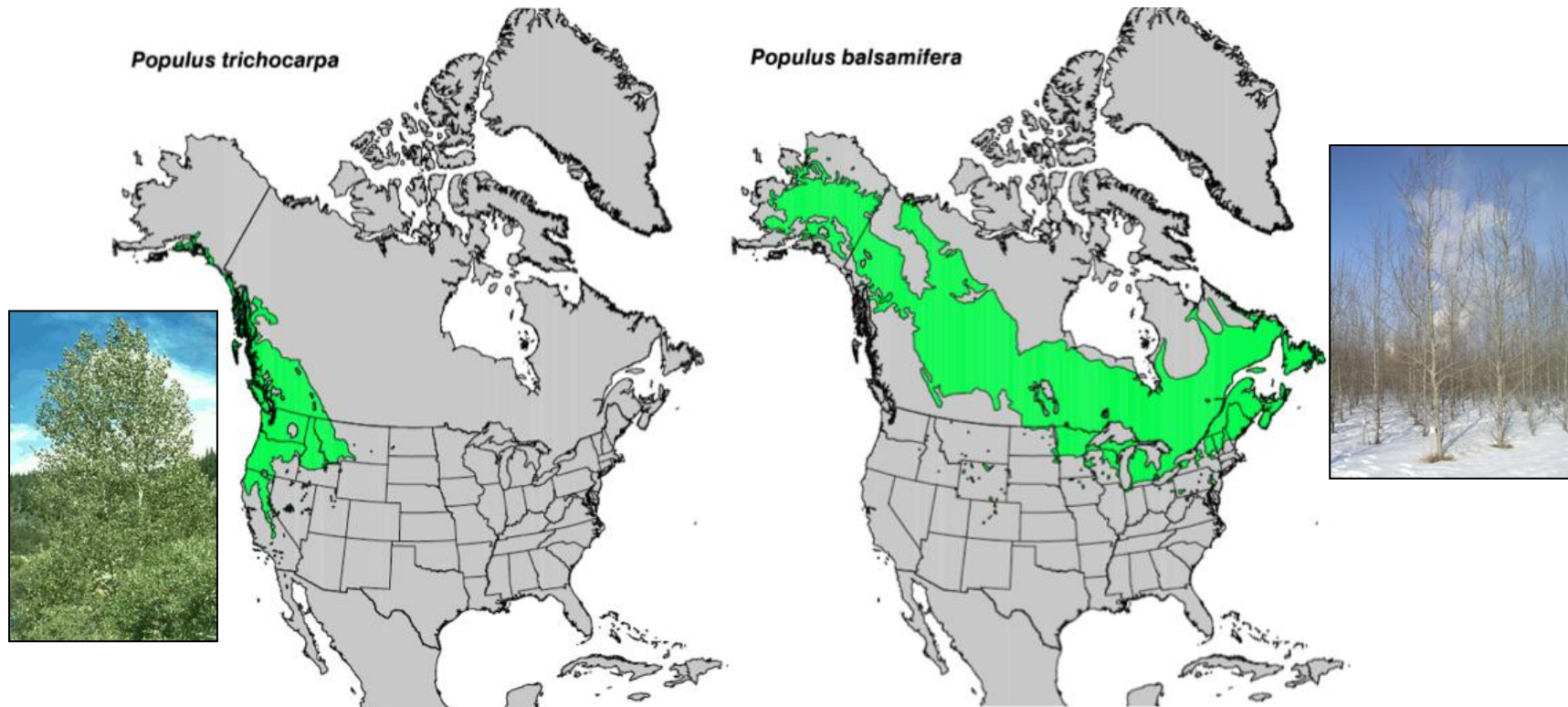
S. Long

Perennial grasses: switchgrass

Comparison of forest tree plantation species

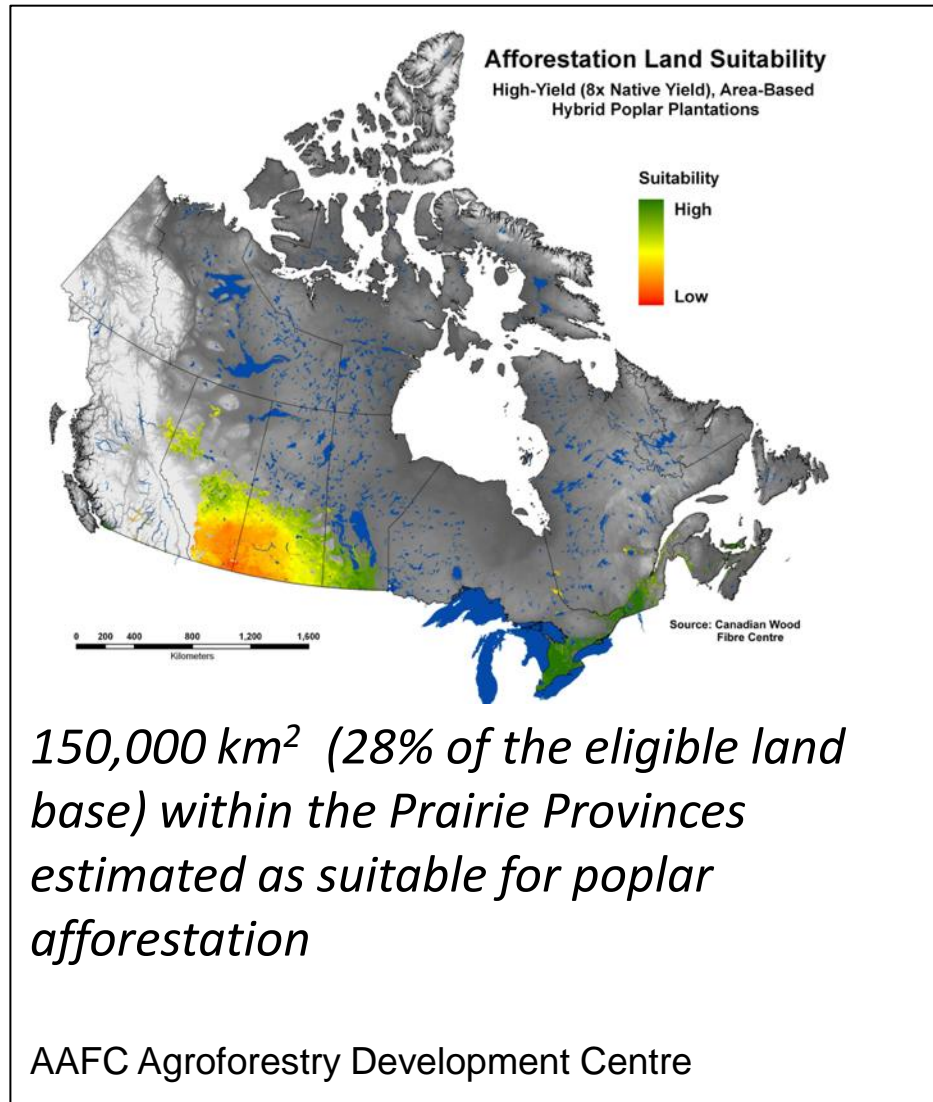


Populus trichocarpa and *P. balsamifera*



Candidate biomass species or as hybrids with other species
for for several Canadian geo climatic regions

Afforestation potential for poplar on Canadian private land



Key traits for improvement

Biomass yield

- productivity, adaptation, pathogen susceptibility

Conversion (lignocellulosics to ethanol)

- sugar release, cell wall structure, polymer biochemistry

Socio-economic challenges

Land use policy

Competition with agriculture for land

Public acceptance of plantation forestry

Land base required for bioenergy biomass plantations is significant



Modeling of impact of genetic improvement on biomass plantation sizes needed

A 28% increase in the conversion-to-ethanol yield (from 234 l/t to 299 l/t) results in:

- Up to 22% reduction in plantation area:

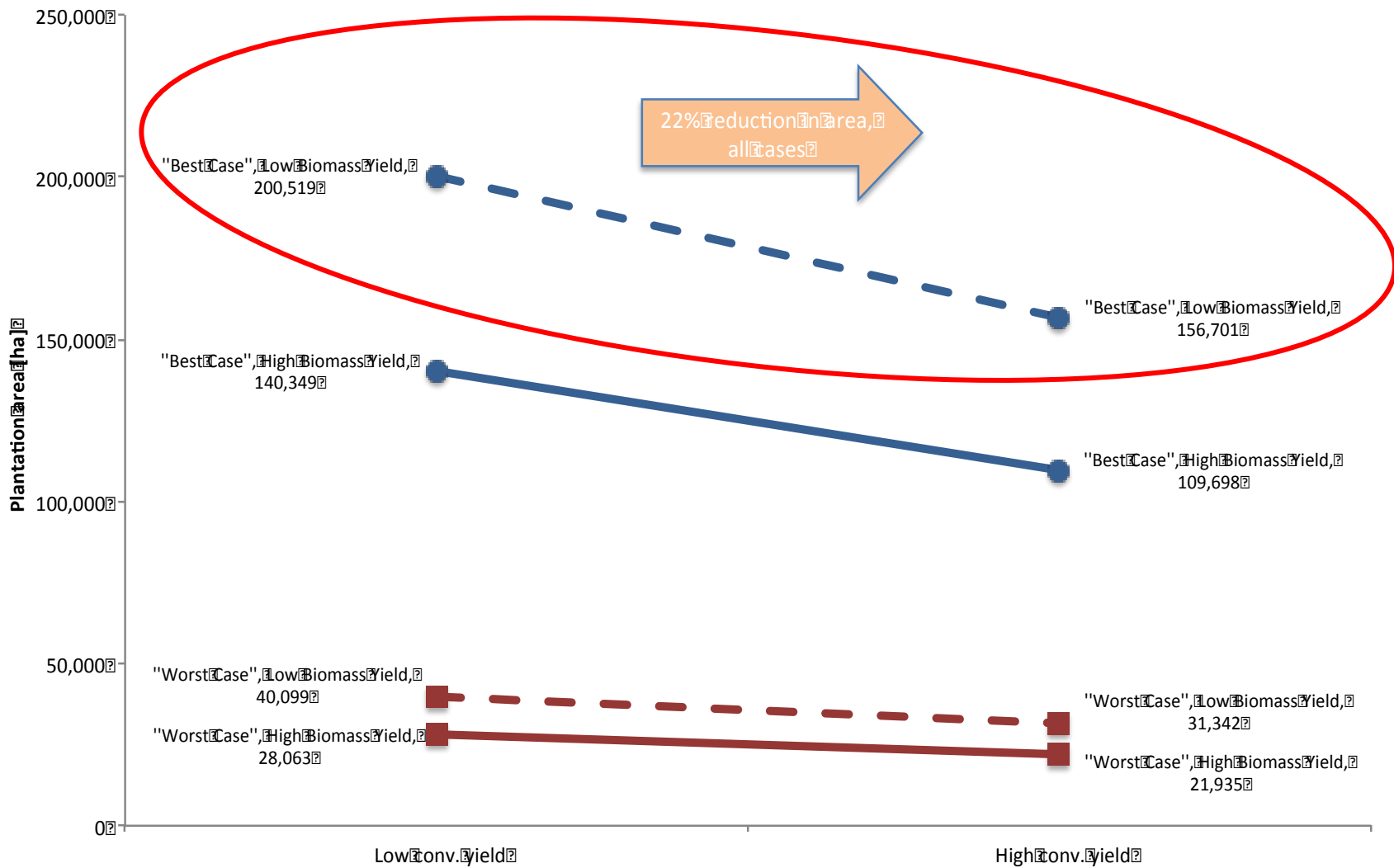
E.g. 43,818 ha reduction (from 200,519 ha to 156,701 ha)

A 43% increase in the biomass productivity growth & yield (from 11.0 t/ha/yr to 15.7 t/ha/yr):

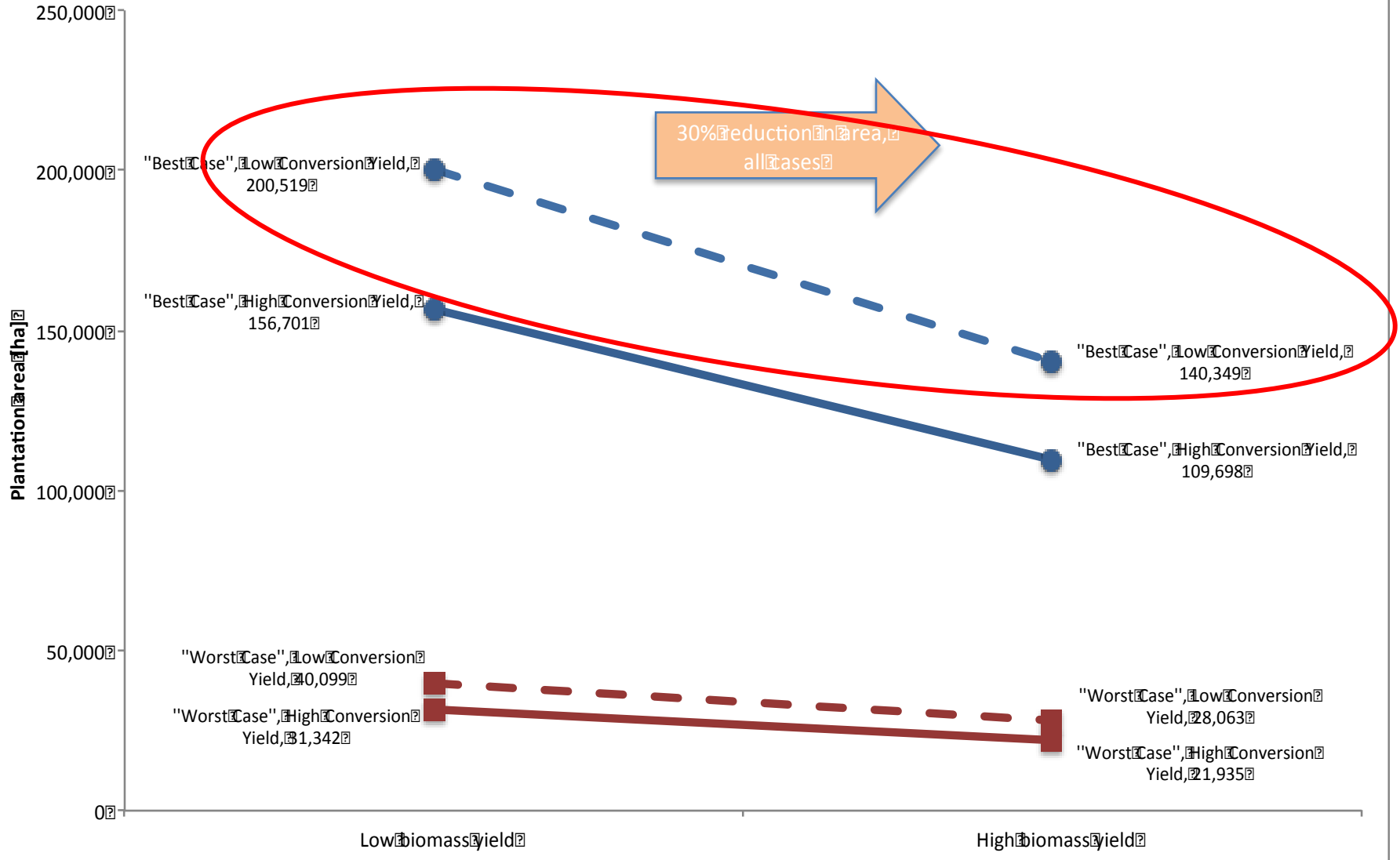
- Up to 30% reduction in plantation area

E.g. 60,170 ha reduction (from 200,519 ha to 140,349 ha)

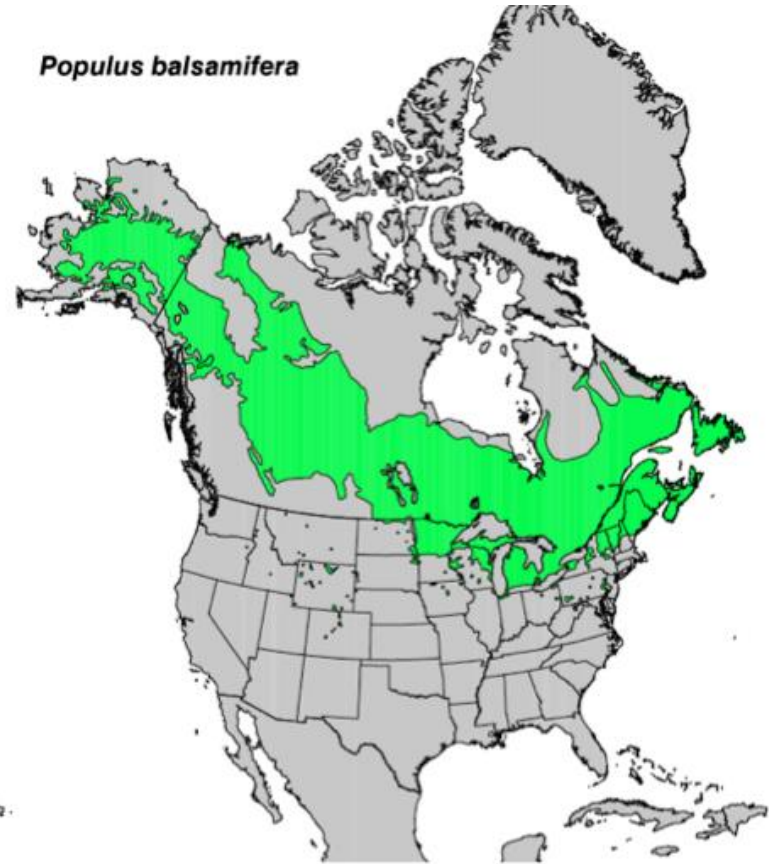
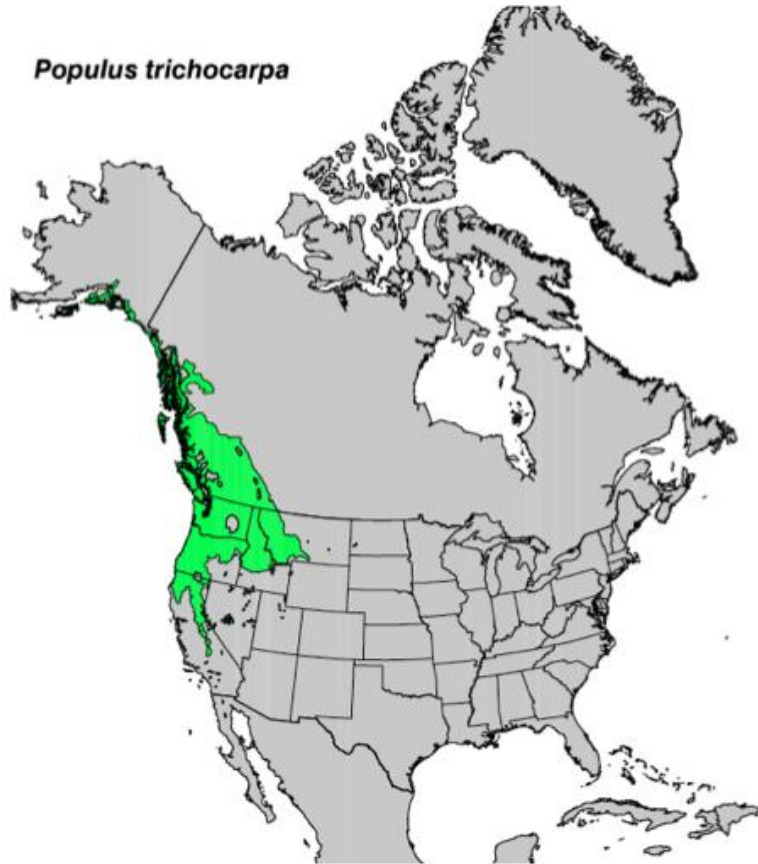
Plantation area needed (holding biomass yield constant, varying conversion yield)



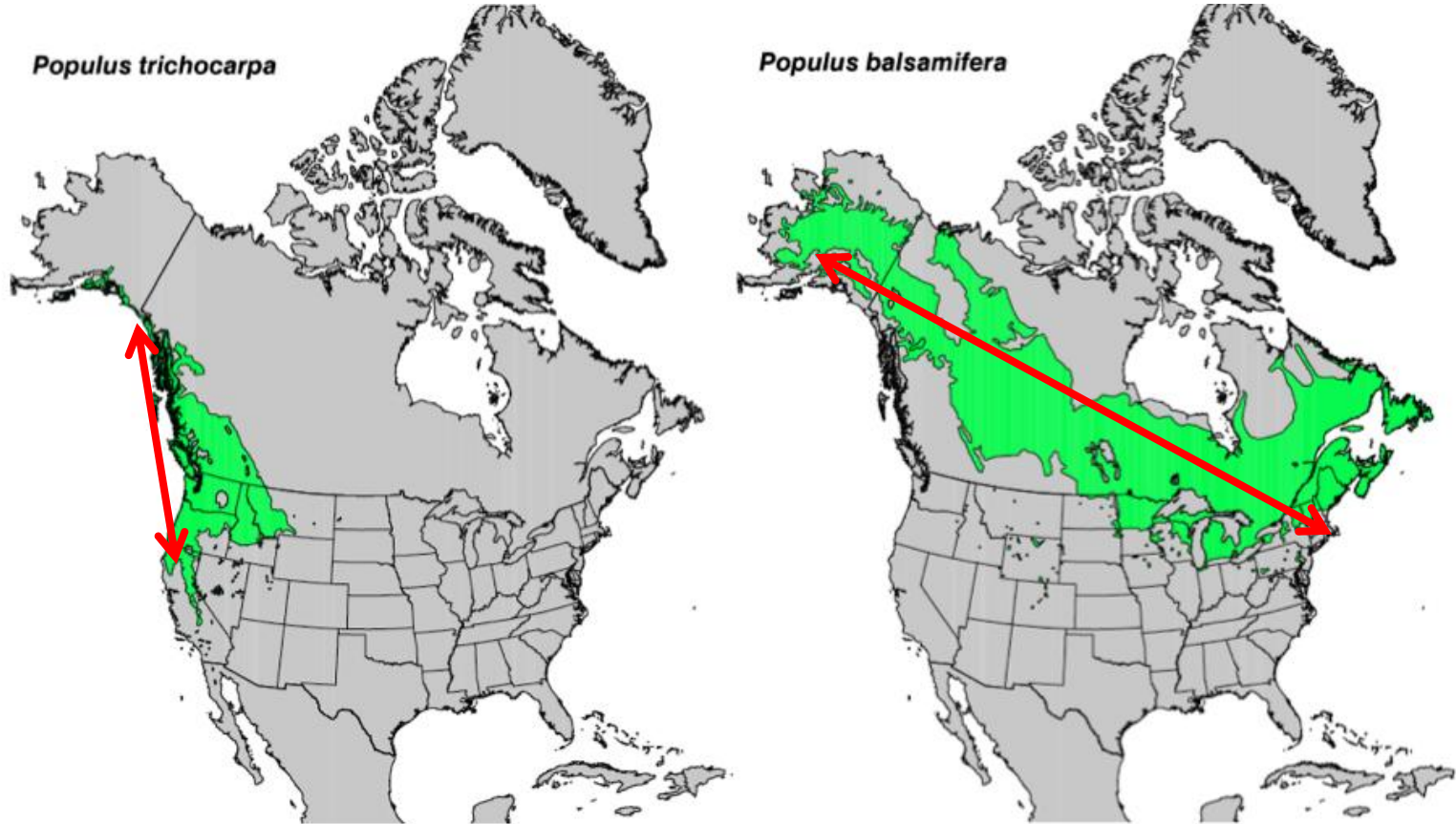
Plantation Area needed (holding conversion yield constant, varying biomass yield)



Populus is largely undomesticated and genetic improvement for bioenergy traits is in its infancy



North American wild populations harbor extraordinary genetic and phenotypic variation

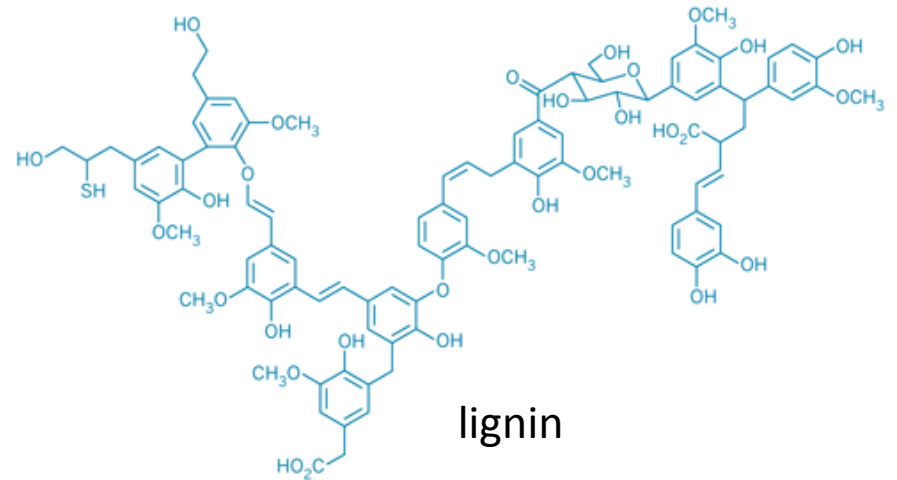
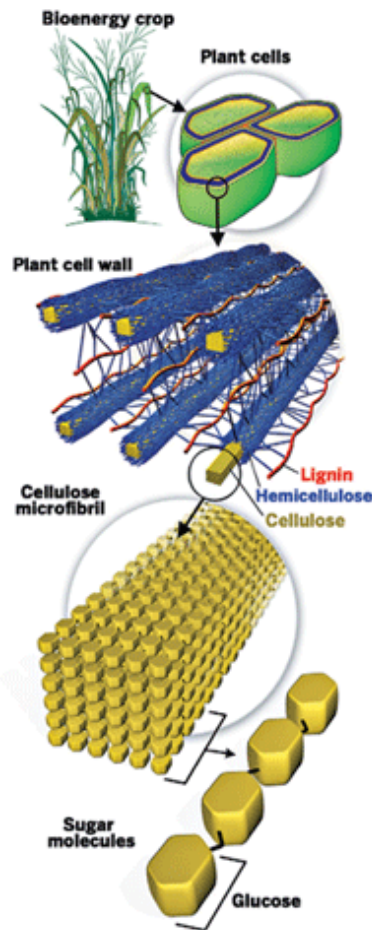


Natural variation in wood chemistry, adaptation to environment, and biomass yield

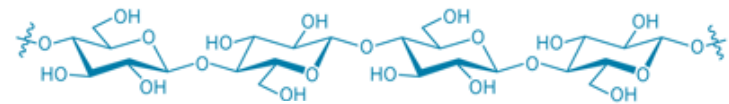
Genomics approaches to understanding genetic basis of trait variation and application to trait improvement are feasible

Lignocellulosic biochemical traits

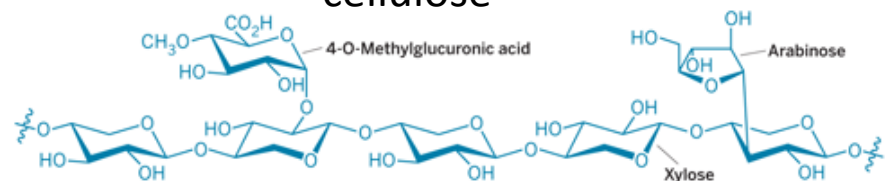
- lignocellulose is a complex biomaterial
- recalcitrant to breakdown into fermentable sugars due to presence of lignin



lignin



cellulose



hemicelluloses

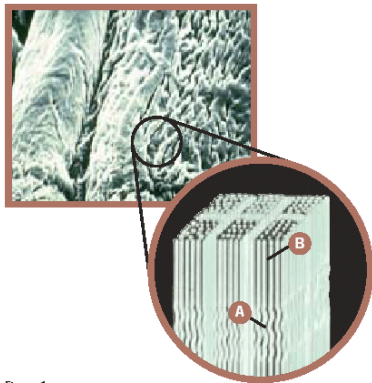
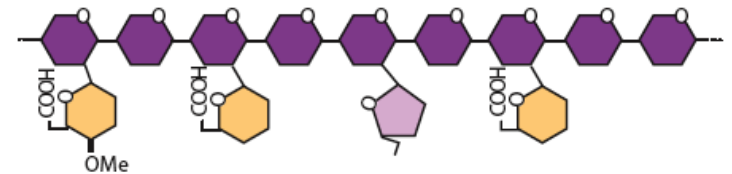
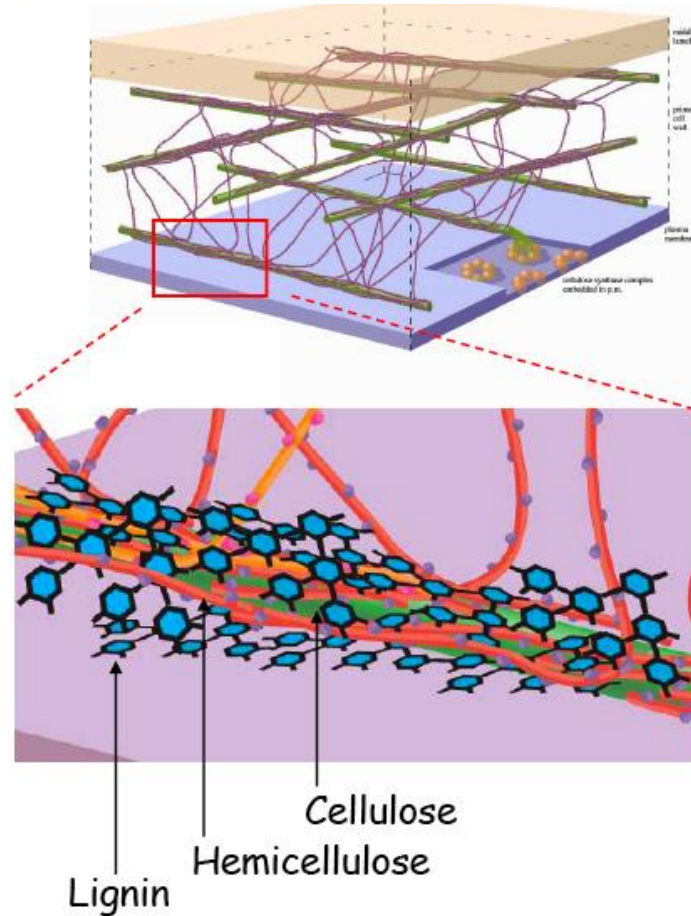
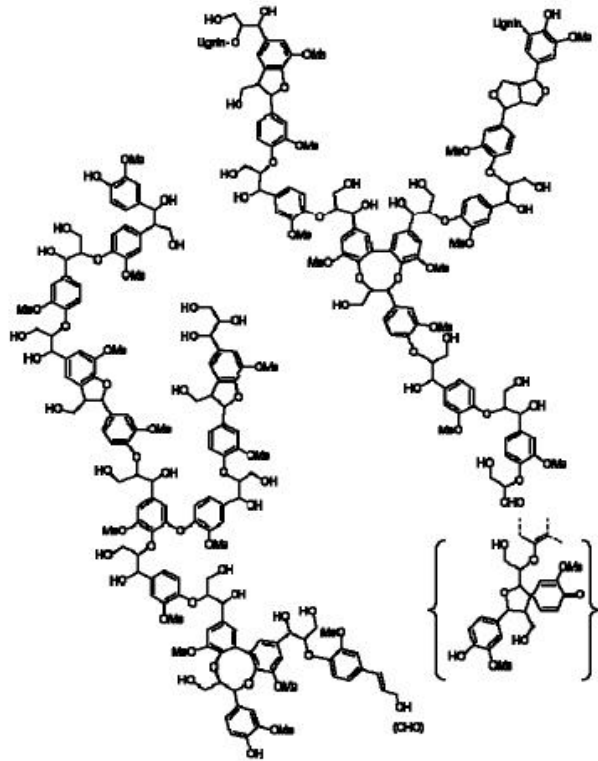


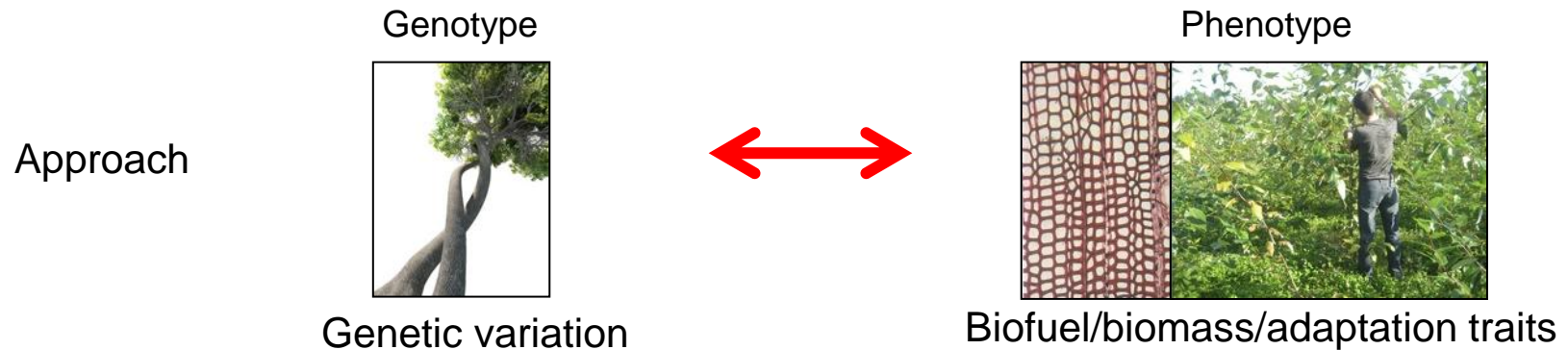
Figure 1
A. The **paracrystalline region**, an amorphous flexible mass of cellulose chains, and
B. The **crystalline region**, which is composed of tight **bundles of microfibrils** in a rigid linear arrangement.

Lignin occludes polysaccharides



Approach for feedstock improvement:

- Use genome-wide analyses to reveal genetic basis for phenotypic variation in wood and biomass/adaptive traits important for feedstock performance
- Develop tool kit of favorable alleles for genetic improvement



Phenotype-genotype associations to identify allelic variation underlying traits of interest

Most traits of interest are complex quantitative traits

Populus trichocarpa – black cottonwood: genomic resources

Genome sequence published in 2006

Nisqually-1 reference genome

~ 40,000 genes

Annotated genome curated at Phytozome.org

Community bioinformatic resources

**Large collections of natural accessions in
common gardens**



Tuskan et al., 2006

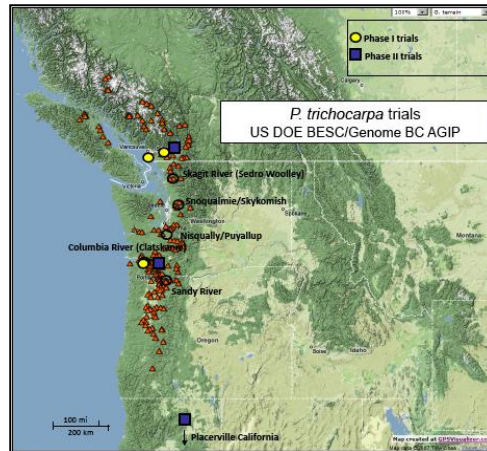
Poplar collections that capture natural variation

P. trichocarpa: BC
MOF/Totem Field



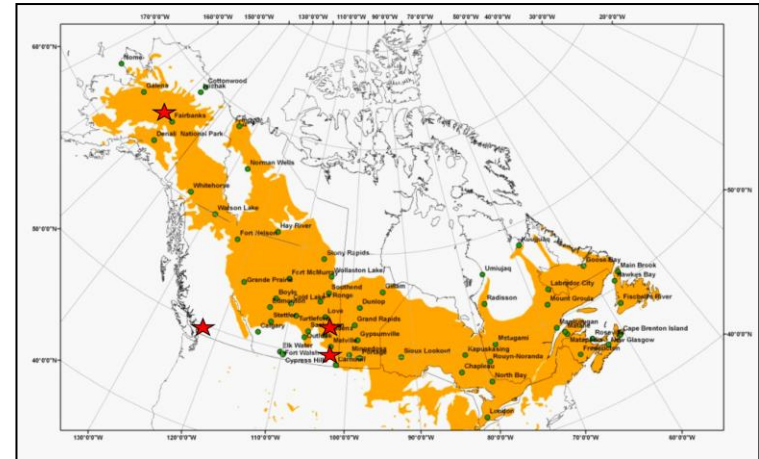
500

P. trichocarpa: BESC-
AGIP Phase II/BC, OR, CA



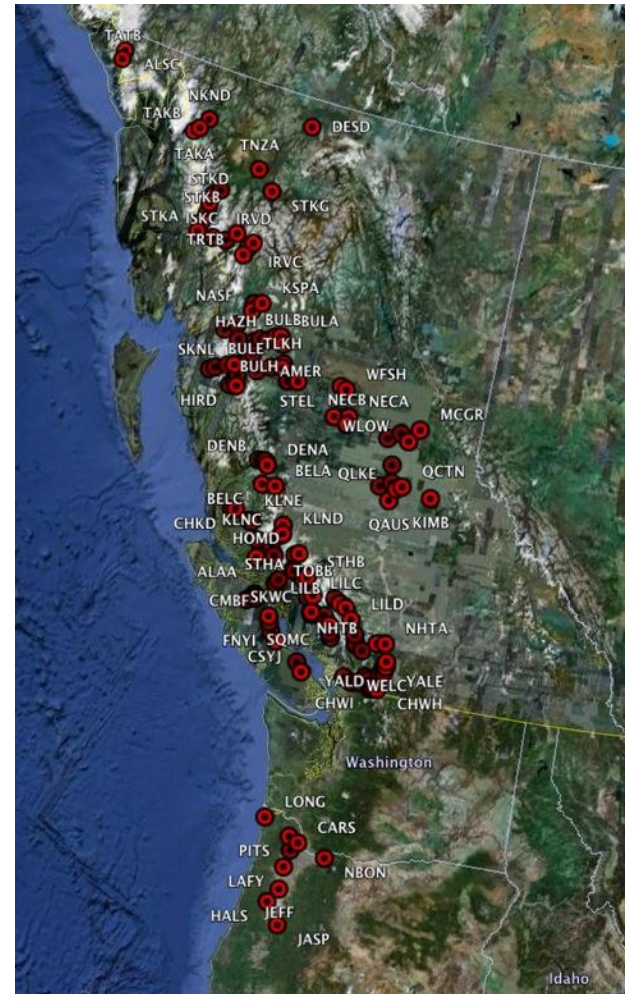
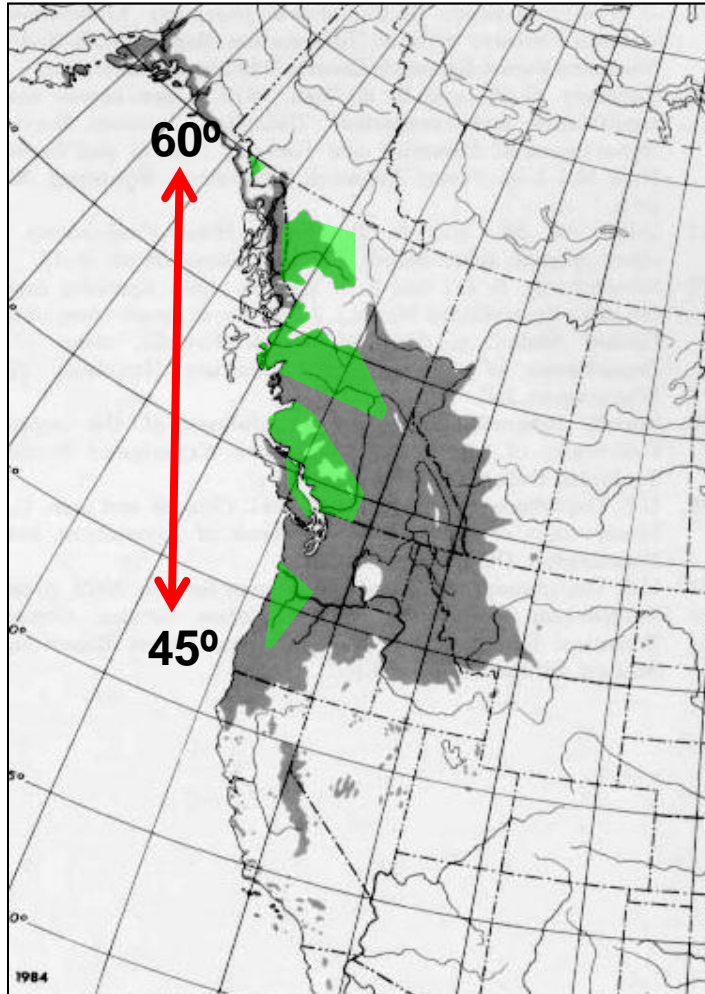
1100

P. balsamifera: AgCanBaP



600

BC Ministry of Forests *P. trichocarpa* collection



Common garden in UBC Totem Field



BC Ministry of Forests collection of ~450 individuals in replicated random design



Populus trichocarpa – Totem Field common garden (2011)



Totem field harvest March 2012

Physiological/biomass phenotyping



Athena McKown

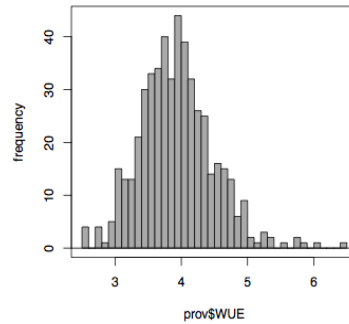
- 456 successful genotype transplants into the Totem Field garden
- 4 - 20 replicates of each genotype (~2088 surviving trees)
- Numerous traits and calculated variables:

	# of traits / variables	years
Phenology	29	2008 - 2011
Biomass / growth related	18	2008 - 2011
Gas exchange related	15 (15 post-bud set)	2009 - 2010
Leaf traits	13	2009 - 2011
Health / disease	9	2008 - 2011
Geography / climate*	23	1971 - 2000

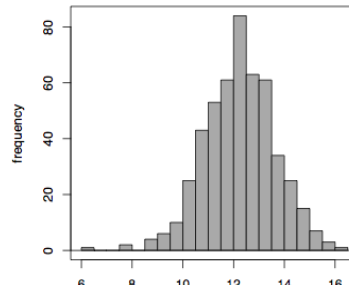
* Derived using ClimateWNA (Wang et al. 2006)

Variation in physiological traits

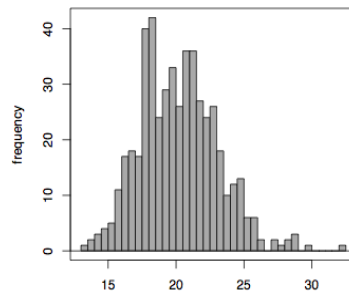
Water use efficiency



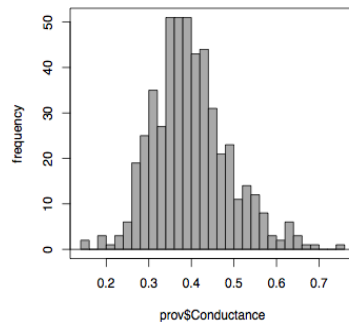
Leaf mass/ area



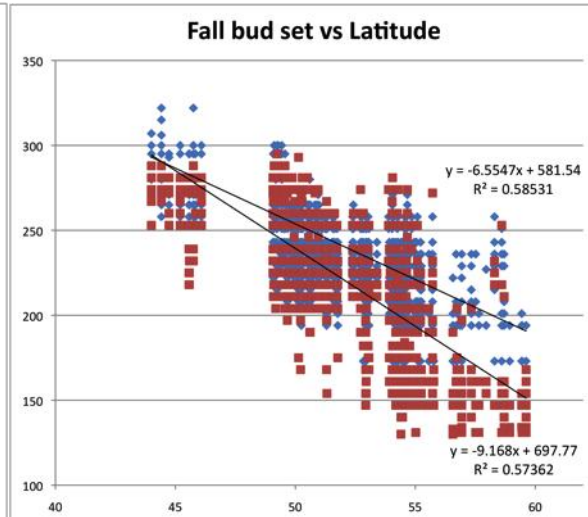
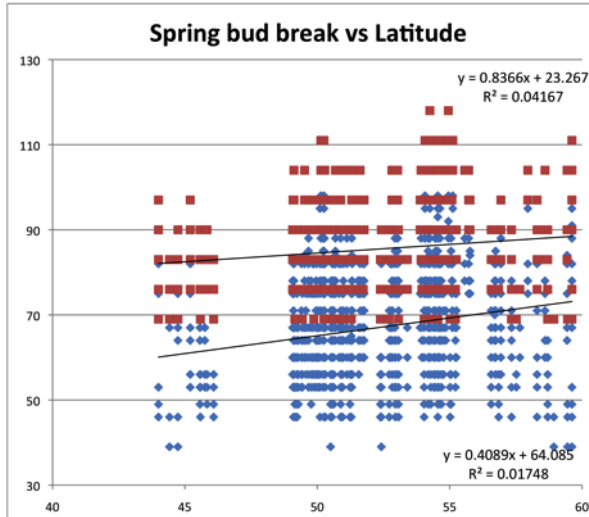
CO₂ assimilation



Stomatal gas exchange

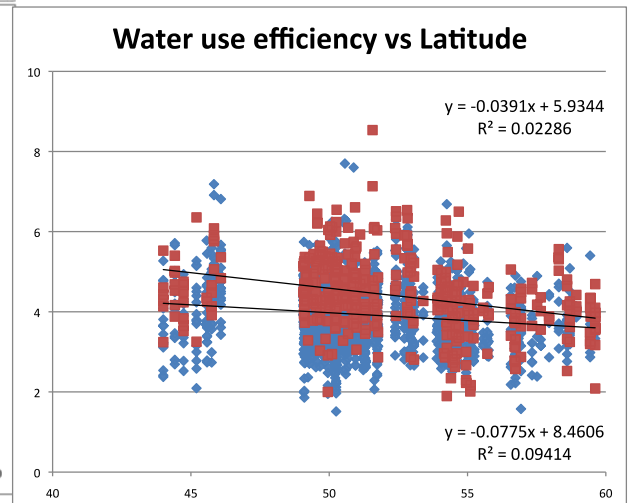
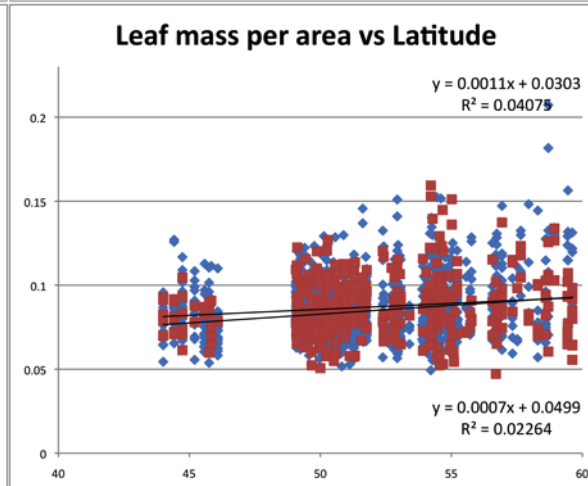
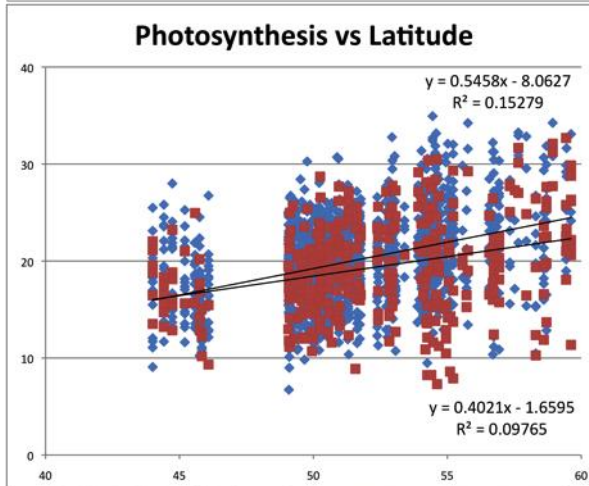


Adaptation in *P. trichocarpa*



- Numerous traits show association with geography of tree origin:

- Latitude
- Longitude
- Elevation



P. trichocarpa provenance trial at Surrey BC – established 2000



Wood Phenotyping

Shawn Mansfield, Ilga Porth, Alex Skyba

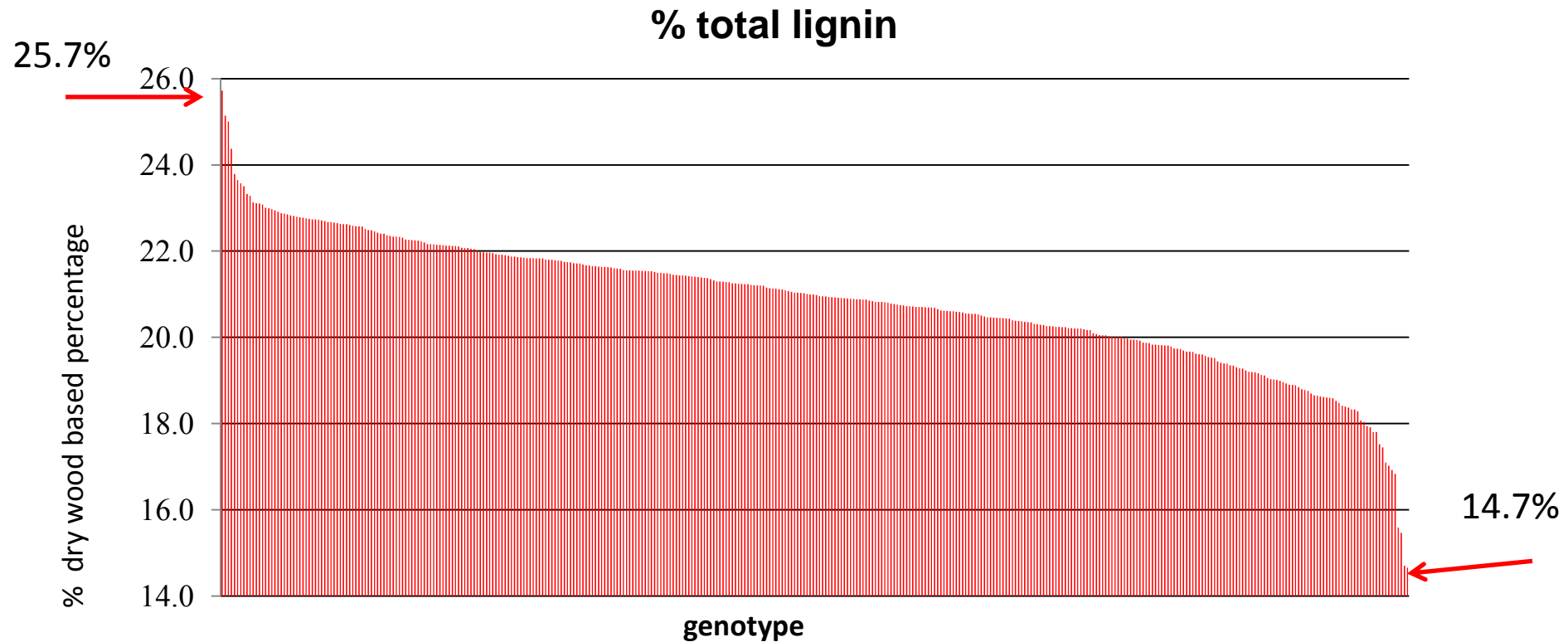
Wood chemistry (13)

1. Arabinose
2. Rhamnose
3. Galactose
4. Glucose
5. Xylose
6. Mannose
7. Insoluble lignin
8. Soluble lignin
9. Total lignin
10. Holocellulose
11. Alpha-cellulose
12. Hemicellulose
13. Syringyl lignin

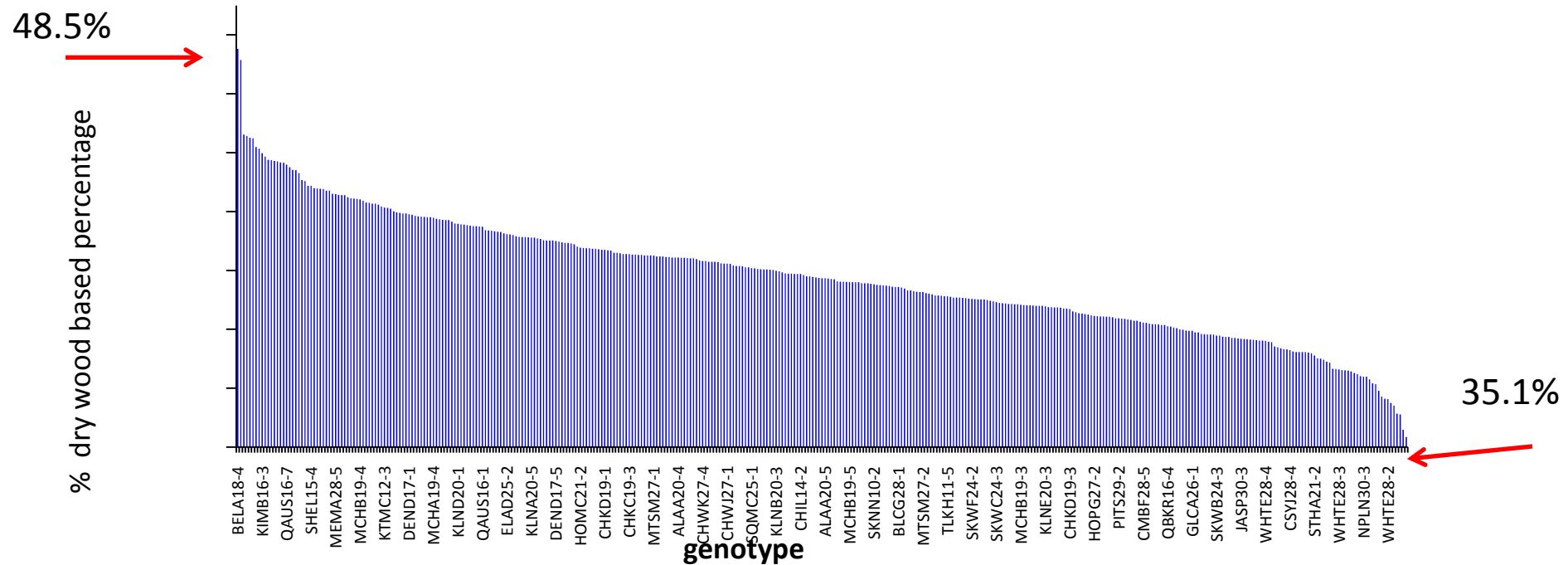
Fiber properties (5)

1. Fibre length
2. MFA (recent growth ring)
3. MFA (first growth ring)
4. Cellulose crystallinity
5. Average wood density

Phenotypic variation – total lignin



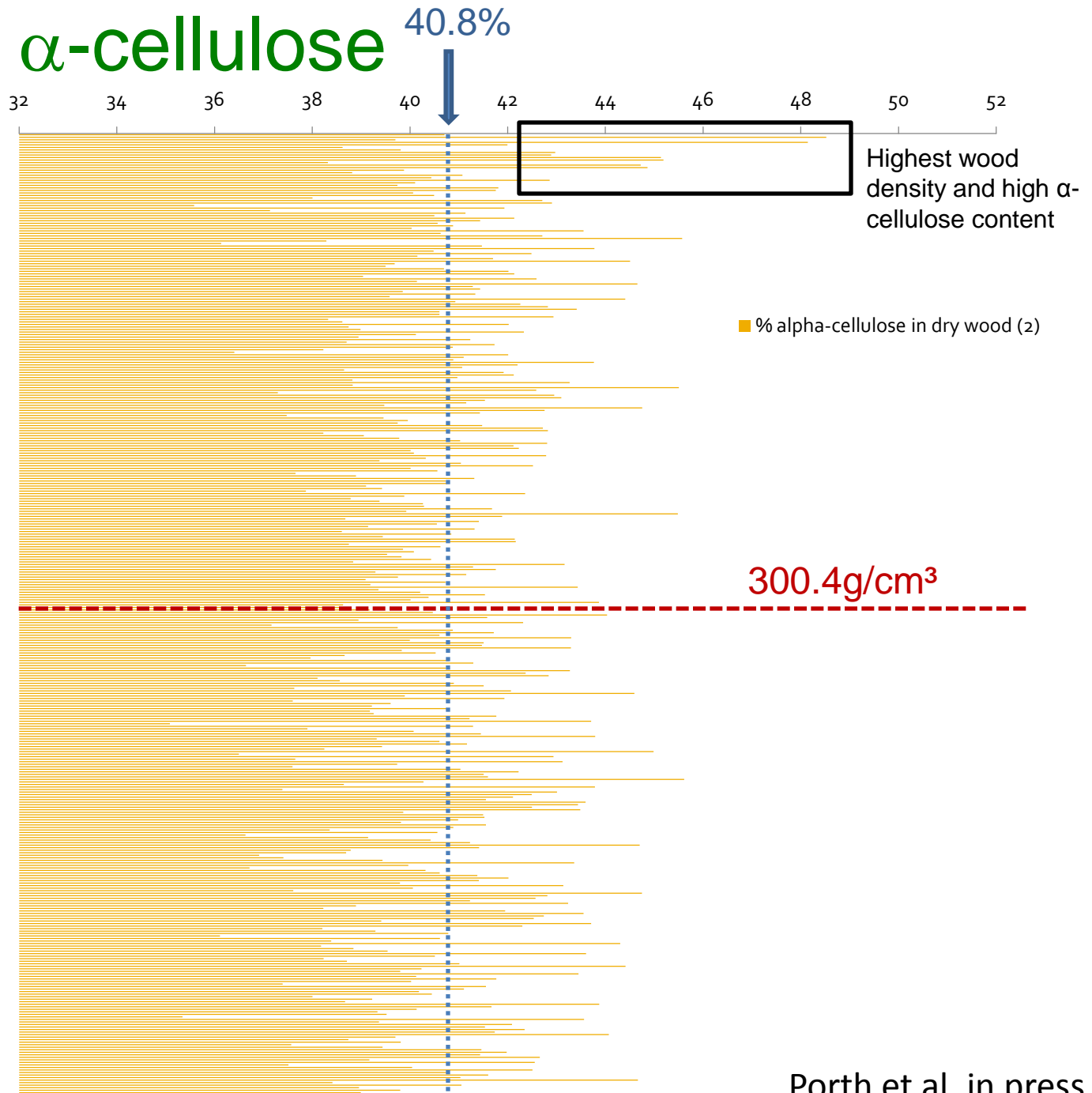
Phenotypic variation – alpha cellulose



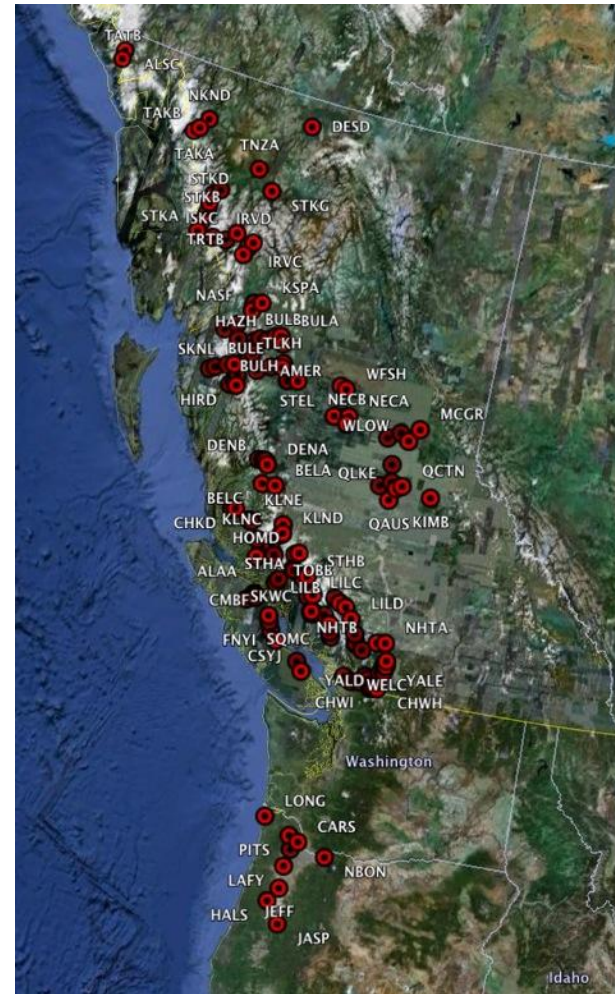
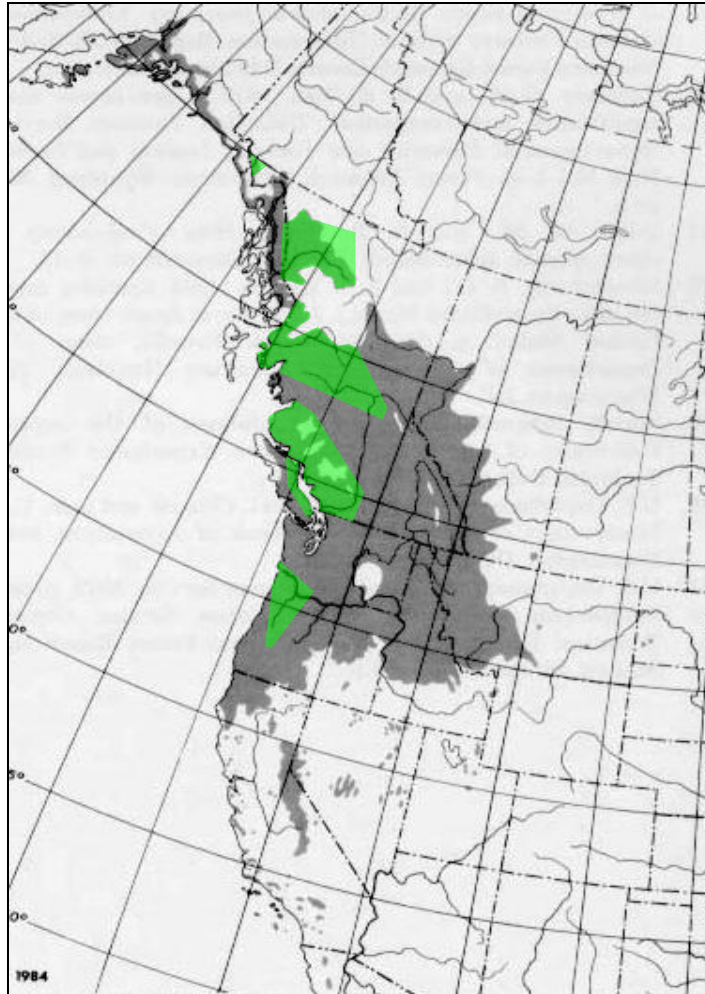
Density and α -cellulose

n=380

wood density



Genetic variation



SNP detection by resequencing



... TACACACGGTCAGACTAGCATCAGTCCGTAATGCT ...

Reference
genome

... CACGGTCAGACGAGCATCAGTCC ...

...CACACGGTCAGACGAGCATCAGT ...

...GGTCAGACGAGCATCAGTCCGTA ...

...CAGACTAGCATCAGTCCGTAATG ...

...CACACGGTCAGACTAGCATCAGT ...

...GGTCAGACTAGCATCAGACCGTA ...

...GGTCAGACTAGCATCAGTCCGTA ...

...GTCAGACTAGCATCAGTCCGTAA ...

Resequencing
reads from one
individual

G/T polymorphism

Illumina Infinium® Poplar 34K Genotyping Array

Resequencing of **36** *P. trichocarpa* genomes and xylem transcriptomes

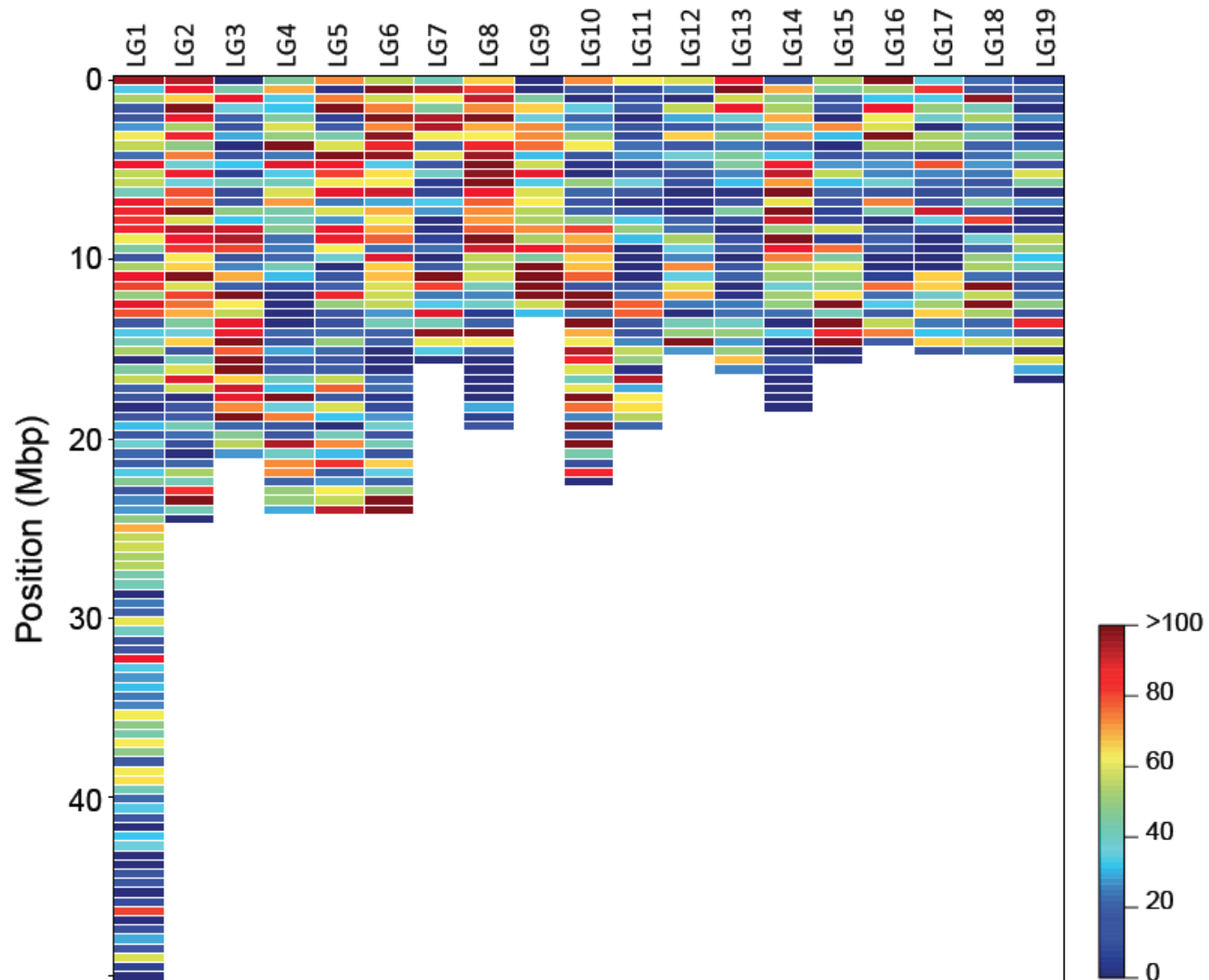
SNP database (~**3M** SNPs) *UBC and BioEnergy Sciences Center (BESC)*

34,000-SNP array developed by Illumina in collaboration with *UBC and BESC*)

SNPs are in **3500 candidate genes** selected for potential involvement in wood chemistry, adaptation, and biomass traits

29,000 SNPs successfully genotyped in **450** individuals from BC MOF association population with minimal missing data

SNP density across the genome

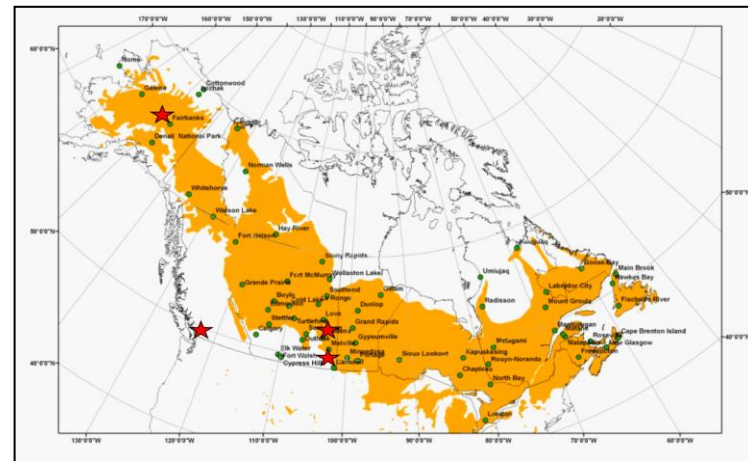


Large-scale whole genome re-sequencing captures total genetic diversity

P. trichocarpa: BC MOF/Totem Field

P. balsamifera: AgCanBaP

Populus trichocarpa



Target genomes: **450**
Completed genomes: **298**
(30X or 15X coverage)

450
178

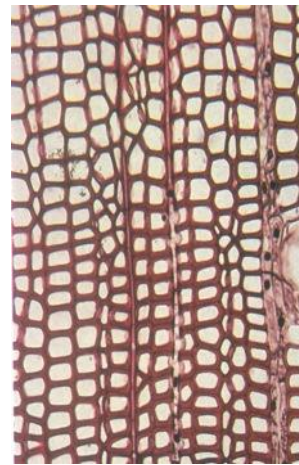
Genomes aligned to reference and SNPs called to date: **460**
Average SNPs/gene called in 98% of individuals sequenced : **~160**
Genes with SNP coverage in 98% of individuals sequenced: **>99.6%**

Analysis of genetic variation underlying trait variation

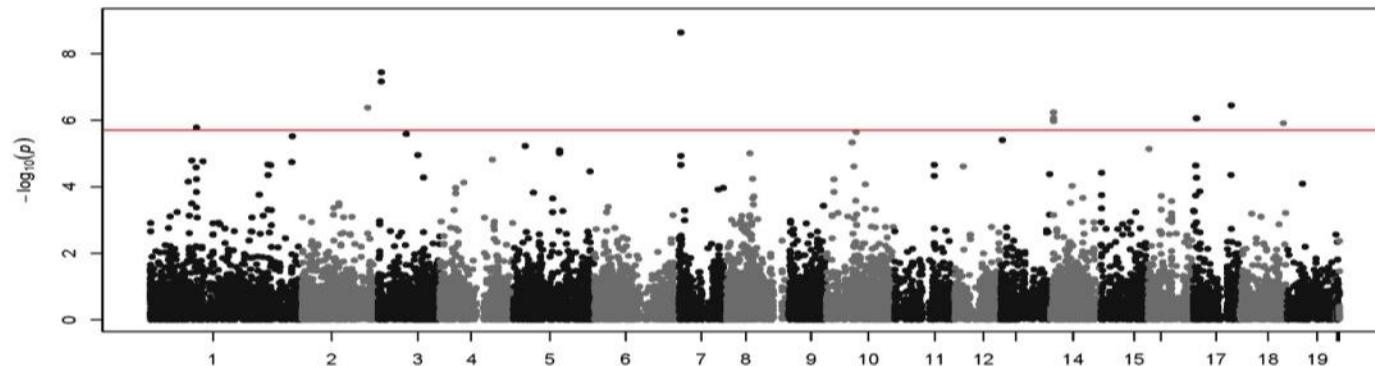
Genotype



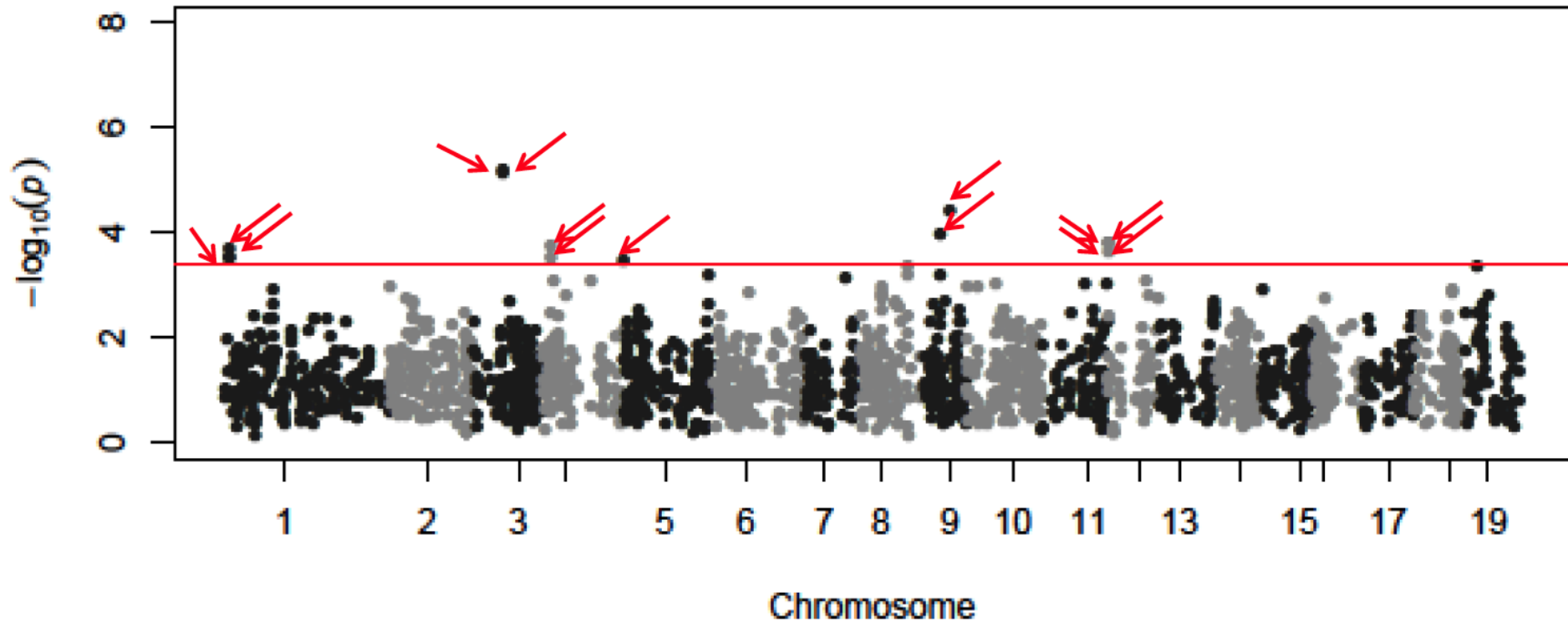
GWAS



Phenotype



34K SNP array: Allelic associations with wood density

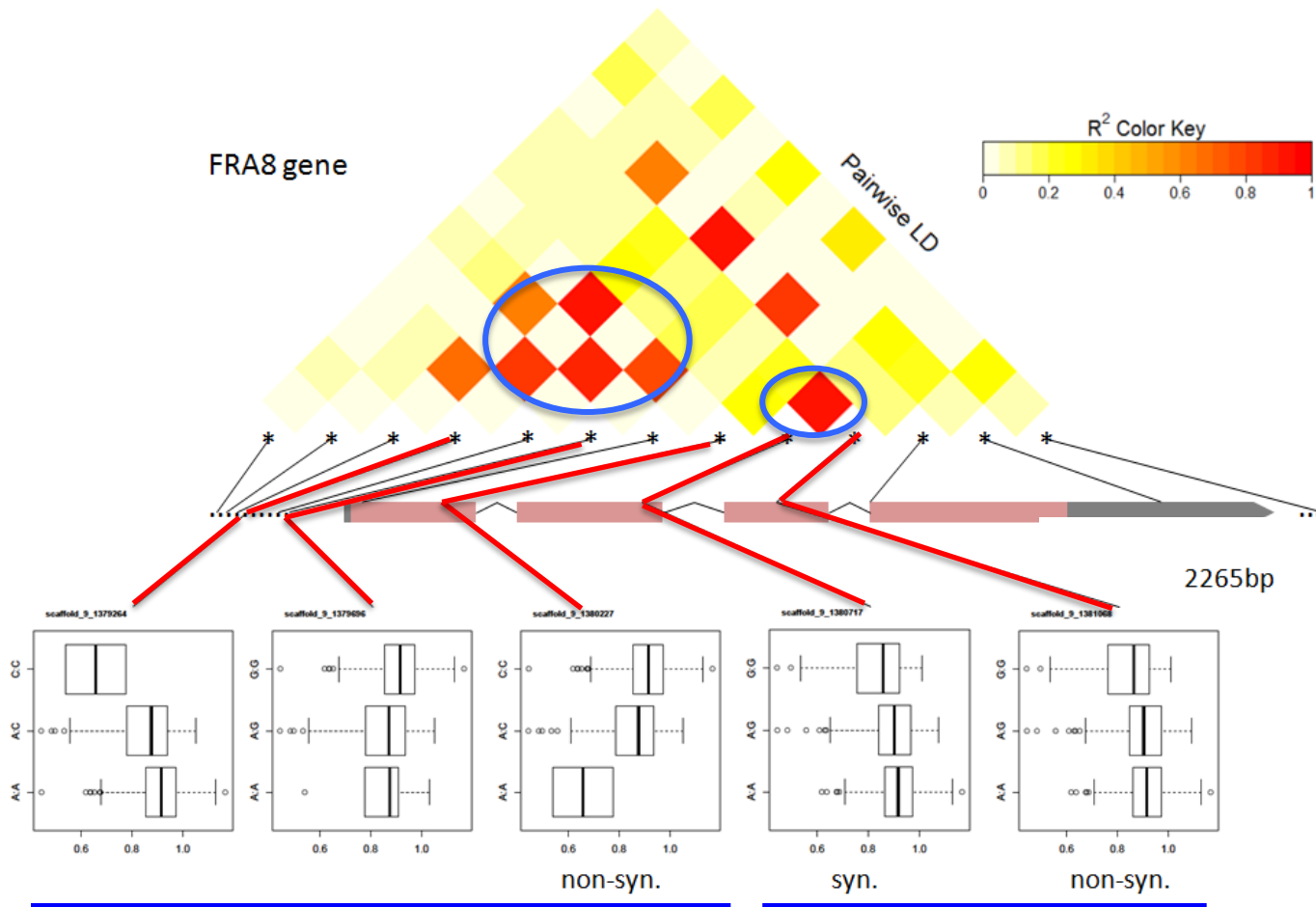


34K SNP array: summary of wood phenotype-SNP associations

Trait	N	% var
Alpha cellulose	12	29.5
Arabinose	3	10.9
Average wood density	14	39.3
Crystallinity	8	25.9
Fibre LW	11	29.1
Galactose	-	-
Glucose	7	24.3
Hemicellulose	9	23.3
Holocellulose	15	45.0
Insoluble lignin	8	29.4
Mannose	12	48.0
MFA1	7	13.7
MFA2	12	42.4
Soluble lignin	3	13.9
Syringyl lignin	12	52.8
Total lignin	6	23.3
Xylose	2	8.5
Total	141	

141 SNPs significantly associated with 16 traits ($\alpha < 0.05$)

Example: Allelic variants of *Fragile Fiber8* gene are associated with fiber length



FRA8 encodes a glycosyl-transferase involved in xylan (hemicellulose) biosynthesis

Heritabilities for 18 wood traits

wood trait	h^2	SE
Arabinose	0.7008	0.0794
Rhamnose	0.7257	0.0853
Galactose	0.5502	0.1036
Glucose	0.3454	0.1195
Xylose	0.5371	0.1011
Mannose	0.522	0.13
Insollignin	0.4285	0.1191
Sollignin	0.7353	0.0642
Totlignin	0.461	0.1155
Syringylligninmonomers	0.6578	0.0952
Holocellulose	0.2966	0.1684
Alphacellulose	0.4943	0.137
Hemicellulose	0.4139	0.1382
FiberlengthLw	0.6004	0.1292
MFA (recent ring)	0.3567	0.1357
MFA (first ring)	0.347	0.1299
Averwooddensity	0.6641	0.0977
PctCrystallinity	0.28	0.1589

Summary and future work

- Phenotype and genotype diversity in *P. trichocarpa* association population is large and provides a foundation for genetic improvement
- 141 significant SNP associations with wood traits relevant to biofuel and solid wood applications
- >190 SNP associations with adaptive/biomass traits (not shown)
- Similar studies in *P. balsamifera* underway
- Moving to genome-wide scans of allelic variation by whole genome re-sequencing to reveal more definitive picture of genetic architecture underlying trait variation
- Allelic variants identified that could be used to significantly improve biomass and conversion traits important for improved feedstocks
- Testing of genetic gain by introduction of selected alleles into hybrid poplar breeding programs

AGIP/POPCAN



AGIP/POPCAN team

Carl Douglas
Shawn Mansfield
Quentin Cronk
Jürgen Ehling
Yousry El-Kassaby
Rob Guy
Malcolm Campbell
Peter Constabel
Richard Hamelin
Marty Luckert
Tom Maness

Shofiul Azam
Hua Bao
Michael Friedmann
Miki Fujita
Armando Gerales
Jan Hannemann
Catalin Ritea
Charles Hefer
Peter Kalynyak
Jaroslav Klapste
Jon LaMantia
Eryang Li
Athena McKown
Ilga Porth
Alex Skyba

Natalia Kolosova
Nima Farzaneh
Shucaï Wang
Daisie Huang
Steffi Fritsche
Arnaud Capron
Mike Barker
Matt King
Heather Ramsay
Chris Grassa



Genome Sciences Centre

Inanc Birol
Reza Falsafi
Steve Jones
Marco Marra
Johnson Pang
Nina Thiessen
Yongjun Zhao

Collaborators

ORNL/BESC – Jerry Tuskan, Steve DiFazio,
Gancho Slavov, Lee Gunter et al.
UPSC/POPLARENERGY – Rishi Bhalerao,
Stefan Jansson
USFS Forest Products Lab - Dan Cullen
BC Ministry of Forests and Range - Alvin Yanchuk
Greenwood Resources - Brian Stanton
Kruger Inc. – Dan Carson
University of Pretoria – Zander Myburg

Applied Genomics Innovation Program

Large Scale Applied Research Program