

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Literature review</b>	<b>4</b>
2.1	Wheat Breeding . . . . .	4
2.2	Wheat Genetics . . . . .	4
2.3	Wheat Genomics . . . . .	4
2.4	Sequencing . . . . .	4
2.5	Sequence analysis . . . . .	5
2.6	Wheat online resources . . . . .	5
<b>3</b>	<b>Genetic mapping of <i>Yr15</i></b>	<b>6</b>
3.1	(Introduction) <i>Yr15</i> . . . . .	6
3.2	Segregating population and resistance essays . . . . .	6
3.3	Sequencing and mapping . . . . .	6
3.4	SNP Calling . . . . .	6
3.5	Bulk Frequency Ratios . . . . .	6
3.6	<i>In silico</i> mapping . . . . .	7
3.7	Assay selection . . . . .	7
3.8	Genetic map . . . . .	7
3.9	Assembly of the transcriptome . . . . .	7
3.10	Conclusions . . . . .	7
<b>4</b>	<b>PolyMarker: A fast polyploid primer design pipeline</b>	<b>8</b>
4.1	Pipeline . . . . .	8
4.2	PolyMarker public web service . . . . .	11
4.3	Applications of PolyMarker . . . . .	11
4.3.1	KASP assays for public sets of SNPs . . . . .	11
4.3.2	Genotyping of <i>Puccinia striiformis</i> f. sp. <i>tritici</i> isolates. . . . .	11
4.3.3	Validation of SNPs in a mutant population . . . . .	13
4.3.4	Validation of deletions on a mutant population . . . . .	13
4.4	Conclusions . . . . .	13
<b>5</b>	<b>Gene expression (expVIP)</b>	<b>15</b>
5.1	Expression experiments (Introduction) . . . . .	15
5.2	Database design . . . . .	15
5.3	Analysis pipeline . . . . .	15
5.4	Graphical interface . . . . .	15

<i>CONTENTS</i>	2
5.5 Conclusions . . . . .	15
<b>6 Conclusions and final remarks</b>	<b>16</b>
<b>A Supplemental tables.</b>	<b>17</b>

# Chapter 1

## Introduction

It defines the objectives and the importance of the research. It focus on the the application of Next Generation Sequencing to molecular biology, wheat genetics and ultimately to breeding programs. It also mentions the current status of the wheat reference genome and other resources (genetic maps, markers) the need of tools to query them effectively.

# Chapter 2

## Literature review

It describes the current status of the wheat genome, genetics and other resources.

### 2.1 Wheat Breeding

An overview of how breeding is carried on currently, the different sources of genetic diversity and the relevance of fixing agriculturally important traits.

### 2.2 Wheat Genetics

The section describes alleles and the concept of gene, both as a locus in the genome (Quantitative Trait Locus, QTL) and an specific transcript (central dogma of molecular biology). Finally, it discusses traditional Mendelian inheritance and the effect of polyploidy.

### 2.3 Wheat Genomics

A description of the current status of the wheat genome (Mayer et al. (2014), Chapman et al. (2015)), the different available assemblies and approaches to sort the scaffolds (Genome Zipper, the various genetic maps).

### 2.4 Sequencing

The importance of the selection of the library preparation and the sequencing platforms available. A brief summary of RNA-Seq, Exome capture, Whole Genome Shotgun, etc. and on which cases are more suitable for different experiments. Mention the new technologies developed during the years of the PhD (Ren-Seq, PacBio?)

## 2.5 Sequence analysis

This section discusses the criteria to decide analysis done after sequencing, when to do re-alignments or *de novo* assemblies, how to do SNP calling in diploid and polyploid organisms and the bulk frequency ratios.

## 2.6 Wheat online resources

A compilation of the currently available resource for wheat genetics and genomics. MAS wheat, CerealsDB, Ensembl, etc.

# Chapter 3

## Genetic mapping of *Yr15*

This section describes in detail than the paper of Ramirez-Gonzalez et al. (2014)

### 3.1 (Introduction) *Yr15*

Breeding importance of *Yr15* and original source (an introgression of *T. diccoides*).

### 3.2 Segregating population and resistance essays

A description of the starting material and how the population was generated.

### 3.3 Sequencing and mapping

RNA-Seq and the decision to call SNPs on gene models rather than the whole reference. Details of the mapping against the Wheat UniGenes Pontius et al. (2002) and the UCW. Krasileva et al. (2013) gene models.

### 3.4 SNP Calling

. Ruby implementation of the methodology described by Trick et al. (2012).

### 3.5 Bulk Frequency Ratios

Results of the simple SNP calls from the progenitors and how the score of the Bulk Frequency Ratios(BFR) improve the location of the SNPs.

### 3.6 *In silico* mapping

Mapping of the gene models to the IWGSC CSS Mayer et al. (2014) reference and the location of the SNPs using the genetic map from Wang et al. (2014).

### 3.7 Assay selection

. The selection criteria to decide which SNPs where selected to produce the genetic map: BFR>6, in the short arm of chromosome group 1 and from the *Yr15* progenitor.

### 3.8 Genetic map

The three versions of the genetic map: With a subset of the F<sub>2</sub> population

### 3.9 Assembly of the transcriptome

A comparison between thef known unigenes and the transcript from the progenitors. Since *Yr15* comes from an introgression with *T. diccocooides*, some novel transcripts can be extracted. Analysis of the gels from Mitaly?

### 3.10 Conclusions

Remarks on how this techinque can be used to do fine-mapping and that if I were to start the project now I would use exome capture or Ren-Seq.

# Chapter 4

## PolyMarker: A fast polyploid primer design pipeline

One of the main challenges of working with polyploid species is the design of genome specific molecular markers. This is particularly true when targeting conserved homoeologue regions, where a primer could bind to a pair, or triplet, of identical sequences. For that reason, designing primers for polyploids require to include bases that are specific to the target, in addition to the physicochemical properties of the primer. The traditional methodology to find primer candidates include a blast search and a local alignment, select the primer candidates manually, and finally, validate the primers with a tool, like **Primer3** (Rozen and Skaletsky, 2000). To reduce the time invested in designed primers I have developed PolyMarker (Ramirez-Gonzalez et al., 2015), a pipeline to automate the primer design for polyploid organisms.

### 4.1 Pipeline

PolyMarker is an automated pipeline that takes as input a list of SNPs and a reference file and produces a list of primer triplets for SNP genotyping. The list of SNPs is first converted to a FASTA file with ambiguity codes (Cornish-Bowden, 1985). The template sequences are aligned with **exonerate** (Slater and Birney, 2005) to find the homoeologous regions to the target sequence. Then, the alignment between homoeologues is refined using **MAFFT** (Katoh and Standley, 2013). A list of candidate variations is produced and used as input for **Primer3** (Rozen and Skaletsky, 2000). Finally, the output of **Primer3** is parsed to find the best primer pair that contains a the targeted SNP and a base that is specific to the target genome (Figure 4.1). The pipeline is written as a Ruby script, using parsers and wrappers from BioRuby (Goto et al., 2010) and bio-samtools (Etherington et al., 2015; Ramirez-Gonzalez et al., 2012). The software is open source and released as a biogem (Bonnal et al., 2012), **bio-polyploid-tools**, the source code is available in github: <https://github.com/TGAC/bioruby-polyploid-tools>.

The PolyMarker input consist on SNP list with: unique name for the marker, the target chromosome and the sequence for the marker. The alterna-



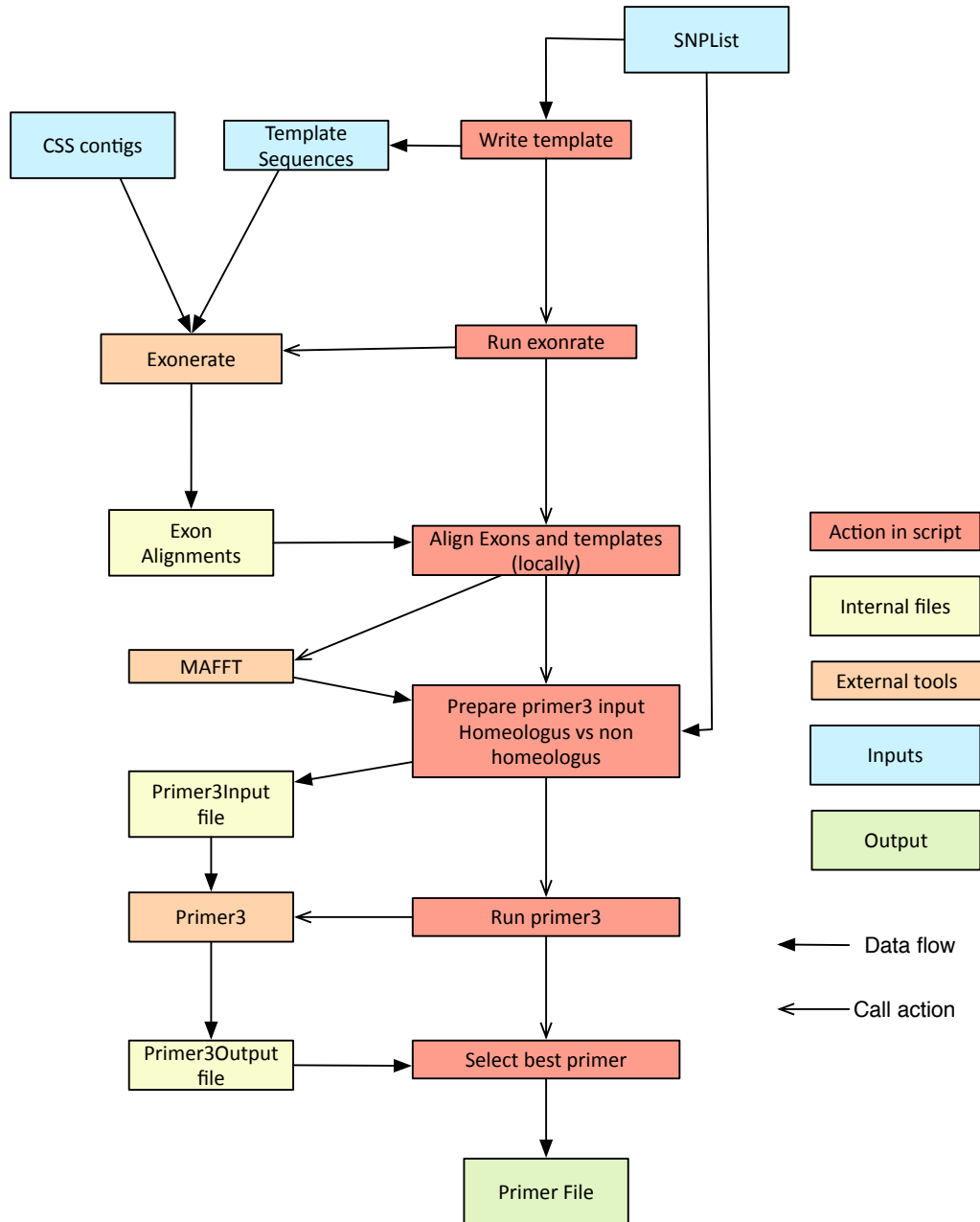


Figure 4.1: Steps and tools called by PolyMarker. The colour of the boxes represent: the step is an action inside the script (red); actions of the script (orange); temporary files (yellow); inputs (blue) and; output (green)

tive alleles are surrounded by square brackets within the sequence. PolyMarker can take a list of several markers and design them in batch (Figure 4.2a). A FASTA file is produced with all the template sequences, with the alternative alleles substituted by the IUAPC ambiguity codes (Cornish-Bowden, 1985). The flanking sequence surrounding the SNP is limited by default to 100bp to reduce the search time and avoid missing regions that diverge near the SNP, as when the variation is near an intron-exon junction.

The template sequences are aligned to the reference using **exonerate** (Slater and Birney, 2005), figure 4.2b. The alignment is refined with the **--model est2genome** option, to allow the search of sequences coming from transcripts, a common source of SNPs (Allen et al., 2011). The exonerate output is formatted with the **--ryo** (roll your own format) to get an output easy to parse. All the hits that contain the SNP are extracted from the reference with a flanking sequence that extend out of the hit, by default, to 100bp on each side of the SNP (Figure 4.2c). The size of the flanking sequence can be set to different sizes to allow the design of different types of primers. Different homoeologues may contain small indels (Figure 4.2d). To enable a comparison base-per-base, a local alignment with **MAFFT** (Katoh and Standley, 2013) is produced (Figure 4.2e).

PolyMarker searches across each base in the local alignment to identify the variations across homoeologues and the target marker. A mask is produced to highlight the bases with a variations (Figure 4.2f) on the following categories:

<b>Specific</b>	Homoeologous polymorphism which is only present in the target genome (upper case).
<b>Semi-specific</b>	Homoeologous polymorphism which is found in 2 of the 3 genomes, hence it discriminates against one of the off-target genomes or when not all the homoeologous sequences were found (lower case).
<b>Non-specific</b>	No variation is found across homoeologues (-).
<b>Homoeologous</b>	The target SNP is present across different chromosomes, so candidate SNP markers on this category are not expected to be reliably identify the allele (:).
<b>Non-homoeologous</b>	The target SNP is not present across chromosomes, so it can be used to identify an allele (&).

PolyMarker was designed to produce SNP assays for KASP genotyping (LGC Genomics, 2013), which requires a common primer and two allele-specific primers. The common primer is selected to start on a position from a: Specific; Semi-specific or; Non-specific, on that priority. This means that the common primer will be as specific as possible in the region. For the allele-specific primers, the starting position of the primer is on the base with the SNP. To ensure that the stability of the candidate primers will be met, the putative starting positions are tested with **Primer3** (Rozen and Skaletsky, 2000).

PolyMarker was designed and validated with the markers described in section 3.8. For wheat, PolyMarker uses the contigs from Mayer et al. (2014), as deposited in Ensembl. As new releases of the wheat genome are made available, different parsers to assign the chromosome to each sequence can be added with little effort to PolyMarker.

## 4.2 PolyMarker public web service

To make PolyMarker accessible to the community, a web server that allow the submission of SNPs was developed. The web interface consists on two virtual machines, one with a web facing interface that stores the queries, and a dedicated node to submit jobs to an HPC cluster. The on-line interface further simplifies the design of KASP assays, a process that used to take a couple of weeks now is done in a couple of hours. Since the release of the public service in July 2014 until August 2016, 1,739 requests to PolyMarker have been done.

## 4.3 Applications of PolyMarker

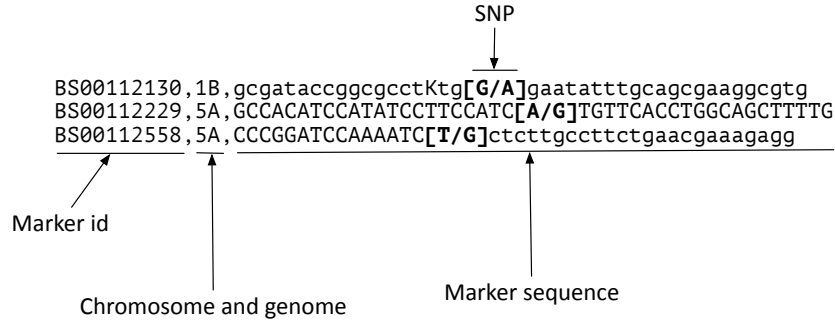
PolyMarker is not restricted to wheat or to KASP assays, the source code is flexible and can be extended for other types of analysis. On each of the following projects, PolyMarker has been adapted to design primers in species where KASP hasn't been used before, the primers are used for regular PCR amplification, or the use of KASP is not the conventional SNP calling.

### 4.3.1 KASP assays for public sets of SNPs

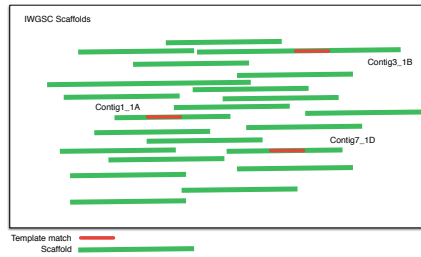
PolyMarker was used to design KASP assays for the 81,587 markers from (Wang et al., 2014), available on the PolyMarker website and in CerealsDB (Wilkinson et al., 2012). Of those markers, 40,267 were designed using the target chromosome using the genetic map published by the genetic map. Genes without a genetic position were aligned to scaffolds sorted by chromosome from the International Wheat Genome Sequencing Consortium (Mayer et al., 2014) with BLAT (Kent, 2002) and the best hit was selected as putative location. 97.5% of the assays were designed and 76% of them are semi-specific or specific, thereby improving their expected performance with respect to randomly designed primers (Table 4.1). A set of the designed assay was used to genotype a mapping population to find resistance to Fusarium head blight (Burt et al., 2015).

### 4.3.2 Genotyping of *Puccinia striiformis* f. sp. *tritici* isolates.

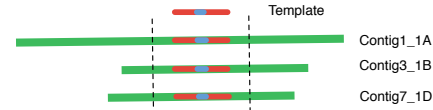
In Hubbard et al. (2015), *Puccinia striiformis* f. sp. *tritici* (PST) isolates were sequenced and assigned to clusters, according to their genotype. The



(a) PolyMarker input. The alternative alleles are surrounded by brackets.



(b) Global search of templates in the reference contigs.



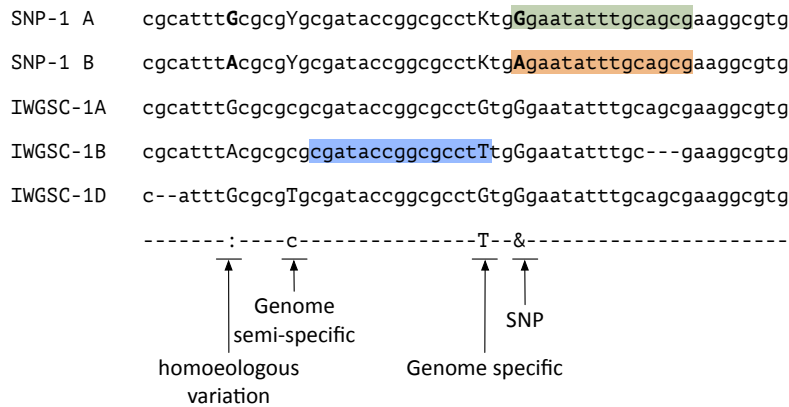
(c) Selected regions around the SNP on every chromosome.

SNP-1 A cgcatttGcgcgYgcgataccggcgccctKtgGgaatatttgcagcgaaggcgtg  
 SNP-1 B cgcatttAcgcgYgcgataccggcgccctKtgAgaatatttgcagcgaaggcgtg  
 IWGSC-1A cgcatttGcgcgcgcgataccggcgccctGtgGgaatatttgcagcgaaggcgtg  
 IWGSC-1B cgcatttAcgcgcgcgataccggcgccctTtgGgaatatttgcagcgaaggcgtg  
 IWGSC-1D c--atttGcgcgTgcgataccggcgccctGtgGgaatatttgcagcgaaggcgtg

SNP-1 A cgcatttGcgcgYgcgataccggcgccctKtgGgaatatttgcagcgaaggcgtg  
 SNP-1 B cgcatttAcgcgYgcgataccggcgccctKtgAgaatatttgcagcgaaggcgtg  
 IWGSC-1A cgcatttGcgcgcgcgataccggcgccctGtgGgaatatttgcagcgaaggcgtg  
 IWGSC-1B cgcatttAcgcgcgcgataccggcgccctTtgGgaatatttgcagcgaaggcgtg  
 IWGSC-1D c--atttGcgcgTgcgataccggcgccctGtgGgaatatttgcagcgaaggcgtg

(d) Sequence of found regions around the SNP.

(e) Local alignment on regions around the SNP detects indels.



(f) Alignment with mask and primer candidates.

Figure 4.2: Alignments done by PolyMarker.

Table 4.1: Count of KASP assays designed for the 40,267 SNP markers located in the genetic map from Wang et al. (2014). 4,228 assays did not align to the target chromosome. Not designed: Primer3 could not find viable primers flanking the SNP.

	Homoeologous variant	Varietal SNP	Percentage
Non-specific	1,765	5,857	21.15%
Semi-specific	7,942	6,907	41.20%
Specific	6,813	5,957	35.43%
Not designed	242	556	2.21%
Total	16,762	19,277	36,039

clusters are useful to monitor the changes in the pathogen population, which can be used to predict if certain wheat lines will be resistant to the isolates in the field. PolyMarker was used to design primers for PST, using the assembly PST-130 Cantu et al. (2011). Out of 15 assays 11 can be used to identify to which cluster of isolates a sample is likely to belong (Supplemental Table A.1).

### 4.3.3 Validation of SNPs in a mutant population

PolyMarker was used to design primers to validate SNPs in a Targeted Induced Local Lesions in Genomes (TILLING) population, an approach to identify the function of genes by mutating them. To calibrate the SNP calling, KASP assays were designed to get the mutations from  $M_2$ ,  $M_3$  and,  $M_5$  mutants King et al. (2015). Then primers were designed for the whole mutant population, consisting of 1,200 Cadenza (Hexaploid) and 1,535 Kronos (Tetraploid) wheat lines Krasileva et al. (submitted 2016). Genome-specific primers 172 and 80 SNP assays on 19 and 8  $M_4$  Cadenza and Kronos lines respectively. Of those, 71(85.5%) Kronos and 147(88.8%) of the Cadenza primers where valid assays (Tables A.2 and A.3).

### 4.3.4 Validation of deletions on a mutant population

Primers

## 4.4 Conclusions

Remarks on the importance of getting the primers right, and the time saved by automating the primer selection. Also mention other primer design tools that have been inspired by polymarker: Ma et al. (2015), Wang et al. (2016)

PolyMarker has been used succesfully to design genome-specific primers in several projects.

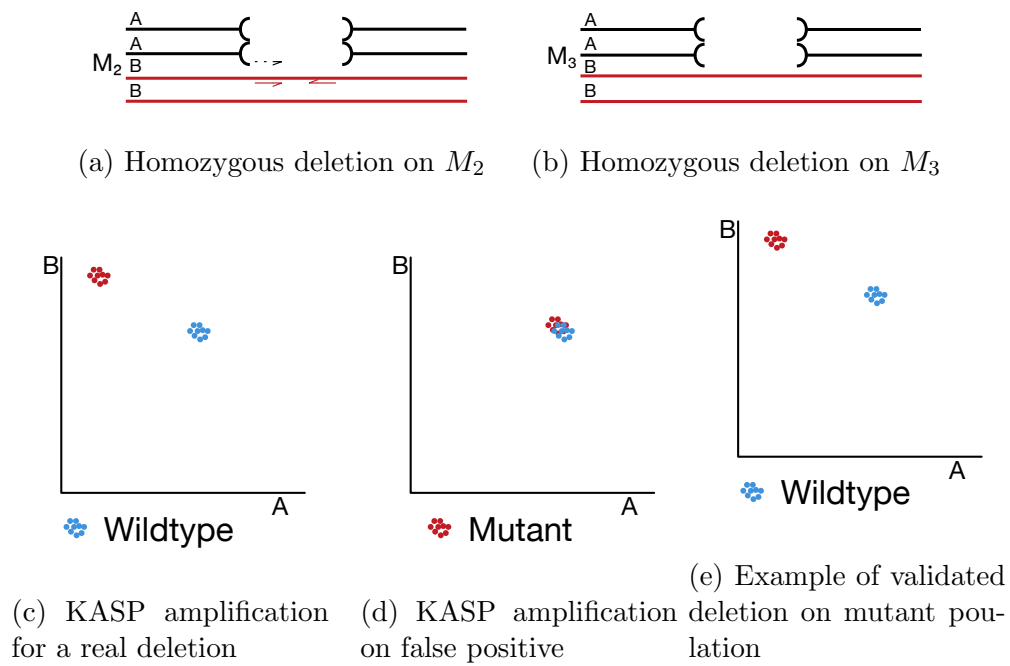


Figure 4.3: PolyMarker used to find primers to detect long deletions in tetraploid wheat.

# Chapter 5

## Gene expression (expVIP)

### 5.1 Expression experiments (Introduction)

Describe the list of previously published expression experiments and how they can potentially be used as a framework for new experiments.

### 5.2 Database design

Description of how the database was designed and the flexibility given by having the factors and units as variables

### 5.3 Analysis pipeline

Implementation of the pipeline, from running kallisto to load the data in the database

### 5.4 Graphical interface

How the expression can be displayed filtered, and sorted

### 5.5 Conclusions

The use of previously published studies is a valuable resource. Also, mention that despite the fact that there are several expression/gene browsers, none of them allow comparisons between species and don't consider polyploids.

# Chapter 6

## Conclusions and final remarks

This section wraps up by showing the relationship and importance of a comprehensive approach to data analysis, from the field, genetics, molecular biology and genomics. I will also remark how the technology and the resources have changed in the last 4 years. As at the references used at beginning where superseded during the PhD.



# Appendix A

## Supplemental tables.

Table A.1: PolyMarker used to genotype PST

Assay	Contig	Position	X	Y	Cluster I isolates		Cluster II isolates		Cluster III isolates			Cluster IV isolates	
					13/26	13/123	CL1	T-13/3	13/09	13/23	13/182	13/36	13/40
1	PST130_14470	268	C	T	X:Y	X:Y	X:X	X:X	X:X	X:X	X:X	X:X	X:X
2	PST130_8160	11876	C	T	Y:Y	Y:Y	X:Y	X:Y	X:Y	X:Y	X:Y	X:Y	X:Y
3	PST130_14628	1712	A	C	X:Y	-	X:X	X:X	X:X	X:X	X:X	X:X	X:X
4	PST130_14898	503	G	A	X:X	X:X	X:Y	X:Y	X:Y	X:Y	-	X:Y	X:Y
5	PST130_28344	2372	A	G	Y:Y	Y:Y	X:Y	X:Y	Y:Y	Y:Y	Y:Y	Y:Y	Y:Y
6	PST130_7634	3463	A	C	Y:Y	Y:Y	X:Y	X:Y	Y:Y	Y:Y	Y:Y	Y:Y	Y:Y
7	PST130_7629	11699	G	A	Y:Y	Y:Y	X:Y	X:Y	Y:Y	Y:Y	Y:Y	Y:Y	Y:Y
8	PST130_10943	2979	C	T	X:Y	X:Y	X:Y	X:Y	X:X	X:X	X:X	X:Y	X:Y
9	PST130_10126	6216	G	T	Y:Y	Y:Y	X:X	X:X	X:X	X:X	-	Y:Y	Y:Y
10	PST130_22010	172	C	T	Y:Y	Y:Y	Y:Y	Y:Y	X:Y	X:Y	-	X:Y	X:Y
11	PST130_16961	1098	C	T	X:X	X:X	X:Y	X:Y	Y:Y	Y:Y	Y:Y	X:Y	X:Y
12	PST130_6915	2710	A	T	Y:Y	Y:Y	Y:Y	Y:Y	Y:Y	X:Y	X:Y	Y:Y	Y:Y
13	PST130_12479	1428	C	T	X:X	X:X	Y:Y	Y:Y	X:X	X:X	X:X	Y:Y	X:X
14	PST130_7634	3883	C	G	X:X	X:X	X:Y	X:Y	X:X	X:X	X:Y	X:Y	X:X
15	PST130_14470	456	T	C	Y:Y	Y:Y	X:Y	X:Y	Y:Y	Y:Y	X:Y	Y:Y	Y:Y

Table A.2: Validation of mutations on  $M_4$  on Cadenza

IWGC contig	Line	Pos	WT	Mut	Predicted	$M_4$	Primer 1 (Cadenza)	Primer 2 (mutant)	Common Primer
IWGC_CSS.3B.scaff.10445294	Cadenza1772	6019	C	T	het	het	caggatAgtGggactgtcaaaG	caggatAgtGggactgtcaaaA	ggagacGGctGtggacatT
IWGC_CSS.3DL.scaff.6955403	Cadenza1772	2418	C	T	het*	hom	tcagCggattgtcgggatG	tcagCggattgtcgggatA	tgctCatgaaTcttgtccacG
IWGC_CSS.4AL.scaff.7106846	Cadenza1772	11277	G	A	hom	hom	tgggatccatgcctacactG	tgggatccatgcctacactA	gatggTGatttgcgctA
IWGC_CSS.4AS.scaff.5991335	Cadenza1772	15710	G	A	hom	hom	ctggccctgcgctgctaC	ctggccctgcgctgctaT	gtggaaGttcagaaggaccaG
IWGC_CSS.4BS.scaff.4956646	Cadenza1772	252	G	A	het*	hom	gcaggttgacttcccggaG	gcaggttgacttcccggaA	tGaggtacgaGcTaaagAagC
IWGC_CSS.4DS.scaff.1715962	Cadenza1772	1225	G	A	hom	hom	cagctgtggTatctcaactgG	cagctgtggTatctcaactgA	CcCtGaaACACcGtttgaT
IWGC_CSS.5AL.scaff.2763407	Cadenza1772	2119	G	A	hom	hom	ggacGaacctcgagatctG	ggacGaacctcgagatctA	gaTggcaAtcgtCgtgcA
IWGC_CSS.5AS.scaff.1548786	Cadenza1772	12625	C	T	het	het	AtaggcacattgctagactgaG	AtaggcacattgctagactgaA	ggattgggtgtgcacG
IWGC_CSS.5BL.scaff.10849226	Cadenza1772	2289	C	T	het*	hom	cctgacatcattgttcacgatC	cctgacatcattgttcacgatT	cactccgaggtgtccatgaT
IWGC_CSS.5BS.scaff.2270737	Cadenza1772	2262	G	A	hom	—	attcCTgtgttggtggCaaatgaG	attcCTgtgttggtggCaaatgaA	taaGcaciaAccctccagctgG
IWGC_CSS.1AL.scaff.3022915	Cadenza1661	891	C	T	hom	hom	ccacagtgcagactcctattgaCG	ccacagtgcagactcctattgaCA	atgtctgattcGtcGtagtcC
IWGC_CSS.1AS.scaff.3297240	Cadenza1661	1970	C	T	het	het	catcccgccGtttctcT	catcccgccGtttctcT	gctcccgatgaagagcT
IWGC_CSS.1BL.scaff.3828996	Cadenza1661	1340	G	A	hom	hom	agccggatgttagtgttaacC	agccggatgttagtgttaacT	agcagcttgTcgcgtaaC
IWGC_CSS.1DS.scaff.1884529	Cadenza1661	10575	G	A	hom	hom	aCagatacaAttgtcatgcaggC	aCagatacaAttgtcatgcaggT	acctgggTTgtccaatactC
IWGC_CSS.2AL.scaff.6318370	Cadenza1661	19142	C	T	het	—	cgtggcCgaatCtcGacG	cgtggcCgaatCtcGacA	ttcttggggagccggcC
IWGC_CSS.2AS.scaff.5213460	Cadenza1661	1358	G	A	hom	hom	gtcacgaaCccgctcagG	gtcacgaaCccgctcagA	aggaagagagagaaaagaGcG
IWGC_CSS.2BS.scaff.5179331	Cadenza1661	5604	G	A	het	het	actctcgtaagaactgatacaG	actctcgtaagaactgatacaA	gcaGagaatgttcttgcaacT
IWGC_CSS.2DS.scaff.5341235	Cadenza1661	4673	G	A	het	het	ggtaggagatctcggagctG	ggtaggagatctcggagctA	gcgcggtctgacaggttG
IWGC_CSS.3AL.scaff.4250995	Cadenza1661	7046	G	A	hom	hom	cCaagaaacgggtggtccaG	cCaagaaacgggtggtccaA	ctgcagctgtccatcatcgT
IWGC_CSS.3B.scaff.10404421	Cadenza1661	4303	G	A	het	het	ccttcgtcgaCaggacctG	ccttcgtcgaCaggacctA	GCcagtaactCacAtgctcC
IWGC_CSS.5DL.scaff.2390496	Cadenza1538	2125	C	T	hom	het	gcagttttatcctcagtagtcttgG	gcagttttatcctcagtagtcttgA	ttctgagaaTgtaattgctGatG
IWGC_CSS.6AL.scaff.5753680	Cadenza1538	3920	C	T	hom	hom	tgtccaaatttgagcacaaTaaC	tgtccaaatttgagcacaaTaaT	aaatgcaagggttaagtttttG
IWGC_CSS.6AS.scaff.4425792	Cadenza1538	4307	G	A	hom	het	agatgcttgtCggGccaG	agatgcttgtCggGccaA	gctgaagcaacgcgatcaaT
IWGC_CSS.6BS.scaff.3003630	Cadenza1538	6933	C	T	het	het	ggcagtaattgtggtgctgagC	ggcagtaattgtggtgctgagT	tTgaCttctgggttggtggcA
IWGC_CSS.6DL.scaff.3246988	Cadenza1538	9186	G	A	het	het	gctaaagaagagcttgagagaattC	gctaaagaagagcttgagagaattT	aatttctgaagagaggtgtgtatG
IWGC_CSS.7AL.scaff.4480114	Cadenza1538	3446	C	T	het	—	gatatctcccacacggcgG	gatatctcccacacggcgA	tgagccactcttgagtttT
IWGC_CSS.7AS.scaff.4193541	Cadenza1538	8359	C	T	hom	het	agcaattctttggctatcaattagC	agcaattctttggctatcaattagT	tcactGTcttaactctactgctG
IWGC_CSS.7BL.scaff.6721572	Cadenza1538	9223	C	T	het	het	gctCaggaggagaaagacaagaaG	gctCaggaggagaaagacaagaaA	tgctatgaagaattccgacctC
IWGC_CSS.7BS.scaff.3152545	Cadenza1538	3960	G	A	hom	—	tcagcaaaatcacctgcGgC	tcagcaaaatcacctgcGgT	gCtgccccatcatcgtttaT
IWGC_CSS.7DS.scaff.3963838	Cadenza1538	2913	G	A	het	het	tCgttgcaagcCttTtgtgC	tCgttgcaagcCttTtgtgT	agaGttaTcaagcTactgtcAc
IWGC_CSS.1AL.scaff.3903380	Cadenza1469	6193	G	A	hom	hom	ctcttcAgagatgaacgcgG	ctcttcAgagatgaacgcgA	tcGtGagatGtggttGTtA
IWGC_CSS.1AS.scaff.3287728	Cadenza1469	3817	C	T	het*	hom	ccgaccaAttcactaacgcG	ccgaccaAttcactaacgcA	acctctttcccAgacatgaT
IWGC_CSS.1BL.scaff.3815304	Cadenza1469	513	G	A	hom	hom	aacatttgctTaCcaaaacGC	aacatttgctTaCcaaaacGT	acacagcaagttataatgCAAGC
IWGC_CSS.1DL.scaff.2266648	Cadenza1469	5926	C	T	het	het	caacatgagacacaacacctT	caacatgagacacaacacctT	gtcaacgcgtgaggttgtC
IWGC_CSS.1DS.scaff.1906671	Cadenza1469	3697	C	T	hom	hom	tggTGTagacacttggcgaG	tggTGTagacacttggcgaA	catggcgaccaccAcctG
IWGC_CSS.2AL.scaff.6337088	Cadenza1469	7334	G	A	het*	hom	acaatgccAagttgacaggttG	acaatgccAagttgacaggttA	gggagttgtgttCagaacaT
IWGC_CSS.2BL.scaff.7972799	Cadenza1469	8995	C	T	het	hom	gTgCtcctcGgcactcttC	gTgCtcctcGgcactcttT	gatcgcGcaaacactacgTG
IWGC_CSS.2DL.scaff.9832343	Cadenza1469	3262	G	A	het	het	TtgtctaAcagcacCGcagG	TtgtctaAcagcacCGcagA	agatctcggtcagcttttC
IWGC_CSS.2DS.scaff.5327939	Cadenza1469	3889	G	A	het	het	ttttTgccttatgtgactctagtaC	ttttTgccttatgtgactctagtaT	gaggccatcacagatagcG
IWGC_CSS.3B.scaff.10395219	Cadenza1469	1292	G	A	hom	—	aggtgcttgtgcttgcgG	aggtgcttgtgcttgcgA	cctctctggggccttataC
IWGC_CSS.3B.scaff.10592217	Cadenza0580	2994	C	T	het	—	acagcagtatcaagccctC	acagcagtatcaagccctT	tgatactgttgTggCggagG
IWGC_CSS.3DS.scaff.2596771	Cadenza0580	1037	G	A	het	het	tgtttatGACaggataatCagG	tgtttatGACaggataatCagA	tggcaaatgtgatgtcattagG
IWGC_CSS.4AL.scaff.7093953	Cadenza0580	9881	C	T	hom	hom	GacaggaagccggtaacaC	GacaggaagccggtaacaT	ctccAgcaggcatgggaT
IWGC_CSS.4BL.scaff.7037448	Cadenza0580	1837	C	T	hom	hom	CgttgaaaaGctgcaagaacttaaC	CgttgaaaaGctgcaagaacttaaT	cagtttcttcTtCaGagagataT
IWGC_CSS.4BS.scaff.4929479	Cadenza0580	10668	G	A	hom	—	tggattttcccgactgttC	tggattttcccgactgttT	gtaaacagggcatttcaagagtcA
IWGC_CSS.4DL.scaff.14359838	Cadenza0580	1408	G	A	hom	—	gCtcAttcagggtTGTcCtaTatG	gCtcAttcagggtTGTcCtaTatA	tgaCagaacagttgggtacacT
IWGC_CSS.4DS.scaff.2276484	Cadenza0580	8034	G	A	hom	hom	gccgtggttgatggAgaG	gccgtggttgatggAgaA	cgctcagattactgatactgcA
IWGC_CSS.5AL.scaff.2756579	Cadenza0580	5278	G	A	het	het	tgaatggatttttctgctccgttC	tgaatggatttttctgctccgttT	ggAATCCTATgCAgaAgaCTG
IWGC_CSS.5BL.scaff.10787208	Cadenza0580	10627	G	A	het	—	gcctctcacatgcggagaC	gcctctcacatgcggagaT	acgatgtcAggtggGegT
IWGC_CSS.5BS.scaff.2282179	Cadenza0580	5267	G	A	het	—	tgatgggtcacgactgC	tgatgggtcacgactgT	tcggcgcccttgaaAtcC

IWGSC contig	Line	Pos	WT	Mut	Predicted	$M_4$	Primer 1 (Cadenza)	Primer 2 (mutant)	Common Primer
IWGSC_CSS_5DL_scaff.4498073	Cadenza0423	4937	C	T	hom	hom	gcaccctctggttggtcatC	gcaccctctggttggtcatT	tgagcagcaAagcagccG
IWGSC_CSS_5DS_scaff.2738970	Cadenza0423	2319	C	T	het	—	cgtgaggtgggtgatttgC	cgtgaggtgggtgatttgT	tggaactagttacactgcagTC
IWGSC_CSS_6AL_scaff.5757109	Cadenza0423	2788	G	A	hom	hom	caggaGcctggcacaataaaGG	caggaGcctggcacaataaaGA	ctttcGcagctctcttagtttcG
IWGSC_CSS_6AS_scaff.4387871	Cadenza0423	2543	G	A	hom	hom	gcaltgctaacaggcgaaaaG	gcaltgctaacaggcgaaaaG	ctcatgctcctgatcttaaggtG
IWGSC_CSS_6BL_scaff.4271391	Cadenza0423	4660	C	T	hom	hom	tacgtgcatgatgtggtagtcgtaC	tacgtgcatgatgtggtagtcgtaT	gtttgaagtgcacatgcagatgTaccA
IWGSC_CSS_6DS_scaff.1880206	Cadenza0423	9159	G	A	het	het	ctgCgaaggctccacaaG	ctgCgaaggctccacaaA	ggatgagaagtttgcatgtctC
IWGSC_CSS_7AS_scaff.4227506	Cadenza0423	952	G	A	het	—	ccatgtgtttccaatgttagagC	ccatgtgtttccaatgttagagT	tgccctagctggtatgcT
IWGSC_CSS_7BL_scaff.6681782	Cadenza0423	1486	C	T	hom	hom	agtaagCGtgacagcaatggG	agtaagCGtgacagcaatggA	AtgtctTtgGtggaagtacatcA
IWGSC_CSS_7BS_scaff.3160328	Cadenza0423	7801	C	T	het	het	tgttaaatGatacagCctgcagC	tgttaaatGatacagCctgcagT	tggaatgggTCgtgttttT
IWGSC_CSS_7DS_scaff.407428	Cadenza0423	2051	G	A	het	het	gtcGCgccatcctgcagaC	gtcGCgccatcctgcagaA	actcatcAggtcagcccaA
IWGSC_CSS_3AL_scaff.442479	Cadenza0364	3198	C	T	het	het	gagtcATaagttggtaagattggC	gagtcATaagttggtaagattggT	GCaGaTaaCaacaggatcacG
IWGSC_CSS_3AL_scaff.4447942	Cadenza0364	11917	G	A	het	het	gtcataaaagattgctcctgtgaaG	gtcataaaagattgctcctgtgaaA	ctcGgatgtgggaggaagA
IWGSC_CSS_3AS_scaff.1557483	Cadenza0364	2547	C	T	het	het	aaagtacacatcatgttaccataaG	aaagtacacatcatgttaccataaA	cgaatccaacgcctcatcA
IWGSC_CSS_3AS_scaff.2648747	Cadenza0364	2688	G	A	het	het	tggAagcAcaaggggccC	tggAagcAcaaggggccT	GccgcgatggagactcG
IWGSC_CSS_3AS_scaff.3304956	Cadenza0364	1017	G	A	het	het	gtcccttgacacagctttG	gtcccttgacacagctttA	cctgctgactacaactcaaT
IWGSC_CSS_3AS_scaff.3321091	Cadenza0364	4585	C	T	het	het	caagaatgATgctgatgttggaG	caagaatgATgctgatgttggaA	acatgctgaatgcgcaatC
IWGSC_CSS_3AS_scaff.3371333	Cadenza0364	538	G	A	het	het	gggaaaCgAgAcgagcgG	gggaaaCgAgAcgagcgA	ccgtgctctcctcacctC
IWGSC_CSS_3AS_scaff.3371815	Cadenza0364	1061	C	T	het	het	atccccacggcacagagA	atccccacggcacagagA	aAttggcccttggtgattcC
IWGSC_CSS_3AS_scaff.3440912	Cadenza0364	4498	G	A	het	het	ccgtaaaactttctgtgcttG	ccgtaaaactttctgtgcttT	atActgacaactacatgatgtgC
IWGSC_CSS_3B_scaff.10343586	Cadenza0364	2242	G	A	het	—	ggttcTgTcctctcttcactG	ggttcTgTcctctcttcactA	tgtgttgaaccgcgaagA
IWGSC_CSS_3AL_scaff.442479	Cadenza0364	3198	C	T	het	het	gagtcATaagttggtaagattggC	gagtcATaagttggtaagattggT	GCaGaTaaCaacaggatcacG
IWGSC_CSS_3AL_scaff.4447942	Cadenza0364	11917	G	A	het	het	gtcataaaagattgctcctgtgaaG	gtcataaaagattgctcctgtgaaA	ctcGgatgtgggaggaagA
IWGSC_CSS_3AS_scaff.1557483	Cadenza0364	2547	C	T	het	het	aaagtacacatcatgttaccataaG	aaagtacacatcatgttaccataaA	cgaatccaacgcctcatcA
IWGSC_CSS_3AS_scaff.2648747	Cadenza0364	2688	G	A	het	het	tggAagcAcaaggggccC	tggAagcAcaaggggccT	GccgcgatggagactcG
IWGSC_CSS_3AS_scaff.3304956	Cadenza0364	1017	G	A	het	het	gtcccttgacacagctttG	gtcccttgacacagctttA	cctgctgactacaactcaaT
IWGSC_CSS_3AS_scaff.3321091	Cadenza0364	4585	C	T	het	het	caagaatgATgctgatgttggaG	caagaatgATgctgatgttggaA	acatgctgaatgcgcaatC
IWGSC_CSS_3AS_scaff.3371333	Cadenza0364	538	G	A	het	het	gggaaaCgAgAcgagcgG	gggaaaCgAgAcgagcgA	ccgtgctctcctcacctC
IWGSC_CSS_3AS_scaff.3371815	Cadenza0364	1061	C	T	het	het	atccccacggcacagagG	atccccacggcacagagA	aAttggcccttggtgattcC
IWGSC_CSS_3AS_scaff.3440912	Cadenza0364	4498	G	A	het	het	ccgtaaaactttctgtgcttG	ccgtaaaactttctgtgcttT	atActgacaactacatgatgtgC
IWGSC_CSS_3B_scaff.10343586	Cadenza0364	2242	G	A	het	—	ggttcTgTcctctcttcactG	ggttcTgTcctctcttcactA	tgtgttgaaccgcgaagA
IWGSC_CSS_5DL_scaff.242342	Cadenza0281	2433	C	T	hom	hom	catggCgacggtGtctG	catggCgacggtGtctA	aAccctcatTTtgGCTACTtCT
IWGSC_CSS_5DL_scaff.4538822	Cadenza0281	1208	G	A	hom	—	acgtcagaacaacccgttgaC	acgtcagaacaacccgttgaT	ttaaattggttgccgccacC
IWGSC_CSS_6AL_scaff.5813297	Cadenza0281	4532	C	T	hom	—	gggagagggagctctcgG	gggagagggagctctcgA	ttctctgccaacagattccG
IWGSC_CSS_6AS_scaff.4378990	Cadenza0281	6748	C	T	hom	hom	cccaggttctgctcttttcC	cccaggttctgctcttttcT	caagtatcaagaaaatgaaggTgT
IWGSC_CSS_6BL_scaff.4360781	Cadenza0281	5426	C	T	het	het	aCtactcaaatgcttGgtgtaG	aCtactcaaatgcttGgtgtaA	tcagtccaacatgTcaagagatT
IWGSC_CSS_7AL_scaff.4488310	Cadenza0281	3808	G	A	hom	hom	gttctctttagtagcagccA	gttctctttagtagcagccA	ggcgctcttctcggcctA
IWGSC_CSS_7BL_scaff.6696509	Cadenza0281	9232	G	A	het	het	gctctaggGgtggcaaAagG	gctctaggGgtggcaaAagA	ggcctGaGgtcGcagtgT
IWGSC_CSS_7BS_scaff.3143575	Cadenza0281	1866	C	T	het	het	agatgttgagaggcgcttT	agatgttgagaggcgcttT	gcttggAtgttggaagtT
IWGSC_CSS_7DL_scaff.3346250	Cadenza0281	1663	G	A	het	het	acgtgcagcaacatcctaaC	acgtgcagcaacatcctaaT	TttccaccaggcccaagA
IWGSC_CSS_7DS_scaff.3933917	Cadenza0281	1243	C	T	het	het	tgCtgagcCttTcaccttgC	tgCtgagcCttTcaccttgT	agaggtttggttcacatGG
IWGSC_CSS_3B_scaff.10626860	Cadenza0148	7847	G	A	het	het	gcagctctgggaaggagG	gcagctctgggaaggagA	gttaatgtacCTcctagctcG
IWGSC_CSS_3DL_scaff.6915683	Cadenza0148	6904	C	T	het	het	cgctcaaCctgtgggcaattG	cgctcaaCctgtgggcaattA	tcatgctcataatgTcatagggT
IWGSC_CSS_4AS_scaff.5929057	Cadenza0148	4238	G	A	hom	hom	gcgcaacgtagCacctacT	gcgcaacgtagCacctacT	ttatctgctgaagtgcaggttCA
IWGSC_CSS_4AS_scaff.5950625	Cadenza0148	10590	C	T	het	het	agaTattCaaaTcggtggAttggC	agaTattCaaaTcggtggAttggT	cctgCtccccctacgtcC
IWGSC_CSS_4AS_scaff.5967119	Cadenza0148	11626	C	T	hom	hom	cgtGgacaccccgagctG	cgtGgacaccccgagctA	gaagcagcactgcagcaC
IWGSC_CSS_4DL_scaff.14455742	Cadenza0148	1946	C	T	hom	hom	gCctgaggagatcgctG	gCctgaggagatcgctT	aaccgGtAaCTGtGgGcA
IWGSC_CSS_4DS_scaff.2318993	Cadenza0148	4000	C	T	hom	hom	tcagtttgacacagattgaatggA	tcagtttgacacagattgaatggA	tgagaTtctgttcccttcacAttG
IWGSC_CSS_5AL_scaff.2750707	Cadenza0148	4603	G	A	het	het	ccttggtgctagccatttcaagTaG	ccttggtgctagccatttcaagTaA	ccaggaTgcAgtgcaatatttcaagG
IWGSC_CSS_5BL_scaff.10794137	Cadenza0148	9235	C	T	hom	hom	gaagctgcttctcggttA	gaagctgcttctcggttA	agtatccctccatataagcagtG
IWGSC_CSS_5BS_scaff.1646558	Cadenza0148	2916	C	T	het	het	gccGtacctacactAtcctttG	gccGtacctacactAtcctttA	gcaaTgtccacttAtcatccT
IWGSC_CSS_1AL_scaff.3883106	Cadenza0110	27536	C	T	het	het	accttccatcactggtgG	accttccatcactggtgA	gtgaagaacaacaggttgaagC
IWGSC_CSS_1BL_scaff.3812829	Cadenza0110	10770	G	A	het*	het	ccccactccattccagG	ccccactccattccagA	gGatgttgtctgtgctggaA
IWGSC_CSS_1DL_scaff.2266648	Cadenza0110	6156	G	A	het	het	actgctgtgttatgggacC	actgctgtgttatgggacT	ccccactggaacacaacA
IWGSC_CSS_1DS_scaff.1889435	Cadenza0110	8826	C	T	hom	hom	aacctgaattactcggacagG	aacctgaattactcggacagA	gcctgaagaattgtatcaaaacA

IWGSC contig	Line	Pos	WT	Mut	Predicted	$M_4$	Primer 1 (Cadenza)	Primer 2 (mutant)	Common Primer
IWGSC_CSS_2AS_scaff.5268634	Cadenza0110	4636	G	A	het	het	gatccatgtgattggcatgtttG	gatccatgtgattggcatgtttA	TgctgtTggatatgcagttacT
IWGSC_CSS_2BL_scaff.7965110	Cadenza0110	15801	C	T	hom	hom	cattgaagcAtacacAattgcAtaC	cattgaagcAtacacAattgcAtaT	gccagagtatccagataaggTttA
IWGSC_CSS_2DL_scaff.9852812	Cadenza0110	13788	G	A	hom	hom	atttttgtatggtctcaatcttcgC	atttttgtatggtctcaatcttcgT	gaacgtTcattcttgcactgtC
IWGSC_CSS_2DS_scaff.5371379	Cadenza0110	2166	C	T	hom	hom	agacacaaaactagtGatgcgC	agacacaaaactagtGatgcgT	gctgctgagaaatgtTgtatttG
IWGSC_CSS_3AL_scaff.4384278	Cadenza0110	1276	C	T	het	het	agcTgaactgccccTgtaG	agcTgaactgccccTgtaA	agggacctCgGtgatgaA
IWGSC_CSS_3AS_scaff.3340122	Cadenza0110	1467	C	T	hom	hom	attcctAgtgttgcggaacatG	attcctAgtgttgcggaacatA	gagaagactagaaagttttcAgcaT
IWGSC_CSS_5DL_scaff.4554222	Cadenza2103	6528	C	T	het*	hom	gctgccctacaagaacaaaattG	gctgccctacaagaacaaaattA	aTcccaactatCGaTtttgcataC
IWGSC_CSS_6AL_scaff.5833640	Cadenza2103	7346	C	T	hom	hom	aagaaaagcccaaatggtttctC	aagaaaagcccaaatggtttctT	aCTctgTcagtggttcccgC
IWGSC_CSS_6AS_scaff.4429974	Cadenza2103	3867	G	A	hom	hom	GagatgaAttattaggcatgtggC	GagatgaAttattaggcatgtggT	ggttccggctgcataaagT
IWGSC_CSS_6DL_scaff.3307626	Cadenza2103	4970	C	T	hom	hom	tgcagatgttgcctgtgtaG	tgcagatgttgcctgtgtaA	ctaggaaggtgatttgtactGtC
IWGSC_CSS_6DS_scaff.2059604	Cadenza2103	5224	G	A	het	—	gctcaatgcatgcTgagtG	gctcaatgcatgcTgagtA	tgtaagattattttctgtctG
IWGSC_CSS_7AL_scaff.4552322	Cadenza2103	1412	C	T	het	het	gcaaaggcTgatactccaacAG	gcaaaggcTgatactccaacA	ggcAAGcAgtataaaagtaaGC
IWGSC_CSS_7BS_scaff.3147455	Cadenza2103	4607	G	A	het	—	gcaccttaggatgtgagTtatG	gcaccttaggatgtgagTtatT	gcattgagggttatttactgttA
IWGSC_CSS_7DL_scaff.3382467	Cadenza2103	3473	C	T	hom	—	GGTtctgCaGTTTCATAActcatC	GGTtctgCaGTTTCATAActcatT	attgaatcaactgatacGaaGactC
IWGSC_CSS_3B_scaff.10457010	Cadenza0277	10599	G	A	het	het	aaccttggccgcagaacacC	aaccttggccgcagaacacT	actgctgcacgagaggG
IWGSC_CSS_3B_scaff.10593852	Cadenza0277	10124	C	T	het	het	tgacaggggacgctatacaG	tgacaggggacgctatacaA	gtctaaCTtACattAcccatcagC
IWGSC_CSS_3DS_scaff.2583390	Cadenza0277	663	G	A	hom	hom	actgcactcatacaatActtCtgC	actgcactcatacaatActtCtgT	tcCacctggacagcaagtG
IWGSC_CSS_4AL_scaff.7093953	Cadenza0277	10004	C	T	hom	hom	cttctgattcaatggaTtgTtttG	cttctgattcaatggaTtgTtttA	ttcccaaaaTaaaagggaagcA
IWGSC_CSS_4AL_scaff.7176064	Cadenza0277	6220	C	T	het	het	gtgccgtaTtcCgcttgA	gtgccgtaTtcCgcttgG	atgttcgaggggaggggG
IWGSC_CSS_4DL_scaff.14122349	Cadenza0277	1010	C	T	hom	hom	gtcgtgctgCttgtgaG	gtcgtgctgCttgtgaA	ggaaaggcccaaggaggG
IWGSC_CSS_5AL_scaff.2736916	Cadenza0277	4296	G	A	het	het	aagaactATgAaaGtaacacacgaC	aagaactATgAaaGtaacacacgaT	ttcGcTtTaaGGeAttCtcG
IWGSC_CSS_5BL_scaff.10883744	Cadenza0277	2080	C	T	hom	hom	gcctctttCtggtTagcctcaG	gcctctttCtggtTagcctcaA	cgacaaggttcgtgatTgcA
IWGSC_CSS_1AL_scaff.3932013	Cadenza0548	11765	C	T	hom	hom	accgcaaaCccaagacaG	accgcaaaCccaagacaA	cccatTAacGcTgcAacG
IWGSC_CSS_1BS_scaff.3417505	Cadenza0548	373	C	T	het	het	gtggtaggaGGgtgGaG	gtggtaggaGGgtgGaA	tggtcGccagtgttgA
IWGSC_CSS_2AS_scaff.5305619	Cadenza0548	2786	C	T	hom	hom	atacagatgccctAAGtggTtC	atacagatgccctAAGtggTtT	ggaaagacaATGctccaggtaC
IWGSC_CSS_2AS_scaff.5306489	Cadenza0548	46953	T	G	het	wt	aggttccatgtccatagaagGT	aggttccatgtccatagaagGG	aggctaATagactcctgtACAgT
IWGSC_CSS_2BL_scaff.7984123	Cadenza0548	11660	G	A	het	het	cattgtggcatagtaatcagtacaG	cattgtggcatagtaatcagtacaA	aatacattgaggaatcaaagccC
IWGSC_CSS_2DL_scaff.9907477	Cadenza0548	1363	C	T	hom	hom	tgctccctttgccagaaC	tgctccctttgccagaaT	ggcaaacctgatgtggcatC
IWGSC_CSS_2DS_scaff.5330886	Cadenza0548	5449	G	A	hom	hom	gcattgccatttatactgaaCgtG	gcattgccatttatactgaaCgtA	catgctgtctcttcggacC
IWGSC_CSS_3AL_scaff.4449951	Cadenza0548	633	C	T	het	het	tccaaacctaacagctctaacactaG	tccaaacctaacagctctaacactaA	gtctgcagTGCaatgtgC
IWGSC_CSS_3B_scaff.10479889	Cadenza0097	3339	C	T	hom	—	ttgTttctGgagaagatgcCG	ttgTttctGgagaagatgcCA	ggtgtcattcaAcGgcA
IWGSC_CSS_3B_scaff.10562262	Cadenza0097	7819	C	T	het	het	agagggggtgctatccatAttgG	agagggggtgctatccatAttgA	agcgatgccaaaggcttcC
IWGSC_CSS_4AL_scaff.7040796	Cadenza0097	10772	G	A	hom	hom	acacaacattgccaccagaG	acacaacattgccaccagaA	CAatCgattgtgtgtTctcC
IWGSC_CSS_4AL_scaff.7063488	Cadenza0097	6360	C	T	het	het	gcctctcacCttAatttgaagctgC	gcctctcacCttAatttgaagctgT	aggcagtggagatgtgaaagtT
IWGSC_CSS_4AL_scaff.7091701	Cadenza0097	5050	G	A	het	het	catgagcatctgggaggaaaatG	catgagcatctgggaggaaaatA	agcaagggaAaatgaacggaaA
IWGSC_CSS_4DS_scaff.1845841	Cadenza0097	7110	G	A	hom	hom	aatgTAGctccccatacCgG	aatgTAGctccccatacCgA	actgaacATgcaactcgtTtatggA
IWGSC_CSS_5AL_scaff.2767581	Cadenza0097	3737	G	A	het	het	gagaggtcctcactAtcggC	gagaggtcctcactAtcggT	cgTcatcacaatattgtctgG
IWGSC_CSS_5BL_scaff.10784643	Cadenza0097	1568	C	T	hom	hom	agaaaTAcatggatggatggaCG	agaaaTAcatggatggatggaCA	catctcCtccaCgGaaaG
IWGSC_CSS_1AL_scaff.3952258	Cadenza2092	8107	C	T	het	—	tgagttagagaaattgacagtgtgG	tgagttagagaaattgacagtgtA	tgccaccattgacatgagaG
IWGSC_CSS_1BL_scaff.3858008	Cadenza2092	10278	G	A	hom	hom	tttgagcaggcaggatcgC	tttgagcaggcaggatcgT	actcacggcctatcActattC
IWGSC_CSS_1DL_scaff.2265172	Cadenza2092	9094	C	T	hom	hom	tgcaTGTcattgttcttatcagC	tgcaTGTcattgttcttatcagT	agtgtccaacttcGttcatC
IWGSC_CSS_2AL_scaff.6435867	Cadenza2092	16201	G	A	hom	hom	tttctGTaccttaacgtcaattgaC	tttctGTaccttaacgtcaattgaT	gtgaggatgatgaggtaagacC
IWGSC_CSS_2AL_scaff.6439430	Cadenza2092	25101	C	T	het	—	caagaaaggcCagCtCagC	caagaaaggcCagCtCagT	tcGtAcTcttctActggtgaA
IWGSC_CSS_2DL_scaff.9760848	Cadenza2092	4733	C	T	het	het	gcaccatgggtctcaggtaC	gcaccatgggtctcaggtaT	tcagtcagtttGCTCTGTCTG
IWGSC_CSS_3AL_scaff.4407012	Cadenza2092	2785	C	T	hom	hom	acatatAgtgttctcatccaccatC	acatatAgtgttctcatccaccatT	acctctctcatgttaatagtttG
IWGSC_CSS_3AS_scaff.3441108	Cadenza2092	541	G	A	het	het	GtgatgaccttgagacGgaG	GtgatgaccttgagacGgaA	aggcaTgacaaCgcgcaA
IWGSC_CSS_3B_scaff.10449827	Cadenza1551	4779	G	A	hom	hom	ggcaaggcTcaagaaacGgtC	ggcaaggcTcaagaaacGgtT	aCagaTggggttagaggcaG
IWGSC_CSS_3B_scaff.10550638	Cadenza1551	3250	C	T	het	het	ctccttcactgtgttgcggC	ctccttcactgtgttgcggT	gcaacATTtgtaactgcaaaG
IWGSC_CSS_3DL_scaff.6945816	Cadenza1551	589	C	T	hom	hom	agcatctcactgcaacCaataC	agcatctcactgcaacCaataT	TgtgcccTctgaAtattttcaTG
IWGSC_CSS_3DL_scaff.6954177	Cadenza1551	3508	C	T	het	het	tgtagcatcacattaactttctC	tgtagcatcacattaactttctA	gcttggtataaacCTtaccagA
IWGSC_CSS_4AS_scaff.5938272	Cadenza1551	19080	G	A	hom	hom	agAcCccgAtcgccatgG	agAcCccgAtcgccatgA	GggAgatAcagggtaaaActcTtcG
IWGSC_CSS_4AS_scaff.5977594	Cadenza1551	11092	C	T	het	het	gccttgattcggaacacaaaC	gccttgattcggaacacaaaT	gcgtctctcagtcctgcA
IWGSC_CSS_5AL_scaff.2671035	Cadenza1551	5859	C	T	het	het	cggtgatattTtagactcgacgC	cggtgatattTtagactcgacgT	ggcagttcagcGaccatT
IWGSC_CSS_5BL_scaff.10889480	Cadenza1551	2530	G	A	hom	hom	gagcttaactcgagatggaG	gagcttaactcgagatggaA	tccatGCAacGccttggT

IWGSC contig	Line	Pos	WT	Mut	Predicted	$M_4$	Primer 1 (Cadenza)	Primer 2 (mutant)	Common Primer
IWGSC_CSS_3B_scaff_10528396	Cadenza2088	8059	G	A	hom	—	cttttcgctccgtaagcaataG	cttttcgctccgtaagcaataA	gtgcaactgttcaggcctgA
IWGSC_CSS_3B_scaff_10637573	Cadenza2088	16815	G	A	het	het	agcaagctttaccGgtctgC	agcaagctttaccGgtctgT	cgagcAactacgagcagctT
IWGSC_CSS_4AL_scaff_7086469	Cadenza2088	6697	G	A	het	het	gccgtctacttcaacgcG	gccgtctacttcaacgcA	ccaGaggcttgTGCattttT
IWGSC_CSS_4AL_scaff_7126302	Cadenza2088	3627	G	A	hom	hom	gttcaaaaaaagtggtAatttgC	gttcaaaaaaagtggtAatttgT	cacaaggatatagaagCTcttctagA
IWGSC_CSS_4BL_scaff_7041808	Cadenza2088	10234	G	A	hom	hom	tcaatggatgagggtgcttC	tcaatggatgagggtgcttT	ccatagcagcatcagccacA
IWGSC_CSS_5AL_scaff_2794167	Cadenza2088	13162	G	A	het	—	agtattcaggacaagcatCttCaG	agtattcaggacaagcatCttCaA	caatgaacctctcgaagaaGaG
IWGSC_CSS_5BL_scaff_10889232	Cadenza2088	3885	G	A	het	het	cTcaaccacaatgggcaAatC	cTcaaccacaatgggcaAatT	tccttcatcaatcatcaattgttgG
IWGSC_CSS_5BS_scaff_2267405	Cadenza2088	11113	C	T	hom	hom	ctttgatgatcctaggcctctTG	ctttgatgatcctaggcctctTA	tgatttggtCTggttAgagtttGA
IWGSC_CSS_3B_scaff_10475354	Cadenza1409	2203	G	A	hom	hom	agCgaacaagagGtcaaacA	agCgaacaagagGtcaaacA	ctgaaacacaCTagaCAattAccG
IWGSC_CSS_3B_scaff_10674115	Cadenza1409	4555	C	T	het	het	gcttcagtgcatgccttcaG	gcttcagtgcatgccttcaA	cttcacaccGagataatGtattG
IWGSC_CSS_4AL_scaff_7153568	Cadenza1409	13073	C	T	hom	hom	tccgaccgAtcaaccttgG	tccgaccgAtcaaccttgA	gaccggaactcctcggcC
IWGSC_CSS_4DL_scaff_14314966	Cadenza1409	2010	G	A	het	hom	gtaggctccctcctCAGgG	gtaggctccctcctCAGgA	cggcgTcacaAgttgCcT
IWGSC_CSS_4DS_scaff_2324074	Cadenza1409	7606	G	A	het	het	tGcatgaaaatgtgtGcaGaG	tGcatgaaaatgtgtGcaGaA	gggtaAGttcAaaactGaatgaaG
IWGSC_CSS_5AS_scaff_1517889	Cadenza1409	3561	G	A	het	het	tctcgacatcttcccggtgaT	tctcgacatcttcccggtgaT	gtcgctggaaacattgcttattA
IWGSC_CSS_5AS_scaff_1523866	Cadenza1409	8054	G	A	hom	—	ggtgatctaccgccaGgaC	ggtgatctaccgccaGgaT	tctcgagCcTctctcA
IWGSC_CSS_5BL_scaff_10917655	Cadenza1409	19073	G	A	hom	hom	caaatgacatgcaaaagaagtG	caaatgacatgcaaaagaagtT	cgcttcatcatcacaAaatatgtcT
IWGSC_CSS_1AL_scaff_3886649	Cadenza1599	5204	C	T	het	het	tgatgcccaaccacaatGcC	tgatgcccaaccacaatGcT	ggactgactgctgacatatttaG
IWGSC_CSS_1BL_scaff_3810267	Cadenza1599	6634	C	T	hom	hom	ccCaggaaatgagcacctC	ccCaggaaatgagcacctT	cgaggcgcaaatgagtgtaTtG
IWGSC_CSS_1DL_scaff_2291677	Cadenza1599	12856	C	T	hom	hom	GgtagacaagtgcgccaG	GgtagacaagtgcgccaA	ctctccttcaacGCcG
IWGSC_CSS_2AL_scaff_6354492	Cadenza1599	7566	G	A	het	het	gGagaatgcaCAgtAacTtctgG	gGagaatgcaCAgtAacTtctgA	tcccggaagaacacaTccTG
IWGSC_CSS_2AS_scaff_5282937	Cadenza1599	9736	G	A	het	het	gctgtagattttatagctgctatG	gctgtagattttatagctgctatT	cacCagaattgttCactgattTC
IWGSC_CSS_2BL_scaff_7952427	Cadenza1599	19249	G	A	hom	hom	cgTccctCcttagcagcaG	cgTccctCcttagcagcaT	aTcactcattagcgagAG
IWGSC_CSS_2DL_scaff_9897981	Cadenza1599	5627	C	T	het	het	cttggtgctTgattgcttactC	cttggtgctTgattgcttactT	gTtggtCtctctgactTtgtG
IWGSC_CSS_3AL_scaff_4446105	Cadenza1599	1765	G	A	hom	—	aaatgctttcctaCcgctagtG	aaatgctttcctaCcgctagtA	ttctAgaggcaatagctTatatgcT

Table A.3: Validation of mutations on  $M_4$  on Kronos

IWGSC contig	Line	Pos	WT	Mut	Predicted	$M_4$	Primer 1 (Kronos)	Primer 2 (mutant)	Common Primer
IWGSC_CSS_1AS_scaff_3284790	Kronos3085	7449	G	A	Het	Het	ccacaccttgagcctcgC	ccacaccttgagcctcgT	gtgattttgcccaggaggA
IWGSC_CSS_1BL_scaff_3897513	Kronos3085	1515	C	T	Het	Het	gcttccactGggtcctgC	gcttccactGggtcctgT	acAaggactgcttcagaGaC
IWGSC_CSS_2AL_scaff_6434745	Kronos3085	3424	C	T	Het	Het	cctcGgttttgcaaatttctatgC	cctcGgttttgcaaatttctatgT	gGCaaTggcataacaacagatA
IWGSC_CSS_3AS_scaff_3408995	Kronos3085	732	C	T	Het	Het	aggccatttcgaattccgG	aggccatttcgaattccgT	ggTgttaTccagAacctgagTG
IWGSC_CSS_3B_scaff_10708748	Kronos3085	2675	G	A	Het	Het	gttgcatgcttcacccagG	gttgcatgcttcacccagA	gtaacaactctgagttctgtagcA
IWGSC_CSS_4AL_scaff_7132733	Kronos3085	1799	C	T	Hom	Hom	caccctgtgagtaccctC	caccctgtgagtaccctT	aCcGcctaGaaagaaagcttC
IWGSC_CSS_5AL_scaff_1534693	Kronos3085	4605	C	T	Het	Het	cagcttctcgggcctcAtC	cagcttctcgggcctcAtT	gtaCctcacAgtaCTgagAG
IWGSC_CSS_6AS_scaff_4361911	Kronos3085	8857	G	A	Het	Het	tcacgaaaagacgacttcaacctC	tcacgaaaagacgacttcaacctT	catgaggtgctgcatctccatcA
IWGSC_CSS_6BS_scaff_3008326	Kronos3085	1528	G	A	Het	Het	ccatgttgtagctggtggtgC	ccatgttgtagctggtggtgT	ggaagcatggCaagtgcA
IWGSC_CSS_7AS_scaff_4214385	Kronos3085	27835	C	T	Hom	Hom	cgtaccttctgttgggaagG	cgtaccttctgttgggaagA	ctcttggtcagctgtataagacT
IWGSC_CSS_1AL_scaff_3929964	Kronos3191	1336	C	T	Het	Het	tttcggccataacctgacatC	tttcggccataacctgacatT	attgcttccagtttctgcaG
IWGSC_CSS_1BL_scaff_3899789	Kronos3191	7925	C	T	Het	Het	actctcacTggcagcagC	actctcacTggcagcagT	caactgtgtgcccacGtA
IWGSC_CSS_2AL_scaff_6426728	Kronos3191	1481	G	A	Hom	Hom	gaaActgcccgagctCgC	gaaActgcccgagctCgT	ccaGcaGctcgtgagaaA
IWGSC_CSS_2BL_scaff_7960273	Kronos3191	690	C	T	Hom	Hom	gccattcatccttaggcgC	gccattcatccttaggcgT	acatgcaattgctgactgactG
IWGSC_CSS_3AS_scaff_3286603	Kronos3191	2975	G	A	Het*	Hom	ccgtgtggtttgtgtggG	ccgtgtggtttgtgtggA	gaaaggaaactgTcaTgcaG
IWGSC_CSS_5AL_scaff_2694249	Kronos3191	2399	C	T	Het	Het	gccttcagatagagccG	gccttcagatagagccG	cgccacatcgacattcctG
IWGSC_CSS_5BL_scaff_10923577	Kronos3191	3713	C	T	Het	Het	gtggattgcctgagcttgC	gtggattgcctgagcttgT	tgggtgctcttgggaC
IWGSC_CSS_6AL_scaff_5823017	Kronos3191	13225	C	T	Hom	Hom	ccctttcgagcctctggaG	ccctttcgagcctctggaA	ttcgagaagccccatcgA
IWGSC_CSS_6BS_scaff_2955394	Kronos3191	1622	C	T	Het*	Hom	gtggagatgaaggtctagcaG	gtggagatgaaggtctagcaA	gataactgTgcaatgggtT
IWGSC_CSS_7BL_scaff_6739382	Kronos3191	12261	G	A	Hom	Hom	gagacaagctttgaattgctcC	gagacaagctttgaattgctcT	CgagtgcactTcatttcccG

IWGSC contig	Line	Pos	WT	Mut	Predicted	$M_4$	Primer 1 (Kronos)	Primer 2 (mutant)	Common Primer
IWGSC.CSS_1AS_scaff.3276389	Kronos3288	9720	C	T	Hom	Hom	aCcaGcaggaccAatgtctC	aCcaGcaggaccAatgtctT	atgatgcaacctcagccaT
IWGSC.CSS_2AL_scaff.6367515	Kronos3288	6976	G	A	Het	Het	caggtcgagTgtctccgG	caggtcgagTgtctccgA	gggggtgatCtgggaaggcC
IWGSC.CSS_2AL_scaff.6422019	Kronos3288	4523	G	A	Het	Het	cgtaggtccctgcataG	cgtaggtccctgcataG	acgcAcgctaagccgtaC
IWGSC.CSS_3AL_scaff.4284850	Kronos3288	7901	C	T	Hom	Hom	tggtttggacaacatcgG	tggtttggacaacatcgA	gttcAgcatcgacagccaG
IWGSC.CSS_4AS_scaff.5962359	Kronos3288	13049	G	A	Het	Hom	ccatcaagaagtacgagttcgaC	ccatcaagaagtacgagttcgaT	accatgcccagcttgtcA
IWGSC.CSS_6AL_scaff.5778773	Kronos3288	6853	G	A	Het	Het	gagtgaccttcccgtcttC	gagtgaccttcccgtcttT	ggagaacagctactcggtC
IWGSC.CSS_6AS_scaff.4392100	Kronos3288	3434	C	T	Het	Het	atggaagcacaggtgaccG	atggaagcacaggtgaccA	ggAagcgaaagtgaacaaacA
IWGSC.CSS_7BL_scaff.6744240	Kronos3288	9772	G	A	Het	Het	agctgtttcttctacttcaaG	agctgtttcttctacttcaaA	caggtcgttcttgagctcC
IWGSC.CSS_1AL_scaff.3887185	Kronos3413	9708	C	T	Hom	Hom	gcacgcctttatcgaggtaaaG	gcacgcctttatcgaggtaaaA	AgaaacagcagagcgcaA
IWGSC.CSS_2BS_scaff.3381362	Kronos3413	5160	C	T	Het*	Hom	caacttctgggctgtagtG	caacttctgggctgtagtA	tgAgaattctgacGcaaaagaC
IWGSC.CSS_3AS_scaff.3296605	Kronos3413	6154	G	A	Het	Het	ctggtcacgggctctagC	ctggtcacgggctctagT	cagcactgagagacatggaC
IWGSC.CSS_3B_scaff.10693516	Kronos3413	12632	C	T	Het	Het	ctaggcttgacaaacaggC	ctaggcttgacaaacaggT	agcttgcatctatgggcatT
IWGSC.CSS_5AS_scaff.1547699	Kronos3413	2686	G	A	Het	Het	gCtacaaccttcaccaatcgC	gCtacaaccttcaccaatcgT	gacggcttgaagtgtcatC
IWGSC.CSS_5BL_scaff.10856077	Kronos3413	5853	G	A	Het	Het	agagcttcaccccatgctC	agagcttcaccccatgctT	acgCacatttAatagctgaagC
IWGSC.CSS_6AL_scaff.5750718	Kronos3413	11046	G	A	Hom	Hom	cacgcTtcccgacttctataG	cacgcTtcccgacttctataA	AgacgatgtgatcaggattcaG
IWGSC.CSS_7AL_scaff.4433177	Kronos3413	3511	C	T	Het	Het	GaTgtccGtcaggctgG	GaTgtccGtcaggctgA	cactactggacaagctcttgG
IWGSC.CSS_7BL_scaff.6742567	Kronos3413	667	C	T	Het	Het	gttgcttgctggcagaC	gttgcttgctggcagaT	cattttgacgtgtgtcTG
IWGSC.CSS_1AL_scaff.3976389	Kronos3935	10941	C	T	Hom	Hom	gggtgaggagatcggCgatG	gggtgaggagatcggCgatA	cagtcacatcacatgagagtgcaG
IWGSC.CSS_1BL_scaff.3873362	Kronos3935	1392	G	A	Het	Het	cagatctgaagcctaGcacatG	cagatctgaagcctaGcacatA	actaccagatcagcacacaaaAC
IWGSC.CSS_2BL_scaff.7882382	Kronos3935	2721	C	T	Het	Het	gcaagctaagatgtaccgtagC	gcaagctaagatgtaccgtagT	gccacagtaggagaaagactT
IWGSC.CSS_3AL_scaff.4242376	Kronos3935	2410	C	T	Het	Het	agaaccctaaacccgTacttaG	agaaccctaaacccgTacttaA	gtagGgtCcaatCtaaagcttG
IWGSC.CSS_3B_scaff.10485067	Kronos3935	3349	C	T	Hom	Hom	gcttgagcaactactccaactG	gcttgagcaactactccaactA	gcaatttctcttaTccgcagT
IWGSC.CSS_4AS_scaff.5984153	Kronos3935	6006	G	A	Het	Het	agCaggtctggccaagtG	agCaggtctggccaagtA	cgaatGtatgaGtaggcgcT
IWGSC.CSS_4BL_scaff.7019402	Kronos3935	9081	C	T	Het	Het	tgcaatcatgtagtgcgtG	tgcaatcatgtagtgcgtA	agcatgatccctagaaCcataC
IWGSC.CSS_5BL_scaff.10842786	Kronos3935	3304	G	A	Het	Het	tggttcccGaaagcctgaaC	tggttcccGaaagcctgaaT	cgcatacttgaacaTGagcAC
IWGSC.CSS_6BS_scaff.3045205	Kronos3935	2293	G	A	Het	Het	aaggaccaagcccaactctcG	aaggaccaagcccaactctcA	agtgatcaagcccaatgtcgcA
IWGSC.CSS_7AL_scaff.4555249	Kronos3935	4487	C	T	Het	Het	cAgtgctcgagatggcgC	cAgtgctcgagatggcgT	cCttgcaacctctctgatT
IWGSC.CSS_1BL_scaff.3918498	Kronos4240	6096	G	A	Het	Het	ttgcatgccccagaagaG	ttgcatgccccagaagaA	tgggcgaactggtaatgtgG
IWGSC.CSS_2BS_scaff.5131713	Kronos4240	5900	G	A	Het	Het	cctttatcgaggaaagagacacC	cctttatcgaggaaagagacacT	caccattgtagggttcttTttC
IWGSC.CSS_5AL_scaff.2769540	Kronos4240	9626	C	T	Het	Het	tgCagtggtgggaacgggA	tgCagtggtgggaacgggA	catgagtGagatcttctgcT
IWGSC.CSS_5BL_scaff.10871091	Kronos4240	7062	G	A	Het	Het	gccaaggAaccataacctgC	gccaaggAaccataacctgT	GgactcttggcAaccggA
IWGSC.CSS_6AL_scaff.5800333	Kronos4240	2360	G	A	Het	Het	cgacagagattgtgagCgC	cgacagagattgtgagCgT	tcagatgctcgaagattcatcT
IWGSC.CSS_7BL_scaff.6716931	Kronos4240	2613	G	A	Het	Het	gGtgGgtattTgcttggtaG	gGtgGgtattTgcttggtaA	tgGtgactgcagacGtGtA
IWGSC.CSS_2BL_scaff.8029221	Kronos4346	2860	G	A	Het	Het	tgcttccgctcttgctcC	tgcttccgctcttgctcT	atTtgcatTCTGAtcgggcC
IWGSC.CSS_3B_scaff.10460714	Kronos4346	14359	C	T	Hom	Hom	ctacctgccatgcgacatG	ctacctgccatgcgacatA	agcaccctcatctcttgacG
IWGSC.CSS_4AS_scaff.5989735	Kronos4346	6404	G	A	Hom	Hom	acgcatgctaacatcagcT	acgcatgctaacatcagcA	actcaagataccaCcgacG
IWGSC.CSS_5BL_scaff.7648030	Kronos4346	6893	C	T	Het	Het	taccttttctactggcagG	taccttttctactggcagA	ttttcagaggaaacacaggtatcA
IWGSC.CSS_6AL_scaff.5755840	Kronos4346	778	C	T	Het	Het	atcgagtaagctgtcacCgC	atcgagtaagctgtcacCgT	acctgcattgtcaCatccaC
IWGSC.CSS_6BS_scaff.2972151	Kronos4346	7876	G	A	Hom	Hom	gcagcaatgtcActgtttgG	gcagcaatgtcActgtttgA	gcttggactgggcatttatG
IWGSC.CSS_7AL_scaff.4542983	Kronos4346	18700	G	A	Het	Het	gcagggctAccggatacC	gcagggctAccggatacT	catctgccGgttaaacatgC
IWGSC.CSS_7BS_scaff.3098098	Kronos4346	5183	C	T	Het	Het	gCgatatggtacttgcaatgaG	gCgatatggtacttgcaatgaA	ttacattgtctataGTTtgCcgG
IWGSC.CSS_1AS_scaff.3259804	Kronos4485	219	C	T	Het	Het	gtcggcacaccccttgC	gtcggcacaccccttgT	gcttctttaaaggaggcgA
IWGSC.CSS_2AL_scaff.6315418	Kronos4485	10490	G	A	Hom	Hom	gccctctctaaCcttctcagC	gccctctctaaCcttctcagT	ttcagacgctCgaggaaatttccC
IWGSC.CSS_2BS_scaff.5181092	Kronos4485	3742	G	A	Het	Het	TggccagcacacctgcaG	TggccagcacacctgcaA	tggacgatgagTgatggAaaT
IWGSC.CSS_3B_scaff.10425015	Kronos4485	2372	C	T	Het	Het	gctactgaagttggCtcGG	gctactgaagttggCtcGA	cttcacatcttgggggTtC
IWGSC.CSS_3B_scaff.10775915	Kronos4485	4701	C	T	Het	Het	ccaagggtcgacagagG	ccaagggtcgacagagA	agacctcagatGtctcC
IWGSC.CSS_5AL_scaff.2754304	Kronos4485	2301	G	A	Het	Het	taacccTgccatcgcccG	taacccTgccatcgcccA	catgtGccagccaTgacT
IWGSC.CSS_5BL_scaff.10919959	Kronos4485	1867	C	T	Hom	Hom	gatgccttgggtggagaagG	gatgccttgggtggagaagA	tcttgttcccgaacatgtcA
IWGSC.CSS_7AS_scaff.4245431	Kronos4485	3402	G	A	Hom	Hom	aaggcgctgtgttttC	aaggcgctgtgttttT	agtaagtggAacagctaagatcaT
IWGSC.CSS_7BL_scaff.6667357	Kronos4485	641	C	T	Het	Het	gatcAgctgctcattcgagG	gatcAgctgctcattcgagA	ttccctgtcaattgatgccC

# Bibliography

- Alexandra M Allen, Gary L a Barker, Simon T Berry, et al. Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (*Triticum aestivum* L.). *Plant biotechnology journal*, 9(9):1086–99, December 2011. ISSN 1467-7652. doi: 10.1111/j.1467-7652.2011.00628.x.
- Raoul J P Bonnal, Jan Aerts, George Githinji, et al. Biogem: An effective tool-based approach for scaling up open source software development in bioinformatics. *Bioinformatics*, 28(7):1035–1037, April 2012. ISSN 13674803. doi: 10.1093/bioinformatics/bts080.
- C. Burt, A. Steed, N. Gosman, et al. Mapping a type 1 fhb resistance on chromosome 4as of triticum macha and deployment in combination with two type 2 resistances. *Theoretical and Applied Genetics*, 128(9):1725–1738, 2015. ISSN 1432-2242. doi: 10.1007/s00122-015-2542-9.
- Dario Cantu, Manjula Govindarajulu, Alex Kozik, et al. Next generation sequencing provides rapid access to the genome of *Puccinia striiformis* f. sp. *tritici*, the causal agent of wheat stripe rust. *PLoS ONE*, 6(8):1–8, 08 2011. doi: 10.1371/journal.pone.0024230.
- Jarrold a Chapman, Martin Mascher, Aydn Buluç, et al. A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. *Genome Biology*, 16(1):1–17, 2015. ISSN 1465-6906. doi: 10.1186/s13059-015-0582-8.
- A Cornish-Bowden. Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. *Nucleic acids research*, 13(9):3021–30, May 1985. ISSN 0305-1048.
- Graham J Etherington, R. H. Ramirez-Gonzalez, and D. MacLean. biosamtools 2: a package for analysis and visualization of sequence and alignment data with SAMtools in Ruby. *Bioinformatics*, pages 1–2, 2015. ISSN 1367-4803. doi: 10.1093/bioinformatics/btv178.
- Naohisa Goto, Pjotr Prins, Mitsuteru Nakao, et al. BioRuby: bioinformatics software for the Ruby programming language. *Bioinformatics (Oxford, England)*, 26(20):2617–9, October 2010. ISSN 1367-4811. doi: 10.1093/bioinformatics/btq475.

- Amelia Hubbard, Clare M Lewis, Kentaro Yoshida, et al. Field pathogenomics reveals the emergence of a diverse wheat yellow rust population. *Genome Biology*, 16(1):1–15, 2015. ISSN 1465-6906. doi: 10.1186/s13059-015-0590-8.
- Kazutaka Katoh and Daron M Standley. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, 30(4):772–80, April 2013. ISSN 1537-1719. doi: 10.1093/molbev/mst010.
- W. J. Kent. BLAT—The BLAST-Like Alignment Tool. *Genome Research*, 12(4):656–664, March 2002. ISSN 1088-9051. doi: 10.1101/gr.229202.
- Robert King, Nicholas Bird, Ricardo Ramirez-Gonzalez, et al. Mutation scanning in wheat by exon capture and next-generation sequencing. *PLoS ONE*, 10(9):1–18, 09 2015. doi: 10.1371/journal.pone.0137549.
- Ksenia Krasileva, Hans Vasquez-Gross, Tyson Howell1, et al. Uncovering hidden variation in young polyploid wheat genomes. *Science (New York, N.Y.)*, submitted 2016.
- Ksenia V Krasileva, Vince Buffalo, Paul Bailey, et al. Separating homeologs by phasing in the tetraploid wheat transcriptome. *Genome biology*, 14(6):R66, June 2013. ISSN 1465-6914. doi: 10.1186/gb-2013-14-6-r66.
- LGC Genomics. <http://www.lgcgroup.com/services/genotyping/>, 2013. URL <http://www.lgcgroup.com/services/genotyping/>.
- Jian Ma, Jiri Stiller, Zhi Zheng, et al. A high-throughput pipeline for detecting locus-specific polymorphism in hexaploid wheat (*triticum aestivum* l.). *Plant Methods*, 11(1), aug 2015. doi: 10.1186/s13007-015-0082-6.
- K. F. X. Mayer, J. Rogers, J. Dole el, et al. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, 345(6194):1251788–1251788, July 2014. ISSN 0036-8075. doi: 10.1126/science.1251788.
- JU Pontius, L Wagner, and GD. Schuler. UniGene: A Unified View of the Transcriptome. In *The NCBI Handbook [Internet]*, chapter 21. National Center for Biotechnology Information (US), October 2002. URL <http://www.ncbi.nlm.nih.gov/books/NBK21083/>.
- R. H. Ramirez-Gonzalez, Cristobal Uauy, and Mario Caccamo. PolyMarker: A fast polyploid primer design pipeline. *Bioinformatics*, pages 2–3, 2015. ISSN 1367-4803. doi: 10.1093/bioinformatics/btv069.
- Ricardo H Ramirez-Gonzalez, Raoul Bonnal, Mario Caccamo, and Daniel Maclean. Bio-samtools: Ruby bindings for SAMtools, a library for accessing BAM files containing high-throughput sequence alignments. *Source code for biology and medicine*, 7(1):6, January 2012. ISSN 1751-0473. doi: 10.1186/1751-0473-7-6.



- Ricardo H Ramirez-Gonzalez, Vanesa Segovia, Nicholas Bird, et al. RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. *Plant Biotechnology Journal*, 12(9):1–12, November 2014. ISSN 14677644. doi: 10.1111/pbi.12281.
- Steve Rozen and Helen Skaletsky. Primer3 on the WWW for general users and for biologist programmers. *Methods in molecular biology (Clifton, N.J.)*, 132: 365–86, January 2000. ISSN 1064-3745.
- Guy St C Slater and Ewan Birney. Automated generation of heuristics for biological sequence comparison. *BMC bioinformatics*, 6:31, January 2005. ISSN 1471-2105. doi: 10.1186/1471-2105-6-31.
- Martin Trick, Nikolai Adamski, Sarah G Mugford, et al. Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. *BMC plant biology*, 12(1):14, January 2012. ISSN 1471-2229. doi: 10.1186/1471-2229-12-14.
- Shichen Wang, Debbie Wong, Kerrie Forrest, et al. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant biotechnology journal*, 12(6):787–796, March 2014. ISSN 1467-7652. doi: 10.1111/pbi.12183.
- Yi Wang, Vijay K. Tiwari, Nidhi Rawat, et al. GSP: a web-based platform for designing genome-specific primers in polyploids. *Bioinformatics*, page btw134, mar 2016. doi: 10.1093/bioinformatics/btw134.
- Paul a Wilkinson, Mark O Winfield, Gary L a Barker, et al. CerealsDB 2.0: an integrated resource for plant breeders and scientists. *BMC bioinformatics*, 13(1):219, January 2012. ISSN 1471-2105. doi: 10.1186/1471-2105-13-219.