Identification of candidate genes for fertility restoration of WA-CMS through mapping and sequencing

C N Neeraja
P Balaji Suresh
B C Viraktamath

Directorate of Rice Research
Hyderabad, India
cnneeraja@gmail.com

Objective

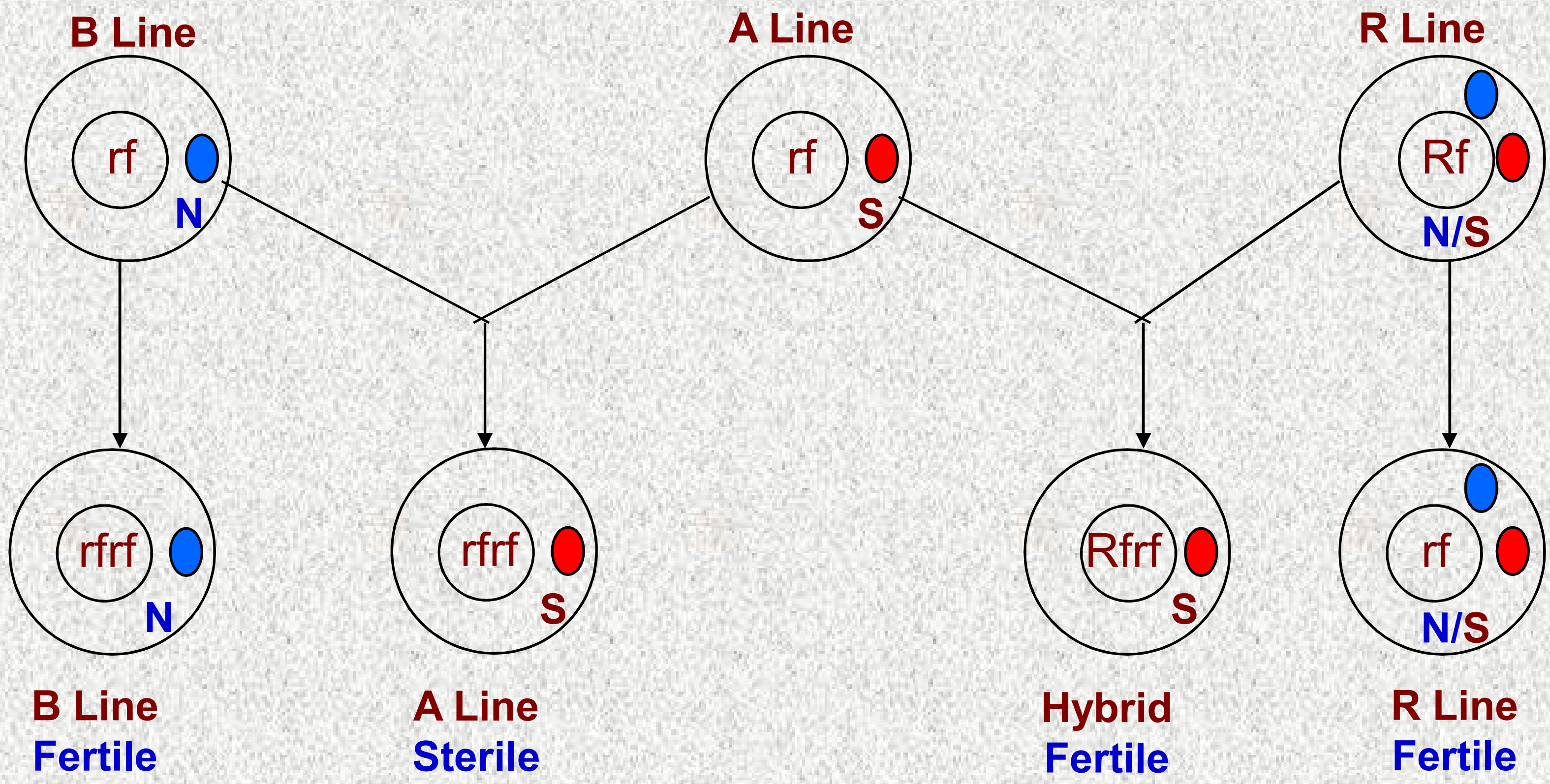
To develop marker system for *Rf* genes for WA-CMS

Outline

- Need of markers
- Genetics
- Mapping
- Bioinformatics
- Problems and Possibilities

Three line approach – Wild Abortive system

Cytoplasmic male-sterile (A line), maintainer (B line) and restorer (R line)



B Line

rf

N

A Line

rf

S

R Line

Rf

N/S

rfrf

N

rfrf

S

Rfrf

S

rf

N/S

B Line
Fertile

A Line
Sterile

Hybrid
Fertile

R Line
Fertile

To identify the potential restorers for WA-CMS

CMS (A Line) X Several genotypes

↓

F1

Evaluate the F_1 progenies for pollen and spikelet fertility

Fully fertile : Restorer, Fully sterile: Maintainer, Partial: Discard

- Two seasons for crossing and evaluation
- Environment dependent
- Efficiency is ~20%

Marker Assisted Selection (MAS) for restorer genes

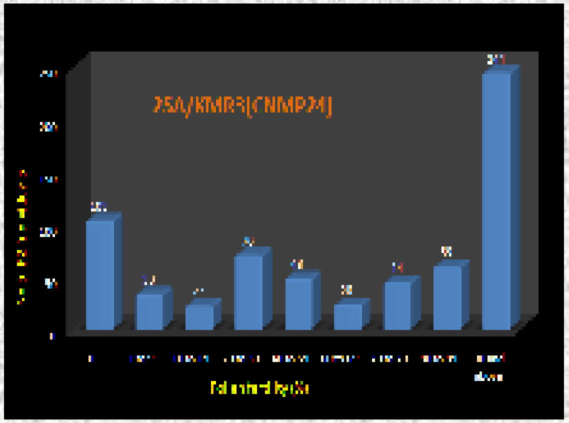
Genetic studies of fertility restoration of WA-CMS

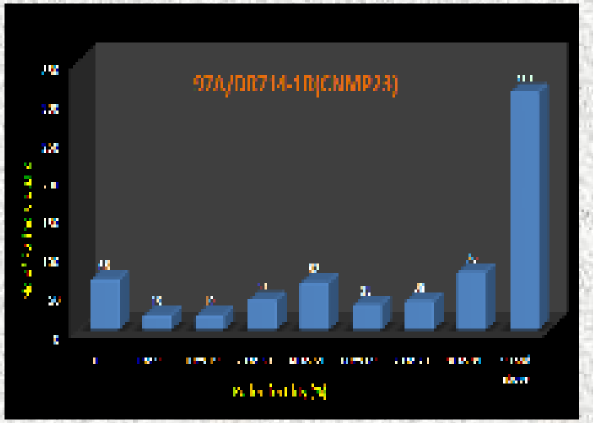
- -monogenic
- -digenic
- -digenic with different types of interactions
- -trigenic
- -trigenic interactions

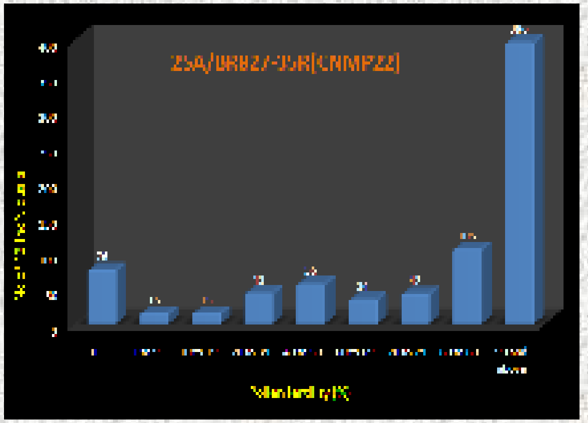
Classification of phenotype

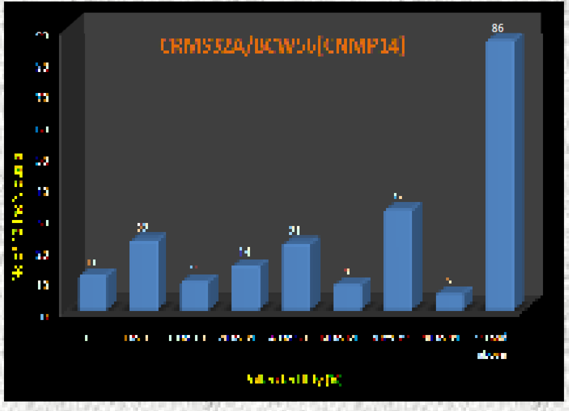
IRRI SES	Continuous
Two classes	>1%
	>40%
Three classes	0%
	1-30% + 30-60%
	60-100%
Four classes	<10%
	10-70%
	70-85%
	>85%

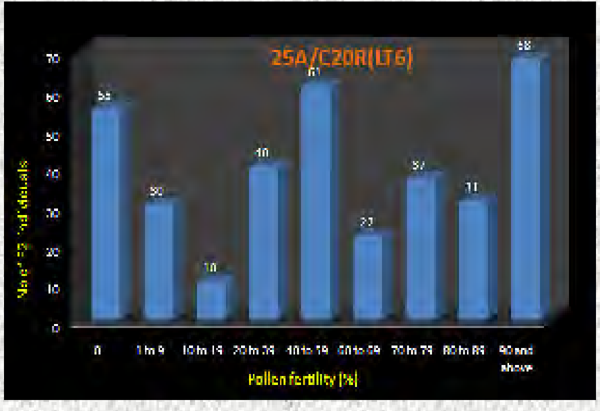
Segregation of pollen fertility in six F_2 mapping populations









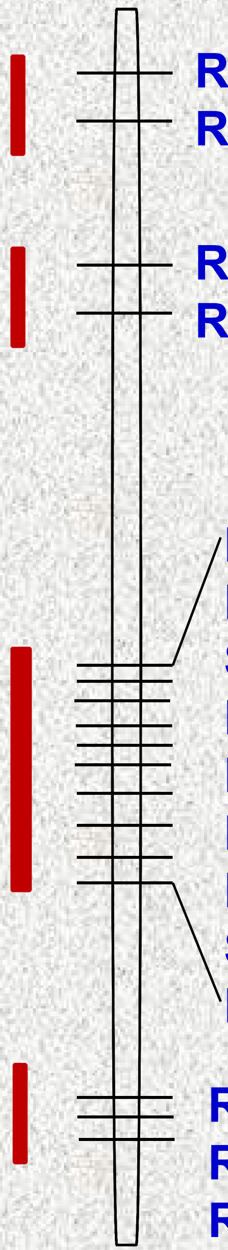


the involvement of major QTL /more than three genes with minor QTL

Mapping of fertility restorer (*Rf*) loci for WA-CMS

A line	R line	Position	Reference
Gui630	02428	Chrs 2,3,4,5	Zhu et al 1996
-	-	Chr1	Shen et al 1996
Zhenshan A	MingHui 63	Chr1, Chr10	Yao et al 1997
Zhenshan A	IR24	Chr 1	Zhang et al 1997
IR58025A	IR36	Chr 1	Zhang et al 1997
RD21A	IR24	Chr10	Tan et al 1998
ZhenshanA	Milyang46	Chrs 1, 7, 10, 11	Zhuang et al 2000
Zhenshan A	IR24	Chr10	Jing et al 2001
Zhenshan A	MingHui 63	Chr10	He et al 2002
IR58025A	IR24, PRR78	Chr10	Mishra et al 2003
Neda A	Amol1	Chr10	Ahmadikhah & Karlov 2006
Neda A	IR62030	Chr10	Ahmadikhah et al. 2007;
IR58025A	IR42686	Chrs 1,7,10,12	Bazarkar et al 2008
IR58025A	IR34686	Chr1, Chr10	Sattari et al 2008
Neda A	IR36	Chr 1	Ahmadikhah & Alavi 2009
IR58025A	KMR3	Chr10	Sheeba et al 2009
Pusa 6A	PRR78	Chr10	Ngangkham 2010
Neda A	DN33-18	Chr1, Chr10	Nematzadeh & Kiani 2010

Rf4 locus



RM244 (0.9 Mb)
RM2309 (1.93)

RM311 (9.48 Mb)
RG257 (12.04)

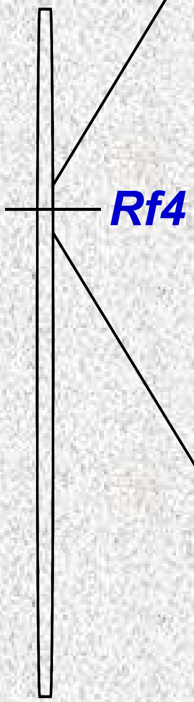
RM269 (18.48 Mb)
RM258 (18.54)
S11148 (18.84)
RM304 (19.18)
RM6100 (19.2)
RM6737 (19.24)
RM171 (19.57)
S10019 (19.58)
RM1108 (19.69 Mb)

RM3128 (22.29 Mb)
RM228 (22.76)
RM591 (23.42)

TMPPR3

Rf4

Nipponbare



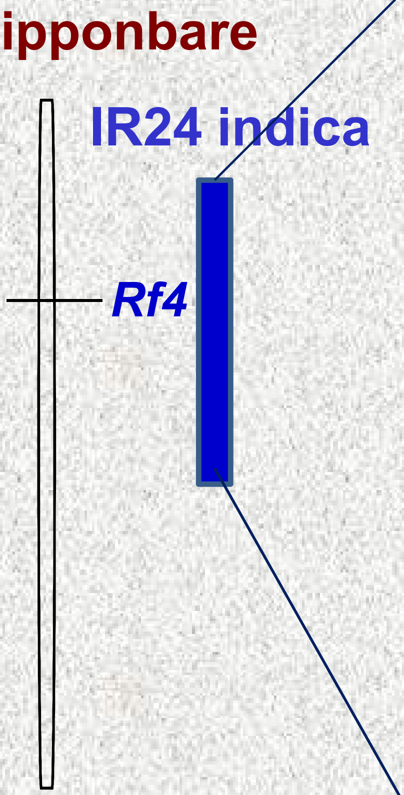
RM25626 18 Mb
RM25636
RM25637
RM25639
RM25640
RM25641
RM25642
RM25649
RM25650
RM25653
RM25654
RM25656
RM25657
RM25659
RM25660
RM25661
RM25663
RM25665
RM25669
RM25672
RM25675
RM25679
RM25681
RM25699 20 Mb

DRRM-RF4-1
DRRM-RF4-2
DRRM-RF4-3
DRRM-RF4-4
DRRM-RF4-5
DRRM-RF4-6
DRRM-RF4-7
DRRM-RF4-8
DRRM-RF4-9
DRRM-RF4-10
DRRM-RF4-11
DRRM-RF4-12
DRRM-RF4-13
DRRM-RF4-14
DRRM-RF415
DRRM-RF4-16
DRRM-RF4-17
DRRM-RF4-18
DRRM-RF4-19
DRRM-RF4-20
DRRM-RF4-21
DRRM-RF4-22
DRRM-RF4-23
DRRM-RF4-24
DRRM-RF4-25
DRRM-RF4-26
DRRM-RF4-27
DRRM-RF4-28
DRRM-RF4-29

DRRM-RF4-30
DRRM-RF4-31
DRRM-RF4-32
DRRM-RF4-33
DRRM-RF4-34
DRRM-RF4-35
DRRM-RF4-36
DRRM-RF4-37
DRRM-RF4-38
DRRM-RF4-39
DRRM-RF4-40
DRRM-RF4-41
DRRM-RF4-42
DRRM-RF4-43
DRRM-RF4-44
DRRM-RF4-45
DRRM-RF4-46
DRRM-RF4-47
DRRM-RF4-48
DRRM-RF4-49
DRRM-RF4-50
DRRM-RF4-51
DRRM-RF4-52
DRRM-RF4-53
DRRM-RF4-54
DRRM-RF4-55
DRRM-RF4-56
DRRM-RF4-57
DRRM-RF4-58

23 RM markers
58 Developed SSRs
10 polymorphic markers/81

Rf4

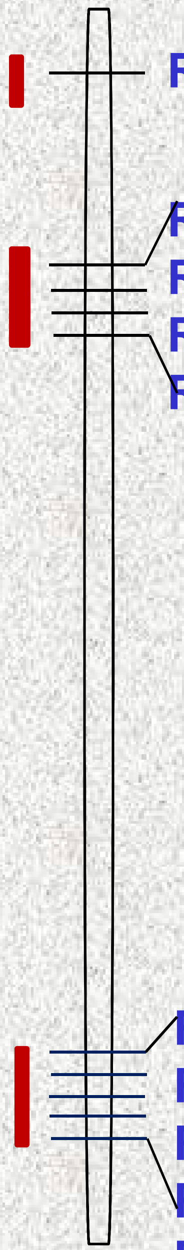


DRRM-RF4-59
DRRM-RF4-60
DRRM-RF4-61
DRRM-RF4-62
DRRM-RF4-63
DRRM-RF4-64
DRRM-RF4-65
DRRM-RF4-66
DRRM-RF4-67
DRRM-RF4-68
DRRM-RF4-69
DRRM-RF4-70
DRRM-RF4-71
DRRM-RF4-72
DRRM-RF4-73
DRRM-RF4-74
DRRM-RF4-75
DRRM-RF4-76
DRRM-RF4-77
DRRM-RF4-78
DRRM-RF4-79
DRRM-RF4-80
DRRM-RF4-81
DRRM-RF4-82
DRRM-RF4-83
DRRM-RF4-84
DRRM-RF4-85
DRRM-RF4-86
DRRM-RF4-87
DRRM-RF4-88
DRRM-RF4-89
DRRM-RF4-90
DRRM-RF4-91
DRRM-RF4-92
DRRM-RF4-93
DRRM-RF4-94
DRRM-RF4-95
DRRM-RF4-96
DRRM-RF4-97
DRRM-RF4-98
DRRM-RF4-99
DRRM-RF4-100
DRRM-RF4-101
DRRM-RF4-102
DRRM-RF4-103
DRRM-RF4-104
DRRM-RF4-105
DRRM-RF4-106
DRRM-RF4-107
DRRM-RF4-108

50 Developed SSRs

5 polymorphic markers/50

Rf3



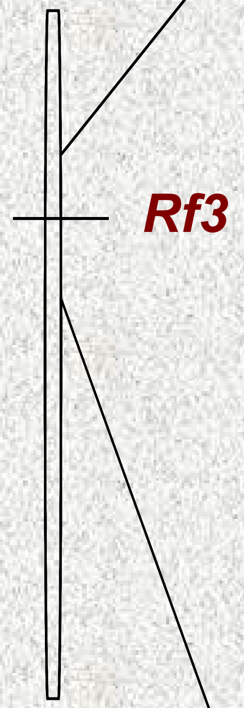
RM3148 (0.74 Mb)

RM1 (4.63 Mb)
RM283 (4.88)
RG140 (5.09)
RG532 (5.79)

RM443 (28.33 Mb)
RM7180 (34.09)
RG394 (34.26)
RG374 (35.22)
RM315 (38.49)

Rf3

Nipponbare



RM10265
RM10277
RM10279
RM10282
RM10283
RM10285
RM10287
RM10289
RM10295
RM10296
RM10298
RM10299
RM10300
RM10301
RM10302
RM10303
RM10305
RM10306
RM10310
RM10311
RM10312
RM10313
RM10314
RM10315

DRRM-RF3-1
DRRM-RF3-2
DRRM-RF3-3
DRRM-RF3-4
DRRM-RF3-5
DRRM-RF3-6
DRRM-RF3-7
DRRM-RF3-8
DRRM-RF3-9
DRRM-RF3-10
DRRM-RF3-11
DRRM-RF3-12
DRRM-RF3-13
DRRM-RF3-14
DRRM-RF3-15
DRRM-RF3-16
DRRM-RF3-17
DRRM-RF3-18
DRRM-RF3-19
DRRM-RF3-20
DRRM-RF3-21
DRRM-RF3-22

DRRM-RF3-23
DRRM-RF3-24
DRRM-RF3-25
DRRM-RF3-26
DRRM-RF3-27
DRRM-RF3-28
DRRM-RF3-29
DRRM-RF3-30
DRRM-RF3-31
DRRM-RF3-32
DRRM-RF3-33
DRRM-RF3-34
DRRM-RF3-35
DRRM-RF3-36
DRRM-RF3-37
DRRM-RF3-38
DRRM-RF3-39
DRRM-RF3-40
DRRM-RF3-41

DRRM-RF3-42
DRRM-RF3-43
DRRM-RF3-44
DRRM-RF3-45
DRRM-RF3-46
DRRM-RF3-47
DRRM-RF3-48
DRRM-RF3-49
DRRM-RF3-50
DRRM-RF3-51
DRRM-RF3-52
DRRM-RF3-53
DRRM-RF3-54
DRRM-RF3-55
DRRM-RF4-56

29 RM markers
56 Developed SSRs
~30 polymorphic
/85 markers

Candidate gene primers

Rf4

Four candidate genes from IR24 – 21 primers

pentatricopeptide (PPR) genes

PPR 794

PPR 683 - Polymorphic

PPR 762 - Polymorphic

PPR 791

Rf3

Four candidate genes from Nipponbare – 23 primers

Mitochondrial-processing peptidase subunit alpha (LOC_Os01g09560)

Pollen-specific protein (LOC_Os01g09670)

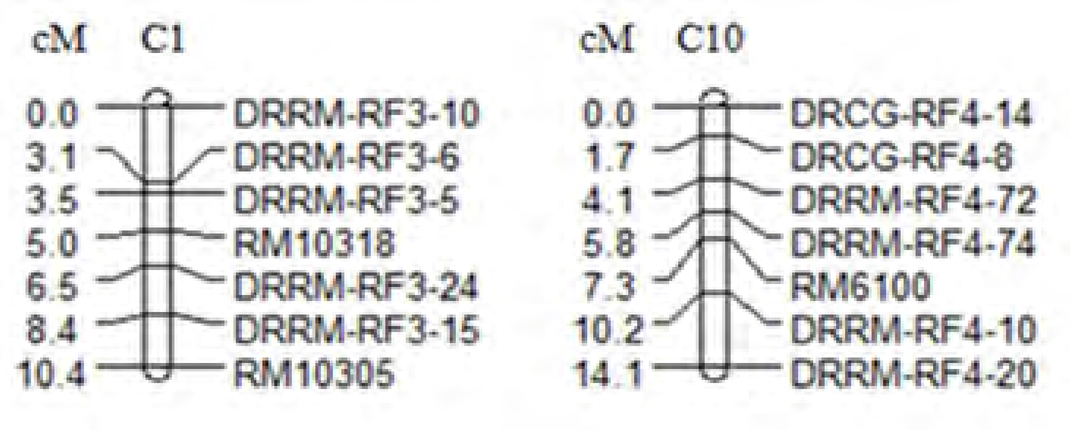
Pentatricopeptide (PPR) genes (LOC_Os01g10090)

Pentatricopeptide (PPR) genes (LOC_Os01g10800)

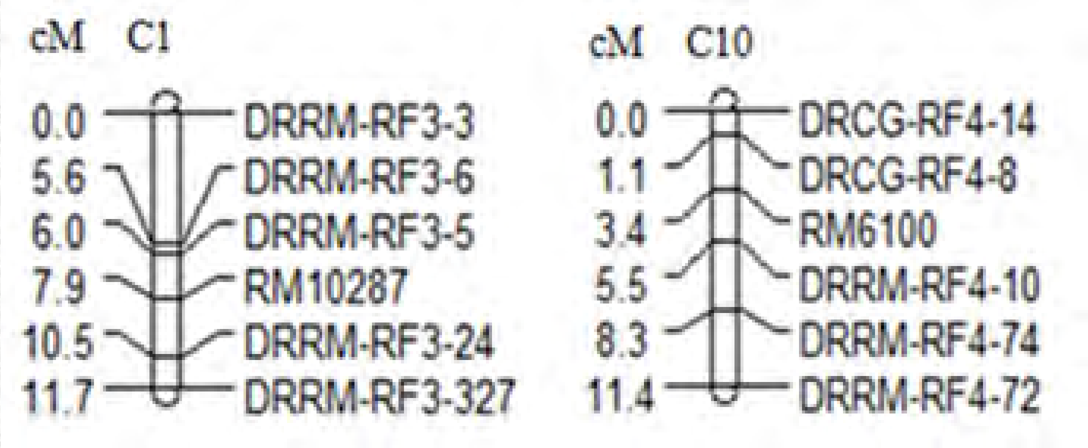
Fine mapping of *Rf3* and *Rf4* loci

Selective genotyping of 48 individuals from each tail group 34 (*Rf3*) and 26 (*Rf4*) markers

KRH2 – F_2 mapping population (~1500) 7 markers for both *Rf3* & *Rf4*

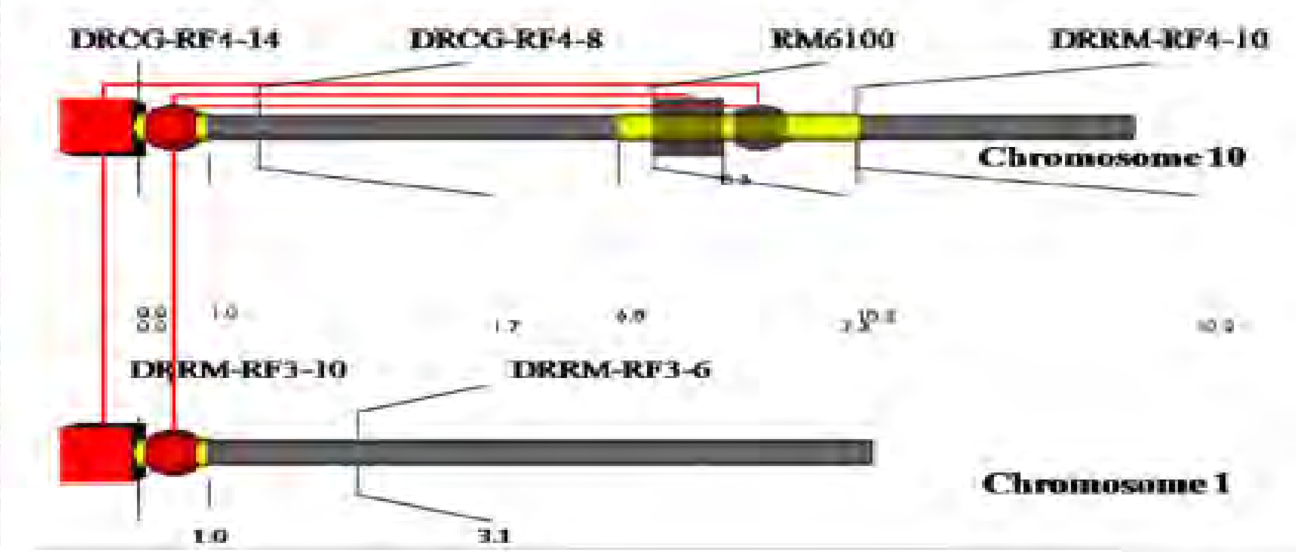


DRRH2 - F_2 mapping population (~1500) 6 markers for both *Rf3* & *Rf4*

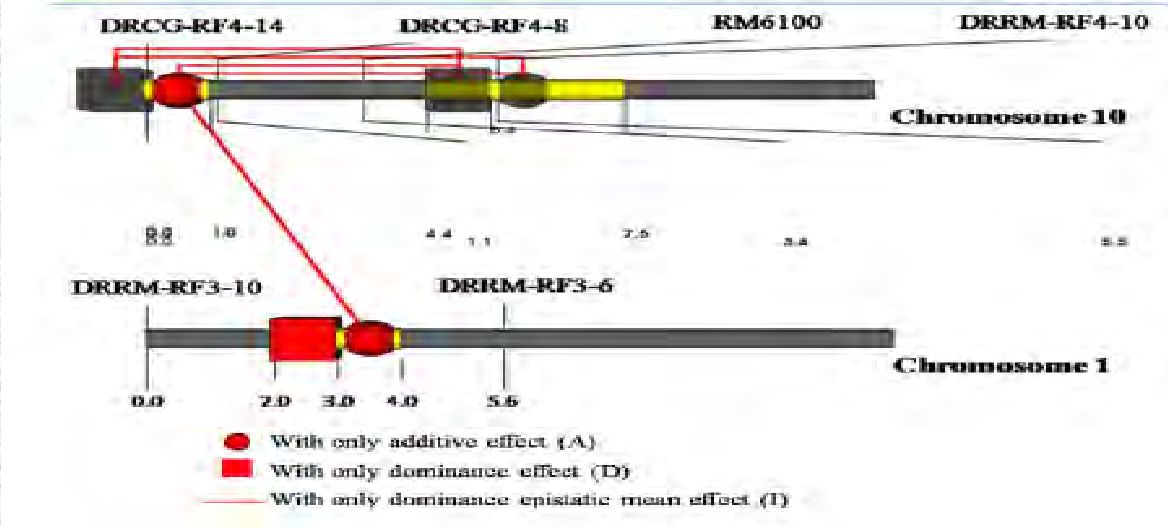


Identification of QTL and their interactions

KRH2



DRRH2



Rf4
Major QTL
41-43% of PV
Minor QTL

Rf3
Major QTL
23-31% of PV

QTL Network

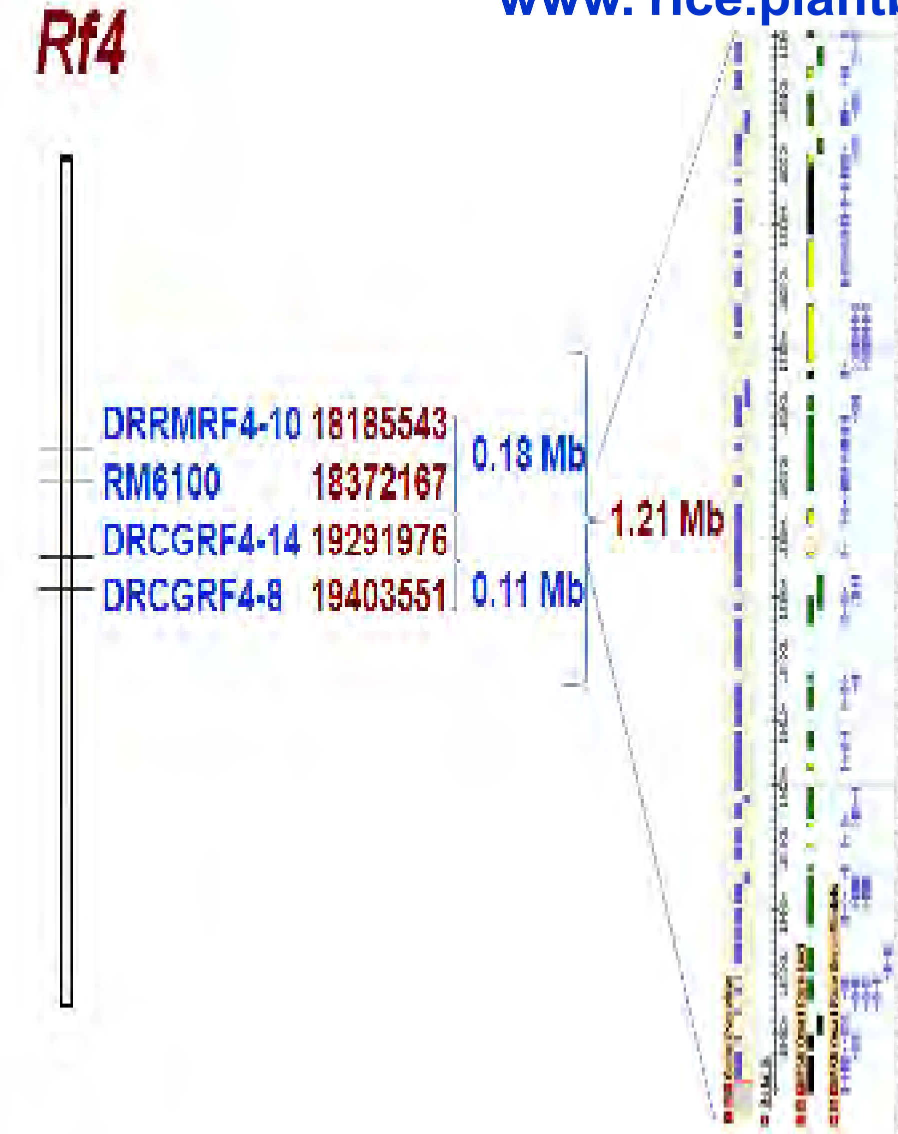
Selection accuracy (%) for combinations of markers for *Rf3* and *Rf4* loci in 212 restorer lines

	Rf3	RG140/pvul*	DRRM-RF3-6	DRRM-RF3-10	DRRM-RF3-5	RM10315	RM10318
		47.3	50.0	57.3	48.4	44.8	44.9
Rf4							
RM6100*	74.7	79.5	81.0	81.0	82.6	80.2	77.2
TMPPR3*	85.2	85.4	88.4	91.9	91.9	86.2	74.8
DRRM-RF4-10	81.0	84.2	85.3	88.4	86.0	84.2	73.0
DRCG-RF4-14	86.8	86.1	89.4	91.9	91.9	87.4	75.4
DRCG-RF4-8	85.2	82.6	89.4	91.9	91.9	87.4	75.4

*Reported markers

Euphytica 2012 187: 421-435

www. rice.plantbiology.msu.edu/



129 genes

38 con.hyp.prt

83 other prt.

9 CGs

-CyMalDH

-PPR

-Rf1 PPR 683 327 bp deletion

-Rf1 PPR 762 106 bp deletion (TMPPR3)

-TPR

-TPR

-TPR

-TPR

-α-glctsdase

/gene="Os10g0478200" Cytoplasmic malate dehydrogenase

```
A line              TAGATAGCTAGATCTCGTTGTGTGGCGGTTGATCTGGTTGGGGGGGGGGGTGCGGATTAT 216
R Line              TAGATAGCTAGATCTCGTTGTGTGGCGGTTGATCTGGTTGGGGGGGGGG-TGCGGATTAT 359
                    ************************************************* **********
  A Line              TTAGTGG---GAGGTCACTGGCGTGATTTGATTTCGGTACTAAGTTTTGTGGTTCGTGGA 273
  R Line              TTAGTGGGAGAAGGTCACTGGCGTGATTTGATTTCGGTACTAAGTTTTGTGGTTCGTGGA 419
                      *******   .*************************************************
```

/gene="Os10g0479200" Tetratricopeptide-like helical domain protein

```
   A Line             GCCCCCAGTAACTTCTATTAGTCACCGAGGGTTTTGTGGGATCGATTTTTTTCCCTCTA 303
   R Line             G-CAGCAGTAACTTCTATTAGTCACCGAGGGTTTTGTGGGATCGATTTTTTTCCCTCTA 539
                      * *. *******************************************************
```

/gene="Os10g0493600 Alpha-galactosidase precursor (EC 3.2.1.22)

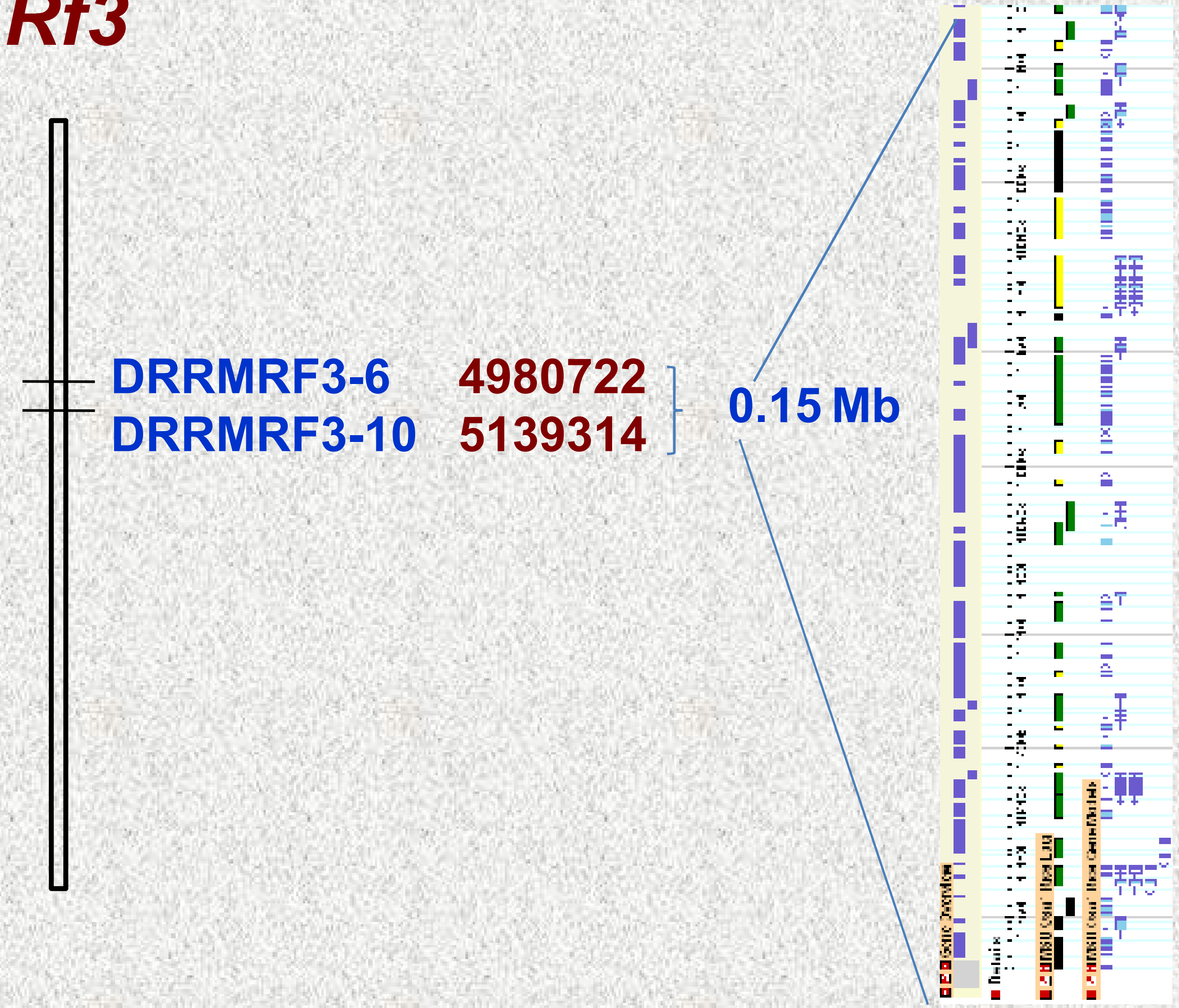
```
A Line               CAGGTACAGGGCACTGTG--AATTTTTTTTCACCCGGAAAAGTGGAAAACTTTCCTGAAT 993
R Line               CAGGTACAGGGCACTGTGTGAATTTTTTTTCACCCGGAAAAGTGGAAAACTTTCCTGAAT 1079
                     ******************  ****************************************

A Line               TCAATGGAAAGCTTAAGAGTTCTGTCACTTGGGTTTTTTTTTTGTACATTTAGGTTCTGT 1053
R Line               TCAATGGAAAGCTTAAGAGTTCTGTCATTTGGGTTTTTTTTTTGTACATTTAGGTTCTGT 1139
                     *************************** ********************************

A Line               CAGACTGTATGAAAG-AAAAAAAATCAGACAATCCTGTTCTCTGTCAGGCTGCAAGAACA 1112
R Line               CAGACTGTATAAAAGAAAAAAAAATCAGACAATCCTGTTCTCTGTCAGGCTGCAAGAACA 1199
                     **********.**** ********************************************
```

/gene="Os10g0493100" Pentatricopeptide-like helical domain protein

```
A Line              ATGCAGTGACGTATGGAGCAGTTATAGGCATACTT-GCAAGTCAGGCAGAGTAGAAGATG 1207
R Line              ATGTAGTGTGCTATGGAACAGTTATAGATGTACTTTGCAAGTCAGGCAGTGTAGATGATG 1379
                    *** ****:  ******.*********. .***** *************:*****:****
```

Rf3

www. rice.plantbiology.msu.edu



DRRMRF3-6	4980722
DRRMRF3-10	5139314

0.15 Mb

18 genes

6 con.hyp.prt

12 other prt.

4 CGs

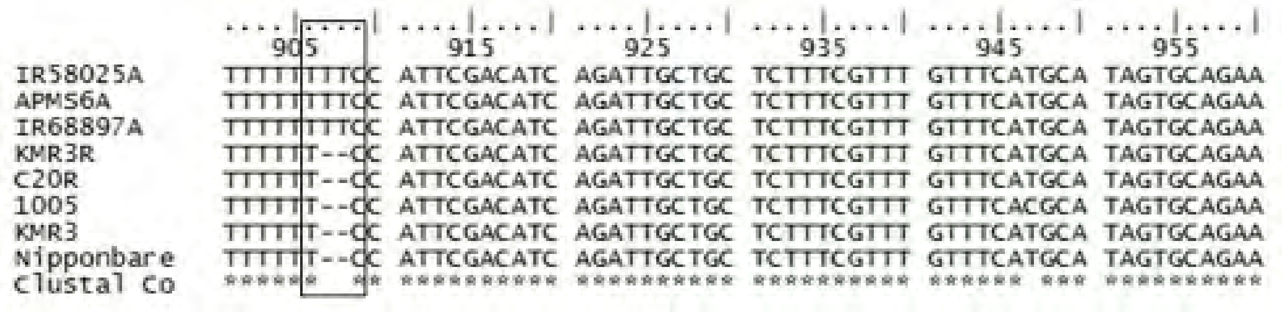
-MTPTdase

-PLNSPrt

-PPR

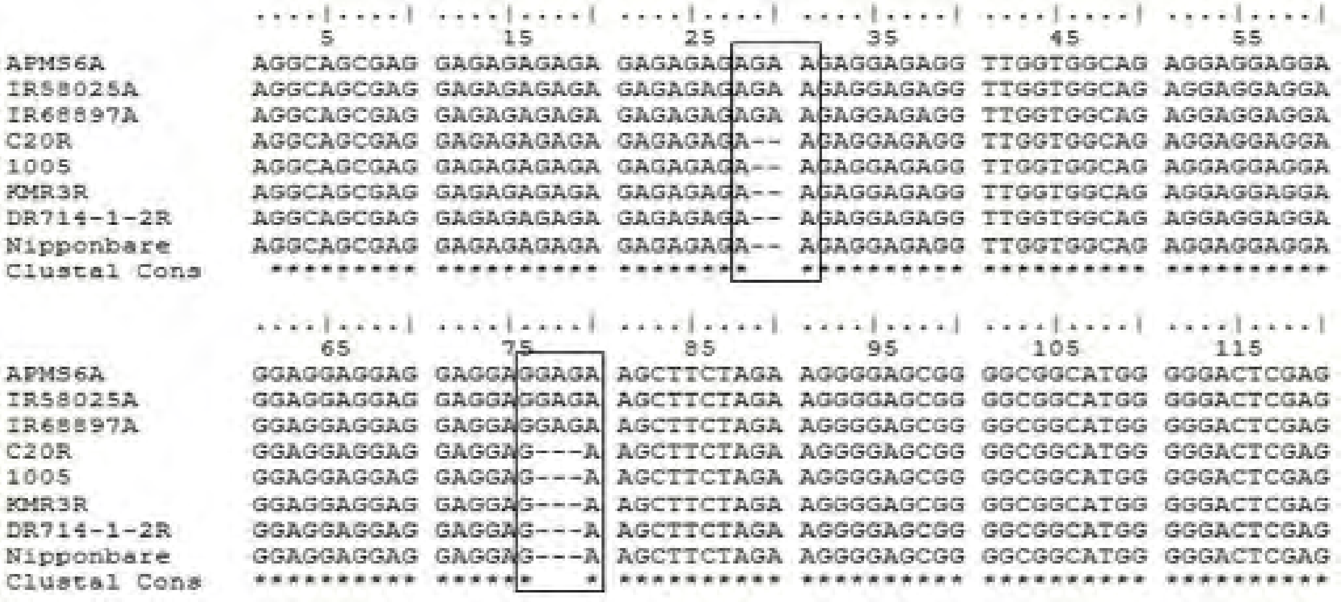
-PPR

Mitochondrial-processing peptidase subunit gene (LOC_Os01g09560)



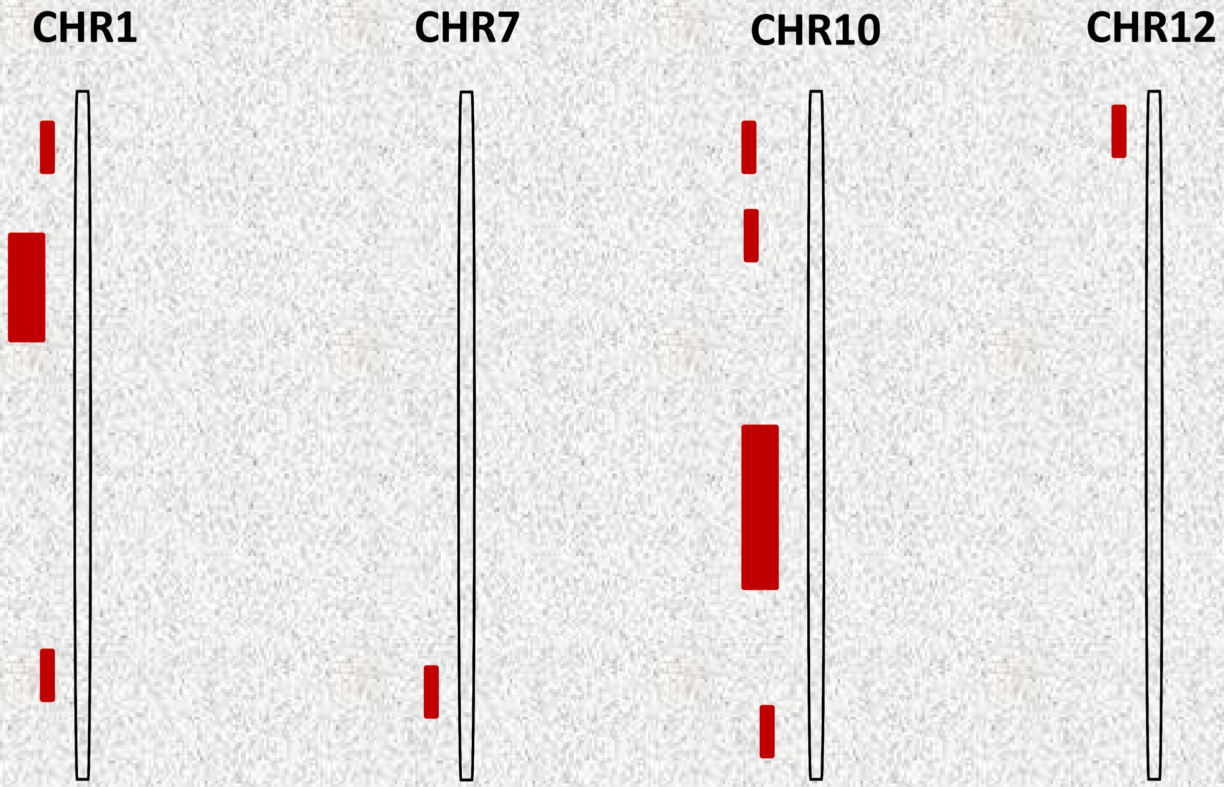
```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                      905        915        925        935        945        955
IR58025A        TTTTTTTTCC ATTCGACATC AGATTGCTGC TCTTTCGTTT GTTTCATGCA TAGTGCAGAA
APMS6A          TTTTTTTTCC ATTCGACATC AGATTGCTGC TCTTTCGTTT GTTTCATGCA TAGTGCAGAA
IR68897A        TTTTTTTTCC ATTCGACATC AGATTGCTGC TCTTTCGTTT GTTTCATGCA TAGTGCAGAA
KMR3R           TTTTTT--CC ATTCGACATC AGATTGCTGC TCTTTCGTTT GTTTCATGCA TAGTGCAGAA
C20R            TTTTTT--CC ATTCGACATC AGATTGCTGC TCTTTCGTTT GTTTCATGCA TAGTGCAGAA
1005            TTTTTT--CC ATTCGACATC AGATTGCTGC TCTTTCGTTT GTTTCACGCA TAGTGCAGAA
KMR3            TTTTTT--CC ATTCGACATC AGATTGCTGC TCTTTCGTTT GTTTCATGCA TAGTGCAGAA
Nipponbare      TTTTTT--CC ATTCGACATC AGATTGCTGC TCTTTCGTTT GTTTCATGCA TAGTGCAGAA
Clustal Co      ******  ** ********** ********** ********** ****** *** **********
```

Pollen-specific protein SF21 (LOC_Os01g0967)



```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    5         15         25         35         45         55
APMS6A          AGGCAGCGAG GAGAGAGAGA GAGAGAGAGA AGAGGAGAGG TTGGTGGCAG AGGAGGAGGA
IR58025A        AGGCAGCGAG GAGAGAGAGA GAGAGAGAGA AGAGGAGAGG TTGGTGGCAG AGGAGGAGGA
IR68897A        AGGCAGCGAG GAGAGAGAGA GAGAGAGAGA AGAGGAGAGG TTGGTGGCAG AGGAGGAGGA
C20R            AGGCAGCGAG GAGAGAGAGA GAGAGAGA-- AGAGGAGAGG TTGGTGGCAG AGGAGGAGGA
1005            AGGCAGCGAG GAGAGAGAGA GAGAGAGA-- AGAGGAGAGG TTGGTGGCAG AGGAGGAGGA
KMR3R           AGGCAGCGAG GAGAGAGAGA GAGAGAGA-- AGAGGAGAGG TTGGTGGCAG AGGAGGAGGA
DR714-1-2R      AGGCAGCGAG GAGAGAGAGA GAGAGAGA-- AGAGGAGAGG TTGGTGGCAG AGGAGGAGGA
Nipponbare      AGGCAGCGAG GAGAGAGAGA GAGAGAGA-- AGAGGAGAGG TTGGTGGCAG AGGAGGAGGA
Clustal Cons    ********** ********** ********   ********** ********** **********

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   65         75         85         95        105        115
APMS6A          GGAGGAGGAG GAGGAGGAGA AGCTTCTAGA AGGGGAGCGG GGCGGCATGG GGGACTCGAG
IR58025A        GGAGGAGGAG GAGGAGGAGA AGCTTCTAGA AGGGGAGCGG GGCGGCATGG GGGACTCGAG
IR68897A        GGAGGAGGAG GAGGAGGAGA AGCTTCTAGA AGGGGAGCGG GGCGGCATGG GGGACTCGAG
C20R            GGAGGAGGAG GAGGAG---A AGCTTCTAGA AGGGGAGCGG GGCGGCATGG GGGACTCGAG
1005            GGAGGAGGAG GAGGAG---A AGCTTCTAGA AGGGGAGCGG GGCGGCATGG GGGACTCGAG
KMR3R           GGAGGAGGAG GAGGAG---A AGCTTCTAGA AGGGGAGCGG GGCGGCATGG GGGACTCGAG
DR714-1-2R      GGAGGAGGAG GAGGAG---A AGCTTCTAGA AGGGGAGCGG GGCGGCATGG GGGACTCGAG
Nipponbare      GGAGGAGGAG GAGGAG---A AGCTTCTAGA AGGGGAGCGG GGCGGCATGG GGGACTCGAG
Clustal Cons    ********** ******   * ********** ********** ********** **********
```

Rf loci reported for WA-CMS



Loci on chrs 7 and 12 and also on chrs 8 and 9 have been identified and are being validated

Observations; quandaries; action plan

1. What is the status of genetics of fertility restoration?

Two major loci (Rf4 & Rf3) are known explaining ~70% of the PV. From genetics of the trait, it appears at least three major loci (~30% PV).

How?

All the loci (major and minor) involved in fertility restoration should be mapped and validated at least in a set of 10-12 restorers

2. How efficient are the available markers for *Rf4* & *Rf3*?

Up to 90% as validated in 212 restorers. If the number is increased the efficiency may go down.

Why?

The markers though based on candidate genes; these are not the exact candidate genes. So they are acting as linked markers but not CG based marker

High sequence variability within the genes is observed

2. Why? (contd…)

The selected putative candidate genes are based on gametophytic type of fertility restoration, the genes for sporophytic type of fertility restoration could be different (CW acyl-carrier protein synthase like domain)

Most of the gene information is derived from Nipponbare, a non-restorer for WA-CMS, thus the gene may not be there

If PPR/TPR genes are involved, the region has many copies of genes, since these genes have degenerate motifs, a functional gene could be operating in indica

How?

Characterization of the complete *Rf4* and *Rf3* loci

3. Can we develop marker system for *Rf* loci of WA-CMS?

Yes. With identification of the all the Rf loci and their characterization, a battery/kit of 3 or 4 markers with multiplex for 3 to 4 loci major can be developed for identifying restorers

When?

IHRS 2016!!!

ACKNOWLEDGEMENT
Department of Science and Technology
Government of India for the research grant

Thank you