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 an 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 discuses 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 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 realignments or *de novo* assemblies, how to do SNP calling in diploid and polyploid organisims and the bulk frequency ratios.

2.6 Wheat online resources

A compilation of the currently available resource for whet genetics and genomics. MAS wheat, CeralsDB, 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. diccocoides*).

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_2 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. diccocoides*, 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`(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 alternative 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

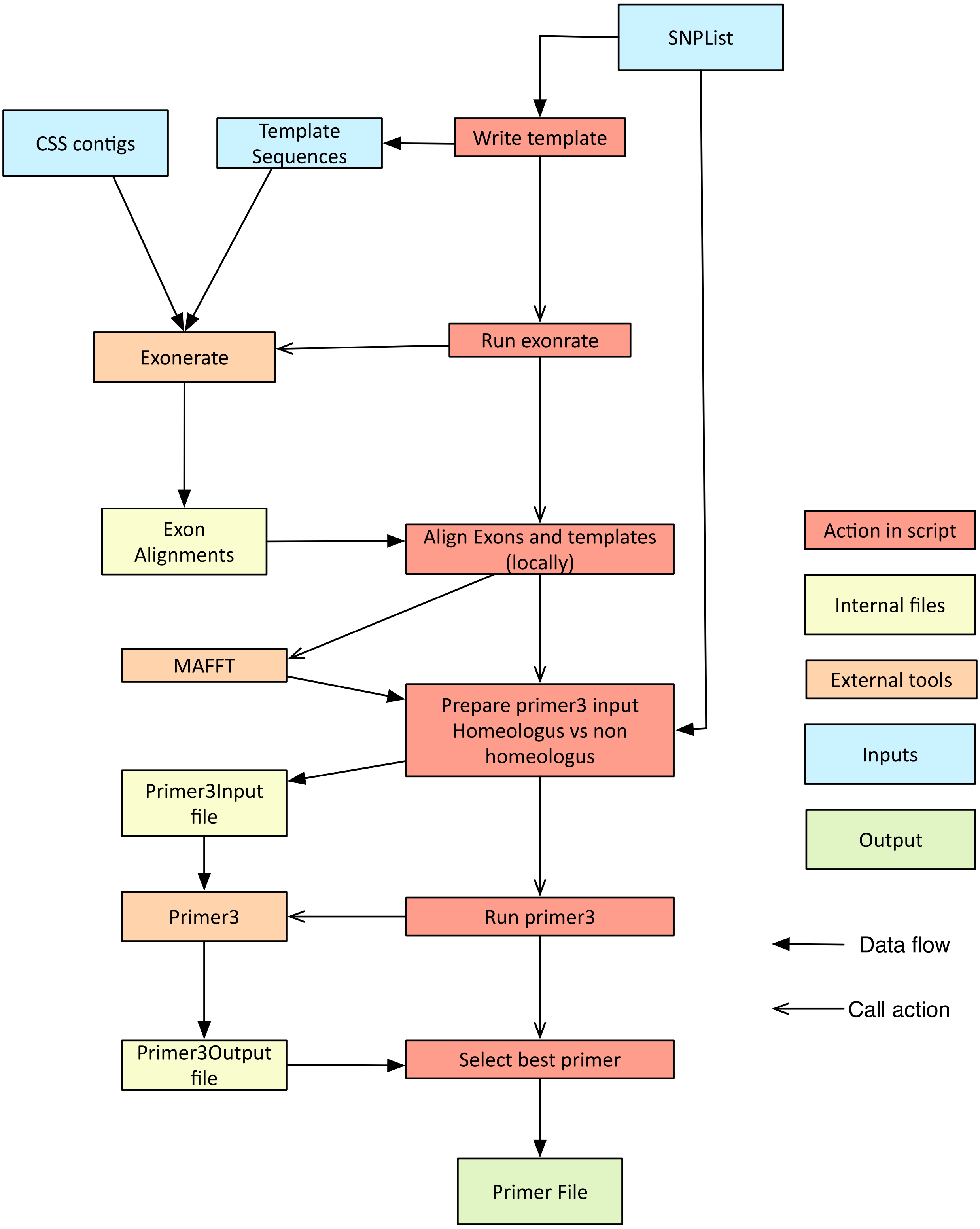


Figure 4.1: Steps and tools called by PolyMarker. The colour of the boxes represent: the step is an action inside the script(red); actions of the script(orange); temporary files(yellow); inputs(blue) and; outpus(green)

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 defualt, 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 comparasion base-per-base, a local alignment with `MAFFT` (Katoh and Standley, 2013) is produced (Figure 4.2e. 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 Primer design tools

In this section, the principles of *in silico* primer design are discussed, and why not simply selecting a genomic variation is enough (thermal stability, primers folding on themselves)

4.3 Primer selection algorithms

Different algorithms to select the best primer.

4.3.1 KASP markers

For KASP markers, the product should be as short as possible with the mutation in the first three bases.

4.4 Designed markers

Details of the generated primers for the 80k iSelect chip and the 820k axiom chip. This section also include counts on how many are genome specific, semi-specific and non specific. Also an analysis of how many are repeated or map to more than one chromosome perfectly.

4.4.1 Regular markers

PolyMarker was designed for KASP assays, but it was later extended to produce regular primers, where both primers start with a genome-specific base. This simplifies the design of primers for regular PCR and capillary sequencing.

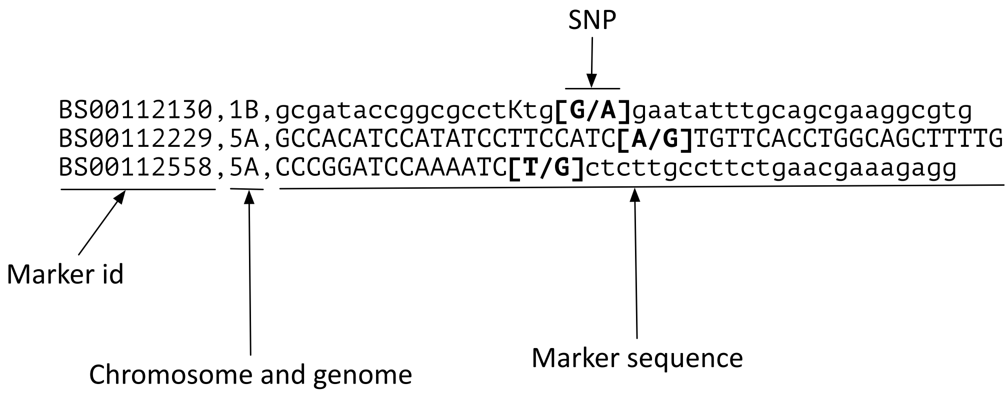
4.4.2 Deletion algorithms

Algorithm to produce KASP for deletions in polyploids.

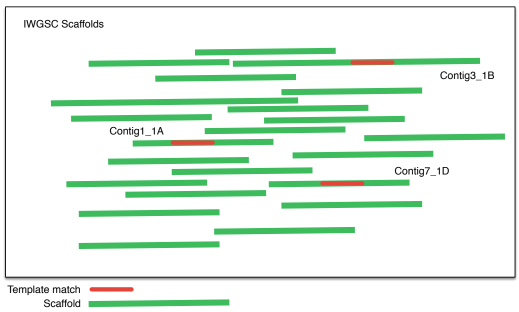
4.5 Conclusions

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)

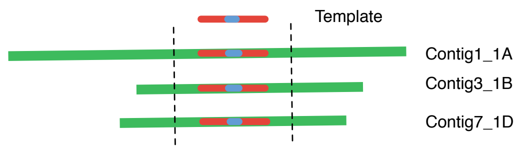
PolyMarker has been used succesfully to design genome-specific primers in several projects.



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



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



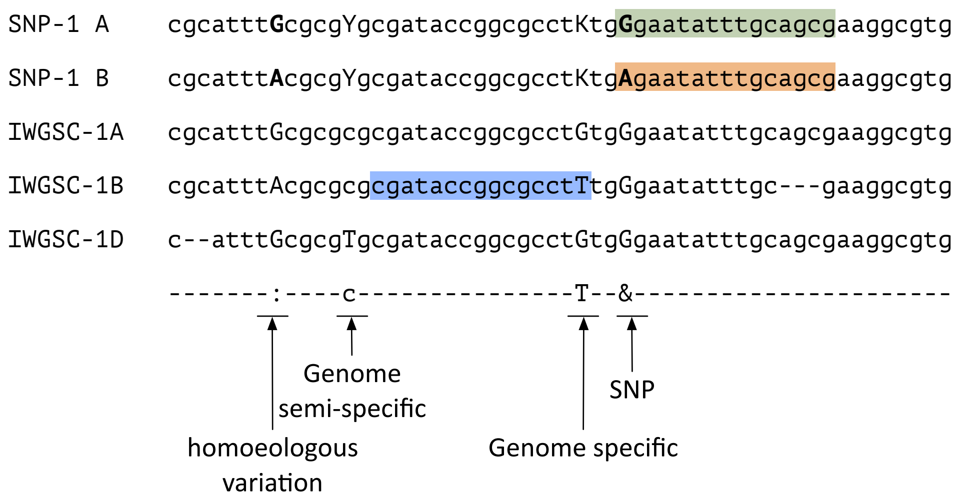
(c) Selected regions around the SNP on every chromosome.

```
SNP-1 A    cgcatttGcgcgYgcgataccggcgcctKtgGgaatatttgcagcgaaggcgtg
SNP-1 B    cgcatttAcgcgYgcgataccggcgcctKtgAgaatatttgcagcgaaggcgtg
IWGSC-1A   cgcatttgcgcgcgcgataccggcgcctgtgggaatatttgcagcgaaggcgtg
IWGSC-1B   cgcatttacgcgcgcgataccggcgcctttgggaatatttgcgaaggcgtg
IWGSC-1D   catttgcgcgTgcgataccggcgcctgtgggaatatttgcagcgaaggcgtg
```

(d) Sequence of found regions around the SNP.

```
SNP-1 A    cgcatttGcgcgYgcgataccggcgcctKtgGgaatatttgcagcgaaggcgtg
SNP-1 B    cgcatttAcgcgYgcgataccggcgcctKtgAgaatatttgcagcgaaggcgtg
IWGSC-1A   cgcatttGcgcgcgcgataccggcgcctGtgGgaatatttgcagcgaaggcgtg
IWGSC-1B   cgcatttAcgcgcgcgataccggcgcctTtgGgaatatttgc---gaaggcgtg
IWGSC-1D   c--atttGcgcgTgcgataccggcgcctGtgGgaatatttgcagcgaaggcgtg
```

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

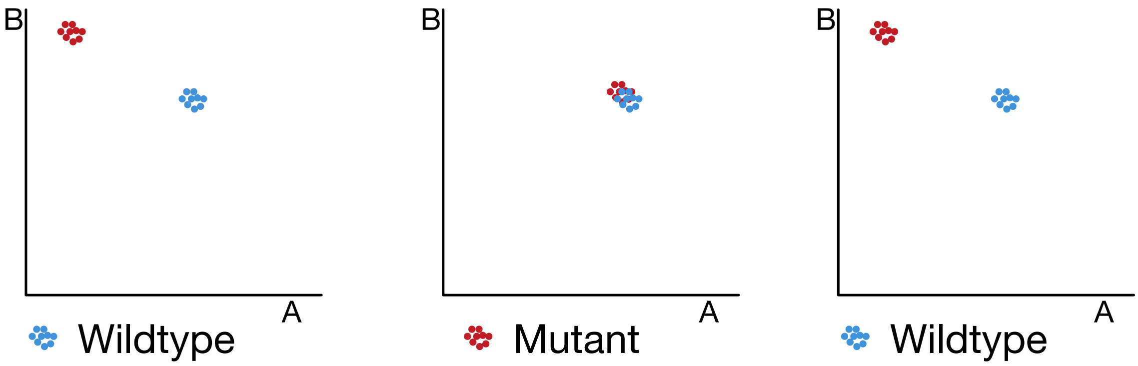


(f) Alignment with mask and primer candidates.

Figure 4.2: Alignments done by PolyMarker.

(a) Homozygous deletion on M_2

(b) Homozygous deletion on M_3



(c) KASP amplification for a real deletion

(d) KASP amplification on false positive

(e) Example of validated deletion on mutant poulation

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

PolyMarker validation

A.1 Validation of mutations on M_4 on Kronos

IWGSC contig	Line	Pos	WT	Mut	Predicted	Called on M_4	Primer 1 (Kronos)	Primer 2 (mutant)	Common Primer
IWGSC_CSS_1AS_scaff_3284790	Kronos3085	7449	G	A	Het	Het	ccacaccttgagcctcgC	ccacaccttgagcctcgT	gtgattttgccaggggagA
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	aggccatttcgaattccgC	aggccatttcgaattccgT	ggTgttaTccagAacctgagTG
IWGSC_CSS_3B_scaff_10708748	Kronos3085	2675	G	A	Het	Het	gttgcatgcttcacccagG	gttgcatgcttcacccagA	gtaacaatctgagttcgtagcaC
IWGSC_CSS_4AL_scaff_7132733	Kronos3085	1799	C	T	Hom	Hom	cacccgtgagtgaccctC	cacccgtgagtgaccctT	aCcGcctaGaaagaaagcttC
IWGSC_CSS_5AS_scaff_1534693	Kronos3085	4605	C	T	Het	Het	cagcttcctggccctcAtC	cagcttcctggccctcAtT	gtaCctcacgAgtcaTgagAG
IWGSC_CSS_6AS_scaff_4361911	Kronos3085	8857	G	A	Het	Het	tcacgaaagacgacttcaacctcC	tcacgaaagacgacttcaacctcT	catgaggtgctgcatctccatcA
IWGSC_CSS_6BS_scaff_3008326	Kronos3085	1528	G	A	Het	Het	ccatgttgtactggtggtgC	ccatgttgtactggtggtgT	ggaagcatggCaagtgcA
IWGSC_CSS_7AS_scaff_4214385	Kronos3085	27835	C	T	Hom	Hom	cgtaccttcgttgggaaagG	cgtaccttcgttgggaaagA	ctcttggtcagctgtataagacT
IWGSC_CSS_1AL_scaff_3929964	Kronos3191	1336	C	T	Het	Het	tttcggccatacctgacatC	tttcggccatacctgacatT	attgcctccagttcttgcaG
IWGSC_CSS_1BL_scaff_3899789	Kronos3191	7925	C	T	Het	Het	actctcacTggcagcagC	actctcacTggcagcagT	caacgtggtgcccatcGtA
IWGSC_CSS_2AL_scaff_6426728	Kronos3191	1481	G	A	Hom	Hom	gaaActgccgcagctCgC	gaaActgccgcagctCgT	ccaGcaGctcgtgagaaA
IWGSC_CSS_2BL_scaff_7960273	Kronos3191	690	C	T	Hom	Hom	gccattcatccttaggcgC	gccattcatccttaggcgT	acatgcaattgctgatgactG
IWGSC_CSS_3AS_scaff_3286603	Kronos3191	2975	G	A	Het*	Hom	ccgtgtggtttgttgtggG	ccgtgtggtttgttgtggA	gaaaggaacgtgTcaTgcaG
IWGSC_CSS_5AL_scaff_2694249	Kronos3191	2399	C	T	Het	Het	gccttccagatagagccGC	gccttccagatagagccGT	cgccacatcgacattcctG
IWGSC_CSS_5BL_scaff_10923577	Kronos3191	3713	C	T	Het	Het	gtggattgcctgagcttgC	gtggattgcctgagcttgT	tggtggccttcttgggaC
IWGSC_CSS_6AL_scaff_5823017	Kronos3191	13225	C	T	Hom	Hom	ccctttcgagcctctggaG	ccctttcgagcctctggaA	ttcgagaaggcccatcgA
IWGSC_CSS_6BS_scaff_2955394	Kronos3191	1622	C	T	Het*	Hom	gtggagatgaaggtctagcaaG	gtggagatgaaggtctagcaaA	gatactcgTgcaatgggtgT
IWGSC_CSS_7BL_scaff_6739382	Kronos3191	12261	G	A	Hom	Hom	gagacaagctttgaattgctcC	gagacaagctttgaattgctcT	CgagtgacctTcatttcccG
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	ggggtgatCtggaagggC
IWGSC_CSS_2AL_scaff_6422019	Kronos3288	4523	G	A	Het	Het	cgctaggtccctgcatagG	cgctaggtccctgcatagA	acgcAcgctaagccgtaC
IWGSC_CSS_3AL_scaff_4284850	Kronos3288	7901	C	T	Hom	Hom	tggctttggacaacatcgG	tggctttggacaacatcgA	tgtcAgcatcgacagccaG
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	gagtgaccttcccgtctttC	gagtgaccttcccgtctttT	ggagaacagctactcggcT
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	agctgttcttctcctacttcaaG	agctgttcttctcctacttcaaA	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	caacttctgggctgtagtgtG	caacttctgggctgtagtgtA	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	ctaggcttggacaaacaggC	ctaggcttggacaaacaggT	agcttgcatctatgggcatT
IWGSC_CSS_5AS_scaff_1547699	Kronos3413	2686	G	A	Het	Het	gCtacaaccttcaccaatcgC	gCtacaaccttcaccaatcgT	gacggctttgaagtgtcatC
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	cacgcTtcccgacttcttataG	cacgcTtcccgacttcttataA	AgacgatgtgatcaggattcaG
IWGSC_CSS_7AL_scaff_4433177	Kronos3413	3511	C	T	Het	Het	GaTgctccGtcaggctgG	GaTgctccGtcaggctgA	cactactggacaagctcttgG
IWGSC_CSS_7BL_scaff_6742567	Kronos3413	667	C	T	Het	Het	gttgcttgcgtggcagaC	gttgcttgcgtggcagaT	cattttgcaccgtgtgtcTG
IWGSC_CSS_1AL_scaff_3976389	Kronos3935	10941	C	T	Hom	Hom	ggtgaggagatcggCgatG	ggtgaggagatcggCgatA	cagtcatctacatgagaggtcaG
IWGSC_CSS_1BL_scaff_3873362	Kronos3935	1392	G	A	Het	Het	cagatctgaagcctaGcacatG	cagatctgaagcctaGcacatA	actaccagaatcagcacaaaaAC
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	agaacccaaaacccgTacttaG	agaacccaaaacccgTacttaA	gtagGgtCcatcCtaaagcttG
IWGSC_CSS_3B_scaff_10485067	Kronos3935	3349	C	T	Hom	Hom	gcttgagcaactactccaactG	gcttgagcaactactccaactA	gcaatttcctttaTccgcagT
IWGSC_CSS_4AS_scaff_5984153	Kronos3935	6006	G	A	Het	Het	agCaggtctggccaagttG	agCaggtctggccaagttA	cgaatGtatgaGtaggcgcT
IWGSC_CSS_4BL_scaff_7019402	Kronos3935	9081	C	T	Het	Het	tgcaatcatgtagtgagctgG	tgcaatcatgtagtgagctgA	agcatgatccctagaaCcataC
IWGSC_CSS_5BL_scaff_10842786	Kronos3935	3304	G	A	Het	Het	tggttcccGaagcctgaaC	tggttcccGaagcctgaaT	cgcatacttgaaacaTGagcAC
IWGSC_CSS_6BS_scaff_3045205	Kronos3935	2293	G	A	Het	Het	aaggaccaagcccaaactctcG	aaggaccaagcccaaactctcA	agtgatcaagcccaatgtcgcA
IWGSC_CSS_7AL_scaff_4555249	Kronos3935	4487	C	T	Het	Het	cAgtgctcgagatggcgC	cAgtgctcgagatggcgT	cCttgcaaccctcctgatT
IWGSC_CSS_1BL_scaff_3918498	Kronos4240	6096	G	A	Het	Het	ttgcatgccccaagaagaG	ttgcatgccccaagaagaA	tgggcgaactggtaatgtgG
IWGSC_CSS_2BS_scaff_5131713	Kronos4240	5900	G	A	Het	Het	cctttatcgaggaaagagacacC	cctttatcgaggaaagagacacT	caccattgtagggttcctt TttC
IWGSC_CSS_5AL_scaff_2769540	Kronos4240	9626	C	T	Het	Het	tgCagtgtgggaaacggaG	tgCagtgtgggaaacggaA	catgagtGagatcttcctgcT
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	cgacaggattgtgagCgC	cgacaggattgtgagCgT	tcagatgctgcaagattcatcT
IWGSC_CSS_7BL_scaff_6716931	Kronos4240	2613	G	A	Het	Het	gGtgGgtattTgcttggtgaG	gGtgGgtattTgcttggtgaA	tgGtggactcgacaGtGtA
IWGSC_CSS_2BL_scaff_8029221	Kronos4346	2860	G	A	Het	Het	tgcttccgctcttgctcC	tgcttccgctcttgctcT	atTtgcatTCgAtcgggcC
IWGSC_CSS_3B_scaff_10460714	Kronos4346	14359	C	T	Hom	Hom	ctaccttgccatgcgacatG	ctaccttgccatgcgacatA	agcaccccagtctttgacG
IWGSC_CSS_4AS_scaff_5989735	Kronos4346	6404	G	A	Hom	Hom	acgcatgctaacatcagcC	acgcatgctaacatcagcT	actcaagataccaCcgcacG
IWGSC_CSS_5BL_scaff_7648030	Kronos4346	6893	C	T	Het	Het	taccctttcctactggcagG	taccctttcctactggcagA	ttttcagaggaacacaggtatcA
IWGSC_CSS_6AL_scaff_5755840	Kronos4346	778	C	T	Het	Het	atcgagtaagctgtcacCgC	atcgagtaagctgtcacCgT	acctgcatgtcaCatccaC
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	ttacattgcttataGTttgCcgG
IWGSC_CSS_1AS_scaff_3259804	Kronos4485	219	C	T	Het	Het	gtcggcacaacccccttgC	gtcggcacaacccccttgT	gcttctttaaggagggcgA
IWGSC_CSS_2AL_scaff_6315418	Kronos4485	10490	G	A	Hom	Hom	gcccctctcaaCcttctcagC	gcccctctcaaCcttctcagT	ttcagacgctCgaggaatttccC
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	cttcacatccttgggggTtC
IWGSC_CSS_3B_scaff_10775915	Kronos4485	4701	C	T	Het	Het	ccaagggctgcagagagG	ccaagggctgcagagagA	agacctcacgatGtcctcC
IWGSC_CSS_5AL_scaff_2754304	Kronos4485	2301	G	A	Het	Het	taacccTgccatcgcccG	taacccTgccatcgcccA	cattgGccagccaTgacT
IWGSC_CSS_5BL_scaff_10919959	Kronos4485	1867	C	T	Hom	Hom	gatgccctttgtggagaagG	gatgccctttgtggagaagA	tcttgttcccgaaacatgtcA
IWGSC_CSS_7AS_scaff_4245431	Kronos4485	3402	G	A	Hom	Hom	aaggcgcctggtgtttcC	aaggcgcctggtgtttcT	agtaagtggaAcagctaagatcaT
IWGSC_CSS_7BL_scaff_6667357	Kronos4485	641	C	T	Het	Het	gatcAgctgctcattcgagG	gatcAgctgctcattcgagA	ttccctgtcaattgatgccC

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	ggagacGGctGtggacatT
IWGSC_CSS_3DL_scaff_6955403	Cadenza1772	2418	C	T	het*	hom	tcagCggattgtcgggatG	tcagCggattgtcgggatA	tgtcCatgaaTcttgtccacG
IWGSC_CSS_4AL_scaff_7106846	Cadenza1772	11277	G	A	hom	hom	tgggatccatgcctacactG	tgggatccatgcctacactA	gatggtGgatttgccgctA
IWGSC_CSS_4AS_scaff_5991335	Cadenza1772	15710	G	A	hom	hom	ctggccctgcgctgctaC	ctggccctgcgctgctaT	gtggaaGttcagaaggaccaG
IWGSC_CSS_4BS_scaff_4956646	Cadenza1772	252	G	A	het*	hom	gcaggttgacttcccggaG	gcaggttgacttcccggaA	tGaggtacgaGcTaaagAaagC
IWGSC_CSS_4DS_scaff_1715962	Cadenza1772	1225	G	A	hom	hom	cagctgtggTatctcaactgG	cagctgtggTatctcaactgA	CcCtGaaACACcGtttggaT
IWGSC_CSS_5AL_scaff_2763407	Cadenza1772	2119	G	A	hom	hom	gcgacGaacctcgagatctG	gcgacGaacctcgagatctA	gaTggcaAtcgtCgtgcA
IWGSC_CSS_5AS_scaff_1548786	Cadenza1772	12625	C	T	het	het	AtaggcacattgctagactgaG	AtaggcacattgctagactgaA	ggattgggtgttgcacgC
IWGSC_CSS_5BL_scaff_10849226	Cadenza1772	2289	C	T	het*	hom	cctgacatcattgttcacgatC	cctgacatcattgttcacgatT	cactccgaggtgtccatgaT
IWGSC_CSS_5BS_scaff_2270737	Cadenza1772	2262	G	A	hom	—	attcCTgtgttgtggCaaatgaG	attcCTgtgttgtggCaaatgaA	taaGcacaaAccctccagctgG
IWGSC_CSS_1AL_scaff_3022915	Cadenza1661	891	C	T	hom	hom	ccacagtgagactcctattgaCG	ccacagtgagactcctattgaCA	atgtctgattcGtcGtagtcC
IWGSC_CSS_1AS_scaff_3297240	Cadenza1661	1970	C	T	het	het	catcccgccGtttcctcC	catcccgccGtttcctcT	gctcgccgatgaagagcT
IWGSC_CSS_1BL_scaff_3828996	Cadenza1661	1340	G	A	hom	hom	agccggatgttagtgttaacC	agccggatgttagtgttaacT	agcagcttgTcgcgttaaC
IWGSC_CSS_1DS_scaff_1884529	Cadenza1661	10575	G	A	hom	hom	aCagatacaAttgtcatgcaggC	aCagatacaAttgtcatgcaggT	acctgggTTgtccaatacttC
IWGSC_CSS_2AL_scaff_6318370	Cadenza1661	19142	C	T	het	—	cgtggcCgaatCtcGacG	cgtggcCgaatCtcGacA	ttcttgtgggagccgggC
IWGSC_CSS_2AS_scaff_5213460	Cadenza1661	1358	G	A	hom	hom	gtcacgaaCccgctcagG	gtcacgaaCccgctcagA	aggaaagagaggaaaagaGcG
IWGSC_CSS_2BS_scaff_5179331	Cadenza1661	5604	G	A	het	het	actctcgtcaagaactgatacaG	actctcgtcaagaactgatacaA	gcaGagaatgttcttgcaacT
IWGSC_CSS_2DS_scaff_5341235	Cadenza1661	4673	G	A	het	het	ggtgaggatctcggagctG	ggtgaggatctcggagctA	gcgcggtcgtacgagttG
IWGSC_CSS_3AL_scaff_4250995	Cadenza1661	7046	G	A	hom	hom	cCaagaaacgggtggtccaG	cCaagaaacgggtggtccaA	ctgcagctgtcccatcatcgT
IWGSC_CSS_3B_scaff_10404421	Cadenza1661	4303	G	A	het	het	ccttcgtcgaCaggacctG	ccttcgtcgaCaggacctA	GCcagtactCacAtgctctC
IWGSC_CSS_5DL_scaff_2390496	Cadenza1538	2125	C	T	hom	het	gcagttttatcctcagtagtcttgG	gcagttttatcctcagtagtcttgA	ttctgagaaTgtaatgtgcGatG
IWGSC_CSS_6AL_scaff_5753680	Cadenza1538	3920	C	T	hom	hom	tgctccaaatttgagcacaaTaaC	tgctccaaatttgagcacaaTaaT	aaatgcaaggggtaagtttttgT
IWGSC_CSS_6AS_scaff_4425792	Cadenza1538	4307	G	A	hom	het	agatgcttgtCggGccaG	agatgcttgtCggGccaA	gctgaagcaacgcgatcaaT
IWGSC_CSS_6BS_scaff_3003630	Cadenza1538	6933	C	T	het	het	ggcagtaatgtggtgctgagC	ggcagtaatgtggtgctgagT	tTgaCttctggtttggtggcA
IWGSC_CSS_6DL_scaff_3246988	Cadenza1538	9186	G	A	het	het	gctaaagaagagcttgagagaattC	gctaaagaagagcttgagagaattT	aatttctgaagagaggtgttgtatG
IWGSC_CSS_7AL_scaff_4480114	Cadenza1538	3446	C	T	het	—	gatatctcccacacggcgG	gatatctcccacacggcgA	tgagccactcttgcagtttT
IWGSC_CSS_7AS_scaff_4193541	Cadenza1538	8359	C	T	hom	het	agcaattctttggctatcaattagC	agcaattctttggctatcaattagT	tcatctGtcttaactctactgctG
IWGSC_CSS_7BL_scaff_6721572	Cadenza1538	9223	C	T	het	het	gctCagggaggaagacaagaaG	gctCagggaggaagacaagaaA	tgctatgaagaattccgacctC
IWGSC_CSS_7BS_scaff_3152545	Cadenza1538	3960	G	A	hom	—	tcagcaaaatcacctgcCgC	tcagcaaaatcacctgcCgT	gCtgccccatcatcgtttaT
IWGSC_CSS_7DS_scaff_3963838	Cadenza1538	2913	G	A	het	het	tCgttgcaagcCttTtgtgC	tCgttgcaagcCttTtgtgT	agaGttaTcaagcTactgtcacA
IWGSC_CSS_1AL_scaff_3903380	Cadenza1469	6193	G	A	hom	hom	ctcttcAgagatgaacgcgG	ctcttcAgagatgaacgcgA	tcGtGagatgGtggtttGTtA
IWGSC_CSS_1AS_scaff_3287728	Cadenza1469	3817	C	T	het*	hom	ccgaccaAttcactaaccgG	ccgaccaAttcactaaccgA	accctctttcccAgacatgaT
IWGSC_CSS_1BL_scaff_3815304	Cadenza1469	513	G	A	hom	hom	aacatttgcctTaCcaaaacGC	aacatttgcctTaCcaaaacGT	acacagcaagttataatgCAAgC
IWGSC_CSS_1DL_scaff_2266648	Cadenza1469	5926	C	T	het	het	caacatgagacacaacaccttC	caacatgagacacaacaccttT	gtcaacgcgtgaggattgtC
IWGSC_CSS_1DS_scaff_1906671	Cadenza1469	3697	C	T	hom	hom	tggTGtagacacttggcgaG	tggTGtagacacttggcgaA	catggcgaccaccAcctG
IWGSC_CSS_2AL_scaff_6337088	Cadenza1469	7334	G	A	het*	hom	acaatgccAagttgacaggttG	acaatgccAagttgacaggttA	gggagtgttggttCagaacaT
IWGSC_CSS_2BL_scaff_7972799	Cadenza1469	8995	C	T	het	hom	gTgCtcctcGgcatccttC	gTgCtcctcGgcatccttT	gatccgGgcaaactacgTG
IWGSC_CSS_2DL_scaff_9832343	Cadenza1469	3262	G	A	het	het	TtgtctaAcagcacCGcagG	TtgtctaAcagcacCGcagA	agatctcggtcagcctttcT
IWGSC_CSS_2DS_scaff_5327939	Cadenza1469	3889	G	A	het	het	ttttTgccttatgtgactctagtaC	ttttTgccttatgtgactctagtaT	gaggccatcacagatagcG
IWGSC_CSS_3B_scaff_10395219	Cadenza1469	1292	G	A	hom	—	aggtgcttgtgcttgctgG	aggtgcttgtgcttgctgA	cctcttctgggggctttataC
IWGSC_CSS_3B_scaff_10592217	Cadenza0580	2994	C	T	het	—	acagcagtatcaagcccctC	acagcagtatcaagcccctT	tgatactgttgTggCggagG
IWGSC_CSS_3DS_scaff_2596771	Cadenza0580	1037	G	A	het	het	tggttatgCAcaggataatCagG	tggttatgCAcaggataatCagA	tggcaaatgtgatgtcattaggT
IWGSC_CSS_4AL_scaff_7093953	Cadenza0580	9881	C	T	hom	hom	GacaggaagccggtaacaC	GacaggaagccggtaacaT	ctccAgcaggcatgggaT
IWGSC_CSS_4BL_scaff_7037448	Cadenza0580	1837	C	T	hom	hom	CgttgaaaaGctgcaagaacttaaC	CgttgaaaaGctgcaagaacttaaT	cagttcttccTtCaGagcagataT
IWGSC_CSS_4BS_scaff_4929479	Cadenza0580	10668	G	A	hom	—	tggattttcccgcactgttC	tggattttcccgcactgttT	gtaaacaaggcatttcaagagtcA
IWGSC_CSS_4DL_scaff_14359838	Cadenza0580	1408	G	A	hom	—	gCtcAttcagggatTGTcCtaTatG	gCtcAttcagggatTGTcCtaTatA	tgaCagaacagttggtcatacT
IWGSC_CSS_4DS_scaff_2276484	Cadenza0580	8034	G	A	hom	hom	gccgtggttgatggAgaG	gccgtggttgatggAgaA	cgtccagattactgatacttgcA
IWGSC_CSS_5AL_scaff_2756579	Cadenza0580	5278	G	A	het	het	tgaatggatttttcgtcccgttC	tgaatggatttttcgtcccgttT	ggAAtCCTATgCAgaAgAaaCTG
IWGSC_CSS_5BL_scaff_10787208	Cadenza0580	10627	G	A	het	—	gcctctcacatgcggagaC	gcctctcacatgcggagaT	acgatgtcAggtggGcgT
IWGSC_CSS_5BS_scaff_2282179	Cadenza0580	5267	G	A	het	—	tgatgggctacgacgtgC	tgatgggctacgacgtgT	tcggcgcccttgaaAtcC
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	tggaactagttacactgcagtTC
IWGSC_CSS_6AL_scaff_5757109	Cadenza0423	2788	G	A	hom	hom	caggaGcctggcaaataaaGG	caggaGcctggcaaataaaGA	ctttcGcagtctcttagtttcG
IWGSC_CSS_6AS_scaff_4387871	Cadenza0423	2543	G	A	hom	hom	gcatgctaacaggcgaaaagG	gcatgctaacaggcgaaaagA	ctcatgctcctgatcttaaggtT
IWGSC_CSS_6BL_scaff_4271391	Cadenza0423	4660	C	T	hom	hom	tacgtgcatgatgtggtagtcgtaC	tacgtgcatgatgtggtagtcgtaT	gtttgaagtgcatcagatgTaccA
IWGSC_CSS_6DS_scaff_1880206	Cadenza0423	9159	G	A	het	het	ctgCgaaggctccacaaG	ctgCgaaggctccacaaA	ggatgagaagtttgcattgctC
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	tggaatggtgCgttgttttT
IWGSC_CSS_7DS_scaff_407428	Cadenza0423	2051	G	A	het	het	gtcGCgccatcctgacaG	gtcGCgccatcctgacaA	actcatcAggtcagcccaA
IWGSC_CSS_3AL_scaff_442479	Cadenza0364	3198	C	T	het	het	gagtcaTtaagttggtaagattggC	gagtcaTtaagttggtaagattggT	GCaGaTaaCaacaggatcacG
IWGSC_CSS_3AL_scaff_4447942	Cadenza0364	11917	G	A	het	het	gtcataaagattgctcctgtgaaG	gtcataaagattgctcctgtgaaA	ctcGgatgtgggaggaagA
IWGSC_CSS_3AS_scaff_1557483	Cadenza0364	2547	C	T	het	het	aaagtcacatcatgcttaccataaG	aaagtcacatcatgcttaccataaA	cgaaatccaacgcctcatcA
IWGSC_CSS_3AS_scaff_2648747	Cadenza0364	2688	G	A	het	het	tggAagcAcaaggggccC	tggAagcAcaaggggccT	GccgccgatggagactcG
IWGSC_CSS_3AS_scaff_3304956	Cadenza0364	1017	G	A	het	het	gtcccttgcacacagctttG	gtcccttgcacacagctttA	cctgctggactacaacttcaaT
IWGSC_CSS_3AS_scaff_3321091	Cadenza0364	4585	C	T	het	het	caagaatgATgctgatgttggaG	caagaatgATgctgatgttggaA	acatgctgaatcgccgaatC
IWGSC_CSS_3AS_scaff_3371333	Cadenza0364	538	G	A	het	het	gggaaaCgAgAcgagcgG	gggaaaCgAgAcgagcgA	ccgtgccttcctcacccT
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	ccgtaaaactttctgtgcttgC	ccgtaaaactttctgtgcttgT	atActgacaaactacatgatgtgC
IWGSC_CSS_3B_scaff_10343586	Cadenza0364	2242	G	A	het	—	ggttcTgTcctctcttccactG	ggttcTgTcctctcttccactA	tgtgttgaacccgcaagcA

IWGSC contig	Line	Pos	WT	Mut	Predicted	Called on M_4	Primer 1 (Cadenza)	Primer 2 (mutant)	Common Primer
IWGSC_CSS_3AL_scaff_442479	Cadenza0364	3198	C	T	het	het	gagtcaTtaagttggtaagattggC	gagtcaTtaagttggtaagattggT	GCaGaTaaCaacaggatcacG
IWGSC_CSS_3AL_scaff_4447942	Cadenza0364	11917	G	A	het	het	gtcataaagattgctcctgtgaaG	gtcataaagattgctcctgtgaaA	ctcGgatgtgggaggaagA
IWGSC_CSS_3AS_scaff_1557483	Cadenza0364	2547	C	T	het	het	aaagtcacatcatgcttaccataaG	aaagtcacatcatgcttaccataaA	cgaaatccaacgcctcatcA
IWGSC_CSS_3AS_scaff_2648747	Cadenza0364	2688	G	A	het	het	tggAagcAcaaggggccC	tggAagcAcaaggggccT	GccgccgatggagactcG
IWGSC_CSS_3AS_scaff_3304956	Cadenza0364	1017	G	A	het	het	gtcccttgcacacagctttG	gtcccttgcacacagctttA	cctgctggactacaacttcaaT
IWGSC_CSS_3AS_scaff_3321091	Cadenza0364	4585	C	T	het	het	caagaatgATgctgatgttggaG	caagaatgATgctgatgttggaA	acatgctgaatcgccgaatC
IWGSC_CSS_3AS_scaff_3371333	Cadenza0364	538	G	A	het	het	gggaaaCgAgAcgagcgG	gggaaaCgAgAcgagcgA	ccgtgccttcctcacccT
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	ccgtaaaactttctgtgcttgC	ccgtaaaactttctgtgcttgT	atActgacaaactacatgatgtgC
IWGSC_CSS_3B_scaff_10343586	Cadenza0364	2242	G	A	het	—	ggttcTgTcctctcttccactG	ggttcTgTcctctcttccactA	tgtgttgaacccgcaagcA
IWGSC_CSS_5DL_scaff_242342	Cadenza0281	2433	C	T	hom	hom	catggCgacggtGtcctG	catggCgacggtGtcctA	aAccctcatTTtggCTACTtCT
IWGSC_CSS_5DL_scaff_4538822	Cadenza0281	1208	G	A	hom	—	acgtcagaacaaccgtttgaC	acgtcagaacaaccgtttgaT	ttaaattggttggcgccacC
IWGSC_CSS_6AL_scaff_5813297	Cadenza0281	4532	C	T	hom	—	gggagagggacgtctcgG	gggagagggacgtctcgA	ttcttctgccaacgattccG
IWGSC_CSS_6AS_scaff_4378990	Cadenza0281	6748	C	T	hom	hom	cccaggttctgcttcttttcC	cccaggttctgcttcttttcT	caagtatcaagaaaatgaagggTgT
IWGSC_CSS_6BL_scaff_4360781	Cadenza0281	5426	C	T	het	het	aCtactcaaatggcttGgtgtaG	aCtactcaaatggcttGgtgtaA	tcagtccaacatgTcaagagatT
IWGSC_CSS_7AL_scaff_4488310	Cadenza0281	3808	G	A	hom	hom	gttctcttgtagtagcagccG	gttctcttgtagtagcagccA	ggcgctttcttcggcctA
IWGSC_CSS_7BL_scaff_6696509	Cadenza0281	9232	G	A	het	het	gctctaggGgtggcaaAagG	gctctaggGgtggcaaAagA	ggcttGaGgtcGcagtgT
IWGSC_CSS_7BS_scaff_3143575	Cadenza0281	1866	C	T	het	het	agatgttgagagggcgcttC	agatgttgagagggcgcttT	gcttggAtggtggcaagtT
IWGSC_CSS_7DL_scaff_3346250	Cadenza0281	1663	G	A	het	het	acgtgcagcaacatcctaaC	acgtgcagcaacatcctaaT	TttcccaccaggcccaagA
IWGSC_CSS_7DS_scaff_3933917	Cadenza0281	1243	C	T	het	het	tgCtgagcCttTcaccttgC	tgCtgagcCttTcaccttgT	agaggtttggttccatcGG
IWGSC_CSS_3B_scaff_10626860	Cadenza0148	7847	G	A	het	het	gcagctctgggaaggagG	gcagctctgggaaggagA	gttaatgtacCTcctagcctcG
IWGSC_CSS_3DL_scaff_6915683	Cadenza0148	6904	C	T	het	het	cgtcaaCctgtgggcaattG	cgtcaaCctgtgggcaattA	tcatgctcataatgTcatagggT
IWGSC_CSS_4AS_scaff_5929057	Cadenza0148	4238	G	A	hom	hom	gcgcaacgtagCacctacC	gcgcaacgtagCacctacT	ttatctggtgaagtgacaggttCA
IWGSC_CSS_4AS_scaff_5950625	Cadenza0148	10590	C	T	het	het	agaTattCaaaTcggtggAttggC	agaTattCaaaTcggtggAttggT	cctgCtcccctcacgtcC
IWGSC_CSS_4AS_scaff_5967119	Cadenza0148	11626	C	T	hom	hom	cgtGgacaccccgagctG	cgtGgacaccccgagctA	gacgacgcactgcacgaC
IWGSC_CSS_4DL_scaff_14455742	Cadenza0148	1946	C	T	hom	hom	gCctgagggagatcgcgC	gCctgagggagatcgcgT	aaccgGtAaCTGtGgGcA
IWGSC_CSS_4DS_scaff_2318993	Cadenza0148	4000	C	T	hom	hom	tccagtttgacacagattgaatggG	tccagtttgacacagattgaatggA	tgagaTtctgtttcctttcacAttG
IWGSC_CSS_5AL_scaff_2750707	Cadenza0148	4603	G	A	het	het	ccttggtgctagccatttcaagTaG	ccttggtgctagccatttcaagTaA	ccaggaTgcAgtgcaatatttcaaG
IWGSC_CSS_5BL_scaff_10794137	Cadenza0148	9235	C	T	hom	hom	gaagctgcttctgcgttG	gaagctgcttctgcgttA	agtatcccttccatataagcagtG
IWGSC_CSS_5BS_scaff_1646558	Cadenza0148	2916	C	T	het	het	gccGtacactcacctAtcctttG	gccGtacactcacctAtcctttA	gcaaTgtccacttAtcatcccT
IWGSC_CSS_1AL_scaff_3883106	Cadenza0110	27536	C	T	het	het	accttccatcactggctgG	accttccatcactggctgA	gtgaagaacaacaggttgaagC
IWGSC_CSS_1BL_scaff_3812829	Cadenza0110	10770	G	A	het*	hom	cccccactccattccagG	cccccactccattccagA	gGatgttgttctgtgctggaA
IWGSC_CSS_1DL_scaff_2266648	Cadenza0110	6156	G	A	het	het	actgcgtggttatgggacC	actgcgtggttatgggacT	ccccatcactgaacacaacA
IWGSC_CSS_1DS_scaff_1889435	Cadenza0110	8826	C	T	hom	hom	aaccatgaattactcggacagG	aaccatgaattactcggacagA	gccctgaagaattgtatcaaaacaG
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	gaacgtTcattcttgtacttgcT
IWGSC_CSS_2DS_scaff_5371379	Cadenza0110	2166	C	T	hom	hom	agacacaaaactagtGatgcgC	agacacaaaactagtGatgcgT	gctgctgagaatgttTtgtatttG
IWGSC_CSS_3AL_scaff_4384278	Cadenza0110	1276	C	T	het	het	agcTgaactgccccTgtaG	agcTgaactgccccTgtaA	agggacctCgGtggatgaA
IWGSC_CSS_3AS_scaff_3340122	Cadenza0110	1467	C	T	hom	hom	attcctAgtgttgtcggaacatG	attcctAgtgttgtcggaacatA	gagaagactagaaagttttcAgcaT
IWGSC_CSS_5DL_scaff_4554222	Cadenza2103	6528	C	T	het*	hom	gctgccctacaaagaaacaaaattG	gctgccctacaaagaaacaaaattA	aTcccaactatCGaTtttgtcataC
IWGSC_CSS_6AL_scaff_5833640	Cadenza2103	7346	C	T	hom	hom	aagaaaagccacaatggtttctC	aagaaaagccacaatggtttctT	aCTctgTcagtgtttcccagC
IWGSC_CSS_6AS_scaff_4429974	Cadenza2103	3867	G	A	hom	hom	GagatgaAtttattgagcatgtggC	GagatgaAtttattgagcatgtggT	ggttccggctgcataagT
IWGSC_CSS_6DL_scaff_3307626	Cadenza2103	4970	C	T	hom	hom	tgcagatgttgtcctgtgtaG	tgcagatgttgtcctgtgtaA	ctaggaaggtgattttgtactGtC
IWGSC_CSS_6DS_scaff_2059604	Cadenza2103	5224	G	A	het	—	gctcaatgcatgcTgagtgG	gctcaatgcatgcTgagtgA	tgtcaagtattattttcctgctctG
IWGSC_CSS_7AL_scaff_4552322	Cadenza2103	1412	C	T	het	het	gcaaaggcTgatactccaacaG	gcaaaggcTgatactccaacaA	ggcAAGccAgtataaaagtaaGC
IWGSC_CSS_7BS_scaff_3147455	Cadenza2103	4607	G	A	het	—	gcaccttaggatgtgagTtatgC	gcaccttaggatgtgagTtatgT	gcatgtagggtttatttgactgttA
IWGSC_CSS_7DL_scaff_3382467	Cadenza2103	3473	C	T	hom	—	GGTtctgCaGTTCATAActcatC	GGTtctgCaGTTCATAActcatT	attgaatcaactgatacGaaGactC
IWGSC_CSS_3B_scaff_10457010	Cadenza0277	10599	G	A	het	het	aaccttggccgcagaacaC	aaccttggccgcagaacaT	actggctgcacgagaggG
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	ccttgtattcaatggaTtgTtttgG	ccttgtattcaatggaTtgTtttgA	ttccccaaaTaaaaaggaagagC
IWGSC_CSS_4AL_scaff_7176064	Cadenza0277	6220	C	T	het	het	gtgccgtaTtcCgcctgG	gtgccgtaTtcCgcctgA	atgttcgaggggatgggG
IWGSC_CSS_4DL_scaff_14122349	Cadenza0277	1010	C	T	hom	hom	gtcgctgctgCttgtgaG	gtcgctgctgCttgtgaA	ggaacaggcccaaggagG
IWGSC_CSS_5AL_scaff_2736916	Cadenza0277	4296	G	A	het	het	aagaactATgAaaGtaacacacgaC	aagaactATgAaaGtaacacacgaT	ttcGcTttTaagGcAttCtcG
IWGSC_CSS_5BL_scaff_10883744	Cadenza0277	2080	C	T	hom	hom	gcctctttCtgttTagcctcaG	gcctctttCtgttTagcctcaA	cgacaaggttcgtgatTgcA
IWGSC_CSS_1AL_scaff_3932013	Cadenza0548	11765	C	T	hom	hom	accgccaaCccaagacaG	accgccaaCccaagacaA	cccattaGccgTgcAacG
IWGSC_CSS_1BS_scaff_3417505	Cadenza0548	373	C	T	het	het	gtggtgaggaGGgtgGaG	gtggtgaggaGGgtgGaA	tggtcgGccagttgttgA
IWGSC_CSS_2AS_scaff_5305619	Cadenza0548	2786	C	T	hom	hom	atacagatgccctAAgtggTtC	atacagatgccctAAgtggTtT	ggaagacaAtGctccaggtaC
IWGSC_CSS_2AS_scaff_5306489	Cadenza0548	46953	T	G	het	wt	aggttccatgtccatagaagGT	aggttccatgtccatagaagGG	aggctaTAgactcctgtACAgT
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	tgcctccctttgccagaaC	tgcctccctttgccagaaT	ggcaaacctgatgtggcatC
IWGSC_CSS_2DS_scaff_5330886	Cadenza0548	5449	G	A	hom	hom	gcatgtccatttatactgaaCgtG	gcatgtccatttatactgaaCgtA	catgctgcttcttctggacC
IWGSC_CSS_3AL_scaff_4449951	Cadenza0548	633	C	T	het	het	tccaaacctaacagtctaacactaG	tccaaacctaacagtctaacactaA	gtctgcagTGCaatgtgC
IWGSC_CSS_3B_scaff_10479889	Cadenza0097	3339	C	T	hom	—	ttgTttctGgagaagatgcCG	ttgTttctGgagaagatgcCA	ggtgctcattcaAcGgcA
IWGSC_CSS_3B_scaff_10562262	Cadenza0097	7819	C	T	het	het	agaggggtgctatccatAttgG	agaggggtgctatccatAttgA	agcgatgccaaggcttcC
IWGSC_CSS_4AL_scaff_7040796	Cadenza0097	10772	G	A	hom	hom	acacaacattgccaccagaG	acacaacattgccaccagaA	CAatCgattgcttgctTctcC
IWGSC_CSS_4AL_scaff_7063488	Cadenza0097	6360	C	T	het	het	gcctctcacCttAatttgaagctgC	gcctctcacCttAatttgaagctgT	aggcagtggagtatgtgaagttT
IWGSC_CSS_4AL_scaff_7091701	Cadenza0097	5050	G	A	het	het	catgagcatctgggaggaaaatG	catgagcatctgggaggaaaatA	agcaagggaAtaatgaacggaaA
IWGSC_CSS_4DS_scaff_1845841	Cadenza0097	7110	G	A	hom	hom	aatgTAgctccccatacCgG	aatgTAgctccccatacCgA	actgaaacTgcaatcgtTtatggA
IWGSC_CSS_5AL_scaff_2767581	Cadenza0097	3737	G	A	het	het	gagaggtcctcactAtcggC	gagaggtcctcactAtcggT	cgTcatcacaaatattgctggG
IWGSC_CSS_5BL_scaff_10784643	Cadenza0097	1568	C	T	hom	hom	agaaaTAcatggatggatggaCG	agaaaTAcatggatggatggaCA	catctcCCttccaCgGaaaG

IWGSC contig	Line	Pos	WT	Mut	Predicted	Called on M_4	Primer 1 (Cadenza)	Primer 2 (mutant)	Common Primer
IWGSC_CSS_1AL_scaff_3952258	Cadenza2092	8107	C	T	het	—	tgagtagagaaattgacagtgtgG	tgagtagagaaattgacagtgtgA	tgccaccattgacatgagaG
IWGSC_CSS_1BL_scaff_3858008	Cadenza2092	10278	G	A	hom	hom	tttgagcaggcaggatcgC	tttgagcaggcaggatcgT	actcacggcctatatcActattC
IWGSC_CSS_1DL_scaff_2265172	Cadenza2092	9094	C	T	hom	hom	tgcaTGTcatttgttcttatcagC	tgcaTGTcatttgttcttatcagT	agtgtccaacttccGttcatC
IWGSC_CSS_2AL_scaff_6435867	Cadenza2092	16201	G	A	hom	hom	tttctgTaccttaacgtcaattgaC	tttctgTaccttaacgtcaattgaT	gtgaggatgatgaggtaagacC
IWGSC_CSS_2AL_scaff_6439430	Cadenza2092	25101	C	T	het	—	caagaaagggCagCtCagC	caagaaagggCagCtCagT	tcGttAcTctttcActggtgaA
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	acctctctcatgttaataggtttgT
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	ggcaaggtcaagaaacGgtC	ggcaaggtcaagaaacGgtT	aCagaGtgggttagaggcaG
IWGSC_CSS_3B_scaff_10550638	Cadenza1551	3250	C	T	het	het	ctccttcacttgttgcggC	ctccttcacttgttgcggT	gcaacAtTttgatactgcaaagG
IWGSC_CSS_3DL_scaff_6945816	Cadenza1551	589	C	T	hom	hom	agcatctcacctgcaaCaataC	agcatctcacctgcaaCaataT	TgtgcccTctgaAtattttcaTG
IWGSC_CSS_3DL_scaff_6954177	Cadenza1551	3508	C	T	het	het	tgtagcatcacattaactttcctG	tgtagcatcacattaactttcctA	gcttggtataaaccCttacgacA
IWGSC_CSS_4AS_scaff_5938272	Cadenza1551	19080	G	A	hom	hom	agAcCccgAtcgccatgG	agAcCccgAtcgccatgA	GggAgatAcaggtaaaActcTtcG
IWGSC_CSS_4AS_scaff_5977594	Cadenza1551	11092	C	T	het	het	gccttgattcggaacaacaaaC	gccttgattcggaacaacaaaT	gcgtctctcagtcctgcA
IWGSC_CSS_5AL_scaff_2671035	Cadenza1551	5859	C	T	het	het	cggtgatattTttagacttcgacgC	cggtgatattTttagacttcgacgT	ggcagttcagcGacccatT
IWGSC_CSS_5BL_scaff_10889480	Cadenza1551	2530	G	A	hom	hom	gagcttaactcgcagatggaG	gagcttaactcgcagatggaA	tccatgCAacGccttggT
IWGSC_CSS_3B_scaff_10528396	Cadenza2088	8059	G	A	hom	—	cttttccgtccgtaagcaataG	cttttccgtccgtaagcaataA	gtgcactgttcaggcctgA
IWGSC_CSS_3B_scaff_10637573	Cadenza2088	16815	G	A	het	het	agcaagcttaccGgtctgC	agcaagcttaccGgtctgT	cgagcAactacgagcagctT
IWGSC_CSS_4AL_scaff_7086469	Cadenza2088	6697	G	A	het	het	gccgtctacttcaacgcG	gccgtctacttcaacgcA	ccaGaggcttgtTGcattttT
IWGSC_CSS_4AL_scaff_7126302	Cadenza2088	3627	G	A	hom	hom	gttcaaaaacaagtggctAatttgC	gttcaaaaacaagtggctAatttgT	cacaaggatatgaagcTcttctagA
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	caatgaaacctctcgaagaaGaG
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	agCgaacaagagGtcaaacG	agCgaacaagagGtcaaacA	ctgaaacacaCtagaCAattAccG
IWGSC_CSS_3B_scaff_10674115	Cadenza1409	4555	C	T	het	het	gcttcagtgcatgccttcaG	gcttcagtgcatgccttcaA	cttcacacccGagataatGtattG
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	gtaggtcccctcctCAggG	gtaggtcccctcctCAggA	cggcgTcacaAgttgCcT
IWGSC_CSS_4DS_scaff_2324074	Cadenza1409	7606	G	A	het	het	tGcatgaaaatgtgtGcaGaG	tGcatgaaaatgtgtGcaGaA	gggtaAgttcAaaactGaagtgaaG
IWGSC_CSS_5AS_scaff_1517889	Cadenza1409	3561	G	A	het	het	tctcgacatcttcccgtgtaC	tctcgacatcttcccgtgtaT	gtgcctggaacattgcttatttA
IWGSC_CSS_5AS_scaff_1523866	Cadenza1409	8054	G	A	hom	—	ggtgatctaccgccaGgaC	ggtgatctaccgccaGgaT	tcctgcagCcTctcctcA
IWGSC_CSS_5BL_scaff_10917655	Cadenza1409	19073	G	A	hom	hom	caaatgacatgcaaaagaagttgC	caaatgacatgcaaaagaagttgT	cgcttcatcactacaAaatatgtcT
IWGSC_CSS_1AL_scaff_3886649	Cadenza1599	5204	C	T	het	het	tgatgccaaccacaatGcC	tgatgccaaccacaatGcT	ggactgactgctgaccatatttaG
IWGSC_CSS_1BL_scaff_3810267	Cadenza1599	6634	C	T	hom	hom	ccCaggaaatgagcacctC	ccCaggaaatgagcacctT	cgcaggcgaagatgtgaTtG
IWGSC_CSS_1DL_scaff_2291677	Cadenza1599	12856	C	T	hom	hom	GgtagacaagtcgccgaG	GgtagacaagtcgccgaA	cctcctccttcaacGCcG
IWGSC_CSS_2AL_scaff_6354492	Cadenza1599	7566	G	A	het	het	gGagaatgcaCAgtAacTtctgG	gGagaatgcaCAgtAacTtctgA	ttccgaagaaccacaTccTG
IWGSC_CSS_2AS_scaff_5282937	Cadenza1599	9736	G	A	het	het	gctgtagattttatagctgctatgC	gctgtagattttatagctgctatgT	cacCagaattgttCactgatttTC
IWGSC_CSS_2BL_scaff_7952427	Cadenza1599	19249	G	A	hom	hom	cgTccctCcctagcacgaC	cgTccctCcctagcacgaT	aTcactccattagcgcgAG
IWGSC_CSS_2DL_scaff_9897981	Cadenza1599	5627	C	T	het	het	cttggtgctTgattgcttactC	cttggtgctTgattgcttactT	gTttgctCtctctgatctTtgtG
IWGSC_CSS_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.

Jarrod 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. bio-samtools 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.

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.

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.

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. Biosamtools: 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.