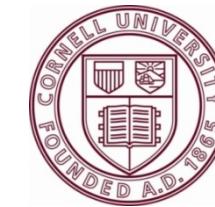




Next generation sequencing of the *Salix purpurea* genome and transcriptome: Tools for the genetic improvement of willow biomass crops

Larry Smart, Michelle Serapiglia, Fred Gouker

Cornell University, NYS Ag Experiment Station, Geneva, NY



Chris Town and Haibao Tang

J. Craig Venter Institute, Rockville, MD



Jerry Tuskan, Kerrie Barry, and Erika Lindquist

DOE Joint Genome Institute, Walnut Creek, CA



Ed Buckler, Sharon Mitchell, and Rob Elshire

USDA-ARS, Institute for Genomic Diversity, Cornell University



Steve DiFazio and Eli Rodgers-Melnick

West Virginia University, Morgantown, WV



lbs33@cornell.edu

www.hort.cornell.edu/smart/

 **WILLOWPEDIA**
<http://willow.cals.cornell.edu/>



Shrub willow (*Salix* spp.) bioenergy crops



Shrub willow biomass is a feedstock for biopower, liquid biofuels, and heat



McNeil Generating Station, Burlington, VT
Photo: US DOE



VeraSun Station, Brookings, SD



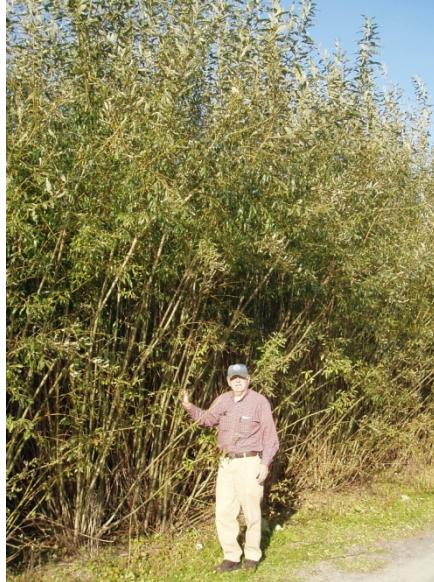
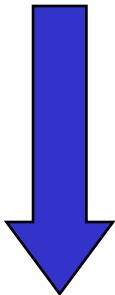
Harman Wood Pellet Stove

Willow Biomass Production Cycle

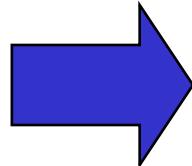
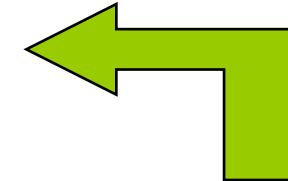
Site preparation



Planting



Three years old
after coppice



First-year growth
Winter coppice

Regrowth after coppice



One year old after
coppice

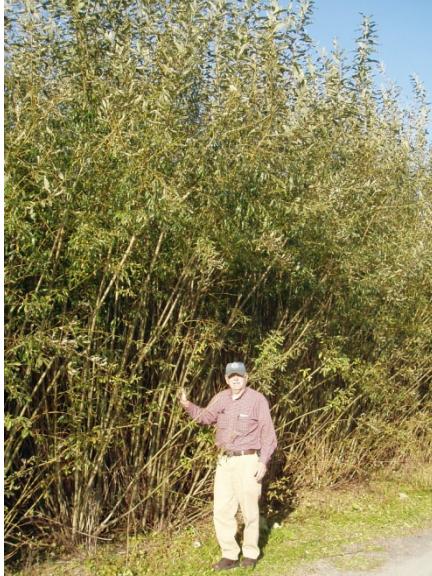


Willow Biomass Production Cycle

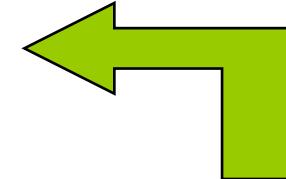
Winter harvest



Fertilizer: ~100 lbs N
per acre after every
harvest (every 3 years)



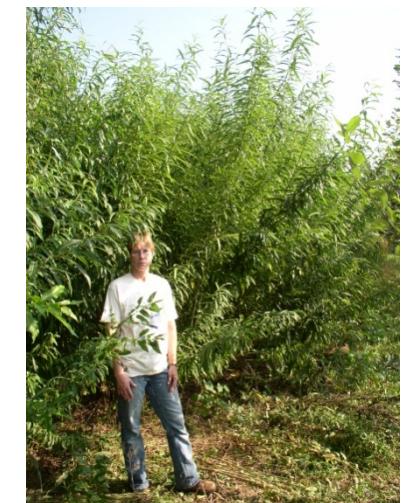
Three years old
after coppice



At least 7
harvests



Regrowth after coppice

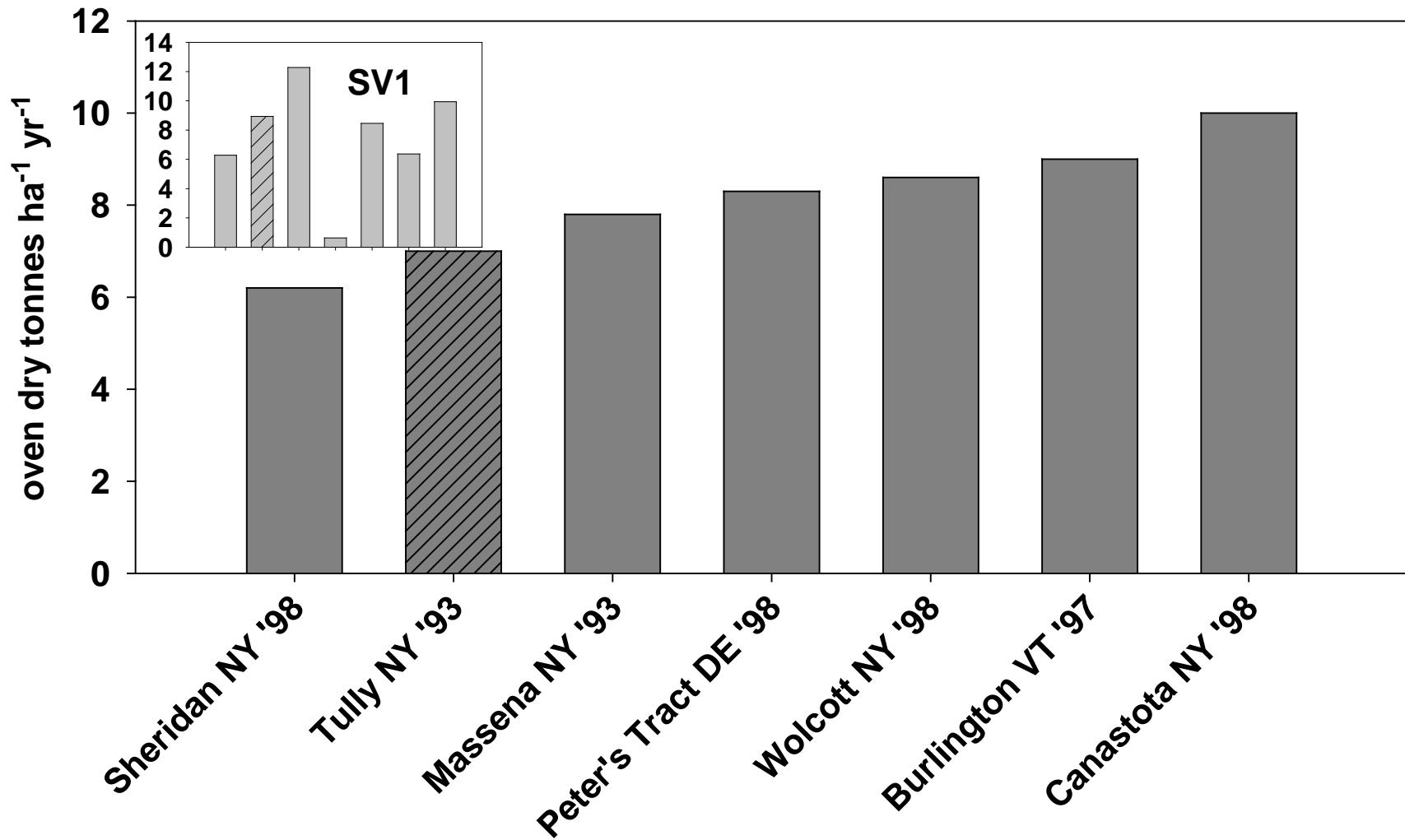


One year old after
coppice



Mean Yield of Top Five Clones in Yield Trials

- Selections from Toronto breeding and natural collections
- First harvest – three-years post-coppice

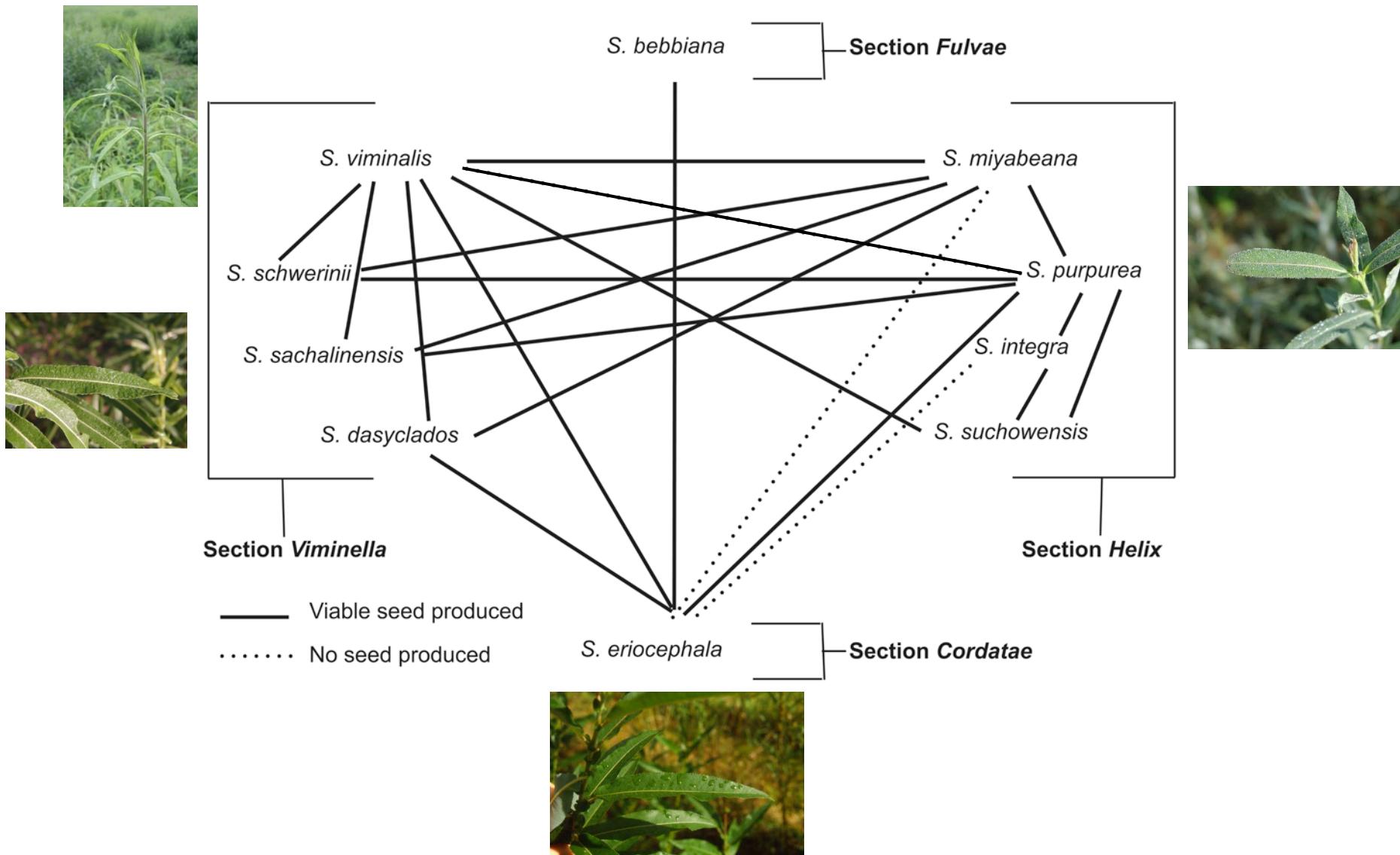


Clones in yield trials: 1993-1998

Clone	No. sites	Species	Origin ^a
S19	6	<i>Salix eriocephala</i>	U of T
S25	10	<i>S. eriocephala</i>	U of T
S34	1	<i>S. eriocephala</i>	U of T
S71	2	<i>S. petiolaris x eriocephala</i>	U of T
S185	2	<i>S. eriocephala</i>	U of T
S287	2	<i>S. eriocephala</i>	U of T
S301	11	<i>S. interior x eriocephala</i>	U of T
S365	10	<i>S. discolor</i>	U of T
S546	5	<i>S. eriocephala</i>	U of T
S557	2	<i>S. eriocephala</i>	U of T
S566	3	<i>S. eriocephala</i>	U of T
S599	2	<i>S. eriocephala x petiolaris</i>	U of T
S625	5	<i>S. eriocephala x interior</i>	U of T
S646	5	<i>S. eriocephala</i>	U of T
S652	3	<i>S. eriocephala</i>	U of T
SA2	7	<i>S. alba</i>	OMNR
SH3	6	<i>S. purpurea</i>	OMNR
SP3	1	<i>S. purpurea</i>	OMNR
SV1	11	<i>S. dasyclados</i>	OMNR
SX61	7	<i>S. sachalinensis</i>	U of T
SX64	5	<i>S. miyabeana</i>	U of T
SX67	5	<i>S. miyabeana</i>	U of T
NM6	9	<i>Populus maximowiczii x nigra</i>	OMNR
NM5	1	<i>P. maximowiczii x nigra</i>	OMNR
PUR12	3	<i>S. purpurea</i>	U of T
PUR34	5	<i>S. purpurea</i>	U of T
94001 (FC185)	3	<i>S. purpurea</i>	SUNY-ESF
94004 (FC188)	3	<i>S. purpurea</i>	SUNY-ESF
94005 (FC189)	2	<i>S. purpurea</i>	SUNY-ESF
94006 (FC190)	4	<i>S. purpurea</i>	SUNY-ESF
94009 (B193)	3	<i>S. purpurea</i>	SUNY-ESF
94011 (B195)	2	<i>S. purpurea</i>	SUNY-ESF

Crosses attempted to capture hybrid vigor

Compilation of results from US, Canadian, and European breeding programs



Intra- and Inter-specific Hybridizations



F_1

- S. integra*
- S. viminalis*
- S. cordata*
- S. nigra*
- S. sachalinensis*
- S. cordata x S. eriocephala*
- S. purpurea x S. eriocephala*
- S. purpurea x S. viminalis*
- S. purpurea x S. sachalinensis*
- S. purpurea x S. gilgiana*
- S. viminalis x S. miyabeana*
- S. viminalis x S. eriocephala*
- S. x dasyclados x S. miyabeana*
- S. x dasyclados x S. eriocephala*
- S. x dasyclados x S. viminalis*
- S. eriocephala x S. purpurea*
- S. matsudana x S. alba*
- S. sachalinensis x S. eriocephala*
- S. sericea x S. purpurea*
- S. sericea x S. eriocephala*
- S. sericea x S. sachalinensis*
- S. discolor x S. cinerea*
- S. discolor x S. eriocephala*

F_1 and F_2

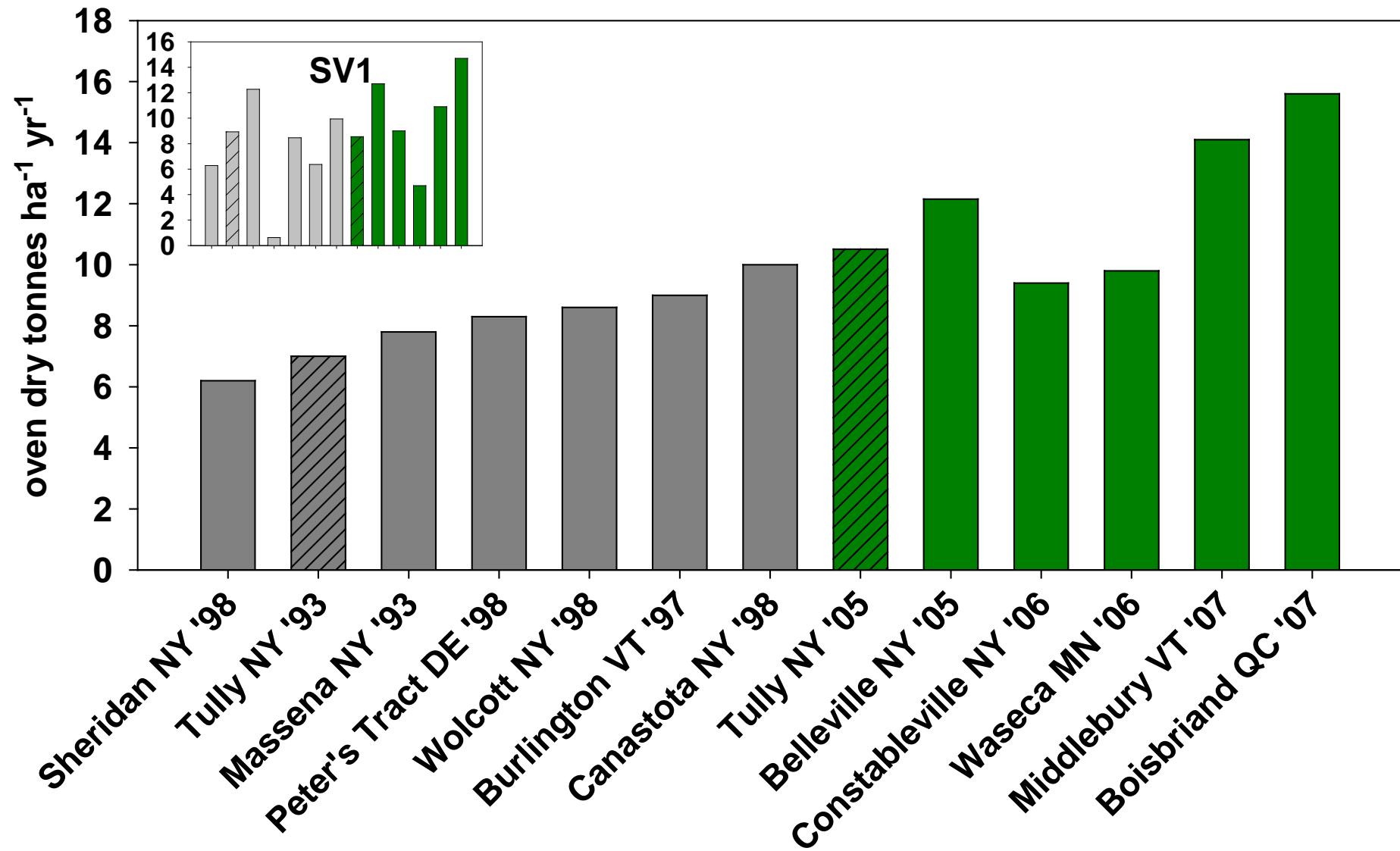
- S. purpurea*
- S. eriocephala*
- S. miyabeana*
- S. sachalinensis x S. miyabeana*
- S. purpurea x S. integra*

Multi-species hybrids

- S. viminalis x (S. sachalinensis x S. miyabeana)*
- S. viminalis x (S. purpurea x S. miyabeana)*
- S. viminalis x (S. viminalis x S. miyabeana)*
- S. purpurea x (S. sachalinensis x S. miyabeana)*
- S. purpurea x (S. purpurea x S. miyabeana)*
- S. purpurea x (S. viminalis x S. miyabeana)*
- (S. sachalinensis x S. miyabeana) x S. purpurea*
- (S. sachalinensis x S. miyabeana) x S. miyabeana*
- (S. sachalinensis x S. miyabeana) x S. viminalis*
- (S. viminalis x S. schwerinii) x S. cinerea*
- (S. sachalinensis x S. miyabeana) x (S. purpurea x S. miyabeana)*
- (S. sach x S. miya) x (S. viminalis x (S. schwerinii x S. viminalis))*

Mean Yield of Top Five Clones in Yield Trials

- New varieties contribute to 38% greater yield overall



Clones in yield trials: 2005-2007

Dennis Rak
Double A Willow



Photo by Larry Abrahamson

<http://www.doubleawillow.com>

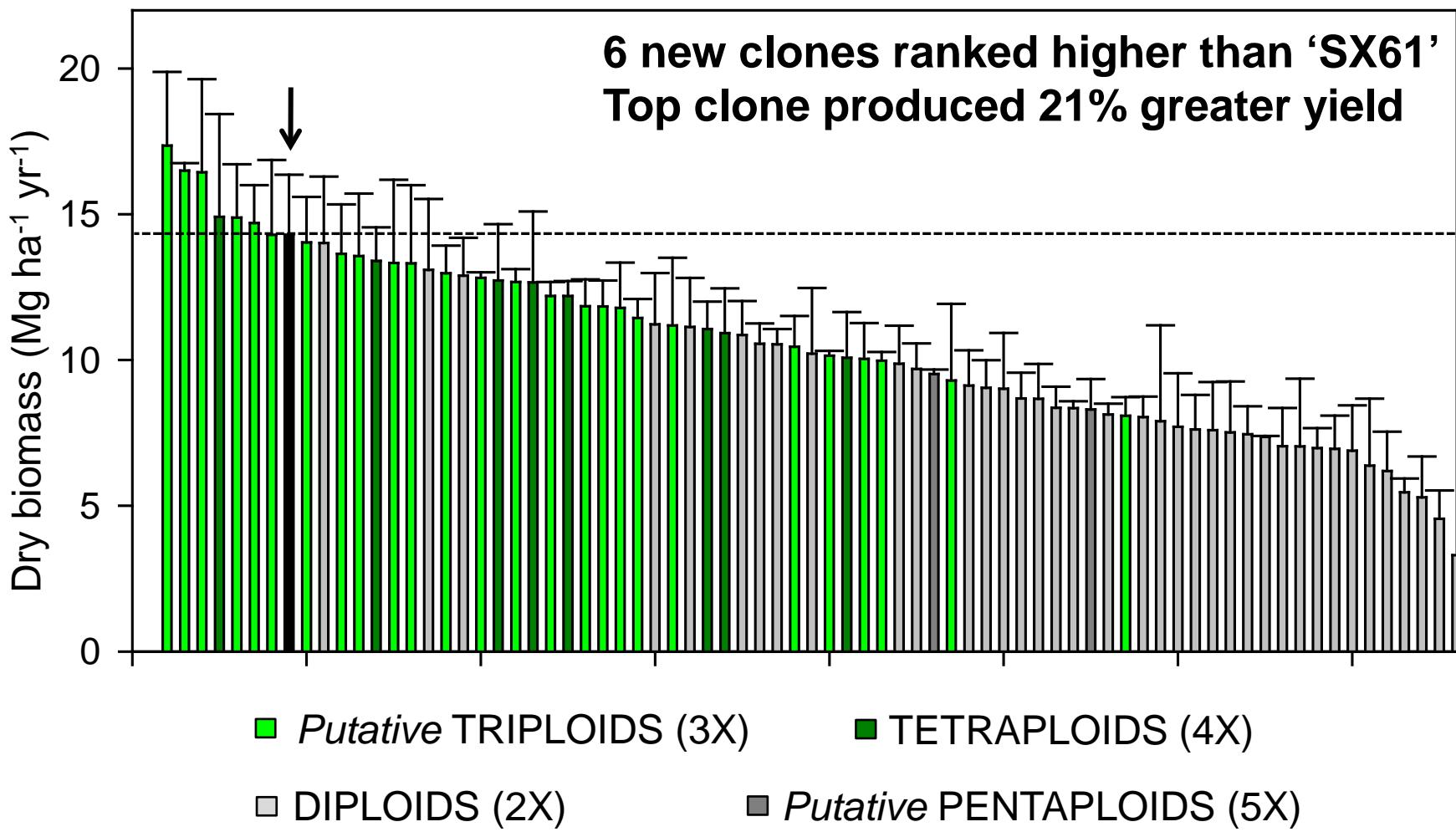
Diversity Group	Species	Variety
1	<i>Salix x dasyclados</i>	SV1
2	<i>S. sachalinensis</i>	SX61
3	<i>S. caprea hybrid</i>	S365
4	<i>S. eriocephala</i>	S25
5	<i>S. miyabeana</i>	SX64
		SX67
6	<i>S. purpurea</i>	Fish Creek
		Onondaga
		Allegany
7	<i>S. sachalinensis x S. miyabeana</i>	Sherburne
		Canastota
8	<i>S. viminalis x S. miyabeana</i>	Tully
		Champion
		Fabius
		Owasco
		Otisco
9	<i>S. purpurea x S. miyabeana</i>	Oneida
		Millbrook

Flow Cytometric Estimation of Nuclear DNA Content

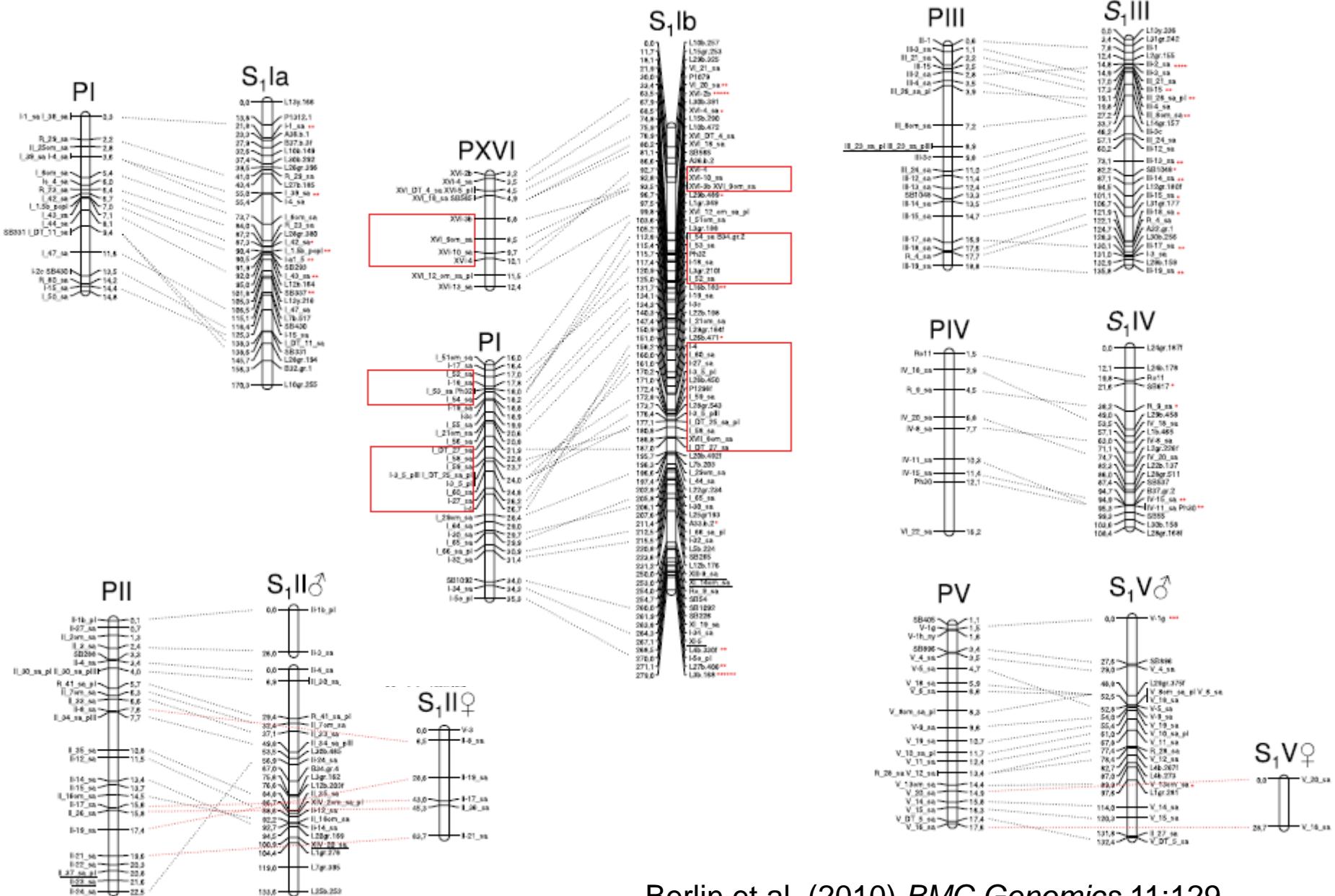
Clone	Species	(pg/2C)
Diploid		
Björn	<i>S. schwerinii</i> x <i>S. viminalis</i>	0.76
Olof	<i>S. viminalis</i> x (<i>S. schwerinii</i> x <i>S. viminalis</i>)	0.80
Jorr	<i>S. viminalis</i>	0.83
P63	<i>S. integra</i>	0.93
94006	<i>S. purpurea</i>	0.94
SV1	<i>S. x dasyclados</i>	0.94
Triploid		
Nimrod	(<i>S. schwerinii</i> x <i>S. viminalis</i>) x <i>S. udensis</i>	1.15
Terra Nova	(<i>S. triandra</i> x <i>S. viminalis</i>) x <i>S. udensis</i>	1.11
Tully Champion	<i>S. viminalis</i> x <i>S. miyabeana</i>	1.31
Oneida	<i>S. purpurea</i> x <i>S. miyabeana</i>	1.33
01X-266-001	<i>S. viminalis</i> x (<i>S. viminalis</i> x <i>S. miyabeana</i>)	1.26
Tetraploid		
SX64	<i>S. miyabeana</i>	1.65
SX61	<i>S. sachalinensis</i> (<i>S. udensis</i>)	1.65
Canastota	<i>S. sachalinensis</i> x <i>S. miyabeana</i>	1.67
05X-284-001	<i>S. purpurea</i> x (<i>S. purpurea</i> x <i>S. miyabeana</i>)	1.87
Pentaploid?		
05X-286-001	<i>S. miyabeana</i> x (<i>S. purpurea</i> x <i>S. miyabeana</i>)	2.17

2008 Genetic Selection Trial - Geneva, NY

- 24-plant plots, 3 replicates, 76 clones
- Biomass harvested from middle 8 plants in Dec. 2011

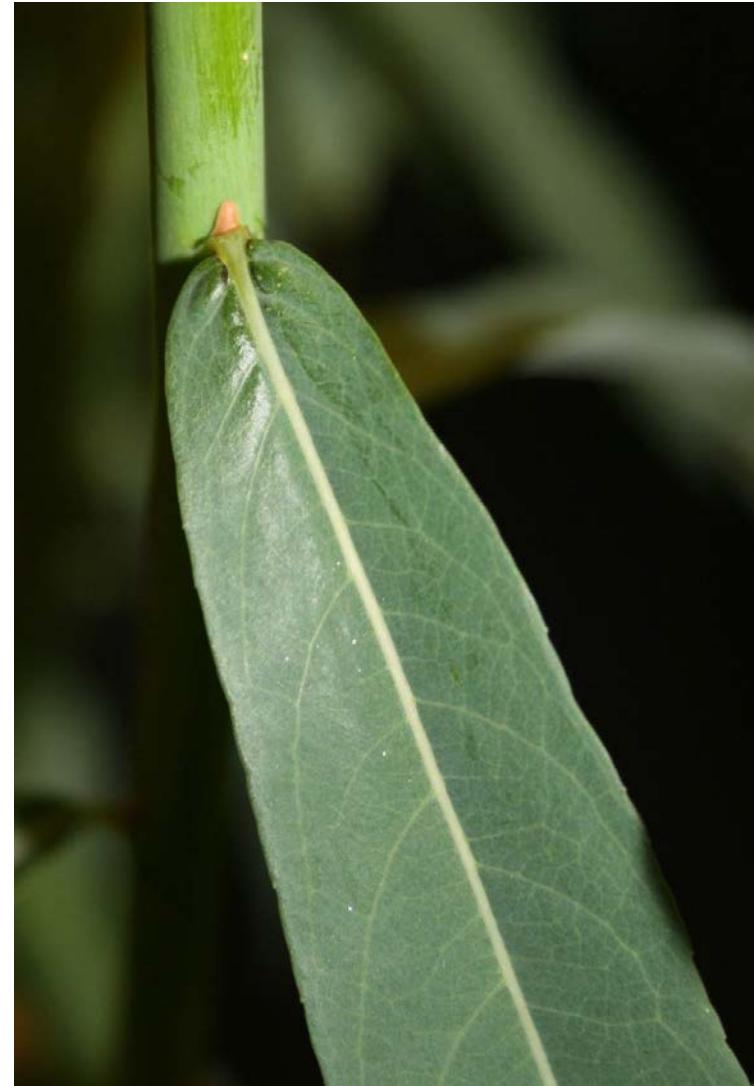


Macrosynteny between *Populus* and *Salix*



Sequencing of the *Salix purpurea* L. genome

- Whole genome shotgun sequencing of clone ID 94006
 - diploid female ($n = 19$, ~485 Mb)
 - underway at DOE JGI
- (PIs: Tuskan, Smart, Town)
- Genomic data to date:
 - 900 Mbp 454 standard
 - 19 Gb Illumina 76 bp standard
 - 3 Gb of 76 bp Illumina from 7 kb-insert jumping library
 - Waiting for ~20 Gb of Illumina 150bp from fragment libraries (250bp, 500bp, 4 kb)
- Two 5X BAC libraries constructed (third on the way)



Draft de novo genome assembly

- Using CLC de novo assembly tool

Number of assembled sequences 290,871

Residue counts:

Number of N's	574791	0.18 %
Total	327,266,967	(327 Mbp)

Sequence lengths:

Minimum	200
Maximum	190101
Average	1125
N50	2224

(50% of bases are in contigs <2224 bp, *this is too small*)

- This is a quick and dirty assembly. We will try to reassemble using ALLPATHS-LG and/or Celera Assembler when we have additional Illumina data

Comparison to poplar genome

- Willow assembled size: 327 Mb (real bases)
- Poplar assembled size: 404 Mb (real bases)
- Compared willow assembly against poplar genome using LASTZ (like BLAST, nucl based)

*65% of the *P. trichocarpa* genome (261.9 Mb) aligned to the draft *S. purpurea* genomic assembly:*

- sequence identity: **88%** (coding and non-coding)

Transcriptome sequencing

- *Salix purpurea* transcriptome data to date:
 - 5 lanes Illumina 2x76 from 8 tissue libraries done at JGI



Shoot tip
44.0 M reads



Leaves
Day: 16.4 M
Nite: 15.3 M
Drought: 13.6



Stem
Node: 18 M reads
Internode: 21.5 M



Roots
51.7 M reads



Catkins
46.1 M reads

- *Salix sachalinensis* 'SX61': stem (8.3 M), shoot tip (7.7 M)
- *Salix miyabeana* 'SX64': stem (6.2 M), shoot tip (7.4 M)
 - these are tetraploid, sequencing done at JCVI

Salix purpurea RNA-Seq assembly

- All reads from the same genotype were pooled and assembled for *S. purpurea* 94006, *S. sachalinensis* SX61, and *S. miyabeana* SX64.
- Transcript assembly using CLC assembler, for *S. purpurea*
Number of assembled sequences 201,926

Residue counts:

Total 79,555,995 (80 Mb)

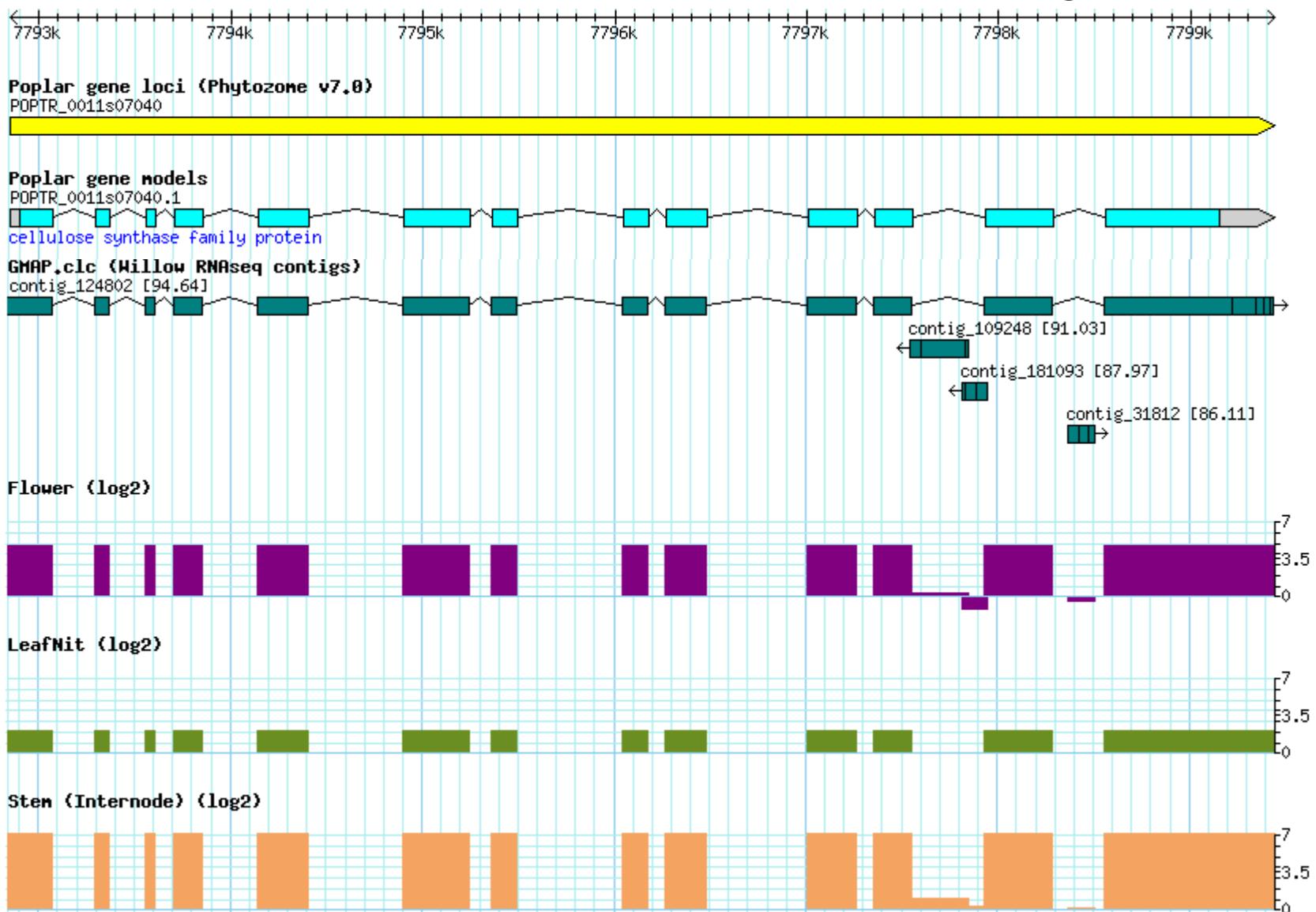
Sequence lengths:

Minimum	100
Maximum	15400
Average	394
N50	871

- 80.5% of the poplar transcripts have at least half of their lengths aligned to the willow transcripts
- *S. purpurea* coding sequence vs. poplar: 92% (genomic: 88%)
- *S. purpurea* coding seq vs. *S. sach* or *S. miya*: 98%

Salix purpurea transcriptome: GBrowse

Salix purpurea de novo transcript assembly aligned to
Populus trichocarpa v.2.0 as the reference genome



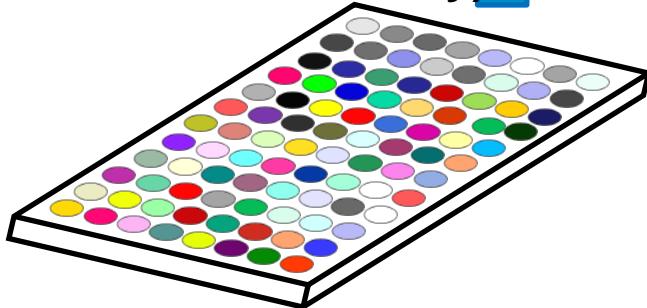
High-density markers: GBS 96-plex protocol

Rob Elshire, Sharon Mitchell, Ed Buckler

Elshire et al. (2011) PLoS One

Digest DNA with
EcoT22I (6-base) or
ApeKI (5 base-cutter)

Ligate barcoded
adapters (may be
done simultaneously)

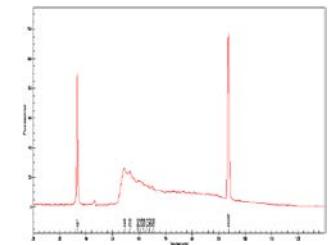


Pool
DNAs

PCR

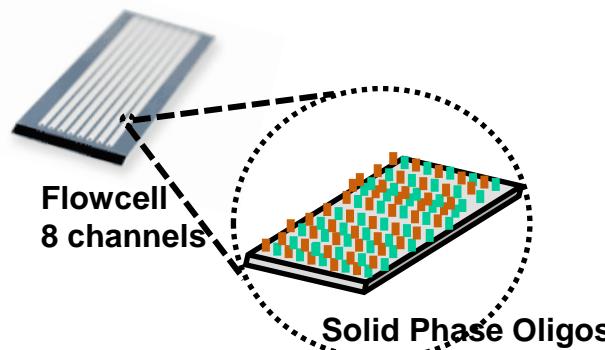
Primers

Clean-up



Optimize
fragment sizes
& adapter
ligation

Sequence on Illumina Platform
~200,000 reads per sample
75,000-100,000 SNPs



Next generation sequencing of the *Salix purpurea* genome and transcriptome: Tools for the genetic improvement of willow biomass crops

Summary

- Excellent hybrid pedigrees have been produced and new hybrid combinations are being explored
- Each round of breeding has generated cultivars with ~20% improved yields over previous standards
- Triploid progeny have shown hybrid vigor for yield and low fertility
- A draft sequence of the *Salix purpurea* genome is nearly complete
- High density GBS markers are being developed

Thanks to...

J. Craig VenterTM
INSTITUTE

