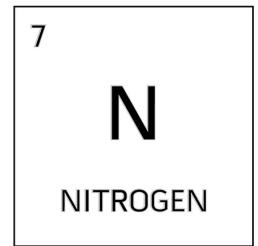
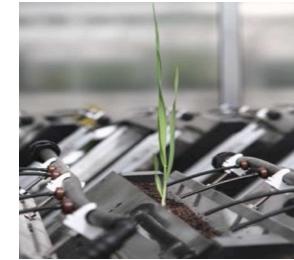
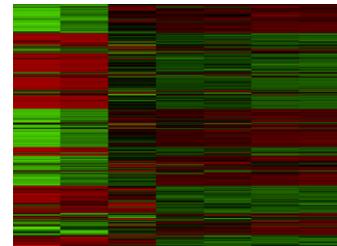
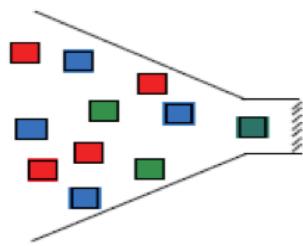


Selection for adaptation to high nitrogen was the driving force of tetraploid wheat domestication

RNA-Seq day, Avignon, November 2016



Authors: Y. Holtz, T. Gioia, S. Santoni, V. Ranwez, R. Beleggia, D. Rau, J. David, R. Papa

-- The [short] history of durum wheat



T. t. dicoccoides
Wild Emmer

Brittle rachis
hulled grain
2 kernels per spikelet

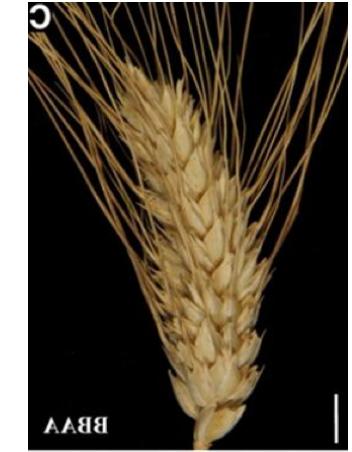
Domestication



T. t. dicoccum
Emmer

Non brittle rachis
hulled grain
2 kernels per spikelet

Selection /
Secondary
domestication



T. t. durum
Durum wheat

Non brittle rachis
naked grain
>2 kernels per spikelet
Large grain
Semi dwarf
Adapted to modern
agronomical practices
(N?)

Dubcovsky et al. 2007

Tanno and Willcox, 2006

Zohary et al. 2012

Maier, 1996

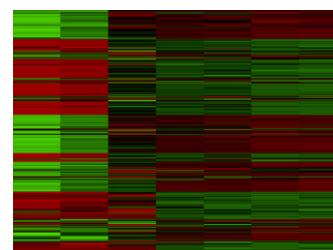
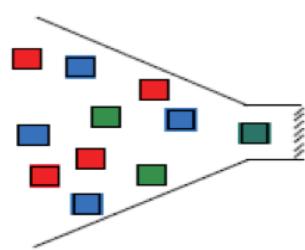
Genotyping level

- ✓ Is RNA-Seq suitable to genotype durum wheat & study evolution?
- ✓ What is the effect of evolution at the nucleotidic level?
- ✓ Pattern of diversity along chromosomes?

Expression level

- ✓ Can we trust the gene expression data coming from RNA-seq?
- ✓ What is the effect of evolution on genes expression?
- ✓ Does Nitrogen play an important role?

The Experiment



7
N
NITROGEN

T. t. dicoccoides

10 accessions

T. t. dicoccum

10 accessions

T. t. durum

12 accessions

32 Accessions

OPEN  ACCESS Freely available online

PLOS ONE

Genetic Diversity and Population Structure of Tetraploid Wheats (*Triticum turgidum* L.) Estimated by SSR, DArT and Pedigree Data

Giovanni Laidò¹, Giacomo Mangini², Francesca Taranto², Agata Gadaleta², Antonio Blanco², Luigi Cattivelli^{1*}, Daniela Marone¹, Anna M. Mastrangelo¹, Roberto Papa¹, Pasquale De Vita^{1*}

¹Consiglio per la Ricerca e la sperimentazione in Agricoltura, Cereal Research Centre, Foggia, Italy, ²Department of Soil, Plant, and Food Sciences, Section of Genetics and Plant Breeding, University of Bari, Via Amendola, Bari, Italy

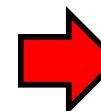
OPEN  ACCESS Freely available online

PLOS ONE

Linkage Disequilibrium and Genome-Wide Association Mapping in Tetraploid Wheat (*Triticum turgidum* L.)

Giovanni Laidò, Daniela Marone, Maria A. Russo, Salvatore A. Coleccchia, Anna M. Mastrangelo, Pasquale De Vita, Roberto Papa*

Consiglio per la Ricerca e la sperimentazione in Agricoltura, Cereal Research Centre, Foggia, Italy



Main aim

Maximization of
the genetic
diversity of
each subspecies

T. t. dicoccoides
10 accessions

T. t. dicoccum
10 accessions

T. t. durum
12 accessions

32 Accessions



2 Treatments

- Low Nitrogen
- High Nitrogen

T. t. dicoccoides
10 accessions

T. t. dicoccum
10 accessions

T. t. durum
12 accessions

32 Accessions



2 Treatments

- Low Nitrogen
- High Nitrogen



2 Repetitions

T. t. dicoccoides
10 accessions

T. t. dicoccum
10 accessions

T. t. durum
12 accessions

32 Accessions



2 Treatments

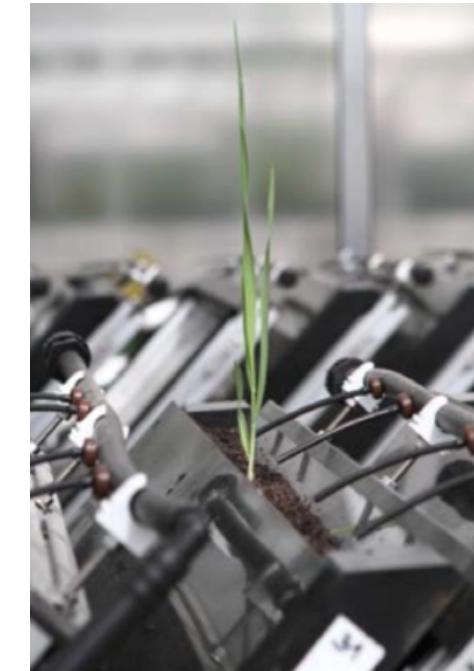
- Low Nitrogen
- High Nitrogen



2 Repetitions

- ✓ 128 accessions in total
- ✓ Possibility to test Nitrogen effect
- ✓ Possibility to analyse domestication and selection effect

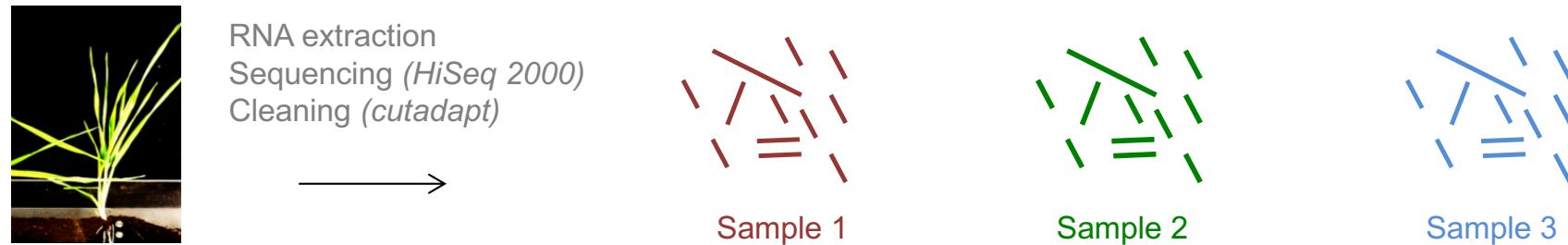
-- Growth condition & phenotyping



 **JÜLICH**
FORSCHUNGSZENTRUM

- ✓ Growth & phenotyping on a growscreen-Rhizo **phenotyping robot**
- ✓ First publication concerning **root architecture** (*Gioia et al. 2015*)
- ✓ **RNA extraction**

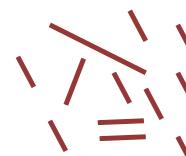
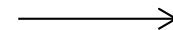
-- Bioinformatics pipeline



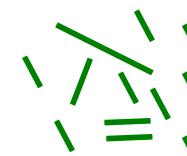
-- Bioinformatics pipeline



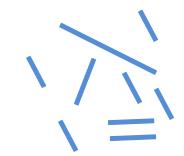
RNA extraction
Sequencing (*HiSeq 2000*)
Cleaning (*cutadapt*)



Sample 1



Sample 2



Sample 3

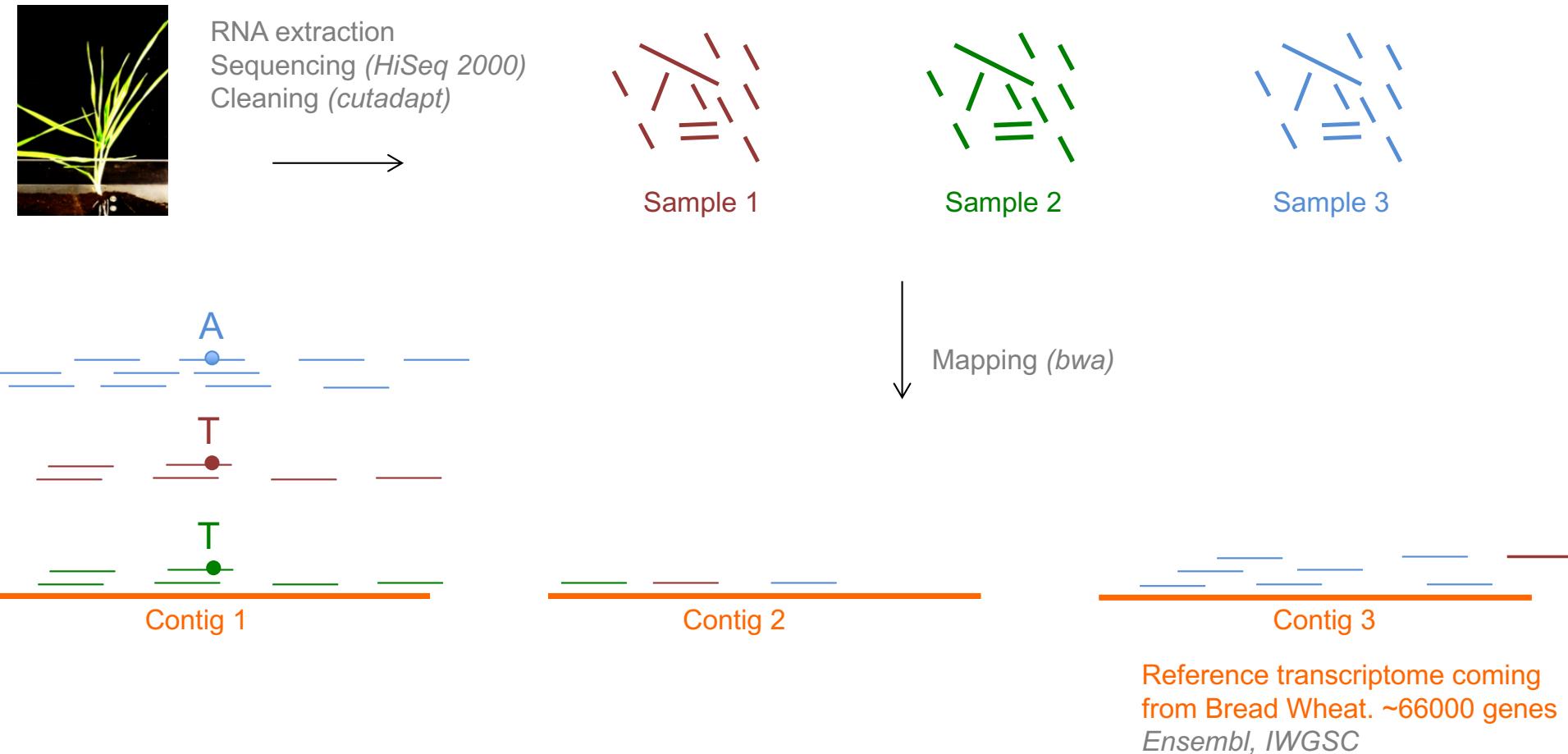
Contig 1

Contig 2

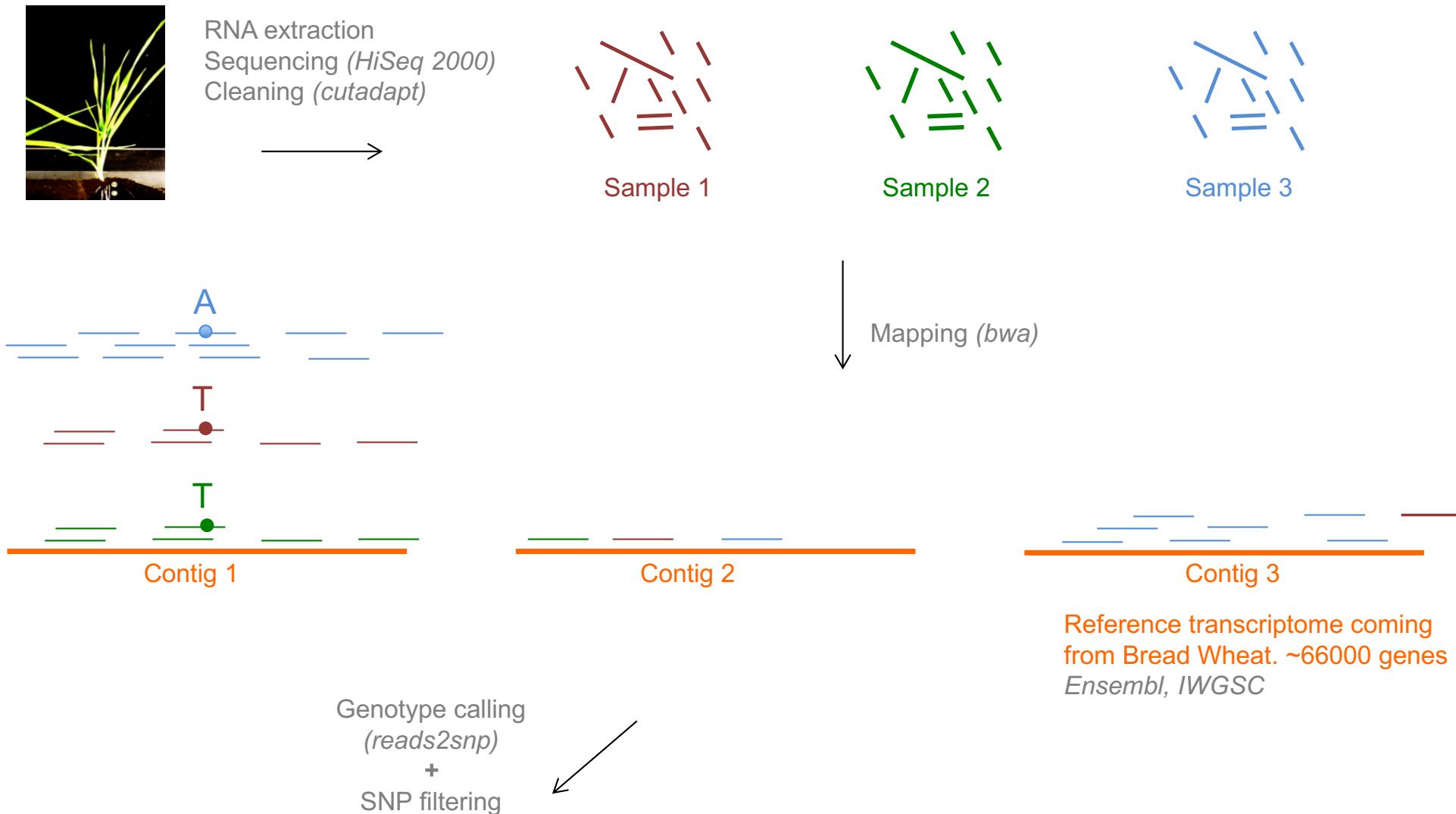
Contig 3

Reference transcriptome coming
from Bread Wheat. ~66000 genes
Ensembl, IWGSC

-- Bioinformatics pipeline

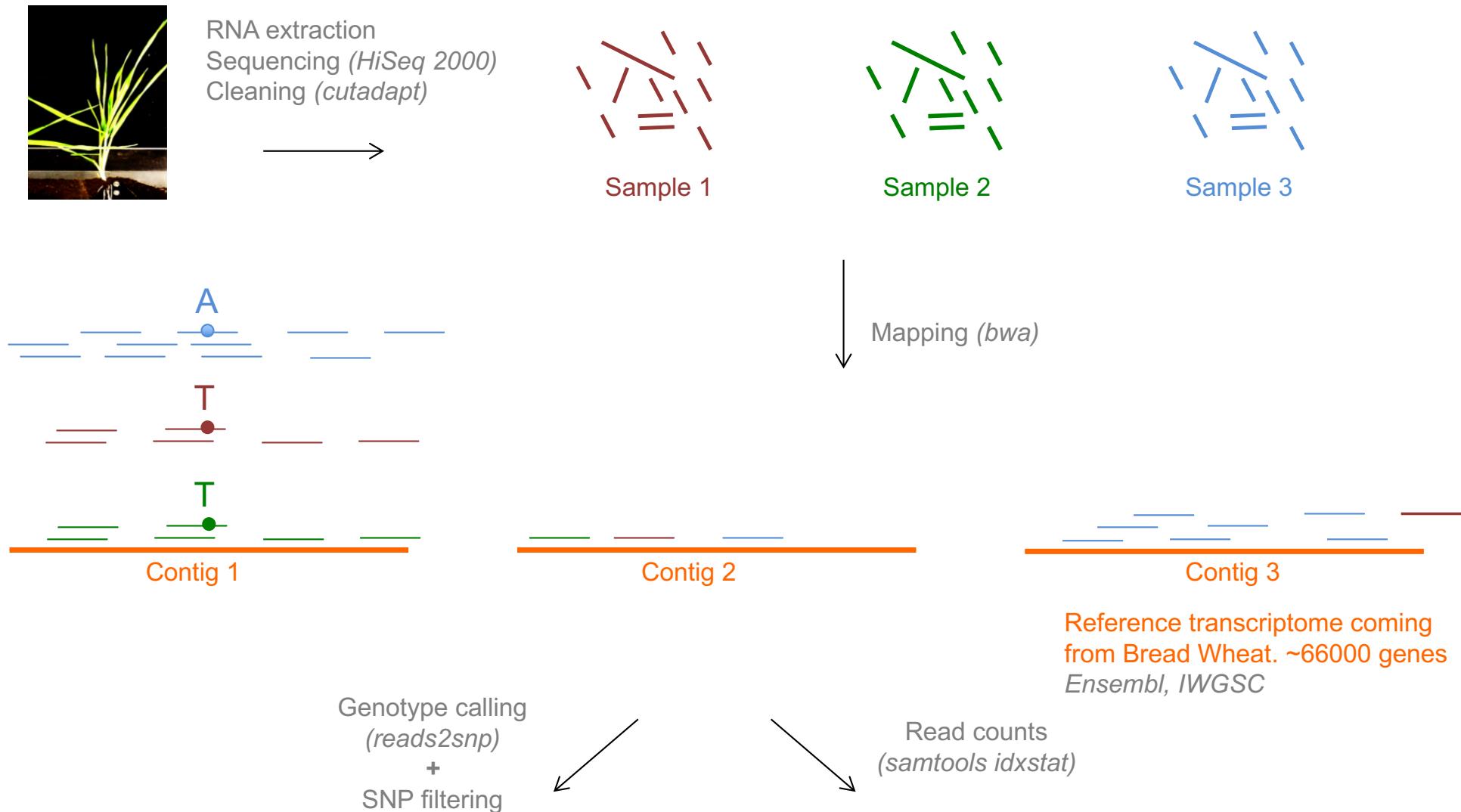


-- Bioinformatics pipeline



	Pos	Sample 1	Sample 2	Sample 3
Contig 1	425	T	T	A
Contig 12	15	A	C	A
Contig 74	362	A	A	C

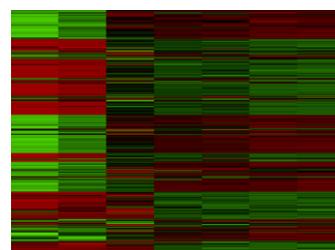
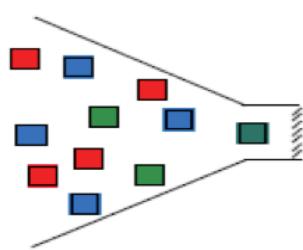
-- Bioinformatics pipeline



	Pos	Sample 1	Sample 2	Sample 3
Contig 1	425	T	T	A
Contig 12	15	A	C	A
Contig 74	362	A	A	C

	Sample 1	Sample 2	Sample 3
Contig 1	200	204	302
Contig 2	2	2	1
Contig 3	0	2	475

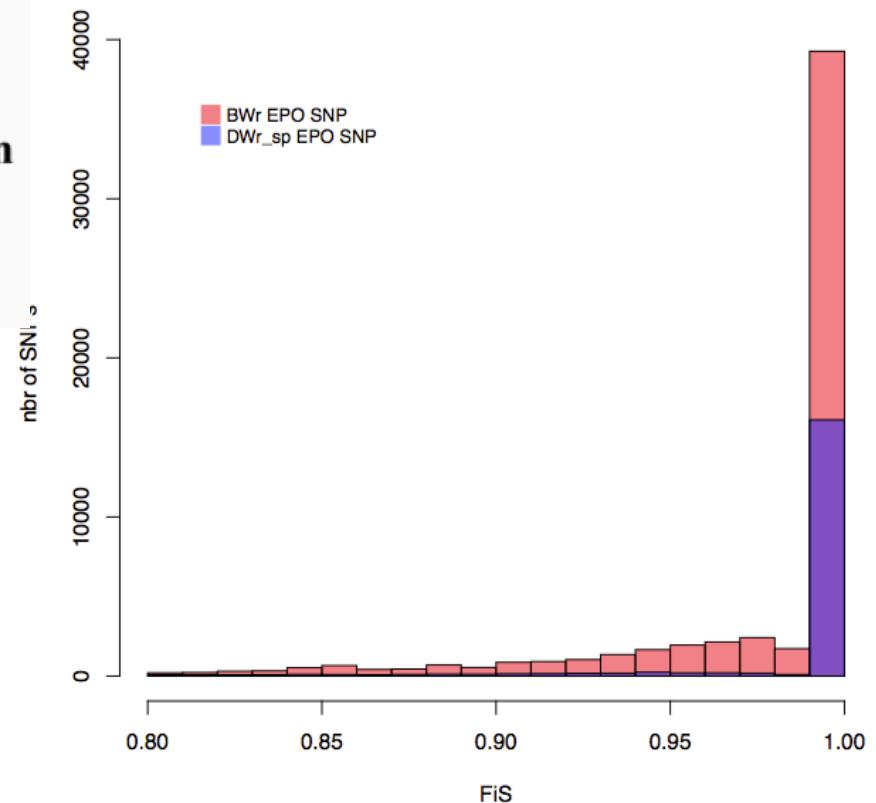
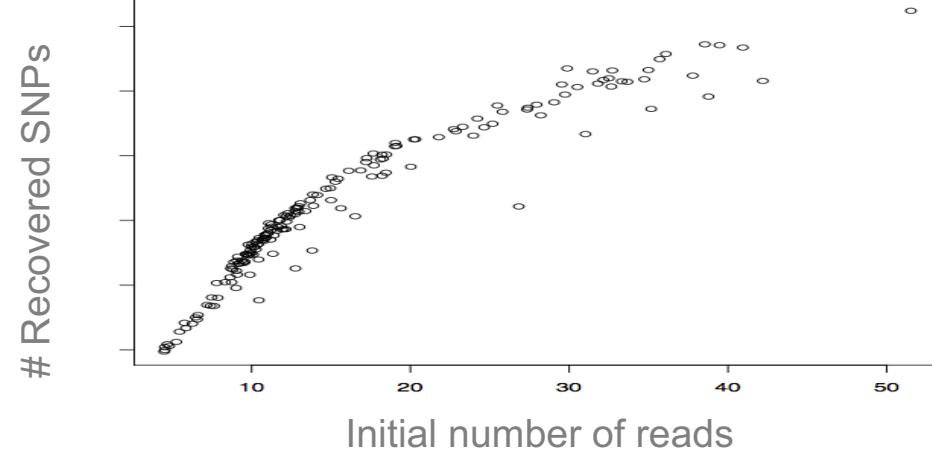
Genomic results



7
N
NITROGEN

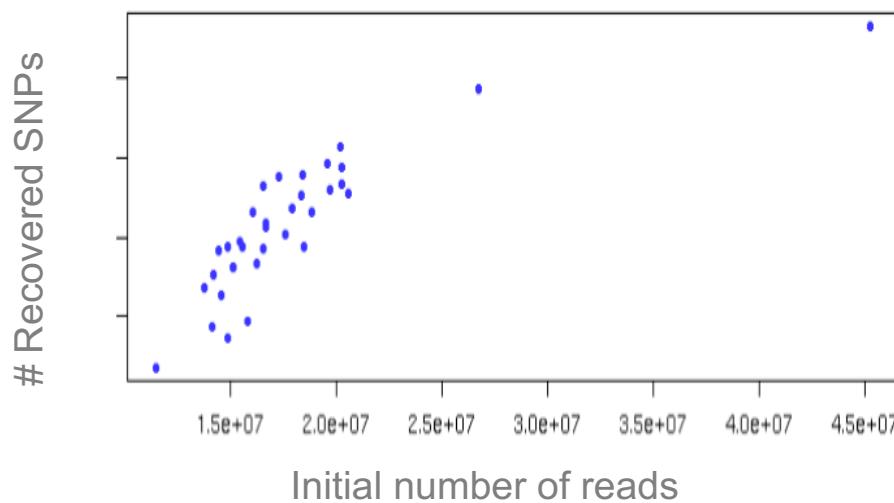
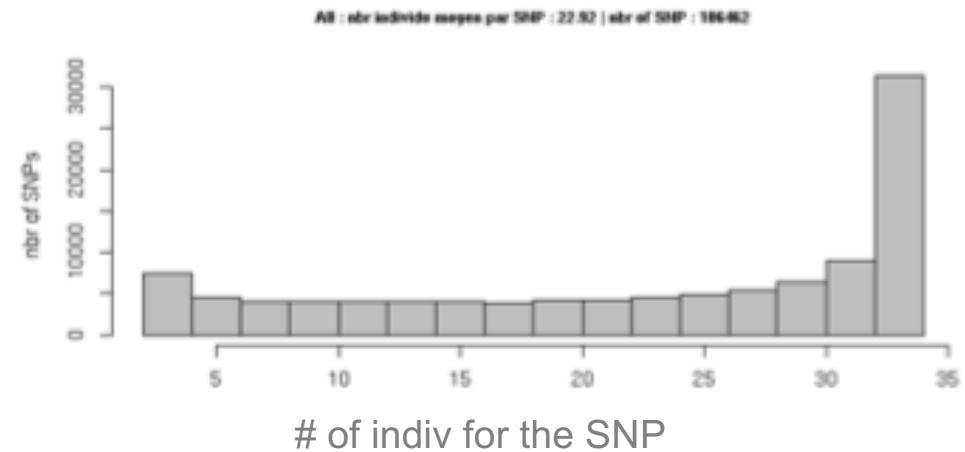
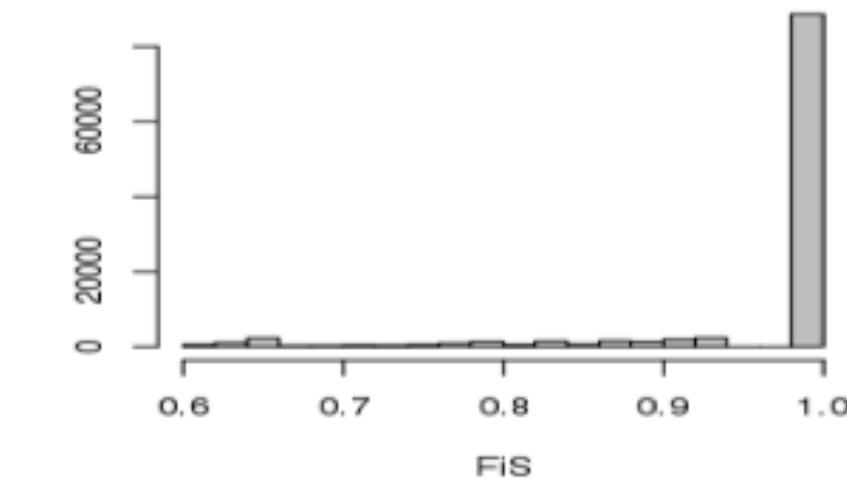
Genotyping by sequencing transcriptomes in an evolutionary pre-breeding durum wheat population

Jacques David · Yan Holtz · Vincent Ranwez · Sylvain Santoni ·
Gautier Sarah · Morgane Ardisson · Gérard Poux · Frédéric Choulet ·
Clémence Genthon · Pierre Roumet · Muriel Tavaud-Pirra



- ✓ Several thousands of SNPs can be detected in durum wheat using RNA-seq
- ✓ The method is highly repeatable, called genotypes can be trusted
- ✓ It is possible to disentangle homeogenomes
- ✓ Quantity of missing data depends of genes expression & initial number of reads

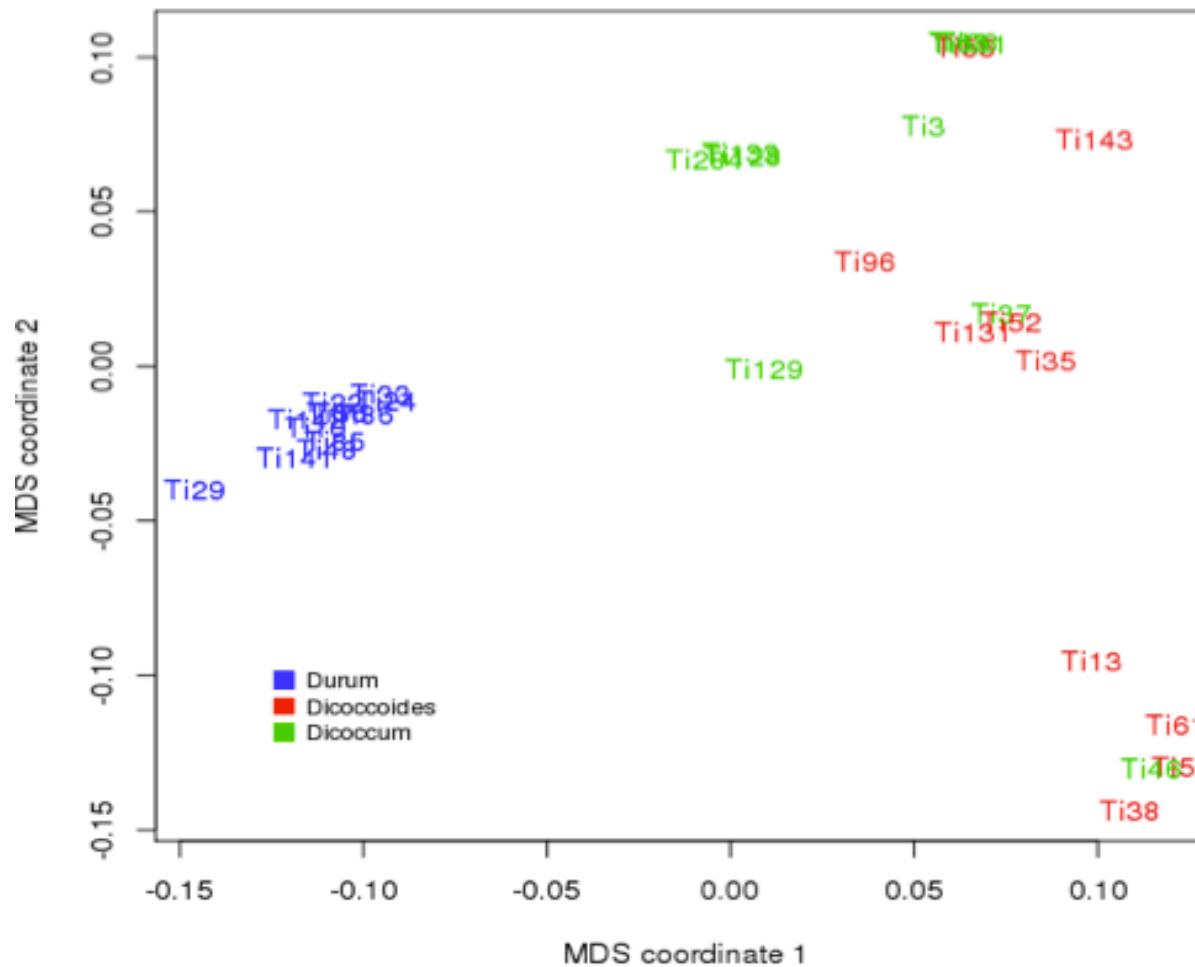
-- RNAseq is suitable to genotype Durum Wheat.



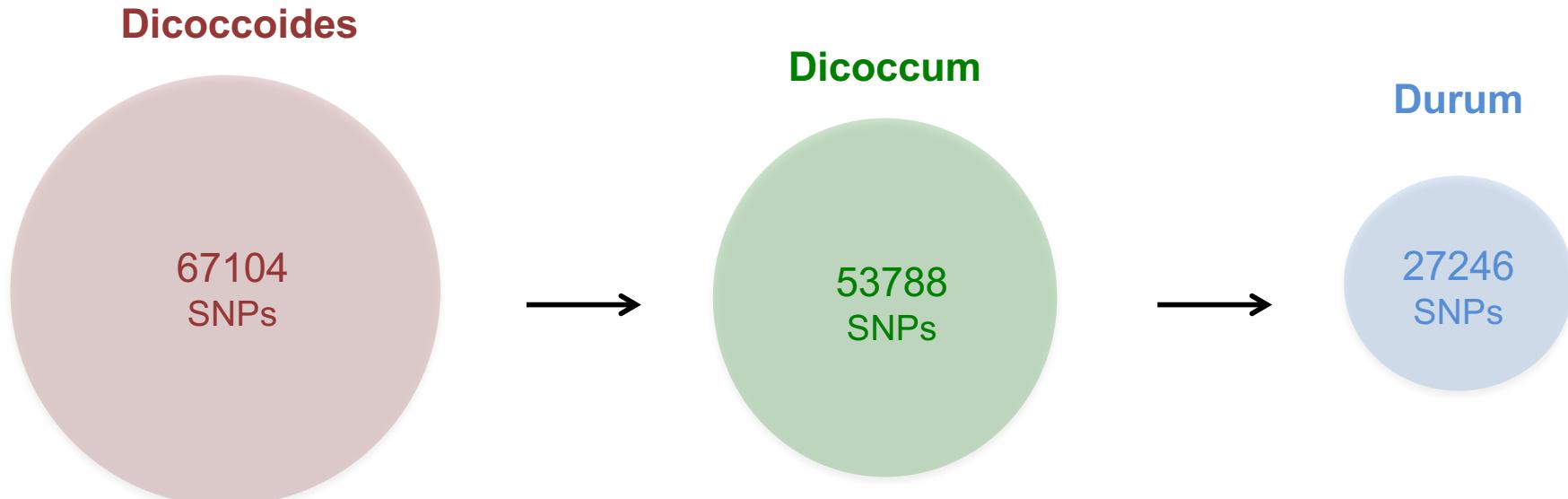
- ✓ 107267 clean SNPs detected
- ✓ Proportion of missing data depends of initial number of reads & gene expression
- ✓ Very few heterozygote loci

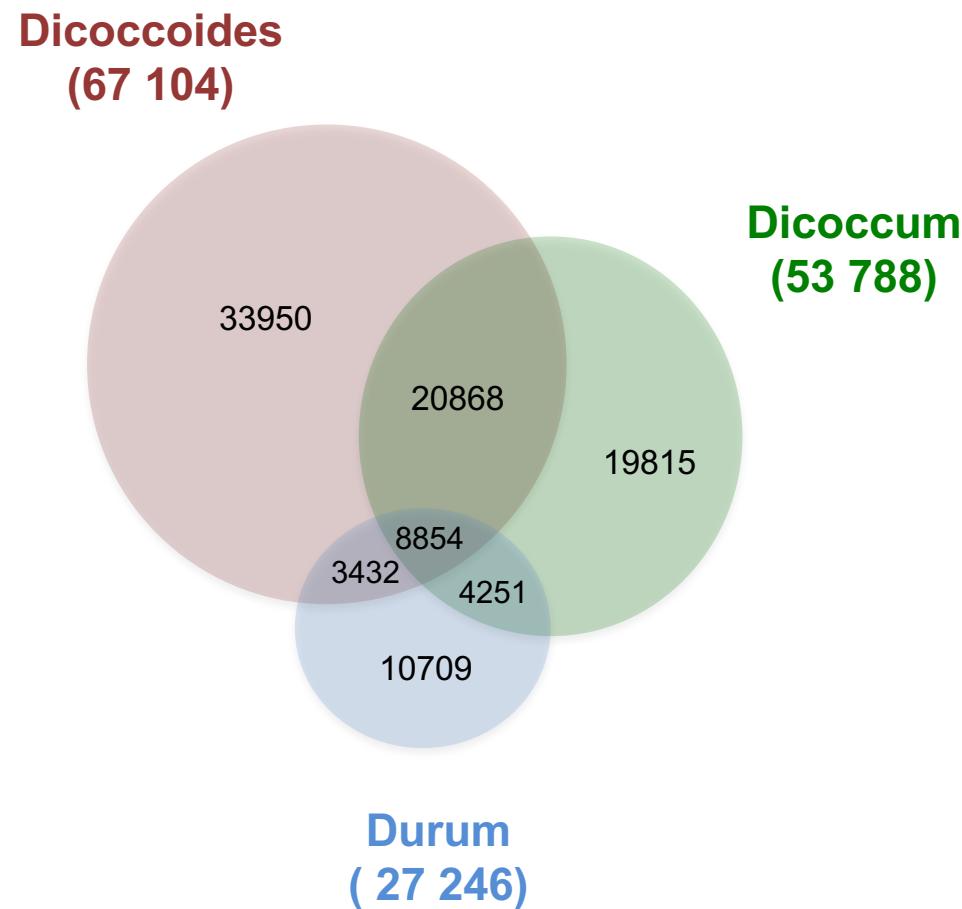
-- Population Structure

Multi Dimensionnal Scaling (MDS) applied to the distance matrix

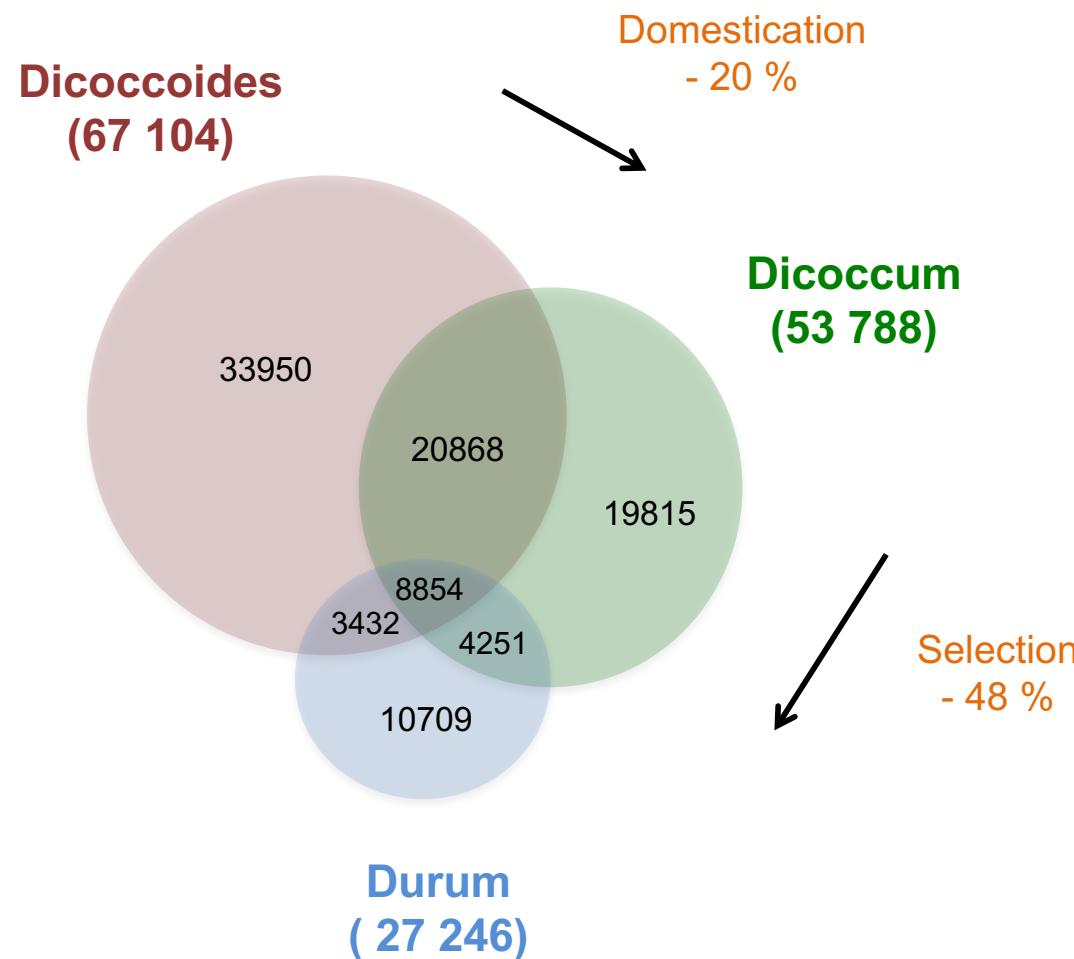


- ✓ Durum is clearly distinct from the other subspecies
- ✓ Durum shows less diversity: every genotypes are really close to each other
- ✓ The distinction between dicoccum and dicoccoides is unclear





- ✓ A huge **loss of diversity** during tetraploid wheat **evolution**
- ✓ Already observed: *Haudry et al. 2007, Reif et al. 2004*
- ✓ But durum still have **specific SNPs...**



- ✓ A huge **loss of diversity** during tetraploid wheat **evolution**
- ✓ Already observed: *Haudry et al. 2007, Reif et al. 2004*
- ✓ But durum still have **specific SNPs**.

Number of SNPs

ind1 AAC**A**GGC**G**TA
Ind2 AAC**C**GGCTTA
Ind3 AAC**C**GGCTTA
Ind4 AAC**C**GGCTTA
Ind5 AAC**C**GGCTTA

SNPs : 2

Number of SNPs / Density of SNPs / N. polymorphic genes

ind1 AAC~~A~~GGC~~G~~T~~A~~

Ind2 AACCGGCTTA

Ind3 AACCGGCTTA

Ind4 AACCGGCTTA

Ind5 AACCGGCTTA

SNPs : 2

Number of SNPs / Density of SNPs / N. polymorphic genes / Pi / PiN / PiS

ind1 AAC~~A~~GGC~~G~~TA
Ind2 AACCGGCTTA
Ind3 AACCGGCTTA
Ind4 AACCGGCTTA
Ind5 AACCGGCTTA

SNPs : 2

ind1 AAC~~A~~GGC~~G~~TA
Ind2 AAC~~A~~GGC~~G~~TA
Ind3 AACCGGCTTA
Ind4 AACCGGCTTA
Ind5 AACCGGCTTA

SNPs : 2

Number of SNPs / Density of SNPs / N. polymorphic genes / Pi / PiN / PiS

$$\Pi = \sum_{ij} x_i x_j \pi_{ij} = \sum_{i=1}^n \left(\sum_{j=1}^i x_i x_j \pi_{ij} \right), \text{ Nei and Li, 1979}$$

ind1 AACAGGC $\textcolor{orange}{G}$ T $\textcolor{black}{A}$
Ind2 AACCGGGCTTA
Ind3 AACCGGGCTTA
Ind4 AACCGGGCTTA
Ind5 AACCGGGCTTA

SNPs : 2

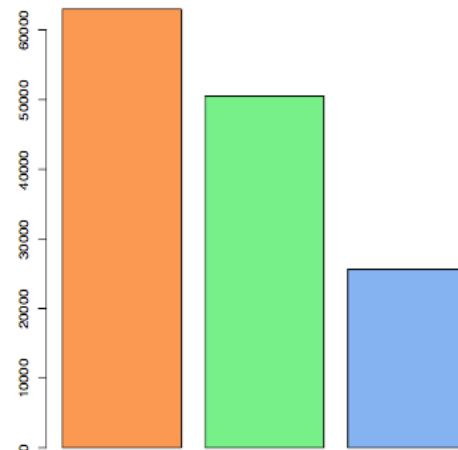
Diversity : $8*0 + 2*0.2 = \textcolor{orange}{0.4}$

ind1 AACAGGC $\textcolor{orange}{G}$ T $\textcolor{black}{A}$
Ind2 AACAGGC $\textcolor{orange}{G}$ T $\textcolor{black}{A}$
Ind3 AACCGGGCTTA
Ind4 AACCGGGCTTA
Ind5 AACCGGGCTTA

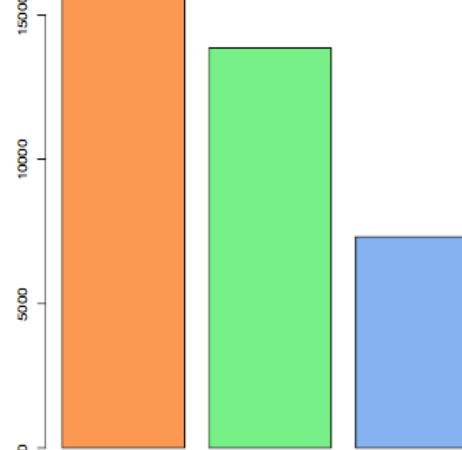
SNPs : 2

Diversity : $8*0 + 2*0.4 = \textcolor{orange}{0.8}$

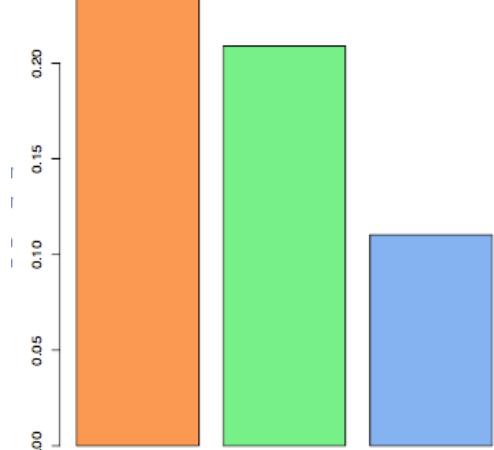
-- The loss of diversity is true for every estimators



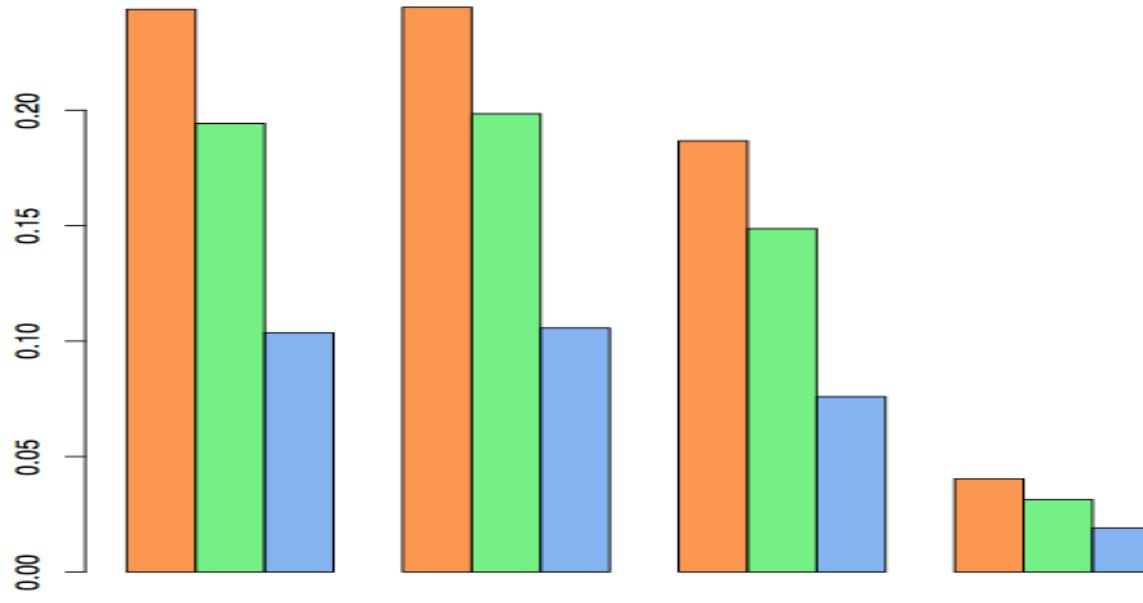
Number of SNPs



N. polymorphic genes



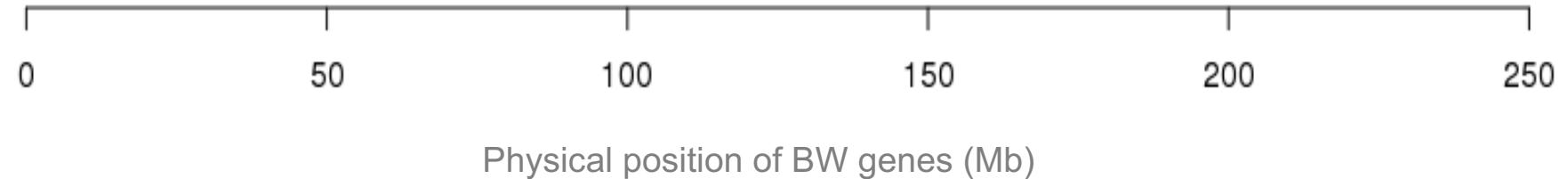
Density of SNPs



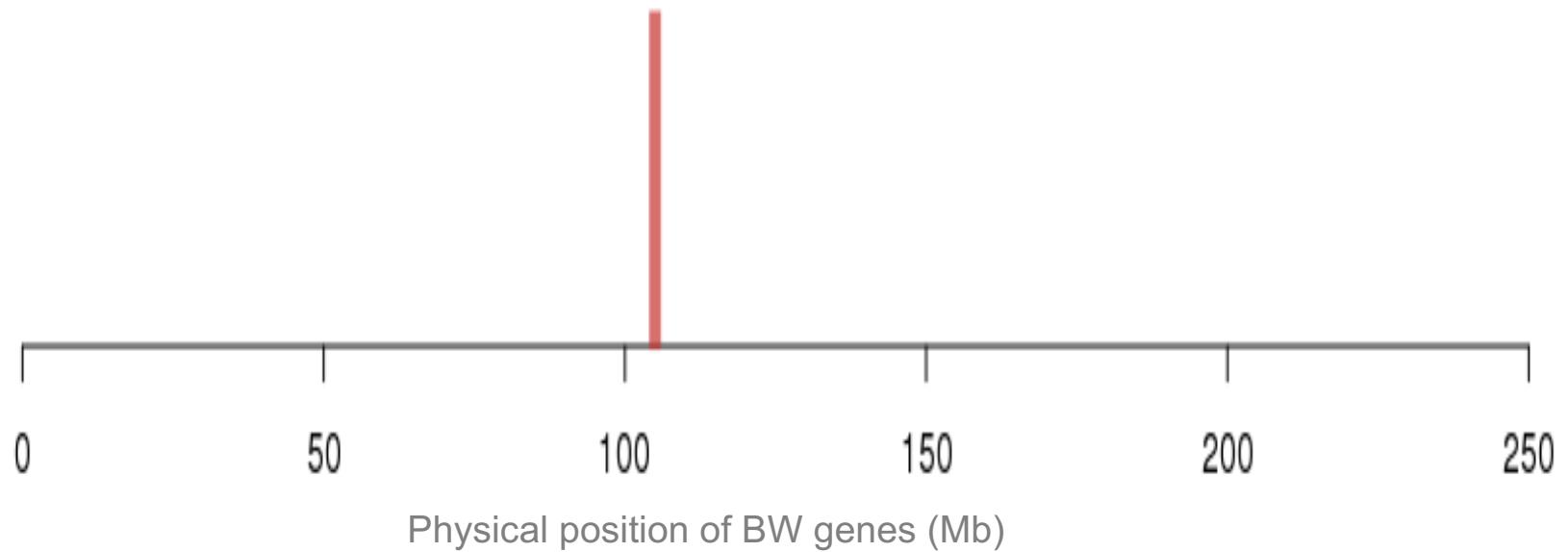
Nucleotidic diversity : Pi

- ✓ Domestication : - 20 %
- ✓ Selection : - 50%
- ✓ True in CDS and UTR

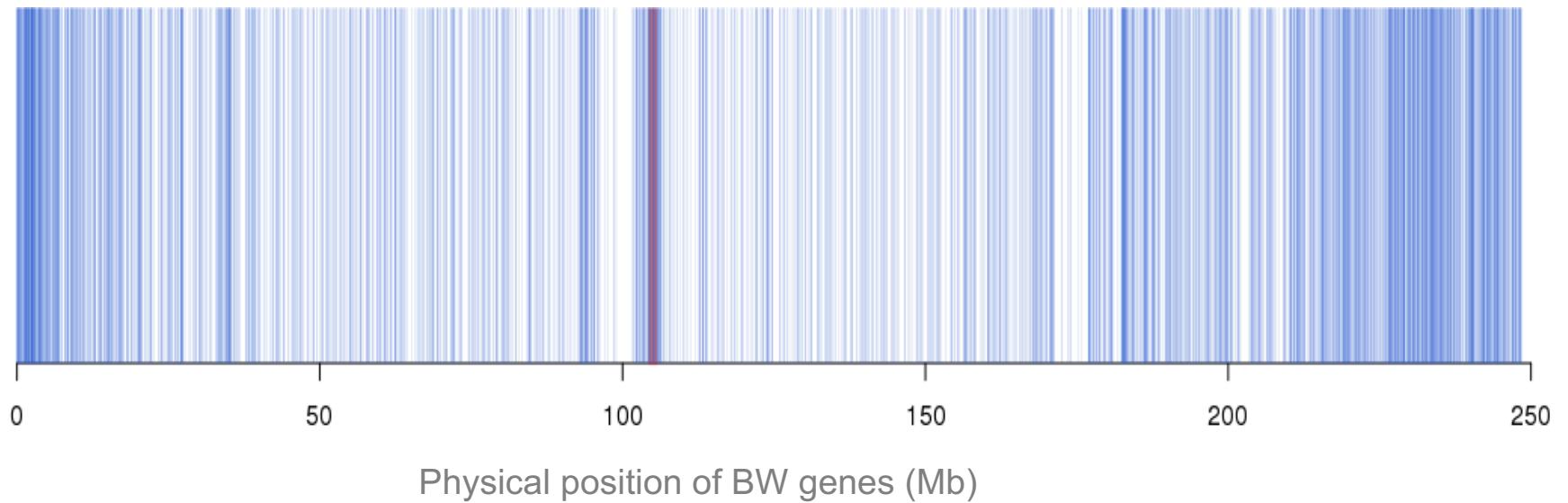
-- Nucleotidic diversity along chromosomes 1A ?



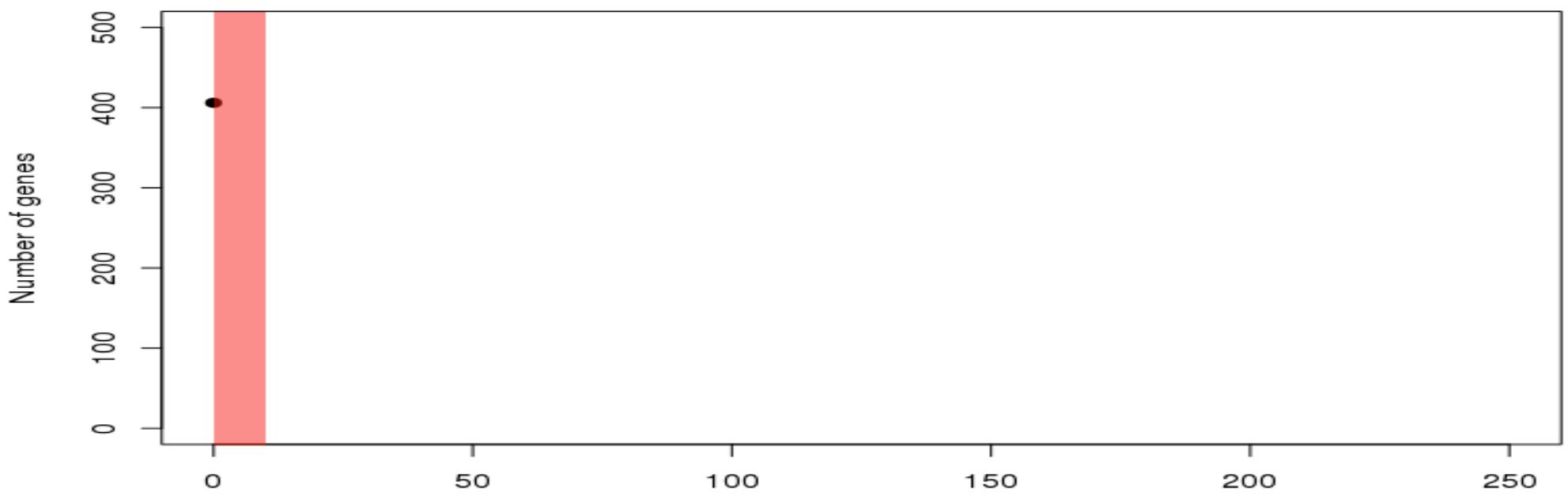
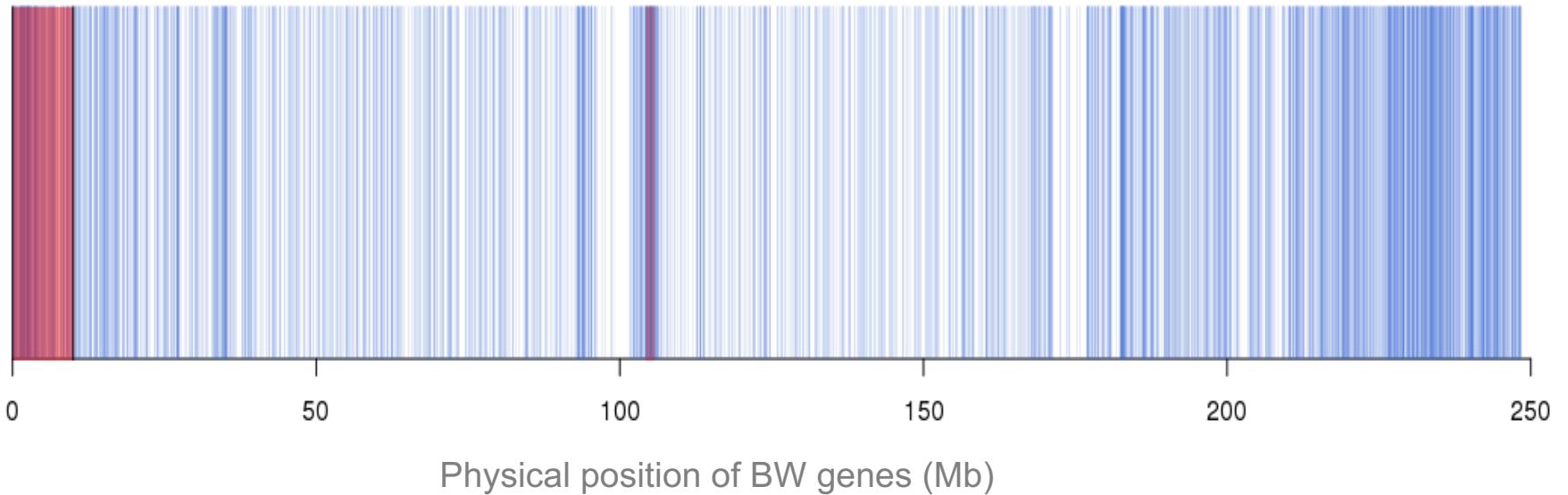
-- Nucleotidic diversity along chromosomes 1A ?



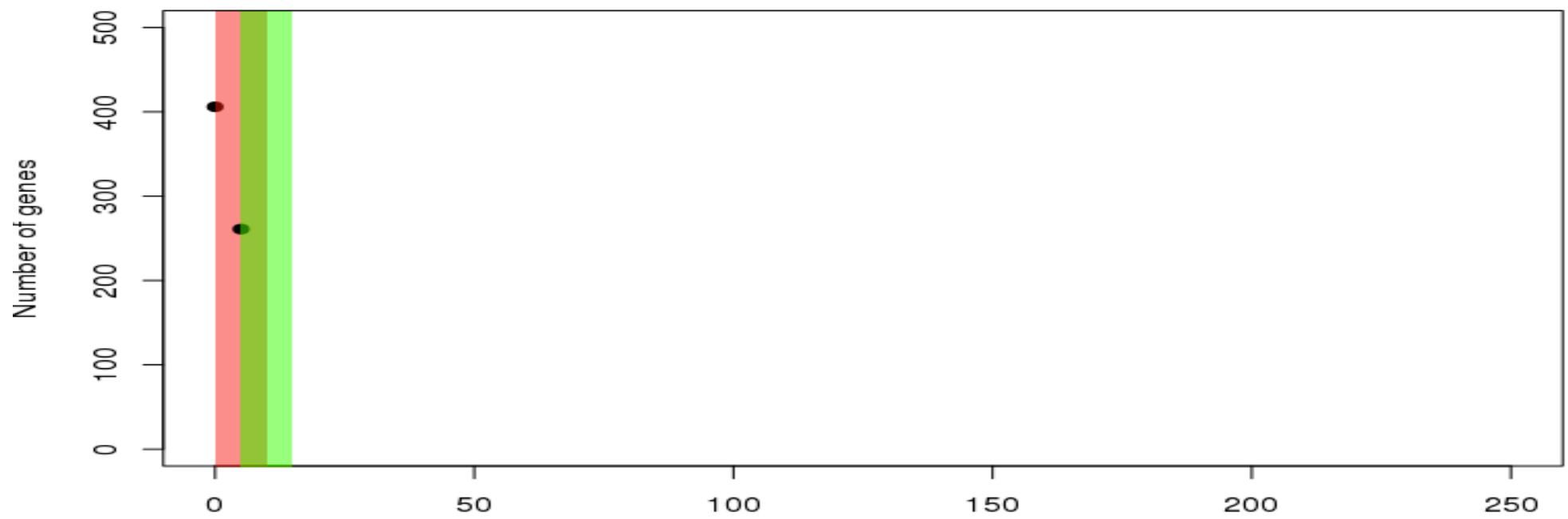
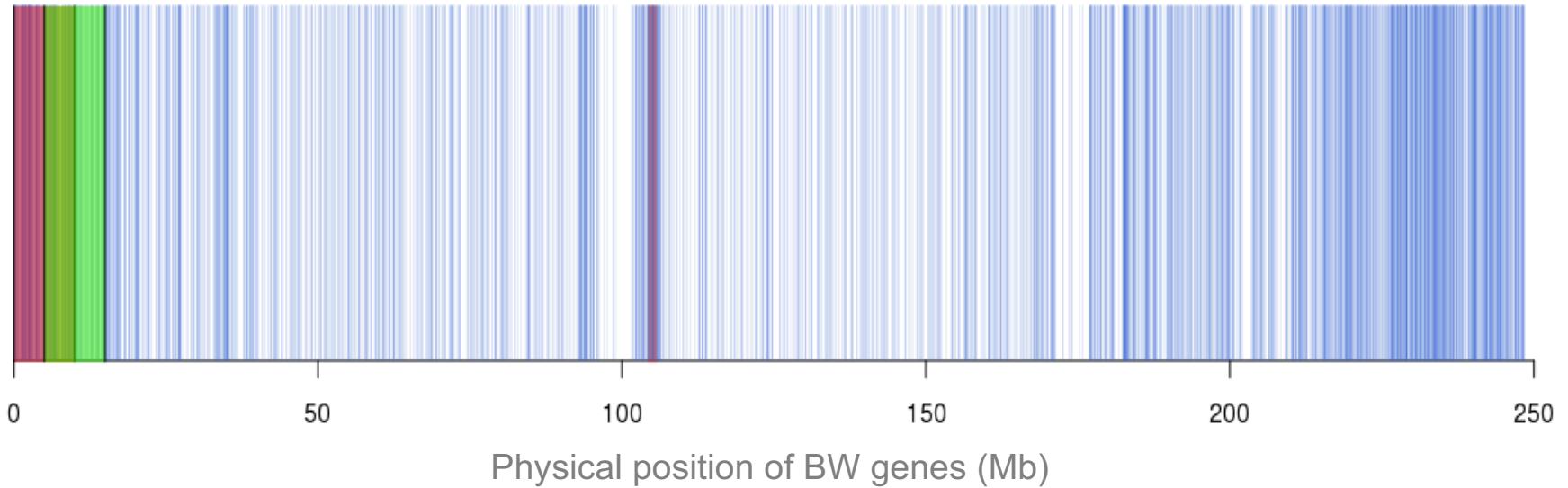
-- Nucleotidic diversity along chromosomes 1A ?



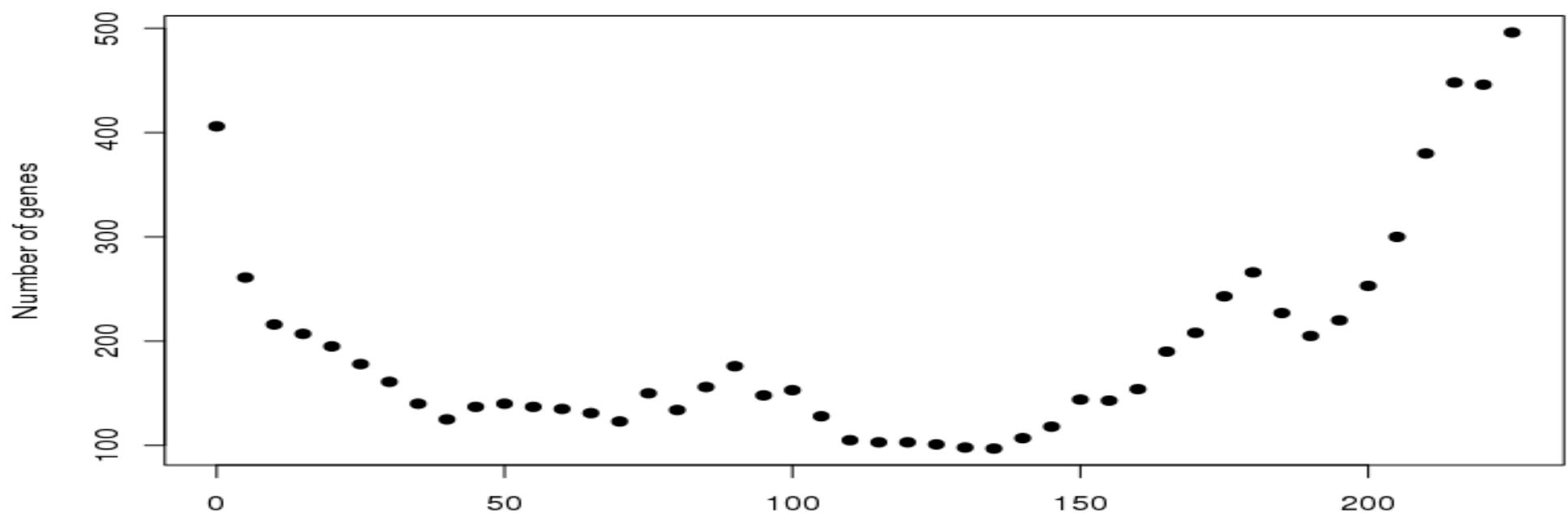
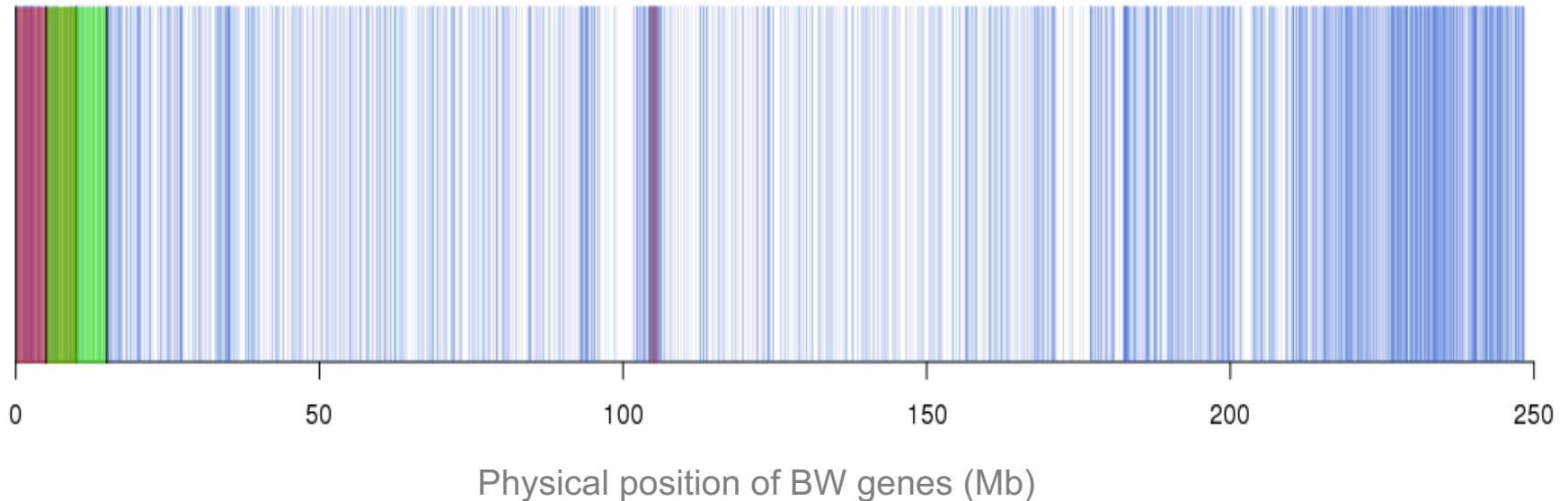
-- Nucleotidic diversity along chromosomes ?



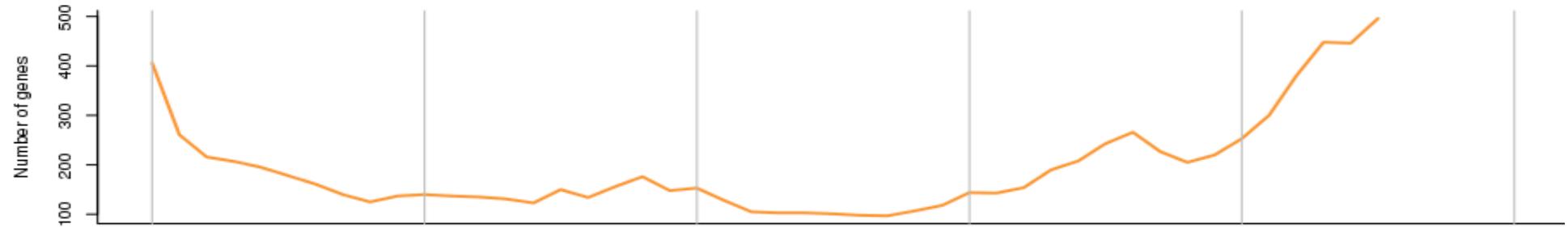
-- Nucleotidic diversity along chromosomes 1A ?



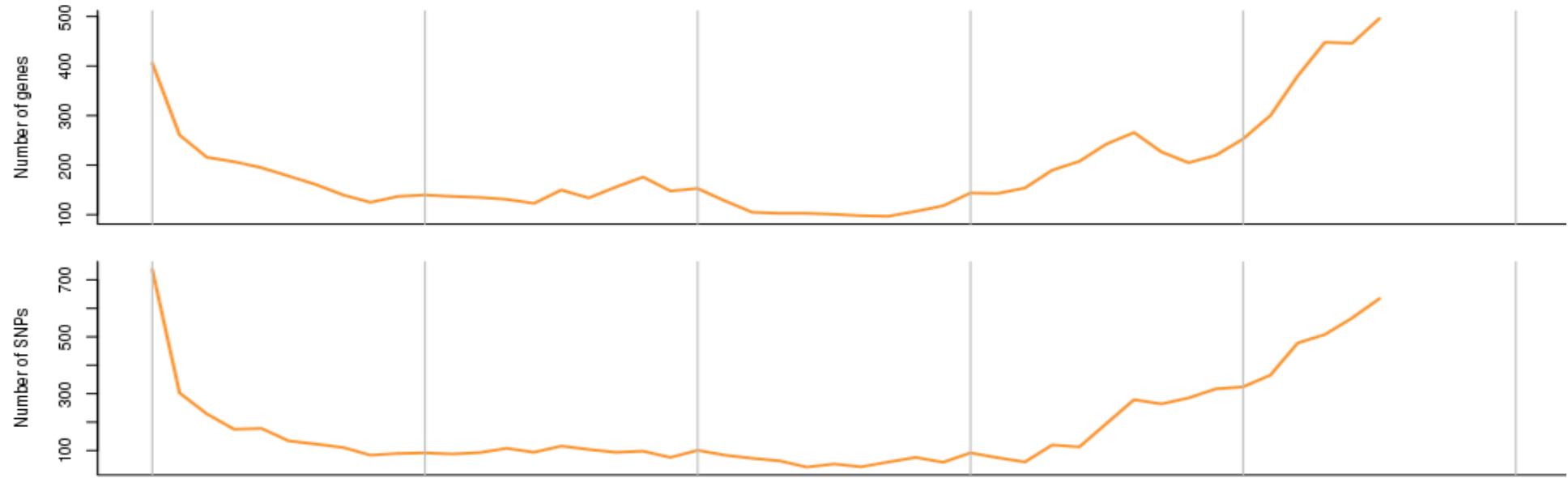
-- Nucleotidic diversity along chromosomes 1A ?



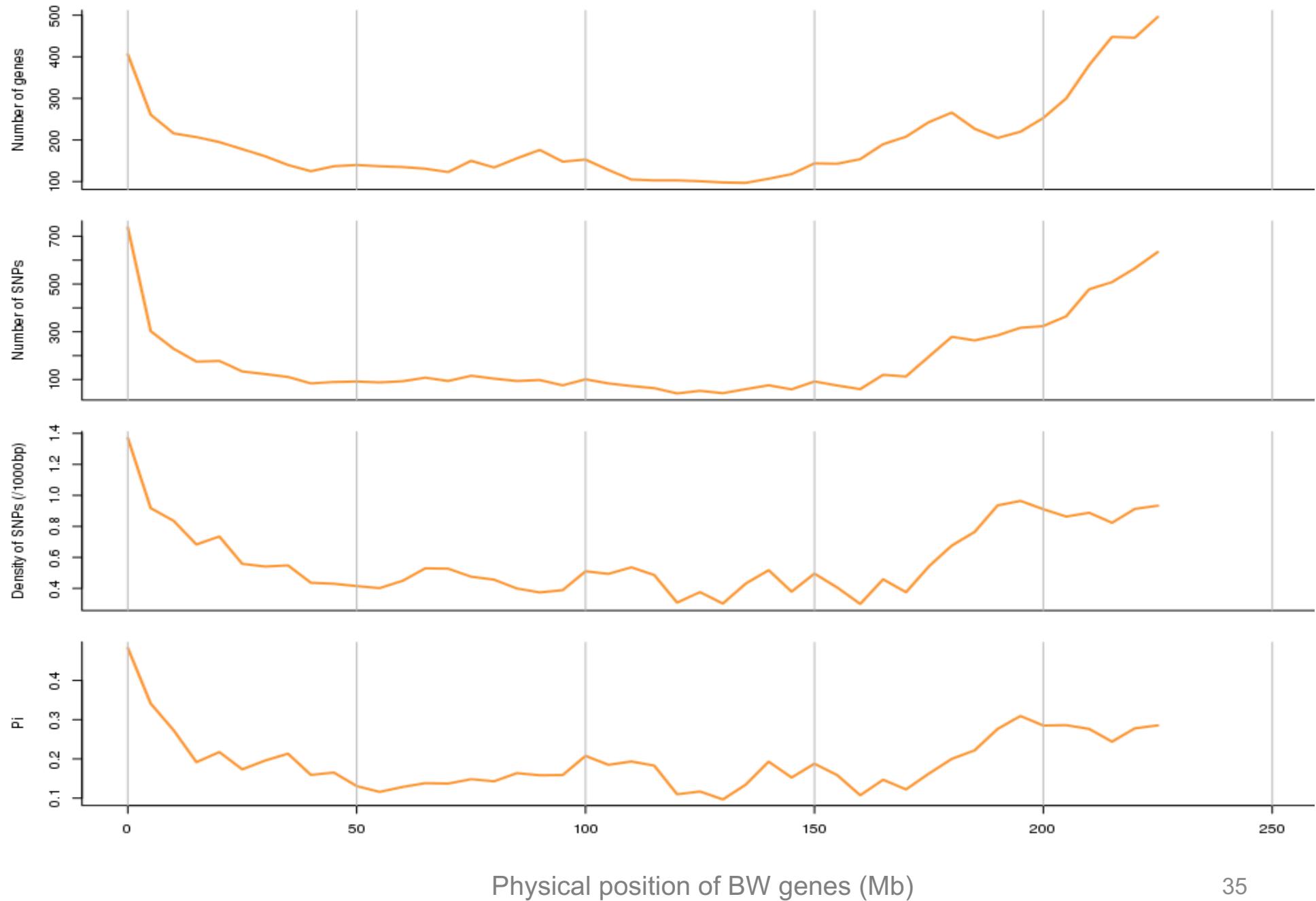
-- Nucleotidic diversity along chromosomes ?

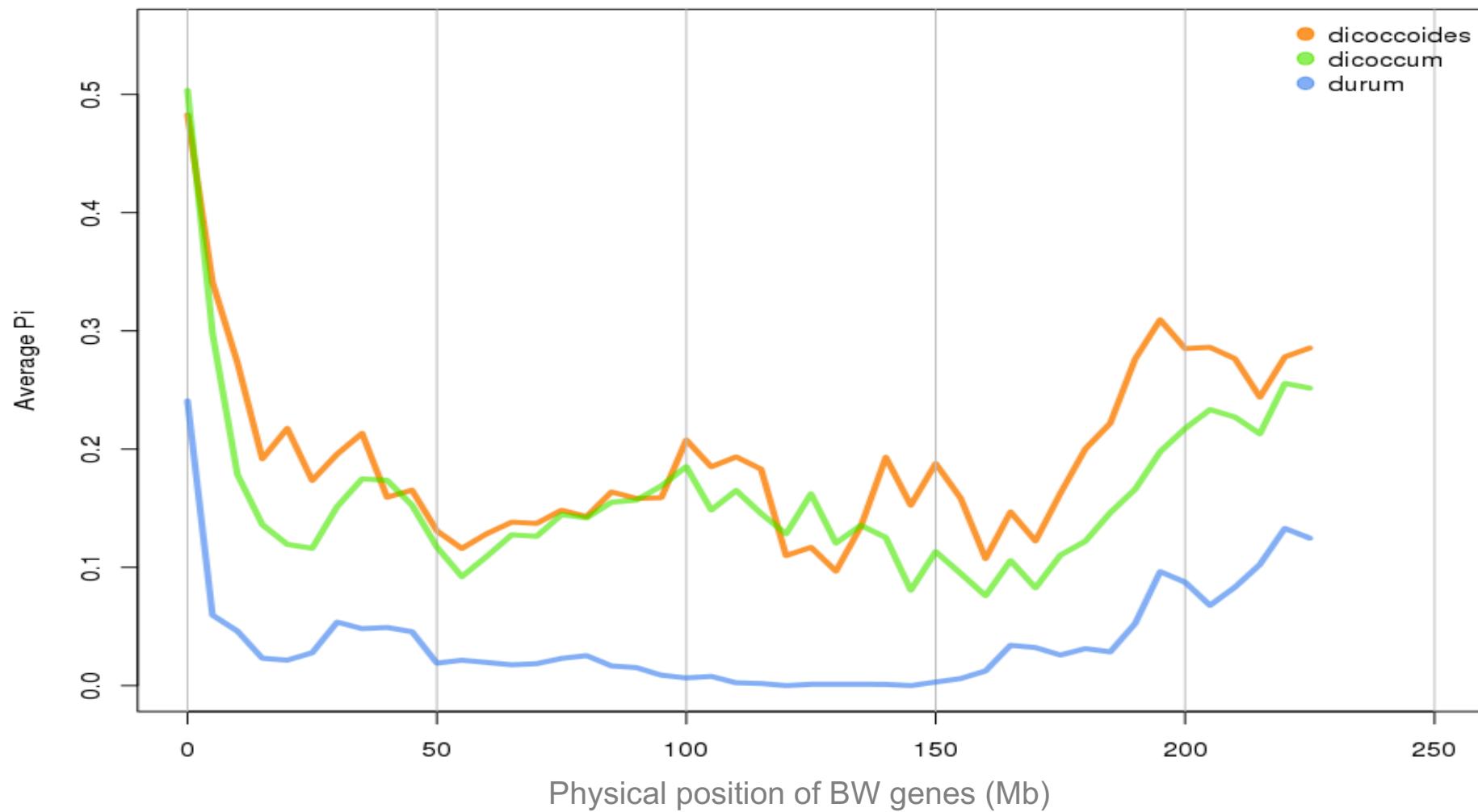


-- Nucleotidic diversity along chromosomes ?



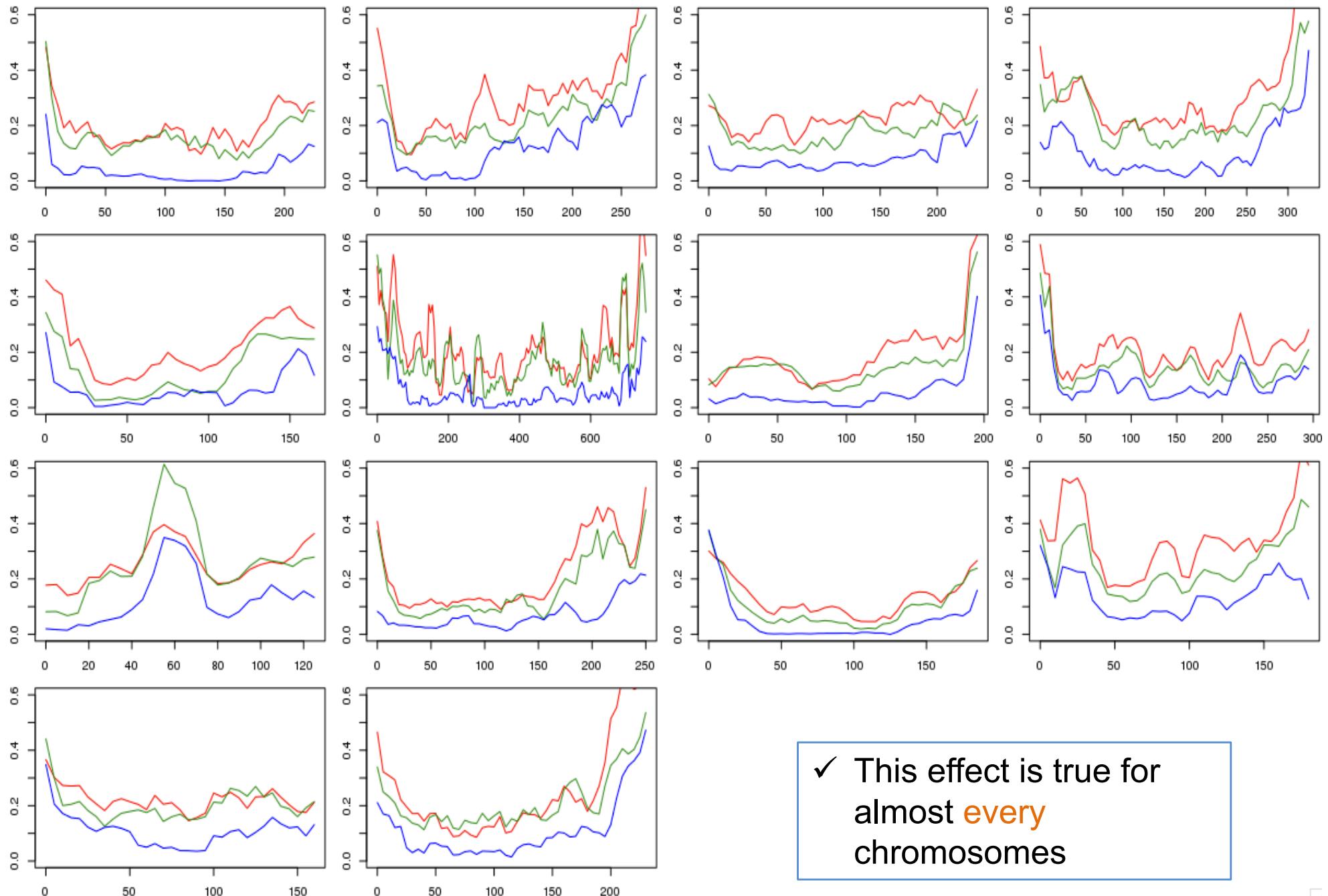
-- More nucleotidic diversity in distal parts !





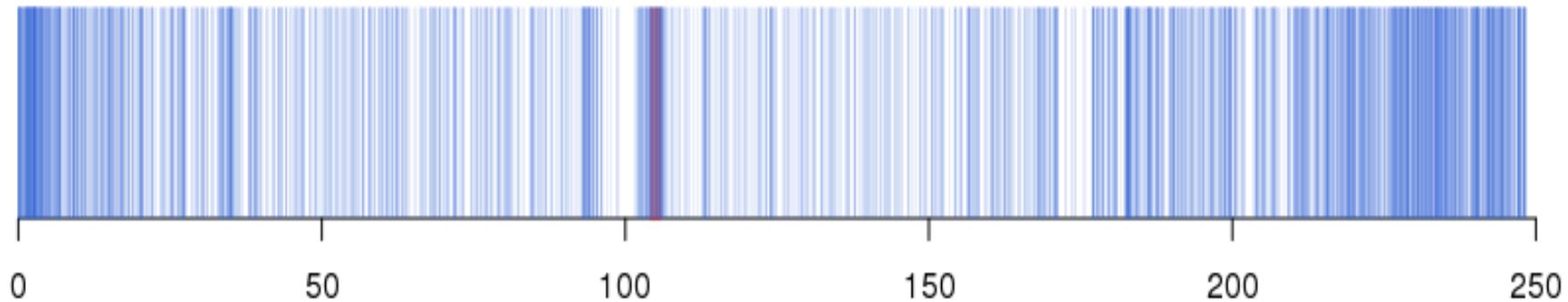
- ✓ More nucleotidic diversity in the **distal parts**
- ✓ Dicoccoides > Dicoccum **>> Durum**
- ✓ Durum has segments with absolutely **no polymorphism**
- ✓ **Selective sweep** visible on durum?

-- Nucleotidic diversity, evolution and chromosome position

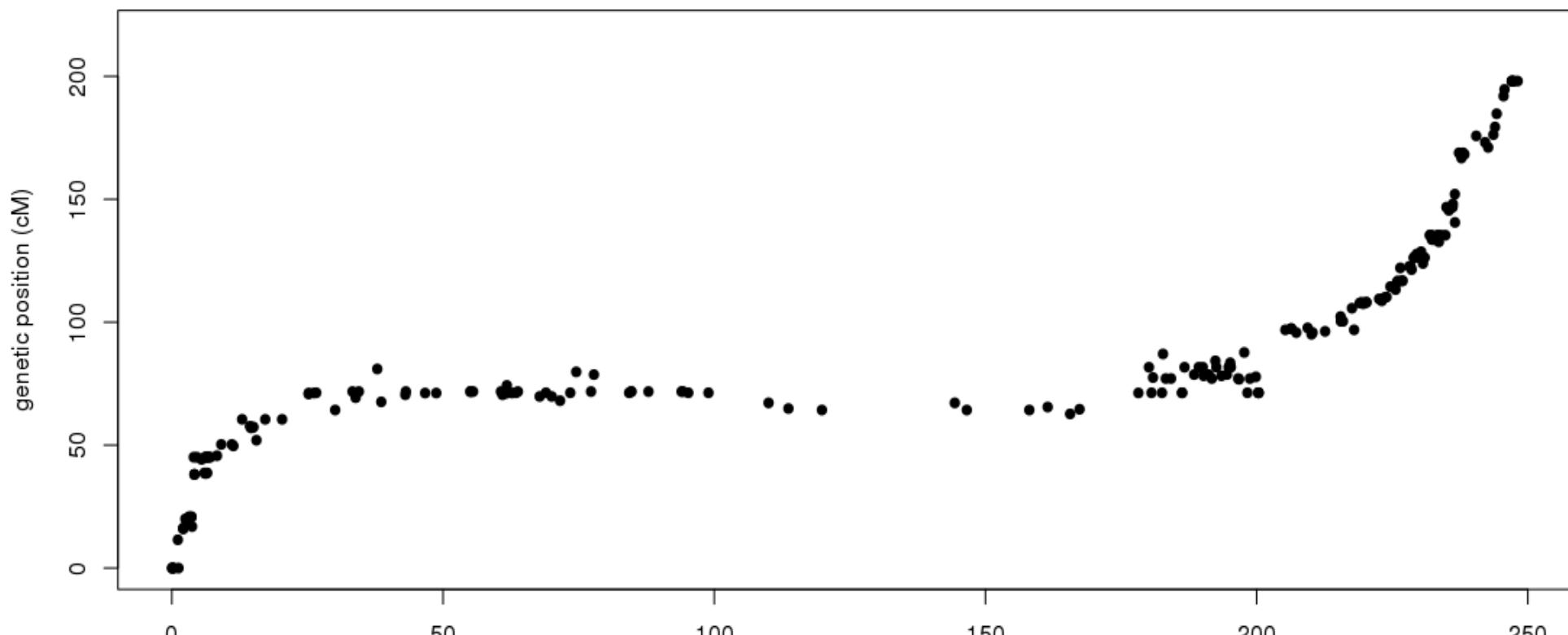


✓ This effect is true for almost **every** chromosome

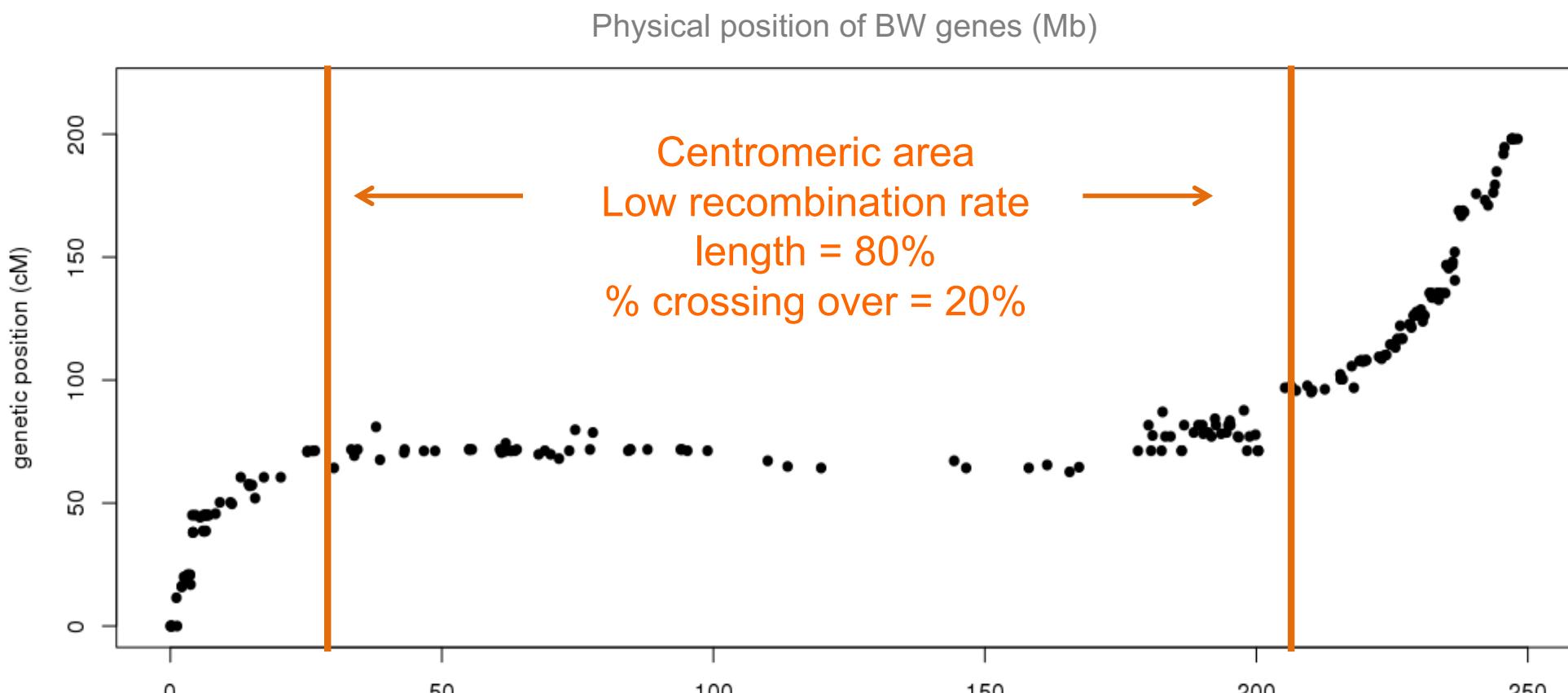
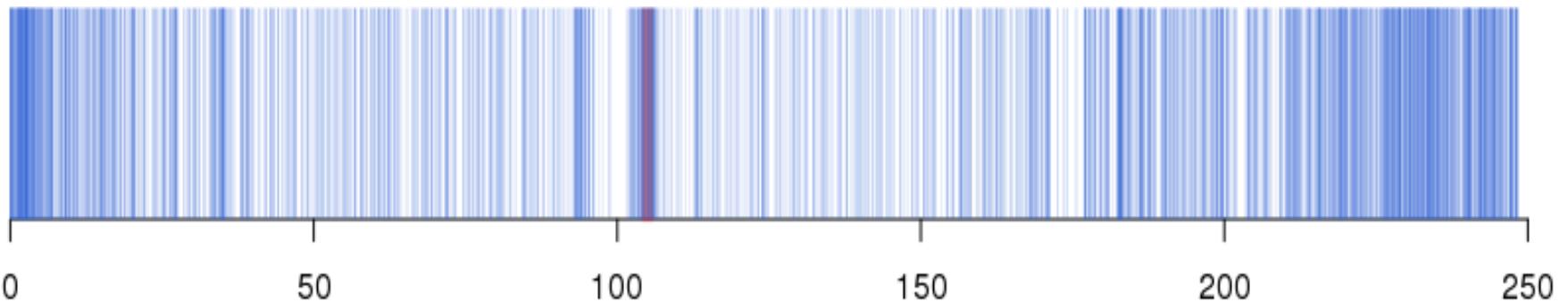
-- Recombination Rate along chromosome



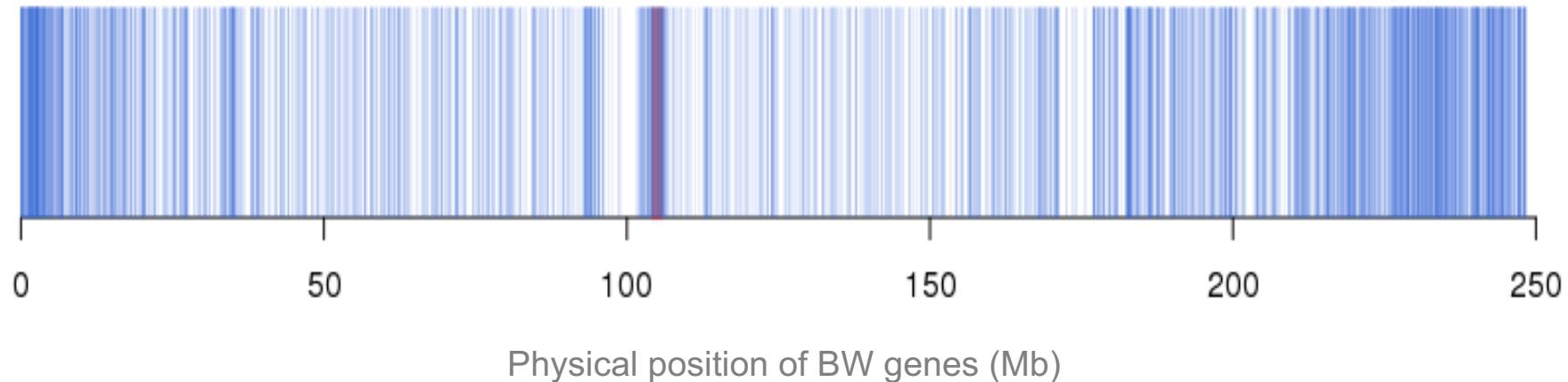
Physical position of BW genes (Mb)



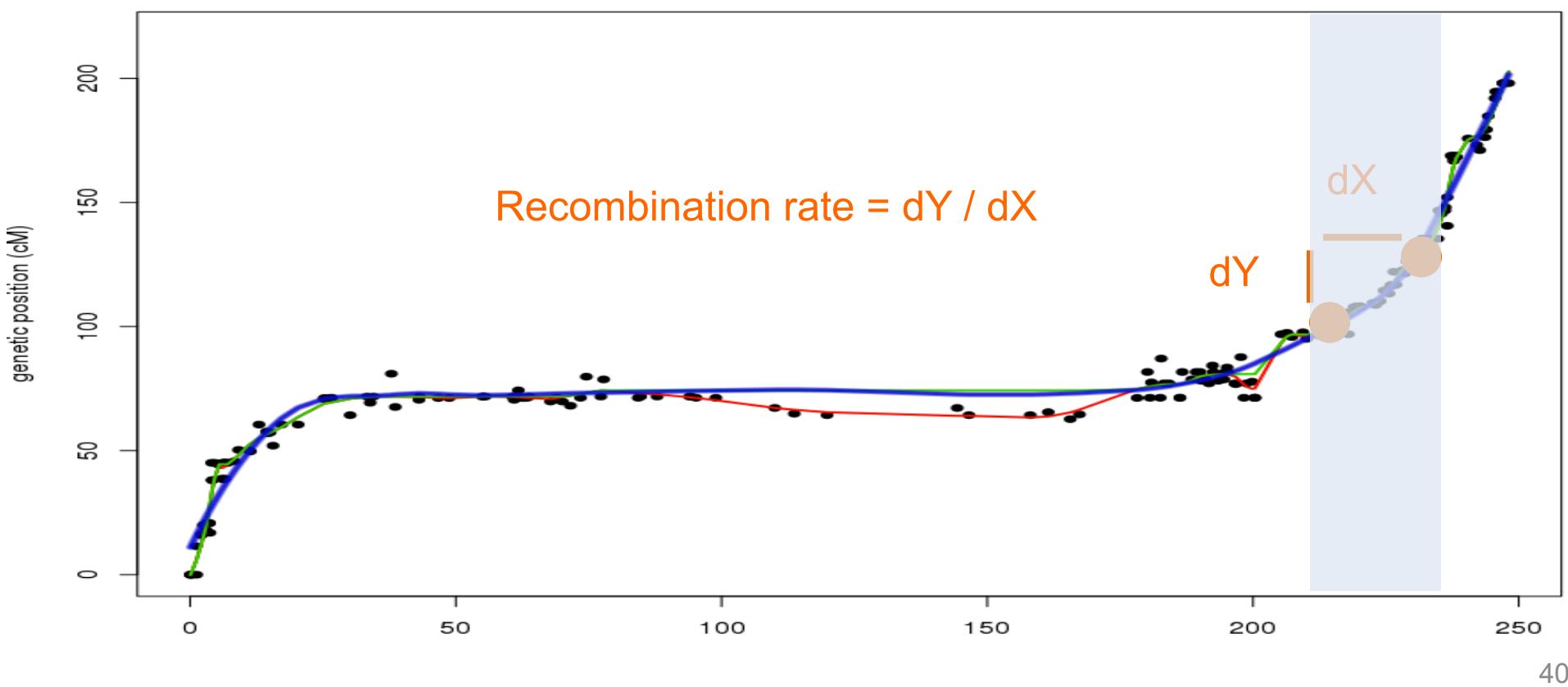
-- Recombination Rate along chromosome



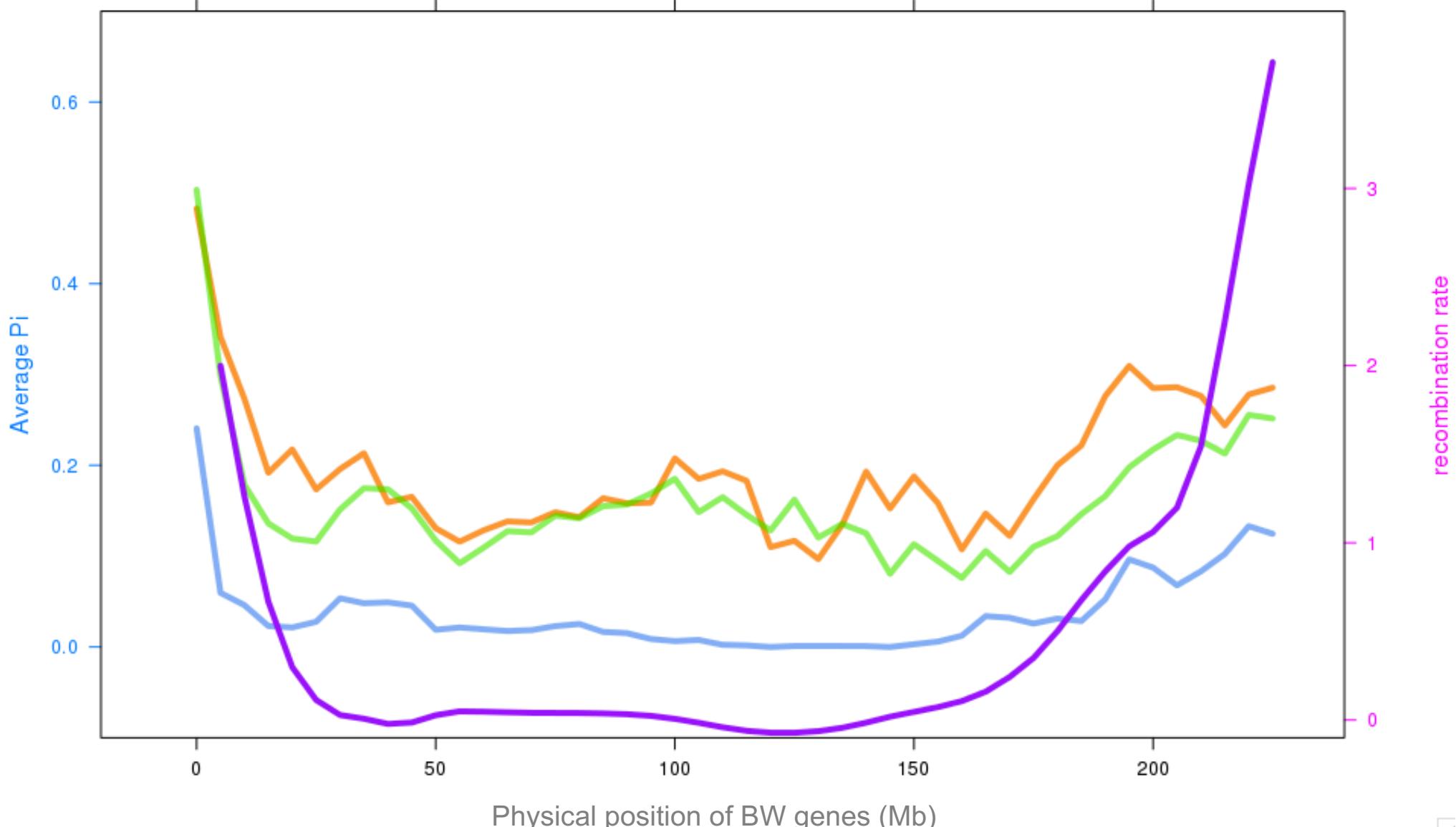
-- Recombination Rate along chromosome



$$\text{Recombination rate} = dY / dX$$

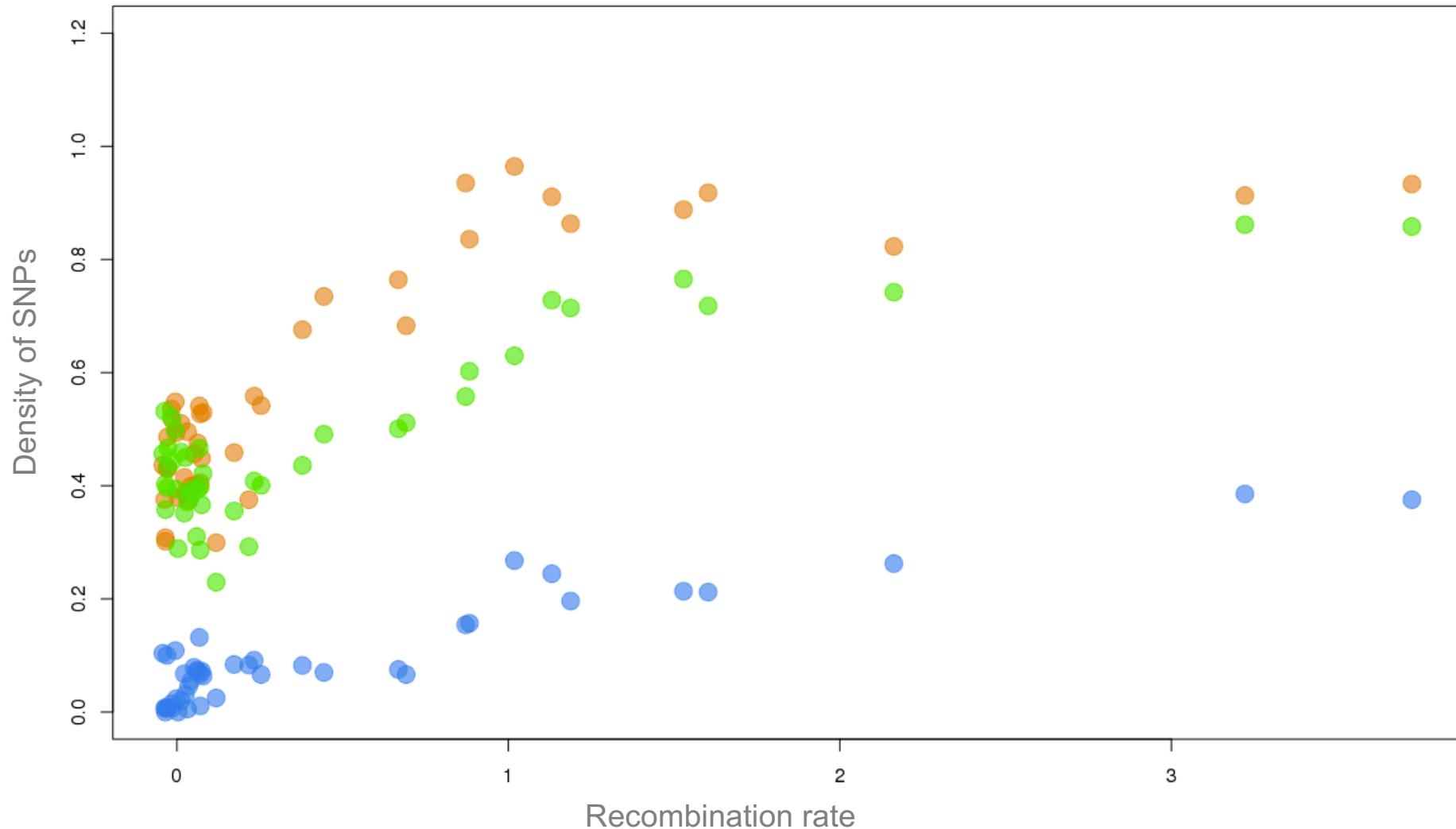


-- Recombination Rate along chromosome



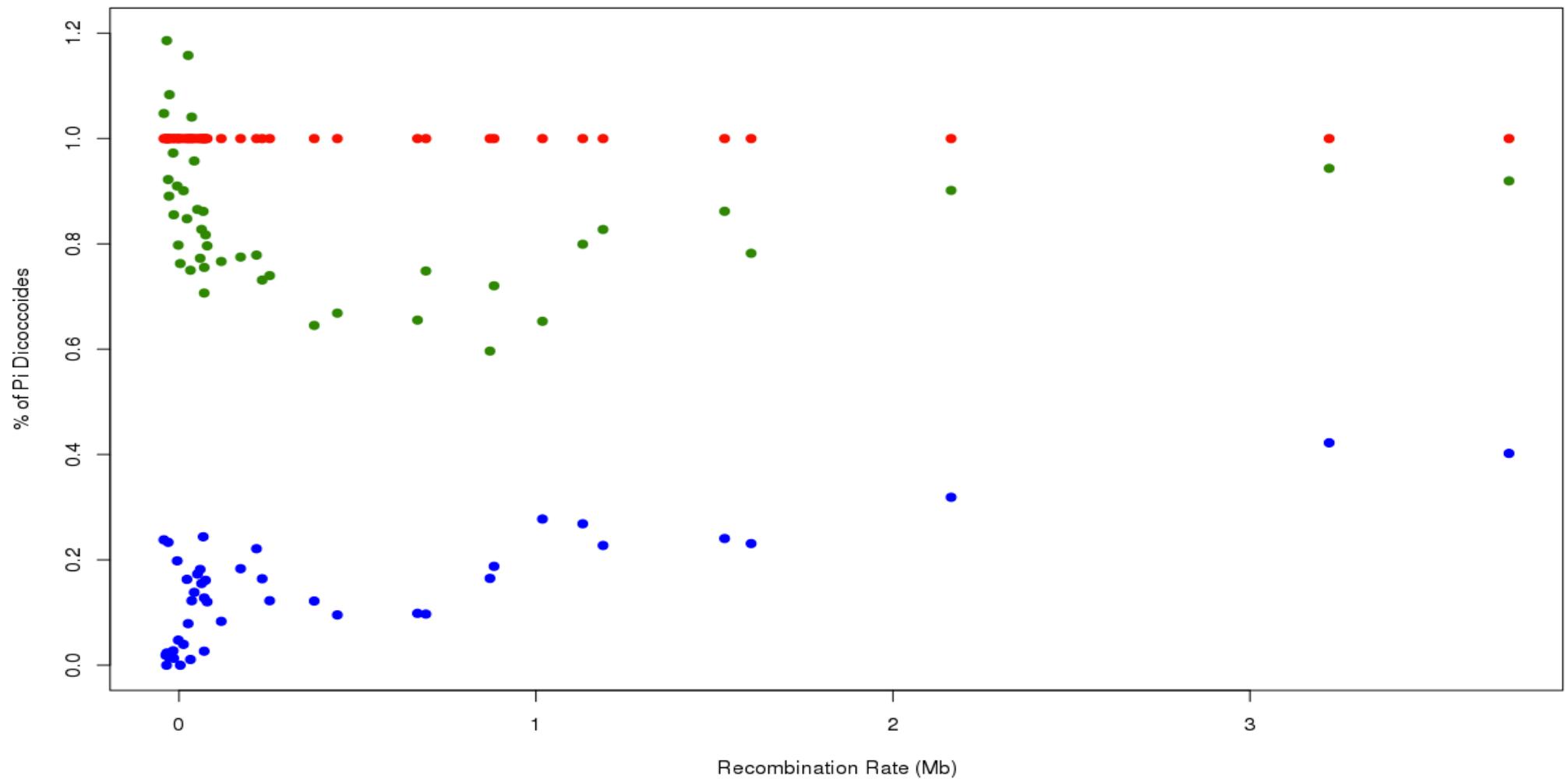
✓ Recombination rate & nucleotidic diversity = **same pattern**

-- Recombination Rate explains diversity



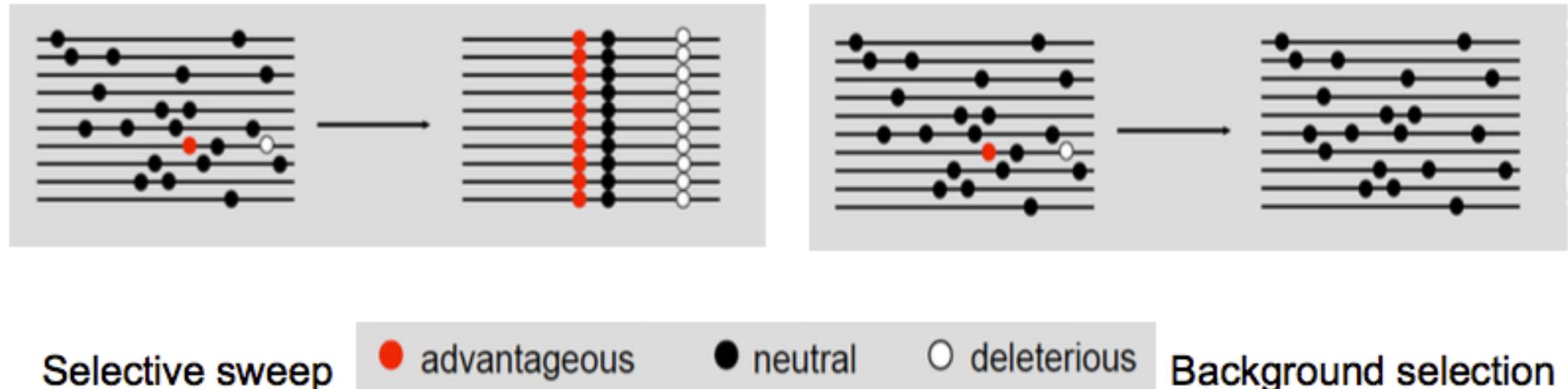
- ✓ No more polymorphism when no recombination for Durum.
- ✓ Relationship plateaus when recombination rate > 2

-- Recombination Rate explains diversity



✓ % of diversity loss in Durum is higher when no recombination.

-- Mecanism linking diversity and recombination

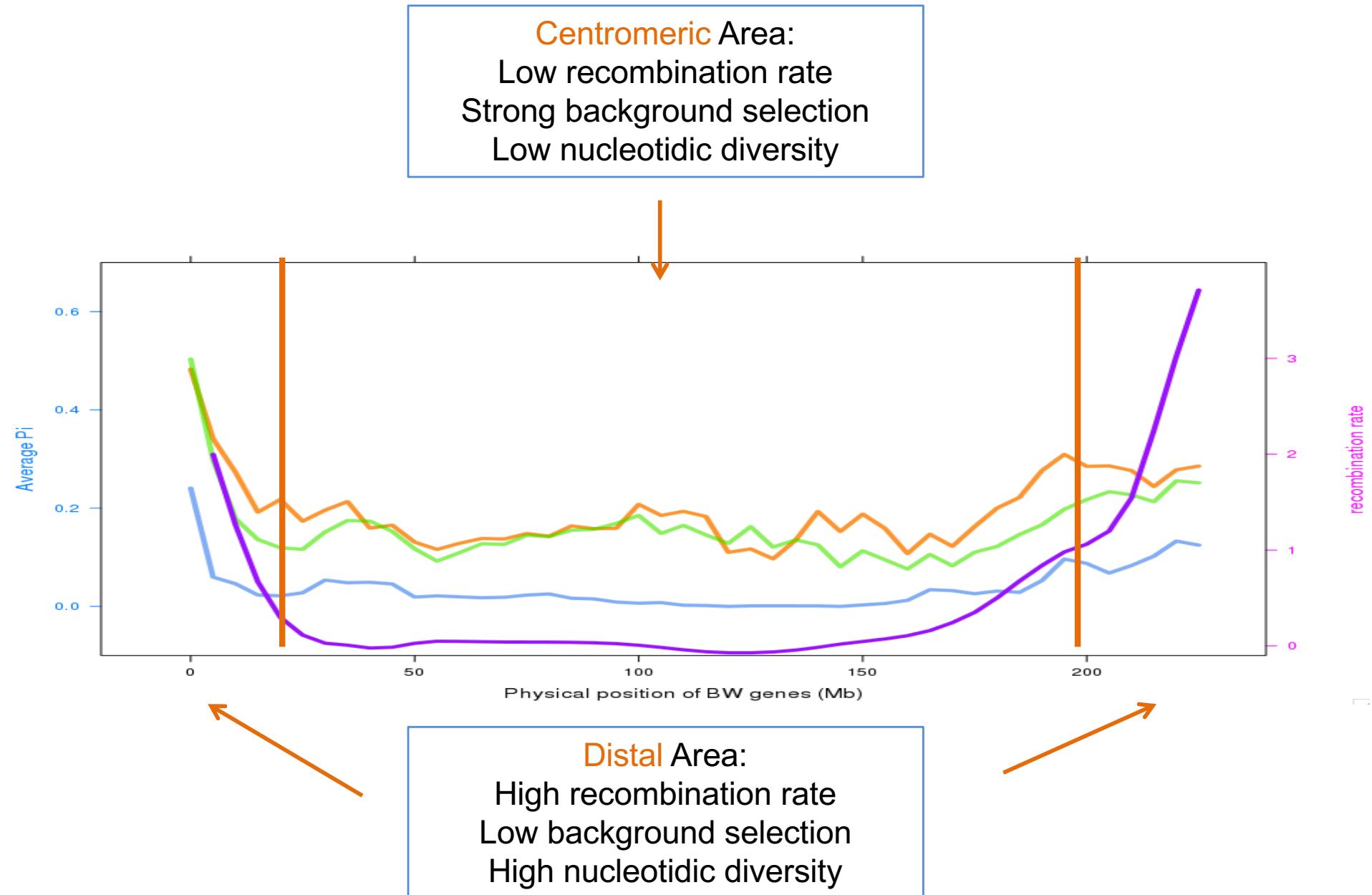


img: Brian Charlesworth

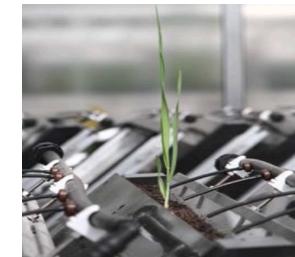
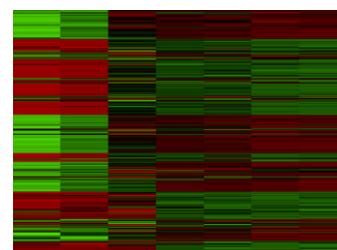
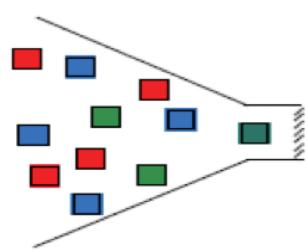
- “Deleterious Mutations Reduce Variation at Linked Sites “
- Low recombination rate : loss of huge fragment of diversity

Charlesworth et al. 1993
Cutter & Payseur 2013⁴⁴
Hudson and Kaplan (1995)

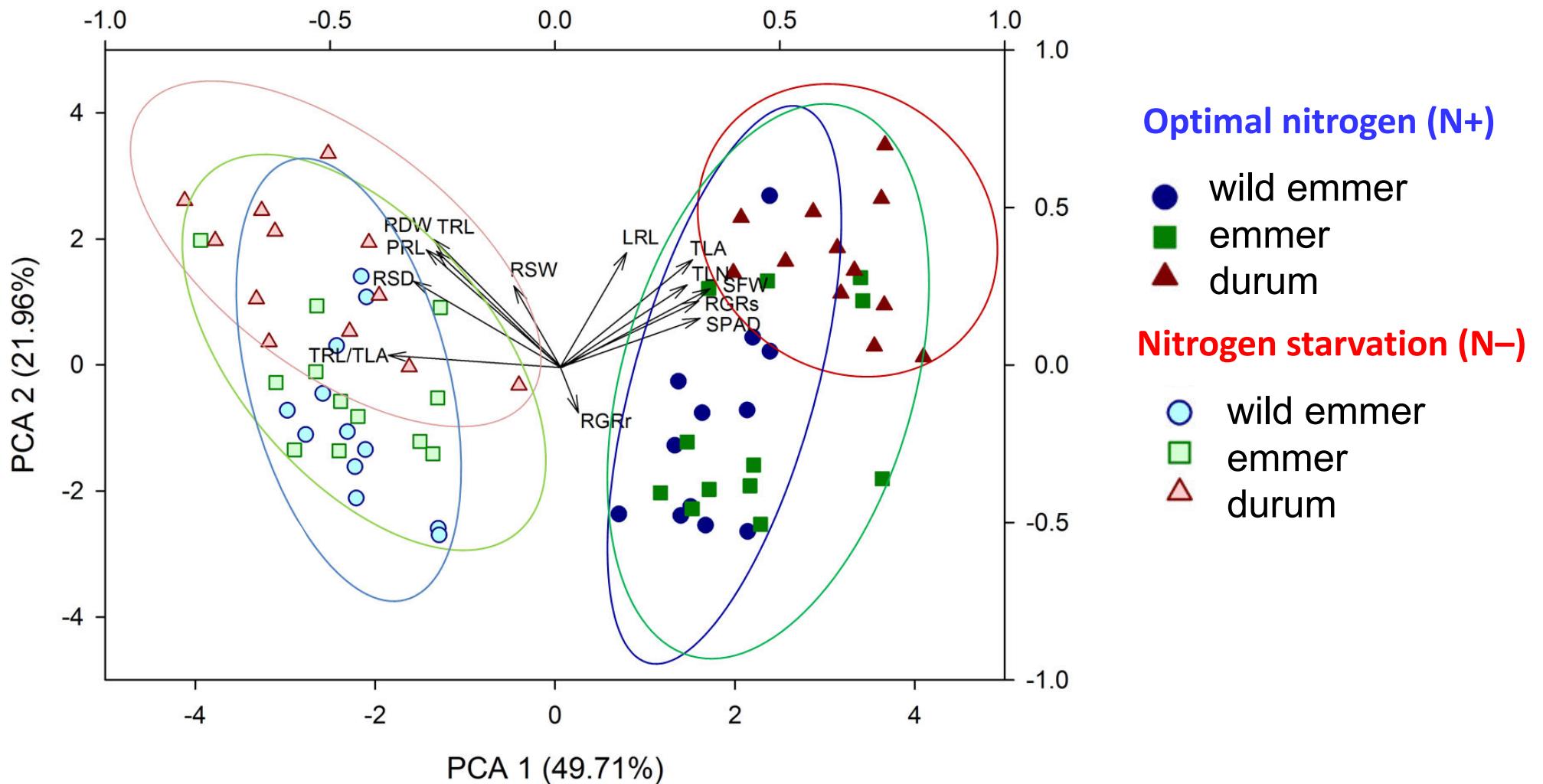
Conclusion



Expression analysis



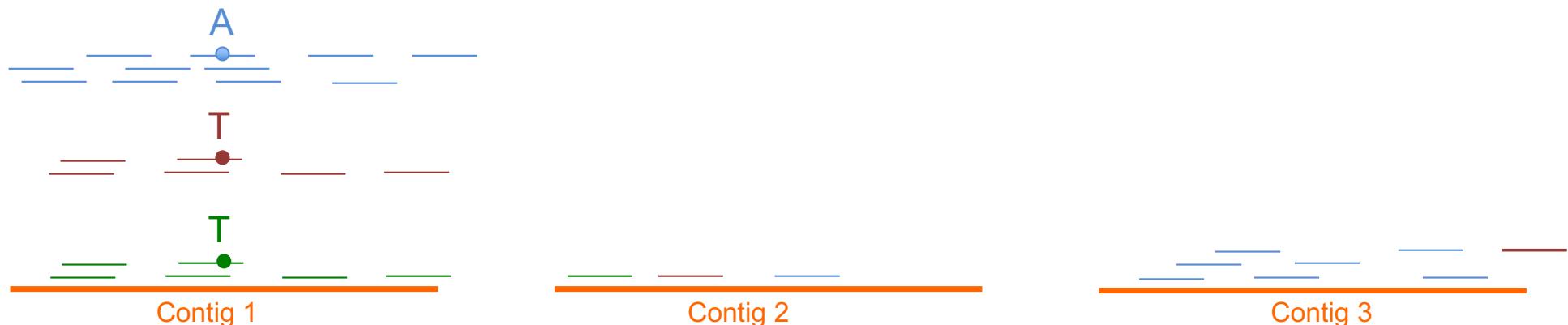
7
N
NITROGEN



Principal component analysis of the morphological and physiological data obtained from the wheat plants grown under different nitrogen level. *Gioia et al. 2015*

- ✓ Nitrogen HAS a strong effect on plant morphology (root, shoot..)
- ✓ Difference in Nitrogen Use Efficiency (NUE) among varieties *Cormier et al. 2016*

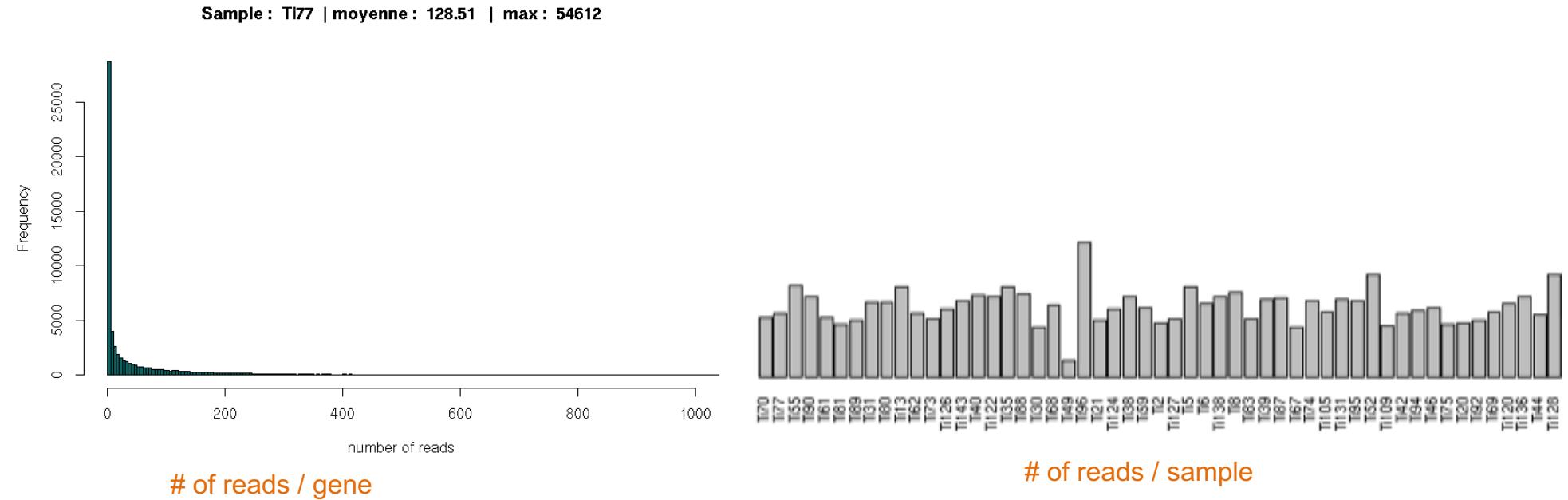
-- Gene expression estimation



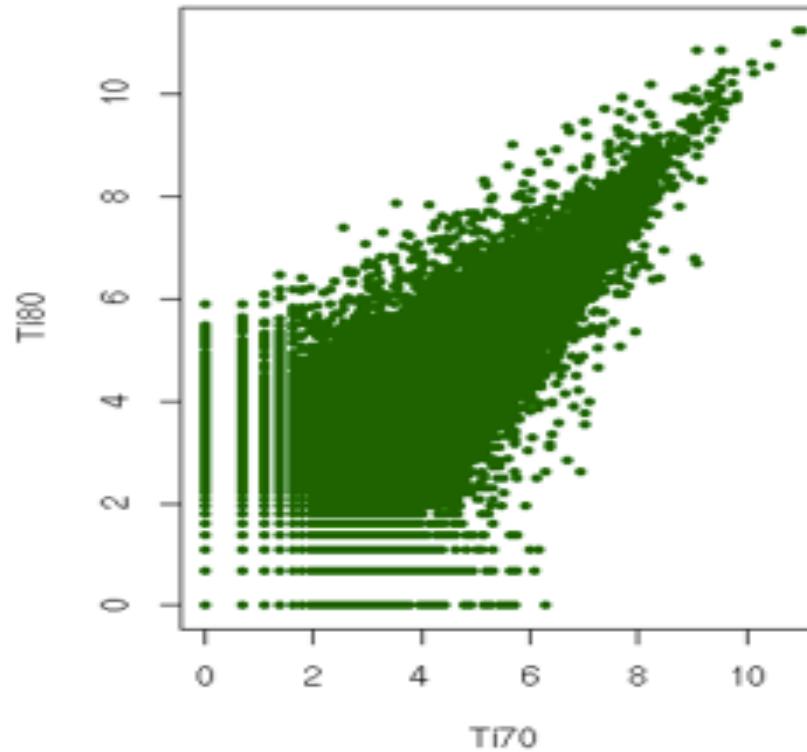
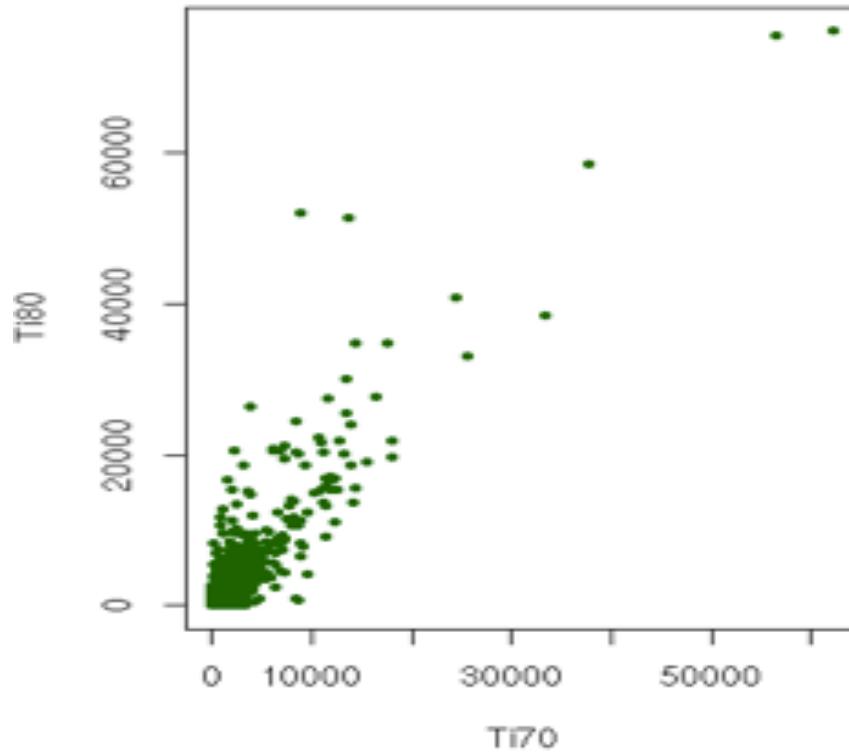
↓
Reads count
(*samtools idxstat*)

	<i>Dicoccoides</i>		<i>Dicoccum</i>		<i>Durum</i>	
	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
Contig 1	200	199	172	204	473	302
Contig 2	2	3	8	2	25	1
Contig 3	0	1	1	2	985	475

Matrix of Read Counts



-- Repetability / heretability ?



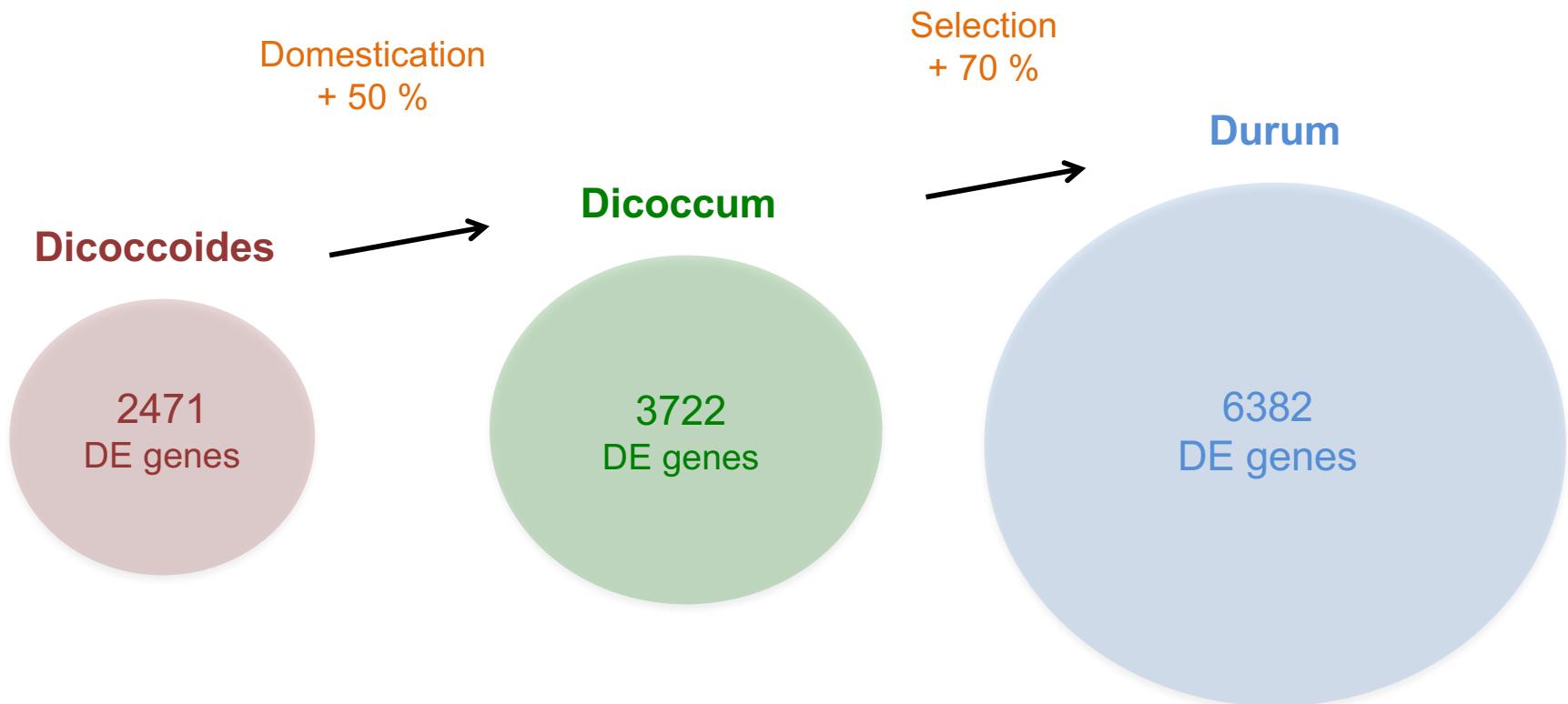
- ✓ Good correlation between expression of 2 samples of a genotypes
- ✓ .. And between 2 samples of different genotypes
- ✓ But a “Poison” noise → influence in low expressed genes
- ✓ Data are not homoskedastic

-- Differentially expressed genes between treatments → How to find them

	<i>Dicoccoides</i>		<i>Dicoccum</i>		<i>Durum</i>	
	Sample 1 (N+)	Sample 2 (N-)	Sample 3 (N+)	Sample 4 (N-)	Sample 5 (N+)	Sample 6 (N-)
Contig 1	200	199	172	204	473	302
Contig 2	2	3	8	2	25	1
Contig 3	0	1	1	2	985	475

- ✓ Use of DESeq2 Love et al. 2014
- ✓ Normalization using the total number of reads / sample
- ✓ Use of shrinkage to avoid low expressed genes
- ✓ Control of false discovery rate
- ✓ Detection by pairs

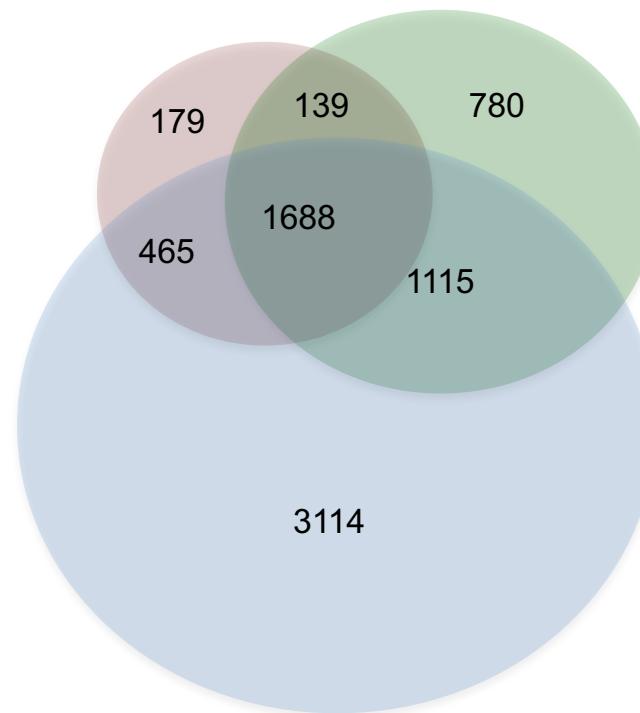
-- DE genes between treatments?



- ✓ 2x more differentially expressed genes for Durum than for old wheat
- ✓ A bit more under-expressed genes in N+ condition (silencing)

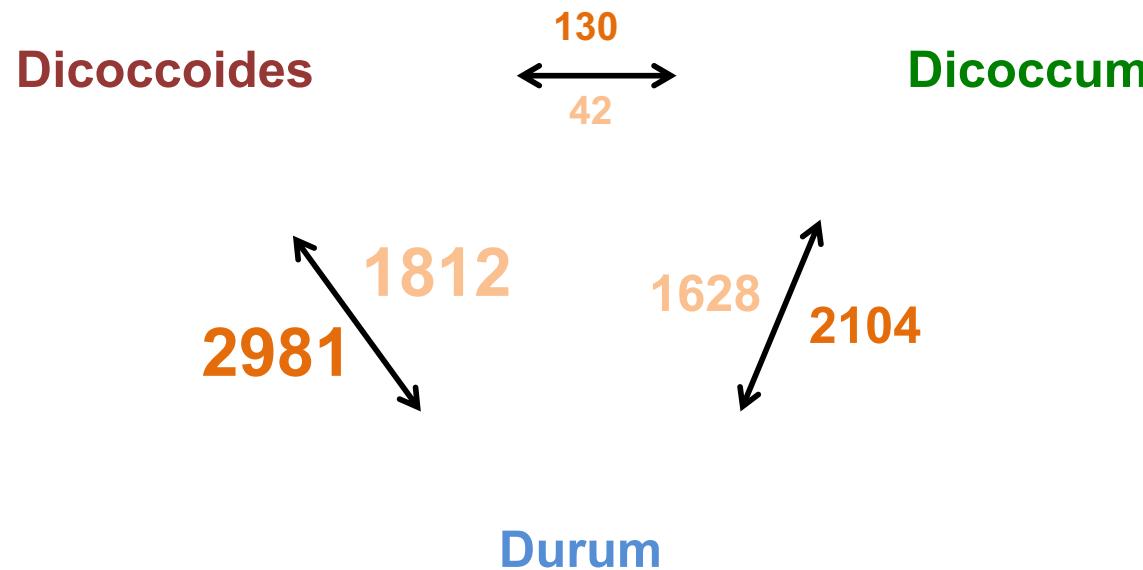
Dicoccoides (2471)

Dicoccum (3722)

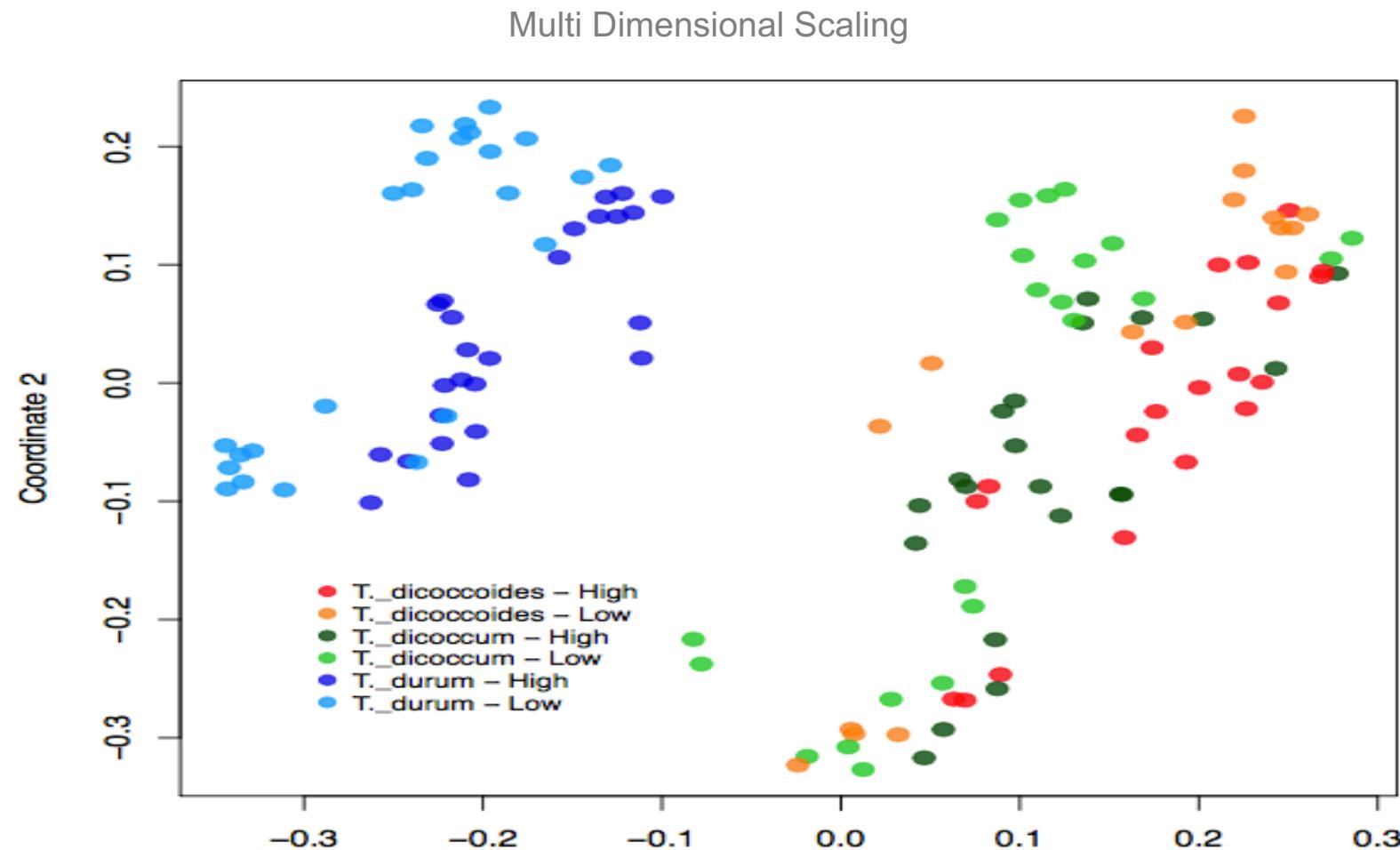


Durum (6382)

- ✓ 2x more differentially expressed genes for Durum than for old wheat
- ✓ A bit more under-expressed genes in N+ condition (silencing)
- ✓ Very few genes DE in old groups only

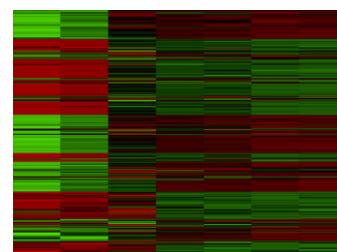
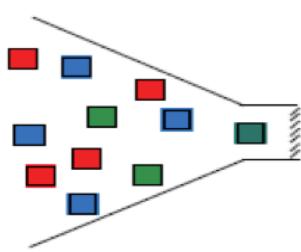


- ✓ Almost no DE genes between Dicoccoides and Dicoccum
- ✓ Lot of DE genes for durum, especially in N+ condition
- ✓ DE genes are more often silenced during evolution



- ✓ Dicoccoides and Dicoccum present the same pattern of expression
- ✓ Durum is clearly distinct
- ✓ Nitrogen has low effect on the global pattern.

Conclusion & Perspectives



7
N
NITROGEN

Genotyping level

- ✓ Pattern of **Synonymous** and non-synonymous **diversity**
- ✓ Use of the Linkage Disequilibrium (LD)

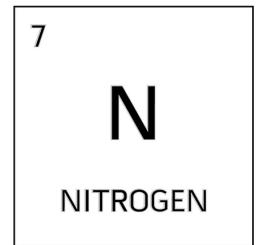
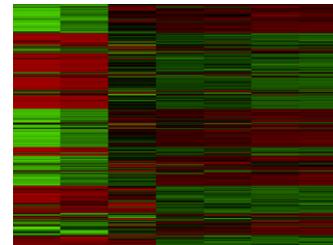
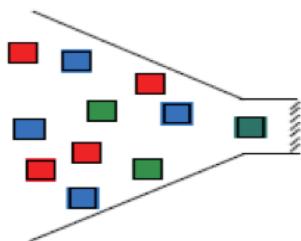
Expression level

- ✓ **Enrichment analysis**
- ✓ **Network analysis**
- ✓ **Position** of differentially expressed genes

Thank you!

Selection for adaptation to high nitrogen was the driving force of tetraploid wheat domestication

RNA-Seq day, Avignon, November 2016



Authors: Y. Holtz, T. Gioia, S. Santoni, V. Ranwez, R. Beleggia, D. Rau, J. David, R. Papa