

Physical Mapping of Chestnut Genome: A Genome Resource for Fagaceae and Other Tree Species

**Eric Fang, Barbara Blackmon, Meg Staton, David Henry,
Megan Monsanto, Stephen Ficklin, Chun-Huai Cheng,
Laura Georgi, Jeff Tomkins, and Albert Abbott**

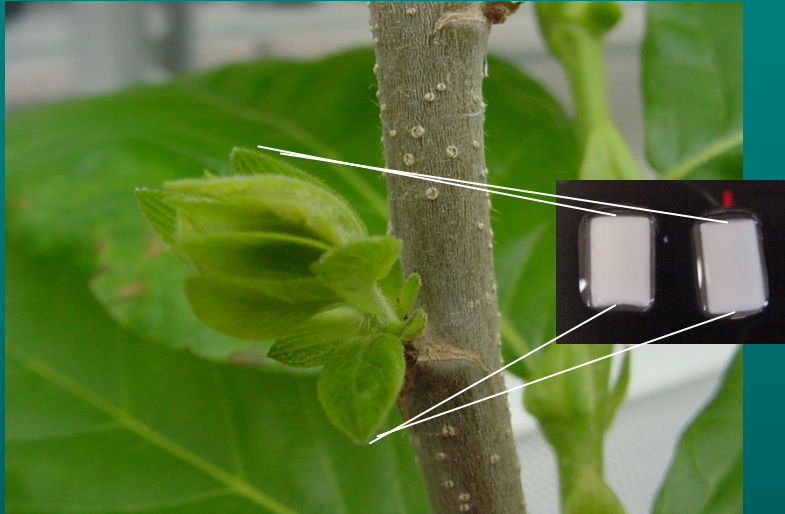
Chestnut as a Model Plant for Fagaceae Tree Species

- Reasonable genome size (800 Mbp/1C)
- Diploid inheritance ($n = 12$)
- Fast growth rate and valuable wood traits
- Available transformation and embryogenesis procedure
- Compatible linkage groups and genetic markers with other Fagaceae plants

Application of the Chestnut Genome Resource

- Marker identification
 - molecular markers for ID chestnut species
 - blight resistance
 - *Phytophthora* root rot resistance
 - wood quality, growth rate....
- Comparative genomics analysis
 - American chestnut vs. Chinese chestnut and other chestnut species
 - chestnut vs. Fagaceae species and other woody plants

Construction of BAC Libraries



High MW gDNA extraction

Partial restriction digest

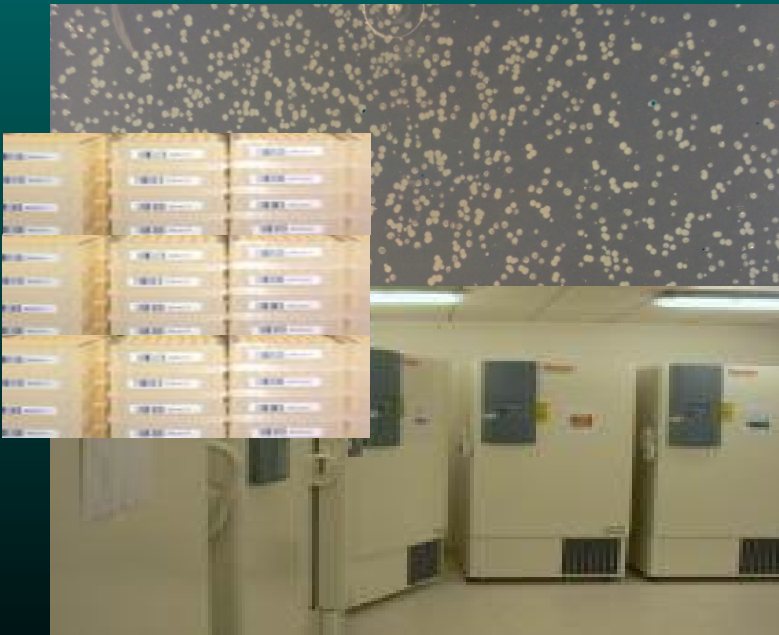
Two rounds of size selection
on CHEF system

Elute DNA and ligate to
BAC vector

Transformation

QC

Mass-transformation, colony picking,
duplication, and storage



Chestnut BAC Libraries

	Library			
	CMCMBb ^a	CMCMBd ^a	CM_MBc ^a	CDC_Ba ^b
Restriction digest	<i>Hind</i> III	<i>Eco</i> RI	<i>Hind</i> III	<i>Hind</i> III
Average insert size	123 kb	115 kb	90 kb	140 kb
# of clones	73,728	92,160	110,592	73,728
Genome equivalents	11X	10X	12X	12X
Filters	4	5	6	4

^aCMCMBb, CMCMBd and CM_MBc libraries were made from Chinese chestnut, *Castanea mollissima* cultivar Vanuxem.

^bThe CDC_Ba library was made from American Chestnut, *C. dentata*.

Construction of A Chestnut Physical Map Using **HICF** Technique

HICF

(A) main components

- SNaPshot labeling
- Capillary electrophoresis

(B) Procedure

1. DNA extracted from BAC clones by 96-well filter plate miniprep.
2. 5-enzyme restriction digest of BAC DNA.
3. Digested DNA labeled with 4 fluorescent dyes.
4. Labeled DNA samples resolved on capillary electrophoresis 3730 DNA analyzer.

Pipeline for Data Processing/Analysis for Chestnut Physical Mapping

3730 Data Collection



GeneMapper



GenoProfiler



FPC

Chestnut FPC v.9.1

- 125,459 clones input,
- 18X genome equivalents,
- Initial setting: **Tolerance: 3**
Cutoff : 1e-60
- Final setting: **Tolerance: 3**
Cutoff: 1e-30
- **4,345** contigs, **17,132** singletons
- **1,104,099** bands for MTP
(length = 1,270 Mb. 1.6X genome size)

Project CMCMBb_d_1e-30_ends1e-45break Clones 125459 Seq 0 Markers 4579 ze... Help

Contigs 4345 Singles 17132

Max 3324, 1312 (>25), 1275 (25:10), 1455 (9:3), 303 (2)

Chr_Remark Search Summary

Contig	Clone	Marker	Seq	Date	Status	Qs	Chr_Remark
151	170	2	-	11sep09 11:20	Ok	~ 9	
208	167	2	-	11sep09 11:20	Ok	~ 5	
4	166	3	-	11sep09 11:20	Ok	~ 8	
178	166	2	-	11sep09 11:20	Ok	~ 7	
185	166	4	-	7sep09 12:46	Ok	~ 5	
13	165	2	-	7sep09 13:08	Ok	~ 2	
81	165	1	-	11sep09 11:20	Ok	~ 4	
425	165	-	-	7sep09 12:45	Ok	~ 1	
1180	164	3	-	11sep09 11:20	Ok	~ 0	
266	162	6	-	11sep09 11:20	Ok	~ 3	
66	161	3	-	7sep09 12:44	Ok	~ 3	
60	159	2	-	11sep09 11:20	Ok	~ 7	
73	157	1	-	11sep09 11:20	Ok	~ 6	
211	156	1	-	11sep09 11:20	Ok	~ 6	
55	155	2	-	11sep09 11:20	Ok	~ 7	
139	154	1	-	11sep09 11:20	Ok	~ 5	
270	148	2	-	11sep09 11:20	Ok	~ 3	
80	143	-	-	11sep09 11:20	Ok	~ 7	
17	142	2	-	11sep09 11:20	Ok	~ 3	
8723	142	3	-	7sep09 12:00	Ok	~ 3	
726	140	-	-	11sep09 11:20	Ok	~ 3	
12	139	3	-	7sep09 12:29	Ok	~ 1	
240	139	2	-	7sep09 12:50	Ok	~ 6	
980	139	-	-	11sep09 11:20	Ok	~ 4	
5694	136	2	-	7sep09 13:08	Ok	~ 6	
757	135	4	-	11sep09 11:20	Ok	~ 0	
82	134	2	-	7sep09 12:33	Ok	~11	
1341	134	1	-	11sep09 11:20	Ok	~ 2	
1879	131	2	-	7sep09 13:02	Ok	~ 1	
32	130	4	-	11sep09 11:20	Ok	~ 5	
1042	130	1	-	11sep09 11:20	Ok	~ 2	
383	129	3	-	11sep09 11:20	Ok	~ 4	
100	128	1	-	11sep09 11:20	Ok	~ 6	
365	128	2	-	11sep09 11:20	Ok	~ 5	
148	127	1	-	11sep09 11:20	Ok	~ 3	
169	127	2	-	7sep09 12:22	Ok	~ 2	
782	127	1	-	7sep09 12:58	Ok	~ 6	
149	126	3	-	7sep09 11:29	Ok	~ 4	
815	126	1	-	11sep09 11:20	Ok	~ 5	
4210	125	-	-	11sep09 11:20	Ok	~ 5	
777	124	2	-	7sep09 12:33	Ok	~ 4	
409	122	2	-	11sep09 11:20	Ok	~ 2	
663	122	-	-	7sep09 11:34	Ok	~ 4	
772	122	3	-	7sep09 12:11	Ok	~ 2	
224	121	2	-	11sep09 11:20	Ok	~ 1	
557	121	4	-	7sep09 13:02	Ok	~ 2	
101	120	2	-	11sep09 11:20	Ok	~ 5	
192	120	-	-	7sep09 12:43	Ok	~ 9	
299	120	3	-	11sep09 11:20	Ok	~ 5	
130	119	4	-	11sep09 11:20	Ok	~ 6	
399	119	2	-	7sep09 12:41	Ok	~ 0	
436	119	2	-	11sep09 11:20	Ok	~ 5	
128	118	1	-	7sep09 12:41	Ok	~ 2	

Integration of the Physical Map and Genetic Map by Marker Hybridization

Goals

- Verify the contig build of the physical map
- Construct a sequence-ready map for map-based cloning, comparative genomics studies....

Approach

- Overgo design
- Pool hybridization
- HybSweeper II
- FPC

High-throughput Overgo Hybridization

EST database

Markers

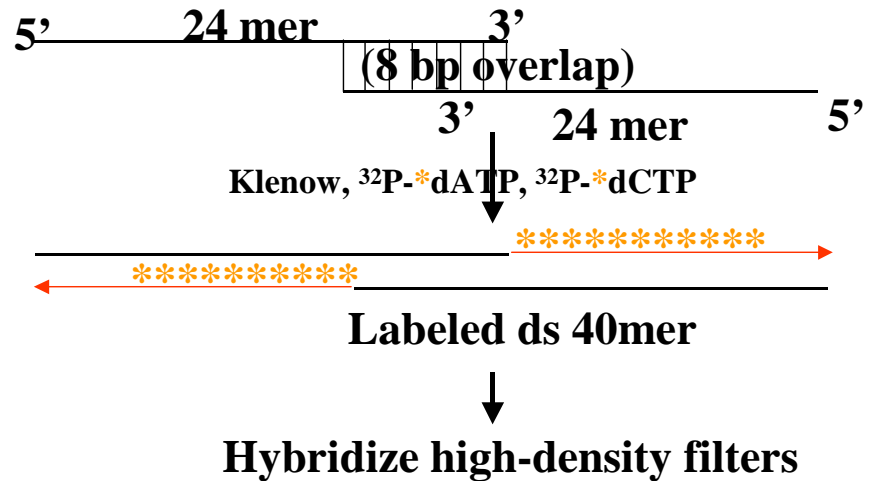
- SSRs, + RFLP markers
- SNPs

Repeat Masker,
Cross Match
Tandem Repeat Finder

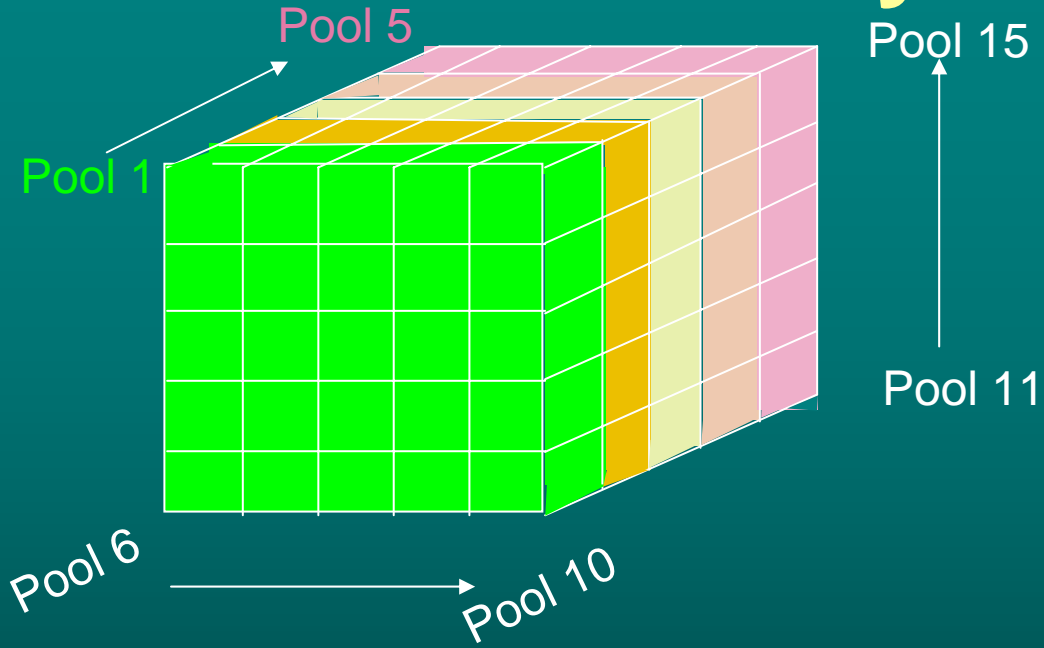
Overgo Maker
OligoSpawn

Overgo probes

Overgo Hybridization



Pool Hybridization



POOL DESIGN
 (15 hyb for 125 probes)
 - 125 probes in 15 pools,
 - 25 probes/pool

21	16	11	6	1
22	17	12	7	2
23	18	13	8	3
24	19	14	9	4
25	20	15	10	5

Pool 1

46	41	36	31	26
47	42	37	32	27
48	43	38	33	28
49	44	39	34	29
50	45	40	35	30

Pool 2

71	66	61	56	51
72	67	62	57	52
73	68	63	58	53
74	69	64	59	54
75	70	65	60	55

Pool 3

96	91	86	81	76
97	92	87	82	77
98	93	88	83	78
99	94	89	84	79
100	95	90	85	80

Pool 4

121	116	111	106	101
122	117	112	107	102
123	118	113	108	103
124	119	114	109	104
115	110	105	100	95

Pool 5

Hybridization Results Were Processed by HybSweeper II

Hybsweeper - revision 1.6

:1-CMCMBb-C-27057-090224-1.gif

Select Image File...

Import... Export...

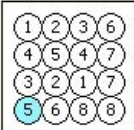
Pool: X1 Library: CMCMBb

C | 97-144

Top Left
 Bottom Right
 Pick Spot

Refresh Help

X: 65	Field: 6
Y: 64	Pos: 5
Plate: 126	Row: P
	Col: 16



Clone prs. saved: 29

Save Coordinates

No Rating Plate Order

Save Report to File...

0134	J	19
0134	K	09

Show Remove

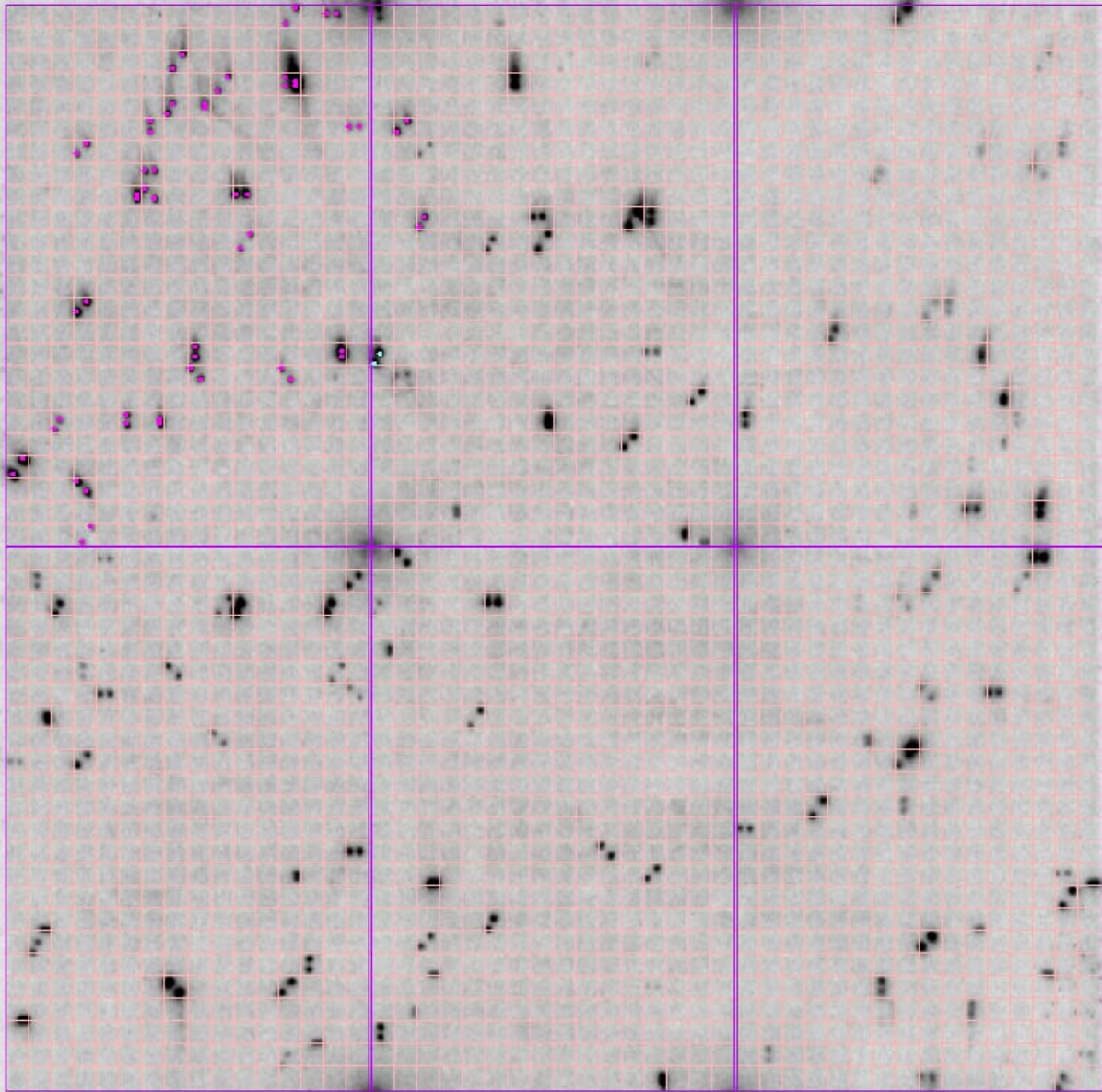
Brightness: 10

Contrast: 0

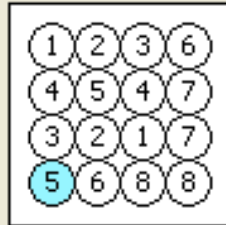
Rotate: 0

Scale: 71

Deconvolution...



X: 65 Field: 6
Y: 64 Pos: 5
Plate: 126 Row: P
 Col: 16



Clone prs. saved: 29

Save Coordinates

No Rating Plate Order

Save Report to File...

0134	J	19
0134	K	09

< >

Show Remove

Brightness: 10

Contrast: 0

Rotate: 0

Scale: 71

Deconvolution...

Current hit pattern

Current clone address

Addresses of selected hits

Deconvolution of the Hybridization results

Chestnut FPC

v.9.1

- 125,459 clones input,
- 1,104,099 bands total,
- 18X genome equivalents,
- Initial setting: Tolerance: 3
Cutoff : 1e-60
- Final setting: Tolerance: 3
Cutoff: 1e-30
- 4345 contigs, 17132 singletons
- 579 markers anchored
- CMap available at
<http://www.fagaceae.org/cmap/>

Project CMCMBb_d_1e-30_ends1e-45break Clones 125459 Seq 0 Markers 579 ze.. Help

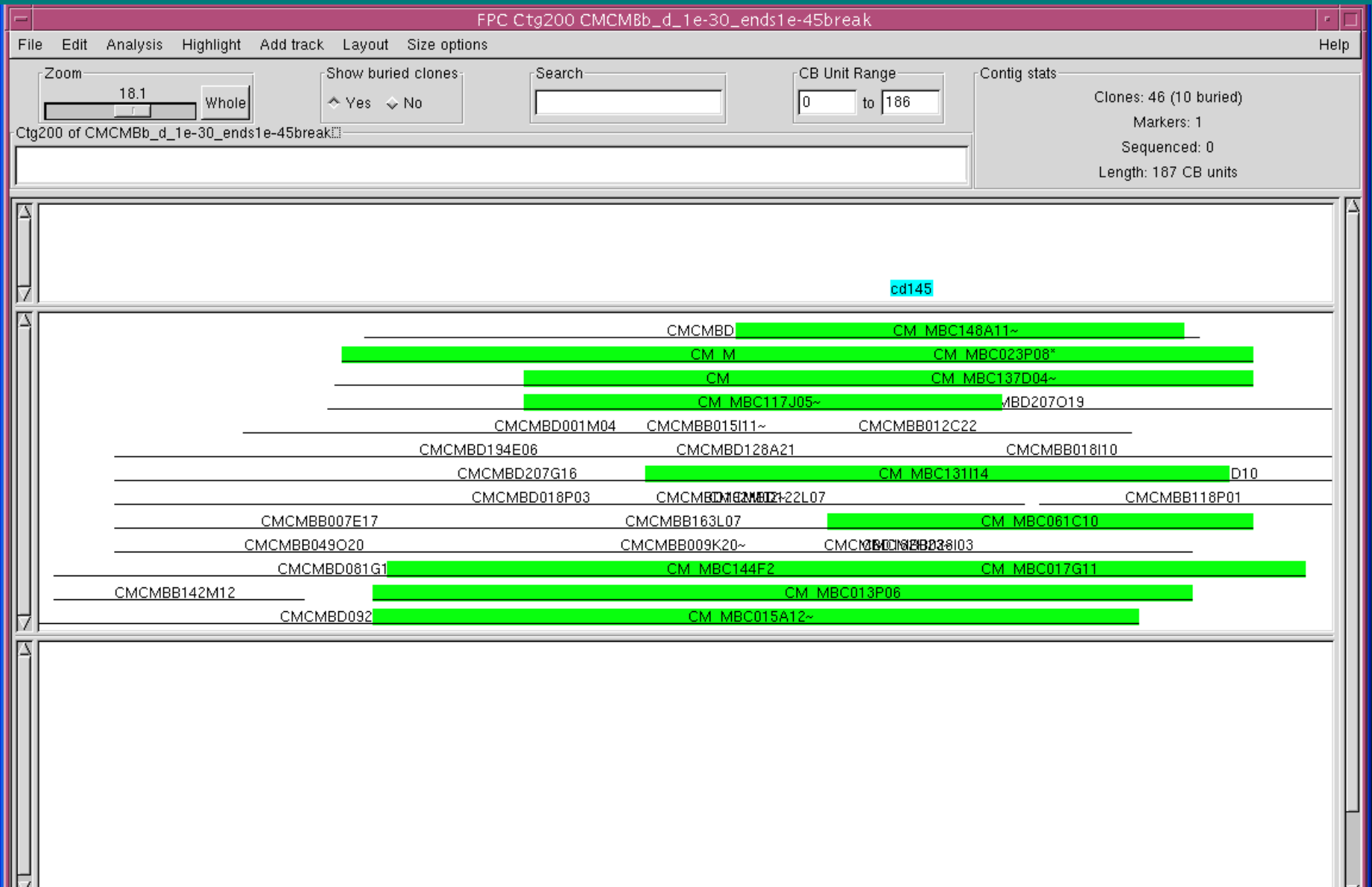
Contigs 4345 Singles 17132

Max 3324, 1312 (>25), 1275 (25:10), 1455 (9:3), 303 (2)

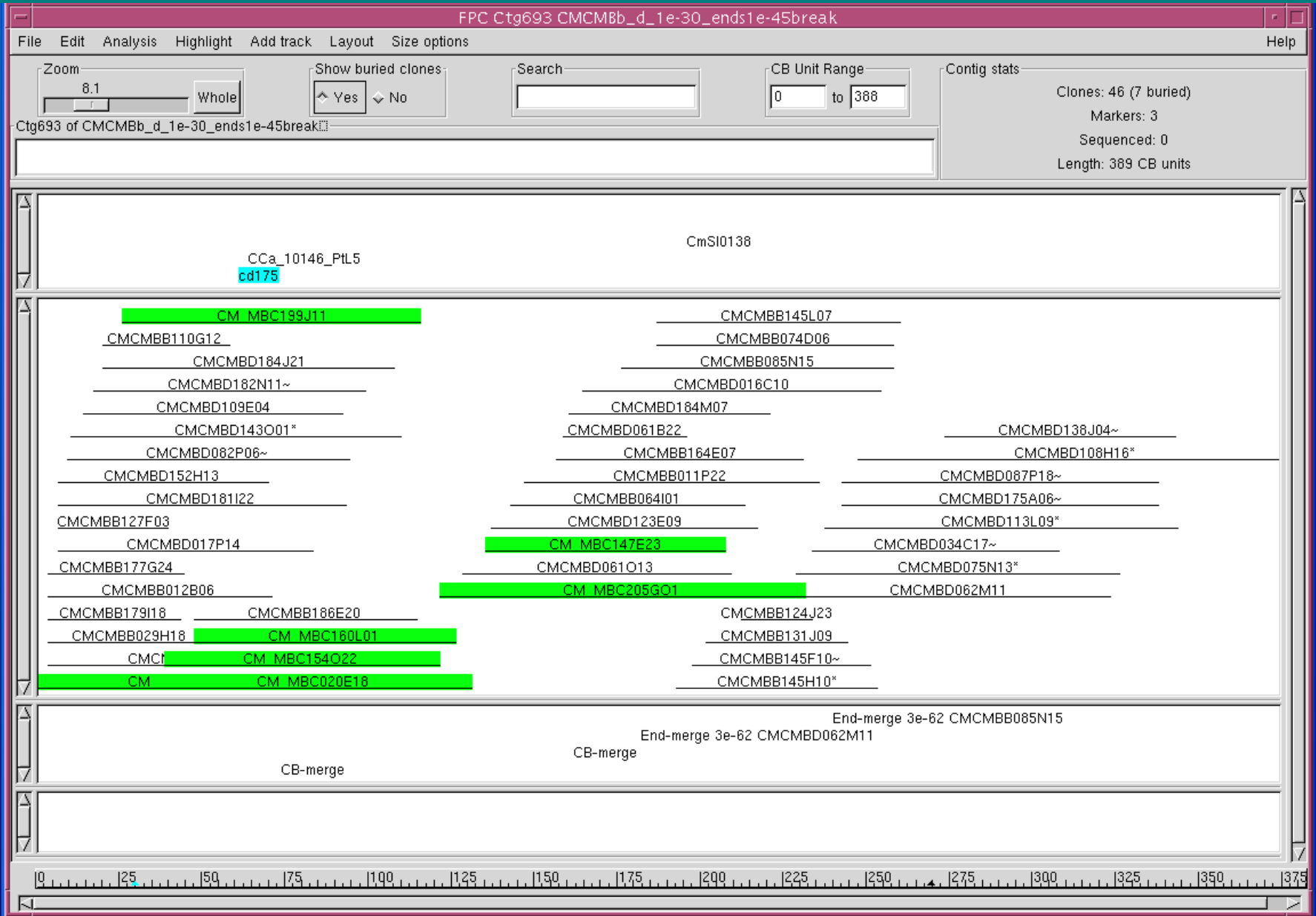
Chr_Remark Search Summary

Contig	Clone	Marker	Seq	Date	Status	Qs	Chr_Remark
151	170	2	-	11sep09 11:20	Ok	~ 9	
208	167	2	-	11sep09 11:20	Ok	~ 5	
4	166	3	-	11sep09 11:20	Ok	~ 8	
178	166	2	-	11sep09 11:20	Ok	~ 7	
185	166	4	-	7sep09 12:46	Ok	~ 5	
13	165	2	-	7sep09 13:08	Ok	~ 2	
81	165	1	-	11sep09 11:20	Ok	~ 4	
425	165	-	-	7sep09 12:45	Ok	~ 1	
1180	164	3	-	11sep09 11:20	Ok	~ 0	
266	162	6	-	11sep09 11:20	Ok	~ 3	
66	161	3	-	7sep09 12:44	Ok	~ 3	
60	159	2	-	11sep09 11:20	Ok	~ 7	
73	157	1	-	11sep09 11:20	Ok	~ 6	
211	156	1	-	11sep09 11:20	Ok	~ 6	
55	155	2	-	11sep09 11:20	Ok	~ 7	
139	154	1	-	11sep09 11:20	Ok	~ 5	
270	148	2	-	11sep09 11:20	Ok	~ 3	
80	143	-	-	11sep09 11:20	Ok	~ 7	
17	142	2	-	11sep09 11:20	Ok	~ 3	
8723	142	3	-	7sep09 12:00	Ok	~ 3	
726	140	-	-	11sep09 11:20	Ok	~ 3	
12	139	3	-	7sep09 12:29	Ok	~ 1	
240	139	2	-	7sep09 12:50	Ok	~ 6	
980	139	-	-	11sep09 11:20	Ok	~ 4	
5694	136	2	-	7sep09 13:08	Ok	~ 6	
757	135	4	-	11sep09 11:20	Ok	~ 0	
82	134	2	-	7sep09 12:33	Ok	~11	
1341	134	1	-	11sep09 11:20	Ok	~ 2	
1879	131	2	-	7sep09 13:02	Ok	~ 1	
32	130	4	-	11sep09 11:20	Ok	~ 5	
1042	130	1	-	11sep09 11:20	Ok	~ 2	
383	129	3	-	11sep09 11:20	Ok	~ 4	
100	128	1	-	11sep09 11:20	Ok	~ 6	
365	128	2	-	11sep09 11:20	Ok	~ 5	
148	127	1	-	11sep09 11:20	Ok	~ 3	
169	127	2	-	7sep09 12:22	Ok	~ 2	
782	127	1	-	7sep09 12:58	Ok	~ 6	
149	126	3	-	7sep09 11:29	Ok	~ 4	
815	126	1	-	11sep09 11:20	Ok	~ 5	
4210	125	-	-	11sep09 11:20	Ok	~ 5	
777	124	2	-	7sep09 12:33	Ok	~ 4	
409	122	2	-	11sep09 11:20	Ok	~ 2	
663	122	-	-	7sep09 11:34	Ok	~ 4	
772	122	3	-	7sep09 12:11	Ok	~ 2	
224	121	2	-	11sep09 11:20	Ok	~ 1	
557	121	4	-	7sep09 13:02	Ok	~ 2	
101	120	2	-	11sep09 11:20	Ok	~ 5	
192	120	-	-	7sep09 12:43	Ok	~ 9	
299	120	3	-	11sep09 11:20	Ok	~ 5	
130	119	4	-	11sep09 11:20	Ok	~ 6	
399	119	2	-	7sep09 12:41	Ok	~ 0	
436	119	2	-	11sep09 11:20	Ok	~ 5	
128	118	1	-	7sep09 12:41	Ok	~ 2	

Ctg200: with CD145 (=3 CM from a blight resistant QTL on LG-F)



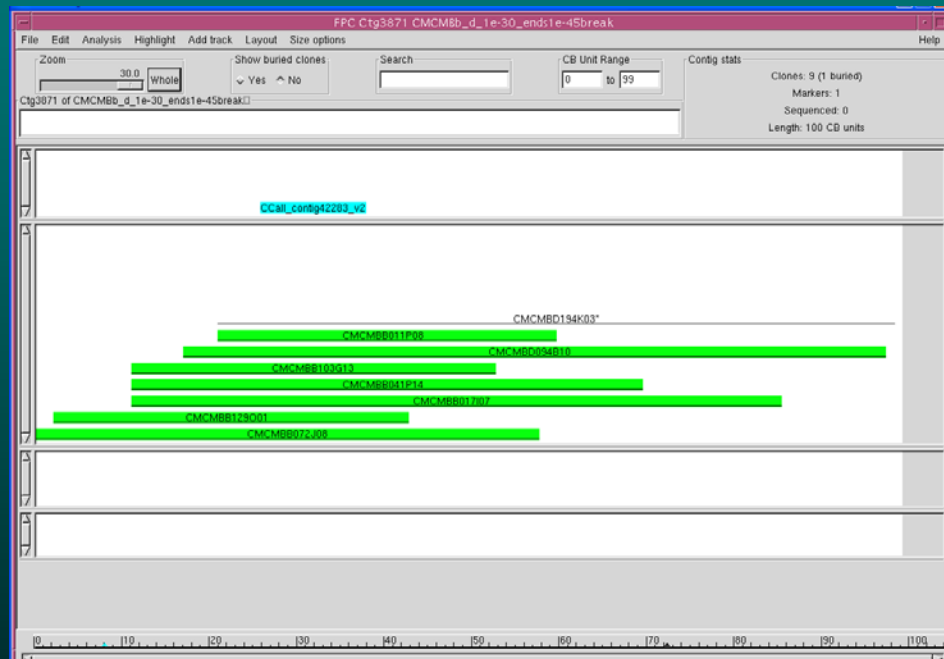
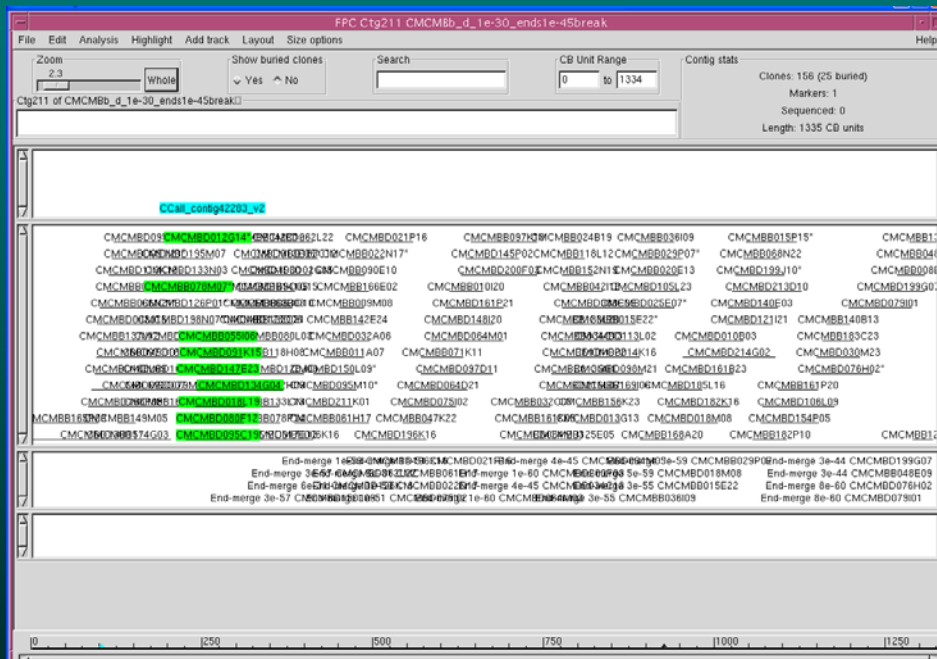
Ctg693: with CD175 (=3 CM away from a blight resistant QTL on LG-B)



Hybridization with a laccase overgo (from CCall_contig 42283_v2) identified two individual contigs.

Ctg211

Ctg3871



At least two loci for the laccase gene

Anchoring Linkage Specific Markers to the Physical Map

LG	Locus	Type of marker	Number of HICF contigs	Ctg# (length in CB unit)
A	CD192	RFLP	1	9033(358)
A	PGI	Isozyme	1	2253(616)
B	CD175	RFLP	1	693(389)
B	SKD	Isozyme	1	102(647)
C	GM49	RFLP	1	185(834)
C	PGM	Isozyme	5	
D	CsCAT02	SSR	1	223(1368)
D	QpZAG58	SSR	1	320(773)
E	5SrDNA	-	n.a.	n.a.
E	CCMC39433	EST	n.a.	n.a.
F	CD145	RFLP	1	200(187)
F	EMCs4	SSR	1	13(1056)
G	CD172	RFLP	2	1707(511)
G	CCMC04326	EST	n.a.	n.a.
H	cystatin	CAPS	2	548(174), 869(455)?
H	CD62	RFLP	2	2293(448), 6546(519)
I	CD143	RFLP	1	223(1368)
I	EMCs14	SSR	1	923(1041)
J	CD60	RFLP	1	1379(270)
J	GOT	Isozyme	1	1180(1358)
K	CsCAT17	SSR	1	5270(99)
K	CCMC09685	EST	n.a.	n.a.
L	EMCs15	SSR	1	2598(195), 1426(261)
L	CCMC40039	EST	n.a.	n.a.

Comparative Genomics of Chestnut and Poplar Genomes: Hybridization with Poplar ortholog Markers

Approach

- **BLAST chestnut ESTs to poplar genome (and peach) database for all orthologs**
- **Remove low complex sequences**
- **Locate the poplar orthologs in poplar genome**
- **Among the poplar orthologs, choose 250 markers evenly spaced across the entire poplar genome, and design overgos from these markers**
- **Multi-dimensional Pool hybridization**
- **Anchor these markers to the physical map**



CC unigene already hybridized
 Gap of > 3M bp on LG
 Gap of > 2M bp on LG

Hybridization with Poplar LG_II Ortholog Markers

KOG gene	CC unigene	Prunus unigene	Populus gene model	Populus Linkage Group	Populus Location	Distance to next marker	Anchored chestnut ctg#(CB unit)
K03665	CCall_contig23188_v2	PUT-171a- Prunus_persica-6479	grail3.0003025501	LG_II	2002641	1831631	2639(227)
K00457	CCall_contig4535_v2	PUT-171a-	fgenes4_pm.C_LG_II000277	LG_II	3834272	1910168	1796(410)
K01698	CCall_contig43327_v2	PUT-171a- Prunus_persica-14968	eugene3.00020758	LG_II	5744440	854901	
K03650	CCall_contig29555_v2	PUT-171a- Prunus_persica-10920	eugene3.00020856	LG_II	6599341	865001	2593, 8744
K03926	CCall_contig18494_v2	PUT-171a- Prunus_persica-15429	grail3.0003090402	LG_II	7464342	395889	
K02492	CCall_contig26790_v2	PUT-171a- Prunus_persica-25980	estExt_Genewise1_v1.C_LG_II0707	LG_II	7860231	1910747	2888, 3634, 2696
K04122	CCall_contig6450_v2	PUT-171a- Prunus_persica-7796	gw1.II.3880.1	LG_II	9770978	1169264	9216, 1015, 3507
K00800	CCall_contig4127_v2	PUT-171a- Prunus_persica-21182	eugene3.00021350	LG_II	10940242	853079	
K02931	CCall_contig17550_v2	PUT-171a- Prunus_persica-13584	gw1.II.2942.1	LG_II	11793321	661935	8822, 12
K07029	CCall_contig41993_v2	PUT-171a- Prunus_persica-89	grail3.0039009801	LG_II	12455256	1987274	
K01363	CCall_contig4884_v2	PUT-171a- Prunus_persica-21088	eugene3.00021714	LG_II	14442530	1393636	1016, 545, 123, 1827
K01887	CCall_contig20184_v2	PUT-171a- Prunus_persica-10426	estExt_Genewise1_v1.C_LG_II3255	LG_II	15836166	2141713	8723(989)
K02257	CCall_contig42873_v2	PUT-171a- Prunus_persica-7783	fgenes4_pg.C_LG_II001967	LG_II	17977879	3708288	
K00604	CCall_contig26937_v2	PUT-171a-	eugene3.00002342	LG_II	21686167	1067797	622(275)
K03189	CCall_contig25908_v2	PUT-171a- Prunus_persica-20190	grail3.0021023901	LG_II	22753964	357392	
K09561	CCall_contig45783_v2	PUT-171a- Prunus_persica-11136	estExt_fgenes4_pm.C_LG_II1150	LG_II	23111356	615609	
K08903	CCall_contig22421_v2	PUT-171a- Prunus_persica-23352	fgenes4_pm.C_LG_II001205	LG_II	23726965	449875	365(1323)
K01520	CCall_contig28752_v2	PUT-171a- Prunus_persica-5761	estExt_fgenes4_pg.C_LG_II2596	LG_II	24176840 24482572	305732	

Summary

- **A total of four BAC libraries were constructed from both Chinese and American chestnut. Each library had 10 to 12X genome equivalents and an average insert size from 90 to 140 kb.**
- **HICF fingerprinting and further assembly by FPC v9.1 from 125,459 BACs from Chinese chestnut libraries led to a physical map with 4,345 contigs, 17,132 singletons, and had a 18X genome coverage.**
- **High-throughput multi-dimensional overgo hybridization of the high density BAC filters with RFLP, SSR, SNP and poplar ortholog markers anchored 579 markers on the physical map.**

Acknowledgements

US Forest Service

Tom Kubisiak

Penn State University

John Carlson

Ali Barakat

SUNY

William A. Powell

Kathleen Baier

NC State University

Ron Sederoff

Chris Smith

Nicholas Wheeler

TACF

Paul Sisco

Fred Hebard

Joe James

Connecticut Agricultural Experiment Station

Sandra Anagnostakis



NSF PGRP 0605135