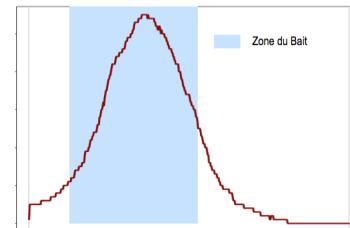
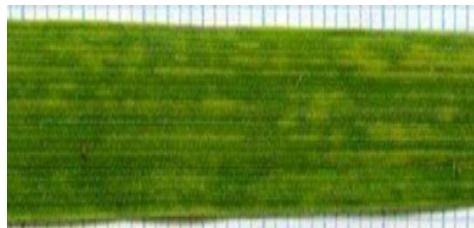
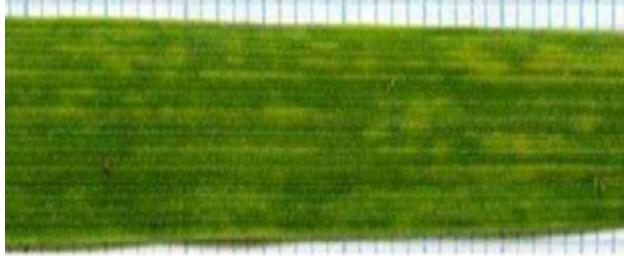


# Genotyping Durum Wheat using specific allelic capture : an application to detect QTL for WSSMV resistance



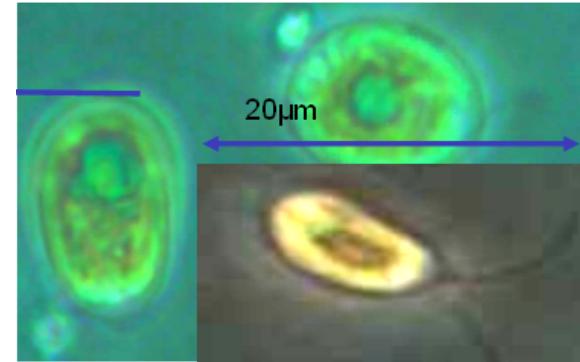
*INRA / Supagro* : Y. Holtz, V. Viader, M. Ardisson, S. Santoni, V. Ranwez, G. Poux, P. Roumet, J. David  
*Arvalis* : M. Bonnefoy, F. Savignard, D. Gouache

## -- Introducing White Spindle Streak Mosaïque Virus (WSSMV)

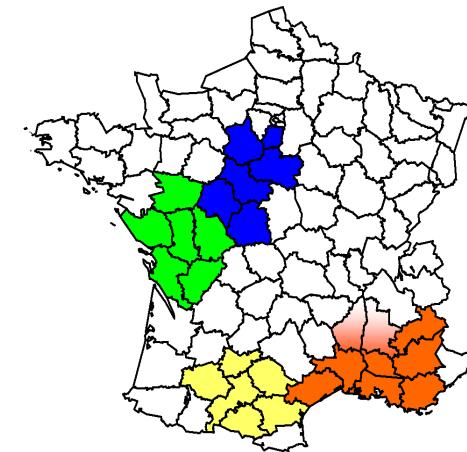


WSSMV Symptoms

- ✓ **WSSMV** has a huge Impact in France
- ✓ **Resistant varieties** are the only way to fight the disease (*Kanyuka et al. 2013*)
- ✓ **Soldur** is resistant and elite but has some default (and sometimes has light symptoms)
- ✓ Germplasm : **Dic2** is a cultivated emmer accession , with a lot of agronomical defaults



*Polomyxa graminis* is the vector of WSSMV



Localization of WSSMV in France

- ✓ Major QTL of resistance at **WSSMV** found on chromosome **2D** in **Bread Wheat**. (*Khaan et al. 2000*)
- ✓ Major QTL of resistance at **SBCMV** found on chromosome **5D** in **Bread Wheat**. (*Perovic et al. 2009*)
- ✓ Major QTL of resistance at **SBCMV** found on chromosome **2BS** and on minor locis for **Durum Wheat** (*Maccaferri et al. 2011*)
- ✓ **No QTL** of resistance at **WSSMV** for **Durum Wheat** has been published yet

→ A QTL of resistance to WSSMV in Durum wheat is still needed !

## -- Our experimentations



Dic2  
*Dicoccum*



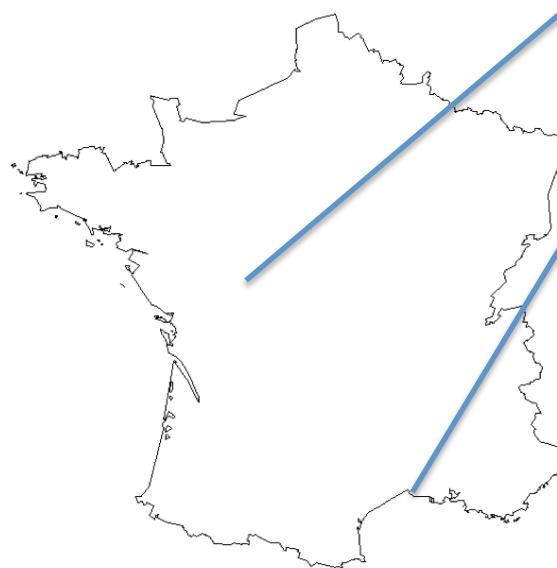
Silur  
*Durum Elite*



**Biparental Population**  
8 Years of Fixation  
180 Rils

### Blois

Field infested by WSSMV only  
Observation in 2012 and 2015  
180 accessions + 80 rep



### Montpellier

Creation of the population in 2007  
Fixation every year

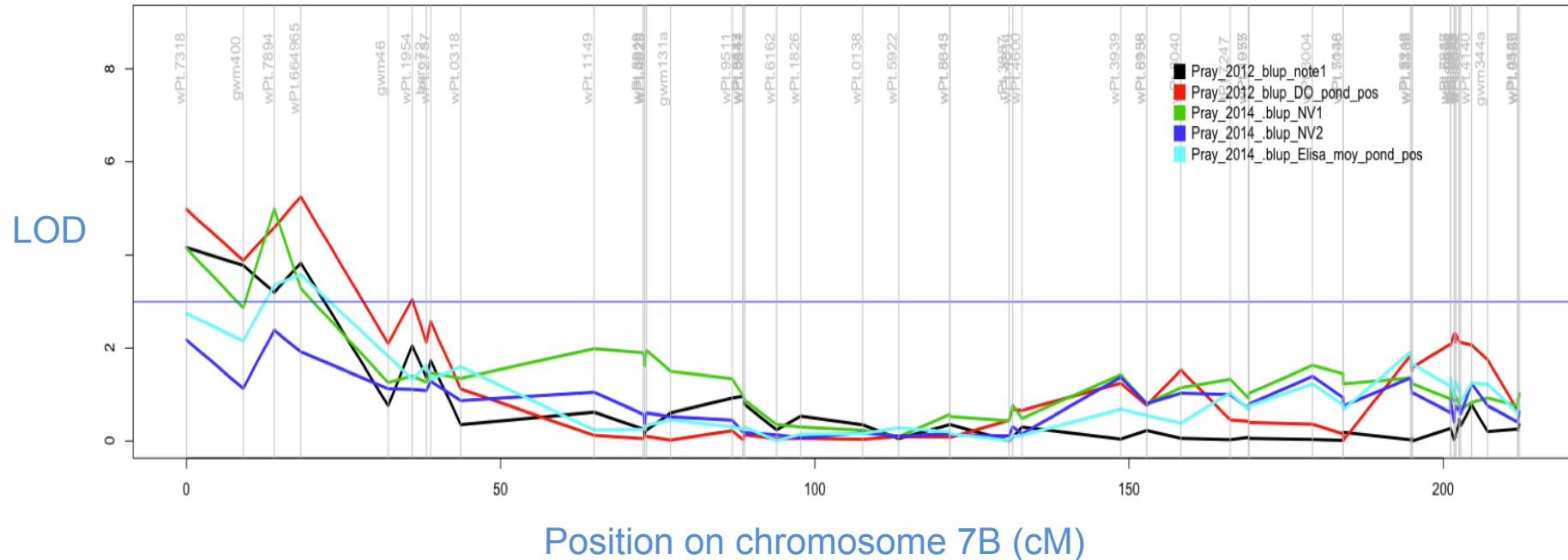


### Phenotyping :

- ✓ **Visual Notation** (Mark between 0 and 5)
- ✓ **ELISA** (Semi quantitative evaluation of WSSMV quantity in leaves)

## -- QTL detection for WSSMV resistance in Durum Wheat

Detection : 180 RILs Dic2 x Silur and 1000 DARTs)



- ✓ A QTL on chromosome 7B
- ✓ Effect consistent in 2012 and 2015
- ✓ Co-localization with SBCMV minor QTL (Maccaferri et al. 2011)
- ✓ Other QTLs detected
- ✓ LOD are low : heterogeneous field contamination, interaction with others traits (earliness ?), distance to the resistance gene ?

-- Strategy of genotyping

## How to Genotype a population Using Specific allelic Capture ?



**Dic2**

*Dicoccum  
Resistant*

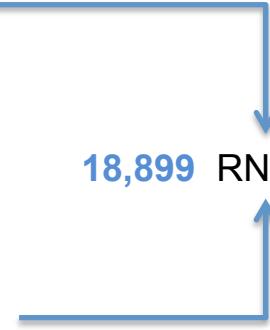


**Silur**

*Durum Elite  
Susceptible*

X

18,899 RNA seq SNP



## -- Strategy of genotyping

### How to Genotype a population Using Specific allelic Capture ?



**Dic2**

*Dicoccum  
Resistant*



**Silur**

*Durum Elite  
Susceptible*

**X**

**18,899 RNA seq SNP**

**EPO**

*(Pre-Breeding population)*

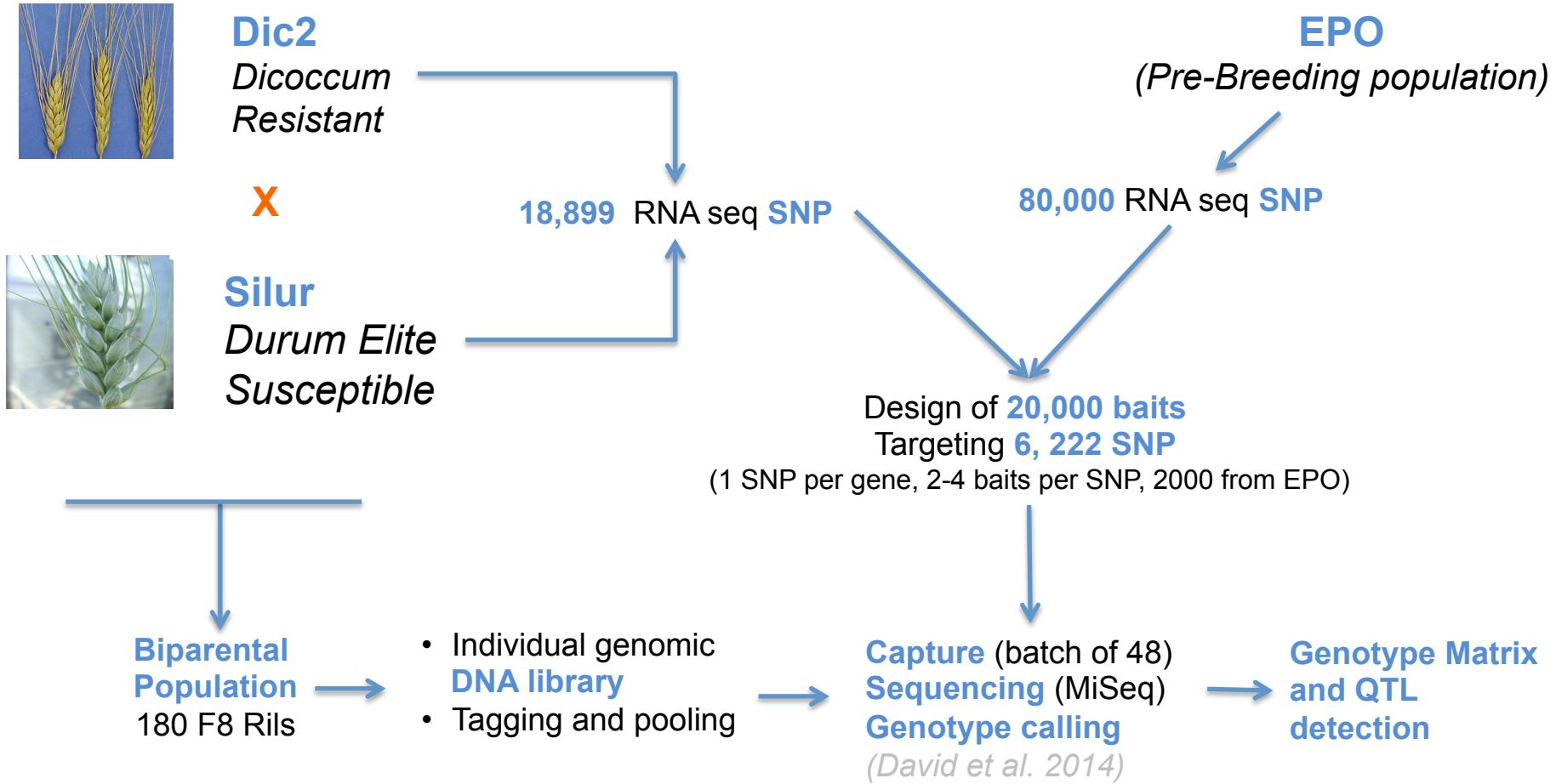
**80,000 RNA seq SNP**

**Design of 20,000 baits  
Targeting 6,222 SNP**

(1 SNP per gene, 2-4 baits per SNP, 2000 from EPO)

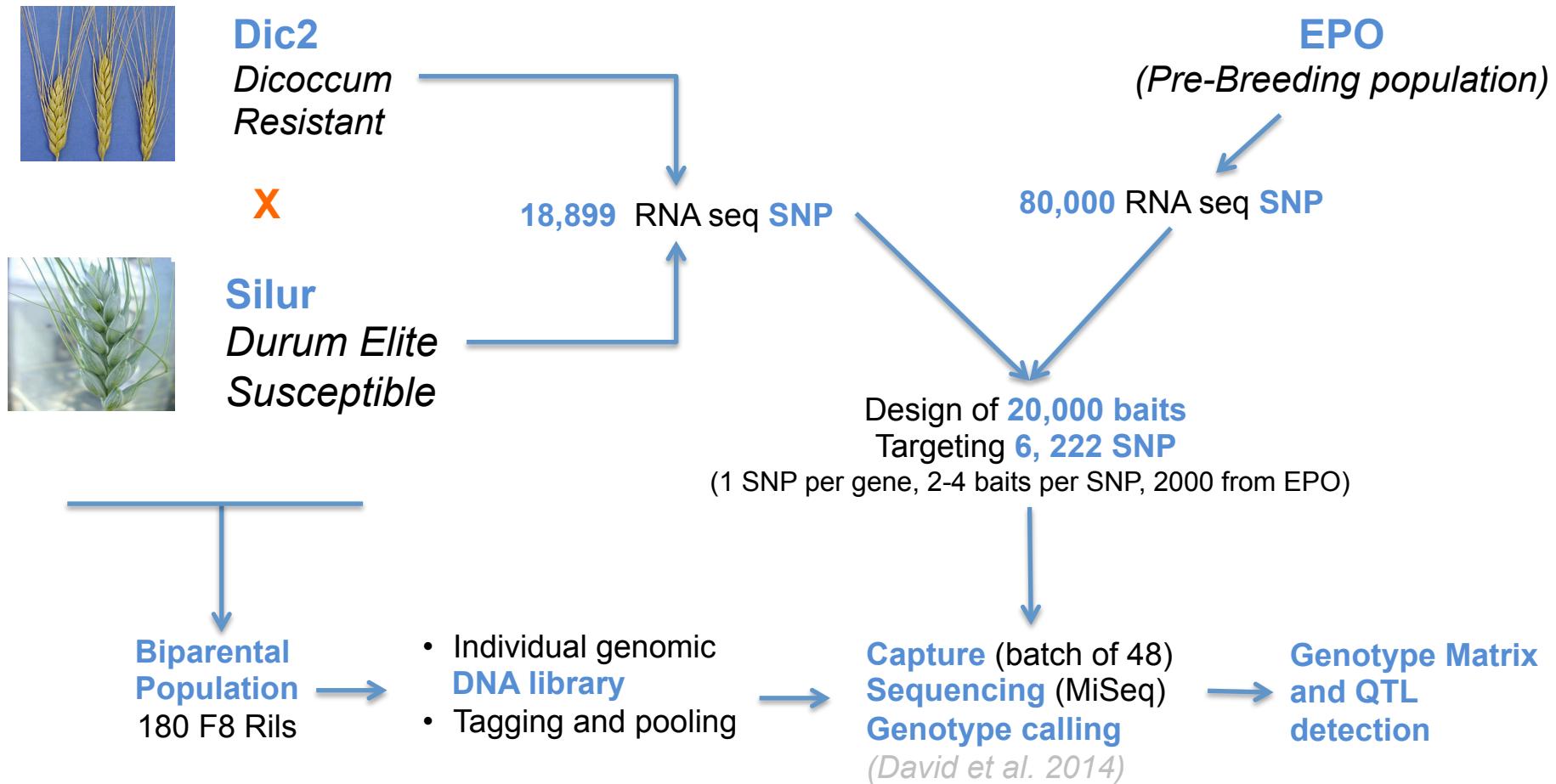
## -- Strategy of genotyping

### How to Genotype a population Using Specific allelic Capture ?



## -- Strategy of genotyping

### How to Genotype a population Using Specific allelic Capture ?



- ✓ Low cost (*Rohland et al., 2012*)
- ✓ Choice of the targeted SNPs (durum and parental specific)
- ✓ Easy design of KASPAR Markers
- ✓ Sequence : stable technic and information, does not depend on a array platform

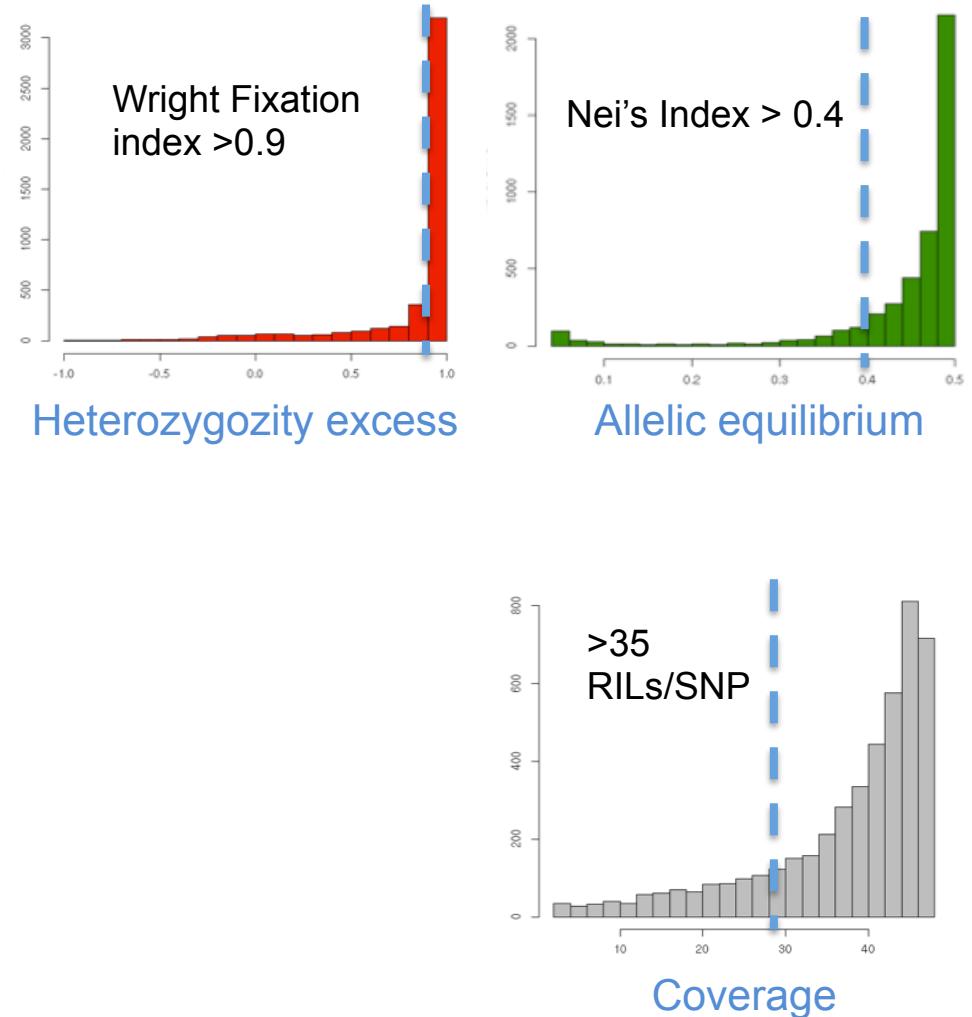
## 6222 designed SNPs :

- ✓ **5200** polymorphic SNPs
- ✓ Parental genotype recovered
- ✓ **4600** redetected SNPs (89%)

## -- A first pilot experiment on a panel of 48 RILs from Dic2 x Silur

### 6222 designed SNPs :

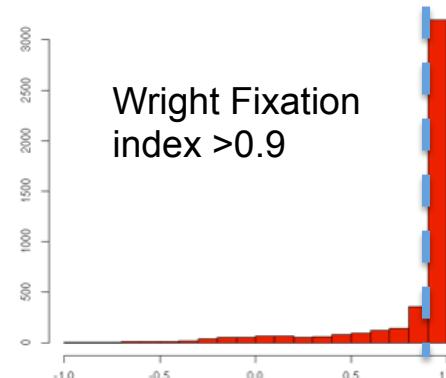
- ✓ 5200 polymorphic SNPs
- ✓ Parental genotype recovered
- ✓ 4600 redetected SNPs (89%)
  
- ✓ 2849 High Quality SNPs
- ✓ 2037 useable on the 48 panel
- ✓ 873 BONUS SNPs



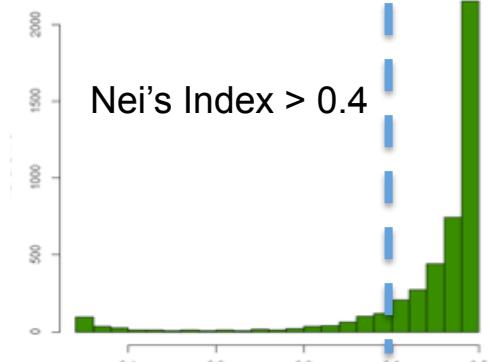
## -- A first pilot experiment on a panel of 48 RILs from Dic2 x Silur

### 6222 designed SNPs :

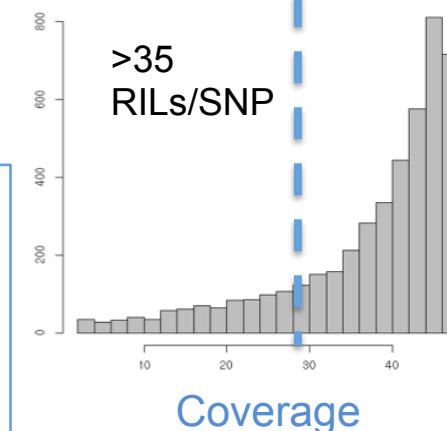
- ✓ 5200 polymorphic SNPs
- ✓ Parental genotype recovered
- ✓ 4600 redetected SNPs (89%)
  
- ✓ 2849 High Quality SNPs
- ✓ 2037 useable on the 48 panel
- ✓ 873 BONUS SNPs



Heterozygosity excess



Allelic equilibrium



Coverage

- ✓ Possibility to use polymorphisms of other populations
- ✓ High rate of recovered SNPs
- ✓ Improvements to come:
  - Decrease of Heterozygosity (*de novo* assembly)
  - Increased coverage :  $10^6$  read pairs/RIL with HiSeq
  - Increased number of targets : 10 000 targeted SNPs

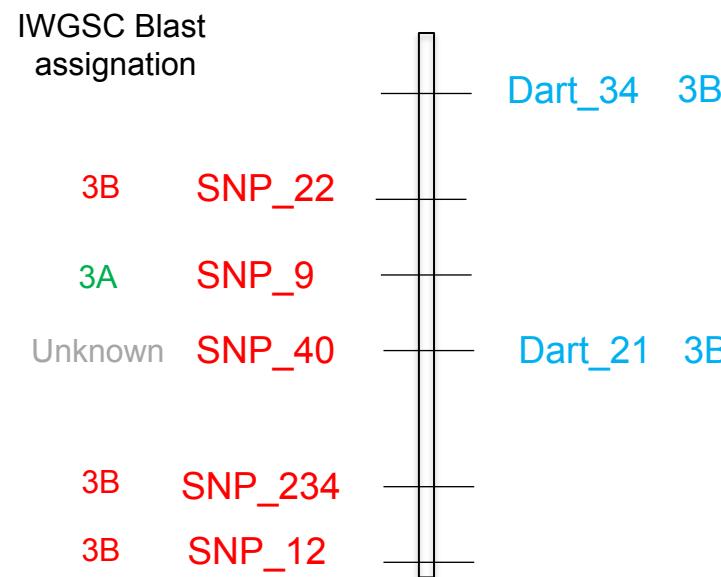
>35  
RILs/SNP

## -- Densifying the Dic2 X Silur genetic map

1037 Darts Markers are available  
(Fusarium GIE Program)

2910 SNP are now available  
(TRAM program)

Creation + Assignation of Linkage Groups to chromosomes ( One Map R package )



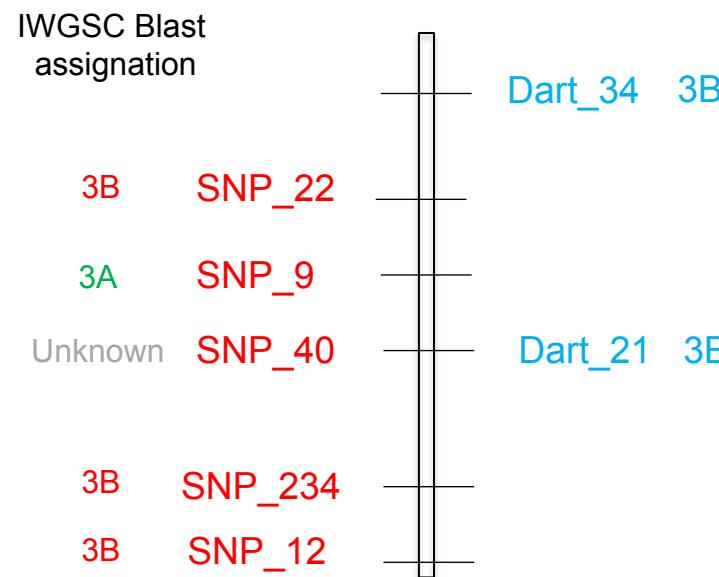
- ✓ SNPs are distributed uniformly among LGs
- ✓ SNPs are linked with Dart Markers

## -- Densifying the Dic2 X Silur genetic map

1037 Darts Markers are available  
(Fusarium GIE Program)

2910 SNP are now available  
(TRAM program)

Creation + Assignation of Linkage Groups to chromosomes ( One Map R package )



✓ **Dart Markers** are attributed to chromosome using **deletion lines**

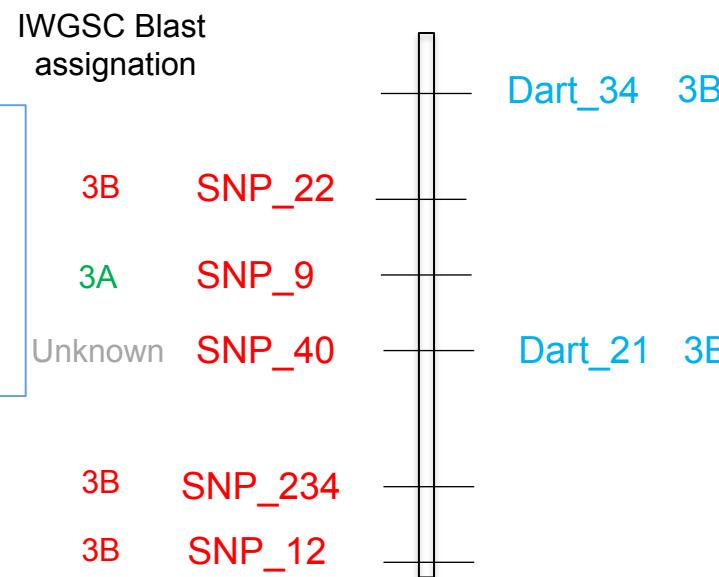
- ✓ SNPs are **distributed** uniformly among LGs
- ✓ SNPs are **linked with Dart Markers**

## -- Densifying the Dic2 X Silur genetic map

1037 Darts Markers are available  
(Fusarium GIE Program)

2910 SNP are now available  
(TRAM program)

Creation + Assignation of Linkage Groups to chromosomes ( One Map R package )



- ✓ SNPs are blasted on the genomic **Bread Wheat reference genome** (IWGSC), with putative **Chromosome assignation** (IWGSC, Mayer et al. 2014)

- ✓ **Dart Markers** are attributed to chromosome using **deletion lines**

- ✓ SNPs are **distributed** uniformly among LGs
- ✓ SNPs are **linked** with Dart Markers

## -- Summary of Chromosomal attributions of SNP

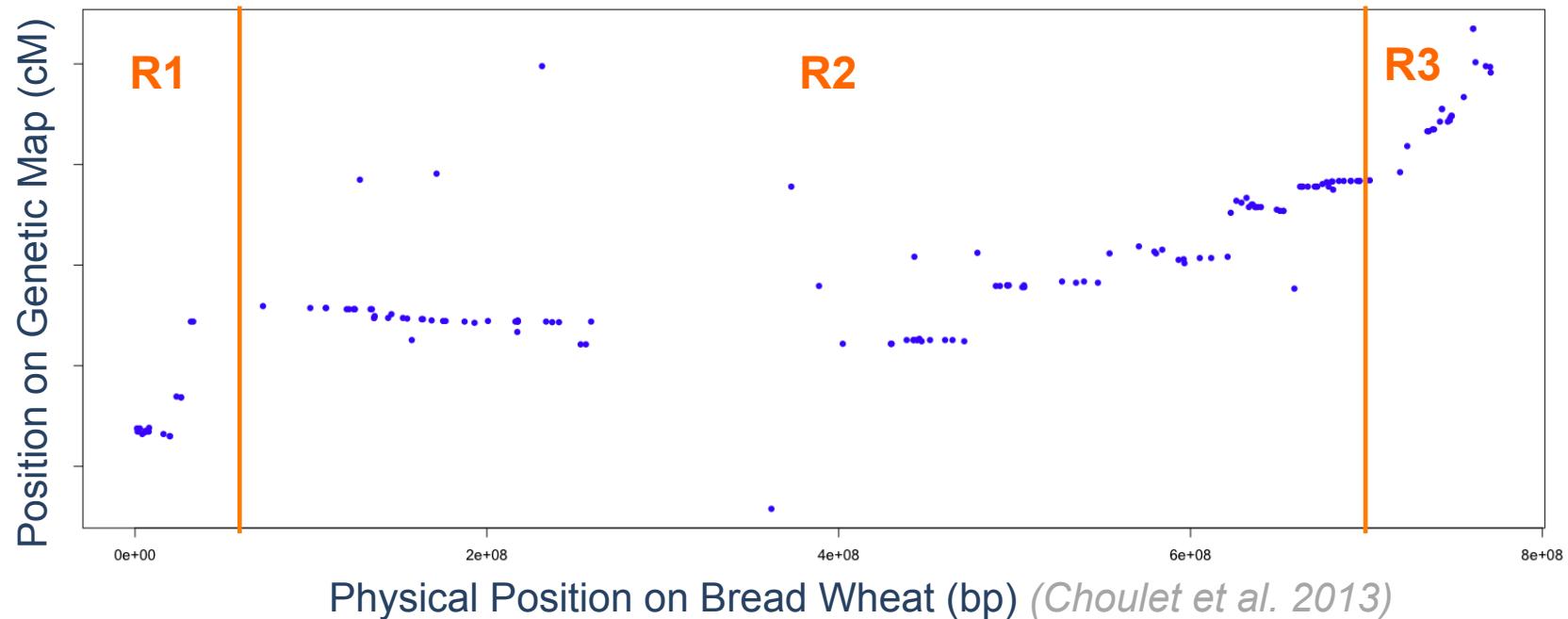
### LG chromosomal Assignation on Bread Wheat

LG DART chromosome assignation

	1A	1B	2A	2B	3A	3B	4A	4B	5A	5B	6A	6B	7A	7B
1A	107	23	0	1	0	0	0	0	0	0	0	0	0	0
1B	41	135	0	0	0	0	2	0	0	2	1	2	0	0
2A	0	0	28	6	0	0	0	0	0	0	0	0	0	0
2B	0	0	14	85	0	0	0	0	0	0	0	1	0	0
3A	1	0	0	0	26	8	0	0	0	0	0	0	0	0
3B	0	0	0	0	29	186	0	1	0	0	0	0	0	0
4A	0	1	0	0	0	0	75	12	0	1	0	0	4	0
4B	0	1	0	0	0	0	27	154	4	0	0	0	0	0
5A	0	0	1	0	0	0	0	1	57	15	0	0	0	0
5B	0	0	0	0	0	0	3	0	15	99	0	0	0	1
6A	0	0	0	0	0	0	0	0	0	1	30	4	0	0
6B	0	0	0	0	0	0	1	1	0	0	6	26	0	0
7A	0	0	0	0	0	0	11	0	0	0	0	1	86	14
7B	0	1	0	0	0	0	0	1	0	2	0	1	14	44

- ✓ **Consistency** of the genetic map: **80 %** of the SNPs have a similar genetic and genomic Blast assignation
- ✓ Confusion with **Homeolog : 16%**
  - ✓ Blast-hit error ?
  - ✓ Difference Durum vs Bread Wheat ?
  - ✓ Absence of certain homeologous copies in the IWGSC reference ?
- ✓ **4%** are divergent : blast error, multigene families.
- ✓ 11 SNP in 4A/7A mixed LG : translocation ?
- ✓ 1058 non assigned (no hit on IWGSC or no DART in their LG)

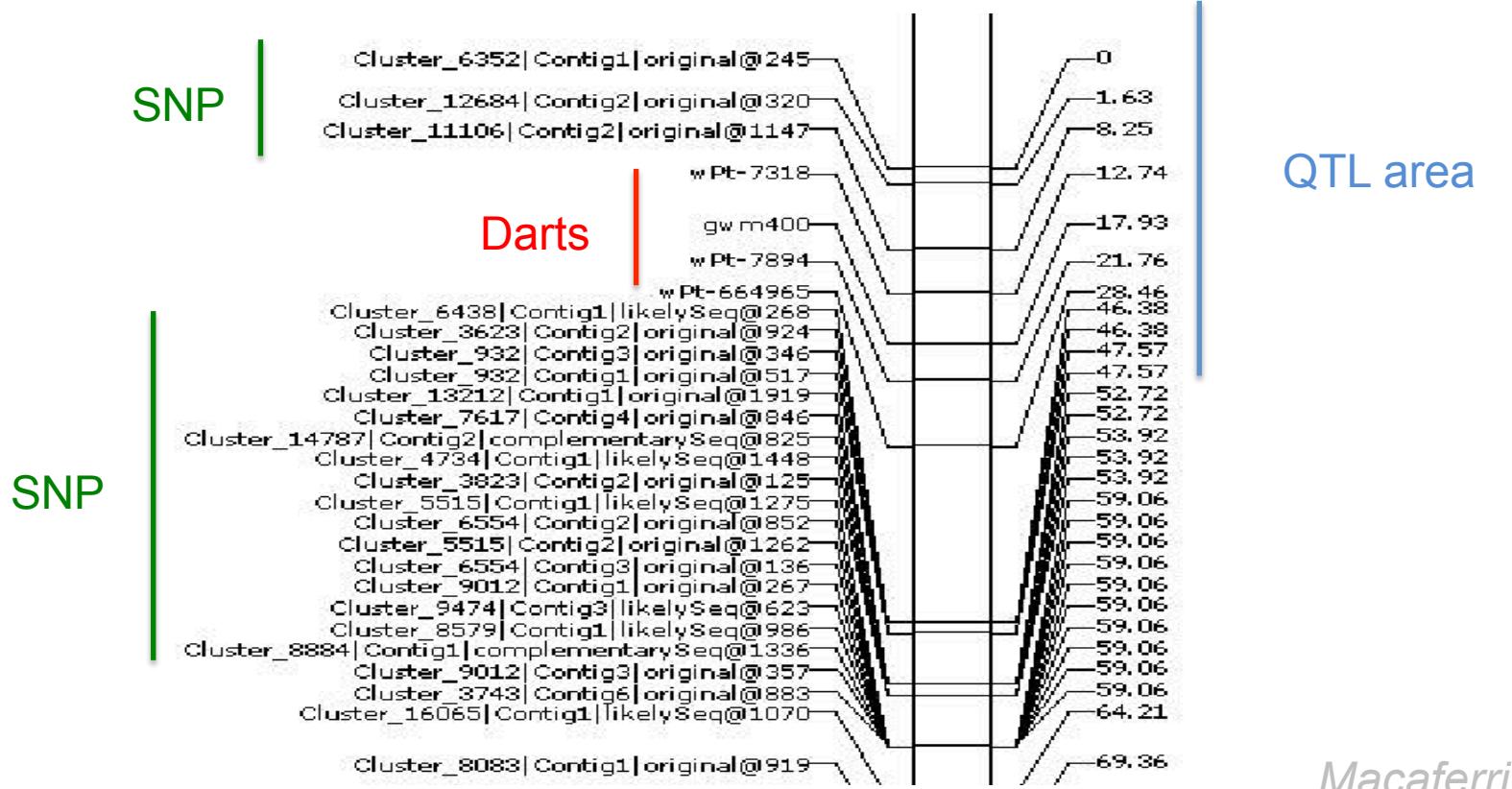
### Consistency of markers order ?



- ✓ Excellent Consistency between physical and genetic positions
- ✓ Increase of the recombination rate in zones R1 and R3 as described by *Saintenac and al., 2013*
- ✓ Few discrepancies (5 / 186)

## -- Densitification of QTL areas ?

### Genetic Map of the highly recombining distal portion of the short arm of chromosome 7B (48 Dic2 \* Silur)



Macaferrari et al. 2014

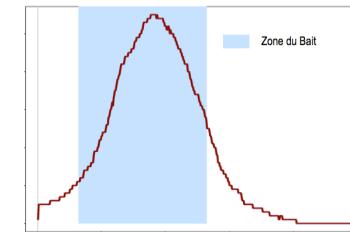
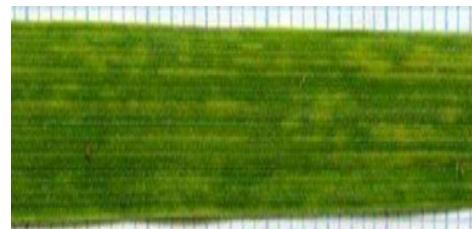
- ✓ **Densification** of the distal part of chromosome 7B
- ✓ How to increase marker density ?

## -- Conclusion and Perspective

- ✓ Specific allelic capture is relevant to genotype a population for targeted SNPs with low costs
- ✓ Applying this protocol on the whole population and with HiSeq technology should give High Quality genotype Matrix (Results in June 2015)
- ✓ Number of markers can be increased improving bio-informatics pipelines and number of Targeted SNPs
- ✓ Applying this Strategy on 2 other populations should permit to precisely localize genes of resistance to WSSMV



Thank You

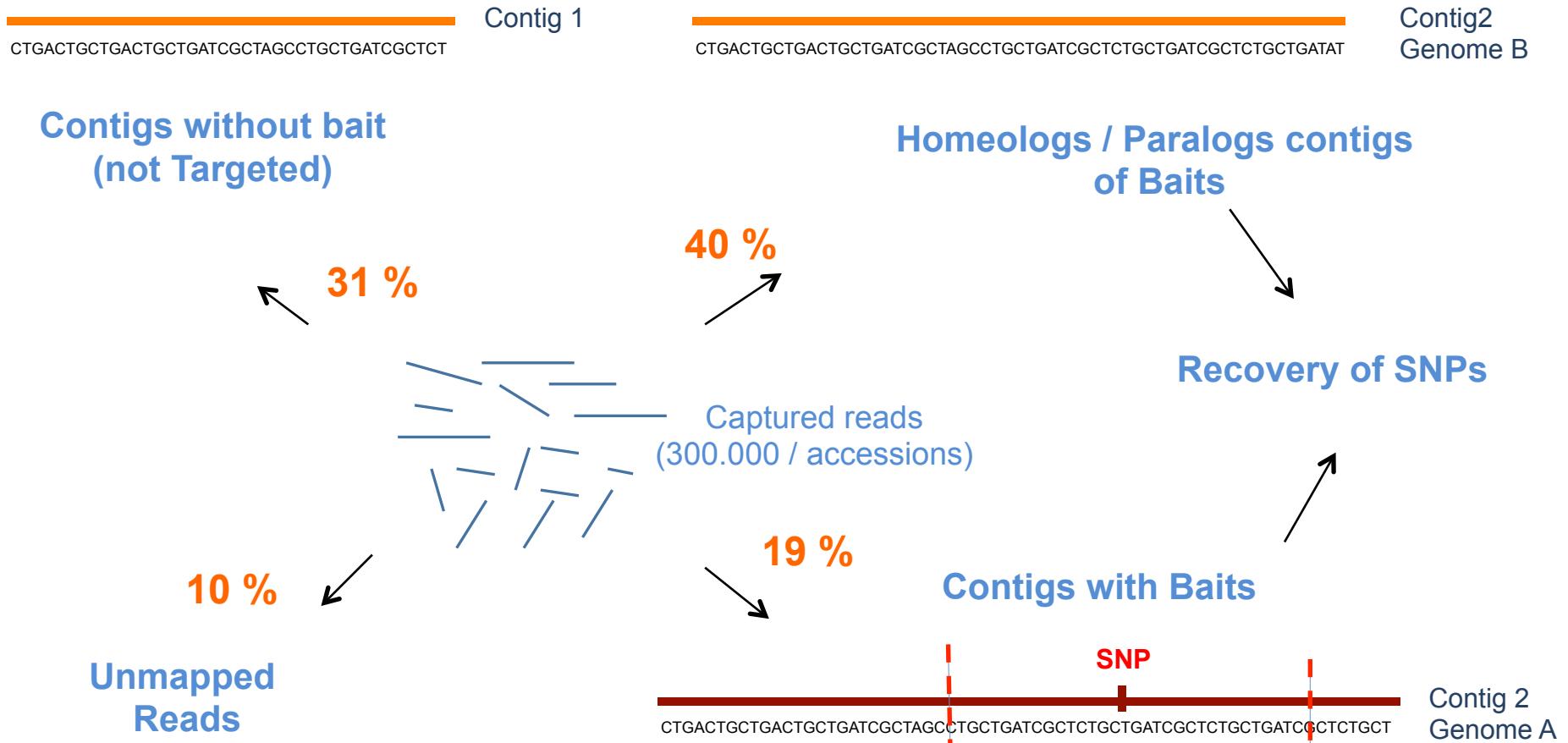


*INRA / Supagro : Y. Holtz, V. Viader, M. Ardisson, S. Santoni, V. Ranwez, G. Poux, P. Roumet, J. David*

*Arvalis : M. Bonnefoy, F. Savignard, D. Gouache*



## -- What did we really captured ?? (Pilot experiment)



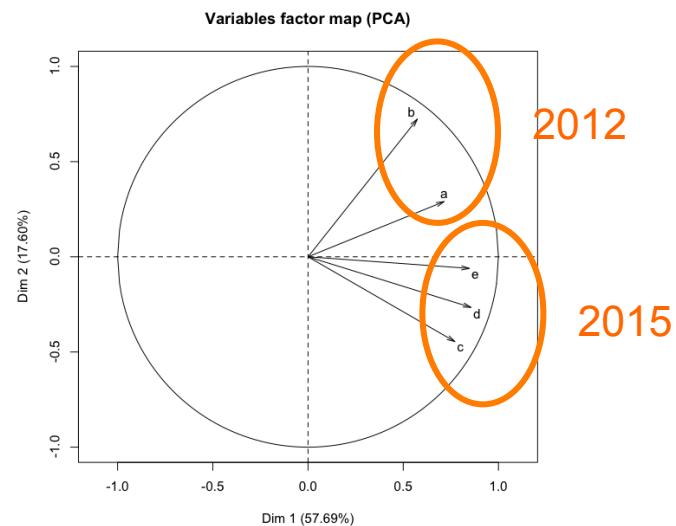
- ✓ **4600** SNPs recovered (92%)
- ✓ **59%** of Enrichment (Useful reads)
- ✓ **“Bonus”** SNPs on homeologs

## -- Correlation between data

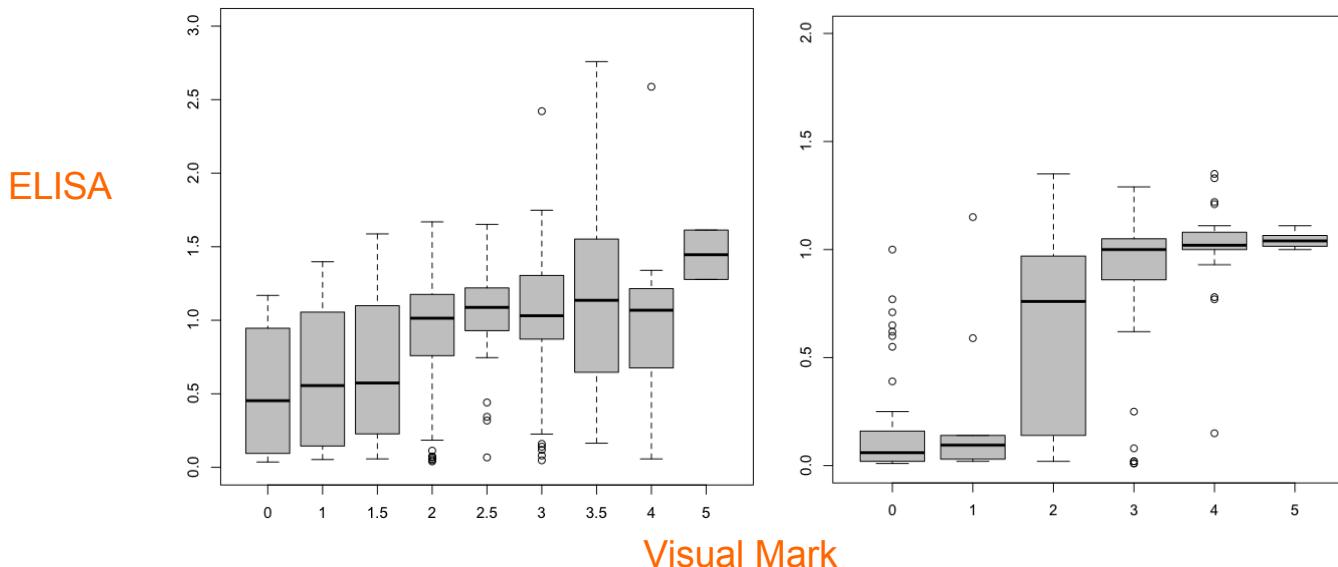
### Heritability

Caractère	2012	2014
Visual Mark 1	-	0,54
Visual Mark 2	0,67	0,65
ELISA	0,45	0,77

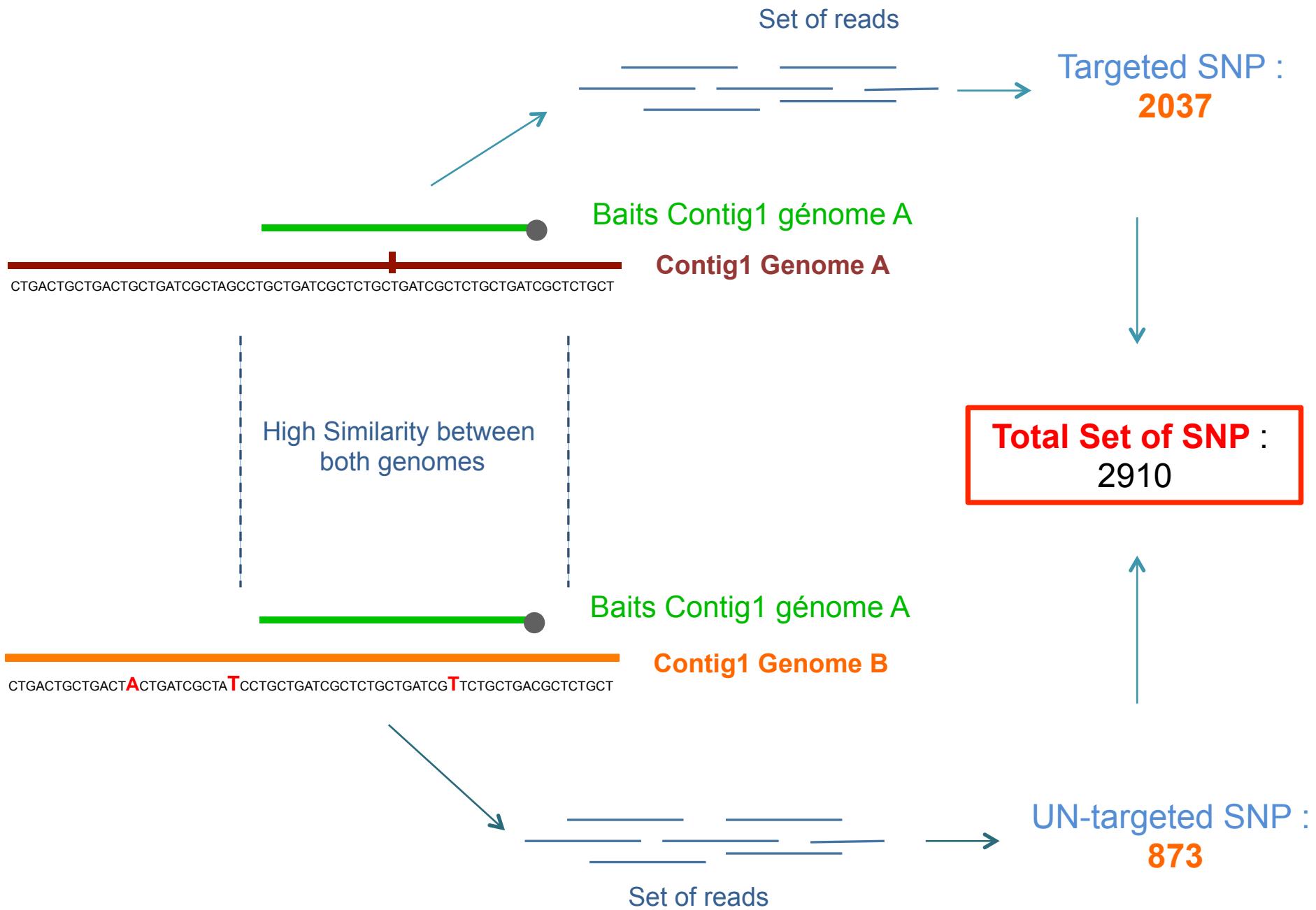
### Correlation between Years



### Correlation between Visual Mark and ELISA



## -- SNP recovered on homeologous contigs

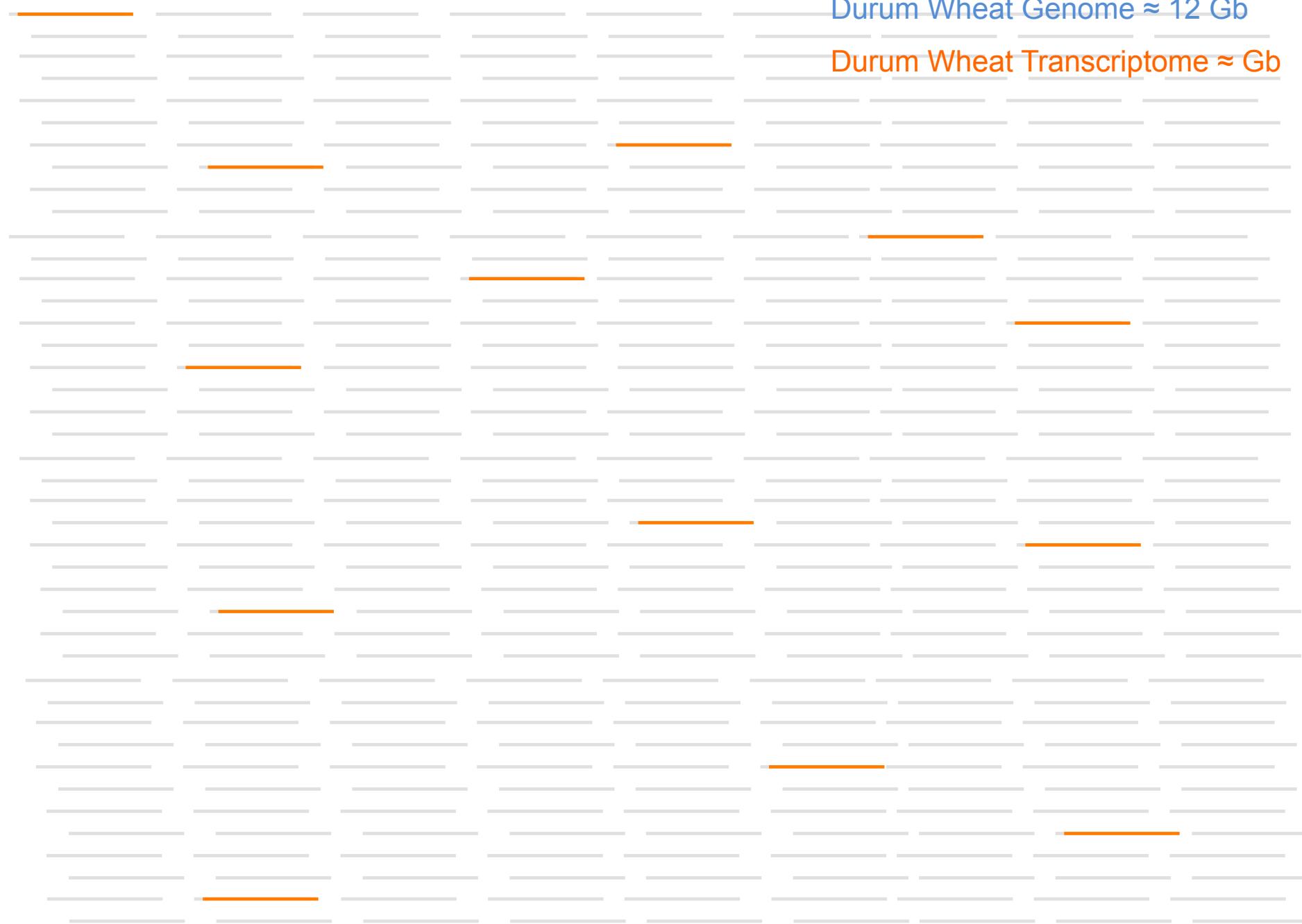


## -- Specific Allelic Capture Proceedings

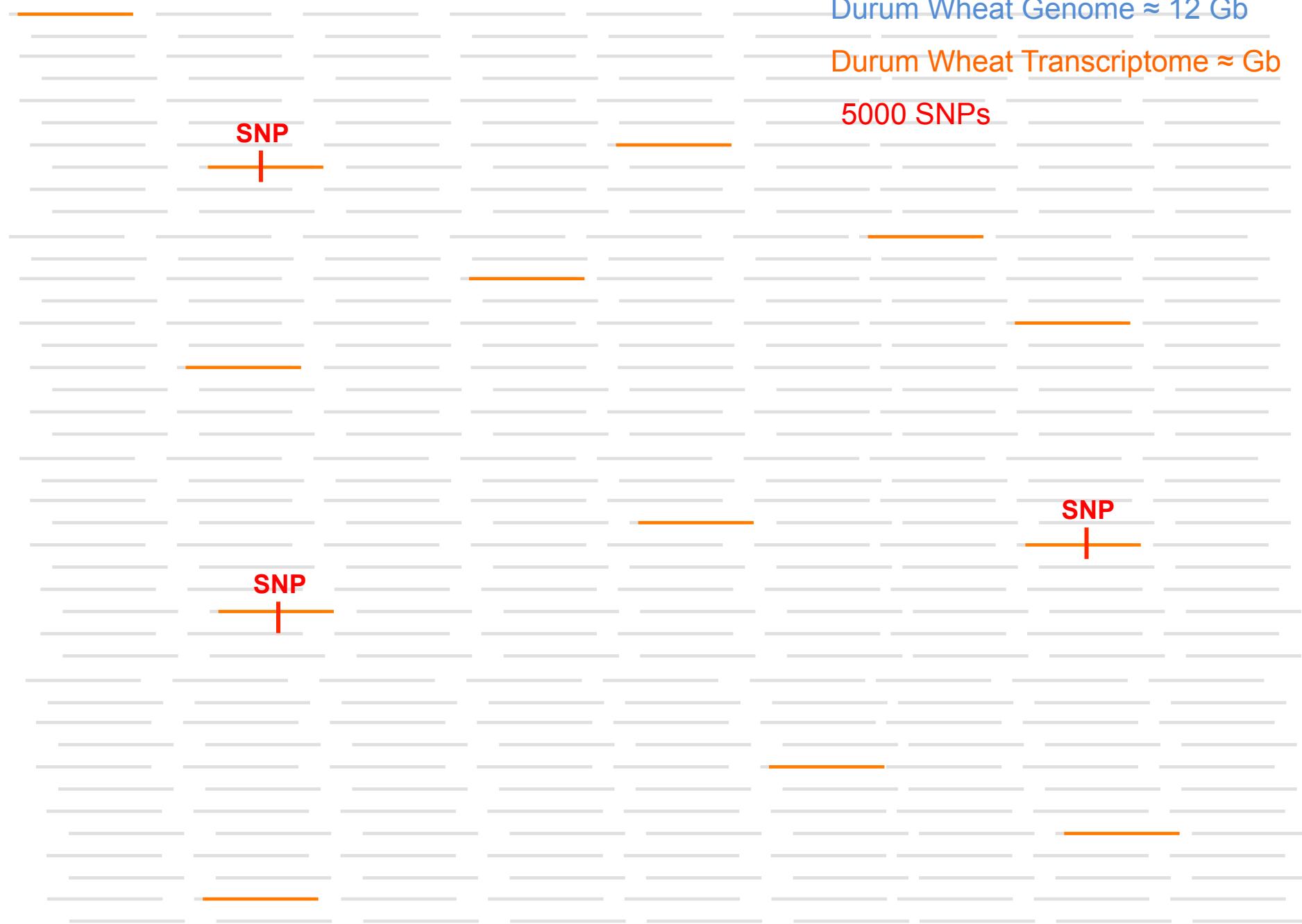
Durum Wheat Genome  $\approx$  12 Gb



## -- Specific Allelic Capture Proceedings



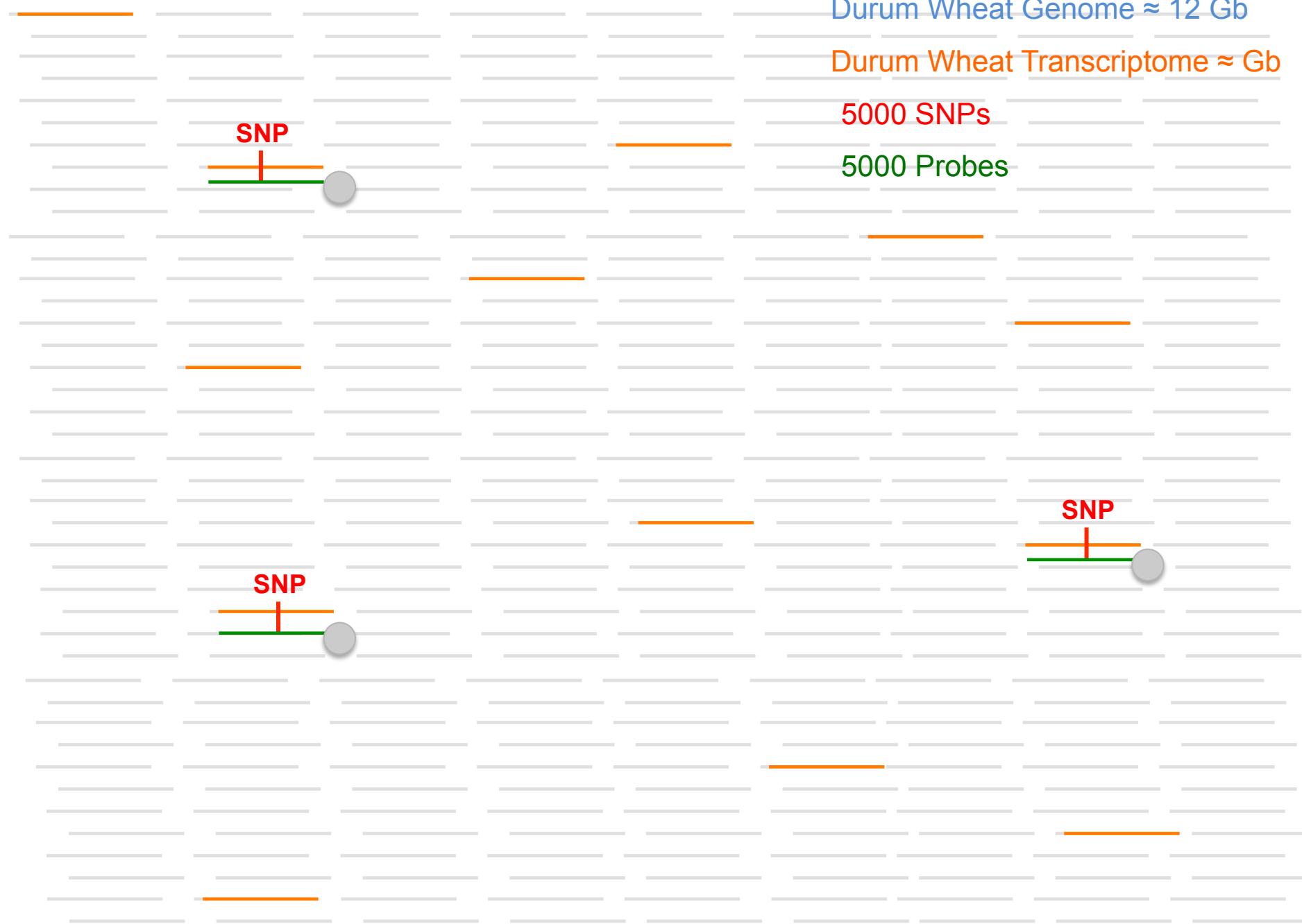
## -- Specific Allelic Capture Proceedings



## -- Specific Allelic Capture Proceedings



## -- Specific Allelic Capture Proceedings



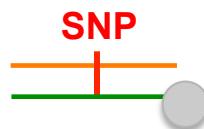
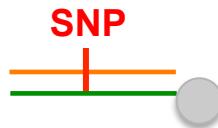
## -- Specific Allelic Capture Proceedings

Durum Wheat Genome  $\approx$  12 Gb

Durum Wheat Transcriptome  $\approx$  Gb

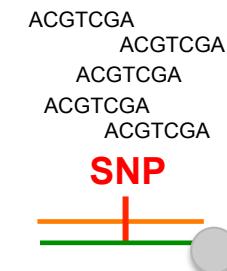
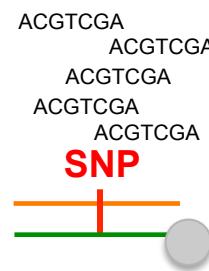
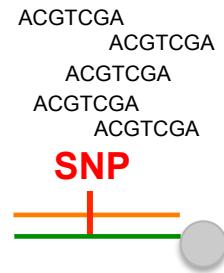
5000 SNPs

5000 Probes



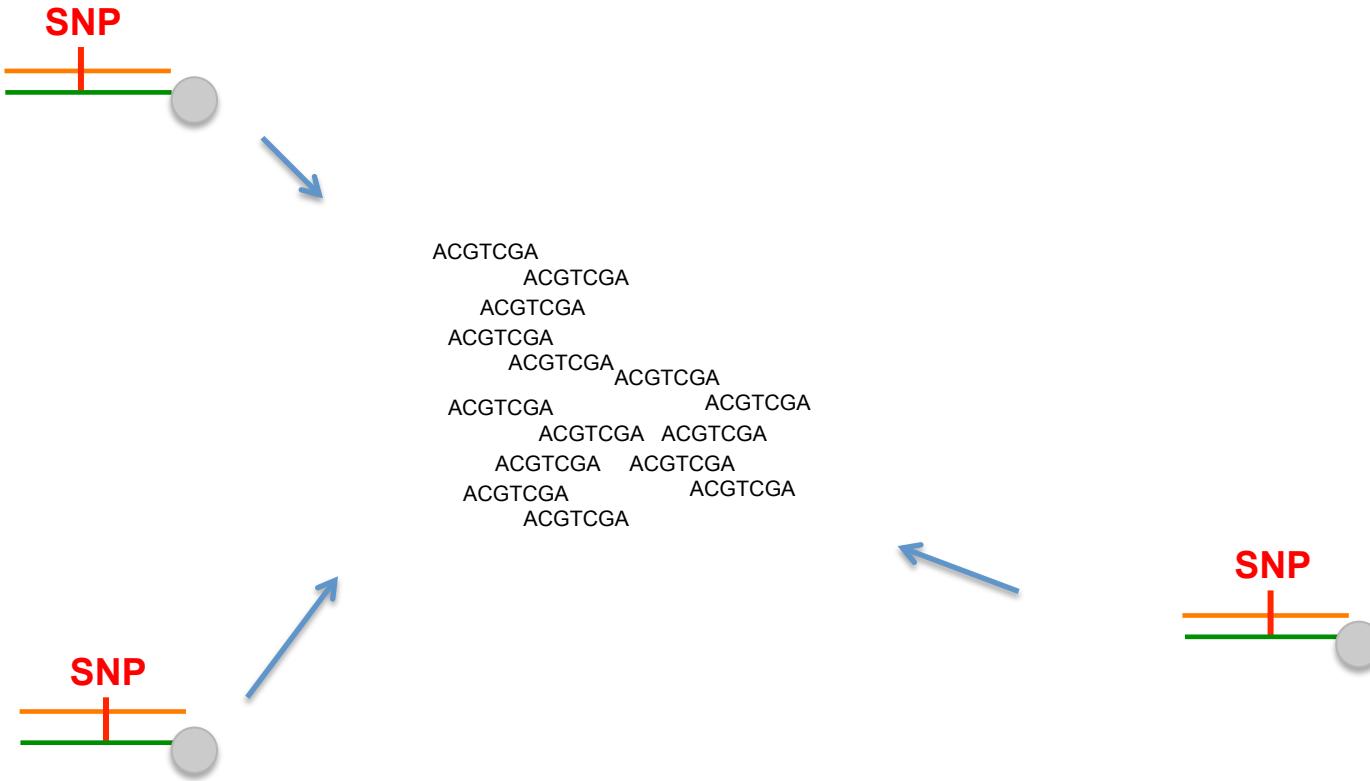
--- > Only Targeted SNPs are kept !

## -- Specific Allelic Capture Proceedings



Sequencing of captured sequences

## -- Specific Allelic Capture Proceedings



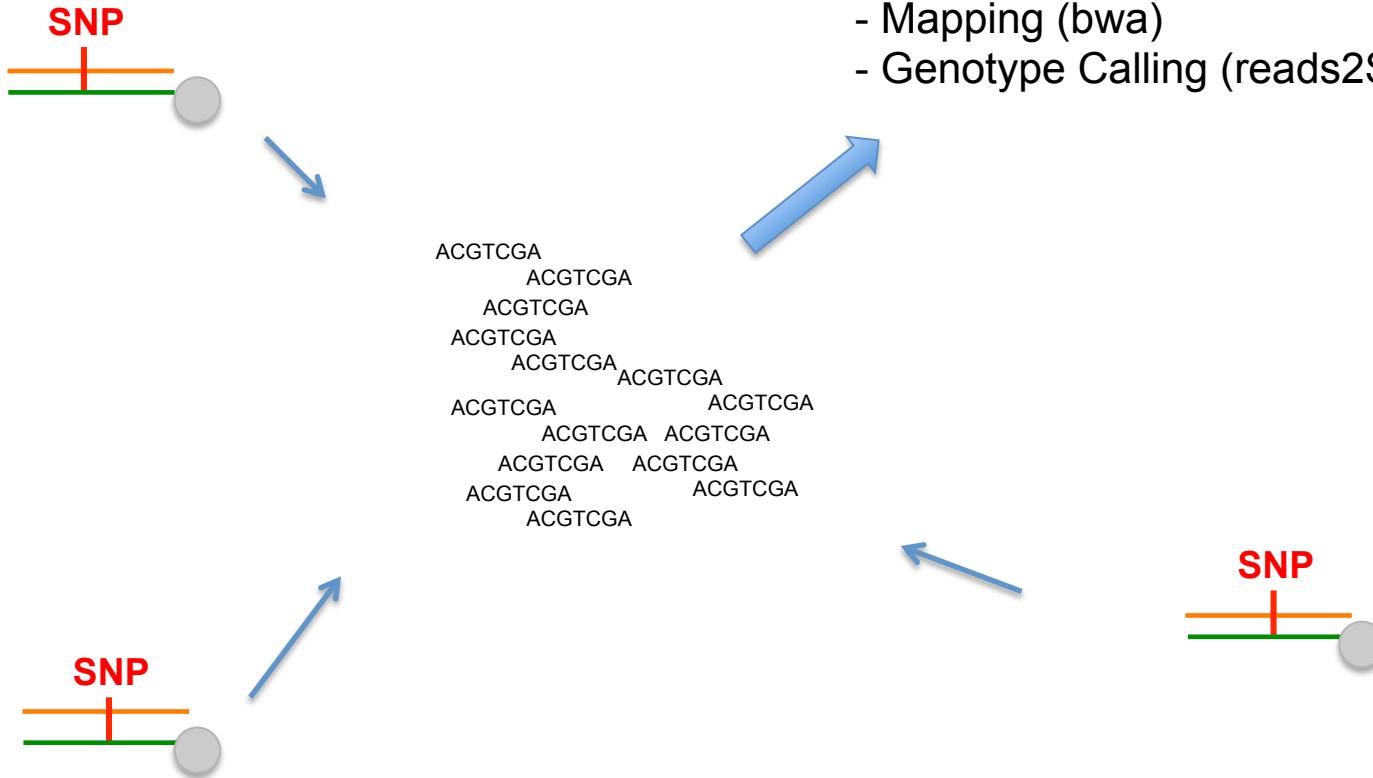
### Sequencing of captured sequences

→ ≈ 300 000 reads / individuals  
→ 48 individuals

## -- Specific Allelic Capture Proceedings

## Classical Pipeline for Genotype calling

- Cleaning reads
  - Mapping (bwa)
  - Genotype Calling (reads2SNP)



## -- How to genotype a Bi-parental Population of Durum Wheat ???

### Specific Allelic Capture

- ✓ Low cost : about 25 Euros / Accession (*Rohland and all, 2012*)
- ✓ Choice of the targeted SNPs
- ✓ Easy creation of KASPAR Markers
- ✓ Enough SNP for a genetic map

### DNA Array

- ✓ High cost : about 60 Euros / Accessions
- ✓ High number of SNPs
- ✓ No Array for Durum Wheat until recently