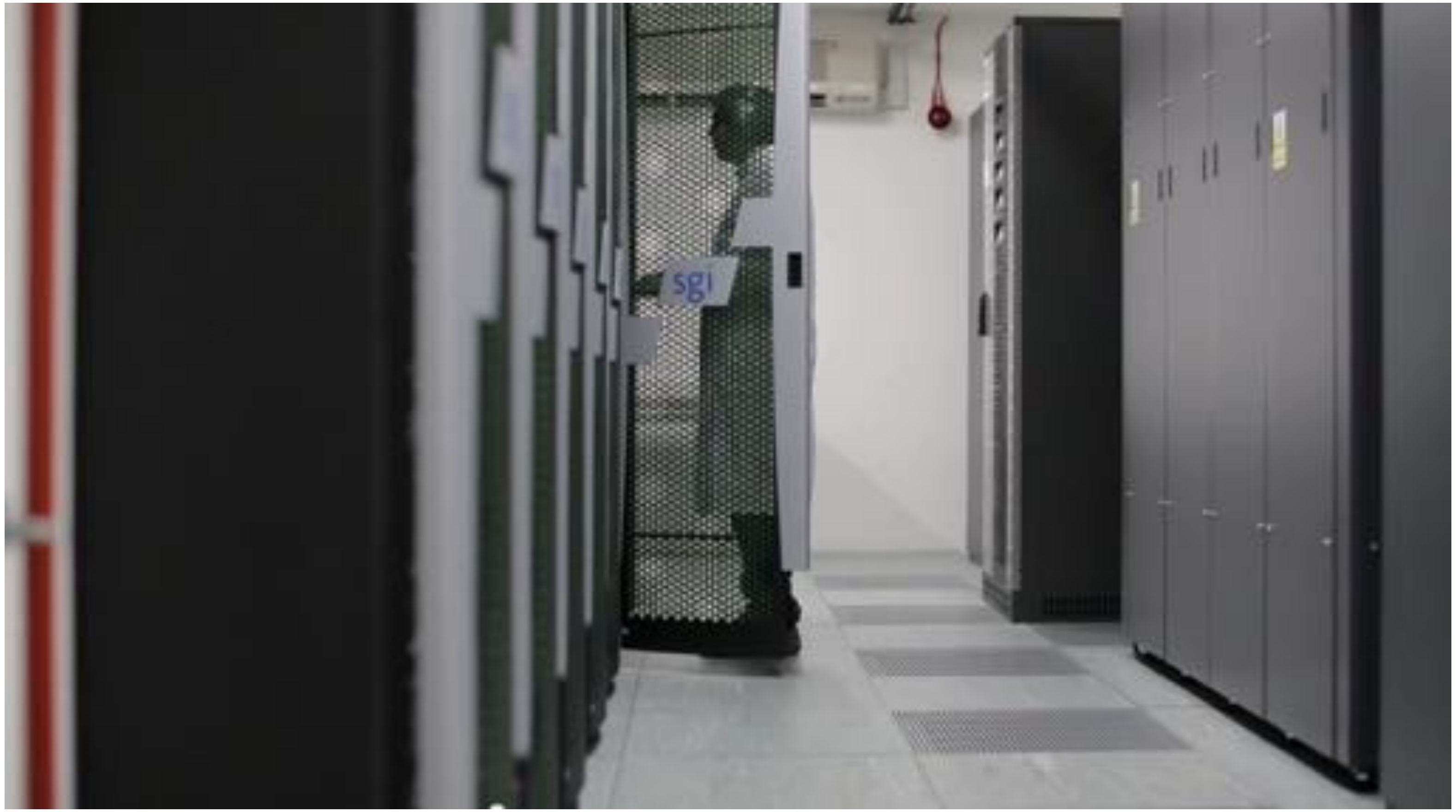
John Innes Centre

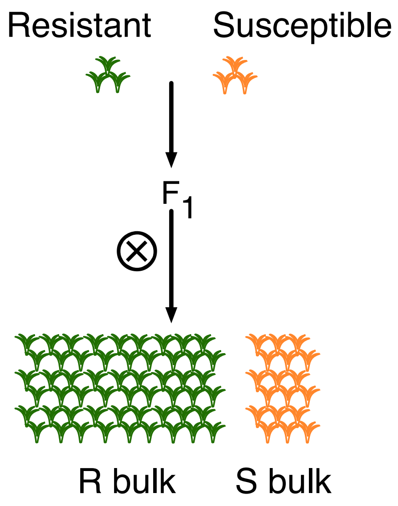
Genome-specific primer design with PolyMarker

RICARDO H. RAMIREZ GONZALEZ

A computer scientist turned to biology



Genotyping applications


Resistant

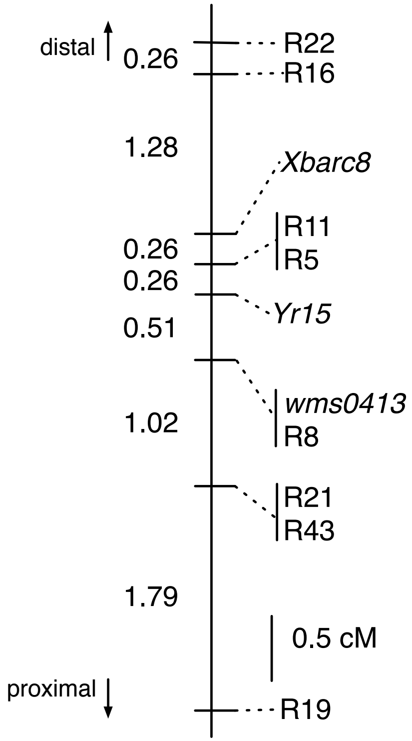
Susceptible

F_1

$\otimes$

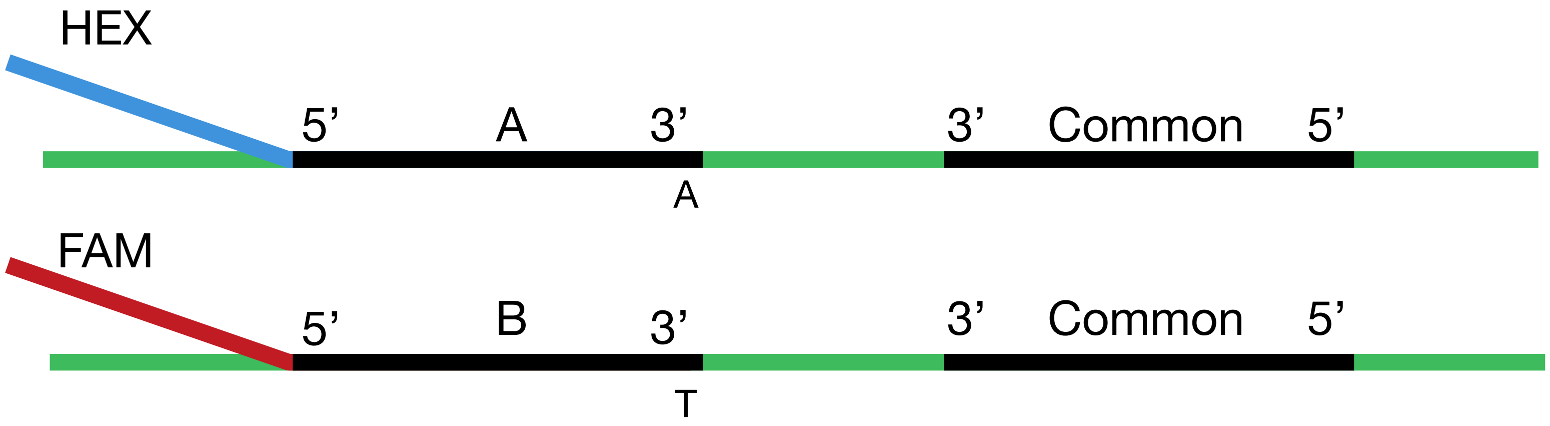
R bulk

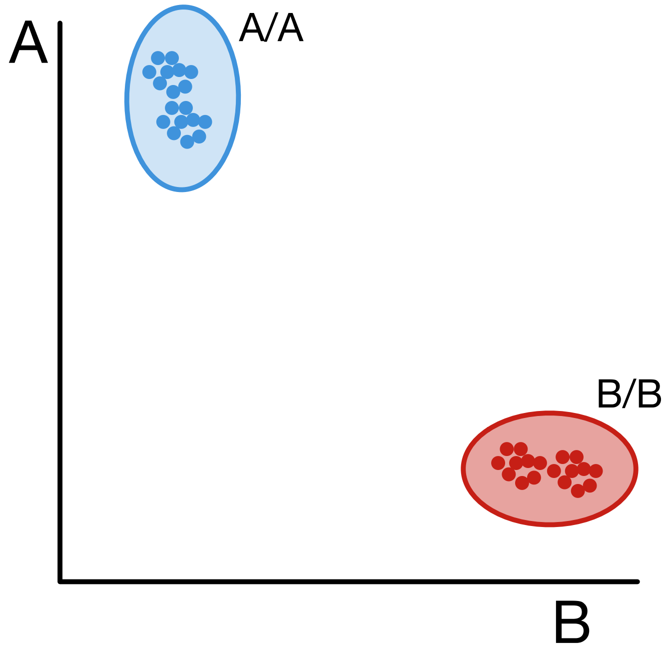
S bulk

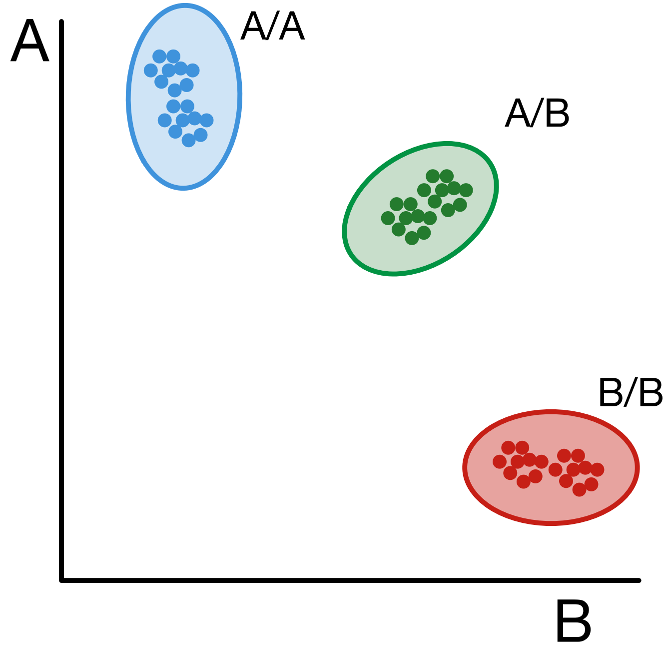


Ramirez-Gonzalez, R, *et al*. (2014). RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. *Plant Biotech. J*, 12(9), 1–12.

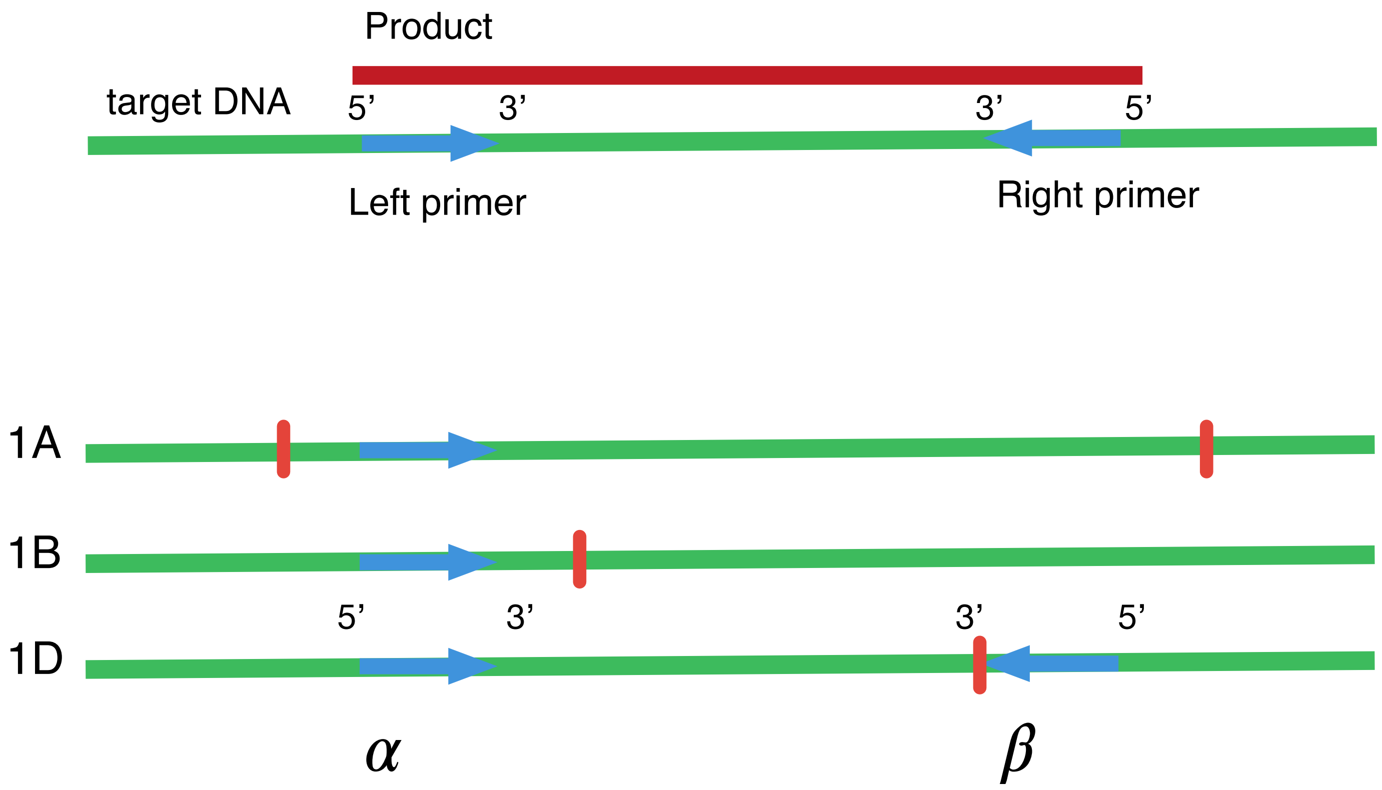
KASP genotyping

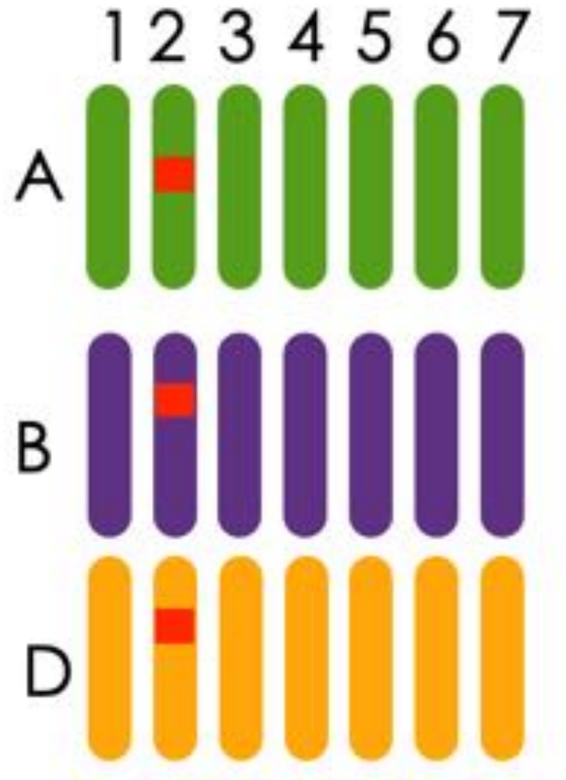




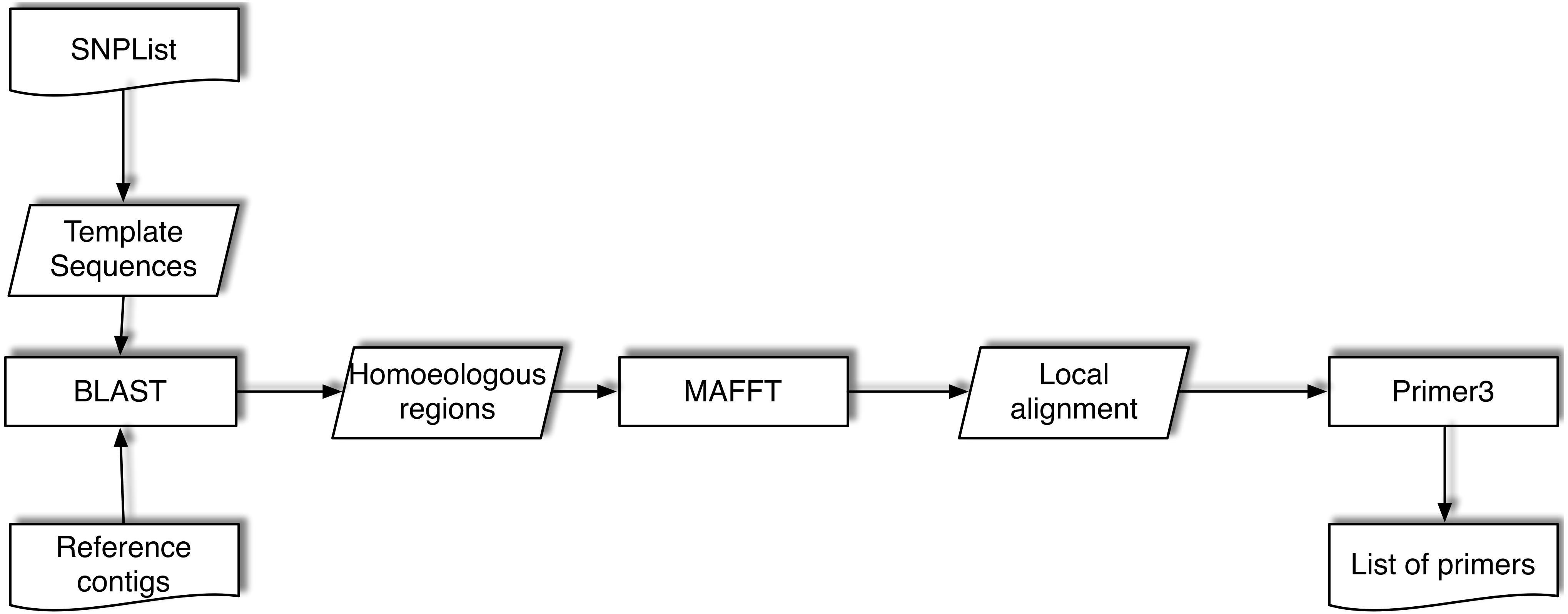


The problem

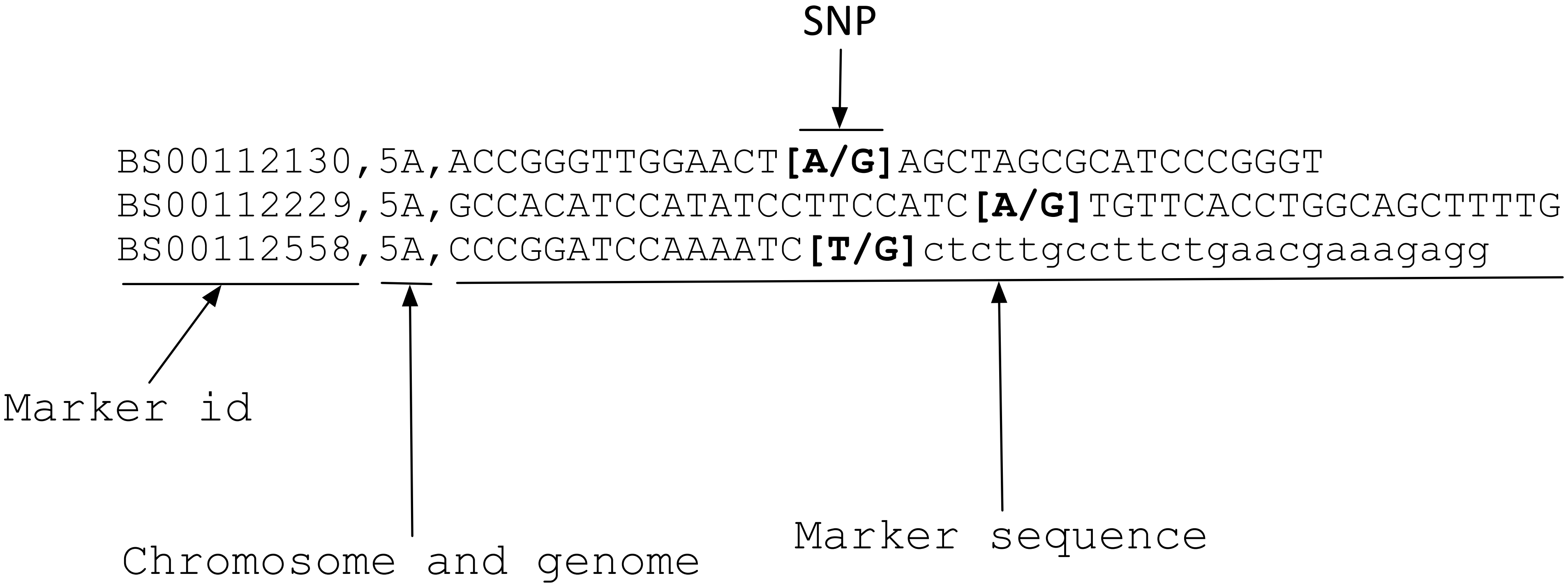

Product
target DNA
5’
3’
3’
5’
Left primer
Right primer
1A
1B
5’
3’
3’
5’
1D
α
β


1 2 3 4 5 6 7
A
B
D

Pipeline overview



Input example



SNP

```
BS00112130,5A,ACCGGGTTGGAACT[A/G]AGCTAGCGCATCCCGGGT
BS00112229,5A,GCCACATCCATATCCTTCCATC[A/G]TGTTCACCTGGCAGCTTTTG
BS00112558,5A,CCCGGATCCAAAATC[T/G]ctcttgccttctgaacgaaagagg
```

Marker id

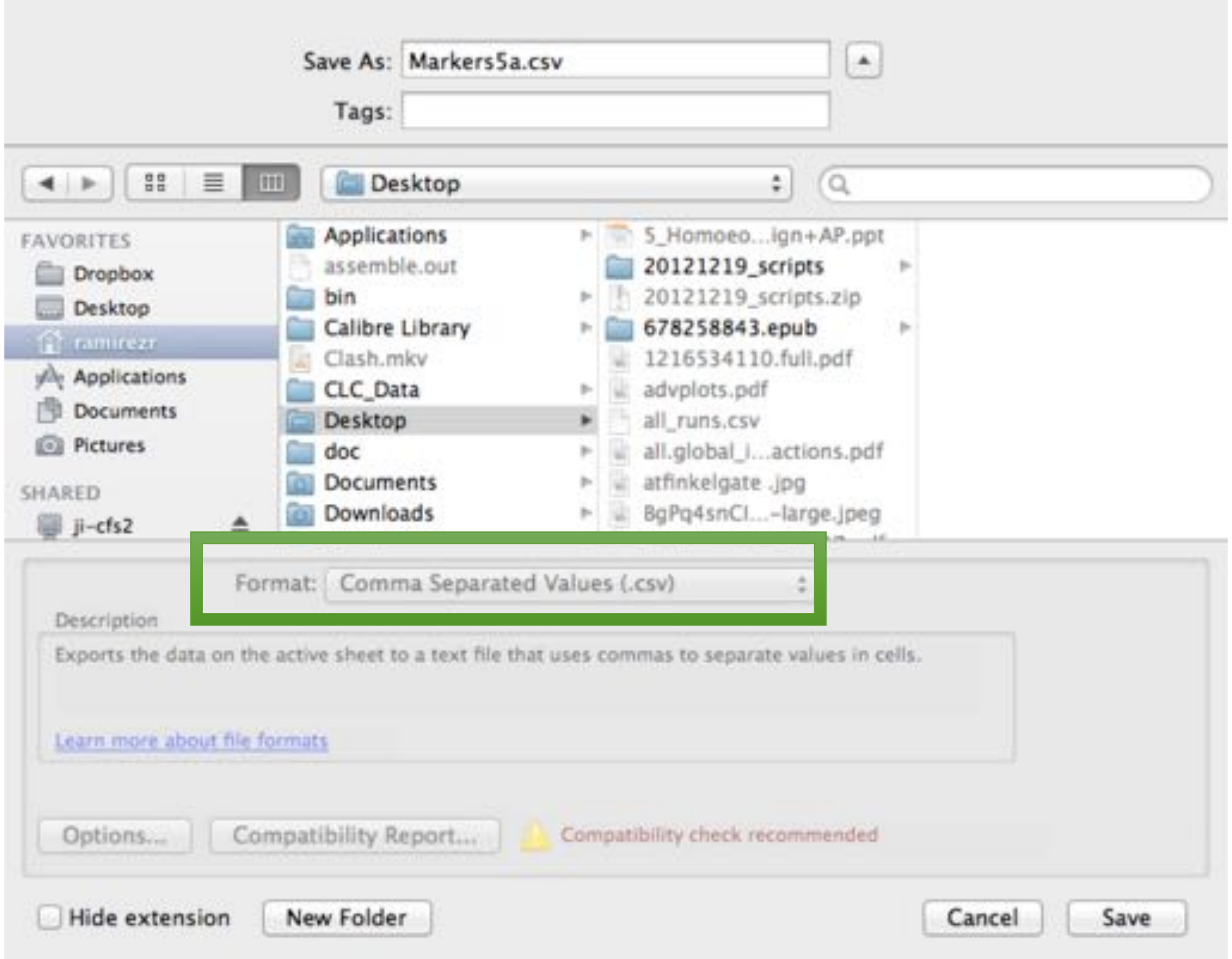
Chromosome and genome

Marker sequence

Input example

SNP ID	Chromosome	Sequence
BS00112130	5A	CTACGCGCCCTCGGGCCCGAACAACCACCACAAGCCACCGGGTTGGAACT[A/G]AGCTAGCGCATCCCGGGTCACAACTTTCATAAAGTTATATCGGTGGTTTC
BS00112229	5A	TCTTCAAATTTGCACTTCATACATCTGTGCCACATCCATATCCTTCCATC[A/G]TGTTCACCTGGCAGCTTTTGGGATAAGAGACGCCTAACACTTATTGTTAG
BS00112558	5A	GGATTCGGCGGGAGAGGGCGGCGGGAGCTGGAGGCCCCGGATCCAAAATC[T/G]ctcttgccttctgaacgaaagaggaaccgcccccatggcagaggttctg
BS00113122	5A	ACAACGCATCACCAATTCGTCCGACAGTGTAAAATGCGTTTGTGTTAGCT[A/G]AGTGAAAATTAGAGATCAGCCTGGTGGTACTTTATTATTTTATTCTATAG
BS00114740	5A	AAAGGGGCATGATGAGACGGGAGCTAGAAAGATTATGCACATTCTTACAC[A/G]AAAGATATATTGAAAAACTGACACTATTTCGCCTCATTGATGGAATCAAC
BS00115880	5A	CTTGCATTCCAAACCTGAAGCATACTTTCATATCAGTTCTTCCTGCCTTC[A/G]cagaggatgaagtggtacccgtgcgtagacttgaatggtgcactagtagc
BS00003618	5A	GTGTGTGAGCTTTGGTGCCATTTTGGGTAGGGCTCTGTATGTGAATTGGG[T/G]AAACTGTGGCCTGGGTGTGTACCACGGTGTCCATTTTCAGTTGTTTTGTT
BS00005311	5A	GAGACTGGTGCCCTTGCAACATGCCAGTGAGCAGGTGCCATGGATCGAAG[A/G]aaacactcctaagccaaagatggacatgggtagagtgatcctagaagatg
BS00009531	5A	ATTCATGCTTTCATTTTTTGTGGTCTCTACCAAAAATGACCCGCAACAGT[A/G]GTGATATTCGCCATCATAGAACCATGGATCCTCCTCGGAGGTCTTCCACC
BS00015653	5A	CACTCCACCAAGAGCGGTGATGAGCTTACCAGCCTCAAGGACTACGTGAC[T/G]AGGATGAAGGAGGGCCAGAACGACATCTACTACATCACCGGTGAGAGCAA
BS00021725	5A	GGTGAACATTTATAATAGACGTGGGCCATTTTCGTGAAATTAGTAATAGC[A/G]GTTATTATTTAACAGTGCATGCACATATGGGAAACATAGTACGAGTATAC

Input example



Save As: Markers5a.csv

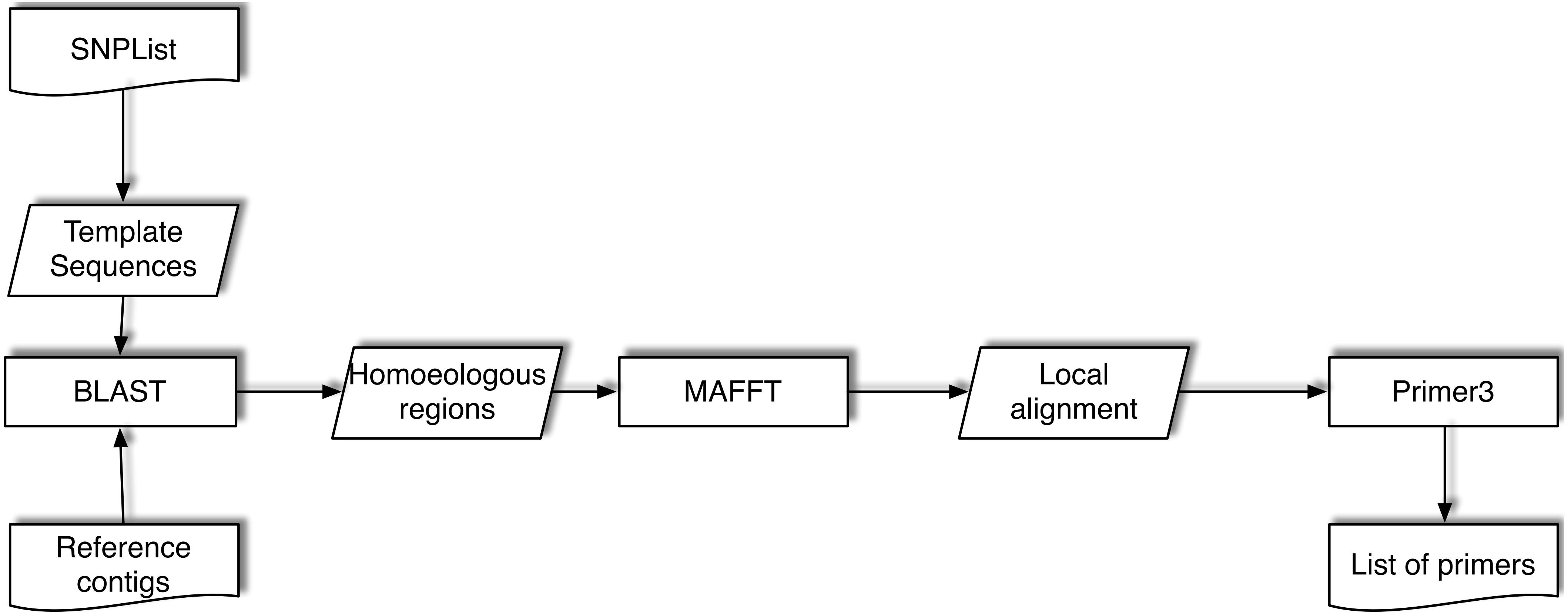
Tags:

Format: Comma Separated Values (.csv)

Description

Exports the data on the active sheet to a text file that uses commas to separate values in cells.

In Silico Primer Design

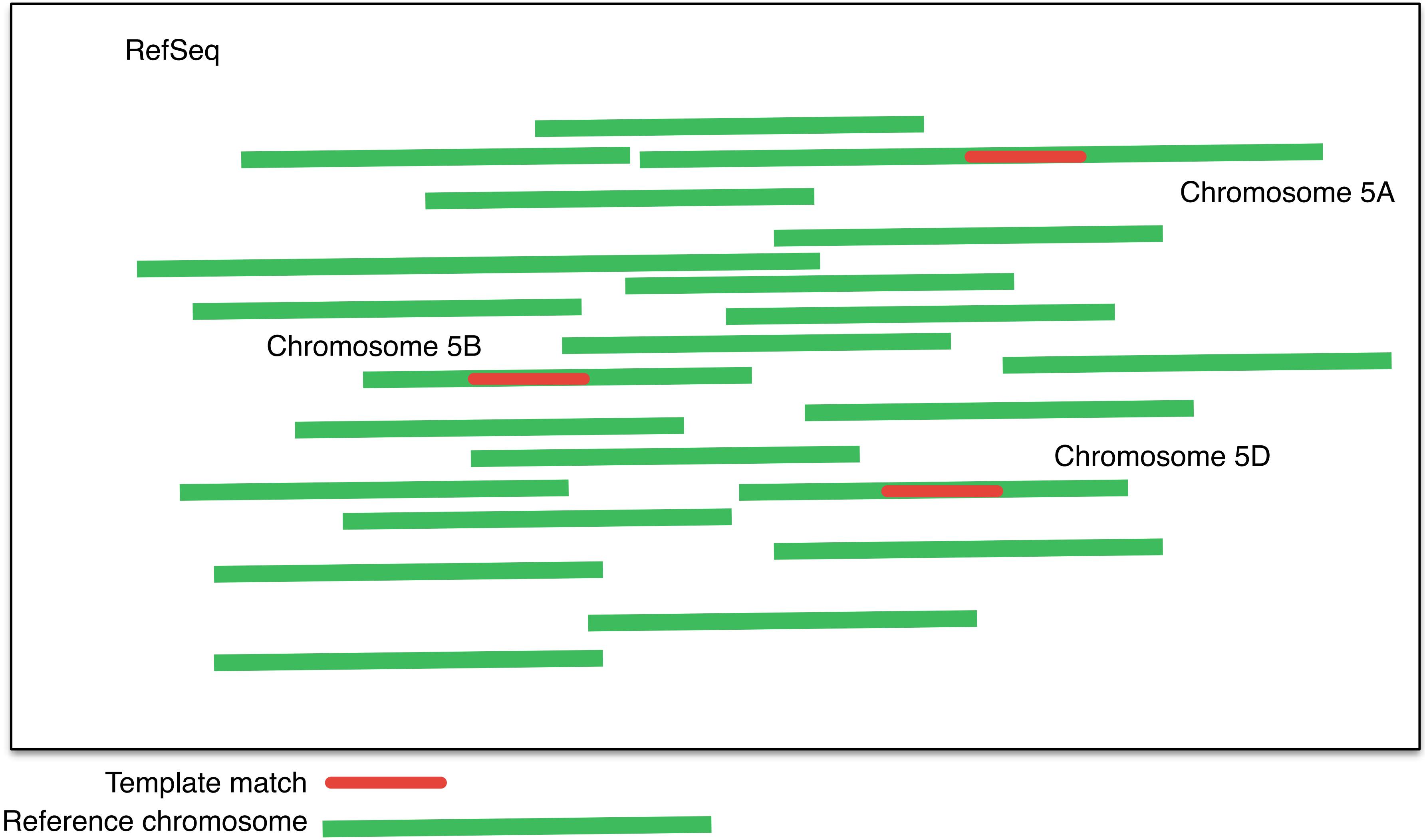


Template sequence

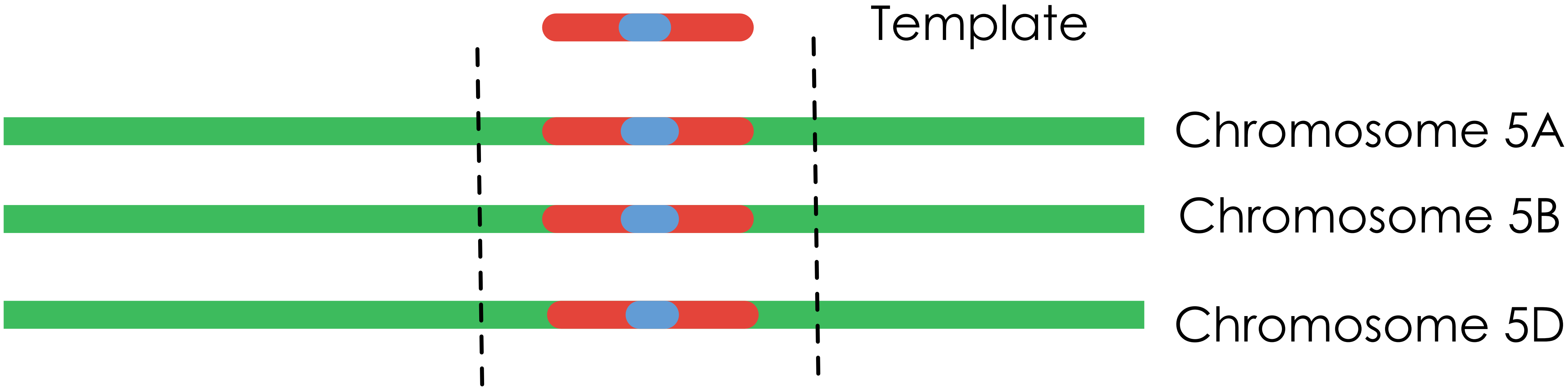
SNP-1 A cgcattt**G**cgcgYgcgataccggcgcctKtg**G**gaatatttgcagcgaaggcgtg

SNP-1 B cgcattt**A**cgcgYgcgataccggcgcctKtg**A**gaatatttgcagcgaaggcgtg

Blast search



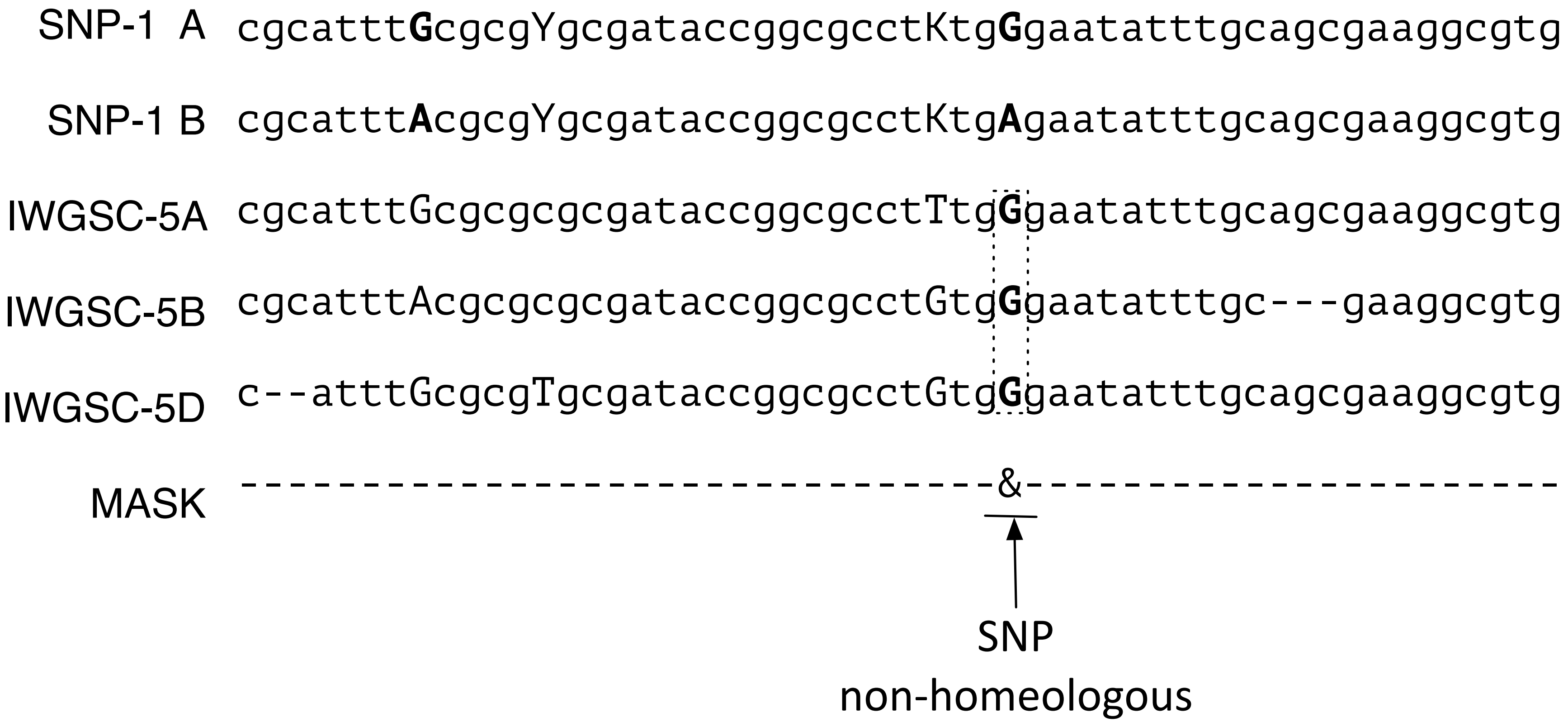
Local Alignment



Local Alignment

SNP-1 A	cgcattt**G**cgcgYgcgataccggcgcctKtg**G**gaatatttgcagcgaaggcgtg
SNP-1 B	cgcattt**A**cgcgYgcgataccggcgcctKtg**A**gaatatttgcagcgaaggcgtg
IWGSC-5A	cgcatttGcgcgcgcgataccggcgcctTtgGgaatatttgcagcgaaggcgtg
IWGSC-5B	cgcatttAcgcgcgcgataccggcgcctGtgGgaatatttgc---gaaggcgtg
IWGSC-5D	c--atttGcgcgTgcgataccggcgcctGtgGgaatatttgcagcgaaggcgtg

Selection of primer candidates



SNP-1 A	`cgcattt`**G**`cgcgYgcgataccggcgcctKtg`**G**`gaatatttgcagcgaaggcgtg`
SNP-1 B	`cgcattt`**A**`cgcgYgcgataccggcgcctKtg`**A**`gaatatttgcagcgaaggcgtg`
IWGSC-5A	`cgcatttGcgcgcgcgataccggcgcctTtg`**G**`gaatatttgcagcgaaggcgtg`
IWGSC-5B	`cgcatttAcgcgcgcgataccggcgcctGtg`**G**`gaatatttgc---gaaggcgtg`
IWGSC-5D	`c--atttGcgcgTgcgataccggcgcctGtg`**G**`gaatatttgcagcgaaggcgtg`
MASK	`-------------------------------&----------------------`

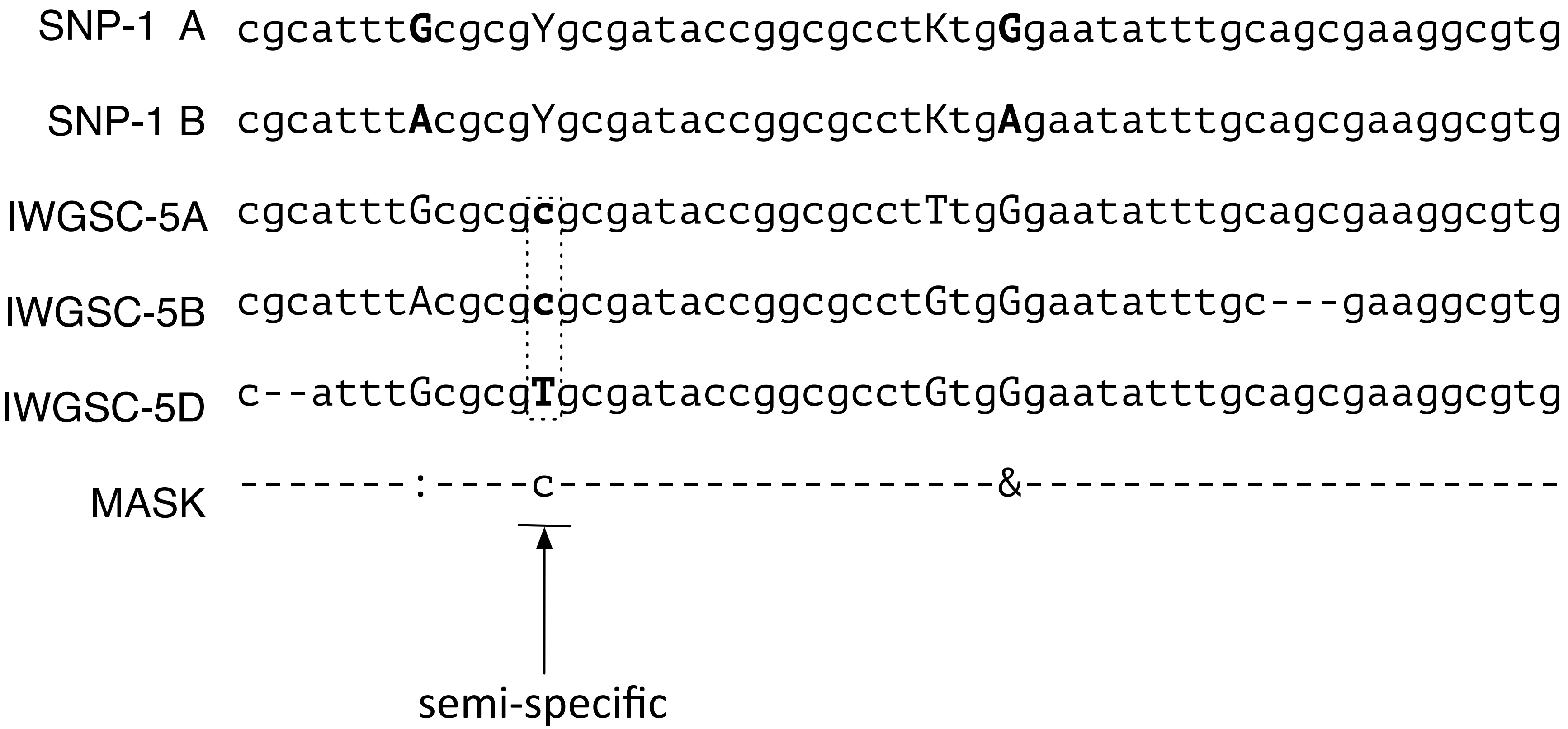
SNP
non-homeologous

Selection of primer candidates

```
SNP-1  A  cgcatttGcgcgYgcgataccggcgcctKtgGgaatatttgcagcgaaggcgtg
SNP-1 B   cgcatttAcgcgYgcgataccggcgcctKtgAgaatatttgcagcgaaggcgtg
IWGSC-5A  cgcatttGcgcgcgcgataccggcgcctTtgGgaatatttgcagcgaaggcgtg
IWGSC-5B  cgcatttAcgcgcgcgataccggcgcctGtgGgaatatttgc---gaaggcgtg
IWGSC-5D  c--atttGcgcgTgcgataccggcgcctGtgGgaatatttgcagcgaaggcgtg
MASK      -------:-----------------------&----------------------
```

SNP
homeologous

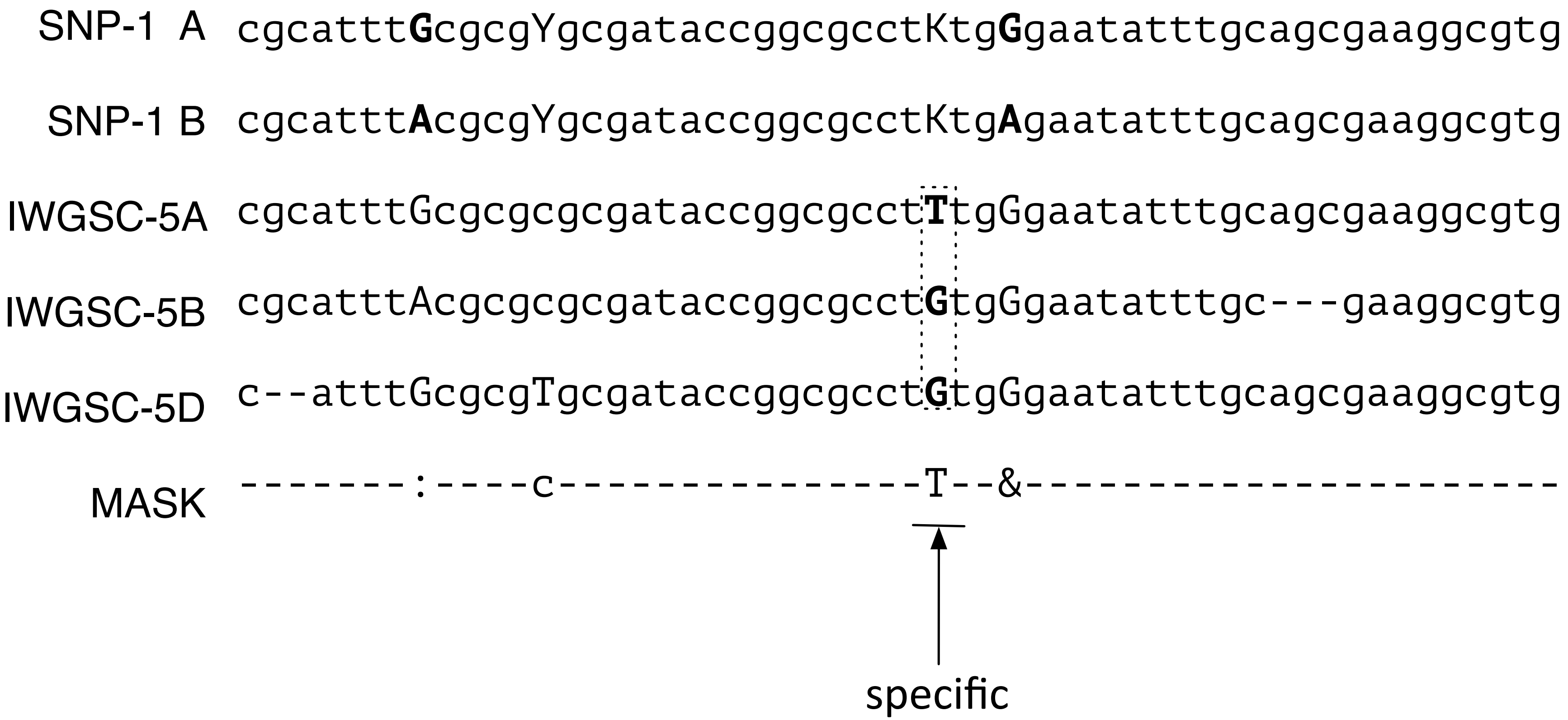
Selection of primer candidates



```
SNP-1  A   cgcatttGcgcgYgcgataccggcgcctKtgGgaatatttgcagcgaaggcgtg
SNP-1 B    cgcatttAcgcgYgcgataccggcgcctKtgAgaatatttgcagcgaaggcgtg
IWGSC-5A   cgcatttGcgcgcgcgataccggcgcctTtgGgaatatttgcagcgaaggcgtg
IWGSC-5B   cgcatttAcgcgcgcgataccggcgcctGtgGgaatatttgc---gaaggcgtg
IWGSC-5D   c--atttGcgcgTgcgataccggcgcctGtgGgaatatttgcagcgaaggcgtg
MASK       -------:----c------------------&----------------------
```

semi-specific

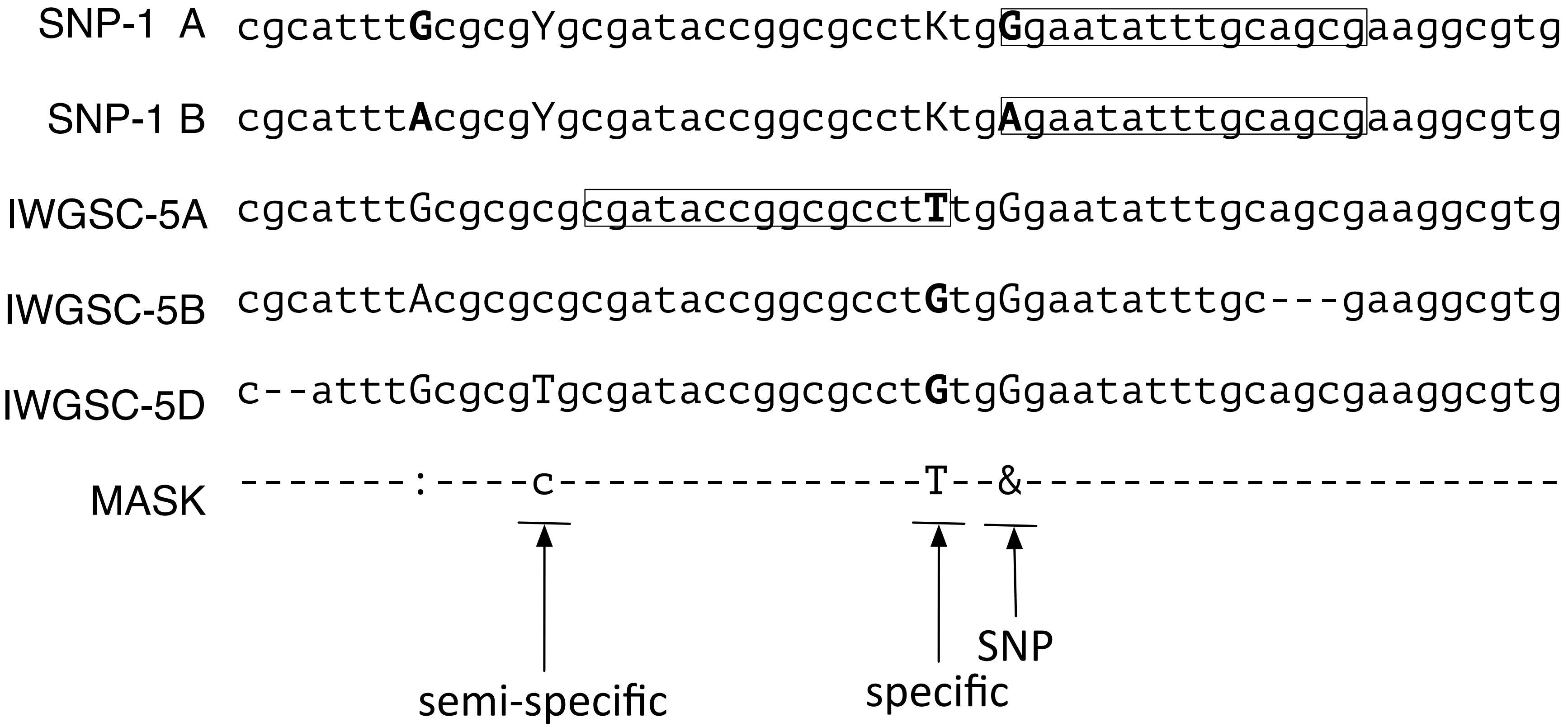
Selection of primer candidates



```
SNP-1  A  cgcatttGcgcgYgcgataccggcgcctKtgGgaatatttgcagcgaaggcgtg
SNP-1 B   cgcatttAcgcgYgcgataccggcgcctKtgAgaatatttgcagcgaaggcgtg
IWGSC-5A  cgcatttGcgcgcgcgataccggcgcctTtgGgaatatttgcagcgaaggcgtg
IWGSC-5B  cgcatttAcgcgcgcgataccggcgcctGtgGgaatatttgc---gaaggcgtg
IWGSC-5D  c--atttGcgcgTgcgataccggcgcctGtgGgaatatttgcagcgaaggcgtg
MASK      -------:----c---------------T--&----------------------
```

specific

Selection of primer candidates



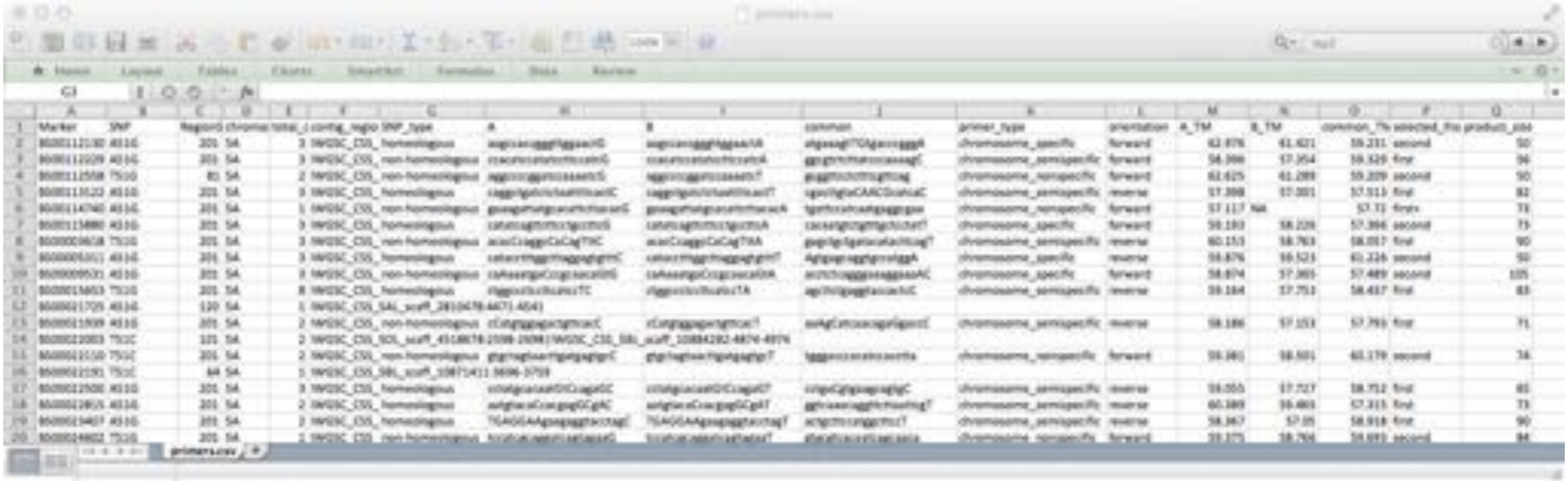
```
SNP-1  A   cgcatttGcgcgYgcgataccggcgcctKtgGgaatatttgcagcgaaggcgtg
SNP-1 B    cgcatttAcgcgYgcgataccggcgcctKtgAgaatatttgcagcgaaggcgtg
IWGSC-5A   cgcatttGcgcgcgcgataccggcgcctTtgGgaatatttgcagcgaaggcgtg
IWGSC-5B   cgcatttAcgcgcgcgataccggcgcctGtgGgaatatttgc---gaaggcgtg
IWGSC-5D   c--atttGcgcgTgcgataccggcgcctGtgGgaatatttgcagcgaaggcgtg
MASK       -------:----c---------------T--&----------------------
```

semi-specific

specific

SNP

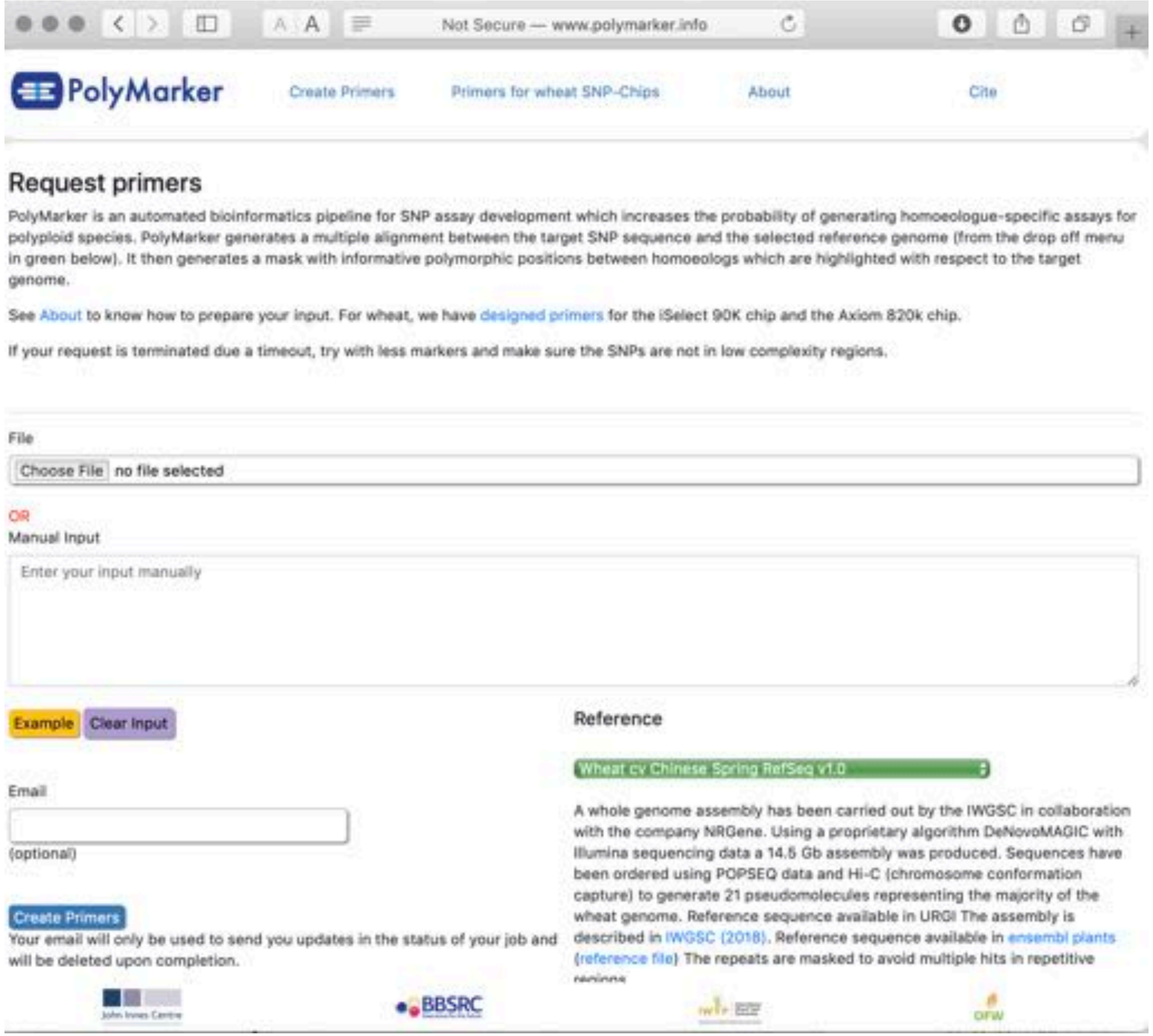
Output



Web interface

- Hexaploid wheat assemblies:
    - IWGSC RefSeq 1.0
    - UK assemblies (Paragon, Cadenza, Robigous, Claire)
- Tetraploid wheat
    - IWGSC RefSeq 1.0 (A&B genomes)
    - Durum wheat cv Svevo
    - UK assembly Kronos
- *Secale cereale* (Rye)
- *Hordeum vulgare* (Barley)
- *Brassica napus*
- *Brassica Rapa*

http://polymarker.info



Request primers

PolyMarker is an automated bioinformatics pipeline for SNP assay development which increases the probability of generating homoeologue-specific assays for polyploid species. PolyMarker generates a multiple alignment between the target SNP sequence and the selected reference genome (from the drop off menu in green below). It then generates a mask with informative polymorphic positions between homoeologs which are highlighted with respect to the target genome.

See About to know how to prepare your input. For wheat, we have designed primers for the iSelect 90K chip and the Axiom 820k chip.

If your request is terminated due a timeout, try with less markers and make sure the SNPs are not in low complexity regions.

Your email will only be used to send you updates in the status of your job and will be deleted upon completion.

A whole genome assembly has been carried out by the IWGSC in collaboration with the company NRGene. Using a proprietary algorithm DeNovoMAGIC with Illumina sequencing data a 14.5 Gb assembly was produced. Sequences have been ordered using POPSEQ data and Hi-C (chromosome conformation capture) to generate 21 pseudomolecules representing the majority of the wheat genome. Reference sequence available in URGI The assembly is described in IWGSC (2018). Reference sequence available in ensembl plants (reference file) The repeats are masked to avoid multiple hits in repetitive

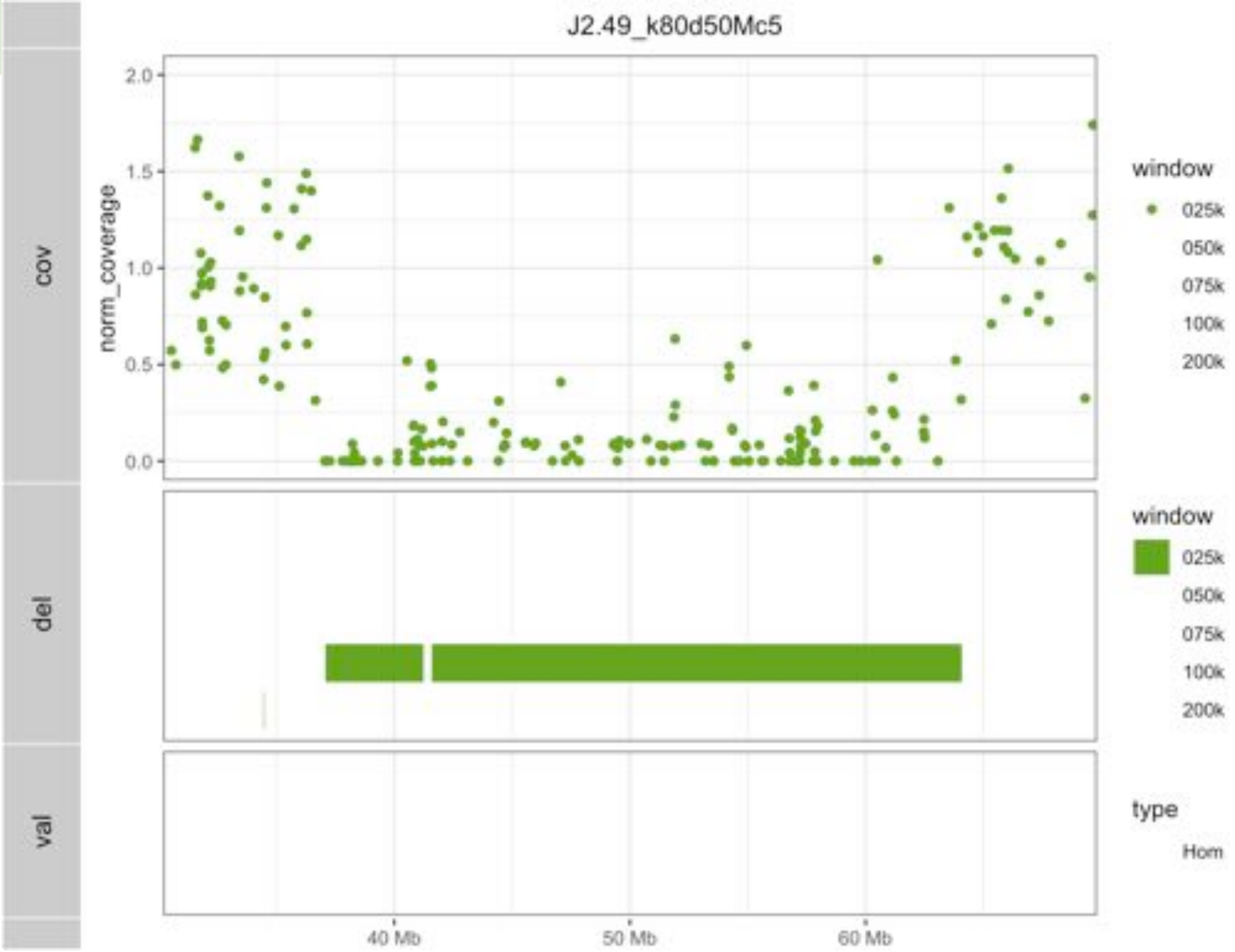
Ramirez-Gonzalez, R. H., Uauy, C., & Caccamo, M. (2015). *PolyMarker: A fast polyploid primer design pipeline*. Bioinformatics, 2–3. doi:10.1093/bioinformatics/btv069

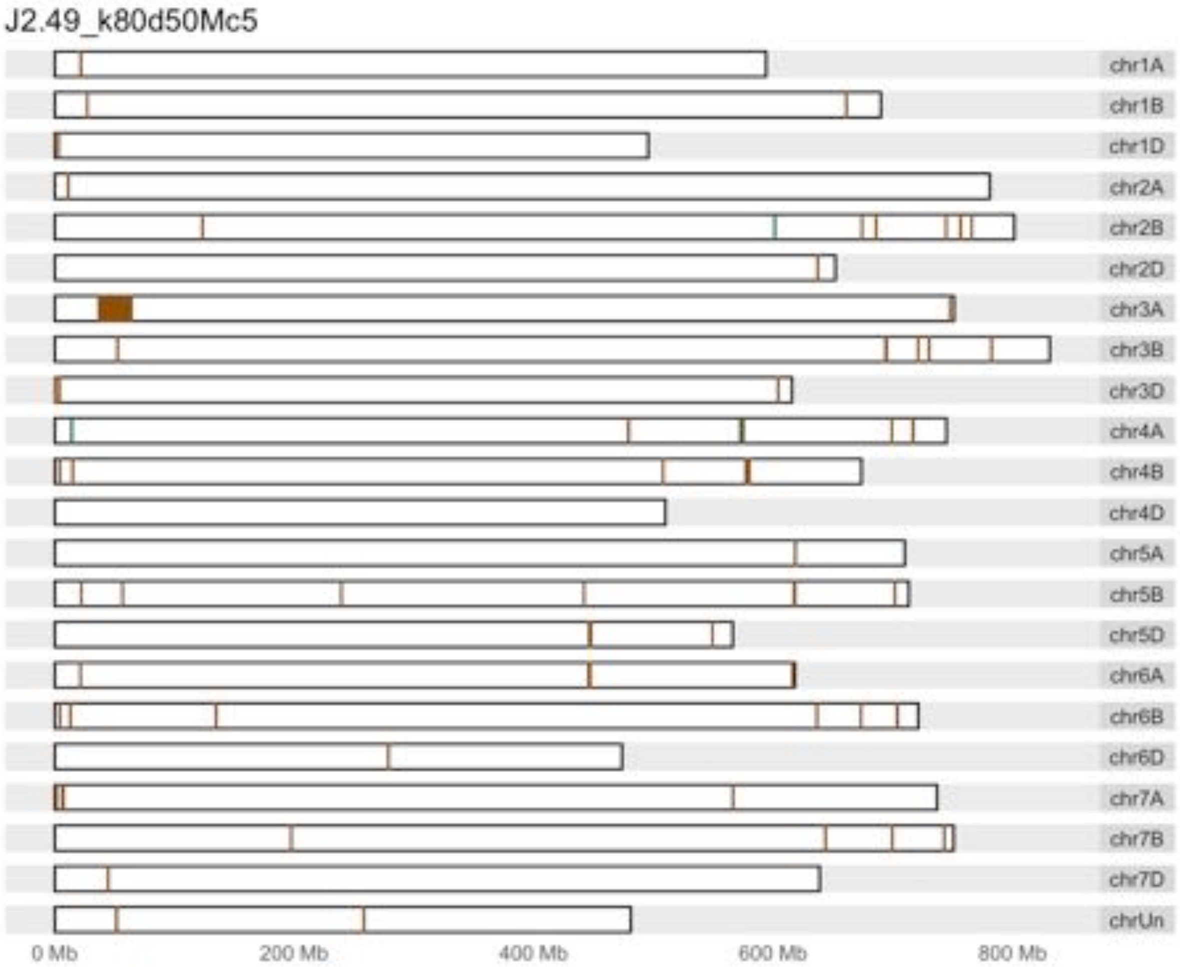
Primer design for deletions

DELETTIONS γ-RADIATION LINES

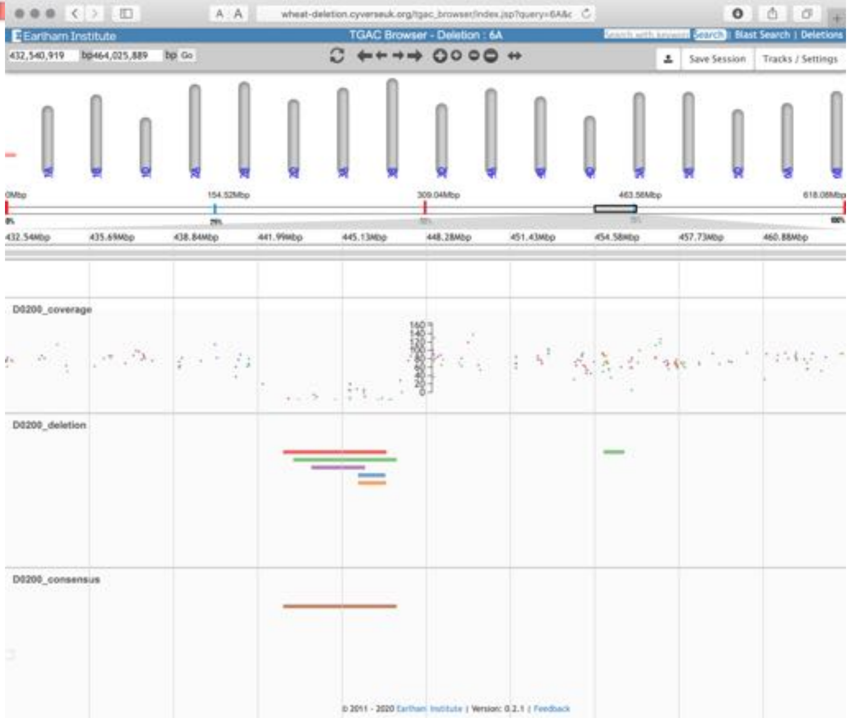
- ~600 Paragon lines with radiation induced deletions
- Identify lines with missing genes to test changes in phenotype
- Low sequencing coverage (0.3x)
- Large window sizes to make up for low coverage
- Collaboration with Paul Nichlson lab.

DELETTIONS γ-RADIATION LINES



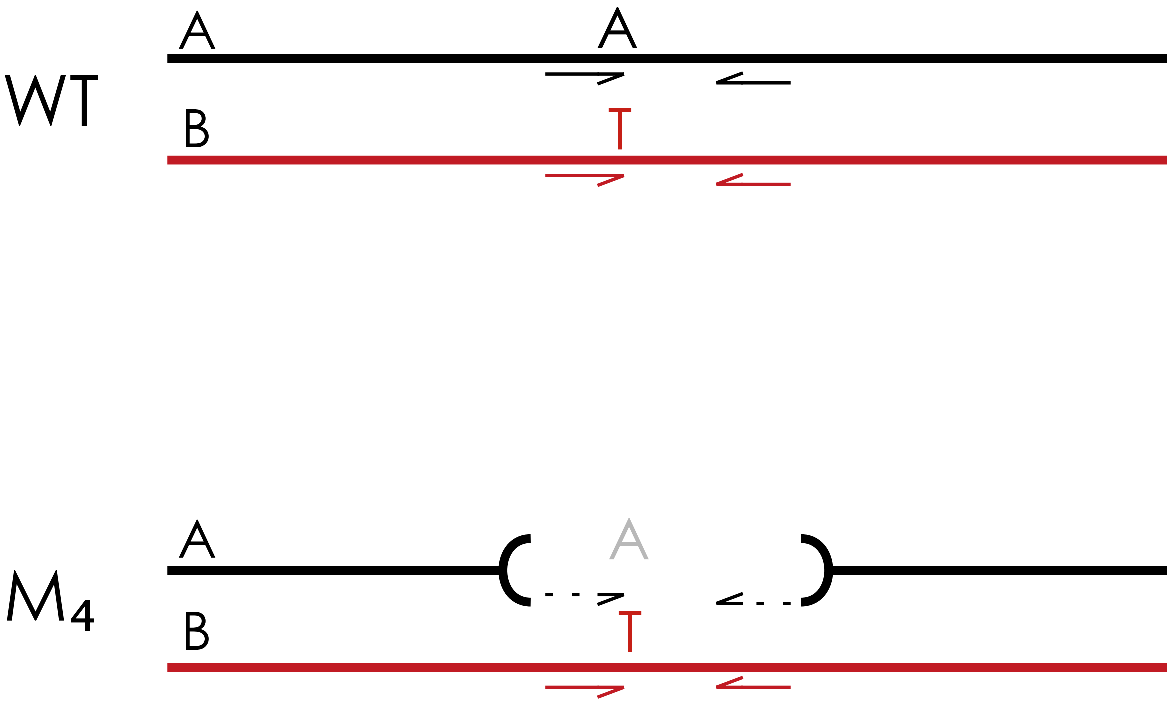


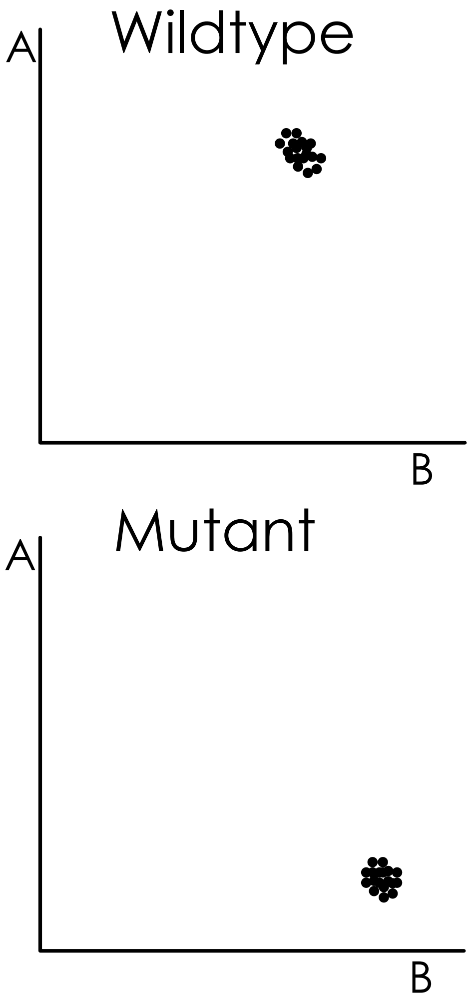
Deletions Browser



http://wheat-deletion.cyverseuk.org/

SNP markers to detect deletions





Deletions primer design

SNP-1 B	cgcatttAcgcgCgcgataccggcgcctTtgAgaatatttgcagcgaaggcgtg
IWGSC-1A	cgcattt**A**cgcgtgcgataccggcgcct**G**tg**A**gaatatttgcagcgaaggcgtg
IWGSC-1B	cgcatttAcgcgtgcgataccggcgcctTtgGgaatatttgcagcgaaggcgtg
IWGSC-1D	c--atttGcgcgtgcgataccggcgcctGtgGgaatatttgcagcgaaggcgtg

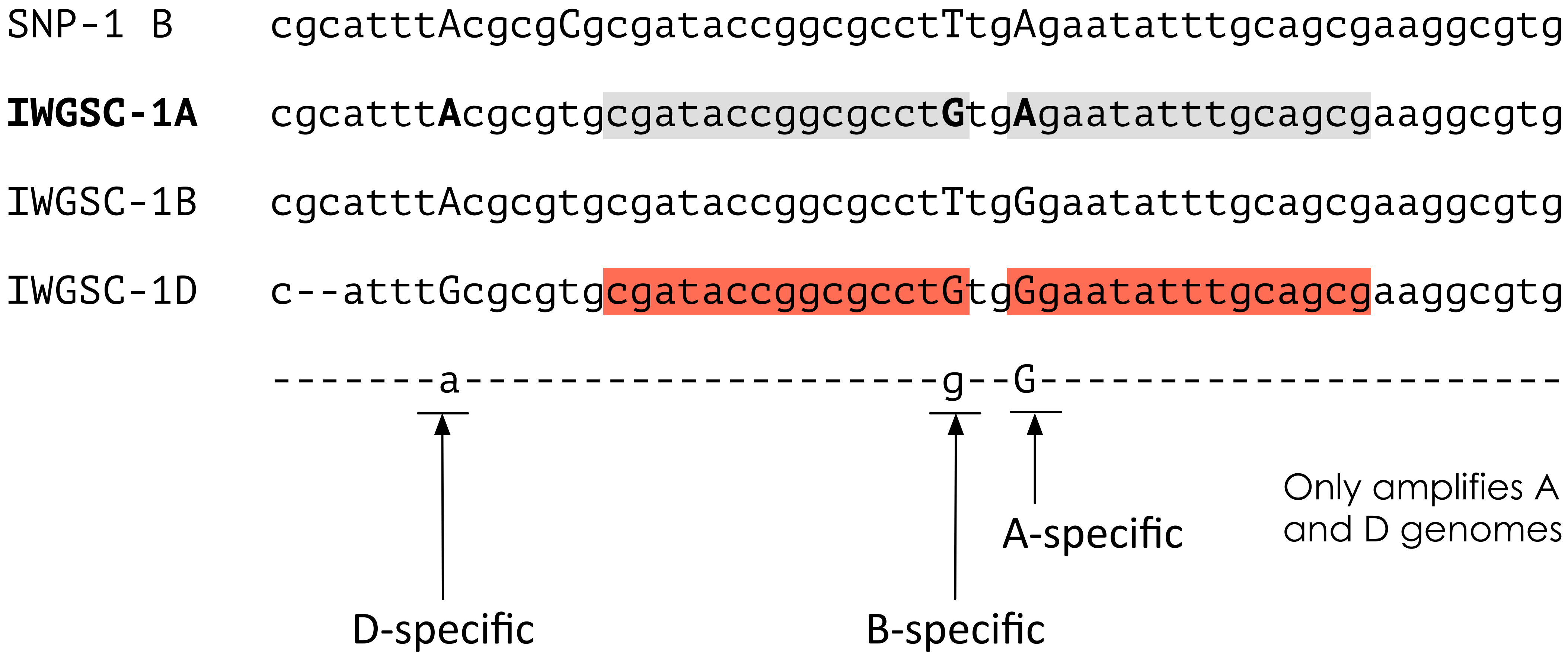
Deletions primer design

```
SNP-1 B    cgcatttAcgcgCgcgataccggcgcctTtgAgaatatttgcagcgaaggcgtg

IWGSC-1A   cgcatttAcgcgtgcgataccggcgcctGtgAgaatatttgcagcgaaggcgtg

IWGSC-1B   cgcatttAcgcgtgcgataccggcgcctTtgGgaatatttgcagcgaaggcgtg

IWGSC-1D   c--atttGcgcgtgcgataccggcgcctGtgGgaatatttgcagcgaaggcgtg

           -------a--------------------g--G----------------------
```

Deletions primer design

```
SNP-1 B     cgcatttAcgcgCgcgataccggcgcctTtgAgaatatttgcagcgaaggcgtg

IWGSC-1A    cgcatttAcgcgtgcgataccggcgcctGtgAgaatatttgcagcgaaggcgtg

IWGSC-1B    cgcatttAcgcgtgcgataccggcgcctTtgGgaatatttgcagcgaaggcgtg

IWGSC-1D    c--atttGcgcgtgcgataccggcgcctGtgGgaatatttgcagcgaaggcgtg

            -------a--------------------g--G----------------------
```

Deletions primer design



```
SNP-1 B     cgcatttAcgcgCgcgataccggcgcctTtgAgaatatttgcagcgaaggcgtg

IWGSC-1A    cgcatttAcgcgtgcgataccggcgcctGtgAgaatatttgcagcgaaggcgtg

IWGSC-1B    cgcatttAcgcgtgcgataccggcgcctTtgGgaatatttgcagcgaaggcgtg

IWGSC-1D    c--atttGcgcgtgcgataccggcgcctGtgGgaatatttgcagcgaaggcgtg

            -------a--------------------g--G----------------------
```

A-specific

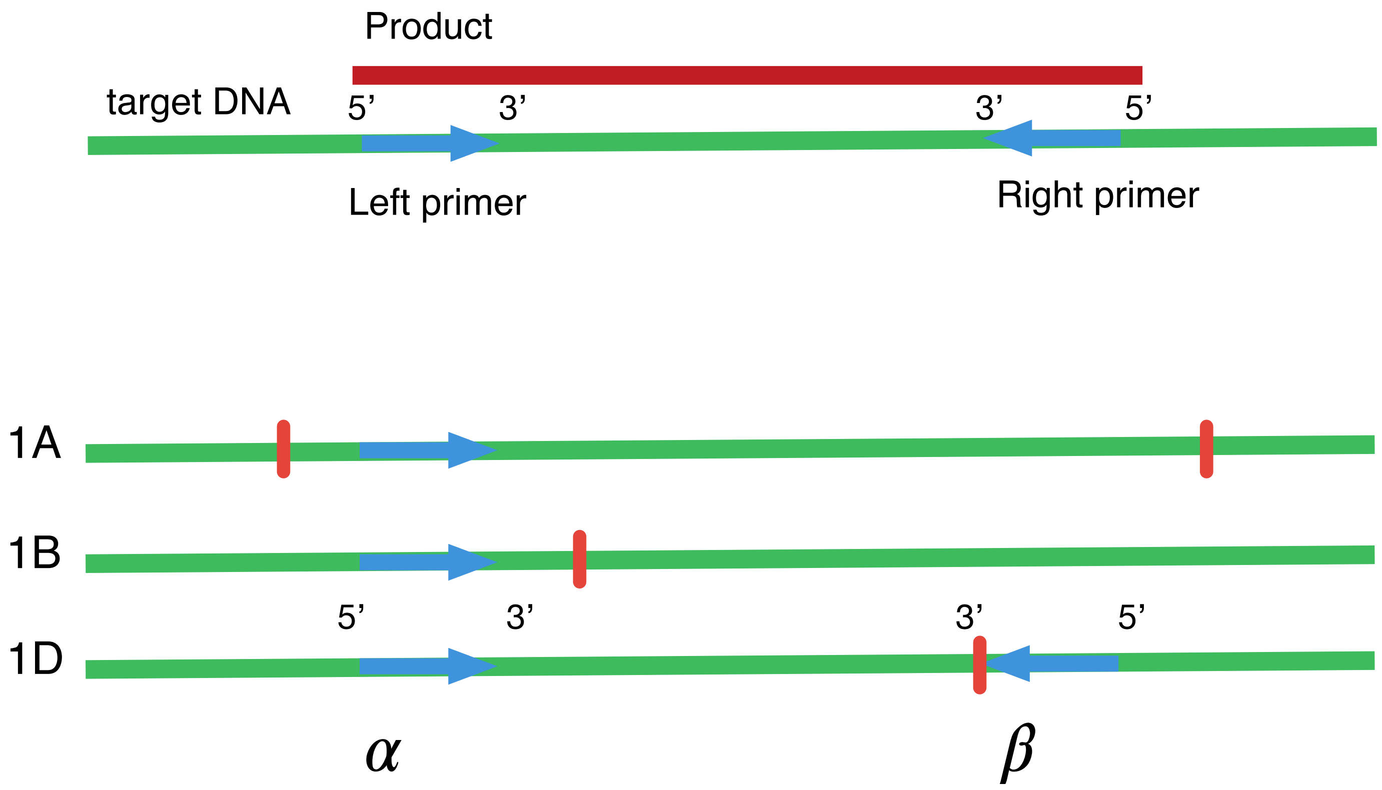
D-specific

B-specific

Only amplifies A and D genomes

Amplicon primers and regular PCR primers

Amplicon primers



Amplicon primers

Region 1B	cgcatttAcgcgCgcgataccggcgcctTtgAgaatatttgcagcgaaggcgtg
IWGSC-1A	cgcatttAcgcgtgcgataccggcGcctTtgAgaatatttgcagcgaaggcgtg
IWGSC-1B	cgcattt**A**cgcgtgcgataccggc**A**cct**G**tg**G**gaatatttgcagcgaaggcgtg
IWGSC-1D	c--atttGcgcgtgcgataccggcGcctGtgAgaatatttgcagcgaaggcgtg

Amplicon primers

```
Region 1B  cgcatttAcgcgCgcgataccggcgcctTtgAgaatatttgcagcgaaggcgtg

 IWGSC-1A  cgcatttAcgcgtgcgataccggcGcctTtgAgaatatttgcagcgaaggcgtg

 IWGSC-1B  cgcatttAcgcgtgcgataccggcAcctGtgGgaatatttgcagcgaaggcgtg

 IWGSC-1D  c--atttGcgcgtgcgataccggcGcctGtgAgaatatttgcagcgaaggcgtg

           -------a----------------A---g--g----------------------
```

Amplicon primers

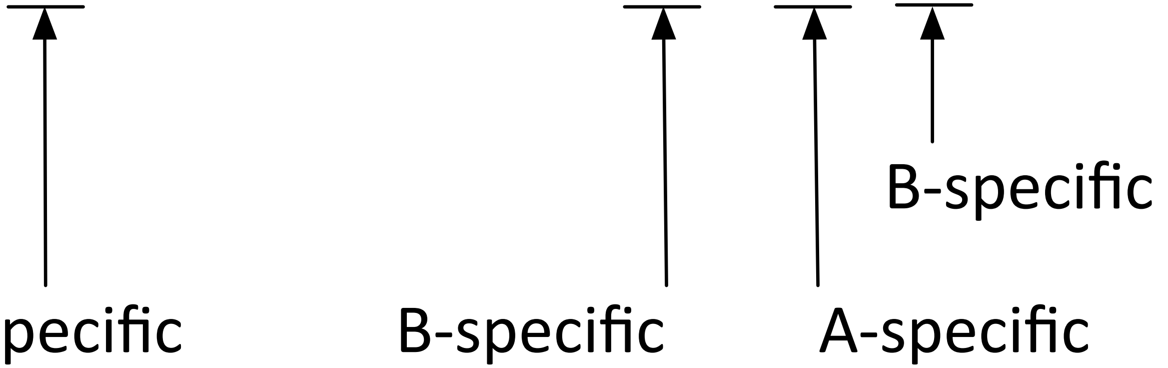
Region 1B cgcatttAcgcgCgcgataccggcgcctTtgAgaatatttgcagcgaaggcgtg

IWGSC-1A cgcatttAcgcgtgcgataccggcGcctTtgAgaatatttgcagcgaaggcgtg

IWGSC-1B cgcattt**A**cgcgtgcgataccggc**A**cct**G**tg**G**gaatatttgcagcgaaggcgtg

IWGSC-1D c--atttGcgcgtgcgataccggcGcctGtgAgaatatttgcagcgaaggcgtg

-------a----------------A---g--g----------------------



B-specific

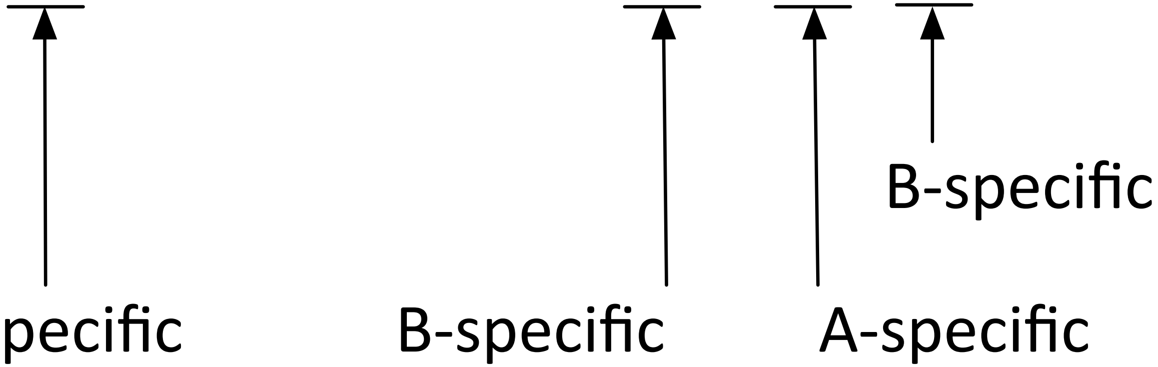
D-specific

B-specific

A-specific

Amplicon primers

```
Region 1B  cgcatttAcgcgCgcgataccggcgcctTtgAgaatatttgcagcgaaggcgtg

 IWGSC-1A  cgcatttAcgcgtgcgataccggcGcctTtgAgaatatttgcagcgaaggcgtg

 IWGSC-1B  cgcatttAcgcgtgcgataccggcAcctGtgGgaatatttgcagcgaaggcgtg

 IWGSC-1D  c--atttGcgcgtgcgataccggcGcctGtgAgaatatttgcagcgaaggcgtg

           -------a----------------A---g--g----------------------
```



B-specific

D-specific

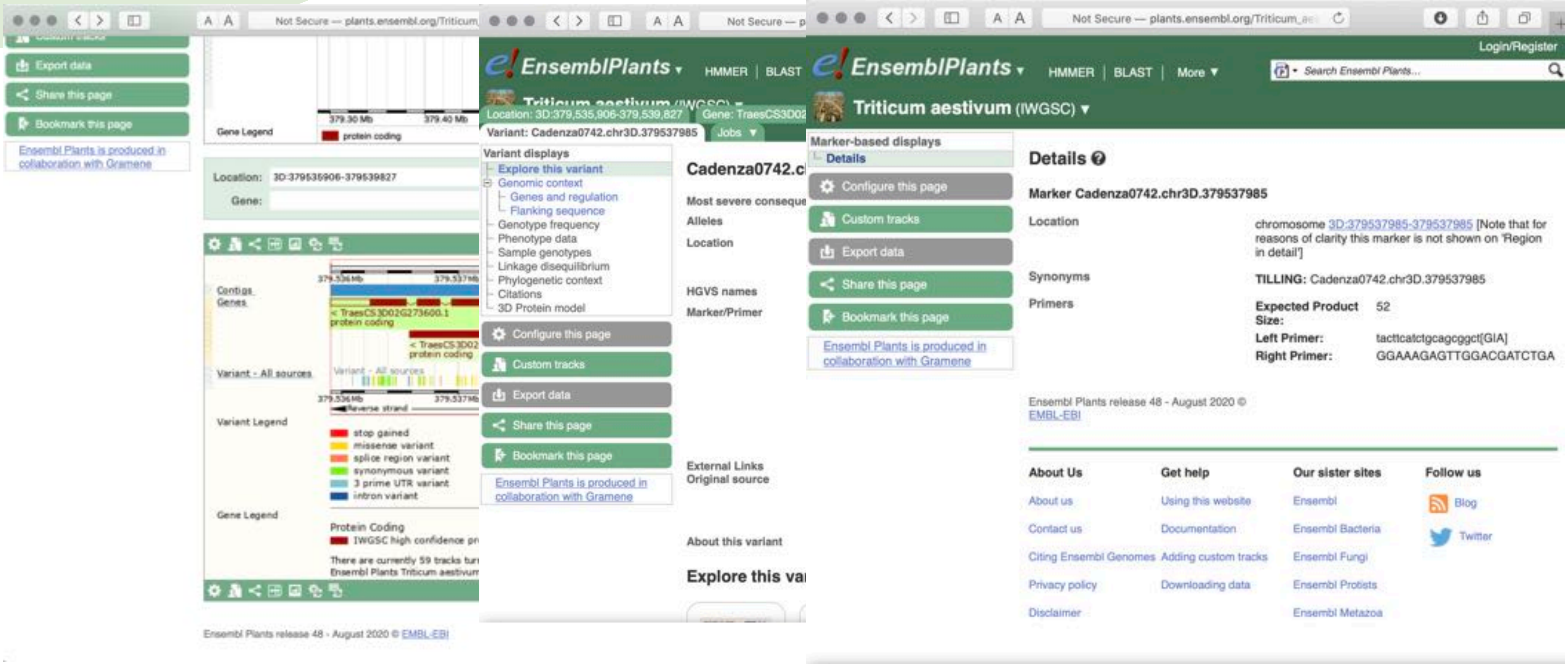
B-specific

A-specific

PolyMarker used for…

http://plants.ensembl.org/Triticum_aestivum

Tilling mutants in ensembl plants



Location: 3D:379535906-379539827

Gene:

There are currently 59 tracks turn

Ensembl Plants Triticum aestivum

Ensembl Plants release 48 - August 2020 © EMBL-EBI

Location: 3D:379,535,906-379,539,827

Gene: TraesCS3D02

Variant: Cadenza0742.chr3D.379537985

Cadenza0742.c

Most severe conseque

Alleles

Location

HGVS names

Marker/Primer

External Links

Original source

About this variant

Explore this va

Details

Marker Cadenza0742.chr3D.379537985

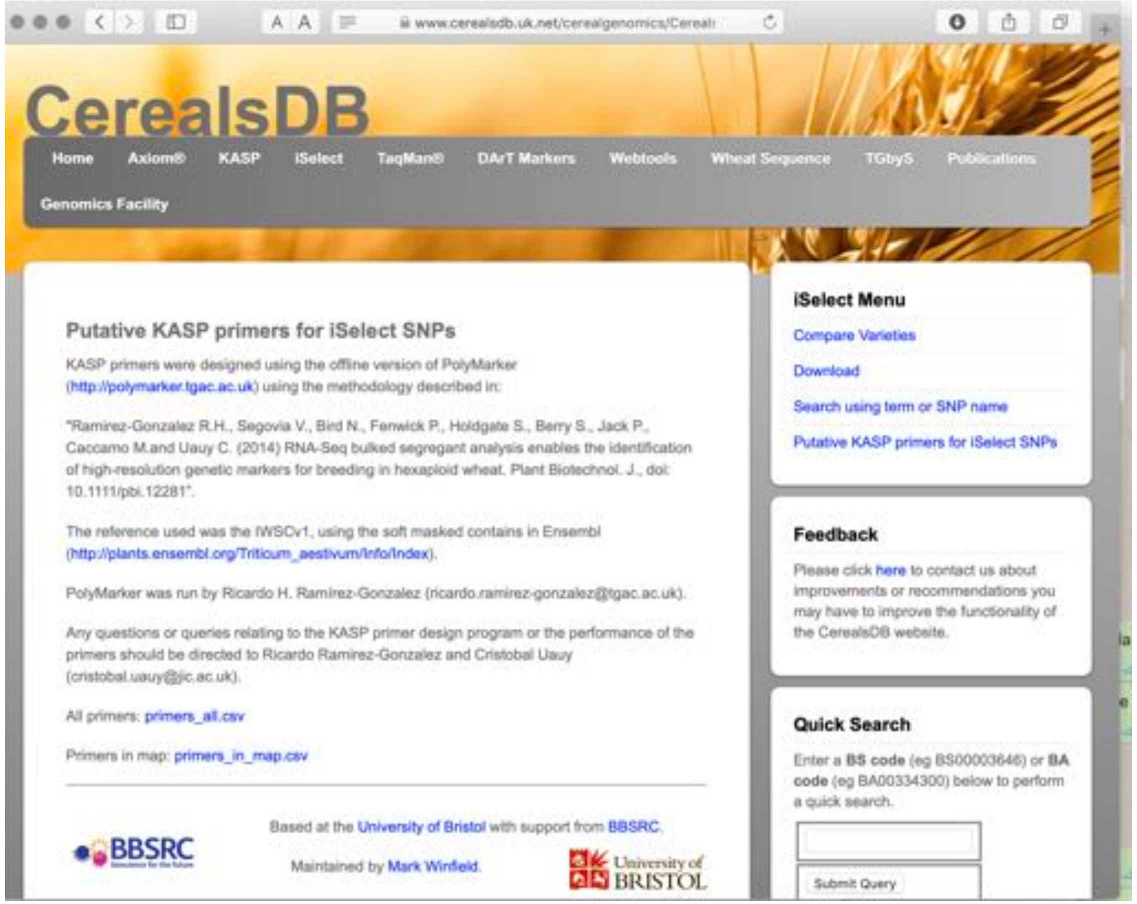
Location	chromosome 3D:379537985-379537985 [Note that for reasons of clarity this marker is not shown on 'Region in detail']
Synonyms	**TILLING:** Cadenza0742.chr3D.379537985
Primers	**Expected Product Size:** 52 **Left Primer:** tacttcatctgcagcggct[G/A] **Right Primer:** GGAAAGAGTTGGACGATCTGA

Ensembl Plants release 48 - August 2020 © EMBL-EBI

Krasileva, K. V, et al (2017). Uncovering hidden variation in polyploid wheat. PNAS, 114(6), E913-E921. doi:10.1073/pnas.1619268114

Markers in CeralsDB

- Primers for iSelect 80K chip
- Primers for Axiom 820K chip



Putative KASP primers for iSelect SNPs

KASP primers were designed using the offline version of PolyMarker (http://polymarker.tgac.ac.uk) using the methodology described in:

"Ramirez-Gonzalez R.H., Segovia V., Bird N., Fenwick P., Holdgate S., Berry S., Jack P., Caccamo M.and Uauy C. (2014) RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. Plant Biotechnol. J., doi: 10.1111/pbi.12281".

The reference used was the IWSCv1, using the soft masked contains in Ensembl (http://plants.ensembl.org/Triticum_aestivum/Info/Index).

PolyMarker was run by Ricardo H. Ramirez-Gonzalez (ricardo.ramirez-gonzalez@tgac.ac.uk).

Any questions or queries relating to the KASP primer design program or the performance of the primers should be directed to Ricardo Ramirez-Gonzalez and Cristobal Uauy (cristobal.uauy@jic.ac.uk).

All primers: primers_all.csv

Primers in map: primers_in_map.csv

Based at the University of Bristol with support from BBSRC.

Maintained by Mark Winfield.

iSelect Menu

Compare Varieties

Download

Search using term or SNP name

Putative KASP primers for iSelect SNPs

Feedback

Please click here to contact us about improvements or recommendations you may have to improve the functionality of the CerealsDB website.

Quick Search

Enter a **BS code** (eg BS00003646) or **BA code** (eg BA00334300) below to perform a quick search.

https://www.cerealsdb.uk.net/cerealgenomics/CerealsDB/KASP_primers_for_iSelect.php

Use your own reference

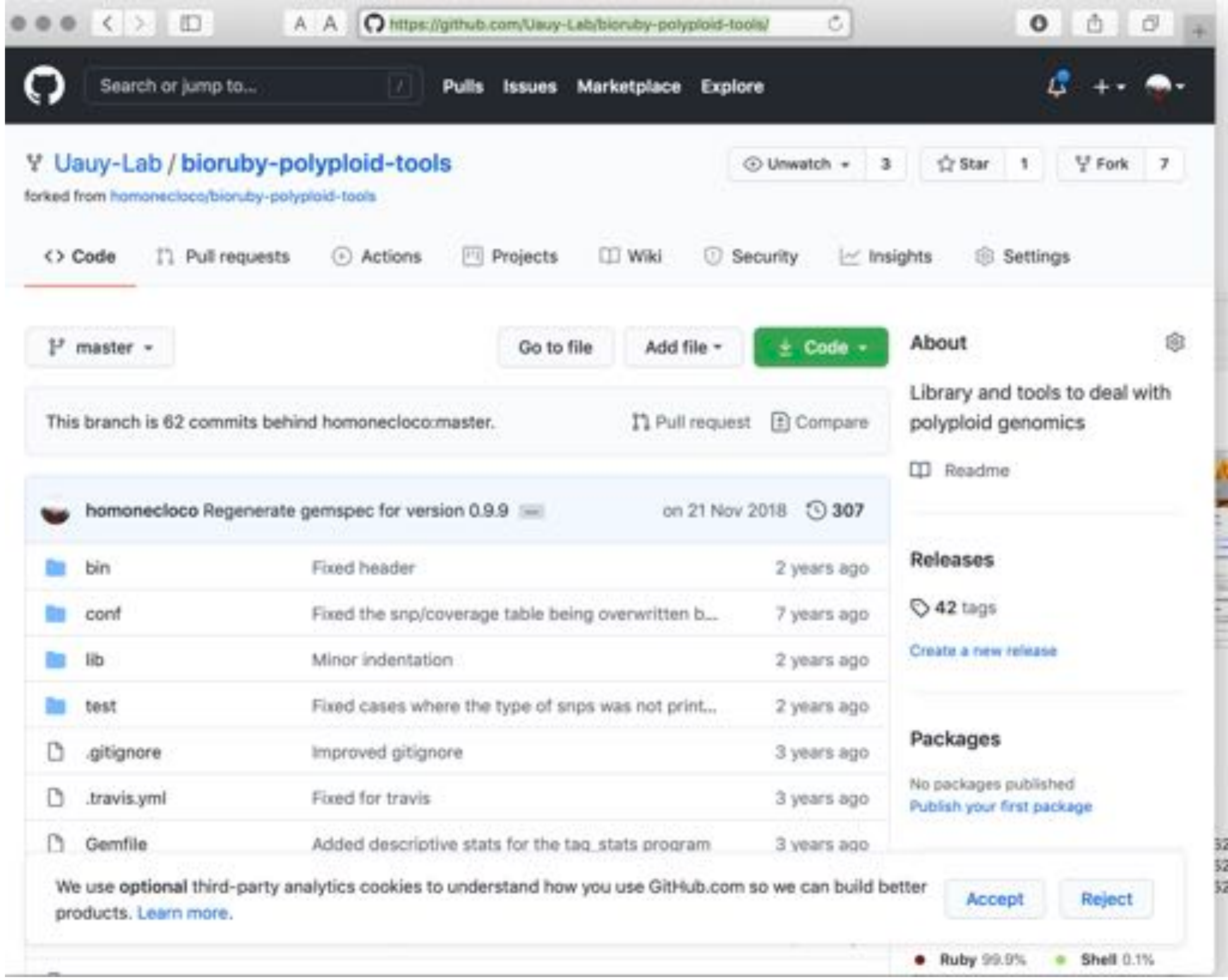
- PolyMarker is open source!

- Run from the command line:

https://github.com/Uauy-Lab/bioruby-polyploid-tools/

- Install your own web service:

https://github.com/Uauy-Lab/pangenome-web



Uauy-Lab / bioruby-polyploid-tools

forked from homonecloco/bioruby-polyploid-tools

This branch is 62 commits behind homonecloco:master.

homonecloco Regenerate gemspec for version 0.9.9 on 21 Nov 2018 307

bin	Fixed header	2 years ago
conf	Fixed the snp/coverage table being overwritten b...	7 years ago
lib	Minor indentation	2 years ago
test	Fixed cases where the type of snps was not print...	2 years ago
.gitignore	Improved gitignore	3 years ago
.travis.yml	Fixed for travis	3 years ago
Gemfile	Added descriptive stats for the tag_stats program	3 years ago

About

Library and tools to deal with polyploid genomics

Releases

42 tags

Packages

No packages published

We use **optional** third-party analytics cookies to understand how you use GitHub.com so we can build better products. Learn more.

Future developments

- Do further optimisations on the amplicon and deletion markers.
- Integration with the 10+ Wheat genome pangenome
- Make the amplicon and deleltion versions of polymarkers
- Other applications that the community may need?

Acknowledgments

- Deletions
  - Ben Hales, Paul Nicholson, Nikolai Adamski, Anil Thanki
- Ensembl plants
  - Guy Naamati
- CerealsDB
  - Mark Winfield.
- Uauy Lab

Resources

- Website: http://polymarker.info
- Source code: https://github.com/Uauy-Lab/bioruby-polyploid-tools/
- Ensembl plants: http://plants.ensembl.org/Triticum_aestivum
- Gamma radiation deletions database: http://wheat-deletion.cyverseuk.org/

Ramirez-Gonzalez, R. H., Uauy, C., & Caccamo, M. (2015). *PolyMarker: A fast polyploid primer design pipeline*. Bioinformatics, 2–3. doi:10.1093/bioinformatics/btv069