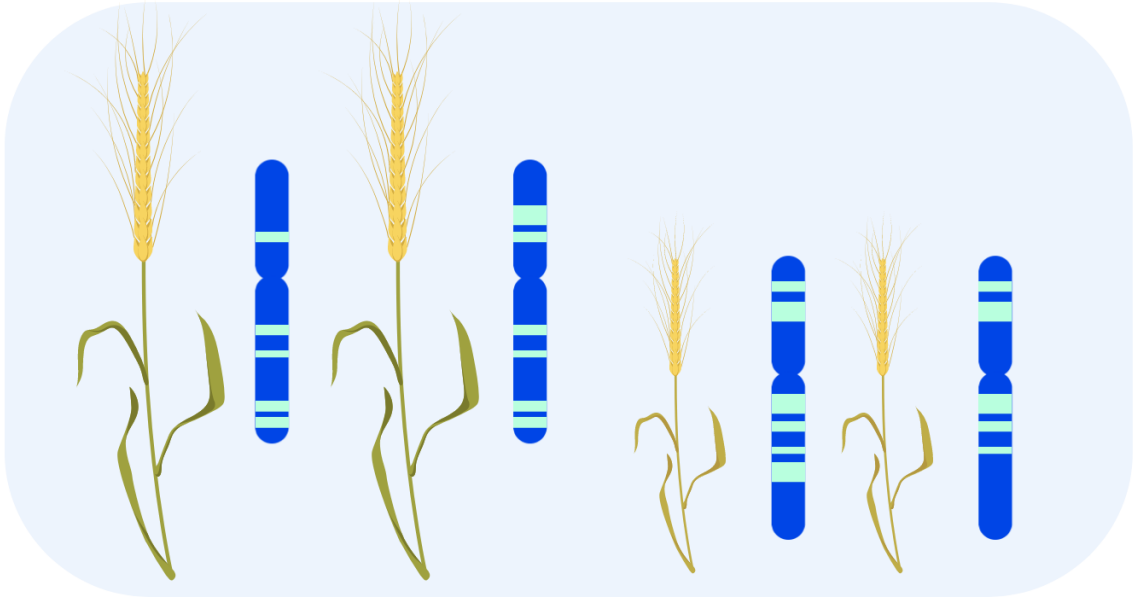
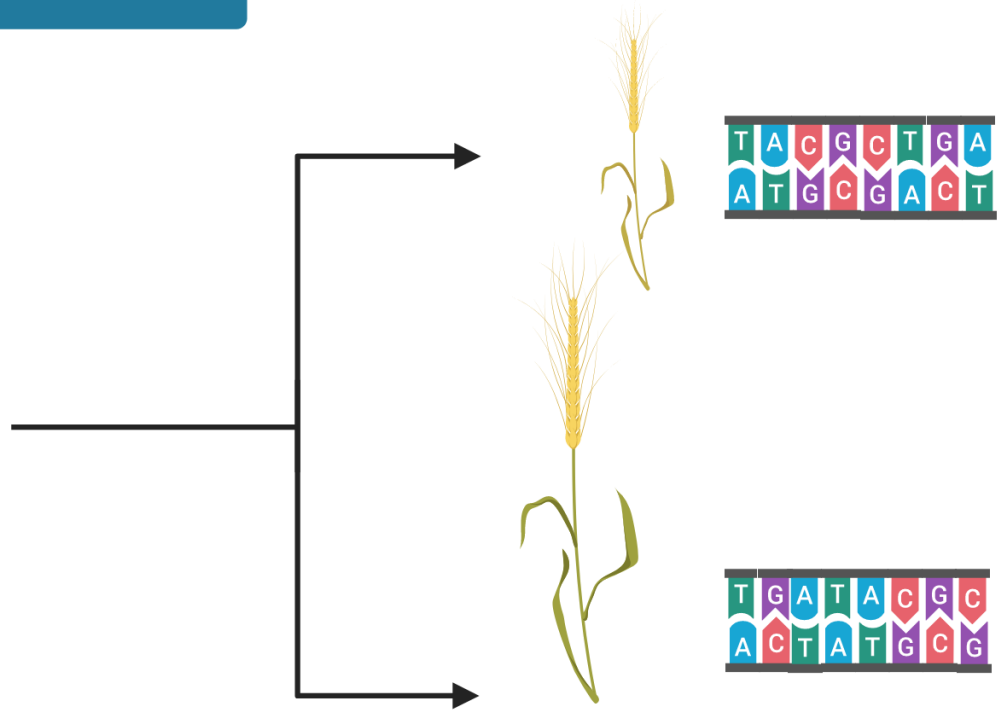
Fine mapping and cloning genes

Ellie Taagen & Shantel A Martinez

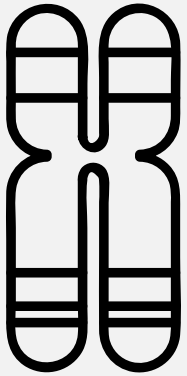
October 24, 2019



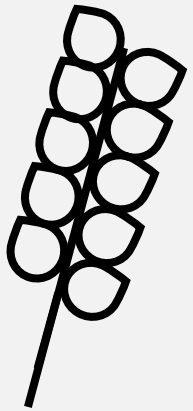


From phenotype and genotype associations

To identificaiton of a trait's causal variant sequence



Background and Basics



Practical Example

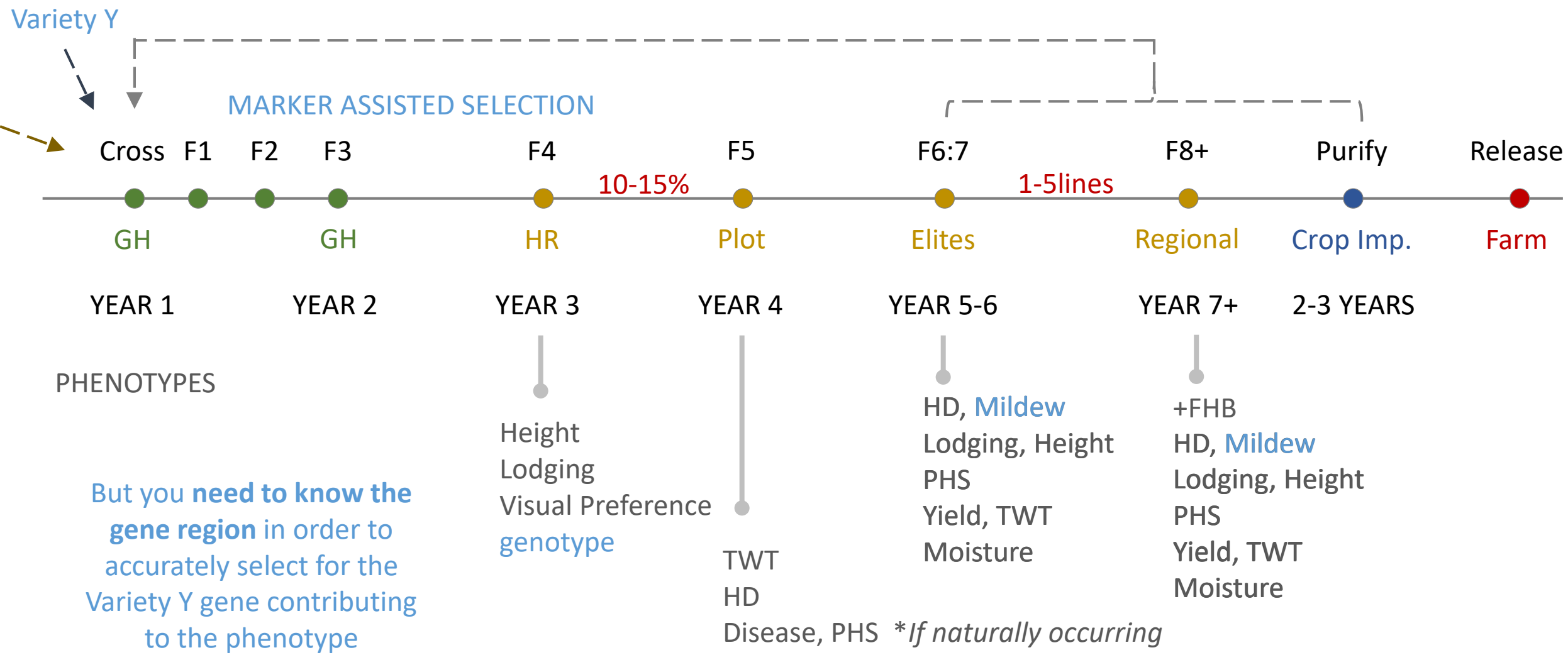
Conceptual Activity

Challenges and Solutions

Fine mapping and cloning is a method for identifying causal genetic variants driving phenotypic variants

This allows for precise selection for desirable traits in the breeding germplasm

Application: An example of traditional wheat breeding in Northeast



But you **need to know the gene region** in order to accurately select for the Variety Y gene contributing to the phenotype

Definitions of key concepts

Genetic linkage analysis

Use statistical probability to test whether a phenotype is associated with a genetic marker by chance or because the phenotype's causal variant and the genetic marker are closely located in the same genomic segment

- The phenotype is something you can measure: disease susceptibility, grain weight, gene expression, etc.
- The resolution of the linkage analysis is limited by the distribution of the genetic markers and position of genetic recombination events

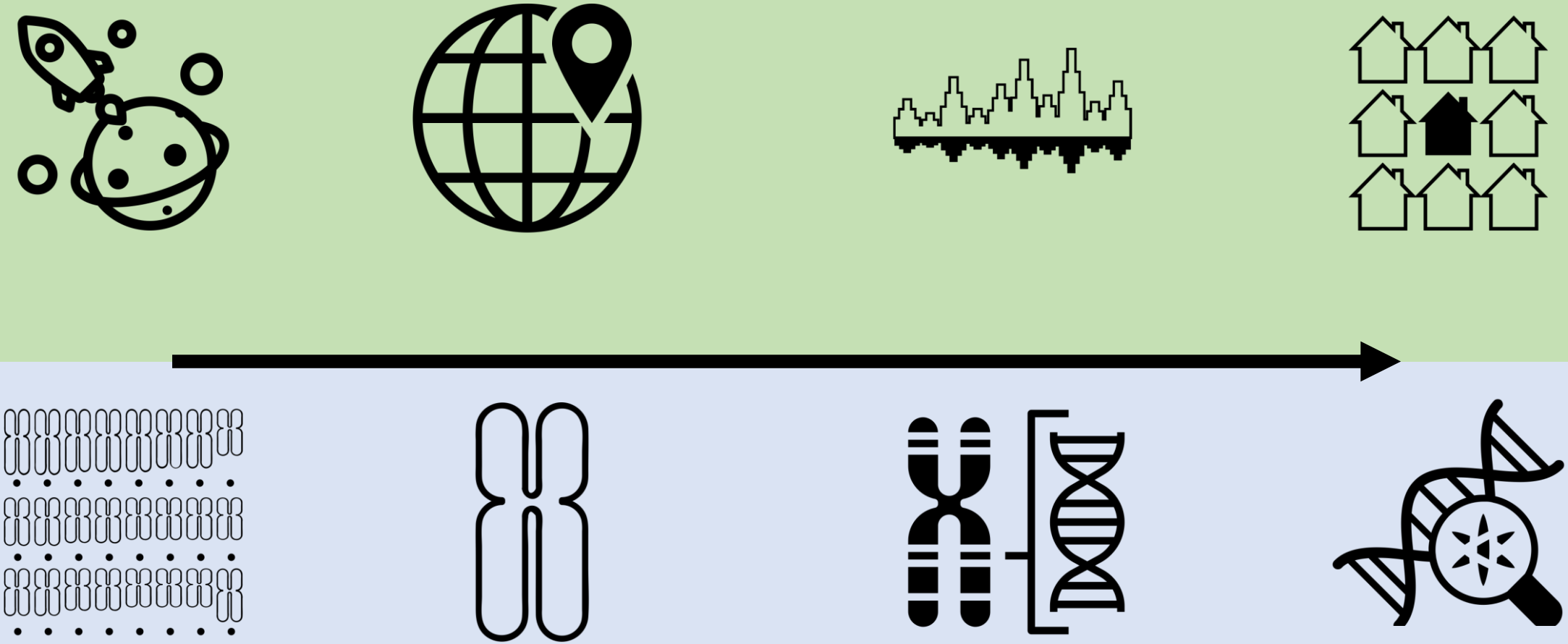
Correlation ..... *Causation*

Positional cloning

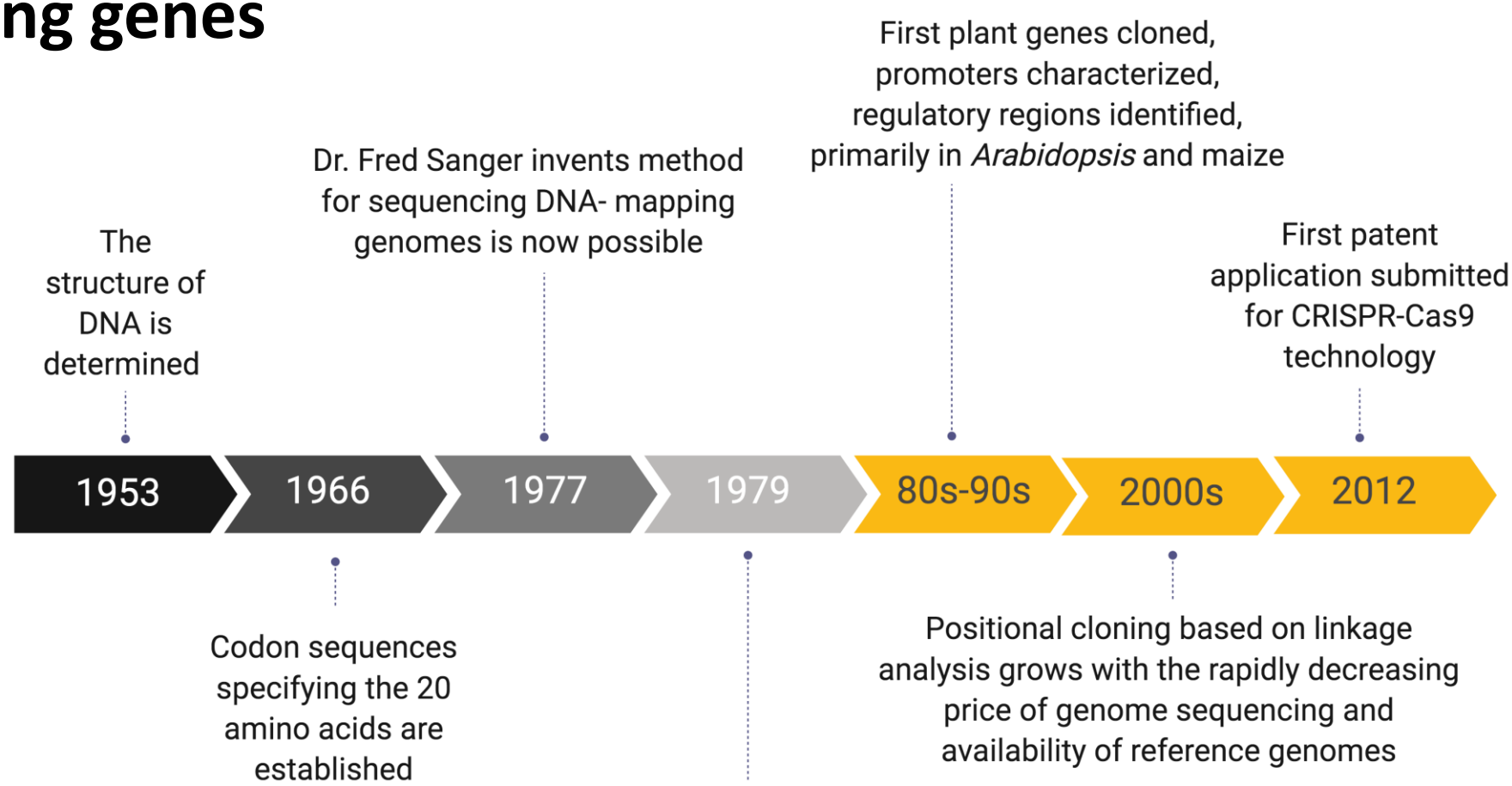
Based on the *physical position* of genetic sequence on chromosome, identify the gene and characterize the causal variant base pair(s) responsible for the phenotype

- Create a perfect marker for screening breeder's germplasm
- Vector based cloning system
  - DNA targeting tools such as the CRISPR-Cas9 system can facilitate more rapid introgression

The logic of mapping a gene

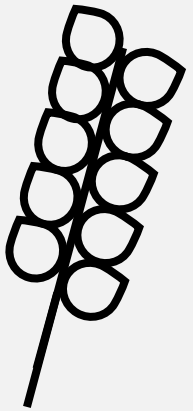


A brief history of mapping and cloning genes



Sources & recommended reading

Goldberg , Plant Physiology, 2001
The Gene, Siddhartha Mukherjee



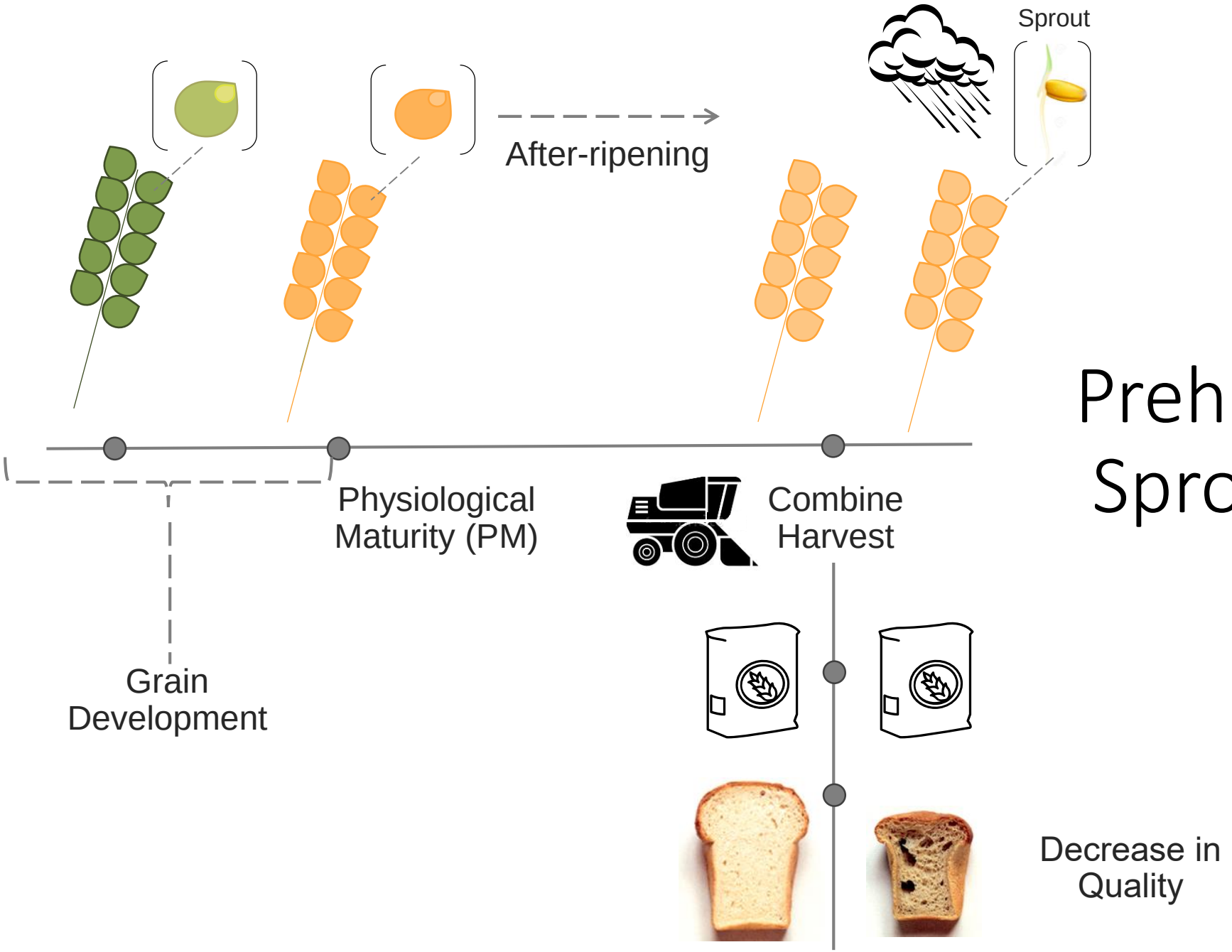
Background and Basics

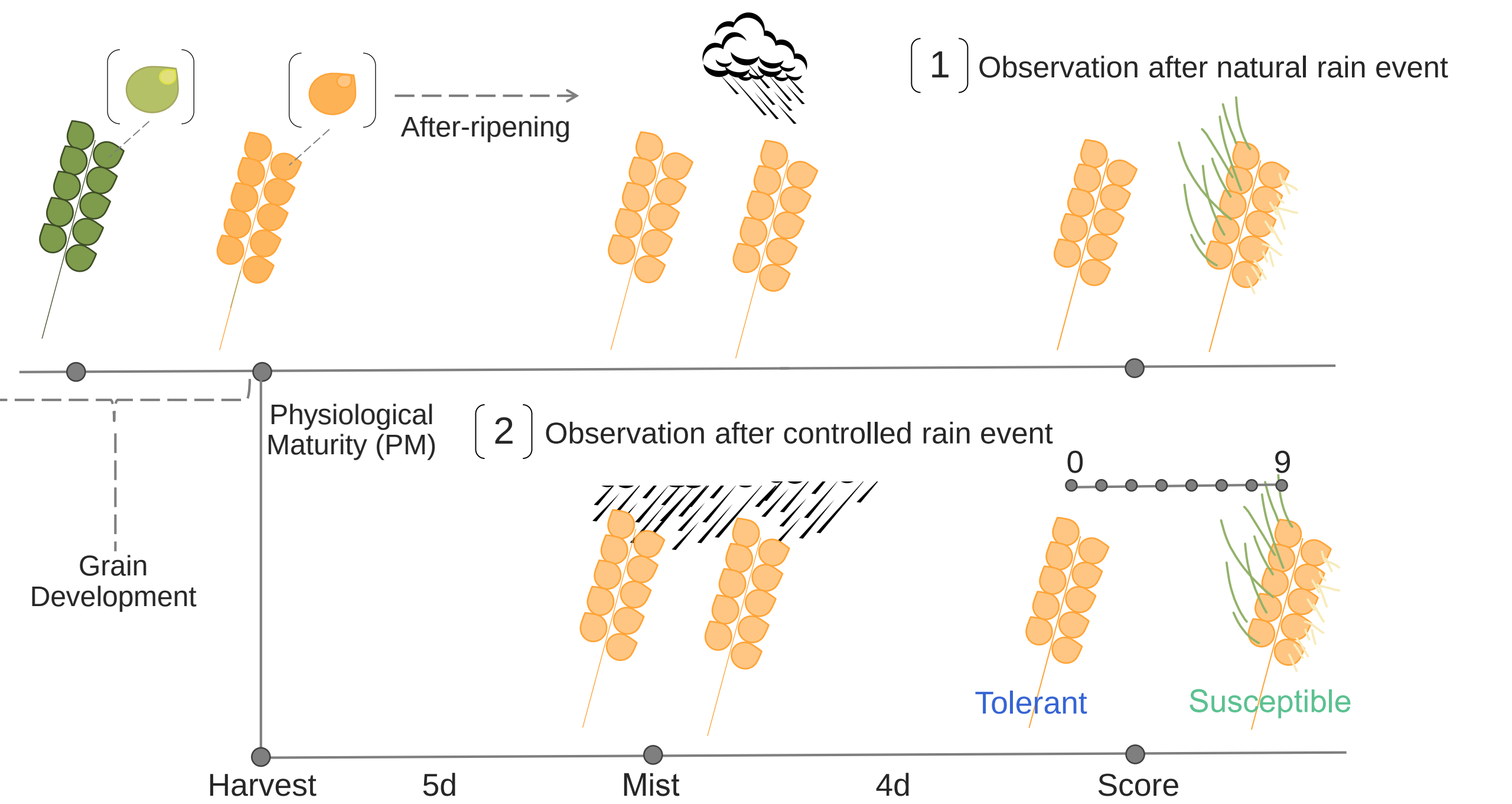
Practical Example

Conceptual Activity

Challenges and Solutions

Preharvest Sprouting




After-ripening
1 Observation after natural rain event
Physiological Maturity (PM)
2 Observation after controlled rain event
0
9
Grain Development
Tolerant
Susceptible
Harvest
5d
Mist
4d
Score

QTL mapping a PHS tolerance gene

Tolerant Line Y X Susceptible Line Z

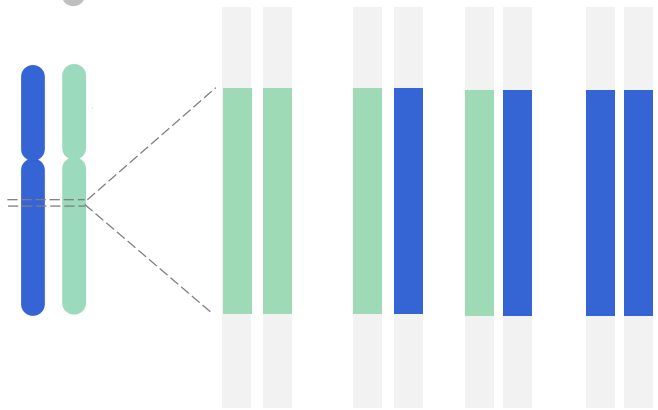
Identifies statistical association of the phenotype with the genotype

Cross F1 F2 F3 F4 F5 F6:X

GH GH HR Plot/HR Plot/HR

200-400 F2 Derived RILs

Recombination, Heterozygosity, Homozygosity

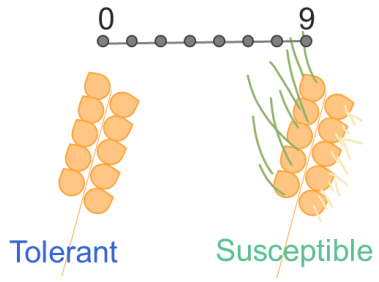


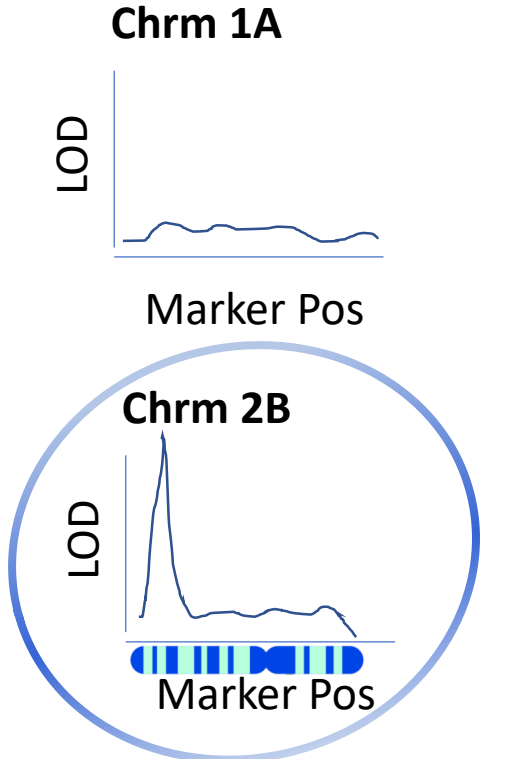
GENOTYPE + PHENOTYPE → QTL MAPPING

Rqtl is a great package to use in R

EX: Progeny 27

PHS Score
Multiple Reps/Env





QTL mapping a PHS tolerance gene

Tolerant Line Y X Susceptible Line Z

Identifies statistical association of the phenotype with the genotype

GENOTYPE + PHENOTYPE → QTL MAPPING

	BM	BN	BO	BP	BQ	BR	BS	BT	BU	BV	BW	BX	BY	BZ	CA	CB	CC	CD	CE	CF	CG	CH	CI	CJ	CK	CL	CM	CN	CO	CP
1	A43_4A_1	A12618_4	A12257_4	A7863_4A	A3343_4A	A26051_4	A25624_4	A19350_4	A471_4A_	A509_4A_	A273_4A_	A292_4A_	A10491_4	A7946_4A	A7575_4A	A1512_4A	Primer1	Primer2	A11335_4	A3059_4A	A11702_4	A11329_4	A12321_4	A12683_4	A13603_4	A13101_4	A4615_4A	A10545_4	A10092_4	A500_4A_	A28
2	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B
3	58.022	59.584	60.477	63.012	65.83	68.08	68.844	70.954	72.651	74.483	194.499	202.802	218.457	241.042	246.52	261.918	266.194	270.545	271.802	272.827	274.658	275.423	276.982	277.107	278.264	279.689	280.846	281.871	282.376	284.765	28
4	1.5	0.5	0.0	0.1	0.5	0.4	0.0	0.1	0.1	0.8	4.7	4.3	4.3	4.8	5.0	5.1	4.7	4.2	5.2	5.0	0.4	0.3	0.7	0.0	0.1	0.3	0.2	0.2	0.1	0.1	
5	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
6	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
7	A	A	A	A	A	A	A	A	A	A	B	B	B	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
8	B	H	H	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	B	B	B	B	B	B	A	A	A	
9	A	A	A	A	A	A	A	A	A	A	B	B	B	H	H	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	
10	A	A	A	A	A	A	A	A	A	A	H	H	B	A	A	A	A	A	A	A	H	H	H	H	B	B	B	B	H	H	B
11	B	B	B	B	B	B	B	B	B	B	A	A	B	H	H	B	H	.	A	H	B	B	B	B	B	B	B	B	B	B	
12	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	
13	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	
14	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	
15	B	B	B	B	B	B	B	B	B	B	B	B	B	H	H	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	
16	B	B	B	B	B	B	B	B	B	B	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	
17	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	
18	B	B	B	B	B	B	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	
19	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
20	A	A	A	A	A	A	A	A	A	A	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
21	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	B	B	H	H	H	A	A	A	B	B	B	B	H	A
22	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	.	B	B	A	A	B	B	B	B	B	B	B	B	
23	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	
24	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	
25	A	A	A	A	H	A	A	A	A	A	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
26	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
27	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	
28	B	B	B	B	B	B	B	B	B	B	A	A	A	H	H	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
29	B	B	B	B	B	B	B	B	B	B	A	A	B	B	B	B	.	.	B	B	B	B	B	B	B	B	B	B	B	B	
30	A	A	A	A	A	A	A	H	A	A	B	B	B	A	A	A	A	H	H	H	B	B	H	H	H	H	A	A	A	B	B
31	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	H	A	A	A	A	A	A	A	A	A	A	A	A	A	A	

Fine mapping the 2B QTL

Tolerant Line Y X Susceptible Line Z

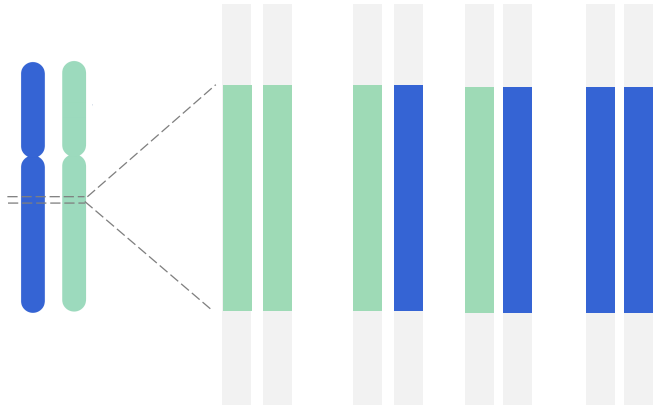
Cross F1 F2 F3 F4 F5 F6:X

GH GH HR Plot/HR Plot/HR

GENOTYPE + PHENOTYPE → FINE MAPPING

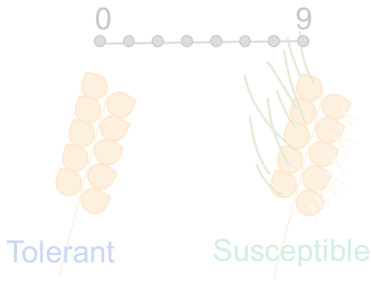
1000s F2 Derived RILs

Recombination,
Heterozygosity,
Homozygosity



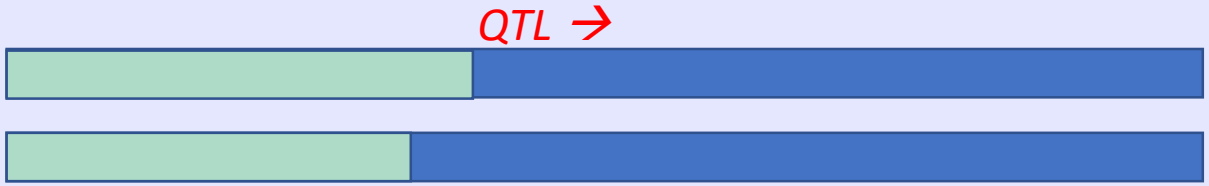
EX: Progeny 27

PHS Score
Multiple Reps/Env



Why does the mapping population **size** matter?

- There are only 1-3 crossovers per chromosome per meiosis
  - The chance that a crossover happens somewhere informative for fine mapping the QTL decreases as we zoom-in!
  - A larger population can better measure the variability of the phenotype



Not a useful crossover for fine-mapping the QTL

Why does the mapping population **structure** matter?

- Biparental (generally)

Select for heterozygosity →
CO provide finer resolution of the QTL →
heterogeneous inbred families (HIFs)

Background and Basics

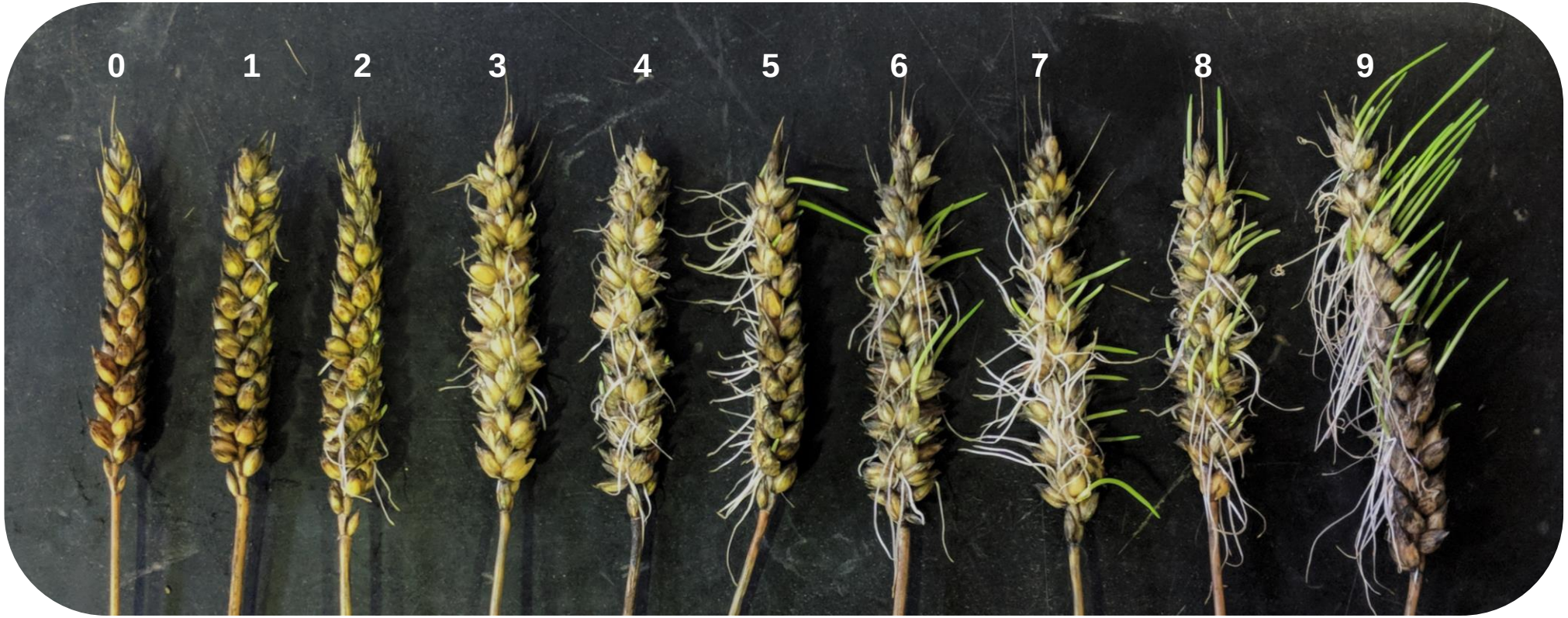
Practical Example

Conceptual Activity

Challenges and Solutions

Step 1: which varieties are PHS tolerant/susceptible?

PHS Tolerant ⟷ PHS Susceptible

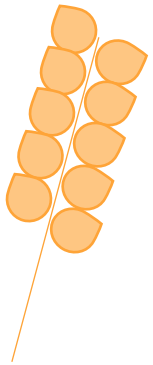


[HANDOUT 1: VARIETIES]

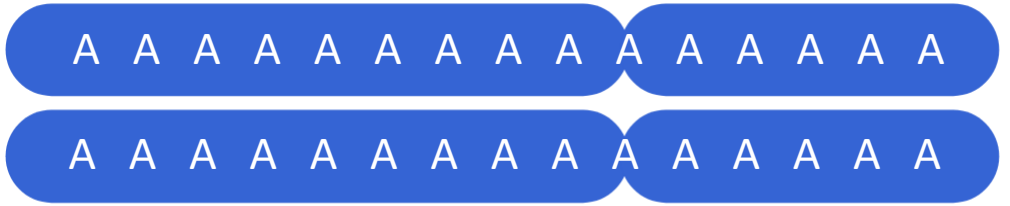
McMaster & Derera *et al*., 1976

Step 1: which varieties are PHS tolerant/susceptible?

Tolerant

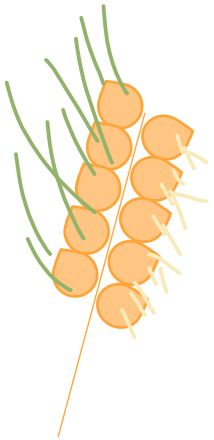


2B

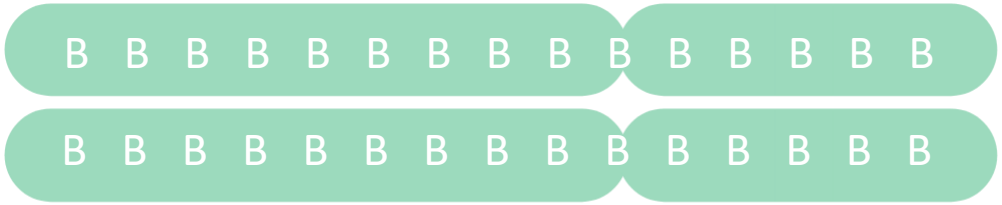


Causal Variant

Susceptible



2B

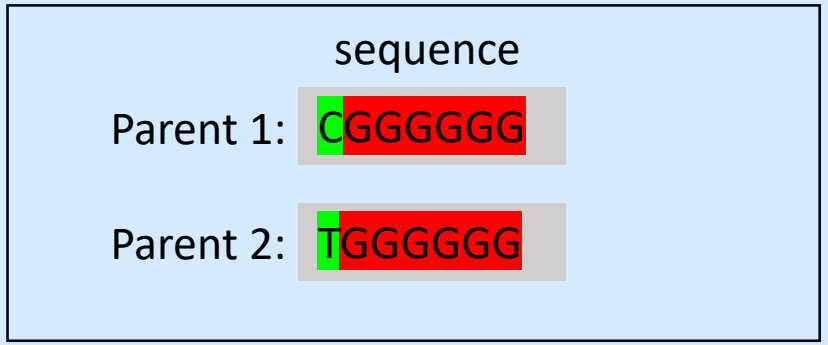


[HANDOUT 1: VARIETIES]

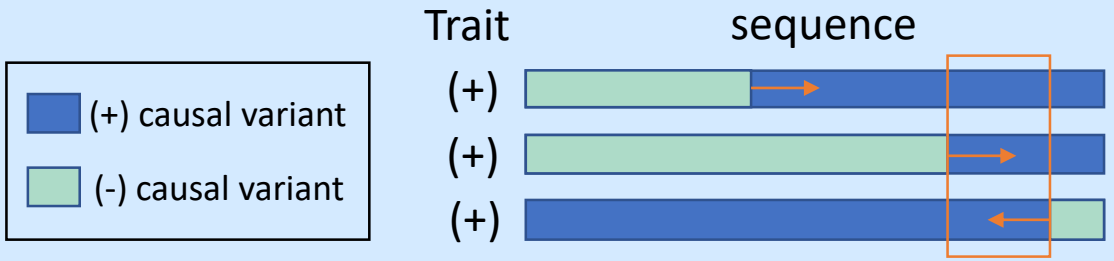
Genetic resolution can be a limiting factor

- Distribution of genetic markers (ie. SNP)

- Polymorphism



- Crossovers



Step 2: order the lines into groups based on genotypes

100 Mbp	200 Mbp	300 Mbp	400 Mbp	500 Mbp	600 Mbp	700 Mbp
SNP	SNP	SNP	SNP	SNP	SNP	SNP
B	B	B	B	A	A	A

Recombination event
(crossover event)

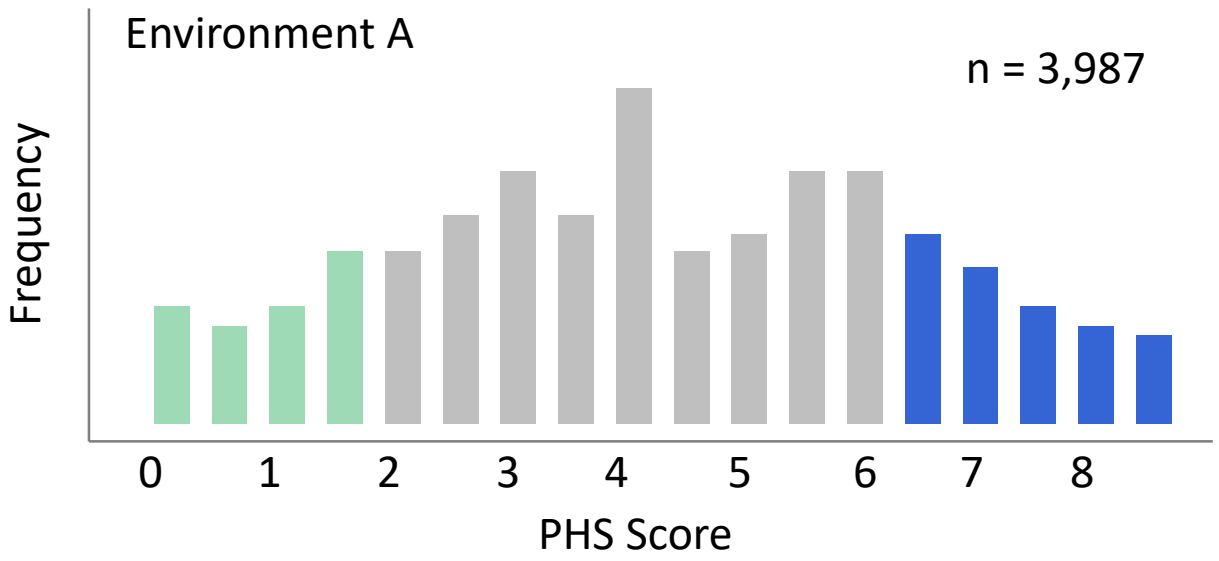
Do lines have similar recombination events?

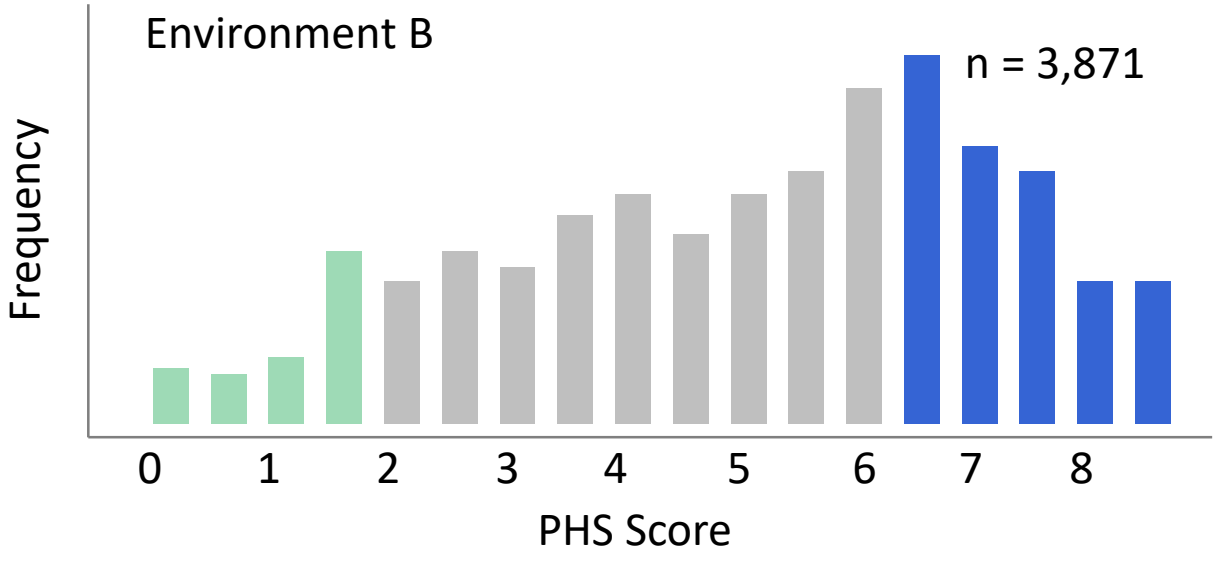
Can the recombination events be ordered along the chromosome?

We suspect AAA lines to be tolerant and the BBB lines to susceptible but we don't know yet the about the recombinants

[HANDOUT 2: GENO]

Step 3: combine geno/pheno data





The handout is only looking at a small subset of these 4,000 progeny

Write the phenotype per line on the ordered genotypes strips

[HANDOUT 3: PHENO]

Step 3: combine geno/pheno data

		AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BP	BQ	BR	BS	BT	BU	BW	BY		
		150,707,718	150,709,720	151,657,263	153,223,783	153,586,336	154,252,949	154,270,857	158,618,787	158,622,113	158,624,647	159,205,657	159,892,498	172,687,314	175,640,531	178,761,042	180,555,421	181,609,374	184,366,025	184,403,048	184,512,531	184,661,746	269,375,959	BC1F5	BC1F5	BC1F6	BC1F7	BC1F7	BC1F8	BC1F8	BC1F9		
		2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B	2B										
Source	Entry	snp1763	tarc2079	tarc2082	AX-94884418	tarc2090	AX-95022624	AX-94750768	AX-95018481	AX-94633037	AX-94549797	ss26.3/Cnl414	tarc2074	WMC474	AX-94487047	GBS-178.7	GBS-180.5	GBS-181.6	GBS-184.36	GBS-184.4	GBS-184.5	AX-94451130	AX-94668928	Ket09	Sny09	McG10	Ket11	Sny11	Hel18	Sny18	Cal19		
1815-3	BC1F5-181(HR)	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	A	A	3.6	2.5	4.3	4.2	3.2	3.7	NA	3		
1839-1	BC1F5-878(HR)	B	B	B	B	B	B	B	B	B	B	B	B	A	B	A	A	A	A	A	A	B	B	2.4	4.6	3.7	5.5	NA	NA	NA	6.6		
1527-5	BC1F5-307(HR)	A	A	A	A	A	A	A	A	A	A	A	A	B	A	B	B	B	B	B	B	A	A	3.2	3.0	4.9	5.8	1.6	8.2	7.8	NA		
0068-4	BC1F5-465(HR)	B	B	B	B	B	B	B	B	B	B	B	B	B	A	B	B	B	B	B	B	A	A	0.8	3.8	5.3	5.6	6.0	NA	NA	6.4		
0213-2	BC1F5-417(HR)	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	B	A	A	A	A	A	A	2.8	3.8	5.7	4.4	6.4	5.2	9.0	7.4		
0472-3	BC1F5-183(HR)	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	1.8	3.4	2.5	3.0	2.2	4.8	NA	1.6		
0035-4	BC1F5-169(HR)	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	H	B	B	1.2	1.4	2.3	1.4	0.3	4.4	4.4	4.4		
0158-3	BC1F5-738(NR-cay)	A	A	A	A	A	A	A	A	A	A	-	A	-	A	A	A	A	B	A	A	A	A	2.6	0.8	2.3	1.8	4.4	4.6	4.0	2.6		
0394-3	BC1F5-306(HR)	B	B	B	B	B	B	B	B	B	B	A	B	A	A	-	B	B	B	B	B	A	B	3.6	3.8	6.1	3.4	4.0	8.0	6.6	7.6		
1832-1	BC1F5-105(HR)	A	A	A	A	A	A	A	A	A	A	A	B	B	B	A	A	A	A	A	A	B	B	2.2	4.6	5.7	5.0	4.4	7.2	3.6	NA		
1531-1	BC1F5-422(HR)	A	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	A	A	A	A	B	B	4.0	3.9	4.7	5.4	5.0	8.2	7.4	NA		
1527-3	BC1F5-320(HR)	B	B	B	B	B	B	B	B	B	B	A	B	B	B	A	A	-	A	A	A	B	B	4.0	2.7	4.5	4.8	6.5	7.0	4.0	NA		
0035-2	BC1F5-135(HR)	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	-	A	A	A	3.0	0.6	1.9	0.2	0.6	5.6	1.6	3.4		
0551-1	BC1F5-238(HR)	B	B	B	B	B	B	B	B	B	B	B	B	A	A	B	B	B	B	B	B	A	B	2.4	3.4	3.1	5.2	6.2	4.8	7.2	7.2		
0454-4	BC1F5-703(HR)	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	A	A	3.6	4.4	NA	2.4	0.6	7.0	5.0	3.8		
0157-4	BC1F5-807(HR)	A	A	A	A	A	H	A	A	A	A	A	-	A	A	A	A	A	A	A	A	B	A	2.0	2.1	4.1	3.8	2.6	4.0	7.4	4.2		
1527-4	BC1F5-251(HR)	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A	A	A	A	A	A	A	NA	3.5	4.9	4.8	6.0	NA	NA	2.5		
0996-1	BC1F5-325(HR)	A	A	A	A	A	A	A	A	A	A	B	A	B	A	A	A	A	A	A	A	A	A	2.8	2.3	4.3	3.2	3.9	4.0	4.8	2.8		
0574-3	BC1F5-331(HR)	B	B	B	B	B	B	B	A	A	A	B	A	B	A	A	A	A	A	A	A	A	A	3.8	3.2	5.7	2.2	5.4	7.4	5.8	5		
0435-1	BC1F5-353(HR)	B	B	B	B	B	B	B	A	A	A	B	A	B	A	A	A	A	A	A	A	A	A	2.4	4.4	2.9	3.1	2.3	7.2	6.6	6		
0213-4	BC1F5-174(NR-cal)	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A	A	A	A	A	A	A	3.4	5.8	4.9	3.6	0.6	3.8	4.0	5		
0213-1	BC1F5-45(NR-cal)	B	B	B	B	B	B	B	A	A	A	A	A	B	A	A	A	A	A	A	A	A	A	4.6	4.8	5.7	2.6	3.6	5.0	6.4	4		
0082-5	BC1F5-790(HR)	A	A	A	A	A	A	A	A	A	A	B	A	B	A	A	A	A	A	A	A	A	B	5.0	5.4	6.7	6.6	3.8	4.6	NA	7.6		
0278-3	BC1F5-809(NR-cal)	A	A	A	A	A	A	A	A	A	A	-	A	B	A	A	A	A	A	A	A	A	A	1.4	4.8	2.9	4.6	3.5	6.8	NA	2.2		
1276-5	BC1F5-429(HR)	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	2.8	2.2	4.5	3.8	2.4	7.8	NA	6.6		
0551-5	BC1F5-6(HR)	B	B	B	B	B	B	B	B	B	B	B	B	A	B	B	B	B	B	B	B	B	A	2.0	3.2	6.5	5.4	6.0	6.8	5.8	6.6		
0551-3	BC1F5-415(HR)	B	B	B	B	B	B	B	B	B	B	B	B	A	B	B	B	B	B	B	B	B	A	NA	2.6	2.3	3.4	2.9	7.6	7.2	2.8		
	BC1F5-31(HR)	B	B	B	B	B	B	B	B	B	B	B	B	A	B	B	B	B	B	B	B	B	A	3.8	4.4	4.7	5.2	6.8	6.0	8.2	7.8		
	BC1F5-209(HR)	B	B	B	B	B	B	B	B	B	B	A	B	A	B	B	B	B	B	B	B	B	B	2.6	3.2	4.7	4.2	3.4	8.6	NA	8		
	BC1F5-182(HR)	B	B	B	B	B	B	B	B	B	B	B	B	A	B	B	B	B	B	B	B	B	B	2.0	5.0	5.3	2.8	5.3	4.0	6.0	NA		
	BC1F5-933(HR)	B	B	B	B	B	B	B	B	B	B	A	B	A	B	B	B	B	B	B	B	B	A	1.0	1.4	3.9	2.8	3.8	5.0	NA	3.8		
	BC1F5-499(HR)	B	B	B	B	B	B	B	B	B	B	A	B	A	B	B	B	B	B	-	B	B	B	1.6	2.8	4.1	2.8	2.2	6.0	5.4	3.6		
	BC1F5-674(HR)	B	B	B	B	B	B	B	B	B	B	A	B	A	B	B	B	B	B	B	B	B	A	4.0	2.8	4.7	5.6	3.6	4.2	NA	8.4		
	BC1F5-189(HR)	B	B	B	B	B	B	B	B	B	B	A	B	A	B	B	B	B	B	B	B	B	B	2.2	4.2	4.5	4.6	3.0	NA	NA	6.2		
0378-3	BC1F5-900(HR)	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	1.4	2.4	4.5	3.0	0.4	4.4	5.6	6		
0378-2	BC1F5-548(HR)	B	B	B	B	B	B	B	B	B	B	B	B	A	A	A	A	A	A	A	A	A	A	1.4	2.2	3.3	2.4	2.3	6.0	5.4	6		
1405-5	BC1F5-132(HR)	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	-	B	B	B	3.2	4.5	5.3	7.0	2.7	7.4	7.2	6.8		
0969-1	BC1F5-88(HR)	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	3.2	3.5	5.3	6.0	2.5	6.4	NA	NA		
0752-4	BC1F5-211(HR)	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	4.4	3.9	5.9	5.4	4.7	NA	NA	8.8		
0752-2	BC1F5-365(HR)	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	0.6	4.4	5.9	2.8	3.3	8.0	6.2	NA		
0035-3	BC1F5-139(NR-cal)	B	B	B	B	B	B	B	B	B	B	-	B	B	B	B	B	B	B	B	B	B	B	4.4	6.2	5.7	5.2	1.8	NA	NA	7.4		
Cayuga	Cayuga	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1.6	0.2	3.7	2.6	3.4	3.0	2.2	2		
Cayuga	Cayuga	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	A	A	A	A	-	-	2.2	1.0	1.5	1.8	2.8	1.6	1.4	2.4		
Cayuga	Cayuga	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1.8	2.2	5.3	0.6	0.6	4.5	1.8	2.6		
Cayuga	Cayuga	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1.8	0.4	1.9	NA	2.5	2.8	3.0	2		
Cayuga	Cayuga	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	-	A	A	2.6	2.0	3.3	2.2	0.2	4.6	0.8	3.2		
Cayuga	Cayuga	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1.2	1.6	3.1	0.4	1.8	1.7	1.0	1		
Caledo	Caledonia	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	-	H	B	B	B	3.2	5.4	4.9	2.4	4.2	8.4	NA	7.6		
Caledo	Caledonia	-	-	-	-	-	-	-	-	-	-	-	-	-	-	B	B	B	B	B	B	-	-	4.8	5.0	5.3	4.2	7.2	5.2	6.0	9		
Caledo	Caledonia	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	3.8	4.0	6.5	7.2	5.4	7.4	7.4	8.8		
Caledo	Caledonia	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	5.2	6.8	7.1	5.4	3.8	7.0	8.2	9		
Caledo	Caledonia	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	4.4	6.8	5.1	5.5	7.6	8.0	7.6	8.6		
1905-1	BC1F5-627(NR-cay)	A	A	A	A	A	A	A	A	A	A	-	A	-	A	A	A	A	A	A	A	A	B	1.2	1.3	1.9	4.6	3.9	2.0	NA	NA		
1887-2	BC1F5-822(NR-cay)	A	A	A	A	A	A	A	B	B	B	-	B	-	B	B	B	B	B	B	B	B	B	1.6	1.3	3.9	2.8	2.5	1.6	3.4	NA		

Step 4: examine the progeny's geno/pheno

- What phenotype pattern do you see in association with the genotypes? Which haplotype is always susceptible? Tolerant?

- Can you identify where there is a critical crossover event? Highlight the two flanking genetic markers.

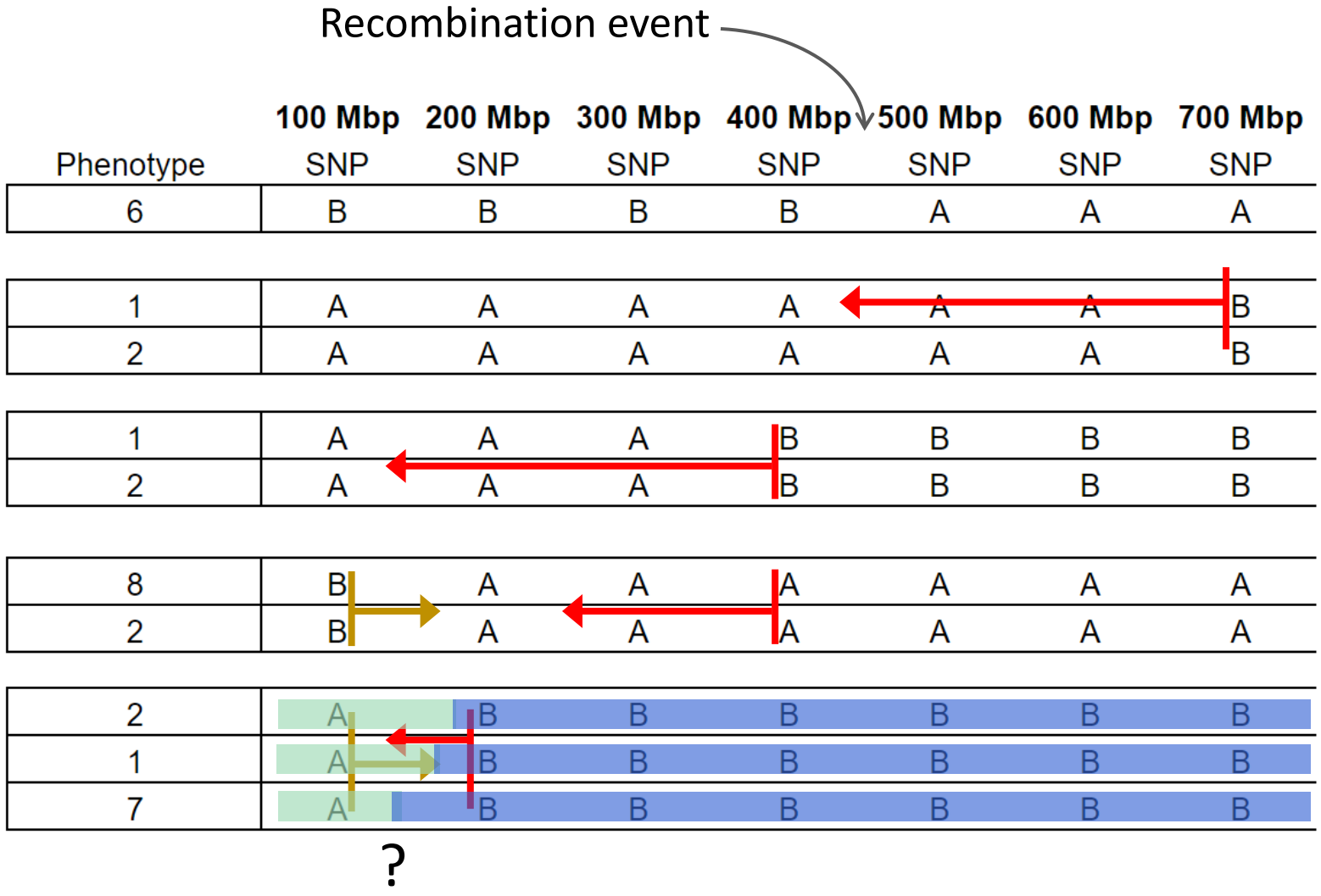
[HANDOUT 3: PHENO]

Having Cayuga markers (A) up until 700Mb results in a Cayuga PHS Tolerant phenotype (1,2)

Same effect: up to 400Mb

Recombination event between 100-200Mb shows either a Tol or Susc phenotype

Suggesting that a gene between 100-200Mb is associated with PHS tolerance



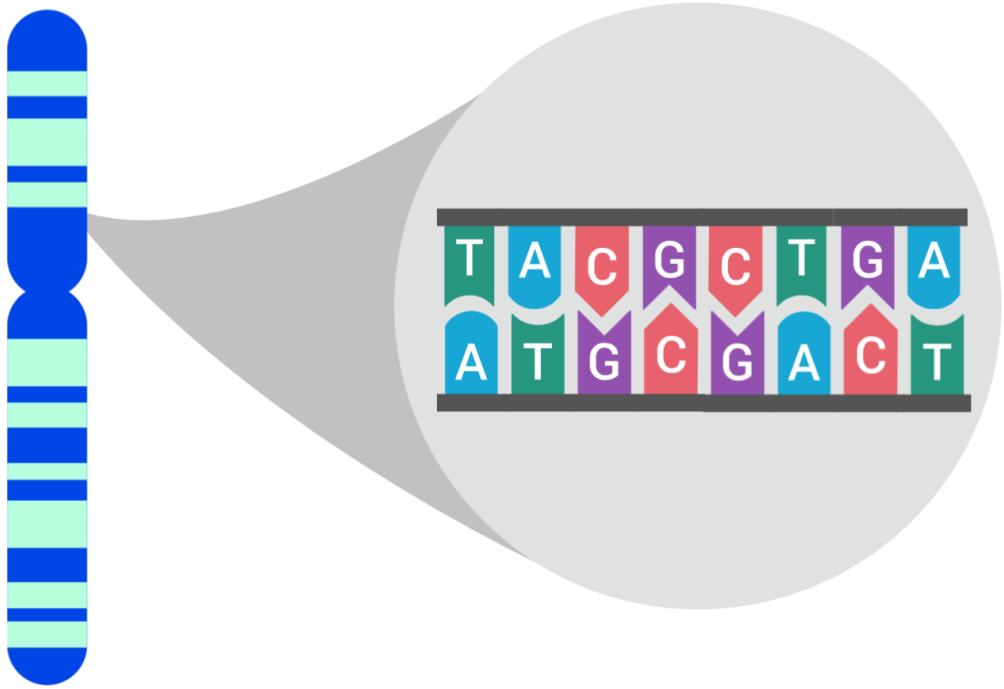
However, we currently do not have enough markers in that region to see which gene in the region provides tolerance

Step 4: examine the progeny's geno/pheno

- What phenotype pattern do you see in association with the genotypes? Which haplotype is always susceptible? Tolerant?
- Can you identify where there is a critical crossover event? Highlight the two flanking genetic markers.
- You don't have the resources to phenotype everything - so pick only the lines that have a crossover between to critical flanks or are heterozygous

Designing genetic markers

We want to zoom in on the sequence between the critical flanking positions …
…AKA identify positions where we can design genetic markers!



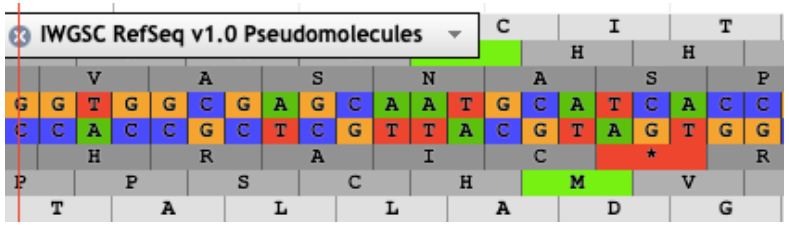
Useful genetic markers are both:

- Polymorphic for your parent lines
- Chromosome specific (and genome specific for polyploid species)

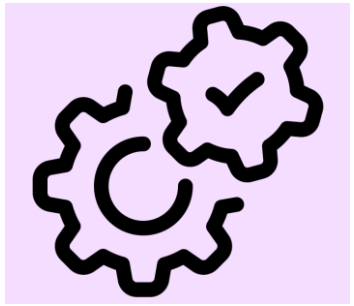
Designing genetic markers

- We put our resources into developing 10 new genetic markers between the critical flanking positions

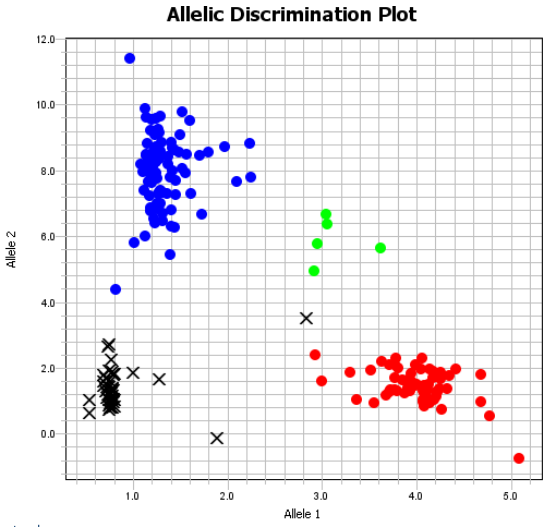
Quick how-to for KASP markers:



After you ID a polymorphic & site specific SNP, ie. 90K, grab sequence surrounding SNP from reference genome



Use tools like Oligo-Calc and NIH Primer Blast to optimize primer design



Test the markers- 3 primer fluorescent system detects alleles A, B and AB

Step 5: the mapping process is iterative

- We genotype the progeny of the plants we advanced with the 10 new markers and measure their PHS phenotype
- Look at your new results - what phenotype pattern do you see in association with the genotypes? Which haplotype is always susceptible? Tolerant?
- Can you identify where there is a critical crossover event? Highlight the two flanking genetic markers.

[HANDOUT 4: GENO+PHENO]

Step 6: prioritize list of genes

A: Look at the **polymorphism between your parental lines** in the EcoTILLING browser.
Prioritize genes that have truncation mutations or changes with negative BLOSUM 62 scores in conserved regions (THIS CAN MISS REGULATORY DIFFERENCES)

B: Look within **expression databases** for the expression profile of the genes.
expVIP WheatExp Wheat eFP
Prioritize genes that are expressed in your target tissue and developmental time

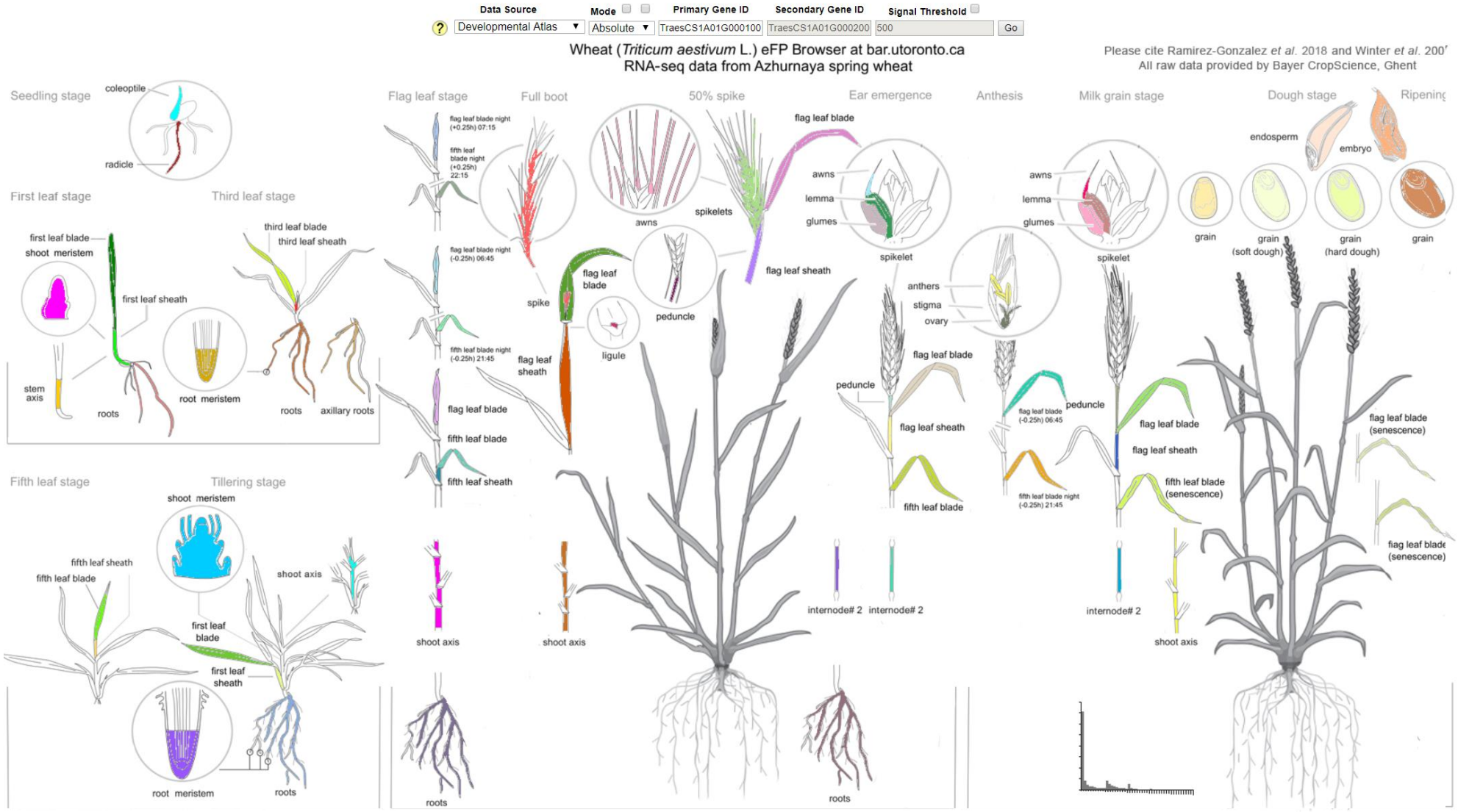
C: Find the **collinear** and closest genes in Rice, Brachy, Barley, Maize.
Find the closest gene in Arabidopsis. Record the names and numbers. If there are close paralogs, you may need to generate a phylogenetic tree to find the closest homologs.
Do a literature search to learn about the function of this gene or gene family:
Starting with barley->brachy->rice->maize and then Arabidopsis.
Transcription factors are genes regulating protein stability are always good candidates
Most of the QTLs cloned in wheat so far are in these categories

D: Prepare a small **document** for each gene describing:
Conserved domains, function, references, expression profile,
polymorphisms between parental lines.
Integrate that info into a Table and prioritize your list

Step 6: prioritize list of genes bit.ly/35UqEXP

Search the gene name(s):
Do some of the genes show no expression (yellow) in the tissue you would except it to be present in?

Check the entire list of genes

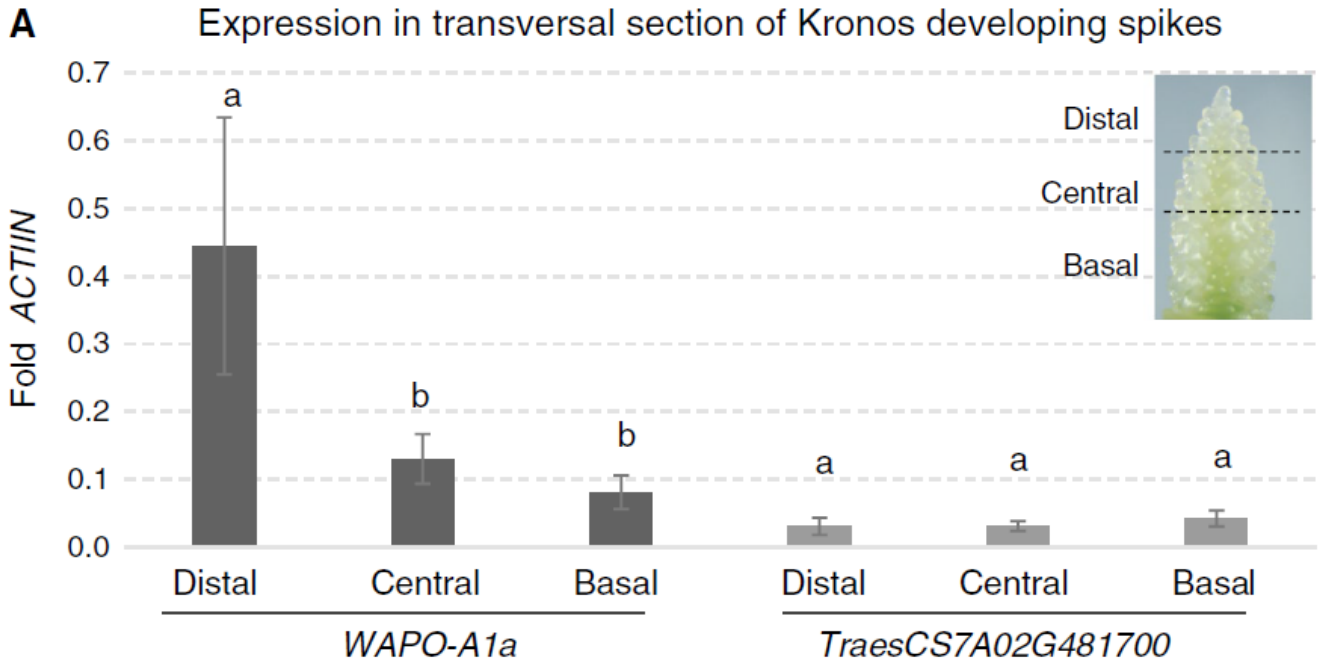


[HANDOUT 5: GENE LIST]

Step 7: gene validation

Check for gene expression within your lines and tissue relevant to the trait

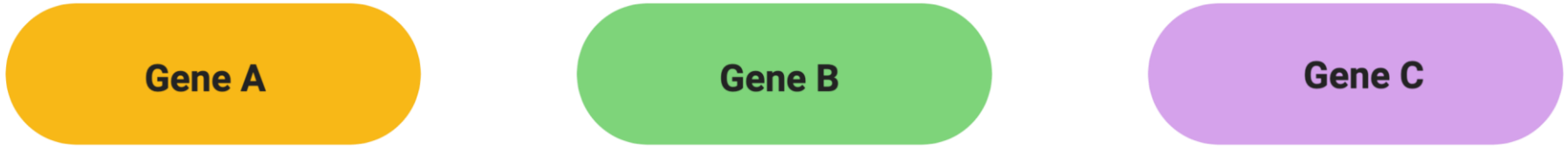
Fig. 4 Transcript levels in developing spikes at the floret primordia stage relative to *ACTIN* as endogenous control. **a** *WAPO-A1* and *TraesCS7A02G481700* in basal, central and distal sections of Kronos developing spikes. The experiment was repeated twice, and data were analyzed together using experiment as block. Means within each gene were compared using Tukey tests (=0.05). Bars are SE of the means. **b** Transcript levels of *WAPO-A1* in hexaploid near isogenic HIF lines homozygous for the *WAPO-A1b* and *WAPO-*



Kuzay et al., 2019 (TAG)

Step 7: gene validation

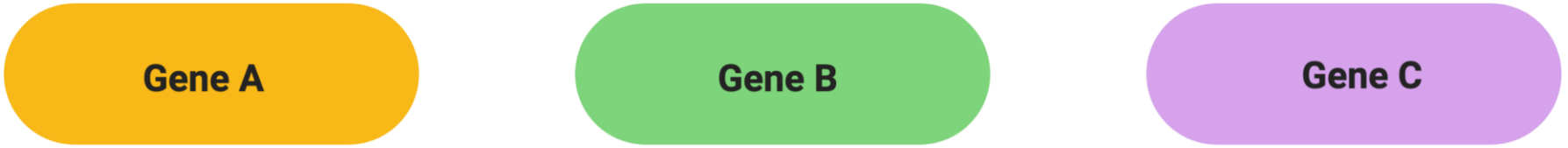
- You look at your research budget and there is only enough funding to pursue **three** genes for knockouts with CRISPR-Cas9



- Pick the plants you want to transform: which allele are we trying to knockout?

- What phenotype do you expect?

Step 7: gene validation



Phenotype	Treatment	Treatment	Treatment

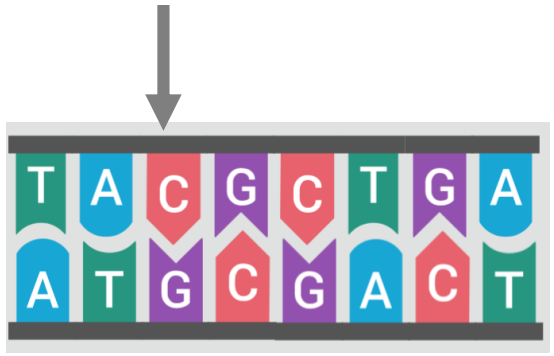
Step 7: gene validation

Phenotype	Treatment	Treatment	Treatment
1	**KO**	**WT**	**WT**
7	**WT**	**KO**	**WT**
2	**WT**	**WT**	**KO**

Step 8: positional cloning

- Sequence the gene from both + and - phenotype
- Identify the causal variant in the sequence (ie. is it a nucleotide shift in the promoter)

Tolerant +



Susceptible -

Theor Appl Genet (2014) 127:855–866

TaSdr-B1a GCCCTTCACTGGACGCCACTGGAATCCACAGTCCTCTCCCTCCAAAGCAGCGCGCCCCGC 60
TaSdr-B1b GCCCTTCACTGGACGCCACTGGAATCCACAGTCCTCTCCCTCCAAAGCAGCGCGCCCCGC 60

Sdr-5 Forward primer

TaSdr-B1a GACTCGCCTC**CGCCTACGTGTCGGCCC**CGTCCCCGCCCGCTCGCCCACGTACCCCGCGCC 120
TaSdr-B1b GACTCGCCTCCGCCTACGTGTCGGCCCCGTCCCCGCCCGCTCGCCCACGTACCCCGCGCC 120

***Pml*I recognition site**

TaSdr-B1a TCGTTCC*CACGTG*CCCCTCCCTCTGCGCGCATCCGATTGGCCGCCCACGCCTTCTTAAGC 180
TaSdr-B1b TCGTTCC*CACGTG*CCCCTCCCTCTGCGCGCATCCGATTGGCCGCCCACGCCTTCTTAAGC 180

TaSdr-B1a CGGCACGGCACCGGGACCCAACGCCGTGCACTCCGTCCACCCCCGTCAGCAGACTTCGAC 240
TaSdr-B1b CGGCACGGCACCGGGACCCAACGCCGTGCACTCCGTCCACCCCCGTCAGCAGACTTCGAC 240

***Pml*I recognition site** **initiation codon**

TaSdr-B1a TCG*CACGTG*CACGCA**ATG**GCCATGGTGCAGCCGGCGGACATGGCCGTCAAGGCCAACGAG 300
TaSdr-B1b TCGCGCGTGCACGCA**ATG**GCCATGGTGCAGCCGGCGGACATGGCCGTCAAGGCCAACGAG 300

TaSdr-B1a ATCCTCGCGCGGTTCCGGCCCATCGCGCCCAAGCCCGCCCTGCCGGCGTCGCCGGTGCAG 360
TaSdr-B1b ATCCTCGCGCGGTTCCGGCCCATCGCGCCCAAGCCCGCCCTGCCGGCGTCGCCGGTGCAG 360

TaSdr-B1a GCGATCGACGGCGCCGCCGACCGCGTGCTCTGCCACCTGCAGAACAGGCCGTGCCGGGCA 420
TaSdr-B1b GCGATCGACGGCGCCGCCGACCGCGTGCTCTGCCACCTGCAGAACAGGCCGTGCCGGGCA 420

TaSdr-B1a AGGAAGCGCGGGCGCCCGAGCGCCGTGCCGGTGTCCGCGCCGGCCGCTGCCGCCAAGAGG 480
TaSdr-B1b AGGAAGCGCGGGCGCCCGAGCGCCGTGCCGGTGTCCGCGCCGGCCGCTGCCGCCAAGAGG 480

TaSdr-B1a AAGAGGGCGGCGTACCCGGTGCCGCTCCGGTGCGCGGCGGCGGCGGCCACCGACGCGGTG 540
TaSdr-B1b AAGAGGGCGGCGTACCCGGTGCCGCTCCGGTGCGCGGCGGCGGCGGCCACCGACGCGGTG 540

TaSdr-B1a GTGTCCACCGCGACGAGGGCCTATGTGTCCGTGCCGGGCAGTGCATGCATGCCGTTTGCG 600
TaSdr-B1b GTGTCCACCGCGACGAGGGCCTATGTGTCCGTGCCGGGCAGTGCATGCATGCCGTTTGCG 600

TaSdr-B1a TCCCTCCCGCCGGCGACCGCGAGTACCGGCGGGAATCTGACGATGCTCTCGACGACCATG 660
TaSdr-B1b TCCCTCCCGCCGGCGACCGCGAGTACCGGCGGGAATCTGACGATGCTCTCGACGACCATG 660

TaSdr-B1a GTGGCGGGCGACGACGAAGAGGAGGAGAGGGACGTCCCCGTGGAGCGCGACCTGCTGCGG 720
TaSdr-B1b GTGGCGGGCGACGACGAAGAGGAGGAGAGGGACGTCCCCGTGGAGCGCGACCTGCTGCGG 720

TaSdr-B1a AAGCTGCTGGAGCCCAAGGTGATCTCGCCGCGGGCGATGCGCCCCGTAGGATCCACCATC 780
TaSdr-B1b AAGCTGCTGGAGCCCAAGGTGATCTCGCCGCGGGCGATGCGCCCCGTAGGATCCACCATC 780

TaSdr-B1a CACGTCGAATCCATCGTCCCCGGCGCCGTCGACGCGACCAGCACCGCCGCCTCGAAGACG 840
TaSdr-B1b CACGTCGAATCCATCGTCCCCGGCGCCGTCGACGCGACCAGCACCGCCGCCTCGAAGACG 840

Sdr-5 Reverse primer

TaSdr-B1a GCGGAGGAGGTGGAGGCGGAGGTGGAGACCGACGCGCT**CCCGGCGGTCGTCACGGA**CTCG 900
TaSdr-B1b GCGGAGGAGGTGGAGGCGGAGGTGGAGACCGACGCGCTCCCGGCGGTCGTCACGGACTCG 900

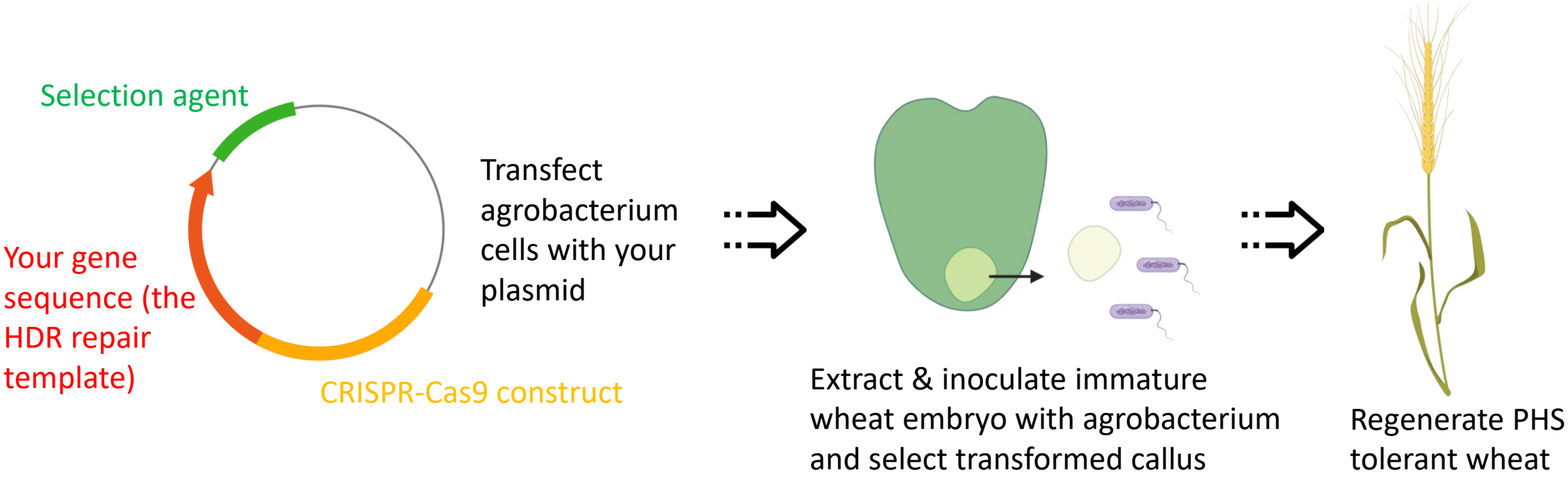
TaSdr-B1a AGCAACCGCGTCCGGCTGGTGAACGACGCGTACAAGGAGATGGTGGGCGCGCCCGAGTGC 960
TaSdr-B1b AGCAACCGCGTCCGGCTGGTGAACGACGCGTACAAGGAGATGGTGGGCGCGCCCGAGTGC 960

TaSdr-B1a CTGTGGCTCGGCGCGGTGGCCGCGTCGAGGAGGATCAGCGGGGAGGTGGCGCTGGTGGTG 1020
TaSdr-B1b CTGTGGCTCGGCGCGGTGGCCGCGTCGAGGAGGATCAGCGGGGAGGTGGCGCTGGTGGTG 1020

Zhang et al., 2014 (TAG)

Step 9: is my gene sufficient to determine the phenotype?

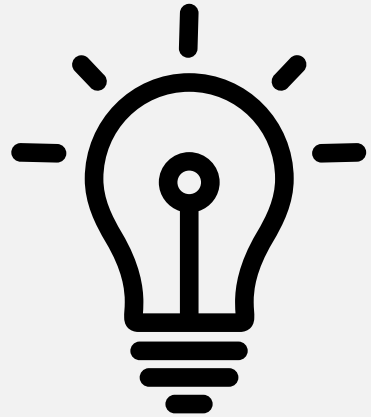
- Create a HDR construct with your tolerant gene sequence and use agrobacterium to knock-in the trait into a PHS susceptible variety



Background and Basics

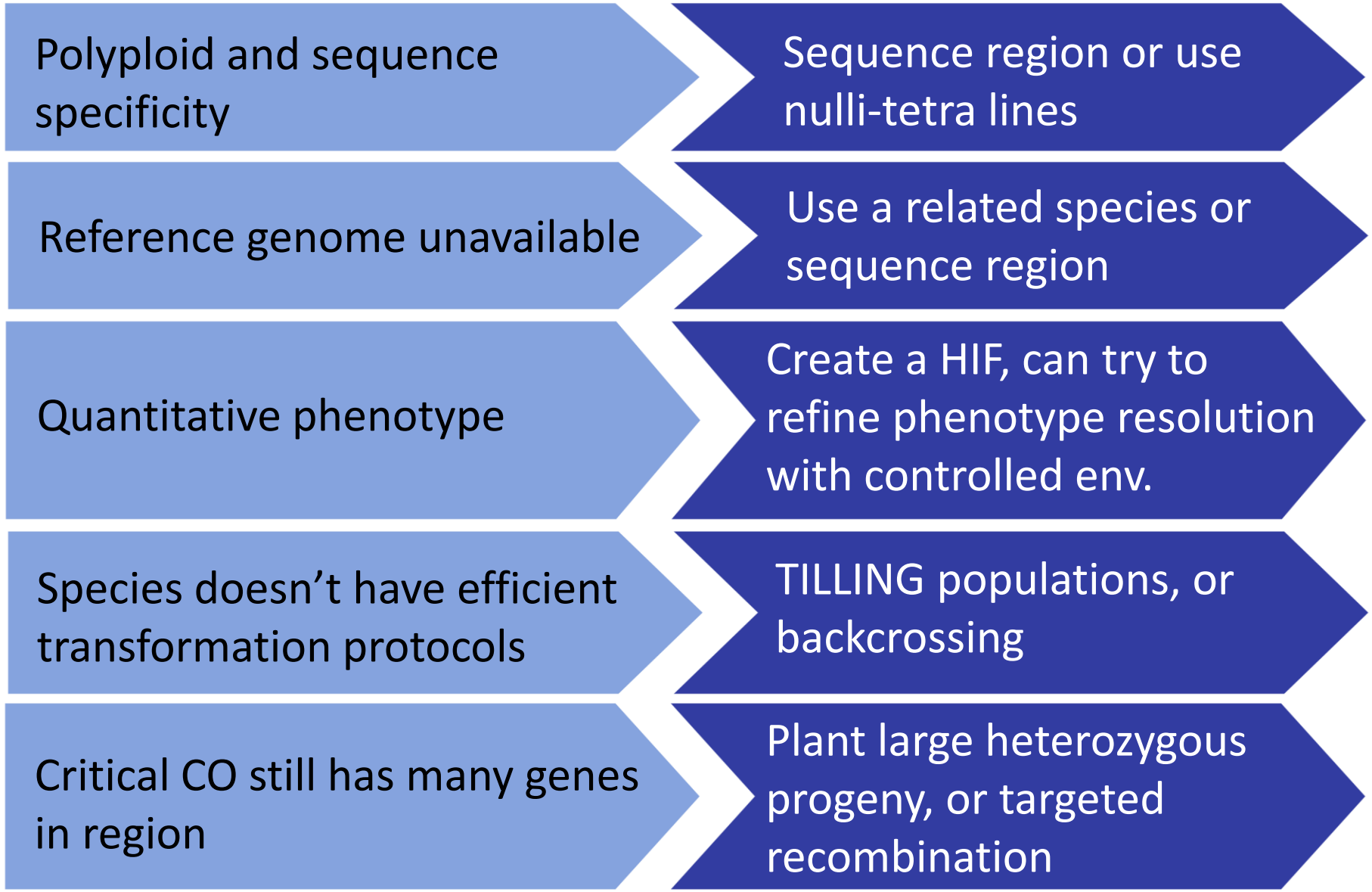
Practical Example

Conceptual Activity

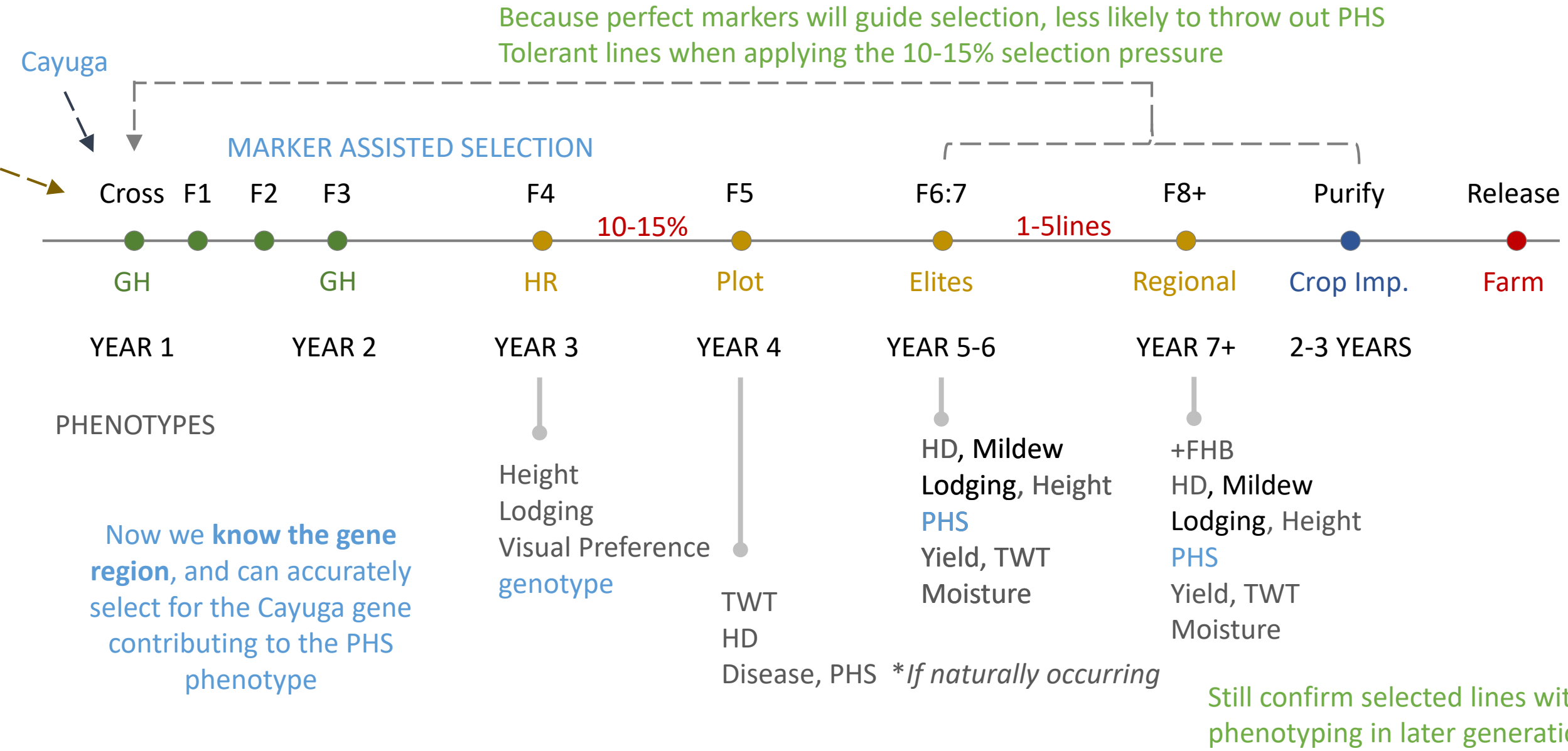


Challenges and Solutions

Challenges / Solutions



Breeding after fine mapping & cloning

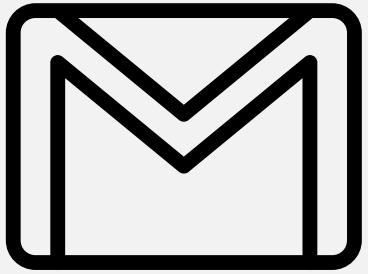


LinkedIn

ella-taagen

shantel-a-martinez

Email



et395

sam594

Slides/Resources

shantel-martinez.github.io/talks

REFERENCES

- Goldberg RB (2001) From Cot Curves to Genomics. How Gene Cloning Established New Concepts in Plant Biology. Plant Physiology 125:4–8. https://doi.org/10.1104/pp.125.1.4
- Kuzay S, Xu Y, Zhang J, et al (2019) Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. Theor Appl Genet 132:2689–2705. https://doi.org/10.1007/s00122-019-03382-5
- McMaster GJ, Derera NF (1976) Methodology and sample preparation when screening for sprouting damage in cereals. Cereal Research Communications 4:251–254
- Mukherjee S (2016) The gene: an intimate history, First Scribner hardcover edition. Scribner, New York
- Ramírez-González RH, Borrill P, Lang D, et al (2018) The transcriptional landscape of polyploid wheat. Science 361:eaar6089. https://doi.org/10.1126/science.aar6089
- Zhang Y, Miao X, Xia X, He Z (2014) Cloning of seed dormancy genes (TaSdr) associated with tolerance to pre-harvest sprouting in common wheat and development of a functional marker. Theor Appl Genet 127:855–866. https://doi.org/10.1007/s00122-014-2262-6

MAPPING/CLONING RESOURCES

Wheat genome sequence browser: URGI Blast https://urgi.versailles.inra.fr/blast_iwgsc/blast.php
URGI JBrowse
https://urgi.versailles.inra.fr/jbrowseiwgsc/gmod_jbrowse/?data=myData%2FIWGSC_RefSeq_v1.0&loc=chr1A%3A59558565..534840209&tracks=HighConfidenceGenesv1.0&highlight=

Primer Design: Oligo-Calc http://biotools.nubic.northwestern.edu/
NIH Primer Blast https://www.ncbi.nlm.nih.gov/tools/primer-blast/ PolyMarker http://polymarker.tgac.ac.uk

SNP functional changes: EcoTILLING browser https://dubcovskylab.ucdavis.edu/eco-tilling-blast

Expression databases: expVIP http://www.wheat-expression.com WheatExp https://wheat.pw.usda.gov/WheatExp/
Wheat eFP http://bar.utoronto.ca/efp_wheat/cgi-bin/efpWeb.cgi
Ensembl Plants https://plants.ensembl.org/index.html

Gene literature review: Google Scholar https://scholar.google.com PubMed https://www.ncbi.nlm.nih.gov/pubmed/

Design guide RNA: Cermak et al., 2017 The Plant Cell http://www.plantcell.org/content/29/6/1196
Voytas Lab Plant Genome Engineering http://cfans-pmorrell.oit.umn.edu/CRISPR_Multiplex/
Wheat CRISPR https://crispr.bioinfo.nrc.ca/WheatCrispr/

Sequence Alignment: ClustalOmega https://www.ebi.ac.uk/Tools/msa/clustalo/

HANDOUT RESOURCES

Google Sheet with handouts: bit.ly/2JdxFIN

PRINTING INSTRUCTIONS:
PartA: Print Handout 1 for n groups (there is enough room for multiple per page)
PartB: Print Handout 2+3 for n groups; Cut apart Handout 2 from 3
Cut Handout 2 in strips, one strip for each Line ID (only one will keep the header, that is okay)
PartC: Print Handout 4 for n groups
Cut handout in half
PartD: Print Handout 5 for n groups

NOTE: be sure to print "fit to page". Some worksheets may be too long or wide by default.

NOUN PROJECT ATTRIBUTES

All graphics were created by the presenters in PowerPoint or BioRender, except for the following:


Emily Akins


counloucon

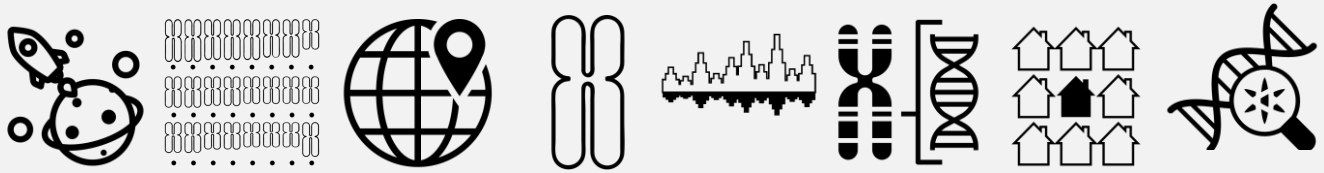

Parkjisun


Chameleon

Wilson Joseph

Kokota

Ellie’s slide icons artists: Maxim Kulikov, Rflor, Bismillah, Juan Pablo Bravo, Gregor Cresnar, Saifurrijal, Cornelia Scheitz



123rf.com