

Contents

1	Introduction	3
2	Literature review	4
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
3	Genetic mapping of <i>Yr15</i>	6
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
4	PolyMarker: A fast polyploid primer design pipeline	8
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 f. sp. 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
5	Gene expression (expVIP)	15
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
6 Conclusions and final remarks	16
A Supplemental tables.	17
A.1 Validation of mutations on M_4 on Kronos	18
A.2 Validation of mutations on M_4 on Cadenza	19

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 F₂ 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 sequences are searched on the genomic reference using **exonerate** (Slater and Birney, 2005) to find the homoeologue regions to the target sequence. Then, the alignment between homoeologues is refined using **MAFFT** (Kato 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-

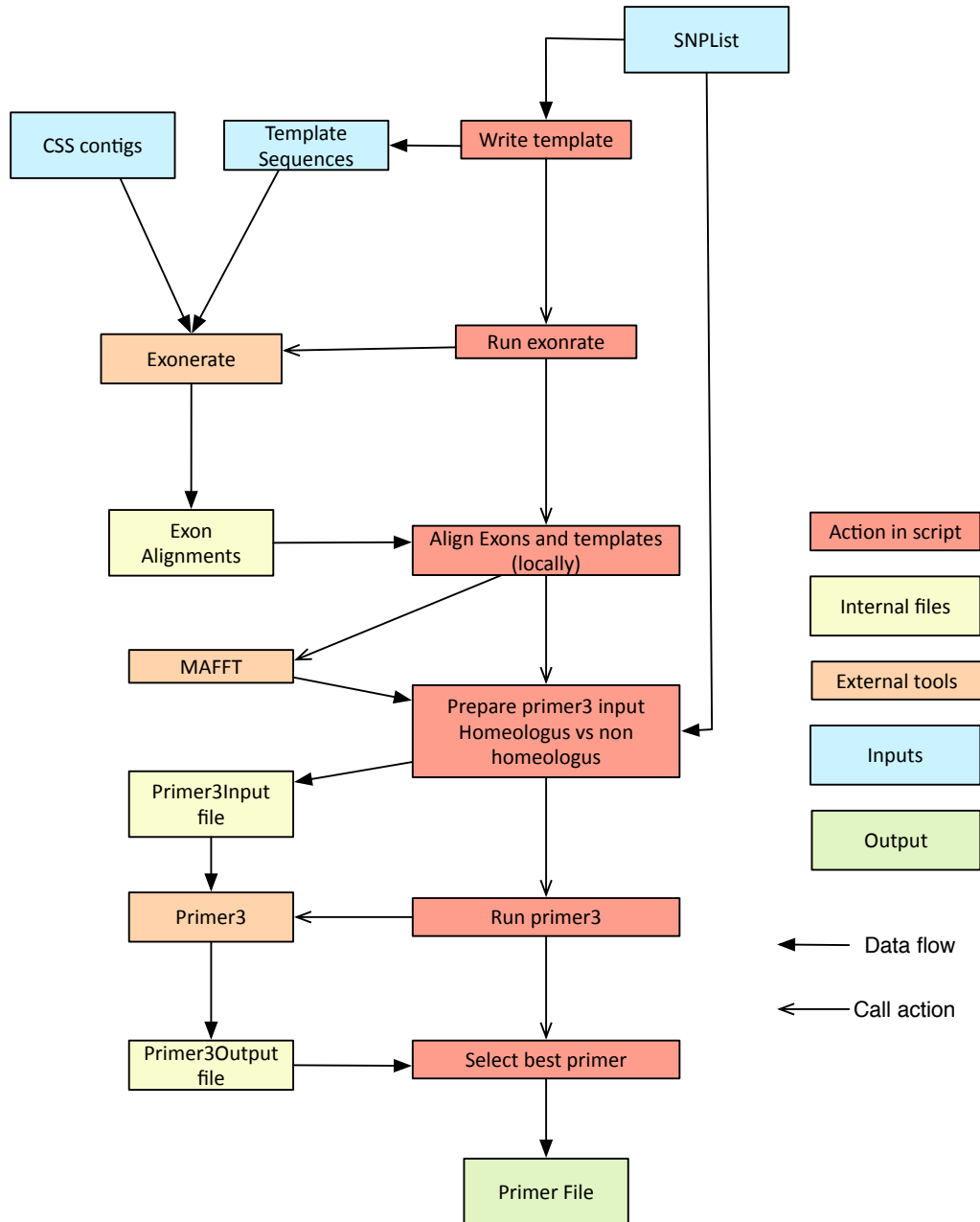


Figure 4.1: Steps and tools called by PolyMarker. The colour of the boxes represent: the step is an action inside the script (red); actions of the script (orange); temporary files (yellow); inputs (blue) and; output (green)

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 searched in the reference sequence using `exonerate` (Slater and Birney, 2005), figure 4.2b. The alignment is run 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:

Specific	Homoeologous polymorphism which is only present in the target genome (upper case).
Semi-specific	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).
Non-specific	No variation is found across homoeologues (-).
Homoeologous	The target SNP is present across different chromosomes, so candidate SNP markers on this category are not expected to be reliably identify the allele (:).
Non-homoeologous	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

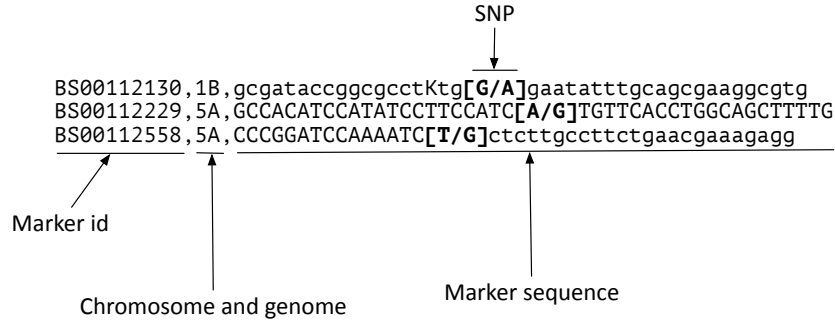
PolyMarker is not restricted to wheat or to KASP assays, the source code is flexible and can be extended for other types of analysis. On each of the following projects, PolyMarker has been adapted to design primers in species where KASP hasn't been used before, the primers are used for regular PCR amplification,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

4.3.1 KASP assays for public sets of SNPs

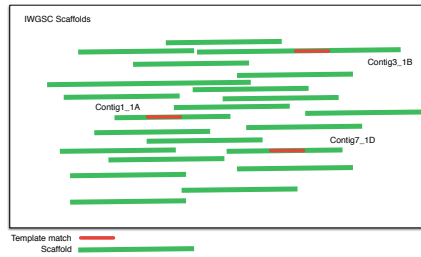
PolyMarker was used to design KASP assays for the 81,587 markers from (Wang et al., 2014), available on the PolyMarker website and in CerealsDB (Wilkinson et al., 2012). Of those markers, 40,267 were designed using the target chromosome using the genetic map published by the genetic map. Genes without 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.

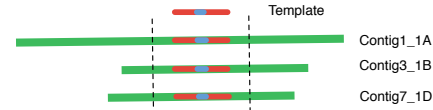
In Hubbard et al. (2015), *Puccinia striiformis* f. sp. *tritici* (PST) isolates were sequenced and assigned to clusters, according to their genotype. The clusters are useful to monitor the changes in the pathogen population, which can be used to predict if certain wheat lines will be resistant to the isolates in



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



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



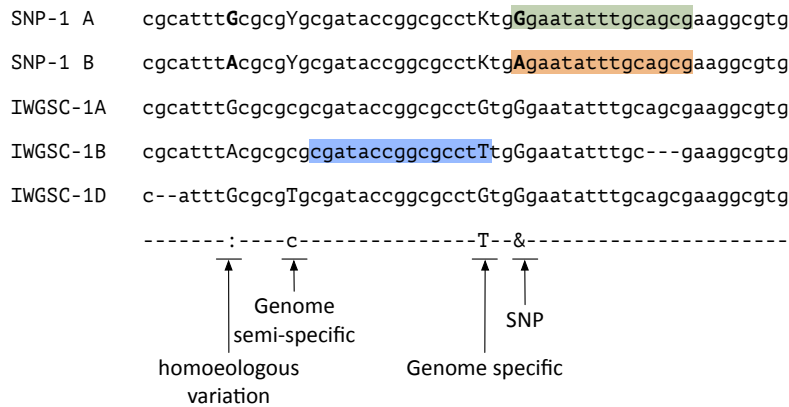
(c) Selected regions around the SNP on every chromosome.

SNP-1 A cgcatttGcgcgYgcgataccggcgccctKtgGgaatatttgcagcgaaggcgtg
 SNP-1 B cgcatttAcgcgYgcgataccggcgccctKtgAgaatatttgcagcgaaggcgtg
 IWGSC-1A cgcatttGcgcgYgcgataccggcgccctGtgGgaatatttgcagcgaaggcgtg
 IWGSC-1B cgcatttAcgcgYgcgataccggcgccctTtgGgaatatttgcagcgaaggcgtg
 IWGSC-1D c--atttGcgcgTgcgataccggcgccctGtgGgaatatttgcagcgaaggcgtg

SNP-1 A cgcatttGcgcgYgcgataccggcgccctKtgGgaatatttgcagcgaaggcgtg
 SNP-1 B cgcatttAcgcgYgcgataccggcgccctKtgAgaatatttgcagcgaaggcgtg
 IWGSC-1A cgcatttGcgcgYgcgataccggcgccctGtgGgaatatttgcagcgaaggcgtg
 IWGSC-1B cgcatttAcgcgYgcgataccggcgccctTtgGgaatatttgcagcgaaggcgtg
 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

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

(Krasileva et al., submitted 2016)

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 successfully to design genome-specific primers in several projects.

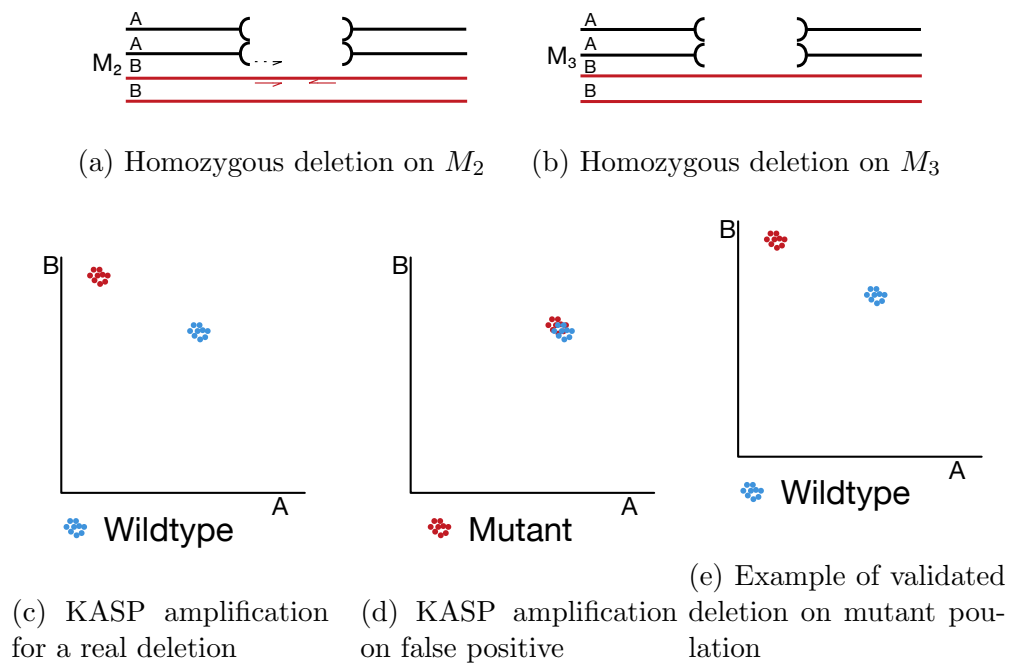


Figure 4.3: PolyMarker used to find primers to detect long deletions in tetraploid wheat.

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

A.1 Validation of mutations on M_4 on Kronos

Line	Pos	WT	Mut	Predicted	Called on M_4	Primer 1 (Kronos)	Primer 2 (mutant)	Common Primer
IWGSC_CSS.1_AS.scaff.3284790	7449	G	A	Het	Het	cacaccttgagctcgcG	ccacacttgagctcgcT	gtgatgttccagggagA
IWGSC_CSS.1_LBL.scaff.3897513	1515	C	T	Het	Het	gettcactGggtctcG	gcttcaactGggtcgcT	acAaggaactgttcagaGaC
IWGSC_CSS.2_AS.scaff.6434745	3424	C	T	Het	Het	aggcatttgcgaattcctcG	ctcCGgtttgcaattttctatgT	ggGtaaTgcagAACcttgagTG
IWGSC_CSS.3_AS.scaff.3408995	732	C	T	Het	Het	gttcattgtctaccacgG	gttcattgtctaccacgT	gttaacattcgtgctctgagcaC
IWGSC_CSS.3_LBL.scaff.10708748	2675	G	A	Het	Het	caccgttgagtgacctcG	caccgttgagtgacctcT	aCcGctaGaagaagcttC
IWGSC_CSS.4_AS.scaff.7132733	1799	C	T	Hom	Hom	cagcttctggcctcATC	cagcttctggcctcAT	gttCctcagAGtaTGagAG
IWGSC_CSS.5_AS.scaff.1534693	4605	C	T	Het	Het	taagaagaagcacttcaactcC	taagaagaagcacttcaactcT	catagagtgctgcatctcctaA
IWGSC_CSS.6_AS.scaff.4361911	8857	C	A	Het	Het	caatgttactgttggttG	caatgttactgttggttG	ggaagcattggCaaagtG
IWGSC_CSS.6_BS.scaff.3008326	1528	G	A	Het	Het	ctgtactcttctgggaagG	ctgtactcttctgggaagG	cttcttggtcagctgtataagacT
IWGSC_CSS.7_AS.scaff.4214385	27835	C	T	Hom	Hom	tttggccataactgacatC	tttggccataactgacatT	atttgcctcagcttcttgcaG
IWGSC_CSS.1_AS.scaff.3929964	1336	C	T	Het	Het	actctacTggcagcagT	actctacTggcagcagT	caacgttggtgcccatGtA
IWGSC_CSS.1_LBL.scaff.3899789	7925	C	T	Het	Het	gaaActgcgcagctCgG	gaaActgcgcagctCgT	ccaGtaGctgtagaA
IWGSC_CSS.2_AS.scaff.6426728	1481	G	A	Hom	Hom	gcattcatctctagcgcG	gcattcatctctagcgcT	acatgcaattgtcagtgactG
IWGSC_CSS.2_LBL.scaff.7960273	690	C	T	Hom	Hom	cgtgtgggtttgttggG	cgtgtgggtttgttggG	gaaaggaacgtTcaTgaG
IWGSC_CSS.3_AS.scaff.32860603	3191	2975	G	A	Het*	ccgttcagatagagccGC	ccgttcagatagagccGT	cgccacatgcacatctcG
IWGSC_CSS.5_AS.scaff.2694249	2399	C	T	Het	Het	gttgattgcctgagcttG	gttgattgcctgagcttG	ttgtggcttcttggagC
IWGSC_CSS.5_LBL.scaff.10923577	3713	C	T	Het	Het	gfgattgcctgagcttG	gfgattgcctgagcttG	ttcgagaagcgcctcG
IWGSC_CSS.6_AS.scaff.5823017	13225	C	T	Hom	Hom	cccttcgagcctctggaG	cccttcgagcctctggaA	gatactcgTgcaatgggT
IWGSC_CSS.6_BS.scaff.2955394	1622	C	T	Het*	Het	gfgagatgaaggtctgcaagG	gfgagatgaaggtctgcaaaA	CgagtgactTcaatttccG
IWGSC_CSS.7_AS.scaff.6739382	12261	G	A	Hom	Hom	gagacaagcttgaattgtctcC	gagacaagcttgaattgtctcT	agatgacaaactcagcaT
IWGSC_CSS.1_LBL.scaff.3276389	9720	C	T	Hom	Hom	aCaGgaagcaAatgtctC	aCaGgaagcaAatgtctT	gggggtgtCtgggaaggC
IWGSC_CSS.2_AS.scaff.6367515	6976	G	A	Het	Het	caggttcagTgtctcgcG	caggttcagTgtctcgcA	acgcAagtaagcgttaC
IWGSC_CSS.2_LBL.scaff.6422019	4523	G	A	Het	Het	cgttaggtctcctcagatG	cgttaggtctcctcagatG	ttgtcAagcagcagcaG
IWGSC_CSS.3_AS.scaff.4284850	7901	C	T	Hom	Hom	tgcttttgacaacatcgG	tgcttttgacaacatcgA	accatgccagctgttcaA
IWGSC_CSS.4_AS.scaff.5962359	13049	G	A	Het	Hom	ccatcaagaagtaagtttgaC	ccatcaagaagtaagtttgaT	ggagaacagctactcggcT
IWGSC_CSS.6_AS.scaff.5778773	6853	G	A	Het	Het	gagtgactctcctcttctC	gagtgactctcctcttctT	ggAagcaaatgtgaacaaA
IWGSC_CSS.6_BS.scaff.4392100	3434	C	T	Het	Het	aggtgtctctcacttcaagG	aggtgtctctcacttcaaaA	caggtgtcttctgagctC
IWGSC_CSS.7_LBL.scaff.6744240	9772	C	A	Het	Het	cagccttctatgcagtaaaG	cagccttctatgcagtaaaA	AgaacacagcagcgcaA
IWGSC_CSS.1_AS.scaff.3887185	9708	C	T	Hom	Hom	caactctgggtctagttgG	caactctgggtctagttgA	tgAgaattctgaCcaaaagC
IWGSC_CSS.2_LBL.scaff.3881362	5160	C	T	Het*	Het	ctgttcacgggctctagC	ctgttcacgggctctagT	cagcactgagagacatggagC
IWGSC_CSS.3_AS.scaff.3296605	6154	C	A	Het	Het	ctaggcttggaacaaaggC	ctaggcttggaacaaaggT	agcttgcactatggggcatT
IWGSC_CSS.3_BS.scaff.10693516	12632	C	T	Het	Het	gCtaacacttcaaatcgcC	gCtaacacttcaaatcgcT	gagcgttctgaagtgtcatC
IWGSC_CSS.5_AS.scaff.1547699	2686	G	A	Het	Het	agagcttcaaccactctC	agagcttcaaccactctT	acgCacattAAtagctgaagC
IWGSC_CSS.5_LBL.scaff.10856077	3413	5853	C	A	Het	cagctTtccgactcttataG	cagctTtccgactcttataA	AgaagtgatgacagtgattcaG
IWGSC_CSS.6_LBL.scaff.5750718	11046	G	A	Hom	Hom	GaTgtctcCtgaagctgG	GaTgtctcCtgaagctgA	caacttgacaacgtcttgG
IWGSC_CSS.7_AS.scaff.4433177	3511	C	T	Het	Het	gttgtctgctggcagaC	gttgtctgctggcagaT	catttgcacgtgtgtcTG
IWGSC_CSS.7_LBL.scaff.6742567	667	C	T	Het	Het	ggTggaggtctggcGatG	ggTggaggtctggcGatA	cagtcactctacagtaggtaG
IWGSC_CSS.1_AS.scaff.3970389	10941	C	T	Hom	Hom	cagatctgaagcttaGcaatA	cagatctgaagcttaGcaatA	actacaggaactcagcaaaaAC
IWGSC_CSS.1_LBL.scaff.3873362	1392	G	C	T	Het	gcaagtgaagtgtaccgtagC	gcaagtgaagtgtaccgtagT	gccaagtaggagaagaactT
IWGSC_CSS.2_LBL.scaff.7882382	2721	C	T	Het	Het	agaacccaacaaactcTacttaG	agaacccaacaaactcTacttaA	gtagGtCcatcCtgaagcttG
IWGSC_CSS.3_AS.scaff.4242376	2410	C	T	Het	Het	gcttgagcaactcactcaatG	gcttgagcaactcactcaatA	gcaatttctcttaTccgcagT
IWGSC_CSS.3_LBL.scaff.10485067	3349	C	T	Hom	Hom	agCaggtctggcaagttG	agCaggtctggcaagttA	agcaatGtaTgaGtgaagcT
IWGSC_CSS.4_AS.scaff.5984133	6006	C	A	Het	Het	tgaatctatgtgtgagctgG	tgaatctatgtgtgagctgA	agcactatctgaacaTcataC
IWGSC_CSS.4_LBL.scaff.7019402	9081	C	T	Het	Het	tggttcccGaaactgaaC	tggttcccGaaactgaaT	cgcatactgaaacaTgagcAC
IWGSC_CSS.5_LBL.scaff.10842786	3304	G	A	Het	Het	anggacccaagccaactctcG	anggacccaagccaactctcA	agtgatcaagccaactgtcgcaC
IWGSC_CSS.6_BS.scaff.3045205	2293	C	T	Het	Het	cAggtctcagatgtggcG	cAggtctcagatgtggcT	cCtTgcaacctctctgatT
IWGSC_CSS.7_AS.scaff.4555249	4487	C	T	Het	Het	tgtcatgcccaagaagA	tgtcatgcccaagaagA	tgggcgaactgtgttaagtG
IWGSC_CSS.1_LBL.scaff.3918498	6096	G	A	Het	Het	ctttatcaggaaagagacacC	ctttatcaggaaagagacacT	caccattgtagggttctctTtnc
IWGSC_CSS.2_BS.scaff.5131713	5900	C	A	Het	Het	tgCagTgtgggaacaggG	tgCagTgtgggaacaggA	catgagtGagatctctcgcT
IWGSC_CSS.5_AS.scaff.2769540	9626	C	T	Het	Het	gccaaggAaccataactcG	gccaaggAaccataactcT	GgaactctggcAaccggA
IWGSC_CSS.5_LBL.scaff.10871091	7062	G	A	Het	Het	cgaacaggtgttggagCgC	cgaacaggtgttggagCgT	tcagatgctgcaagattcctC
IWGSC_CSS.6_AS.scaff.5800333	2360	G	A	Het	Het	gGtgGgtattTgtgttggagG	gGtgGgtattTgtgttggagA	tgGtgagctcgacaGtGtA
IWGSC_CSS.7_LBL.scaff.6716931	2613	G	A	Het	Het	tgcttctgtgtctcC	tgcttctgtgtctcT	atTtgcattTCgAtcgggcC
IWGSC_CSS.2_LBL.scaff.8029221	2860	C	T	Het	Het	ctacttgcgatgcagatG	ctacttgcgatgcagatA	agacccagctctttgagC
IWGSC_CSS.3_BS.scaff.10460714	14359	C	T	Hom	Hom	agcagctatcaatcagcC	agcagctatcaatcagcT	actcaagataccaCcgacC
IWGSC_CSS.4_AS.scaff.5989735	6404	G	A	Hom	Hom	taccttcttactgagG	taccttcttactgagA	tttttcagaggaacacaggtatcA
IWGSC_CSS.5_LBL.scaff.7648030	6893	C	T	Het	Het	atcagtaagtgtctacCgT	atcagtaagtgtctacCgT	accttgatgtaCaatccaC
IWGSC_CSS.6_BS.scaff.5755840	778	C	T	Het	Het	gaagcaatgtcActgtttgG	gaagcaatgtcActgtttgA	gcttggaactggcatttatG
IWGSC_CSS.6_AS.scaff.29792151	4346	G	A	Hom	Hom	gaagggtA ccgataacC	gaagggtA ccgataacT	catctgccGgttaaacatgC
IWGSC_CSS.7_AS.scaff.4542983	18700	C	A	Het	Het	gCgatatggacttgaatgaG	gCgatatggacttgaatgaA	ttactgtcttataGTtgcGg
IWGSC_CSS.7_BS.scaff.3098098	4346	C	T	Het	Het	gtggcgaacaccttctC	gtggcgaacaccttctT	gtcttctttaaggggcgC
IWGSC_CSS.1_AS.scaff.3259804	219	C	T	Het	Het	gccctctcaacCtcttcagC	gccctctcaacCtcttcagT	ttcagagctCgaggaaattccC
IWGSC_CSS.2_AS.scaff.6315418	10490	G	A	Hom	Hom	TggcagcaactcgaC	TggcagcaactcgaA	ttgacagatgagTgaaggAaaT
IWGSC_CSS.2_BS.scaff.5181092	3742	G	A	Het	Het	gtactgaagtgtcCtCgG	gtactgaagtgtcCtCgA	cttcaactcttggggtTtC
IWGSC_CSS.3_LBL.scaff.10425015	2372	C	T	Het	Het	caaagggtctcagagagG	caaagggtctcagagagA	agacctcagatGtctcC
IWGSC_CSS.3_AS.scaff.10775915	4701	C	T	Het	Het	taaccTggcatcggccG	taaccTggcatcggccA	catgtGcagacaTgacT
IWGSC_CSS.5_AS.scaff.2754304	2301	G	A	Het	Het	gagccttcttgggaagG	gagccttcttgggaagA	tttgttcccgaaacatgtaC
IWGSC_CSS.5_LBL.scaff.10919959	1867	C	T	Hom	Hom	anggccttgggttttC	anggccttgggttttT	anggaatggaACagctaagatcaT
IWGSC_CSS.7_AS.scaff.4245431	3402	G	A	Hom	Hom	gatacAgctgctctgagG	gatacAgctgctctgagA	ttctctgtcaatttgatccC
IWGSC_CSS.7_LBL.scaff.6667357	641	C	T	Het	Het			

A.2 Validation of mutations on M_4 on Cadenza

IWGSC contig	Line	Pos	WT	Mut	Predicted	Called on M_4	Primer 1 (Cadenza)	Primer 2 (mutant)	Common Primer
IWGSC.CSS.3B.scaff.10445294	Cadenza1772	6019	C	T	het	het	caggatAgtGggactgtcaaaG	caggatAgtGggactgtcaaaA	sgagacGcctGiggacAT
IWGSC.CSS.3DL.scaff.6955403	Cadenza1772	2418	C	T	het*	het	tcagCggatgtccggatA	tcagCggatgtccggatA	tgtcCaaGaATctgtccacG
IWGSC.CSS.4AL.scaff.7106846	Cadenza1772	11277	G	A	het	hom	tggatccatcctcaactG	tggatccatcctcaactA	gatgtGgatttccgtcA
IWGSC.CSS.4AS.scaff.5991335	Cadenza1772	15710	G	A	het	hom	ctggcctcgtcgtcA	ctggcctcgtcgtcA	gtgaaGcttcaagaagacaG
IWGSC.CSS.4BS.scaff.4956646	Cadenza1772	252	G	A	het*	hom	ggagcttgacttccgaG	ggagcttgacttccgaA	tGagtaagaGcTaaagAaaG
IWGSC.CSS.4DS.scaff.1715982	Cadenza1772	1225	G	A	hom	hom	cagctgtggTactcaactG	cagctgtggTactcaactG	CcCtGaaACAcGgtttgaT
IWGSC.CSS.5AL.scaff.2763407	Cadenza1772	2119	G	A	hom	hom	ggacGaacctcgagatcG	ggacGaacctcgagatcA	gatTggcaATctgtCgtgcA
IWGSC.CSS.5AS.scaff.1548786	Cadenza1772	12625	C	T	het	het	AtaggcaactgtgactgaG	AtaggcaactgtgactgaA	ggattgggtgtgcaacG
IWGSC.CSS.5BL.scaff.18449226	Cadenza1772	2289	C	T	het*	het	ctcagacatcattgtcacatG	ctcagagatgttccacatT	cactcgaggtgtccacatT
IWGSC.CSS.5BS.scaff.2270737	Cadenza1772	2262	G	A	het	—	atfcCTgtgtgtggCaaatgaG	atfcCTgtgtgtggCaaatgaA	taaGcaacaAccctcagctG
IWGSC.CSS.1AL.scaff.3022915	Cadenza1661	891	C	T	hom	hom	ccacagtgagactcctattgaOG	ccacagtgagactcctattgaCA	atgtctgattcGtcGtattcC
IWGSC.CSS.1AS.scaff.3297240	Cadenza1661	1970	C	T	het	het	catccgccGttttcctC	catccgccGttttcctC	gctcgcgatgaagaagcT
IWGSC.CSS.1BL.scaff.3828996	Cadenza1661	1340	G	A	hom	hom	agccggatgttagtttaacT	agccggatgttagtttaacT	agcagctgtTcggttaaC
IWGSC.CSS.1DS.scaff.1884529	Cadenza1661	10575	G	A	hom	hom	aCagatacaAttgtcatcaggT	aCagatacaAttgtcatcaggT	acctgggTtTgtccaatcC
IWGSC.CSS.2AL.scaff.1831870	Cadenza1661	19142	C	T	het	—	cgtggcCgaatCtCgaC	cgtggcCgaatCtCgaC	tcttgtggagccggcG
IWGSC.CSS.2AS.scaff.5213460	Cadenza1661	1358	G	A	hom	hom	gtcacgaCCgcctcagA	gtcacgaCCgcctcagA	aggaaagagagaanaaGcG
IWGSC.CSS.2BS.scaff.5179331	Cadenza1661	5604	C	A	het	het	actctcgtcaagaactgatacaG	actctcgtcaagaactgatacaA	gcaGaaatgttttgtcaacT
IWGSC.CSS.2DS.scaff.5341235	Cadenza1661	4673	G	A	het	het	ggtaggactctcgagctA	ggtaggactctcgagctA	ggcggtcgtacagtgT
IWGSC.CSS.3AL.scaff.4250995	Cadenza1661	7046	C	A	hom	hom	cCaagaacacgggtgtccaA	cCaagaacacgggtgtccaA	ctcagctgtcccatcattG
IWGSC.CSS.3B.scaff.10404421	Cadenza1661	4303	G	A	het	het	ctctcgtcgaCaggacctA	ctctcgtcgaCaggacctA	GCcagctCacATgctcC
IWGSC.CSS.5DL.scaff.2390496	Cadenza1538	2125	C	T	hom	hom	ggacttttactctcagtagcttgG	ggacttttactctcagtagcttgA	tctcagaaTgttaatgtcGatG
IWGSC.CSS.6AL.scaff.5753680	Cadenza1538	3920	C	T	hom	hom	tgctcnaatttgacacaaTaaC	tgctcnaatttgacacaaTaaT	aaatcgaagggttaagttttgT
IWGSC.CSS.6AS.scaff.4425792	Cadenza1538	4307	C	A	hom	het	agatgctgtCggGccaA	agatgctgtCggGccaA	gctgaagcaacgcgatacT
IWGSC.CSS.6BS.scaff.3003630	Cadenza1538	6933	C	T	het	het	ggcagtaaatgttggctgagT	ggcagtaaatgttggctgagT	tTgaCttcgtgtttgtggcA
IWGSC.CSS.6DL.scaff.3246988	Cadenza1538	9186	G	A	het	het	gctaaagaagccttgagagaatTC	gctaaagaagccttgagagaatTC	aatcttgaagagaggtttgtatG
IWGSC.CSS.7AL.scaff.4480114	Cadenza1538	3446	C	T	het	—	gatctcccaacggcG	gatctcccaacggcA	tggacactcttcagttT
IWGSC.CSS.7AS.scaff.1193541	Cadenza1538	8359	C	T	hom	het	agcaattctttgtctatcaatagC	agcaattctttgtctatcaatagT	tcatcTgttactctactcG
IWGSC.CSS.7BL.scaff.6721572	Cadenza1538	9223	C	T	het	het	gctCaggaggaagaacaaG	gctCaggaggaagaacaaA	tgtatgaagaattcgaacC
IWGSC.CSS.7BS.scaff.3152545	Cadenza1538	3960	C	A	hom	—	tcagcaaatctaccgcCgC	tcagcaaatctaccgcCgT	gCtgcctcatcgtttaT
IWGSC.CSS.7DS.scaff.3963838	Cadenza1538	2913	G	A	het	het	tCgttgcaagcCttTtvtgT	tCgttgcaagcCttTtvtgT	agaGttaTcaagCtactgtcacA
IWGSC.CSS.1AL.scaff.3903380	Cadenza1469	6193	G	A	het	hom	ctcttcAGagatgaacgcG	ctcttcAGagatgaacgcA	tcGtGagagGtggtttGTtA
IWGSC.CSS.1AS.scaff.3287728	Cadenza1469	3817	C	T	het*	hom	ccgaccaAttcaataacgG	ccgaccaAttcaataacgA	acctcttcccAacatgaT
IWGSC.CSS.1BL.scaff.3815304	Cadenza1469	513	G	A	hom	hom	aacattgtcctTaCaaaacGC	aacattgtcctTaCaaaacGT	acacagaagttataatGCAAgC
IWGSC.CSS.1DL.scaff.2266648	Cadenza1469	5926	C	T	het	het	caacatgagacacacactTC	caacatgagacacacactTT	gtacaagcgtgagatgtcG
IWGSC.CSS.1DS.scaff.1906671	Cadenza1469	3697	C	T	hom	hom	tggTGTgaacacttggcgaG	tggTGTgaacacttggcgaA	catggcgacacCAcctG
IWGSC.CSS.2BL.scaff.6337088	Cadenza1469	7334	G	A	het*	hom	acaatgcAGgttgacaggttA	acaatgcAGgttgacaggttA	sggagttgtgttGagaacaT
IWGSC.CSS.2BL.scaff.7972799	Cadenza1469	8995	C	T	het	hom	gTgCtctcGtcatcctT	gTgCtctcGtcatcctT	gatcgcGcaactaagTG
IWGSC.CSS.2DL.scaff.9832343	Cadenza1469	3262	G	A	het	het	TtgtctaAcagcacCGagA	TtgtctaAcagcacCGagA	agatcgtgtcagcttctC
IWGSC.CSS.2DS.scaff.5327939	Cadenza1469	3889	C	A	het	het	ttttTgccttatgtactctagtaC	ttttTgccttatgtactctagtaT	gaggccactcacagctcG
IWGSC.CSS.3B.scaff.10395219	Cadenza1469	1292	C	A	hom	—	agsgtctgtgtcgtcG	agsgtctgtgtcgtcGA	cccttctggsggctttataC
IWGSC.CSS.3B.scaff.10592217	Cadenza0580	2994	C	T	het	—	acagcagtaataaagccctC	acagcagtaataaagccctT	tgaacactgtgTggCggagG
IWGSC.CSS.3DS.scaff.2596771	Cadenza0580	1037	G	A	het	het	tggttatgTAcaggataatCagG	tggttatgTAcaggataatCagA	tgctcaaatgtgtcctataggT
IWGSC.CSS.4AL.scaff.7093953	Cadenza0580	9881	C	T	hom	hom	CacaggaagccggttaacaT	CacaggaagccggttaacaT	ctctAGcagcattgggaT
IWGSC.CSS.4BL.scaff.7037448	Cadenza0580	1837	C	T	hom	hom	CgttgaanaaGctcgaagaacttaaC	CgttgaanaaGctcgaagaacttaaT	cagtctctcTtCaGagcagataT
IWGSC.CSS.4BS.scaff.4929479	Cadenza0580	10668	G	A	hom	—	tggattttccgcactgtcC	tggattttccgcactgtT	gtaaacaagcattcnaagatcA
IWGSC.CSS.4DL.scaff.14359838	Cadenza0580	1408	G	A	hom	—	gCtCAitcaggatTGTcCtaTatG	gCtCAitcaggatTGTcCtaTatA	tgaCagaacagtgtgtcactAT
IWGSC.CSS.4DS.scaff.2276484	Cadenza0580	8034	G	A	hom	hom	gcccgtgtgtgagAG	gcccgtgtgtgagAGaA	cgtccagattactgactactcA
IWGSC.CSS.5AL.scaff.2756579	Cadenza0580	5278	G	A	het	het	tgaatgatttttgcctccgttC	tgaatgatttttgcctccgtT	ggAAtCCTATtGCAgaAaCTG
IWGSC.CSS.5BL.scaff.10787208	Cadenza0580	10627	G	A	het	—	gcctctcaatcgtcgagcA	gcctctcaatcgtcgagcA	acagatgcAGgtcgGcgT
IWGSC.CSS.5BS.scaff.2282179	Cadenza0580	5267	G	A	het	—	tgaatggtctacagctgC	tgaatggtctacagctgT	tggcgccctgaaAAtcC
IWGSC.CSS.5DL.scaff.4498073	Cadenza0423	4937	C	T	hom	hom	gaacctctgtgtgtcattC	gaacctctgtgtgtcattT	tgagacgaAagacgcG
IWGSC.CSS.5DS.scaff.2738970	Cadenza0423	2319	C	T	het	—	cgtgaggtgggtgattgC	cgtgaggtgggtgattgT	tggaaactagttacactcagTTC
IWGSC.CSS.6AL.scaff.3757109	Cadenza0423	2788	G	A	hom	hom	caggGccttgccaaataaaG	caggGccttgccaaataaaGA	ctttcGcagctctatgttctcG
IWGSC.CSS.6AS.scaff.4387871	Cadenza0423	2543	G	A	hom	hom	gcaatgaacaggcgaaagG	gcaatgaacaggcgaaagG	ctcatgctcctgatctaaggT
IWGSC.CSS.6BL.scaff.4271391	Cadenza0423	4660	C	T	hom	hom	tacgtgcagtgtgtgtagtgtaT	tacgtgcagtgtgtgtagtgtaT	ggttgaagtgcactagatGtaccA
IWGSC.CSS.6DS.scaff.11880206	Cadenza0423	9159	G	A	het	het	ctgCgaagctccacaA	ctgCgaagctccacaA	ggatgagaagtttgcatgctC
IWGSC.CSS.7AS.scaff.4227506	Cadenza0423	952	G	A	het	—	ccatgtttccaalgttagcG	ccatgtttccaalgttagcT	tggcctagctgtatgcT
IWGSC.CSS.7BS.scaff.6681782	Cadenza0423	1486	C	T	hom	hom	agtaagCGtgacacnaatggG	agtaagCGtgacacnaatggA	AgctcTtgGtgaagatcacA
IWGSC.CSS.7DS.scaff.3160328	Cadenza0423	7801	C	T	het	het	tgttaaatGataacCctgcagC	tgttaaatGataacCctgcagT	tggaaaggtgTggtgtttT
IWGSC.CSS.7DL.scaff.407428	Cadenza0423	2051	G	A	het	het	gtcGCgcaactctgacaA	gtcGCgcaactctgacaA	atcacAAGgtcagcccaA
IWGSC.CSS.3AL.scaff.442479	Cadenza0364	3198	C	T	het	het	gagtcATaaagtttgtaagattgG	gagtcATaaagtttgtaagattgT	GCAaGATaaCaacagatcacG
IWGSC.CSS.3AS.scaff.4447942	Cadenza0364	11917	G	A	het	het	gtcataaagatgctcctgtgaaG	gtcataaagatgctcctgtgaaA	ctcGgattgtggggaagA
IWGSC.CSS.3AS.scaff.1557483	Cadenza0364	2547	C	T	het	het	aaagtacacatcattctacataaG	aaagtacacatcattctacataaA	cgaataccaacgcctcacA
IWGSC.CSS.3AS.scaff.2648747	Cadenza0364	2688	G	A	het	het	tggAagCAaaggggccC	tggAagCAaaggggccC	GcgcgcagtaggacatcG
IWGSC.CSS.3AS.scaff.3304956	Cadenza0364	1017	G	A	het	het	gtcccttgacacagctttG	gtcccttgacacagctttG	aatcgtggaactaaccttcaA
IWGSC.CSS.3AS.scaff.3321091	Cadenza0364	4585	C	T	het	het	caagaatATgtcgtatgtggaG	caagaatATgtcgtatgtggaA	catcgtggaatccgcgaatC
IWGSC.CSS.3AS.scaff.3371333	Cadenza0364	538	G	A	het	het	gggaaaCgAgAcgagcG	gggaaaCgAgAcgagcA	cgtgtcctcttcacccT
IWGSC.CSS.3AS.scaff.3371815	Cadenza0364	1061	C	T	het	het	atccccgcgacagagG	atccccgcgacagagA	aAttgccctctgtgattcC
IWGSC.CSS.3AS.scaff.3440912	Cadenza0364	4498	G	A	het	het	cgttaaaactttctgtgttG	cgttaaaactttctgtgttG	atActgacaaactacatgattgC
IWGSC.CSS.3B.scaff.10343586	Cadenza0364	2242	G	A	het	—	ggttcTgTtctctcttccactG	ggttcTgTtctctcttccactA	tgtgtgaccccgcaagcA

IWGSC contig	Line	Pos	WT	Mut	Predicted	Called on M_4	Primer 1 (Cadenza)	Primer 2 (mutant)	Common Primer
IWGS.C.SS.3.AL.scaff.442479	Cadenza0364	3198	C	T	het	het	gagctATaagttgtaagattggC	gagctATaagttgtaagattggT	GCaGaTaaCaacagatcacG
IWGS.C.SS.3.AL.scaff.4447942	Cadenza0364	11917	C	A	het	het	gtcataaagattgctctgtgaaG	gtcataaagattgctctgtgaaA	CtGgatgtggagagangA
IWGS.C.SS.3.AS.scaff.1557483	Cadenza0364	2547	C	T	het	het	aaagtcacatcatgcttaccataaA	aaagtcacatcatgcttaccataaA	cgaatccaacgctcatcA
IWGS.C.SS.3.AS.scaff.2648747	Cadenza0364	2688	G	A	het	het	tggAagcAcaaggggccT	tggAagcAcaaggggccT	GccggcagtgagactcG
IWGS.C.SS.3.AS.scaff.3304956	Cadenza0364	1017	G	A	het	het	gtccctgacacacgtttG	gtccctgacacacgtttG	CtctgagactaacattcaaT
IWGS.C.SS.3.AS.scaff.3321091	Cadenza0364	4585	C	T	het	het	caagaatgATgctgatgttggaG	caagaatgATgctgatgttggaA	acatgctgaatcgcgaaTC
IWGS.C.SS.3.AS.scaff.3371333	Cadenza0364	538	G	A	het	het	gggaaacGgAgAagagcgA	gggaaacGgAgAagagcgA	ccgtgctctctacacT
IWGS.C.SS.3.AS.scaff.3371815	Cadenza0364	1061	C	T	het	het	atccccaggagcagagG	atccccaggagcagagG	aAttgcccttggtgatcC
IWGS.C.SS.3.AS.scaff.3440912	Cadenza0364	4498	G	A	het	het	ccgtaaaactttctgtctgtT	ccgtaaaactttctgtctgtT	atAcgtaacaaacatgatgtgC
IWGS.C.SS.3.B.scaff.10343586	Cadenza0364	2242	G	A	het	—	gggttcTgTctctcttccactA	gggttcTgTctctcttccactA	tggttgaaocgeaagcA
IWGS.C.SS.5.DL.scaff.452342	Cadenza0281	2433	C	T	het	—	catggCgacggtGtctcA	catggCgacggtGtctcA	aAccctcATTtggCTACTvCT
IWGS.C.SS.5.DL.scaff.4538822	Cadenza0281	1208	G	A	hom	—	acgtcagaacacgcttgaC	acgtcagaacacgcttgaT	ttaaatgtgtggccacC
IWGS.C.SS.6.AL.scaff.5813297	Cadenza0281	4532	C	T	het	—	ggagaggagcgtcagG	ggagaggagcgtcagG	ttctctgcaacgattccG
IWGS.C.SS.6.AS.scaff.4378990	Cadenza0281	6748	C	T	hom	hom	ccagggttctgtctctttcT	ccagggttctgtctctttcT	caagatatacaagaatgaaggTgT
IWGS.C.SS.6.BL.scaff.4360781	Cadenza0281	5426	C	T	het	het	aCtactcaaatggctTggttaaA	aCtactcaaatggctTggttaaA	tcagtcacaagTcaagagT
IWGS.C.SS.7.AL.scaff.4488310	Cadenza0281	3808	G	A	hom	hom	gttctctgttagcagccA	gttctctgttagcagccA	ggcgtttcttccgctA
IWGS.C.SS.7.BL.scaff.6696509	Cadenza0281	9232	G	A	het	het	gctctaggGtggtgcaaAagA	gctctaggGtggtgcaaAagA	ggcttCaGtgcGagT
IWGS.C.SS.7.BS.scaff.3143575	Cadenza0281	1866	C	T	het	het	agaigttagagaggcgtT	agaigttagagaggcgtT	gttggAigtggcnaagT
IWGS.C.SS.7.DL.scaff.3346250	Cadenza0281	1663	G	A	het	het	acgtgcagacaactctaaT	acgtgcagacaactctaaT	TtccaccaggcccaagA
IWGS.C.SS.7.DS.scaff.393917	Cadenza0281	1243	C	T	het	het	tgCtgagcCttTcaacttgT	tgCtgagcCttTcaacttgT	agaggttgggttccatcGG
IWGS.C.SS.3.B.scaff.10626860	Cadenza0148	7847	G	A	het	het	ggagcgtctgggaaggagG	ggagcgtctgggaaggagG	gttaagtacTctagctcG
IWGS.C.SS.3.DL.scaff.6915683	Cadenza0148	6904	C	T	het	het	cgtcaacCgtgtggcaattA	cgtcaacCgtgtggcaattA	tcatgctcaatgTcataggT
IWGS.C.SS.4.AS.scaff.5929057	Cadenza0148	4238	G	A	hom	hom	ggcacaactagCactacT	ggcacaactagCactacT	tatctggTgaagtgcaggttCA
IWGS.C.SS.4.AS.scaff.5950625	Cadenza0148	10590	C	T	het	het	agaTattCaatTcgggtgAttggC	agaTattCaatTcgggtgAttggT	ccgtCtccctcagtcC
IWGS.C.SS.4.AS.scaff.5967119	Cadenza0148	11626	C	T	hom	hom	cgtGgaaccccgagctG	cgtGgaaccccgagctA	gaagcagcactgcagcG
IWGS.C.SS.4.DL.scaff.14455742	Cadenza0148	1946	C	T	hom	hom	gcCtgaggagagatcgctG	gcCtgaggagagatcgctG	aacggCAaCTGtGcGcA
IWGS.C.SS.3.B.scaff.2318993	Cadenza0148	4000	C	T	hom	hom	tcaggtttgacacagatgaaggA	tcaggtttgacacagatgaaggA	tgagaTctgtttctttctacAttG
IWGS.C.SS.5.AL.scaff.2750707	Cadenza0148	4603	G	A	het	het	ctgtgtgtcagcatttcaagTaG	ctgtgtgtcagcatttcaagTaA	ccaggaTgcAgtgcaatttcaaG
IWGS.C.SS.5.BL.scaff.10794137	Cadenza0148	9235	C	T	hom	hom	gaagcgtctctcgttA	gaagcgtctctcgttA	agatctcttcataaagcagtG
IWGS.C.SS.5.BS.scaff.1646558	Cadenza0148	2916	C	T	het	het	gccGtaactcaactAaccttA	gccGtaactcaactAaccttA	gcaaTgtccacttAtcacctC
IWGS.C.SS.1.AL.scaff.3883106	Cadenza0110	27536	C	T	het	het	accttcaactcagcttgG	accttcaactcagcttgA	gtgaagaacacaggttgaagC
IWGS.C.SS.1.BL.scaff.3812829	Cadenza0110	10770	G	A	het*	hom	ccccactcattccagA	ccccactcattccagA	gGagtgtgtctgctggaA
IWGS.C.SS.1.DL.scaff.2266648	Cadenza0110	6156	G	A	het	het	actgctggttatgggacT	actgctggttatgggacT	ccccactgaagaacacA
IWGS.C.SS.1.DS.scaff.1889435	Cadenza0110	8826	C	T	hom	hom	aacatgaattactcagcagA	aacatgaattactcagcagA	gcccgaagaattgtatcaaaagG
IWGS.C.SS.2.AS.scaff.5268634	Cadenza0110	4636	G	A	het	het	gatccatgtgattggcagtttA	gatccatgtgattggcagtttA	TgctgtTggaatgagttgT
IWGS.C.SS.2.BL.scaff.7965110	Cadenza0110	15801	C	T	hom	hom	catigaagcAtaacAattgcAaaC	catigaagcAtaacAattgcAaaT	gccaagtatccagaatgaagTttA
IWGS.C.SS.2.DL.scaff.9852812	Cadenza0110	13788	C	A	hom	hom	attttgtatggtctcaattcggC	attttgtatggtctcaattcggT	gaactTcaatttctactgtcT
IWGS.C.SS.2.DS.scaff.5371379	Cadenza0110	2166	C	T	hom	hom	agacaaacaaactgGatggcT	agacaaacaaactgGatggcT	gcgtcgagaaagttTgtatttG
IWGS.C.SS.3.AL.scaff.4384278	Cadenza0110	1276	C	T	het	het	agcTgaactgccccTgtaG	agcTgaactgccccTgtaA	aggagactCgGgtggaAa
IWGS.C.SS.3.AS.scaff.3340122	Cadenza0110	1467	C	T	hom	hom	atttctAGtgtgttgggaactA	atttctAGtgtgttgggaactA	gagaagactagaagtTttAgcaT
IWGS.C.SS.5.DL.scaff.4554222	Cadenza2103	6528	C	T	het*	hom	gtcgtctcaagaagaacaaattG	gtcgtctcaagaagaacaaattA	aTcccaactCGaTttgtcataC
IWGS.C.SS.6.AL.scaff.5833640	Cadenza2103	7346	C	T	hom	hom	aagaagaagcacaattgttctC	aagaagaagcacaattgttctC	aCTctTcaagtgttccagC
IWGS.C.SS.6.AS.scaff.4429974	Cadenza2103	3867	C	A	hom	hom	GagatgaAttatttgcacgtggC	GagatgaAttatttgcacgtggT	gggtccggtgcataagT
IWGS.C.SS.6.DL.scaff.3307626	Cadenza2103	4970	C	T	hom	hom	tgcaagattgtctgtgtaG	tgcaagattgtctgtgtaA	cttaggaagtgtttgtactGtC
IWGS.C.SS.6.DS.scaff.2059604	Cadenza2103	5224	G	A	het	—	gtccaatgcagtcTgagtgA	gtccaatgcagtcTgagtgA	gtccaagtattattttctcctcG
IWGS.C.SS.7.AL.scaff.4552322	Cadenza2103	1412	C	T	het	het	gcaaaagcTgatactccaacA	gcaaaagcTgatactccaacA	ggcAAGccAgtataaagtaaGC
IWGS.C.SS.7.BS.scaff.3147455	Cadenza2103	4607	G	A	het	het	gcaacttaggatgtgagTtaagT	gcaacttaggatgtgagTtaagT	gcaatgaggtttatttgaactgtA
IWGS.C.SS.4.AL.scaff.7176064	Cadenza2103	3473	C	T	hom	—	GGTtctgCaGTTCATAActcatC	GGTtctgCaGTTCATAActcatT	attgaatcaactgatacGaaGactC
IWGS.C.SS.3.B.scaff.10457010	Cadenza0277	10599	G	A	het	het	aaccttggccgagaacacT	aaccttggccgagaacacT	acgtgtcagcagagggG
IWGS.C.SS.3.B.scaff.10593852	Cadenza0277	10124	C	T	het	het	tgacaggagcgtatacagA	tgacaggagcgtatacagA	gtctaaCTTAActtAccatcagC
IWGS.C.SS.3.DS.scaff.2583390	Cadenza0277	663	G	A	hom	hom	actgcaactatacaatActtCtG	actgcaactatacaatActtCtG	tcCacttggagcagagT
IWGS.C.SS.4.AL.scaff.3973013	Cadenza0548	11765	C	T	hom	hom	ctctgtattcaatggaTgtTtttgA	ctctgtattcaatggaTgtTtttgA	ttcccaaaTaaanaaggaagC
IWGS.C.SS.1.BS.scaff.317205	Cadenza0548	373	C	T	het	het	gtggtgaggaGGgtgCaaG	gtggtgaggaGGgtgCaaG	atgttcgaggtgaggtggG
IWGS.C.SS.2.AS.scaff.5305619	Cadenza0548	2786	C	T	hom	hom	atacagatgcctAAgtgtTtT	atacagatgcctAAgtgtTtT	ggagaacAAtGctccaggtA
IWGS.C.SS.2.BS.scaff.5306489	Cadenza0548	46953	T	G	het	wt	aggttcaigtccatagaagGT	aggttcaigtccatagaagGG	aggctaTAgactcctgtACAgT
IWGS.C.SS.2.BL.scaff.7984123	Cadenza0548	11660	G	A	het	het	catgtggcatagtaactcagtaA	catgtggcatagtaactcagtaA	aatataatggaggaacaaagcC
IWGS.C.SS.2.DL.scaff.9907477	Cadenza0548	1363	C	T	hom	hom	tgctccttttgcagaaC	tgctccttttgcagaaC	ggcaaacctgagttggcatC
IWGS.C.SS.2.DS.scaff.5307477	Cadenza0548	5449	G	A	hom	hom	gcatgtccattatactgaaCgtG	gcatgtccattatactgaaCgtA	catgtcgtctctcttctggacC
IWGS.C.SS.3.AL.scaff.4449951	Cadenza0548	633	C	T	het	het	tccaaactcaactcagcttaacactA	tccaaactcaactcagcttaacactA	gtctgagTCAagtgtcG
IWGS.C.SS.3.B.scaff.10479889	Cadenza0097	3339	C	T	hom	—	ttgTtctCGagaagatgcCA	ttgTtctCGagaagatgcCA	ggtgctatccaAcGgcA
IWGS.C.SS.3.B.scaff.10562262	Cadenza0097	7819	C	T	het	het	agaggggtgctatcactAttgG	agaggggtgctatcactAttgA	agcgatgccaaaggtctcC
IWGS.C.SS.4.AL.scaff.7040796	Cadenza0097	10772	G	A	hom	hom	acacaacttgcaccagaaA	acacaacttgcaccagaaA	CAatCGatgtgtgtctTctcC
IWGS.C.SS.4.AL.scaff.7063488	Cadenza0097	6360	C	T	het	het	gcctctcaCttAatttgaagctgT	gcctctcaCttAatttgaagctgT	aggctgtgagttgtgaagttT
IWGS.C.SS.4.AL.scaff.7091701	Cadenza0097	5050	G	A	het	het	catgacatctggaggagaaatA	catgacatctggaggagaaatA	agcaaggaATAaatgaaggaaA
IWGS.C.SS.4.DS.scaff.1845841	Cadenza0097	7110	G	A	hom	hom	aatgTAgctccccatacCgG	aatgTAgctccccatacCgA	aciganaacTgcaatcgtTaatggA
IWGS.C.SS.5.AL.scaff.2767581	Cadenza0097	3737	G	A	het	het	gagaggttctactAtcggT	gagaggttctactAtcggT	cgTcaacaaatattcgtggG
IWGS.C.SS.5.BL.scaff.10784643	Cadenza0097	1568	C	T	hom	hom	agaaaTAAatggatggatggaaCA	agaaaTAAatggatggatggaaCA	catctcCcttcaCgGaaagG

IWGS contig	Line	Pos	WT	Mut	Predicted	Called on M_4	Primer 1 (CadENZA)	Primer 2 (mutant)	Common Primer
IWGS.C.SS.1AL.scaff.3952258	Cadenza2092	8107	C	T	het	—	tgatgagaaatgacagctgtgG	ttgagtagaagaatgacagctgtgA	tgccacattgacatagagaG
IWGS.C.SS.1BL.scaff.3858008	Cadenza2092	10278	G	A	hom	hom	tttgagcaggcagctgcG	tttgagcaggcagctgcG	actcagcgtatcActatC
IWGS.C.SS.1DL.scaff.2265172	Cadenza2092	9094	C	T	hom	hom	ttgcaTGTcattgttctatcagC	ttgcaTGTcattgttctatcagT	aggttcacactccGttcatC
IWGS.C.SS.2AL.scaff.6435867	Cadenza2092	16201	G	A	hom	hom	ttctgtTacttaagctcaatigaC	ttctgtTacttaagctcaatigaT	gtgagagatgagagtaagacC
IWGS.C.SS.2AL.scaff.6439430	Cadenza2092	25101	C	T	het	—	caagaagggCagCtCagC	caagaagggCagCtCagT	toGtAAcTcttAcctgtaA
IWGS.C.SS.2DL.scaff.76760848	Cadenza2092	4733	C	T	het	het	gcacatgggtctcaggtaC	gcacatgggtctcaggtaT	tcagtcaattGCTCgtCTG
IWGS.C.SS.3AL.scaff.4407012	Cadenza2092	2785	C	T	hom	hom	acatatAggtttctcatccacatC	acatatAggtttctcatccacatT	acctctcagttaataggtttG
IWGS.C.SS.3AS.scaff.3441108	Cadenza2092	541	G	A	het	het	GtgaagcttgagacGgaG	GtgaagcttgagacGgaA	aggcaTgacaaCgagcaA
IWGS.C.SS.3B.scaff.10449827	Cadenza1551	4779	G	A	hom	hom	ggcaaggtcaagaacGgtC	ggcaaggtcaagaacGgtT	aCagaGtaggttagagcaG
IWGS.C.SS.3B.scaff.1050638	Cadenza1551	3250	C	T	het	het	ctctctactgttggcC	ctctctactgttggcT	gcaacATtTgataAtattticaT
IWGS.C.SS.3DL.scaff.6945816	Cadenza1551	589	C	T	hom	hom	agcatctcaactgcaCaataC	agcatctcaactgcaCaataT	TgtgcccTtTgataAtattticaT
IWGS.C.SS.3DL.scaff.6954177	Cadenza1551	3508	C	T	het	het	tgtagatcaatcaatttctctG	tgtagatcaatcaatttctctA	gcttggtataaacCttagacA
IWGS.C.SS.4AS.scaff.5938272	Cadenza1551	19080	G	A	hom	hom	agAcCcgATcgcatG	agAcCcgATcgcatG	GggAgatAcaggttaaaAcTtccG
IWGS.C.SS.4AS.scaff.5977594	Cadenza1551	11092	C	T	het	het	gcttgatcggaacaaacC	gcttgatcggaacaaacT	ggctctcagctcgcA
IWGS.C.SS.5AL.scaff.2671035	Cadenza1551	5859	C	T	het	het	cagtgatattTtagacttcagcC	cagtgatattTtagacttcagcT	ggcagttcagcGaccatT
IWGS.C.SS.5BL.scaff.10889480	Cadenza1551	2530	G	A	hom	hom	gagcttaactcgcagatggaA	gagcttaactcgcagatggaA	tcctatgCAacGccttggT
IWGS.C.SS.3B.scaff.10528396	Cadenza2088	8059	G	A	hom	—	cttttcgctcgttaagcaataG	cttttcgctcgttaagcaataA	gtgcaactgttcaggcctgA
IWGS.C.SS.3B.scaff.10637573	Cadenza2088	16815	G	A	het	het	agcaagcttaccGgtcgtC	agcaagcttaccGgtcgtT	cgagCAactagcagcagctT
IWGS.C.SS.4AL.scaff.7080469	Cadenza2088	6697	G	A	het	het	ggcgtctactcaacgcA	ggcgtctactcaacgcA	ccaGaggctgtTgcattttT
IWGS.C.SS.4AL.scaff.7126302	Cadenza2088	3627	G	A	hom	hom	gttcaaaaaaagtggtcAatttgC	gttcaaaaaaagtggtcAatttgT	cacaaggtatgaagcTctctagA
IWGS.C.SS.4BL.scaff.7041808	Cadenza2088	10234	G	A	hom	hom	tcaatgagtagaggtgtctT	tcaatgagtagaggtgtctT	ccatagcagcatcagccacA
IWGS.C.SS.5BL.scaff.2794167	Cadenza2088	13162	G	A	het	—	agatctcagcaagcatCttCaG	agatctcagcaagcatCttCaA	caatgaacacttcgaagaaG
IWGS.C.SS.5BL.scaff.10889232	Cadenza2088	3885	C	A	het	het	cTcaacacaaatgggcaAatC	cTcaacacaaatgggcaAatT	tccttcaatcaatcaatgttgG
IWGS.C.SS.5BS.scaff.2267405	Cadenza2088	11113	C	T	hom	hom	cttgaatgctcctaggccttTG	cttgaatgctcctaggccttTA	tgatttggTCgtgtAagtttGA
IWGS.C.SS.3B.scaff.10475354	Cadenza1409	2203	C	A	hom	hom	agCgaacaaagGtcaaacG	agCgaacaaagGtcaaacA	cttcaacacaCragaCAattAccG
IWGS.C.SS.3B.scaff.10674115	Cadenza1409	4555	C	T	het	het	gcttcagtcagctcttcaA	gcttcagtcagctcttcaA	cttcaacccGagataatGtattG
IWGS.C.SS.4AL.scaff.7153568	Cadenza1409	13073	C	T	hom	hom	tcgacgcATcaacttgG	tcgacgcATcaacttgA	gacggaaactctcggcC
IWGS.C.SS.4DL.scaff.14314966	Cadenza1409	2010	G	A	het	hom	gtagtcctcctCaggG	gtagtcctcctCaggA	cggtTcaacAgttgCcT
IWGS.C.SS.4DS.scaff.2324074	Cadenza1409	7606	G	A	het	het	tGagaanaatgtgtGcaGaG	tGagaanaatgtgtGcaGaA	gggtAAgttCAaactGaggtgaaG
IWGS.C.SS.5AS.scaff.1517889	Cadenza1409	3561	G	A	het	het	tcctgacatctccgtgtaC	tcctgacatctccgtgtaT	ggcctgggaacattgcttattA
IWGS.C.SS.5AS.scaff.1523866	Cadenza1409	8054	G	A	hom	—	gggtatctacgcgaGgaT	gggtatctacgcgaGgaT	tcctgagCcTctctcA
IWGS.C.SS.5BL.scaff.10917655	Cadenza1409	19073	G	A	hom	hom	caaatgacatgcaaaagagttgC	caaatgacatgcaaaagagttgT	cgcttctacataAaatgtcT
IWGS.C.SS.1AL.scaff.3886649	Cadenza1599	5204	C	T	het	het	tgatgccacacaaatGcC	tgatgccacacaaatGcT	ggactgactgctgacattattA
IWGS.C.SS.1BL.scaff.3810267	Cadenza1599	6634	C	T	hom	hom	ccCaggaalagacactC	ccCaggaalagacactT	cgacggcgaagatgtaTtG
IWGS.C.SS.1DL.scaff.2291677	Cadenza1599	12856	C	T	hom	hom	GgtagcaagtgcgcaG	GgtagcaagtgcgcaA	ctctctcttaacGCCG
IWGS.C.SS.2AL.scaff.6354492	Cadenza1599	7566	G	A	het	het	gGagaatgaCAgtAacTtctgG	gGagaatgaCAgtAacTtctgA	ttccgaagaacacataTccTG
IWGS.C.SS.2AS.scaff.5282937	Cadenza1599	9736	G	A	het	het	gctgtattttatagctgctatG	gctgtattttatagctgctatT	cacCagaattgttCactgattTC
IWGS.C.SS.2BL.scaff.7952427	Cadenza1599	19249	G	A	hom	hom	cgTccctCcttagcagcaT	cgTccctCcttagcagcaT	aTcactccatttagcagAG
IWGS.C.SS.2DL.scaff.9897981	Cadenza1599	5627	C	T	het	het	cttgggtTgattgcttactC	cttgggtTgattgcttactT	gTtggTcTctgactTtggT
IWGS.C.SS.3AL.scaff.4446105	Cadenza1599	1765	G	A	hom	—	aaatgctttcctaCcgctagtG	aaatgctttcctaCcgctagtA	ttctAgaggcaatagctTatatgcT

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.
- 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.