A comparison of candidate gene-based and genotyping-by-sequencing (GBS) approaches to trait mapping in *Gossypium barbadense* L.

Authors

Carla Jo Logan-Young. Texas A&M University. College Station, TX. United States. *(presenting author)*

John Z. Yu. USDA-ARS. College Station, TX. United States. *(corresponding author)*

Richard G. Percy. USDA-ARS. College Station, TX. United States. *(corresponding author)*

Alan E. Pepper. Texas A&M University. College Station, TX. United States. *(corresponding author)*

Ibrohkim Abdurakhmonov. Center of Genomic Technologies, Academy of Sciences of Uzbekistan. Yuqori Yuz, Qibray region Tashkent, Uzbekistan. *(corresponding author)*

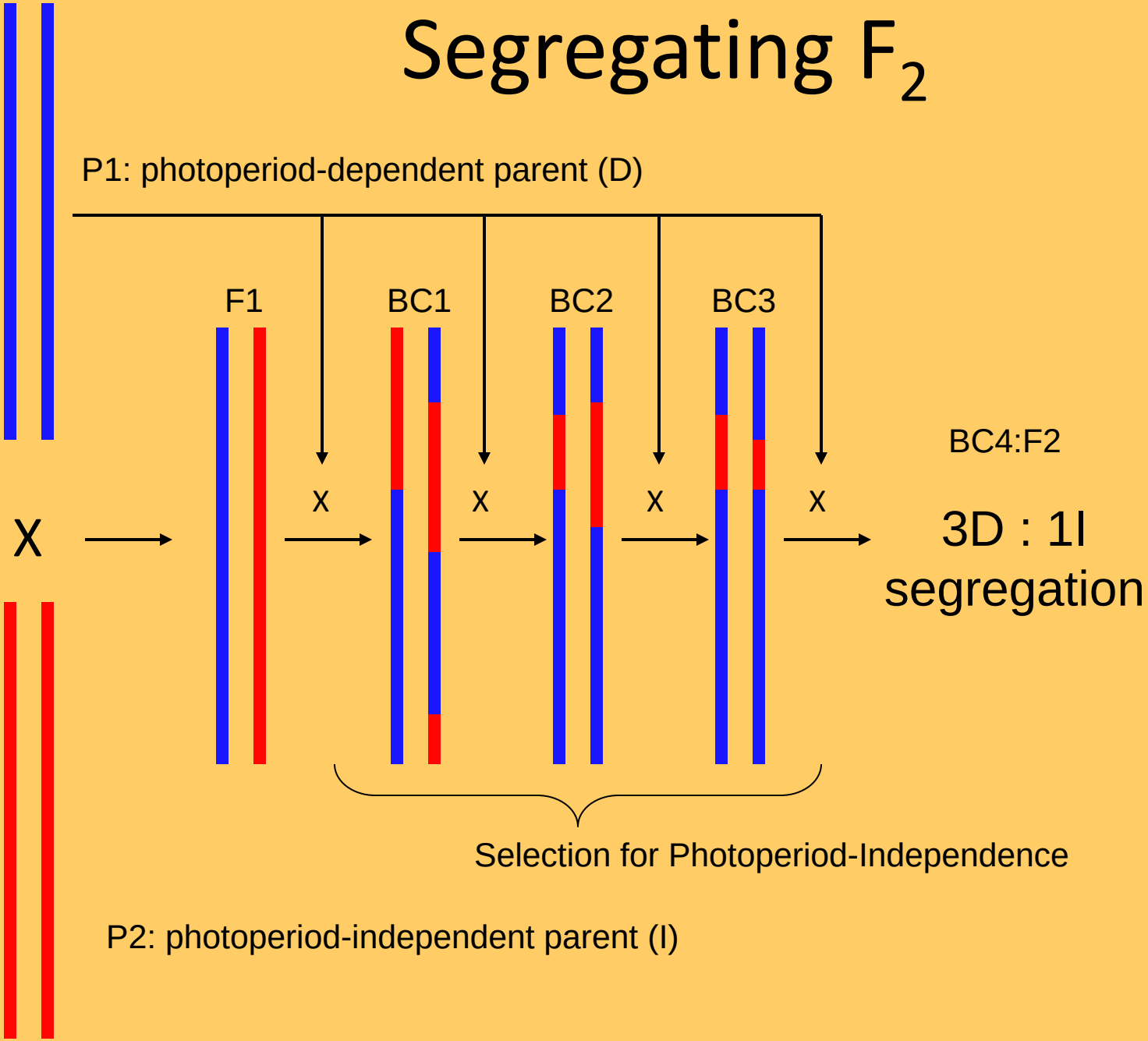
October 11th, 2012

Richard Percy *Gossypium barbadense* Field Trials

Photoperiod Dependent

Photoperiod Independent

Segregating F_2



G. barbadense Phenotypic Results

Photoperiod Dependent : Photoperiod Independent Ratio

		Significance
Argentina - BC4F2 - K158	3:1	<.10
Argentina - BC4F2 - K145	3:1	<.01 **
Belize - BC4F2 - B235	3:1	<.01 **
Martinique - BC4F2 - AS24	3:1	<.01 **
Guadeloupe - BC4F2 - K46	3:1	<.001 ***
South America - BC4F2 - FRYX633	3:1	<.05 *
Belize - BC4F2 - B231	3:1	<.01 **
Dominican Republic - BC4F2 - AS86	3:1	<.01 **

Hypothesis

- A single gene influences photoperiod independence within the floral network in cotton.

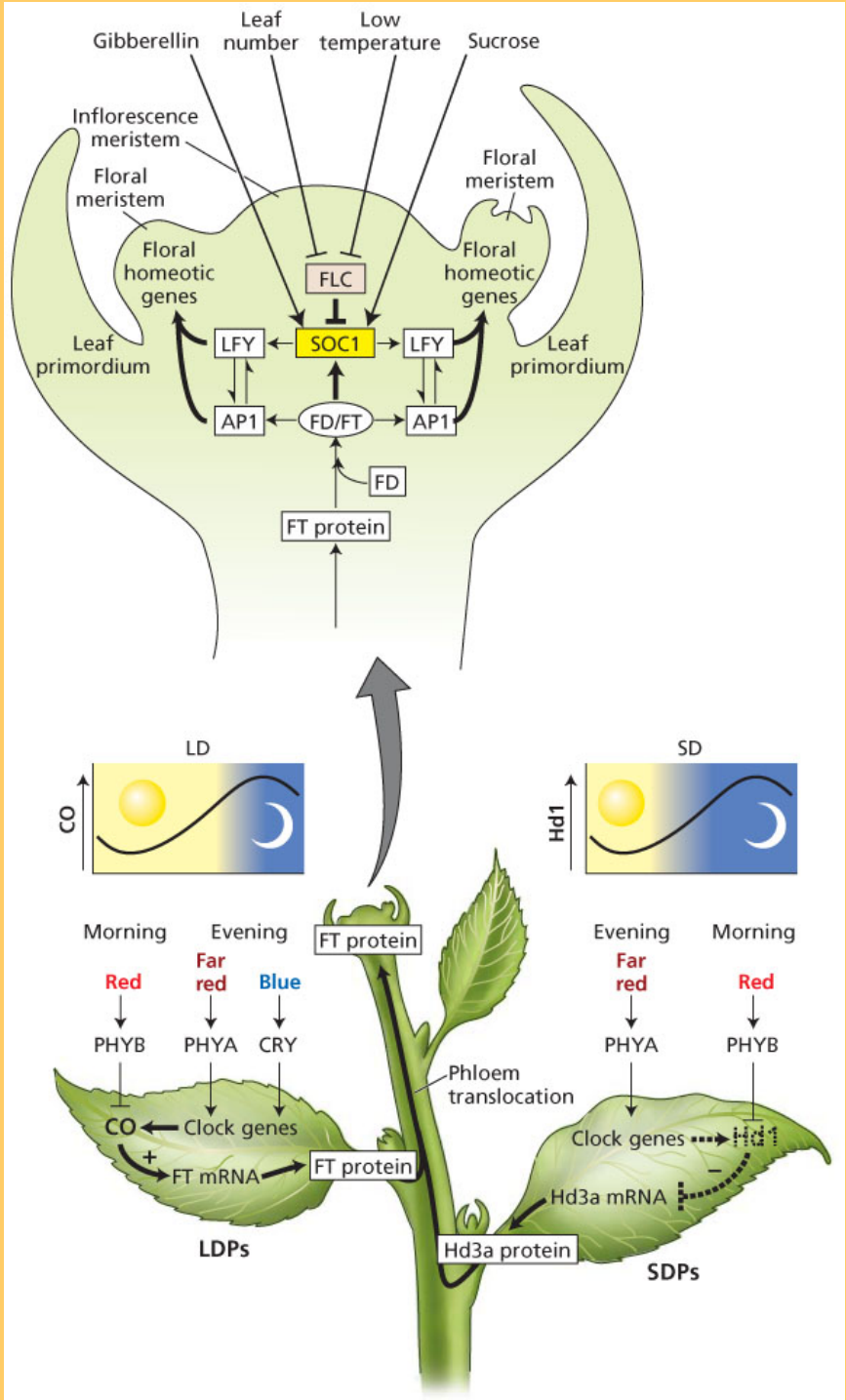
Arabidopsis Flowering Pathway

A Companion to Plant Physiology, Fourth Edition by Lincoln Taiz and Eduardo Zeiger

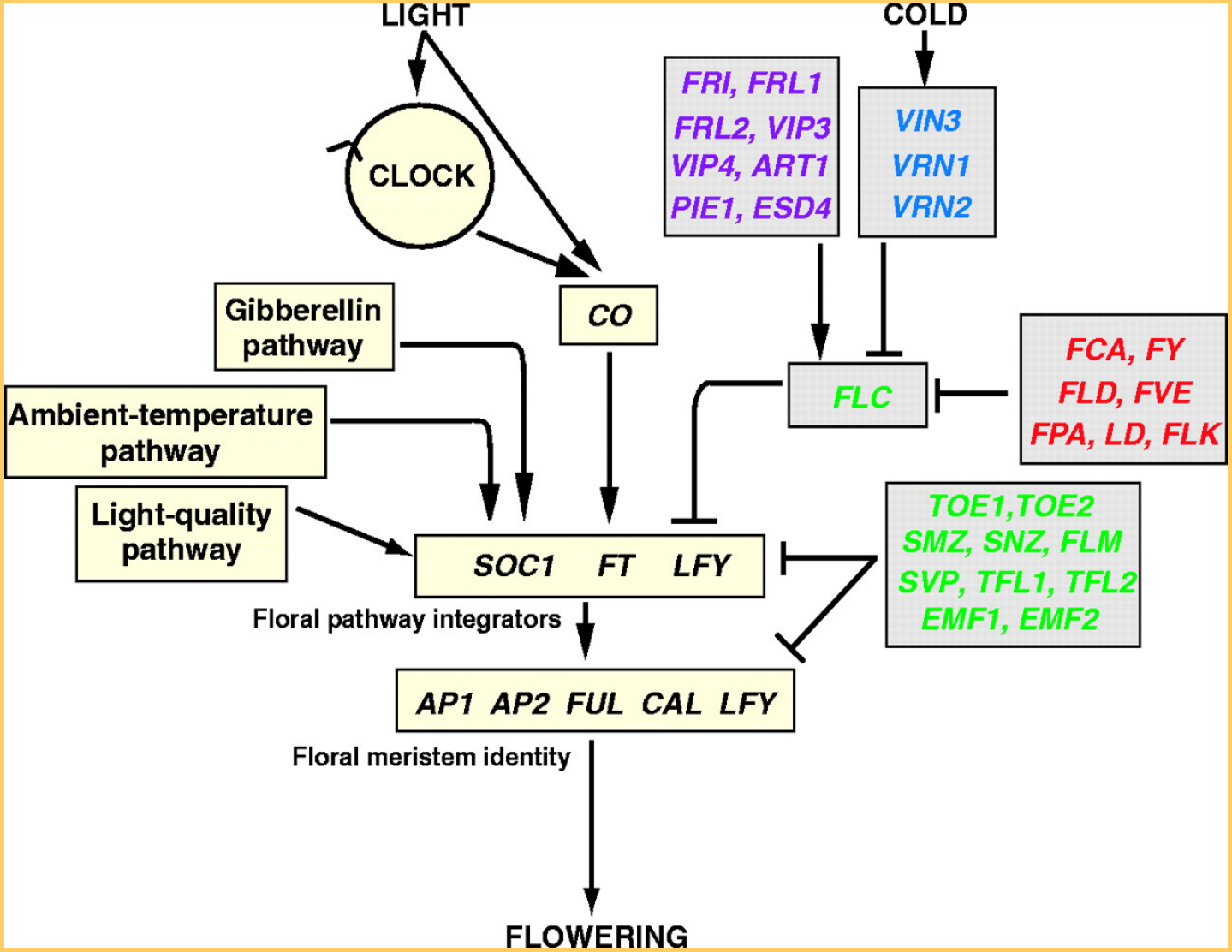
Essay 25.2
FT Protein, not mRNA, is the Phloem-Mobile Signal for Flowering

Jan A. D. Zeevaart, MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI

Figure 1 Multiple developmental pathways for flowering in *Arabidopsis*



Pathways controlling flowering-time in Arabidopsis.



Henderson I R , Dean C (2004) Development 131:3829-3838

Candidate Genes

Circadian Clock Genes

ELF3	Clock Transcription Factor
GI	Clock Transcription Factor
LHY/CCA1	Clock Transcription Factor
TOC1	Clock Transcription Factor
ADO3_FKF1	Signaling Molecule
ZTL	Signaling Molecule

Light Dependent Genes

CRY1	Photoreceptor
CRY2	Photoreceptor
CRY3	Photoreceptor
HY2	Photoreceptor
HY6	Photoreceptor
PHYA	Photoreceptor
PHYB	Photoreceptor
PHYC	Photoreceptor
PHYE	Photoreceptor

Floral MiRNA Activators

MiRNA172c	MiRNA
MiRNA172a	MiRNA
MiRNA172d	MiRNA

Floral Integrator Genes

COP1	Signaling Molecule
DET1	Signaling Molecule
LDW1	Signaling Molecule
LDW2	Signaling Molecule
PIE1	Signaling Molecule
PRR5	Signaling Molecule
PRR7	Signaling Molecule
TFL1	Signaling Molecule
TFL2	Signaling Molecule
ATGRP7	Transcription Factor
CO	Transcription Factor
COL3	Transcription Factor
COL5	Transcription Factor
FD	Transcription Factor

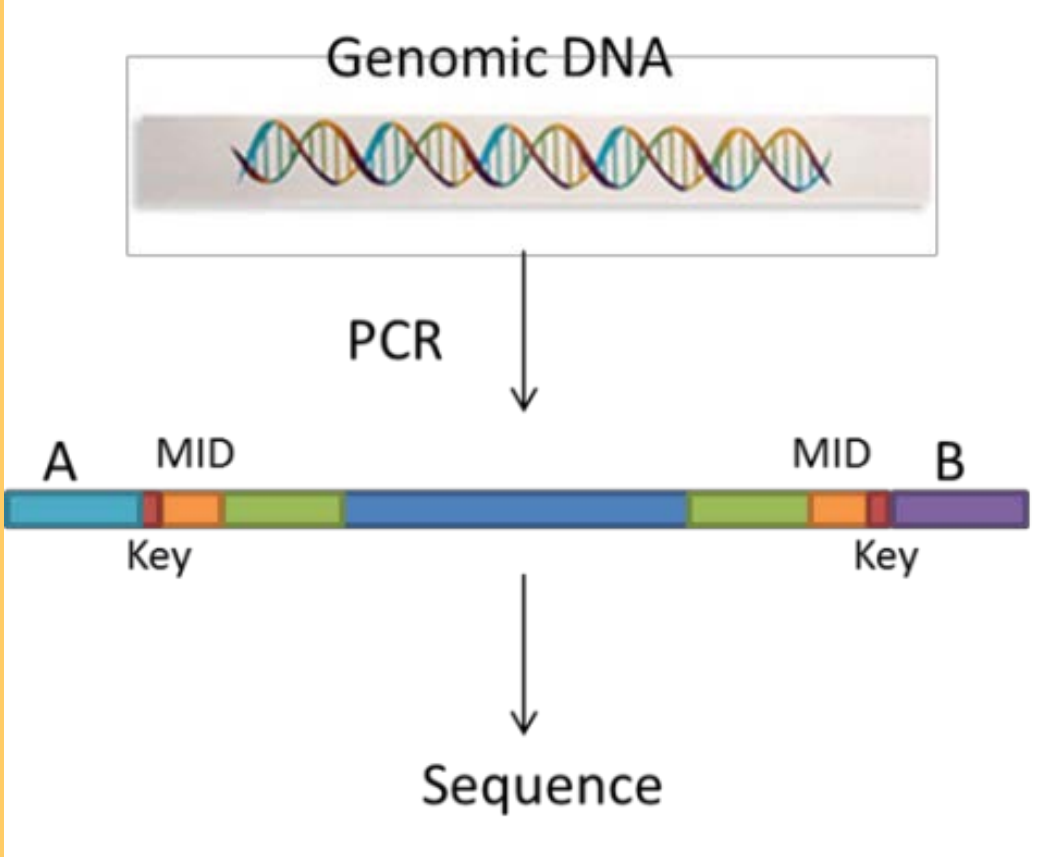
Floral Homeotic Genes

FT	Signaling Molecule
AP1_AGL7	Transcription Factor
LFY	Transcription Factor
SOC1	Transcription Factor

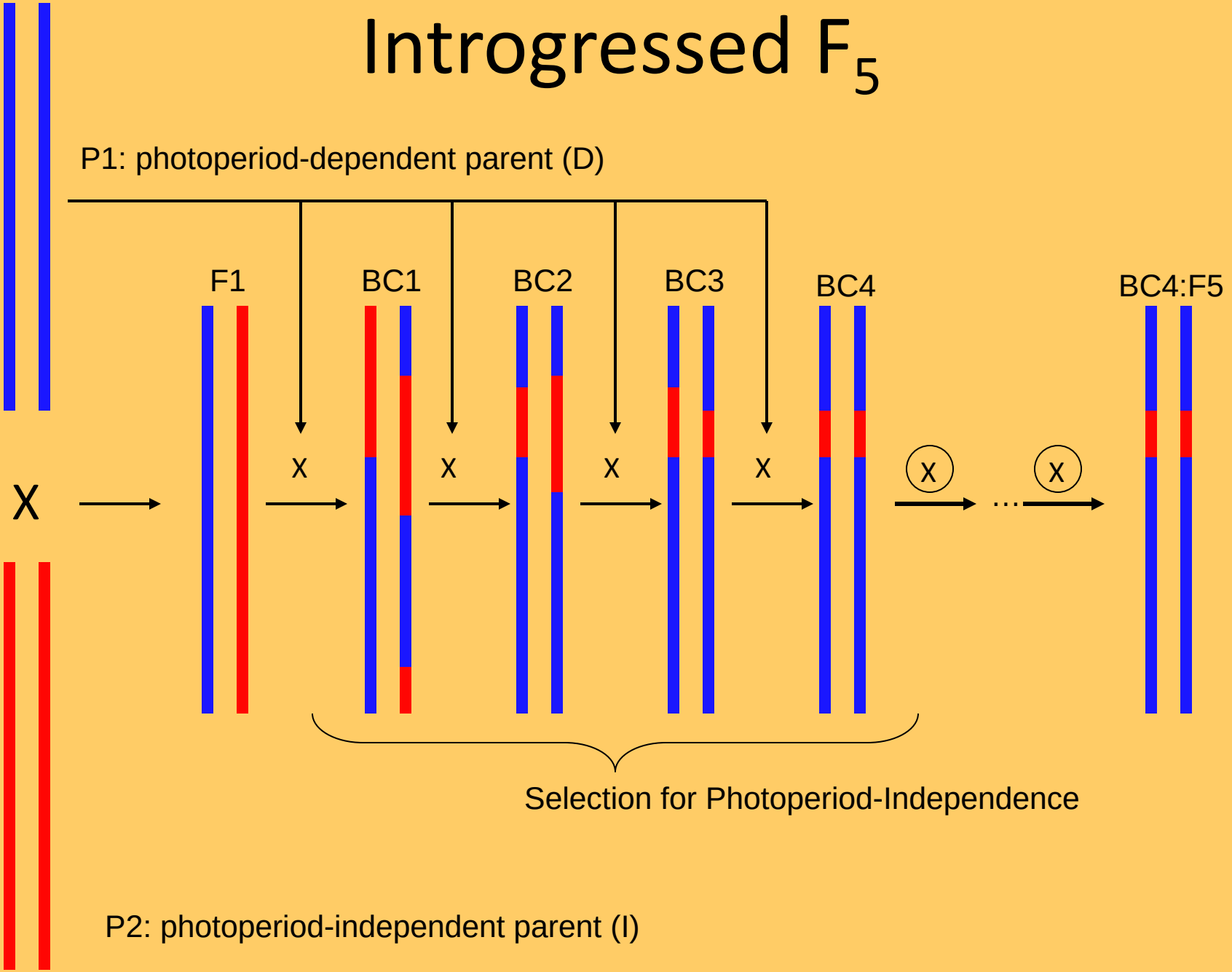
Floral Development Genes

AGL1	Transcription Factor
AGL16	Transcription Factor
AGL2	Transcription Factor
AGL24	Transcription Factor
AGL3_SEP4	Transcription Factor
AGL30	Transcription Factor
AGL32	Transcription Factor
AGL6	Transcription Factor
AGL62	Transcription Factor
AGL65	Transcription Factor
AGL8	Transcription Factor
AGL9_SEP3	Transcription Factor
AP3	Transcription Factor
ARP6	Transcription Factor
Pistillata	Transcription Factor
SPA3	Transcription Factor
SPA4	Transcription Factor

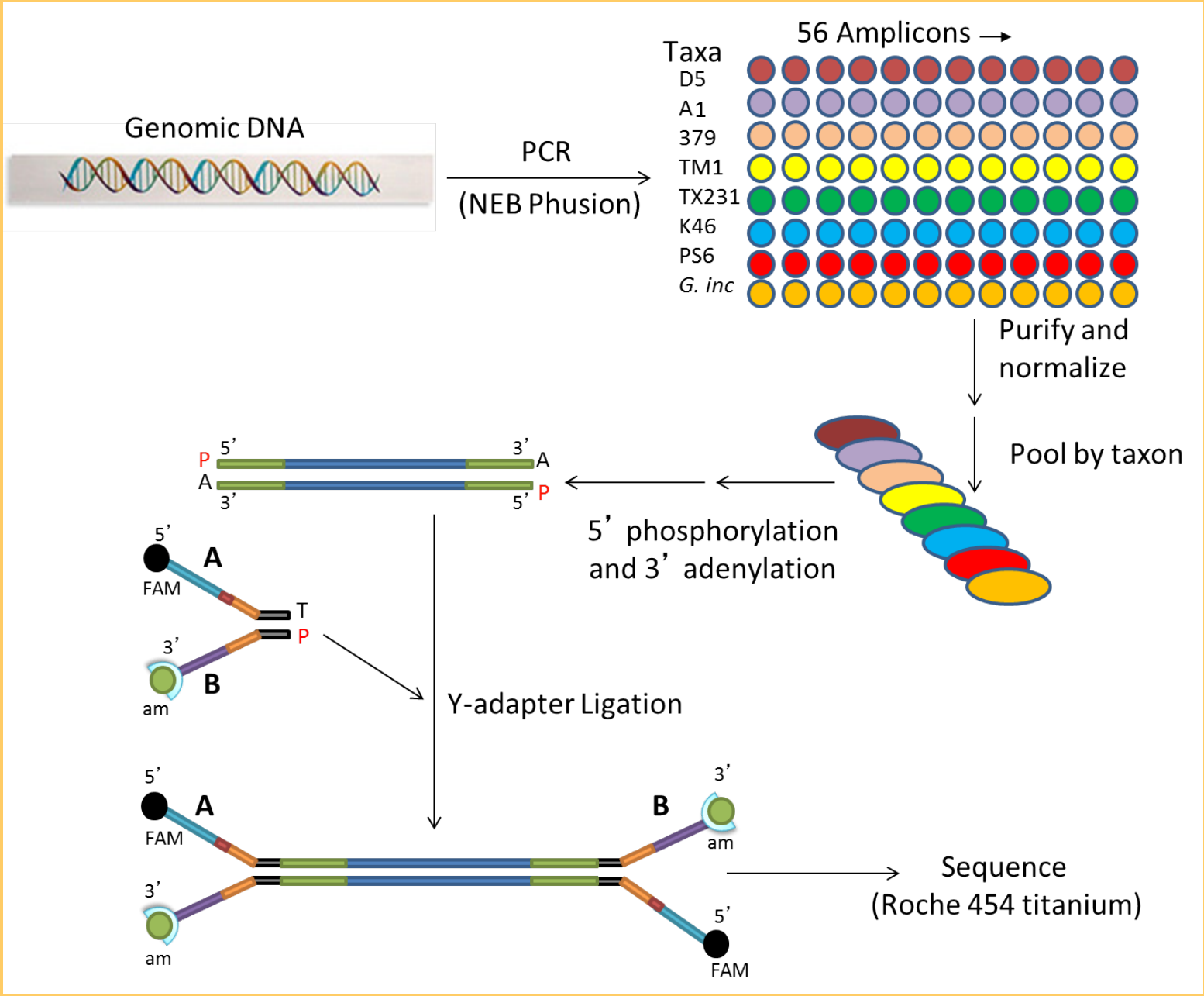
454 Sequencing Product



Introgressed F_5

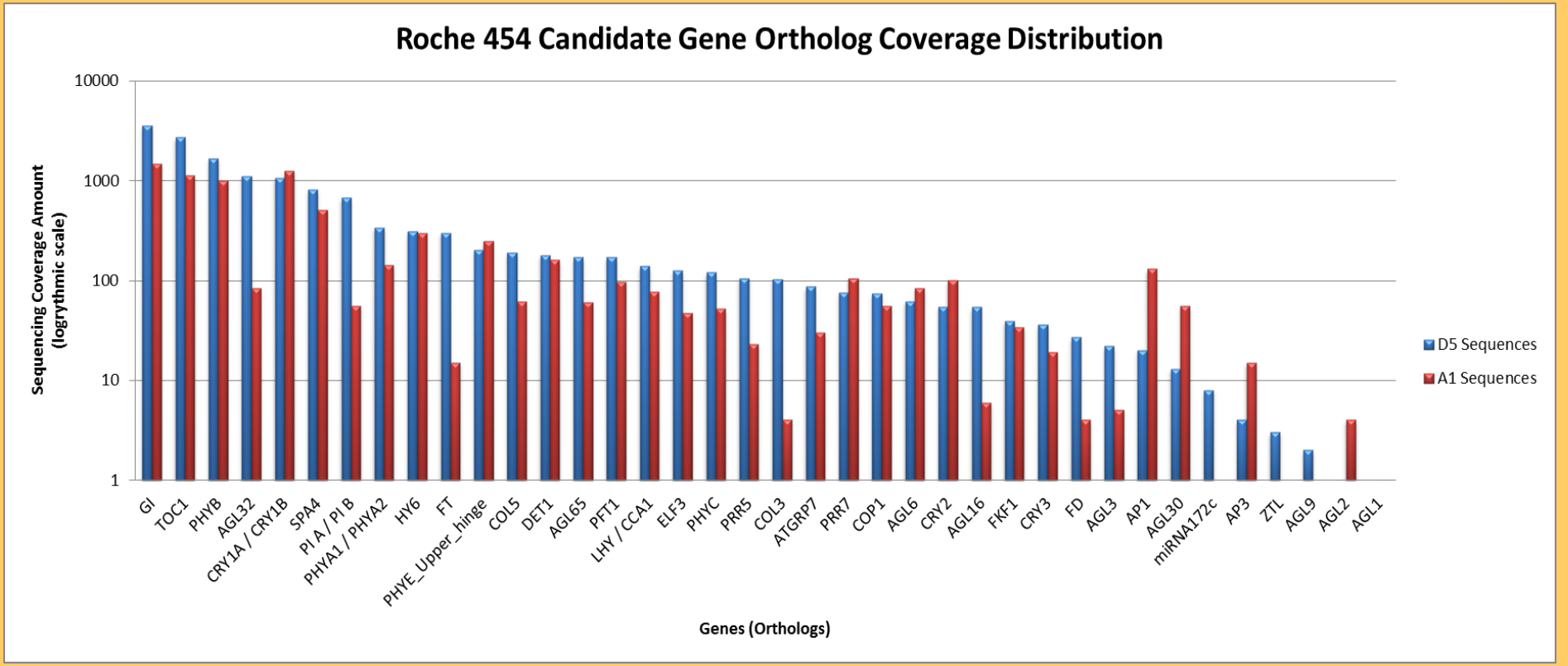


Candidate Gene Exon-Intron Spanning Regions



Candidate Experiment Results

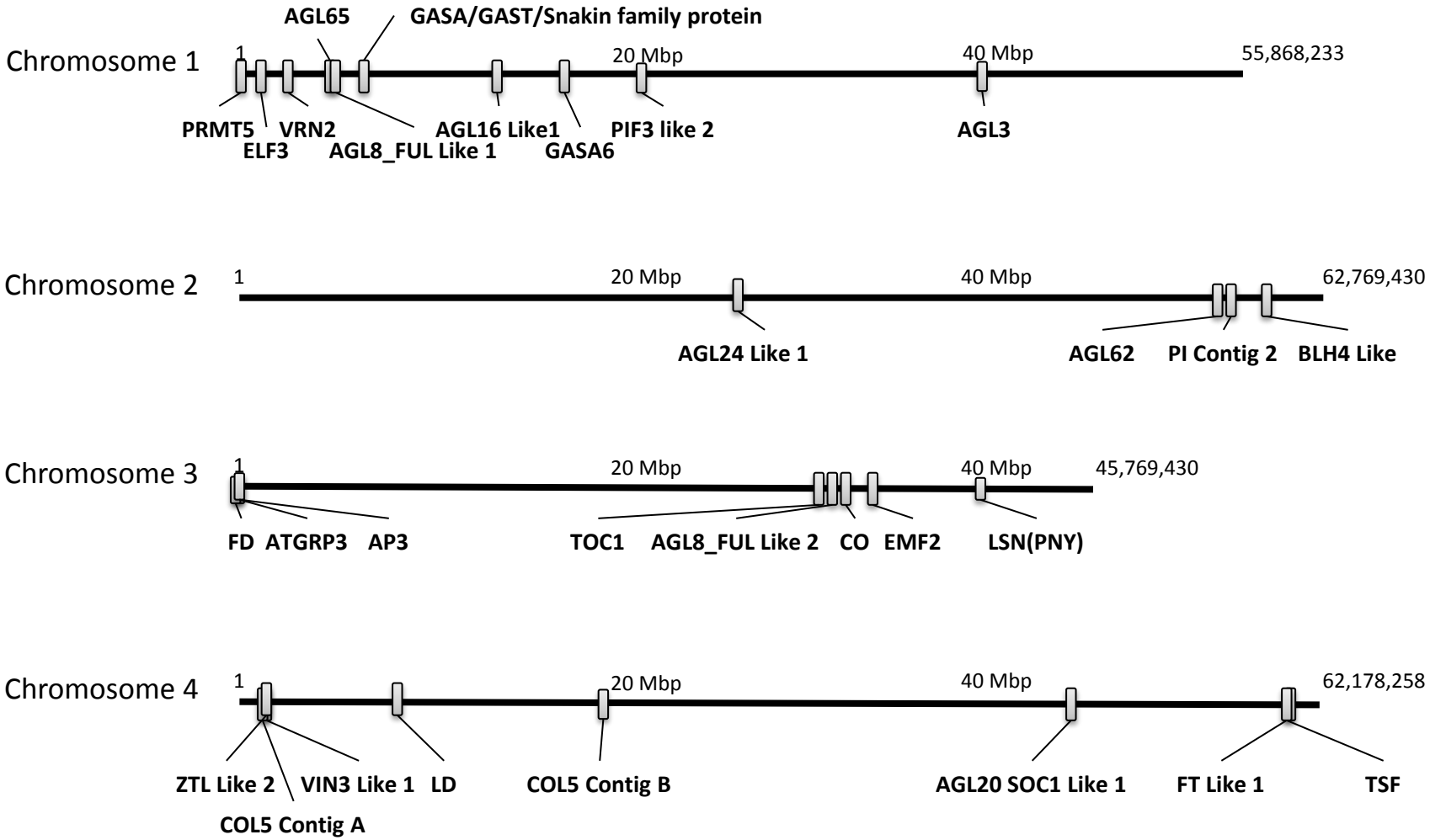
Barcode ID	Barcode	Mid Identifiers			Floral Queue	MID Parsed Sequences
AF1	ACGAGTGCGT	*Gossypium raimondii*	D5	Diploid	Photoperiodic	22693
AF2	ACGCTCGACA	*Gossypium herbaceum*	A1	Diploid	Photoperiodic	11825
AF3	AGACGCACTC	*Gossypium barbadense*	379	Allotetraploid	Photoperiod Independent	23166
AF4	AGCACTGTAG	*Gossypium hirsutum*	TM1	Allotetraploid	Photoperiod Independent	19141
AF5	ATCAGACACG	*Gossypium hirsutum*	TX231	Allotetraploid	Photoperiodic	8152
AF6	ATATCGCGAG	*Gossypium barbadense*	K46	Allotetraploid	Photoperiodic	7010
AF7	CGTGTCTCTA	*Gossypium barbadense*	PS6	Allotetraploid	Photoperiod Independent	2920
AF8	CTCGCGTGTC	*Gossypium incanum*	E4	Diploid	Photoperiodic	4792
		Ungrouped				4531

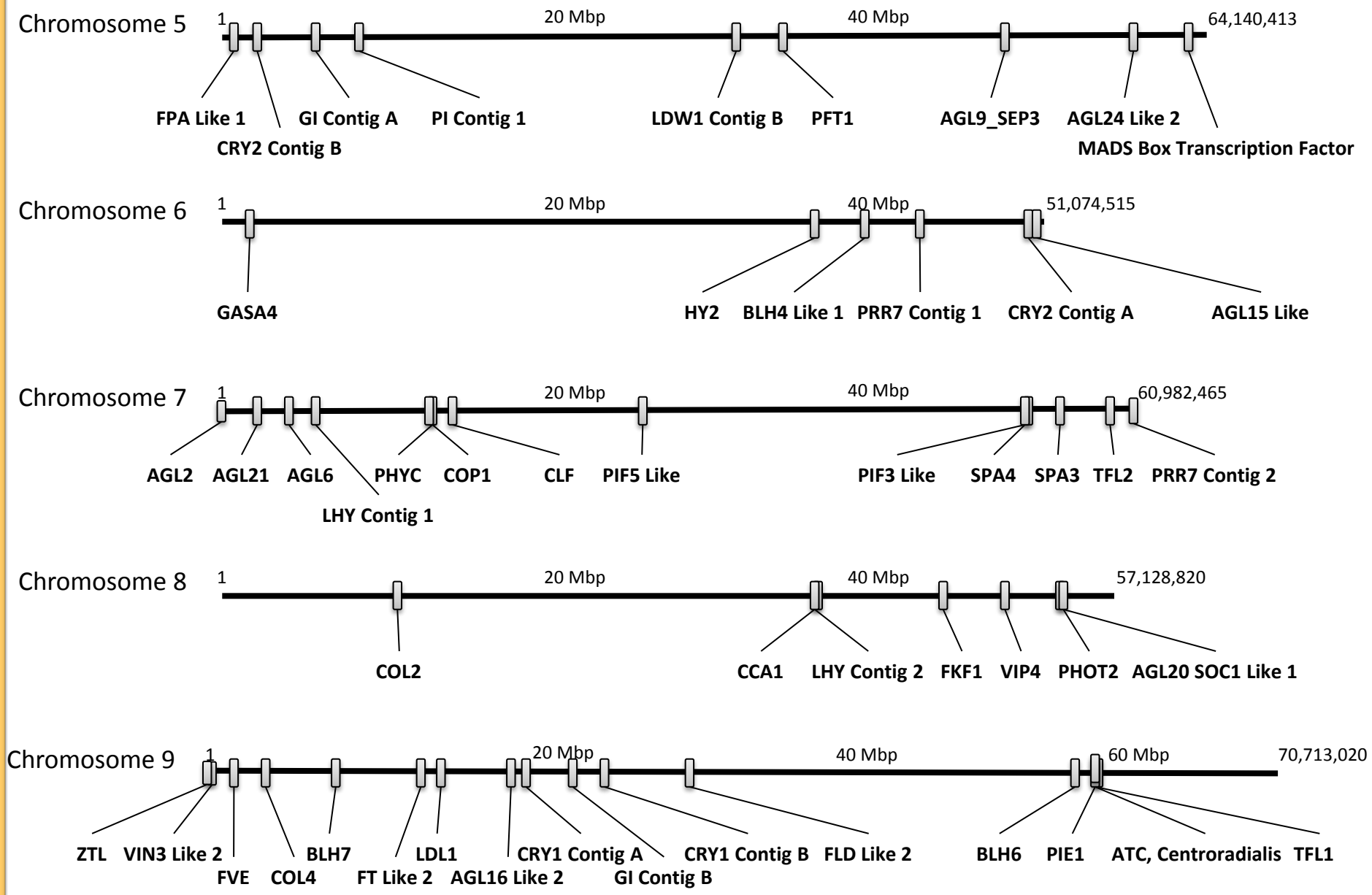


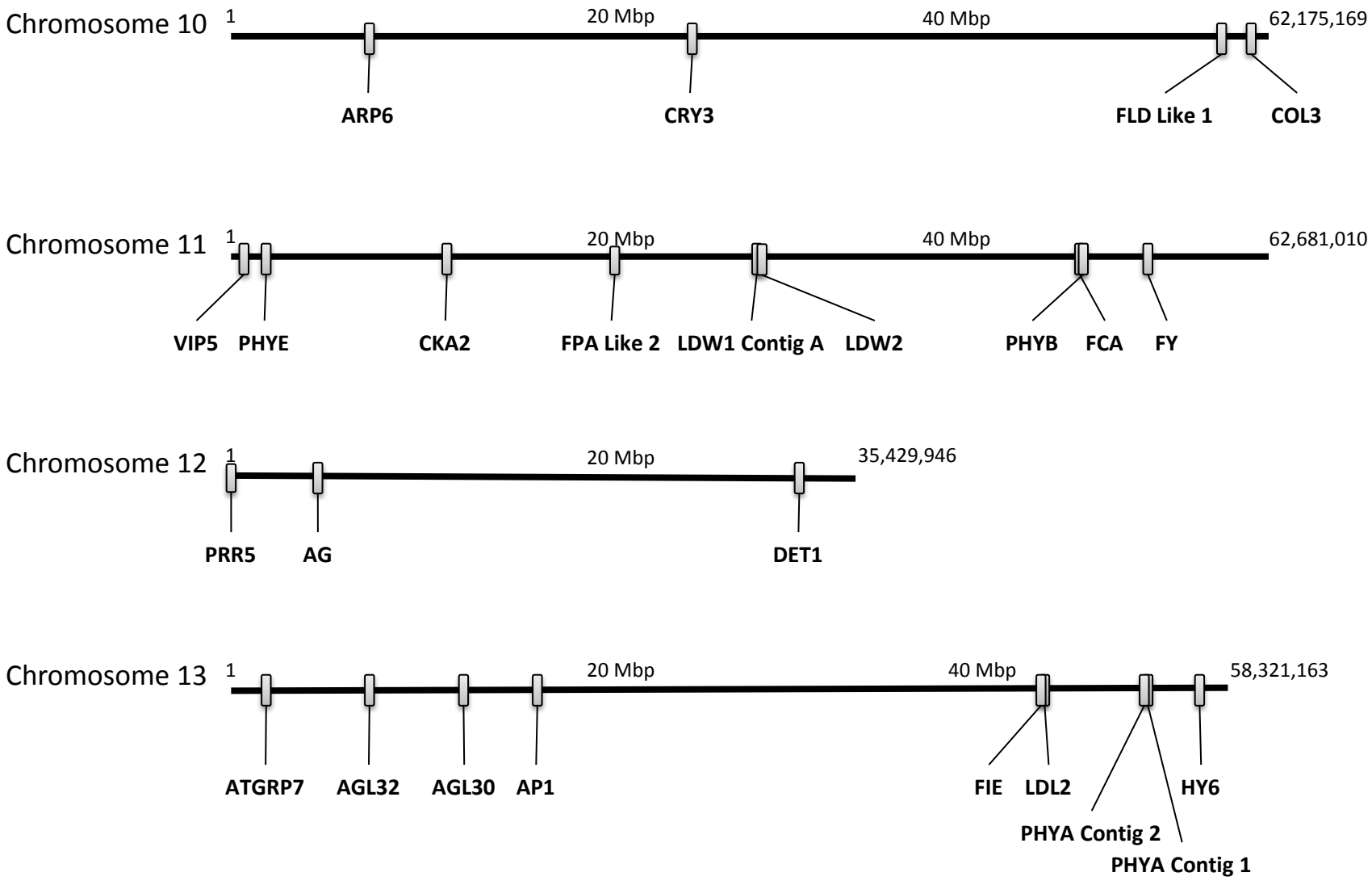
SNP Estimates for 454 Data

- **A/D SNP Ratio**
  - Exon: 1:68 bp
  - Intron: 1:195 bp
- **Intraspecific (cult. vs. wild)**
  - Exon: 1:769 bp
  - Intron: 1:73 bp
- **Interspecfic *(G. hirsutum vs. G. barbadense)***
  - Exon: 1:622 bp
  - Intron: 1:197 bp

D_5 Chromosome Alignment of Candidate Genes




Chromosome 5
1
20 Mbp
40 Mbp
64,140,413
FPA Like 1
CRY2 Contig B
GI Contig A
PI Contig 1
LDW1 Contig B
PFT1
AGL9_SEP3
AGL24 Like 2
MADS Box Transcription Factor
Chromosome 6
1
20 Mbp
40 Mbp
51,074,515
GASA4
HY2
BLH4 Like 1
PRR7 Contig 1
CRY2 Contig A
AGL15 Like
Chromosome 7
1
20 Mbp
40 Mbp
60,982,465
AGL2
AGL21
AGL6
LHY Contig 1
PHYC
COP1
CLF
PIF5 Like
PIF3 Like
SPA4
SPA3
TFL2
PRR7 Contig 2
Chromosome 8
1
20 Mbp
40 Mbp
57,128,820
COL2
CCA1
LHY Contig 2
FKF1
VIP4
PHOT2
AGL20 SOC1 Like 1
Chromosome 9
1
20 Mbp
40 Mbp
60 Mbp
70,713,020
ZTL
VIN3 Like 2
FVE
COL4
BLH7
FT Like 2
LDL1
AGL16 Like 2
CRY1 Contig A
GI Contig B
CRY1 Contig B
FLD Like 2
BLH6
PIE1
ATC, Centroradialis
TFL1


Chromosome 10
1
20 Mbp
40 Mbp
62,175,169
ARP6
CRY3
FLD Like 1
COL3
Chromosome 11
1
20 Mbp
40 Mbp
62,681,010
VIP5
PHYE
CKA2
FPA Like 2
LDW1 Contig A
LDW2
PHYB
FCA
FY
Chromosome 12
1
20 Mbp
35,429,946
PRR5
AG
DET1
Chromosome 13
1
20 Mbp
40 Mbp
58,321,163
ATGRP7
AGL32
AGL30
AP1
FIE
LDL2
HY6
PHYA Contig 2
PHYA Contig 1

Candidate Gene Conclusions

- 48 candidate genes sequenced
- SNPs characterized
- Mapped to D-genome reference sequence

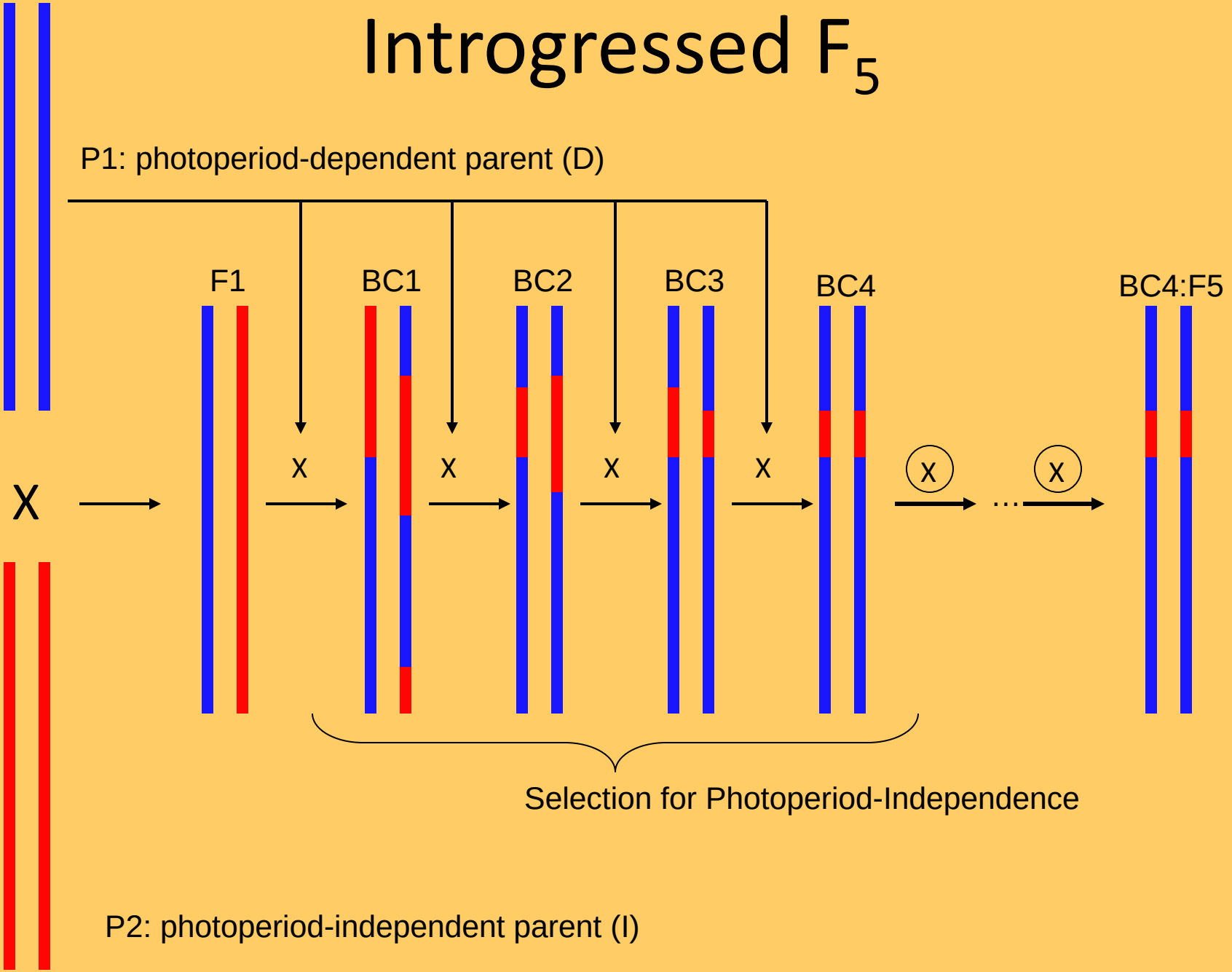
Ongoing Experiments

- Verification of SNPs linked to photoperiodic flowering
- Characterize pattern of sequence evolution of flowering genes

Genotype By Sequencing

- Independent method to narrow down genome regions containing the gene(s) influencing photoperiod independence.
- GBS method is used to sequence specific genomic regions (marker loci) across different samples utilizing restriction enzymes.

Introgressed F_5



Progeny-Dependent SNPs

HinP1I	Total Loci	1 to 3 SNPs Loci 2 parents	1 to 3 SNPs Loci 2 parents w/ progeny	PS-5 Progeny	Undomesticated Progeny
379/TM1	1505	5	n/a	n/a	n/a
379/TX231	2688	76	n/a	n/a	n/a
PS5/K56	39651	129	42	12	21
PS5/PI435242	34108	135	36	17	10
TM1/TX231	1800	2	n/a	n/a	n/a

HinP1I Cross 1: Intraspecific PS-5 x K-56 1.1%
HinP1I Cross 2: Intraspecific PS-5x PI-435242 1.1%

BsrGI	Total Loci	1 to 3 SNPs Loci 2 parents	1 to 3 SNPs Loci 2 parents w/ progeny	PS-5 Progeny	Undomesticated Progeny
379/TM1	15314	1872	n/a	n/a	n/a
379/TX231	19247	2553	n/a	n/a	n/a
PS5/K56	15027	454	364	76	255
PS5/PI435242	10848	472	345	149	148
TM1/TX231	17215	676	n/a	n/a	n/a
A1/D5	16545	300	n/a	n/a	n/a

BsrGI Cross 1: Intraspecific PS-5 x K-56 5.1%
BsrGI Cross 2: Intraspecific PS-5x PI-435242 13.7%

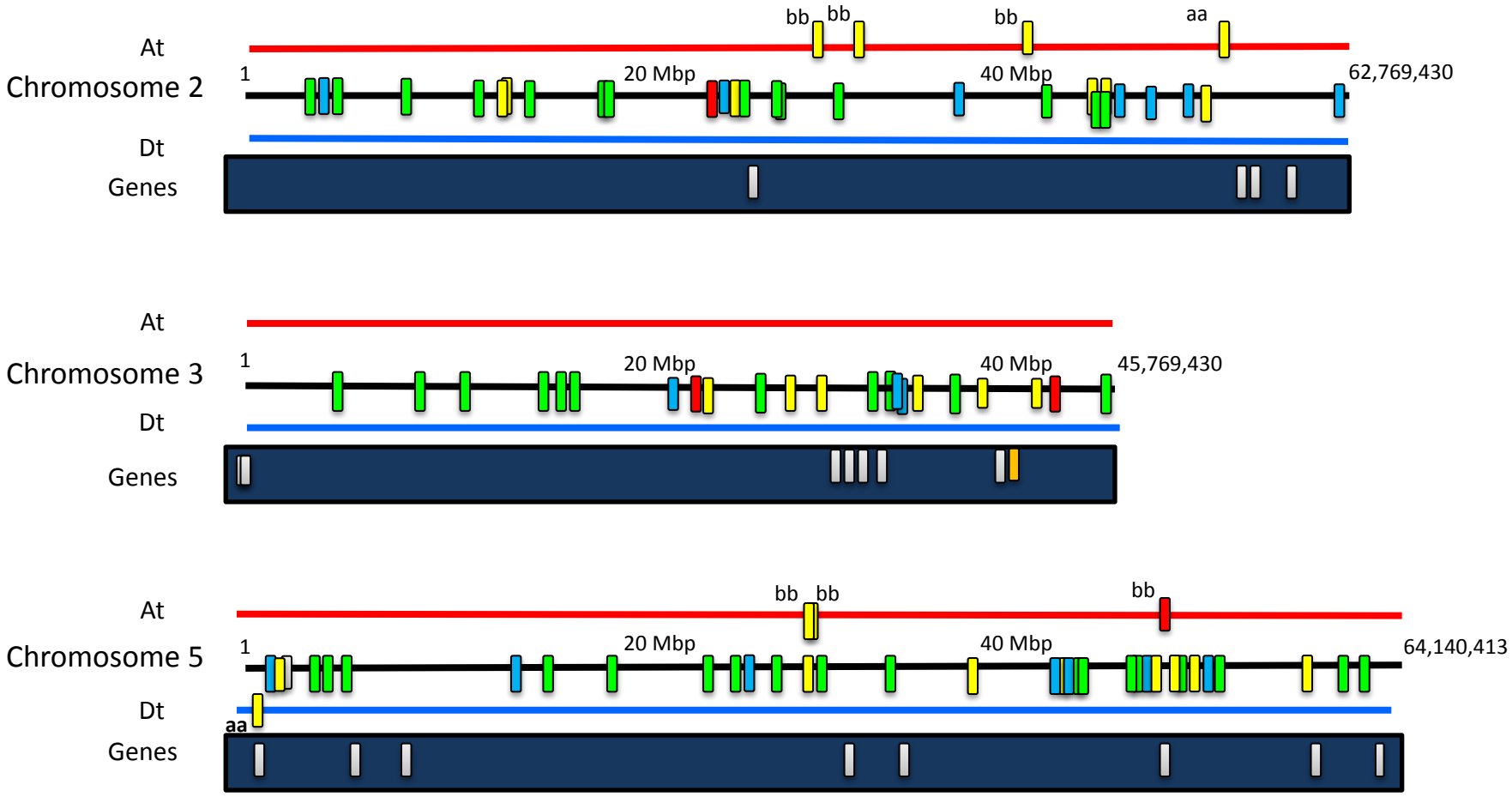
GBS Data for Heat Map

BsrGI

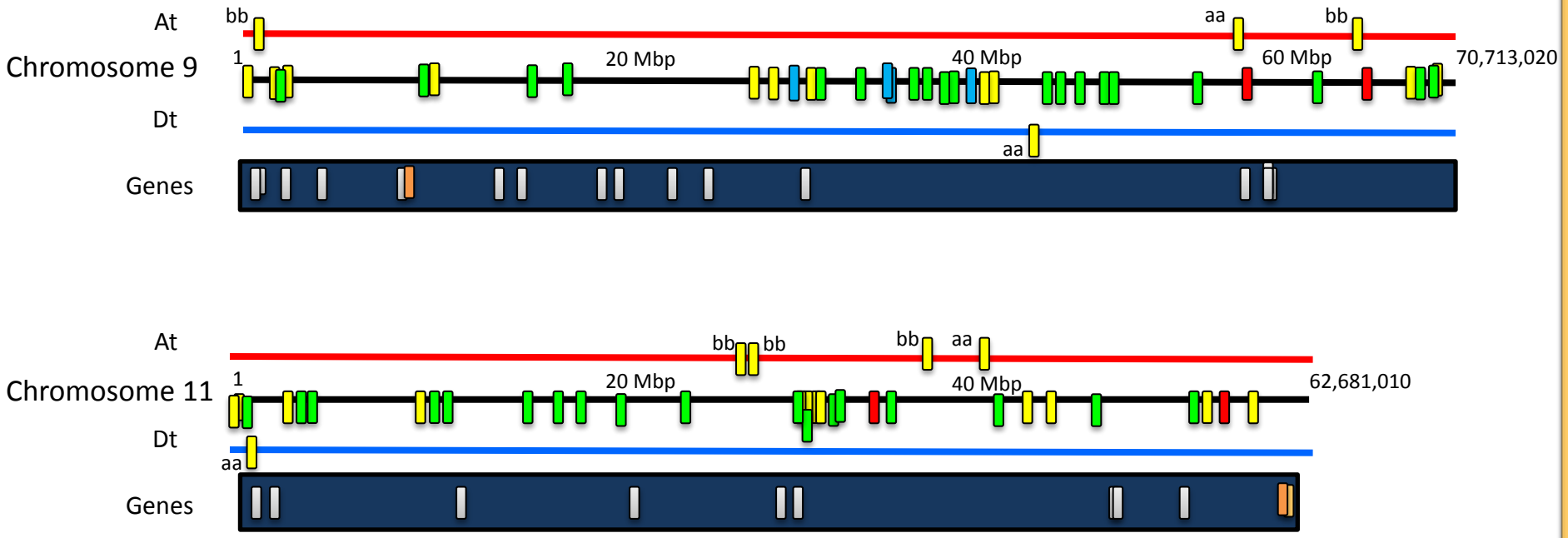
HEAT MAP

	Highest	High	Low	Lowest
SNPs with Progeny	10	211	330	42
A/D	3	64	71	8

SNPs on D_5 Chromosomes 2, 3, and 5



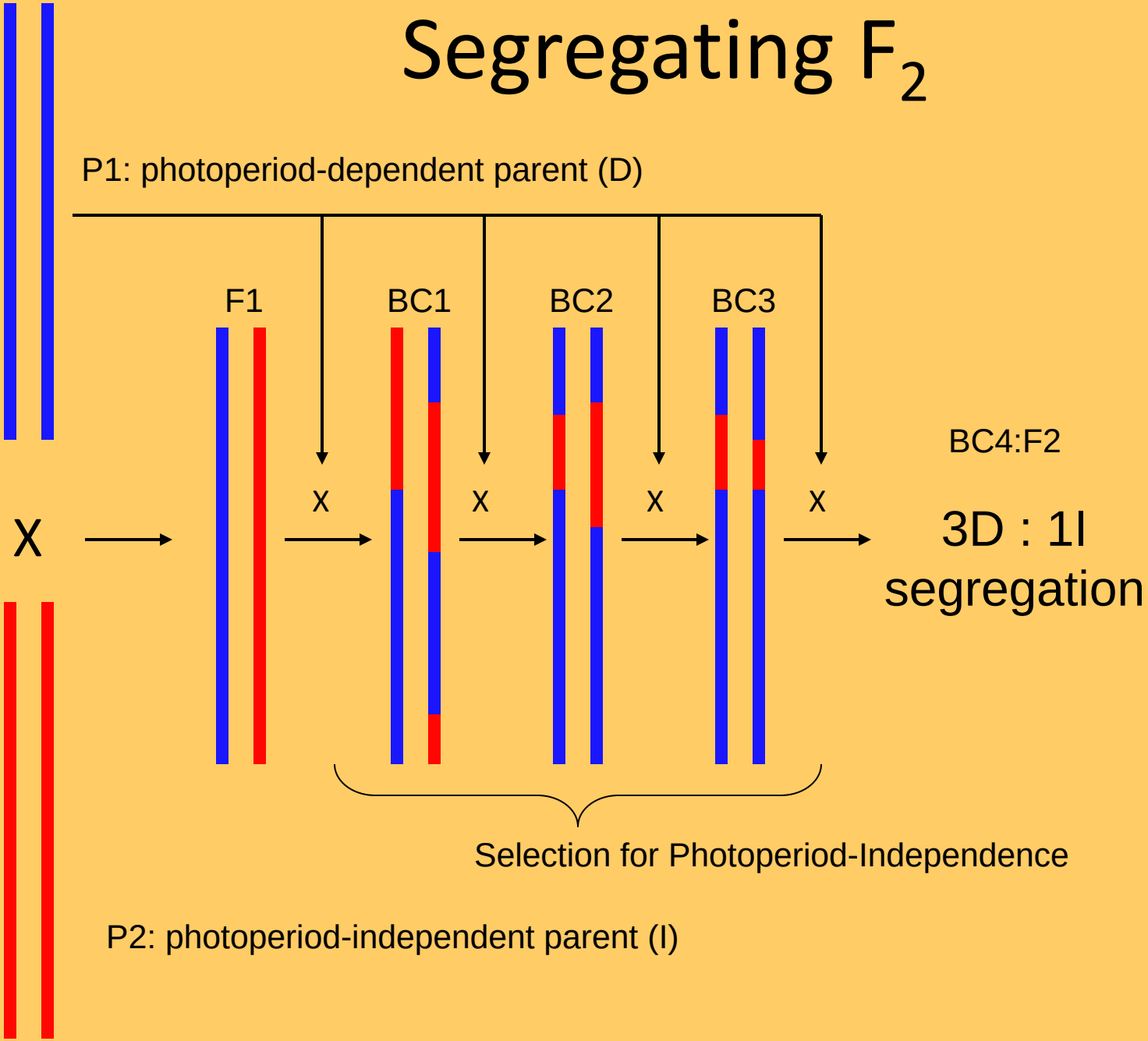
SNPs on D_5 Chromosomes 9 and 11



GBS Conclusions

- Narrow down to 9 regions implicated in photoperiodic flowering
- Allow focus on most important SNPs and discard regions with lowest probablity
- D_5 Reference Genome allows for integration of GBS and candidate gene data
- Ongoing validation of GBS SNPs associated with photoperiodic flowering

Segregating F_2



Acknowledgements

Lab Members

- Angela Hawkins
- Millie Burrell
- Dan Grum

Undergraduate Research Assistants

- Chris Lyle
- Natalie Ware
- Ryan McCormick

Contributing Lab Affliation

- Patricia Klein
- Andy Paterson

Sponsors