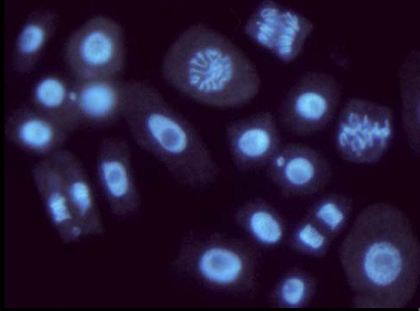
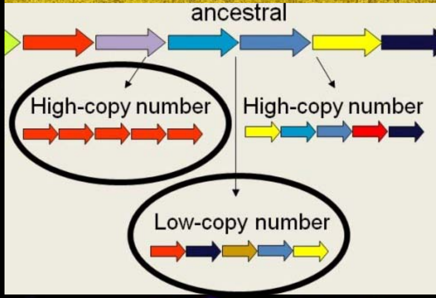
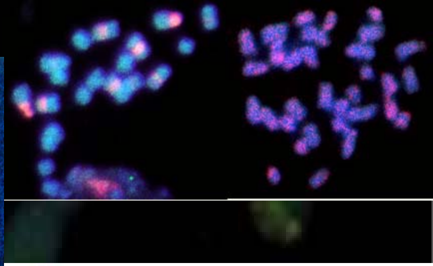
Rapid evolution of repetitive DNA: Chromosome evolution and speciation

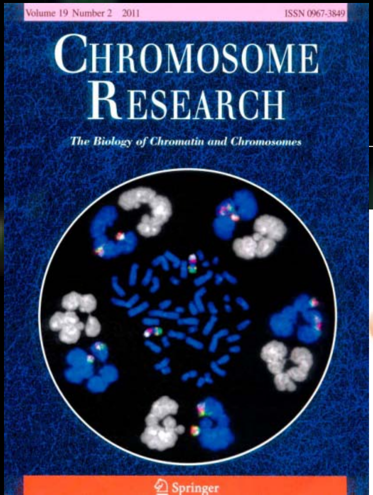
Pat Heslop-Harrison
phh4@le.ac.uk
www.molcyt.com pw/user: ‘visitor’

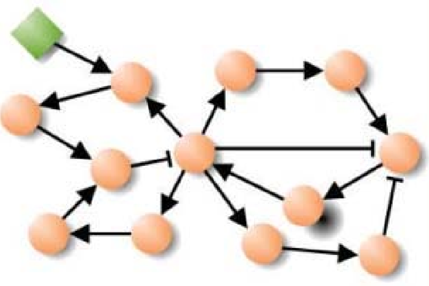
Social media: #ICC18 and Pathh1 on Twitter
Reports: AoBBlog.com and Storify.com/pathh1

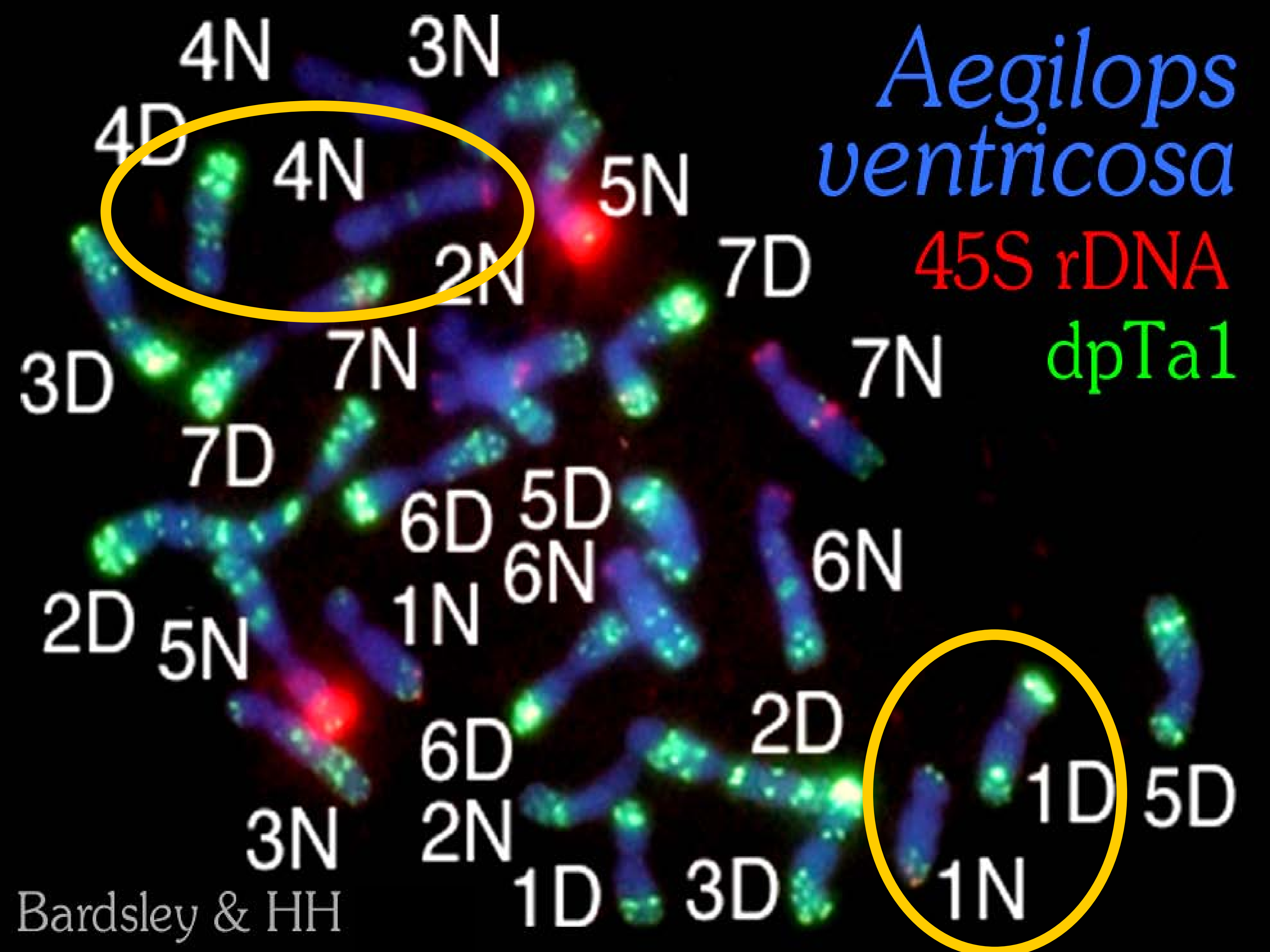







Volume 19 Number 2 2011
ISSN 0967-3849
CHROMOSOME RESEARCH
The Biology of Chromatin and Chromosomes
Springer





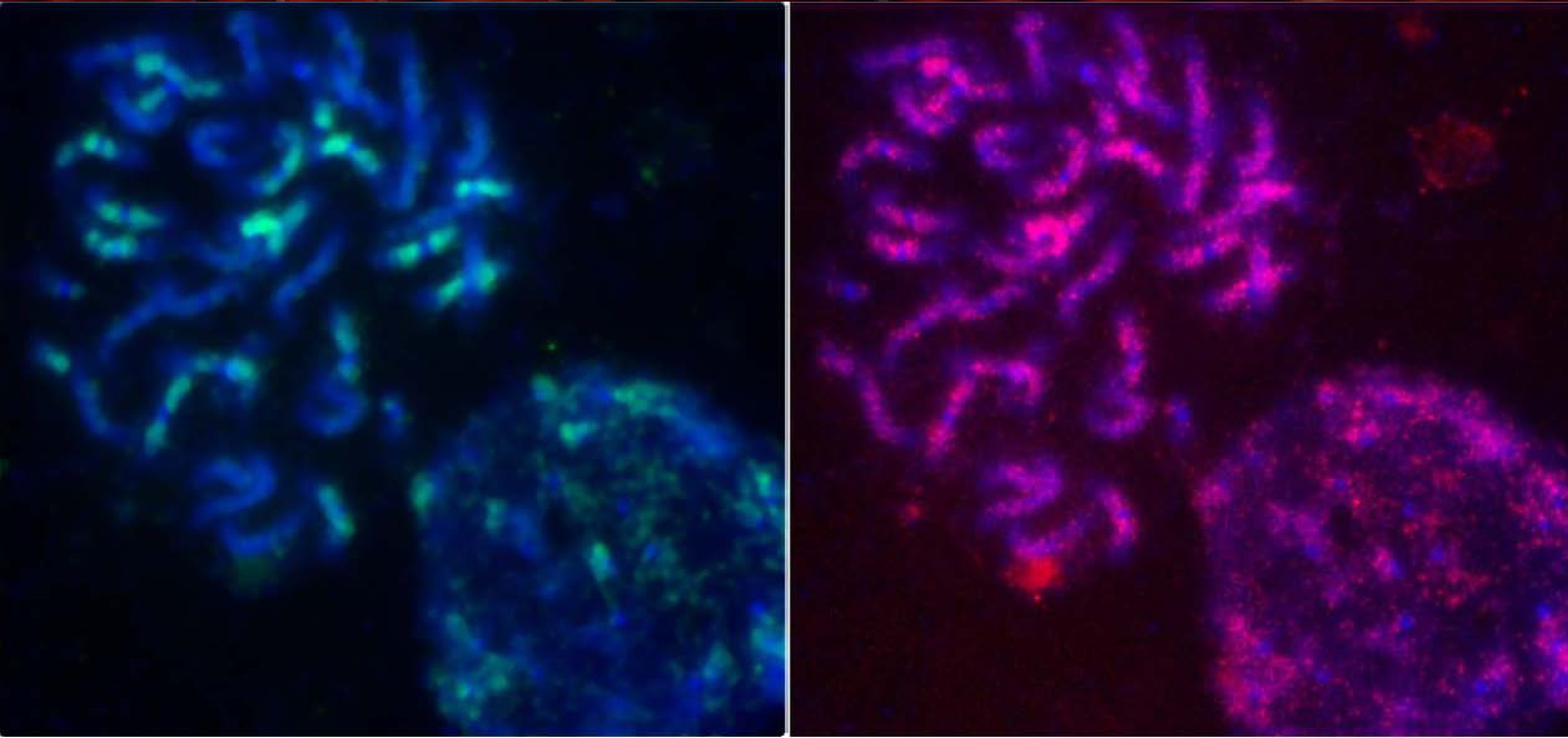
Aegilops ventricosa

45S rDNA

dpTa1

Bardsley & HH

Genome evolution



Arachis hypogaea – 2n=4x=40

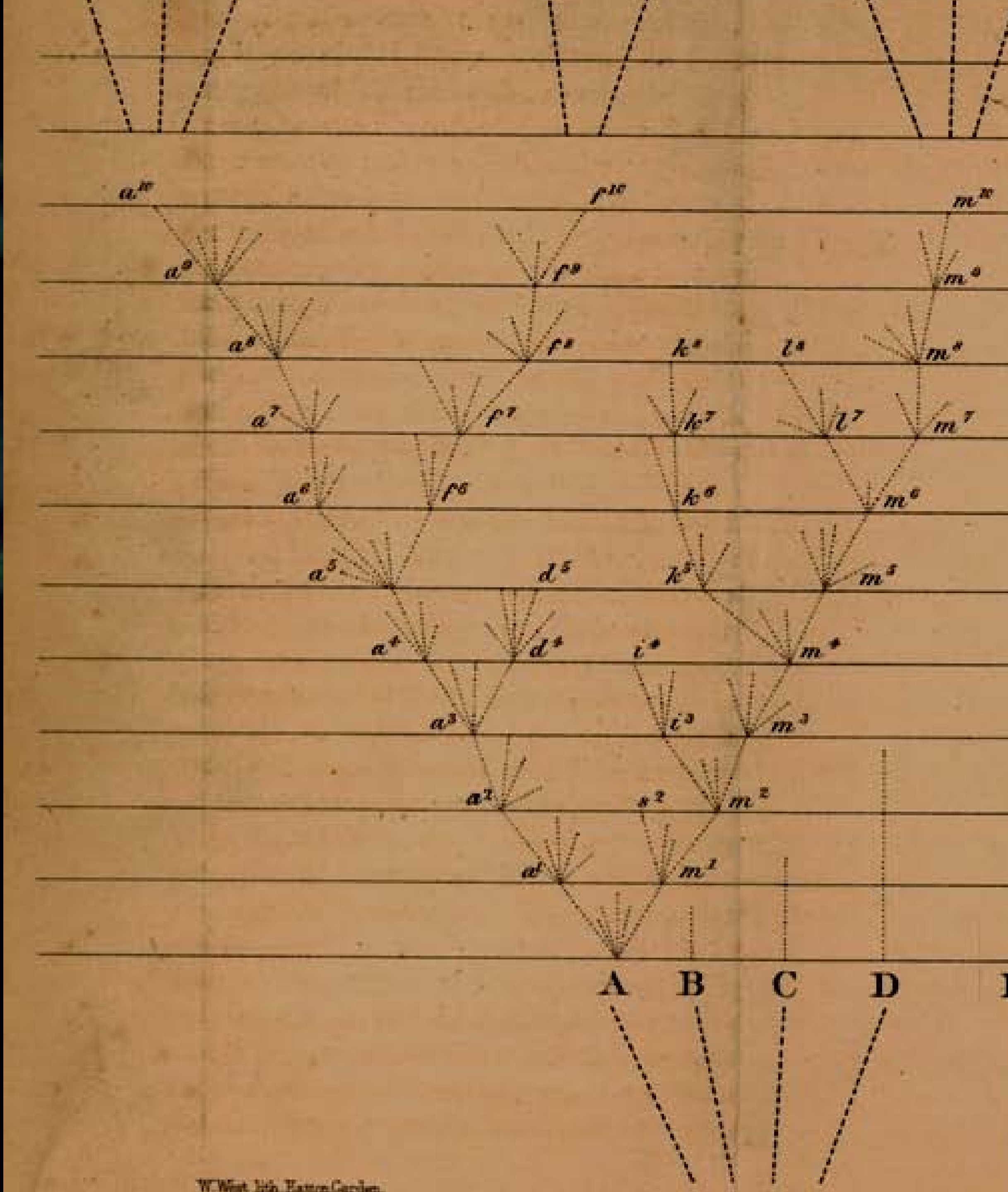
In situ hybridization of two BACs including repeats

Contrasting distribution of their major repeat familie

udia Guerra Araujo et al in prep 2011, EMBRAPA, Brazil

Arachis in situ 6 Slide 2 Meta 7
ADH167F07 red ADH129F24 green

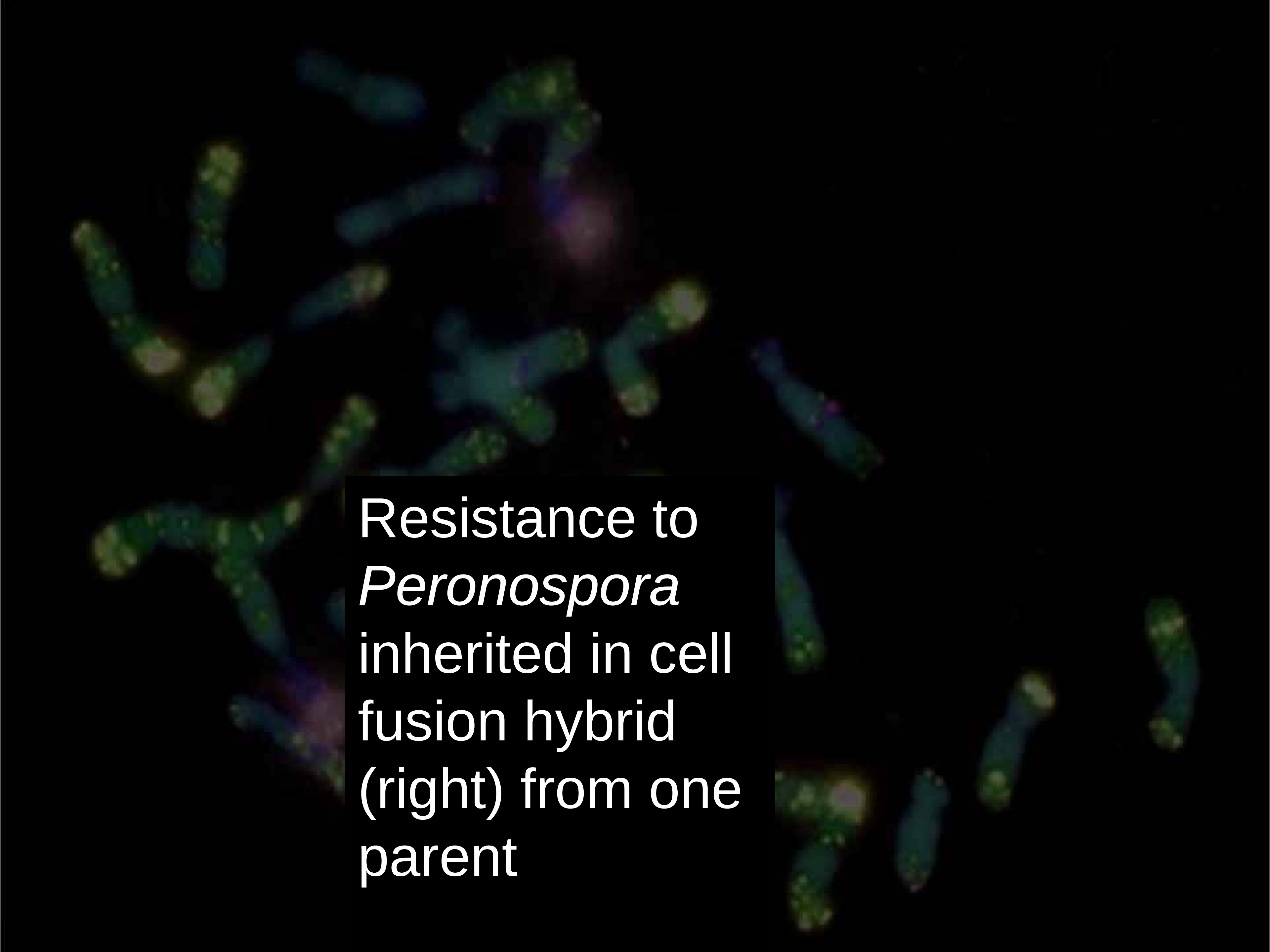
Darwin 1859
The only figure in
“The origin of species”



Chromosome and genome engineering

Cell fusion hybrid of two tetraploid tobacco species

Patel, Badakshi, HH, Davey et al 2011 Annals of Botany in press

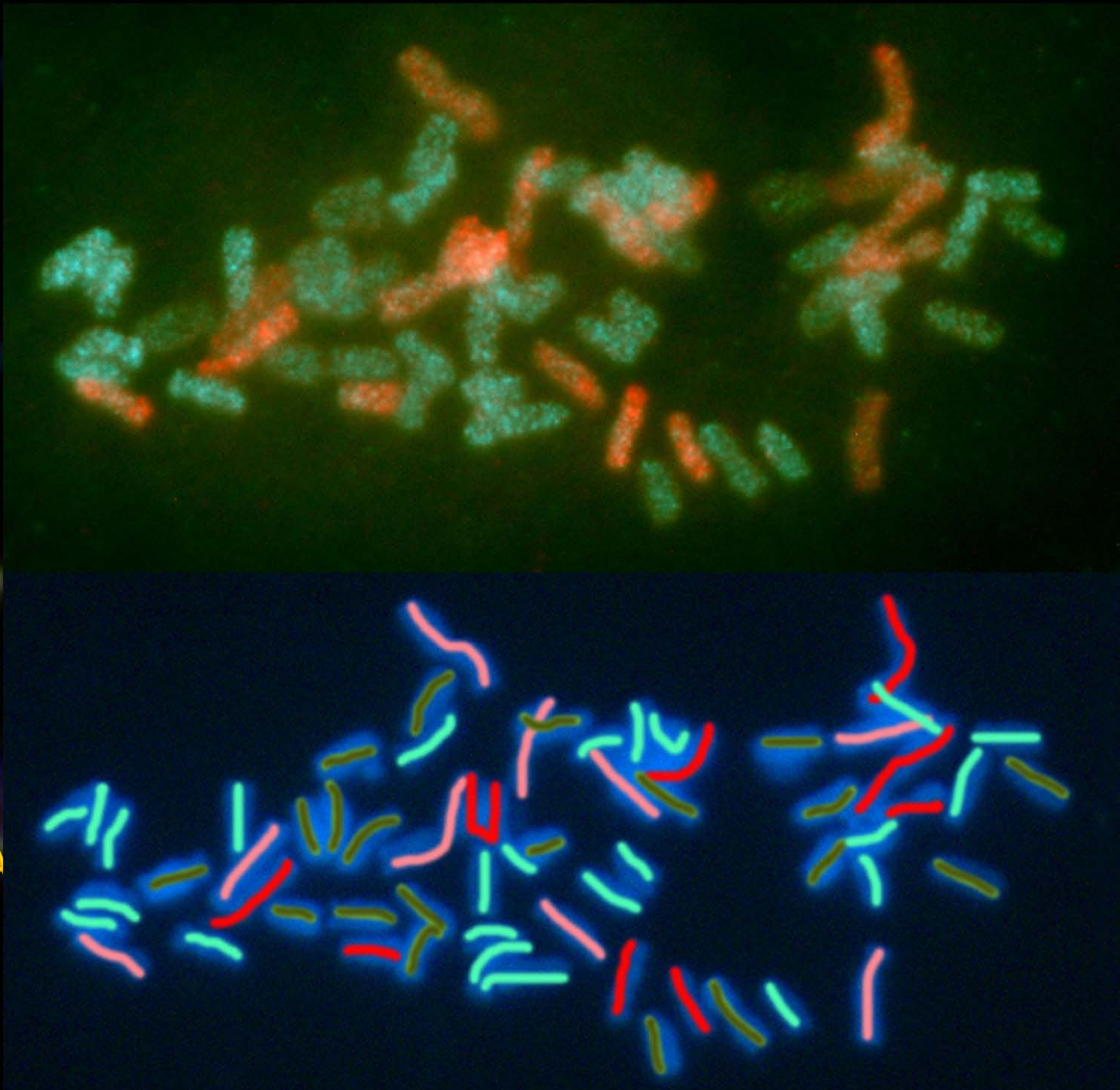


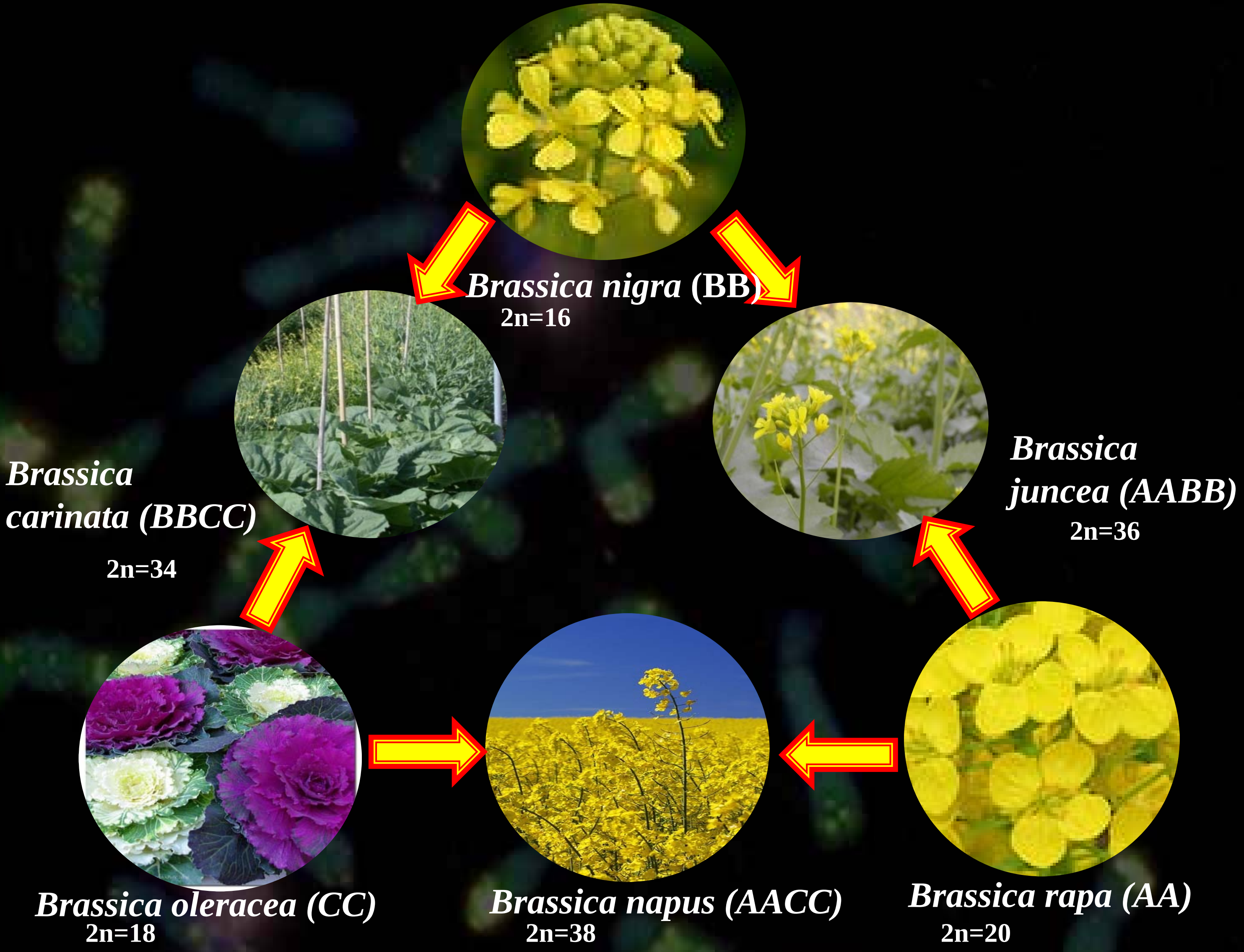
Resistance to *Peronospora* inherited in cell fusion hybrid (right) from one parent

Nicotiana
hybrid
4x + 4x
cell fusions

Each of 4
chromosome
sets has
distinctive
repetitive
DNA when
probed with
genomic DNA

Patel et al
Ann Bot 2011





Brassica nigra (BB)
2n=16

Brassica carinata (BBCC)

2n=34

Brassica juncea (AABB)
2n=36

Brassica oleracea (CC)
2n=18

Brassica napus (AACC)
2n=38

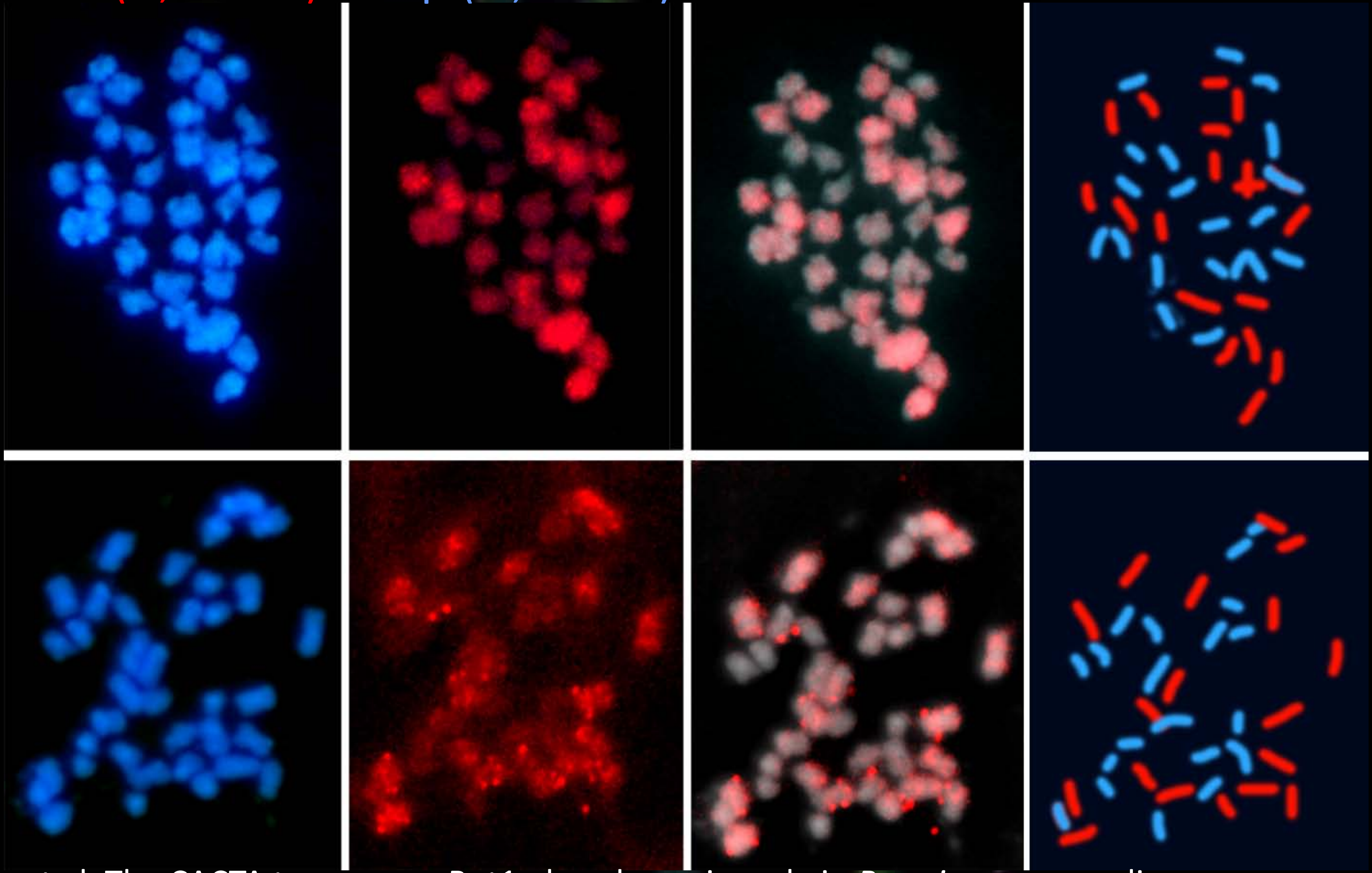
Brassica rapa (AA)
2n=20

The *Brassica* genus is monophyletic, from single common ancestor.
What has changed in the DNA sequences?

Genome Specificity of a CACTA Transposon

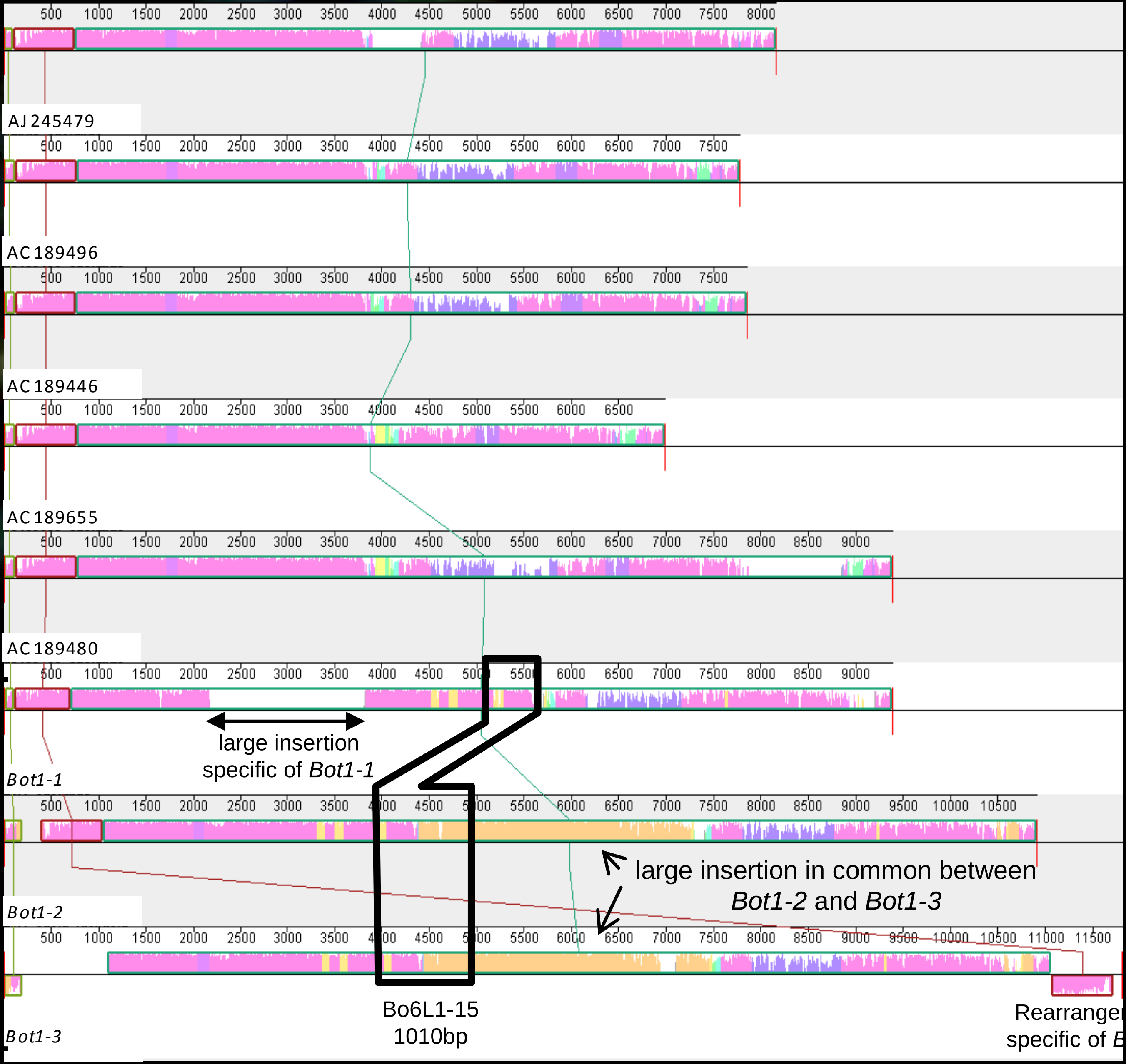
B. napus **(AACC, 2n=4x=38) – hybridized with C-genome CACTA element red**

B. oleracea (CC, 2n=2x=18) **B. rapa (AA, 2n=2x=20)**



Alix et al. The CACTA transposon Bot1 played a major role in *Brassica* genome divergence and gene proliferation. Plant Journal

Genome Specificity of a CACTA (En/Spm) Transposon

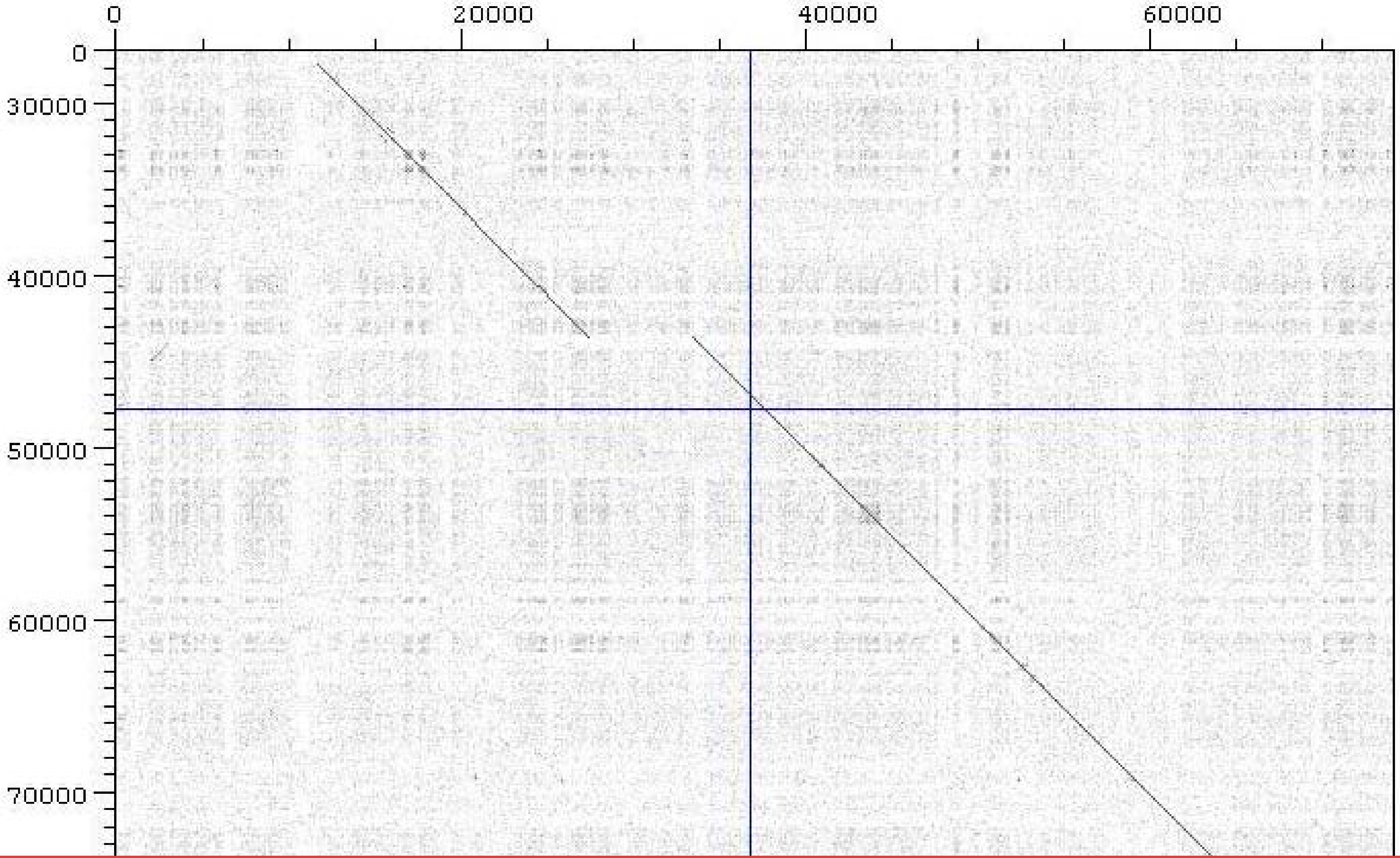


large insertion
specific of *Bot1-1*

large insertion in common between
Bot1-2 and *Bot1-3*

Bo6L1-15
1010bp

Rearrange
specific of *B*

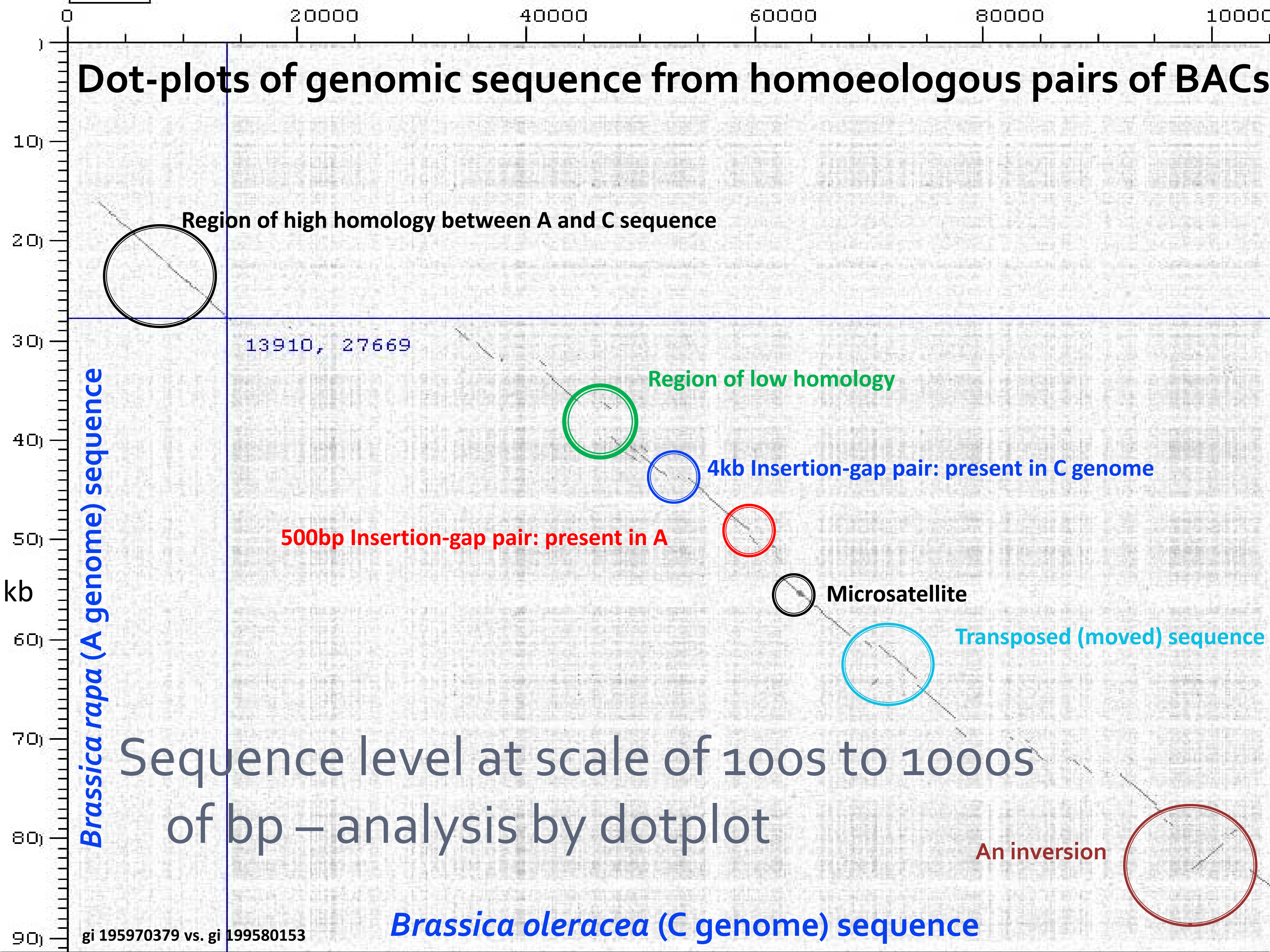


Dotplot comparisons at scale of 10,000s bp

Two *Musa* chromosomes are >95% homologous with gaps

Faisal Nouroz 2012

Dot-plots of genomic sequence from homoeologous pairs of BACs



Region of high homology between A and C sequence

Region of low homology

4kb Insertion-gap pair: present in C genome

500bp Insertion-gap pair: present in A

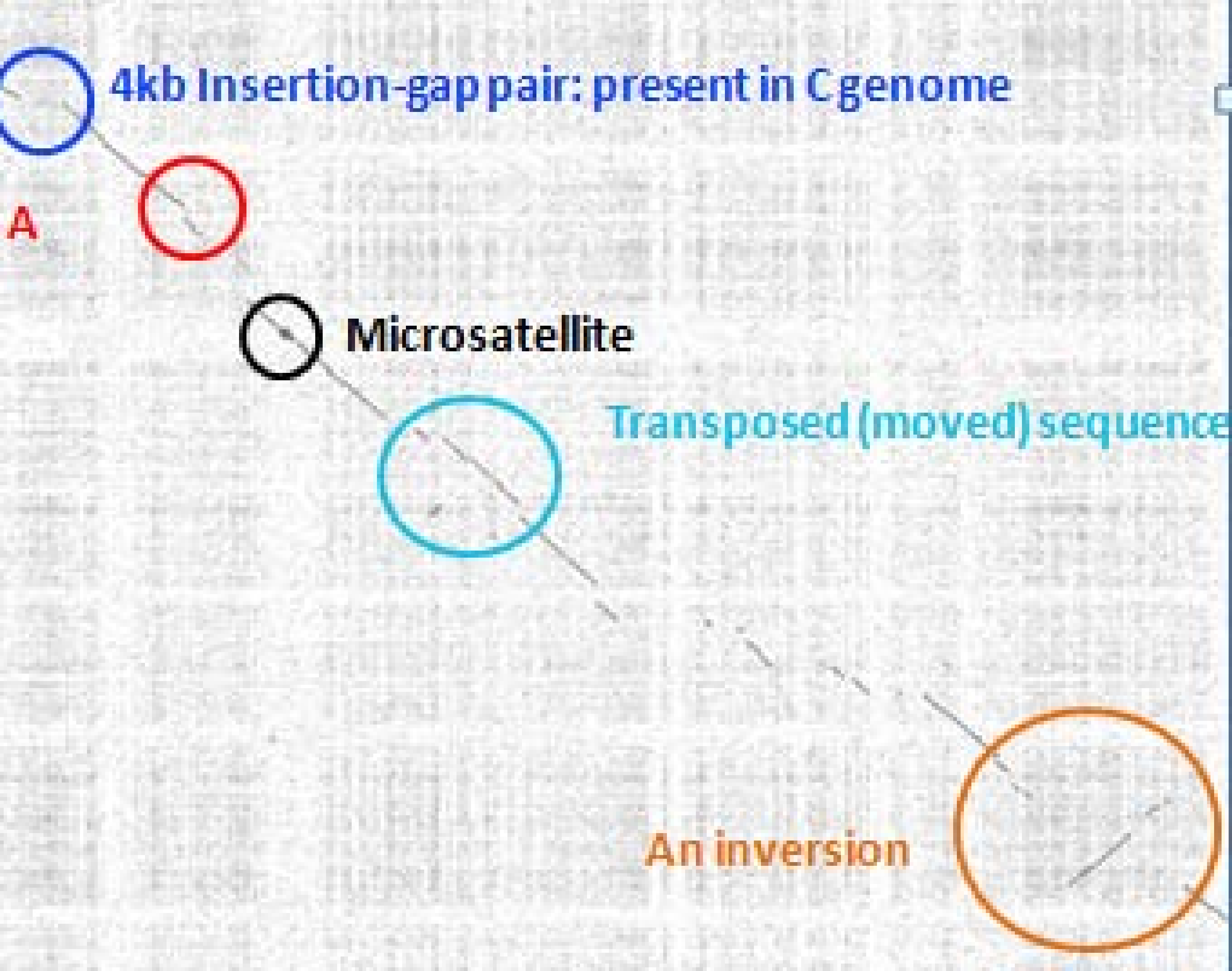
Microsatellite

Transposed (moved) sequence

An inversion

Sequence level at scale of 100s to 1000s of bp – analysis by dotplot

gi 195970379 vs. gi 199580153



4kb Insertion-gap pair: present in C genome

A

Microsatellite

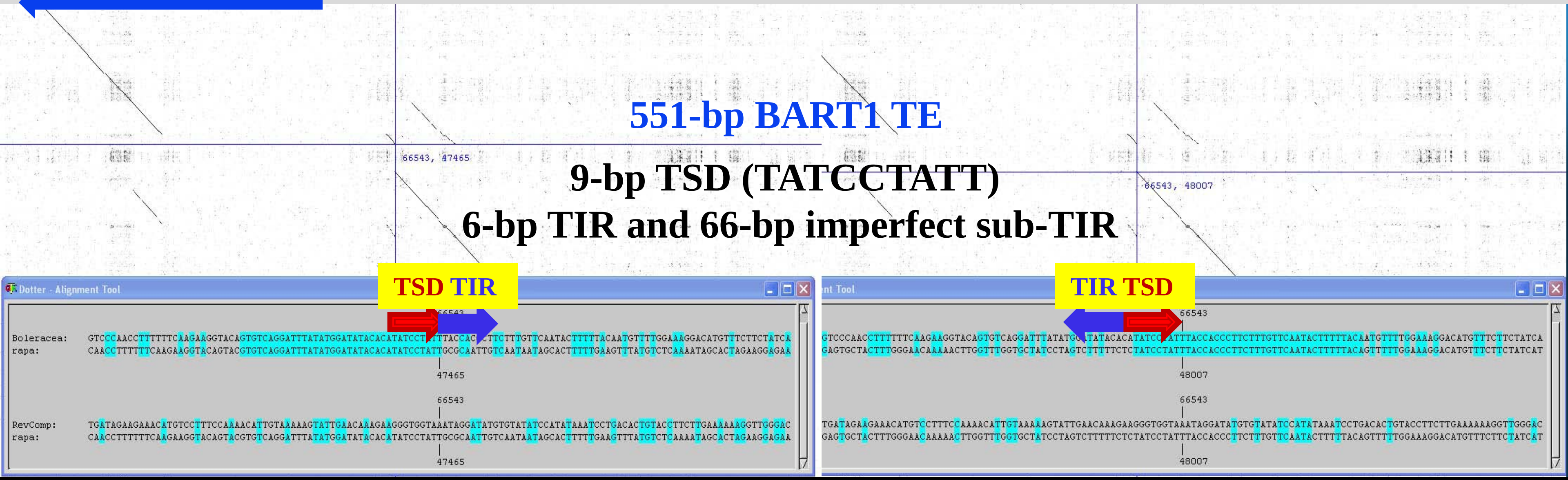
Transposed (moved) sequence

An inversion

AAGTGAATGGATGCTCGCATTAGTTACTATGAGCCGATTCTCGCTCTTGCGAAAGCTAAAGAGGAAAAGGCCTTCGCATTGCAGAAG
AGCTGGCTGCCAGCGAGCAAGAGGTTTTCAATATTGGCTTGTGGAAAATTTGTTGCCACTTTTGCTTTACTAAGGAATGAAATAATAC
TTGTTTTTTTTTTTCATGGTTAATATTAGAAGATATAATTTCCTTTGAAGTTAGATTACGTTTCTTTATGTCGACGAAGTGAAGAAATATT
GTCTTGTTTATGGTTCCTTCTAGTCCCAACCTTTTTTCAAGAAGGTACAGTACGTGTCAGGATTTATATGGATATACACA

TATCCTATTGCGCAATTGTCAATAATAGCACTTTTTTGAAGTTTATGTCTCAAAATAGCACTAGAAGGAGAAAGTCACAAAAATGATATT
CATTAAAGGGTAAAATATCTCTTATATCCTTGGTTTAAAATTAAATAAACAAACAAAAAATAAATAAAAAATAAATAAAAAAAAATGAAAAAAA
AAGAAATTTTTTTTATAGTTTCAGATTATATGTTTTCAGATTCGATTTTTTTTTTATTTTTTTATTTTTTTCGAAATTTTTTTTTTATTTTTTTTCA
AATTTTCTTTTTATAATTTAAAAATACTTTTTGAAACTGTTTTTTTAATTTTTATTTTTTTATTTTAGTATTTATTTTTTTATAAAATTTTAAACCCT
AATTCCTAAACCCCCACCCCTTAACTCTAAACCCTAAGGTTTGGATTAATTAACCCAATGGATATAAGTGTATATTTACCTCTTTAATGA
AACCTATTTTTGTGACTTTGAATCTTGAGTGCTACTTTGGGAACAAAAACTTGGTTTGGTGCTATCCTAGTCTTTTTCTC**TATCCTATT**

TACCACCCTTCTTTGTTCAATACTTTTTACAGTTTTTGGAAAGGACATGTTTCTTCTATCATCACTTAATGGTTATATATGTATGAGAAG
TTTGAAAGAGATTACACTGTTTTGGAATATTAAAAAAAAAAGATATTACAAGATCTGATTTTGTTTGTATTTTAAAATTCTACCAAATC
TCTCCTCAAAATCTTGGTCAAAGTCCAAAAATCCAAATATCTCAGTTAAATTCCACCAAATATGAAATCCTAAAACTTTTCCAAAATA
GTTCAATAAGCCCTTAGTGTTTGGTG



551-bp BART1 TE

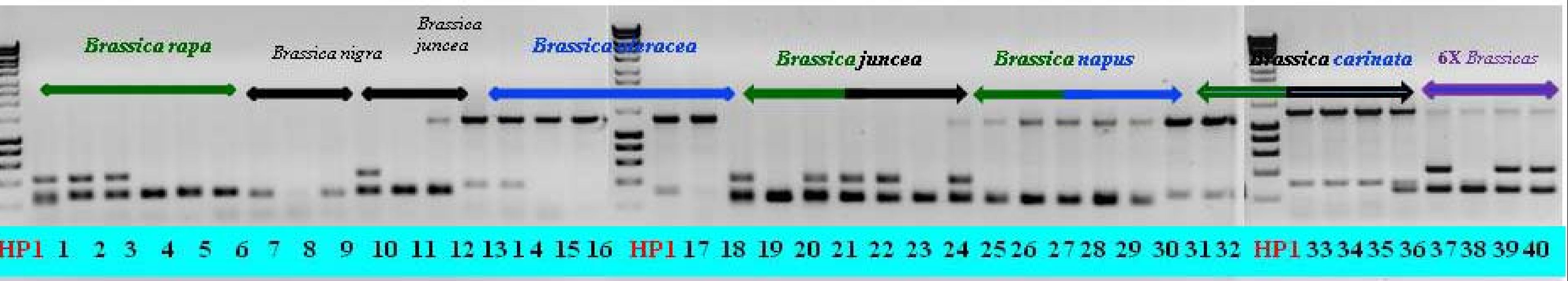
9-bp TSD (TATCCTATT)

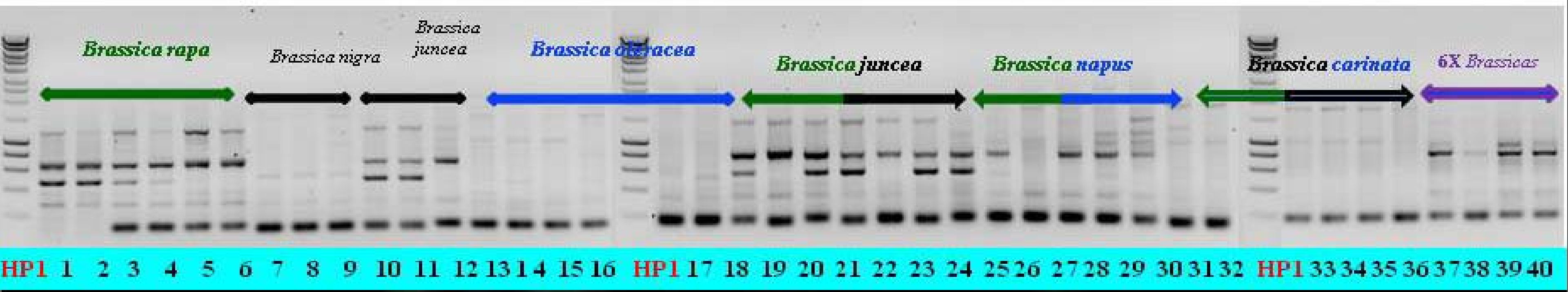
6-bp TIR and 66-bp imperfect sub-TIR

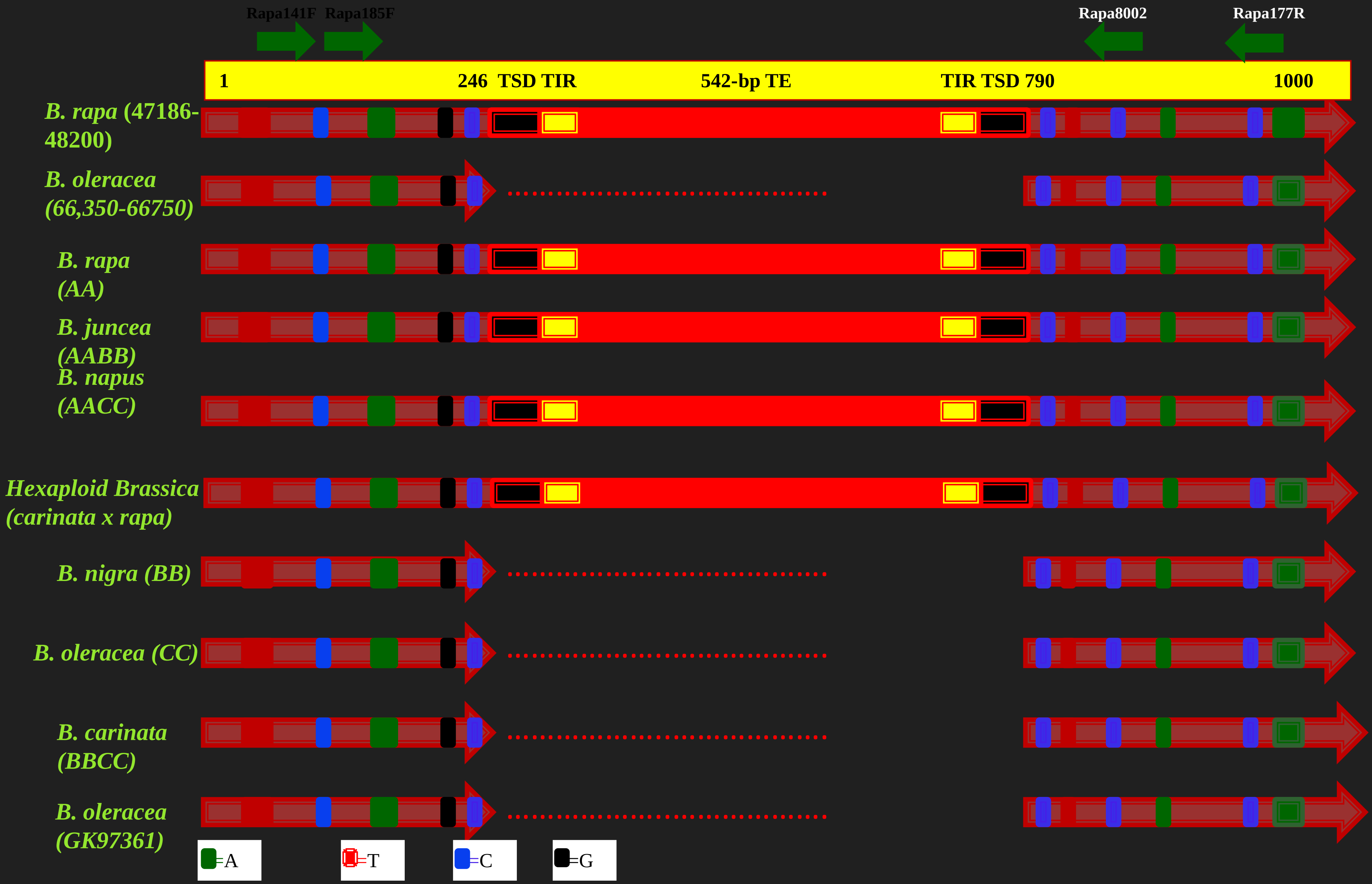
```
Boleracea:   GTCCCAACCTTTTTTCAAGAAGGTACAGTGTCAGGATTTATATGGATATACACATATCCT TTACCAC TTCTTTGTTCAATACTTTTTACAATGTTTTGGAAAGGACATGTTTCTTCTATCA
rapa:        CAACCTTTTTTCAAGAAGGTACAGTACGTGTCAGGATTTATATGGATATACACATATCCTATTGCGCAATTGTCAATAATAGCACTTTTTGAAGTTTATGTCTCAAAATAGCACTAGAAGGAGAA
                                                                            |
                                                                          47465

                                                                          66543
                                                                            |
RevComp:     TGATAGAAGAAACATGTCCTTTCCAAAACATTGTAAAAGTATTGAACAAAGAAGGGTGGTAAATAGGATATGTGTATATCCATATAAATCCTGACACTGTACCTTCTTGAAAAAAGGTTGGGAC
rapa:        CAACCTTTTTTCAAGAAGGTACAGTACGTGTCAGGATTTATATGGATATACACATATCCTATTGCGCAATTGTCAATAATAGCACTTTTTGAAGTTTATGTCTCAAAATAGCACTAGAAGGAGAA
                                                                            |
                                                                          47465
```

```
                                                        66543
GTCCCAACCTTTTTTCAAGAAGGTACAGTGTCAGGATTTATATGG ATATACACATATCC ATTTACCACCCTTCTTTGTTCAATACTTTTTACAATGTTTTGGAAAGGACATGTTTCTTCTATCA
GAGTGCTACTTTGGGAACAAAAACTTGGTTTGGTGCTATCCTAGTCTTTTTCTCTATCCTATTTACCACCCTTCTTTGTTCAATACTTTTTACAGTTTTTGGAAAGGACATGTTTCTTCTATCAT
                                                          |
                                                        48007

                                                        66543
                                                          |
TGATAGAAGAAACATGTCCTTTCCAAAACATTGTAAAAGTATTGAACAAAGAAGGGTGGTAAATAGGATATGTGTATATCCATATAAATCCTGACACTGTACCTTCTTGAAAAAAGGTTGGGAC
GAGTGCTACTTTGGGAACAAAAACTTGGTTTGGTGCTATCCTAGTCTTTTTCTCTATCCTATTTACCACCCTTCTTTGTTCAATACTTTTTACAGTTTTTGGAAAGGACATGTTTCTTCTATCAT
                                                          |
                                                        48007
```

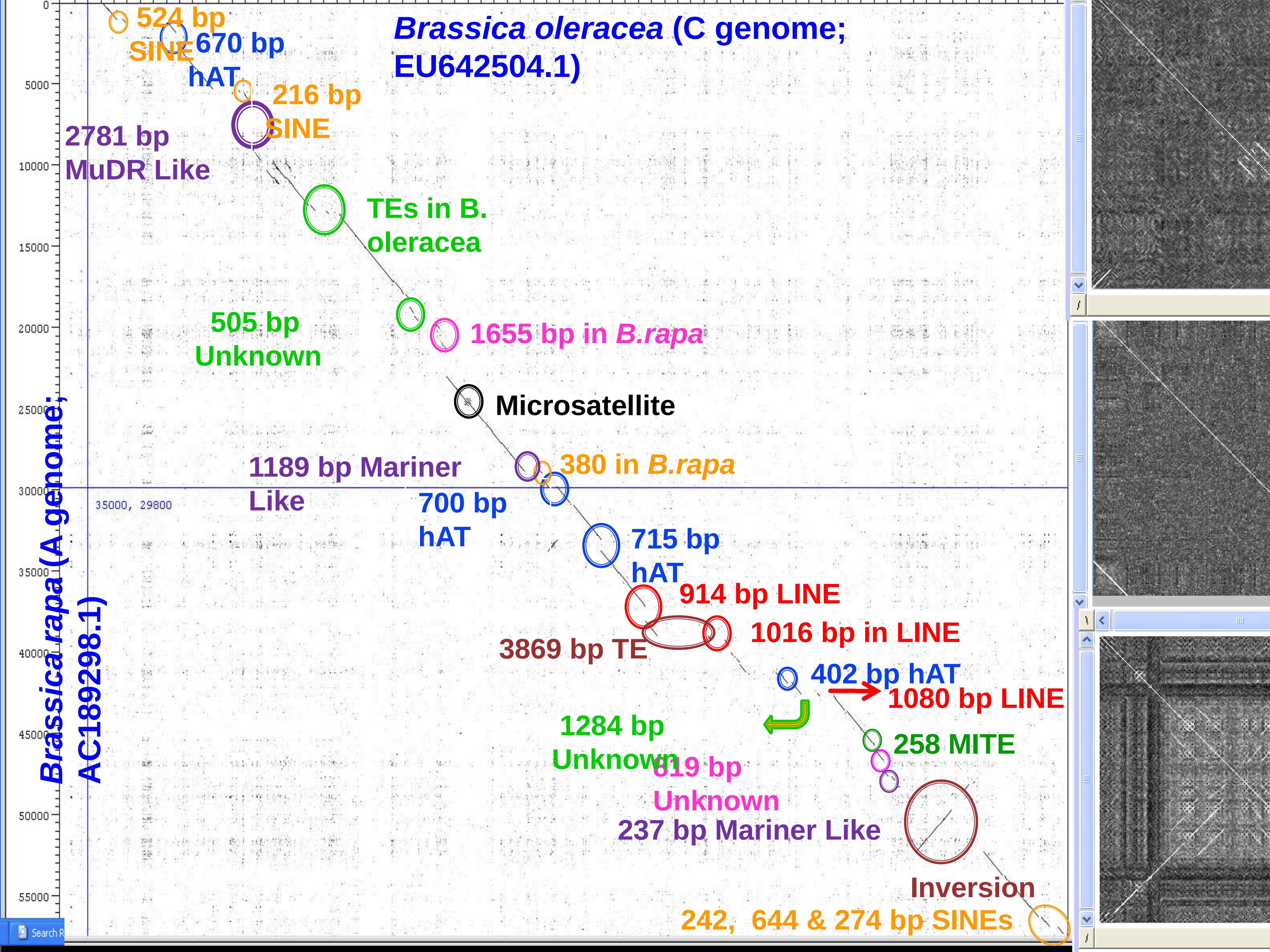
***Brassica rapa* with inserted 542bp sequence not present in *B. oleracea*. 9bp TSD (red letters and arr**
and TIR (blue). Flanking primers used in PCR (next slide) as blue arrows on sequence
Faisal Nouroz 2011

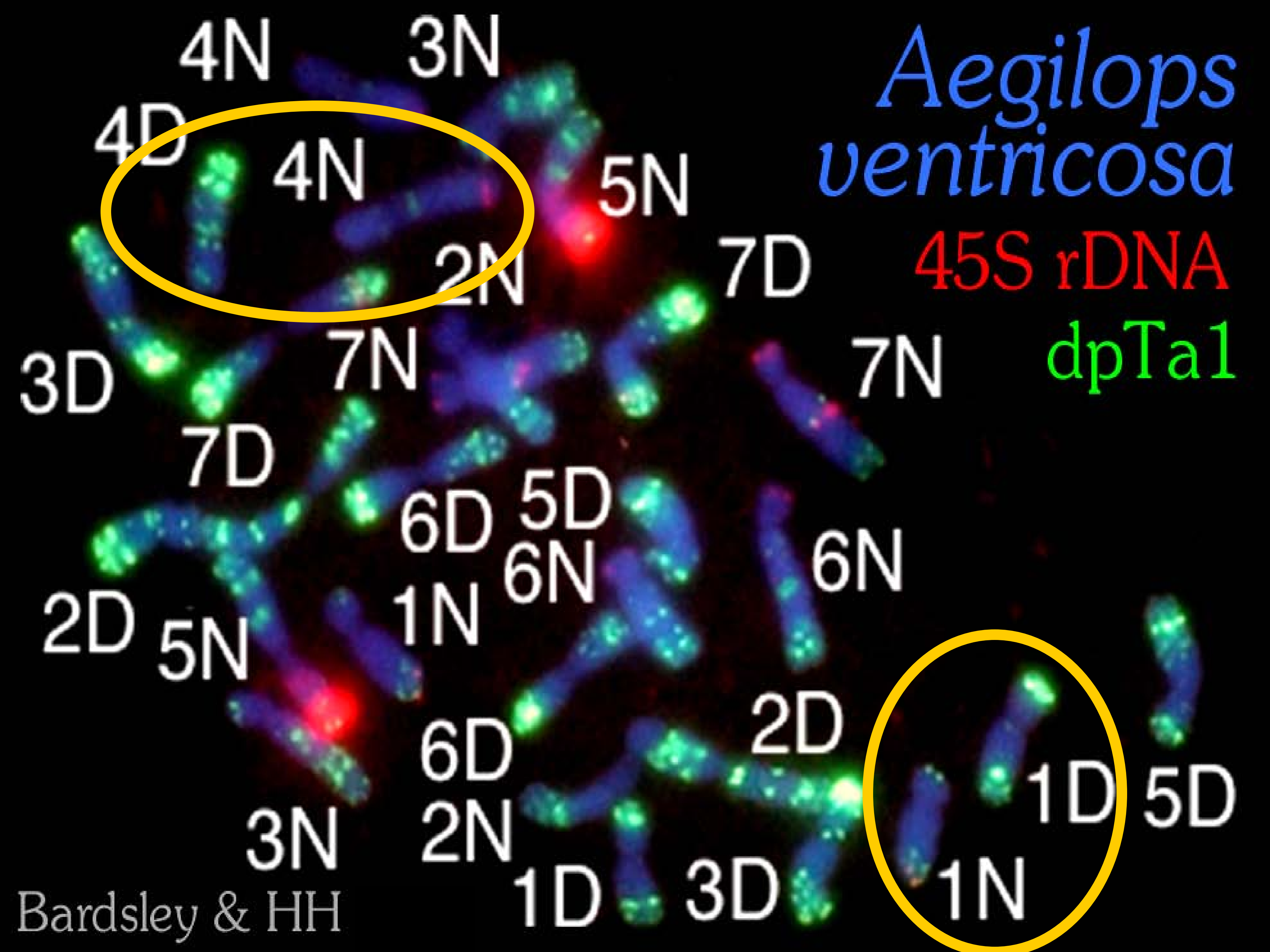

Brassica rapa
Brassica nigra
Brassica juncea
Brassica oleracea
Brassica juncea
Brassica napus
Brassica carinata
6X Brassicas
HP1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 HP1 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 HP1 33 34 35 36 37 38 39 40


Brassica rapa
Brassica nigra
Brassica juncea
Brassica oleracea
Brassica juncea
Brassica napus
Brassica carinata
6X Brassicas
HP1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 HP1 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 HP1 33 34 35 36 37 38 39 40



Schematic representation of insertion in Brassica rapa and other Brassica genomes. Green, red, blue and black boxes showing DNA motifs.


Brassica oleracea (C genome; EU642504.1)
Brassica rapa (A genome; AC189298.1)
0
5000
10000
15000
20000
25000
30000
35000
40000
45000
50000
55000
35000, 29800
524 bp SINE
670 bp hAT
216 bp SINE
2781 bp MuDR Like
TEs in B. oleracea
505 bp Unknown
1655 bp in B.rapa
Microsatellite
1189 bp Mariner Like
380 in B.rapa
700 bp hAT
715 bp hAT
914 bp LINE
1016 bp in LINE
3869 bp TE
402 bp hAT
1080 bp LINE
1284 bp Unknown
258 MITE
819 bp Unknown
237 bp Mariner Like
Inversion
242, 644 & 274 bp SINEs
Search R

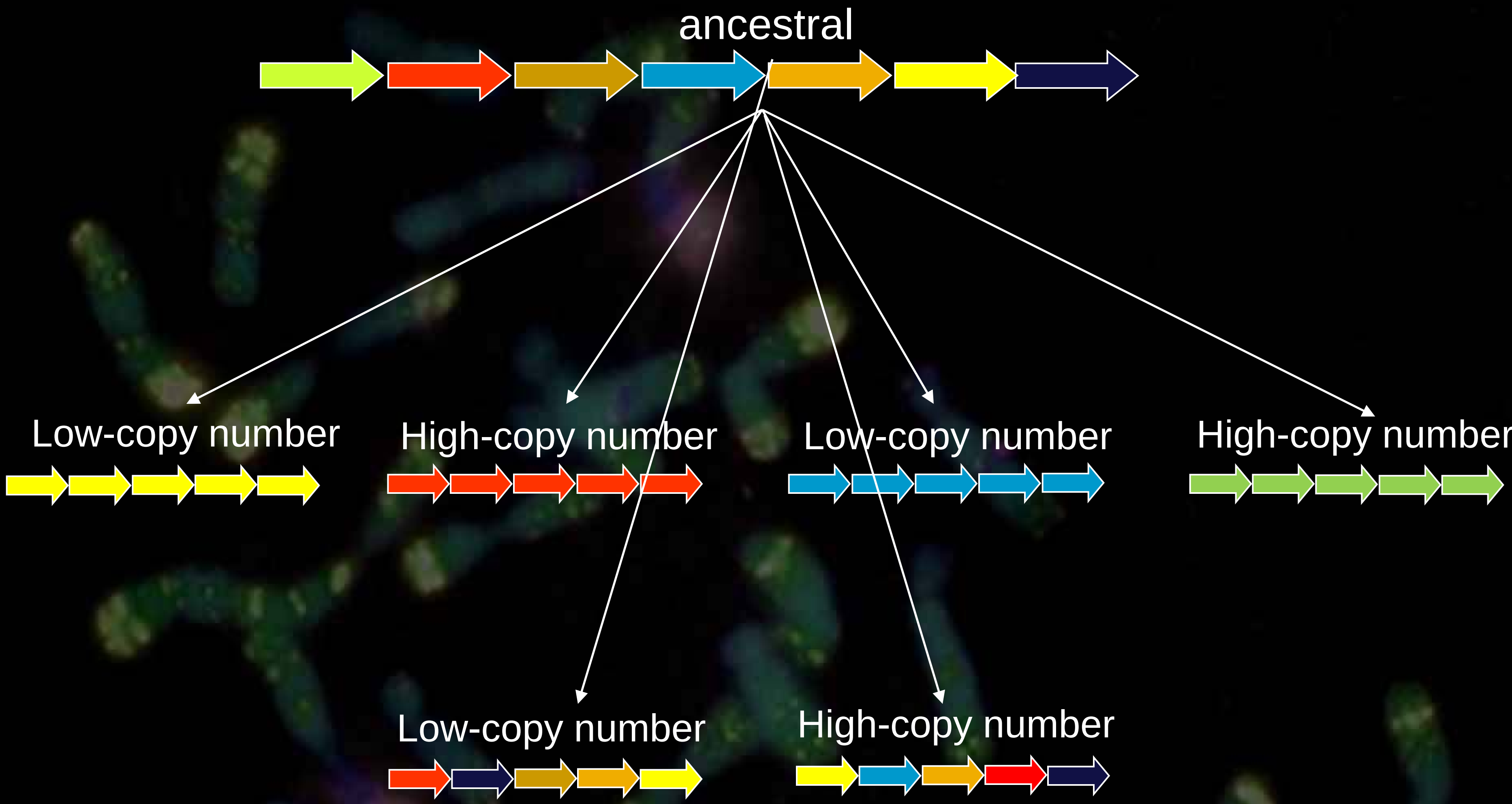


Aegilops ventricosa

45S rDNA

dpTa1

Bardsley & HH



ancestral

Low-copy number

High-copy number

Low-copy number

High-copy number

Low-copy number

High-copy number

High copy spp: homogenized, amplification from a limited number of master copies

Low copy spp: much variation

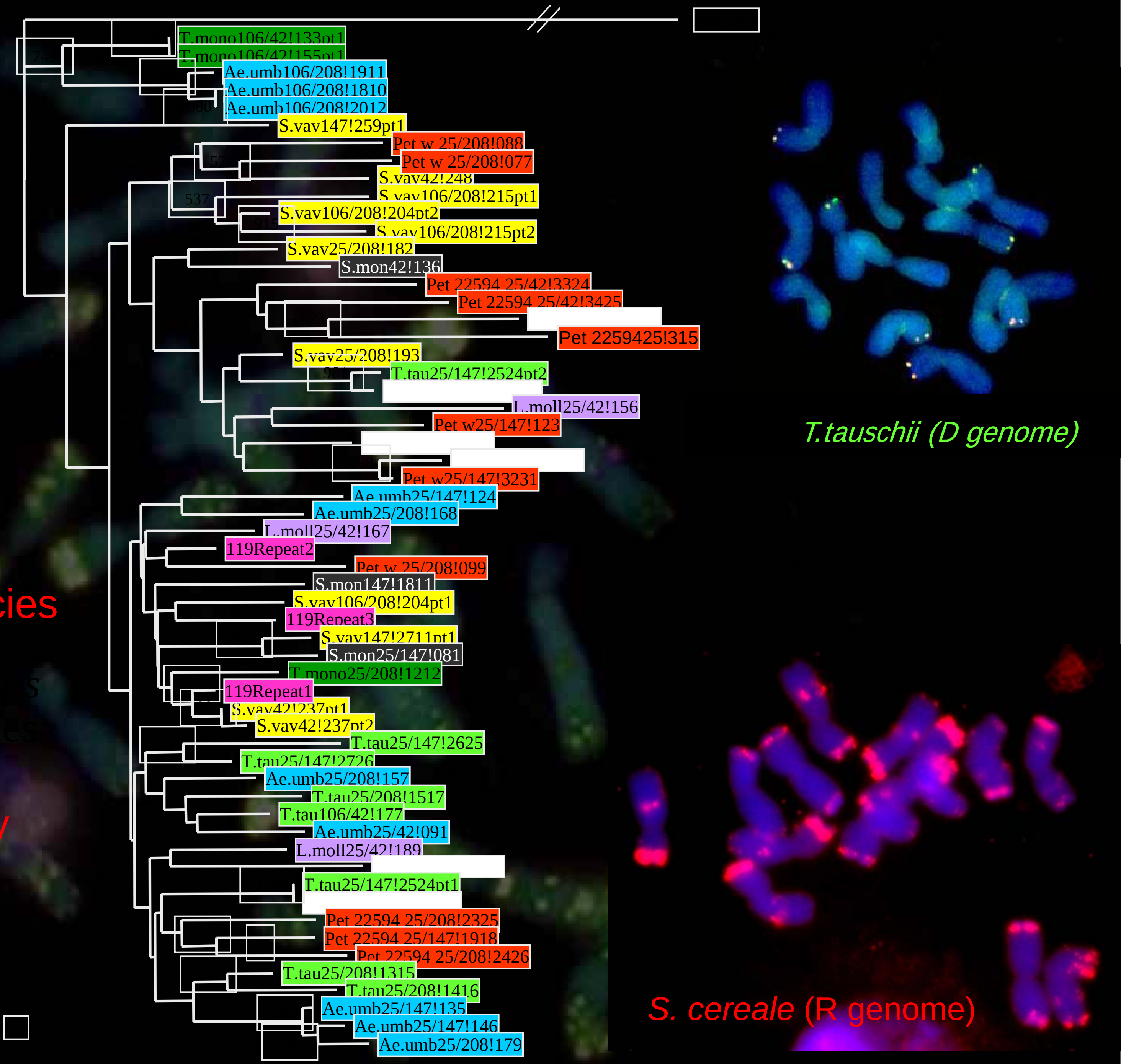
120bp repeat unit family

in *Triticum*, *Aegilops* and *Secale* species

Colour block represents species

High copy,

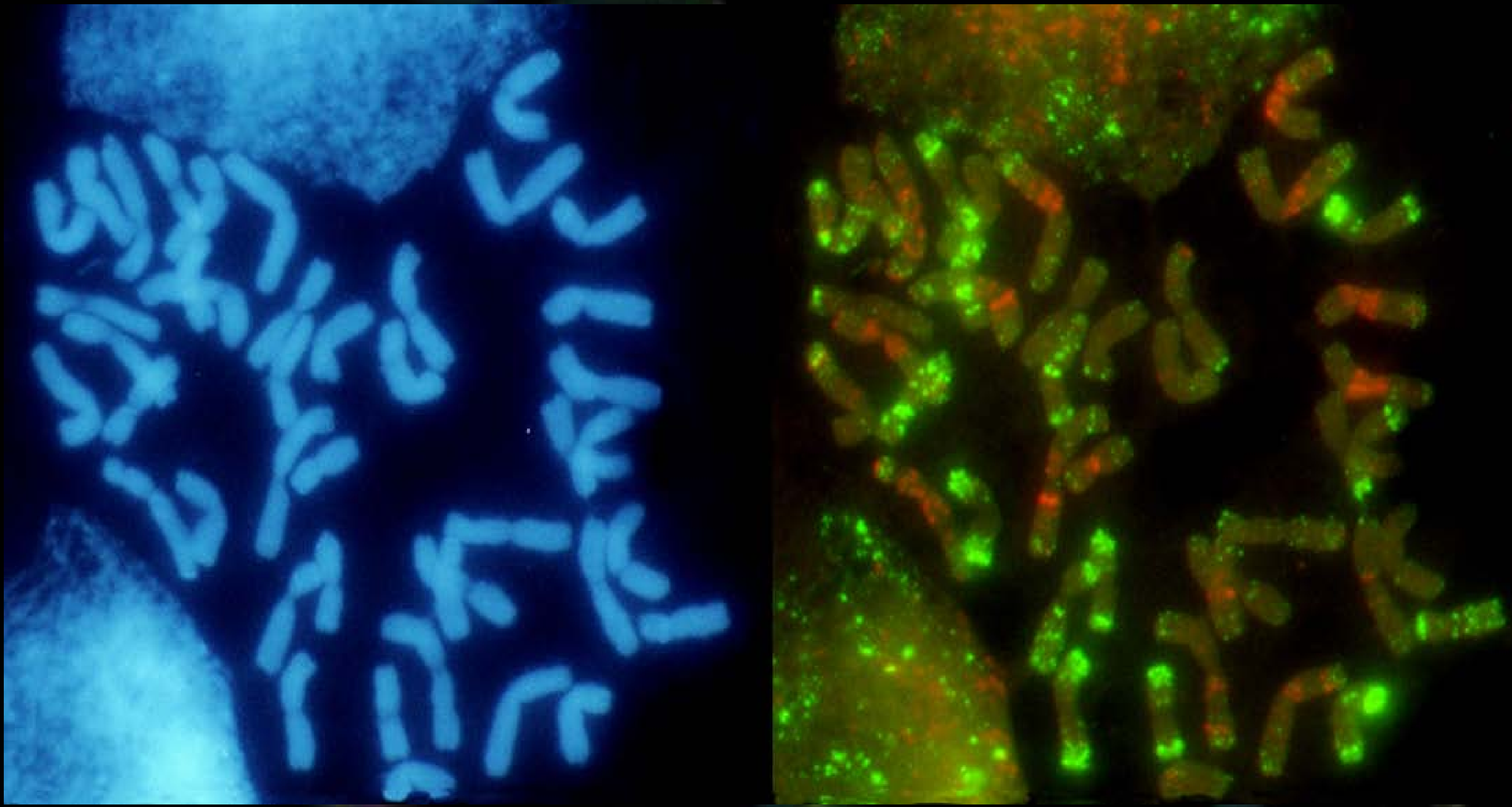
High diversity



T.tauschii (D genome)

S. cereale (R genome)

Triticum aestivum 2n=6x=42

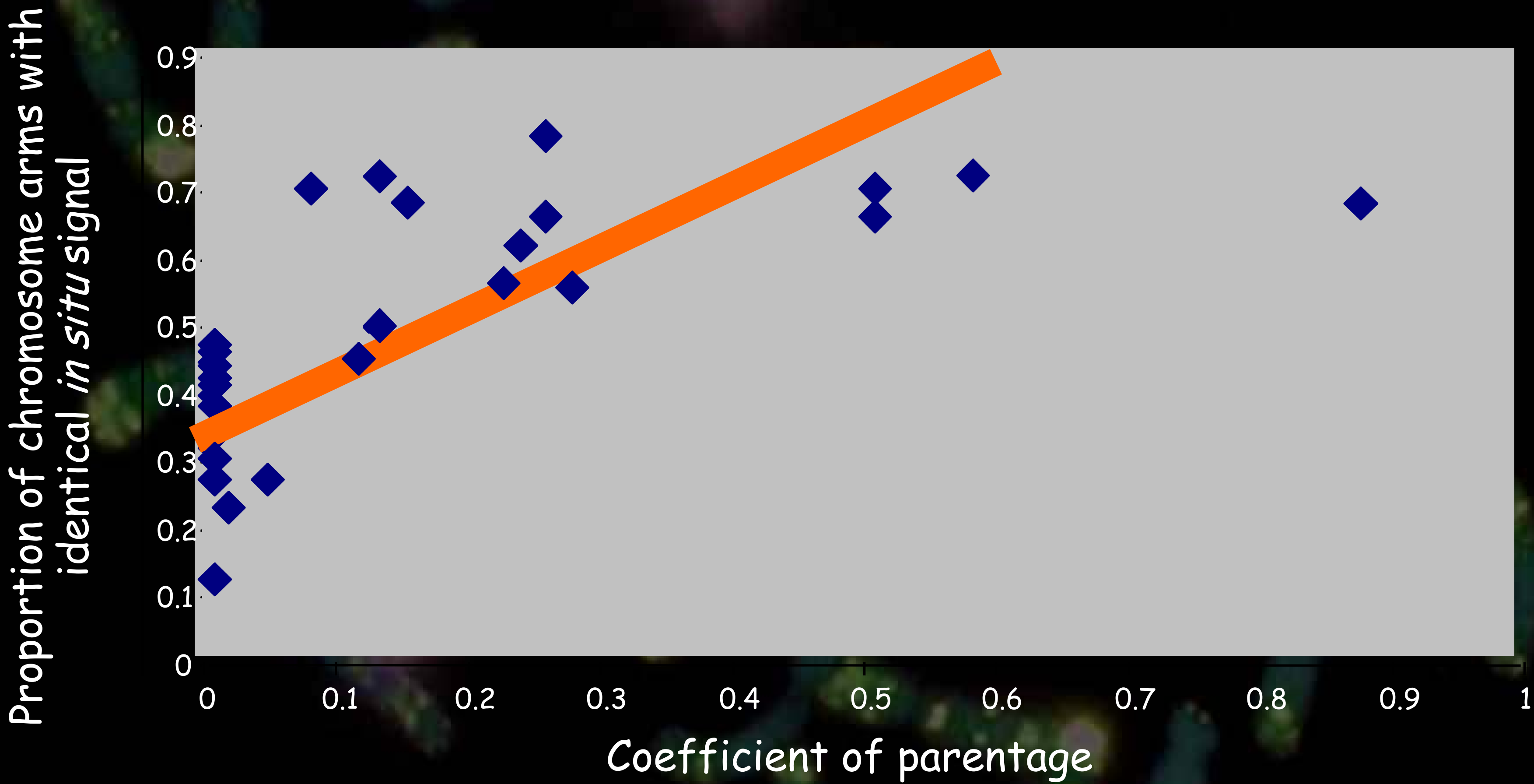


DAPI

pSc119.2 dpTa1

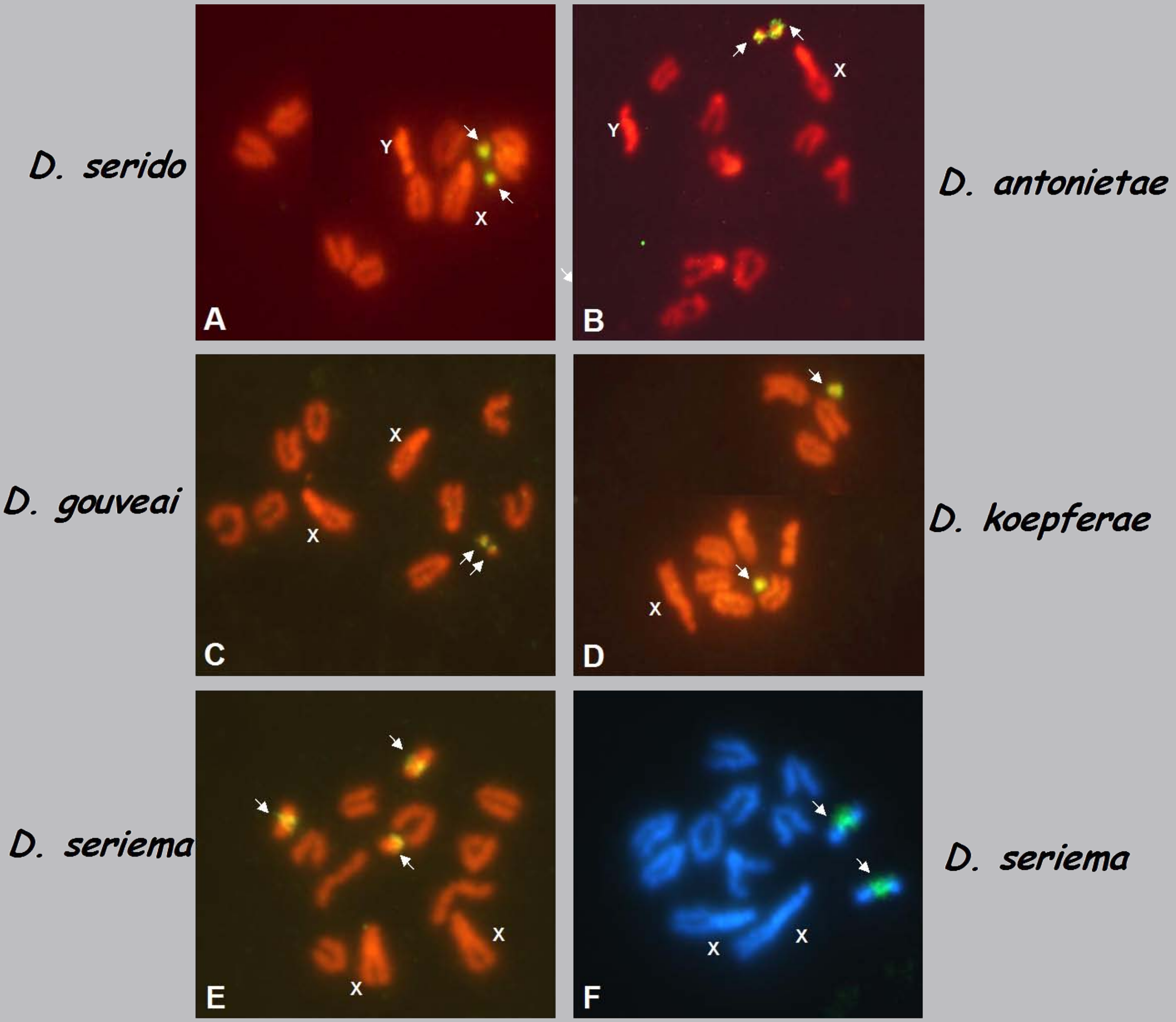
High copy number – low diversity in each of 3 genomes

Correlation between genetic relationships and similarity of dpTa1 hybridization

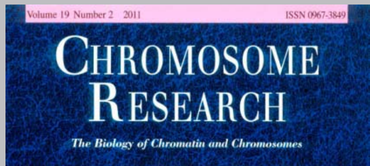


Chromosomal location of DBC-150 repeats

Micro- or dot chromosomes



Kuhn (Belo Horizonte, Brazil) et al. *Chromosome Research*


Volume 19 Number 2 2011 ISSN 0967-3849

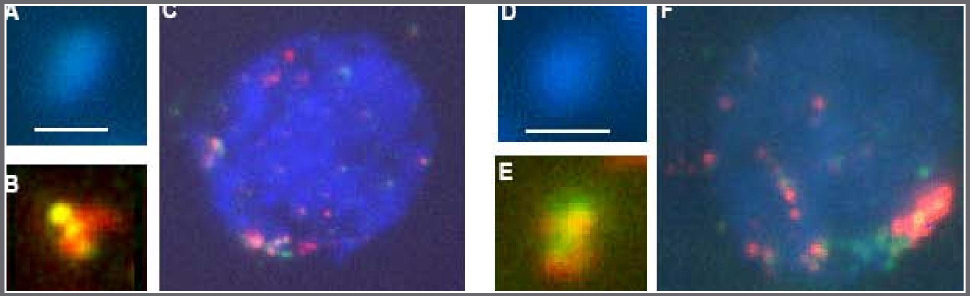
CHROMOSOME RESEARCH

The Biology of Chromatin and Chromosomes

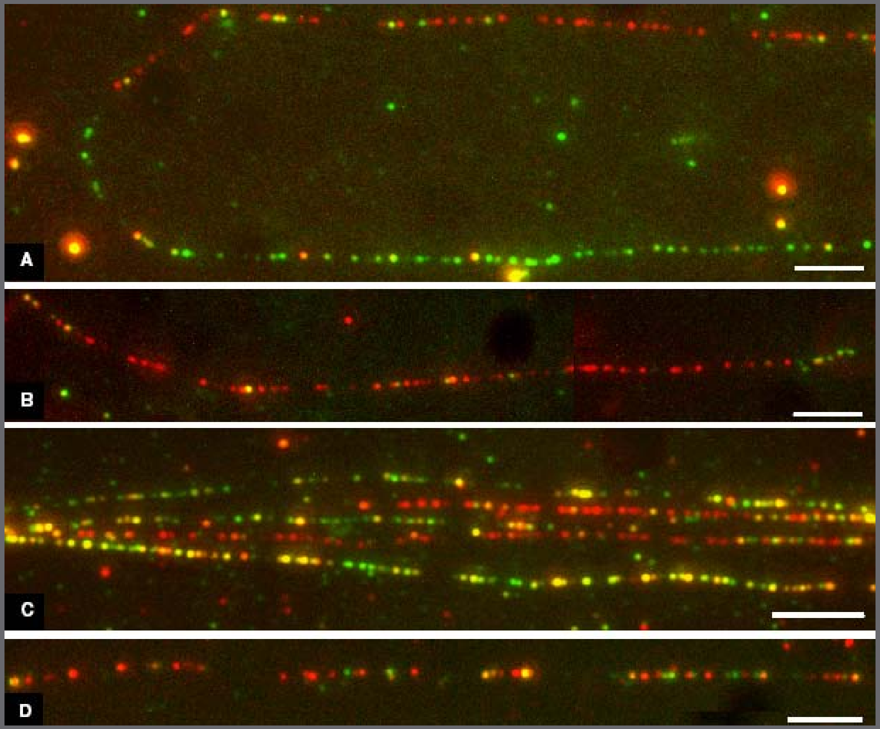
Interspersion of pBuM and DBC-150

D. gouveai

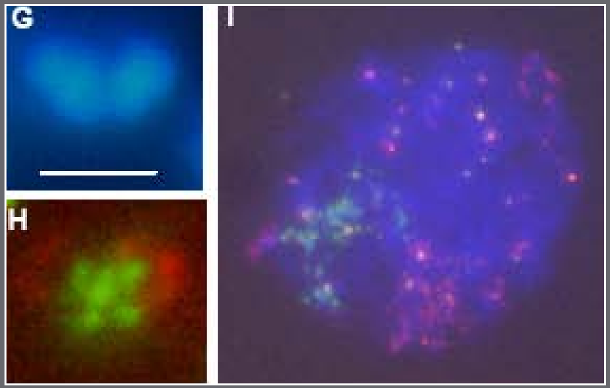
D. antoneita



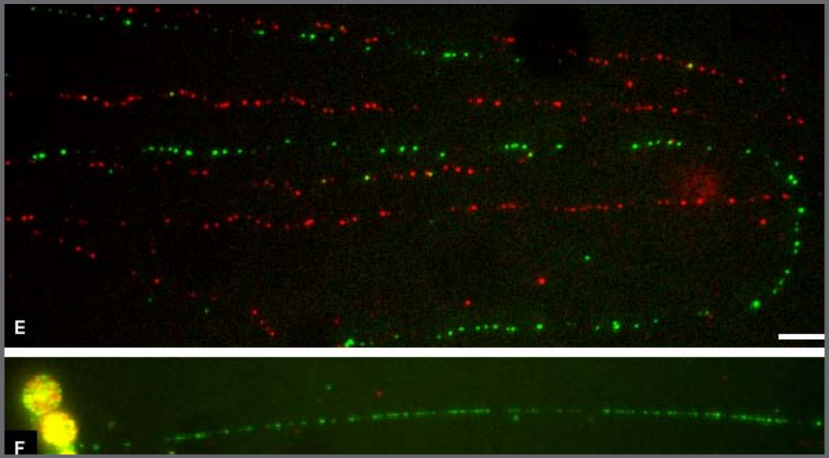
High interspersion



D. seriema

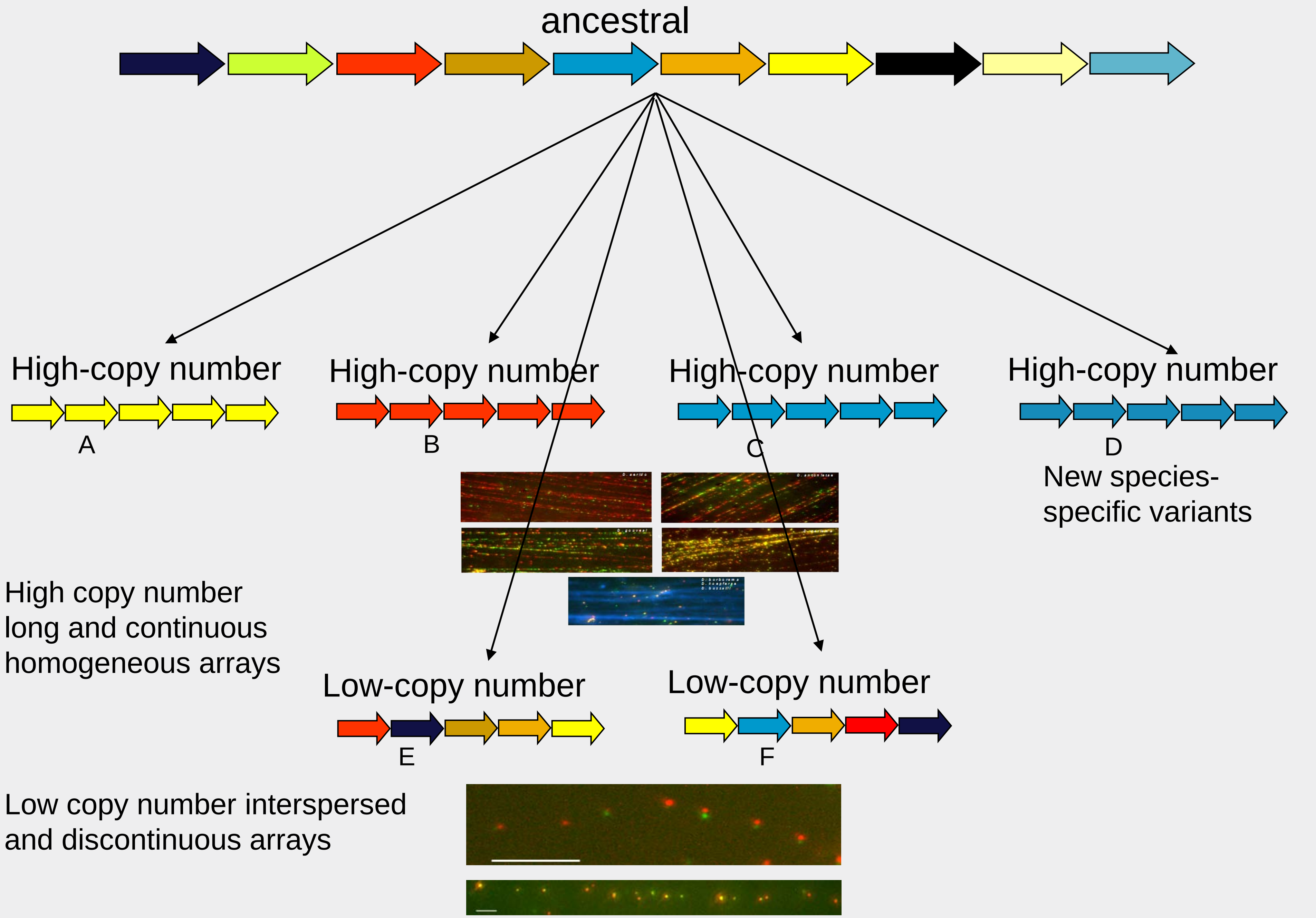


Low interspersion



Non-homologous repeats

pBuM DBC-150



New species-specific variants

High copy number long and continuous homogeneous arrays

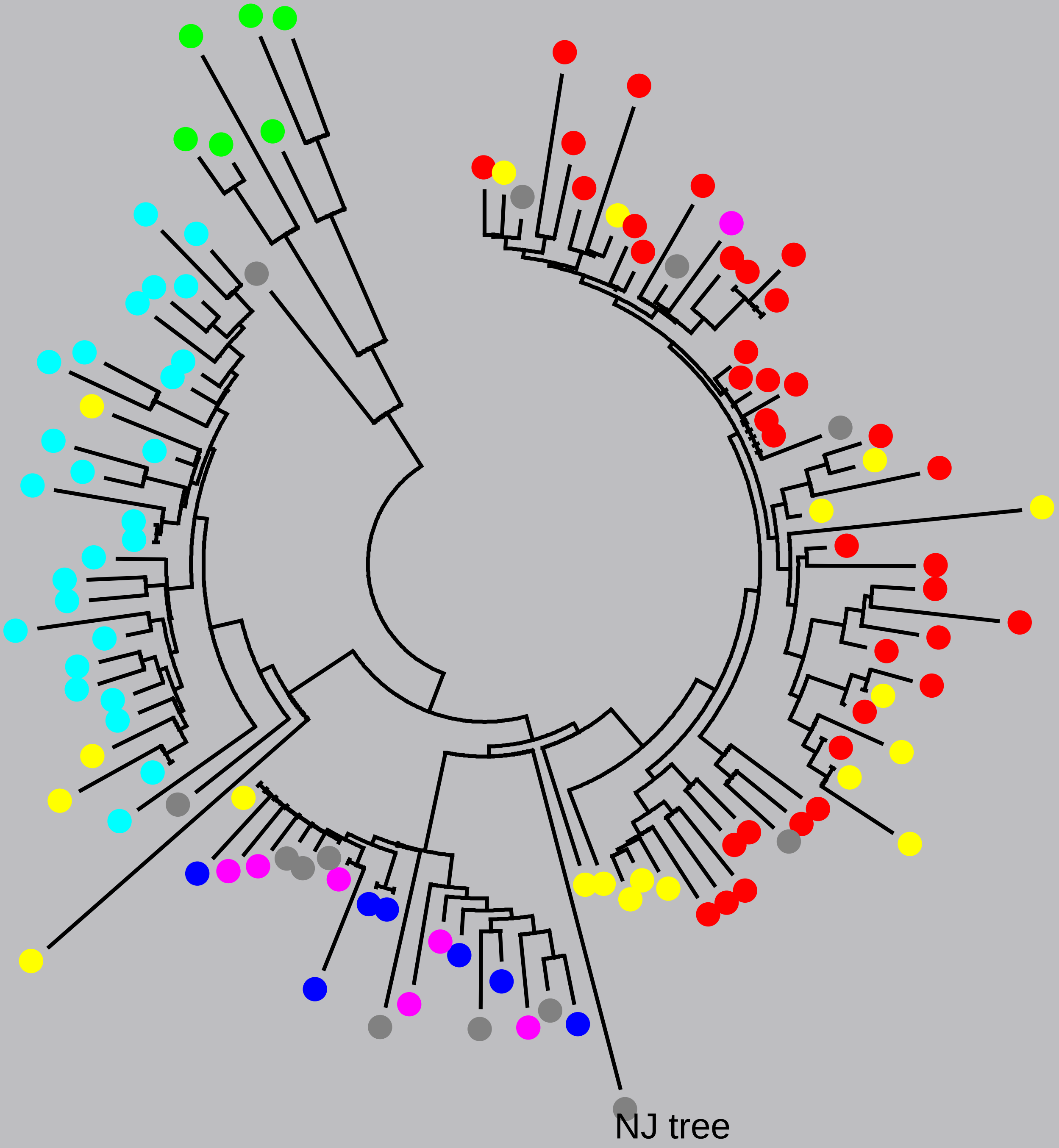
Low copy number interspersed and discontinuous arrays

From species level to populations …

Alpha/beta repeat variation

●
●
●
●
High copy species

●
●
●
Low copy species



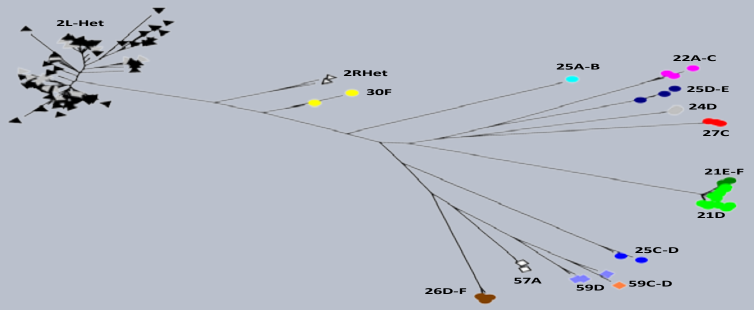
Total = 119 repeats

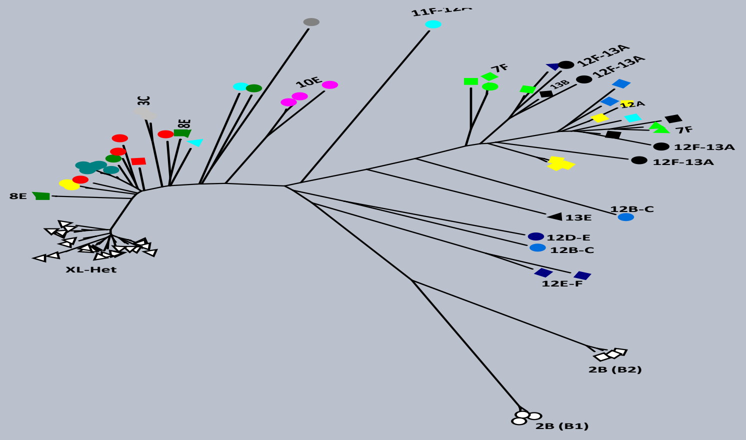
1.688 tandem repeats in *Drosophila melanogaster*

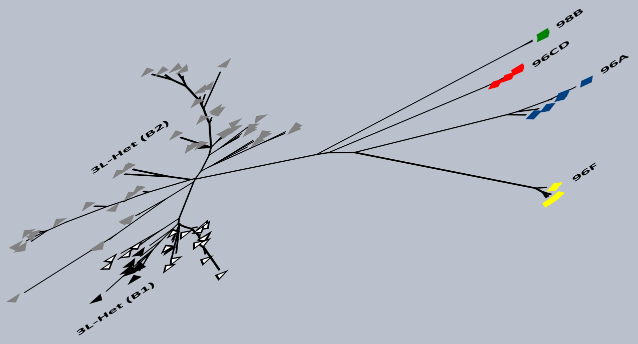
Large arrays n heterochromatin of chromosomes 2, 3 and X
Short arrays are found in the euchromatin

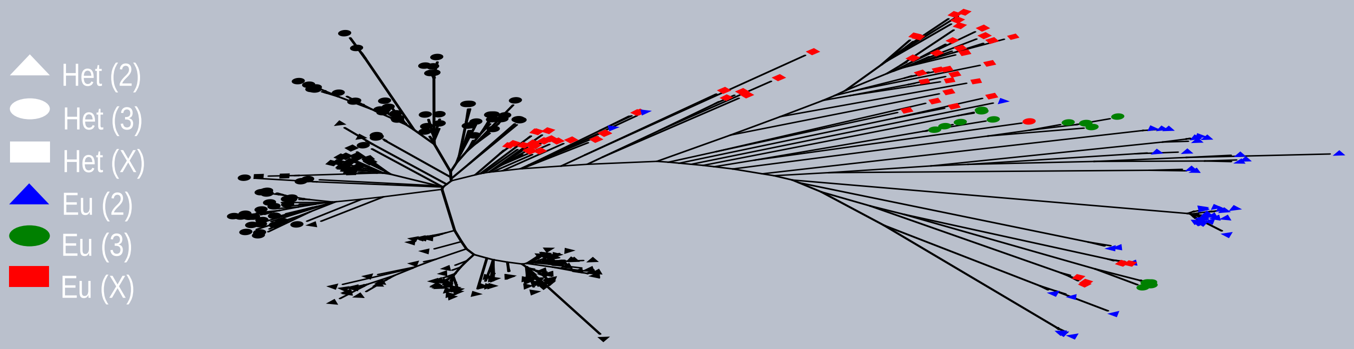
Kuhn et al. 2011 in press Mol Biol Evol

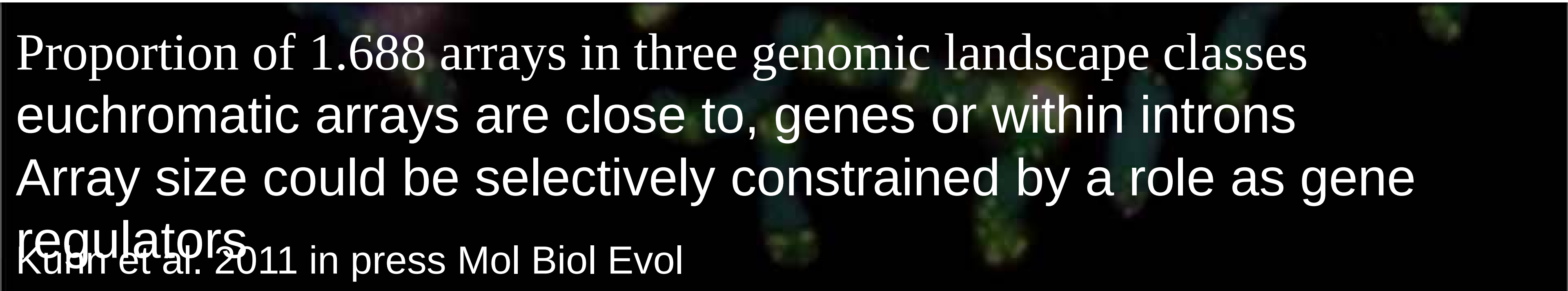
Homogenization of arrays: differential for hetero- and eu-chromatin











Proportion of 1.688 arrays in three genomic landscape classes

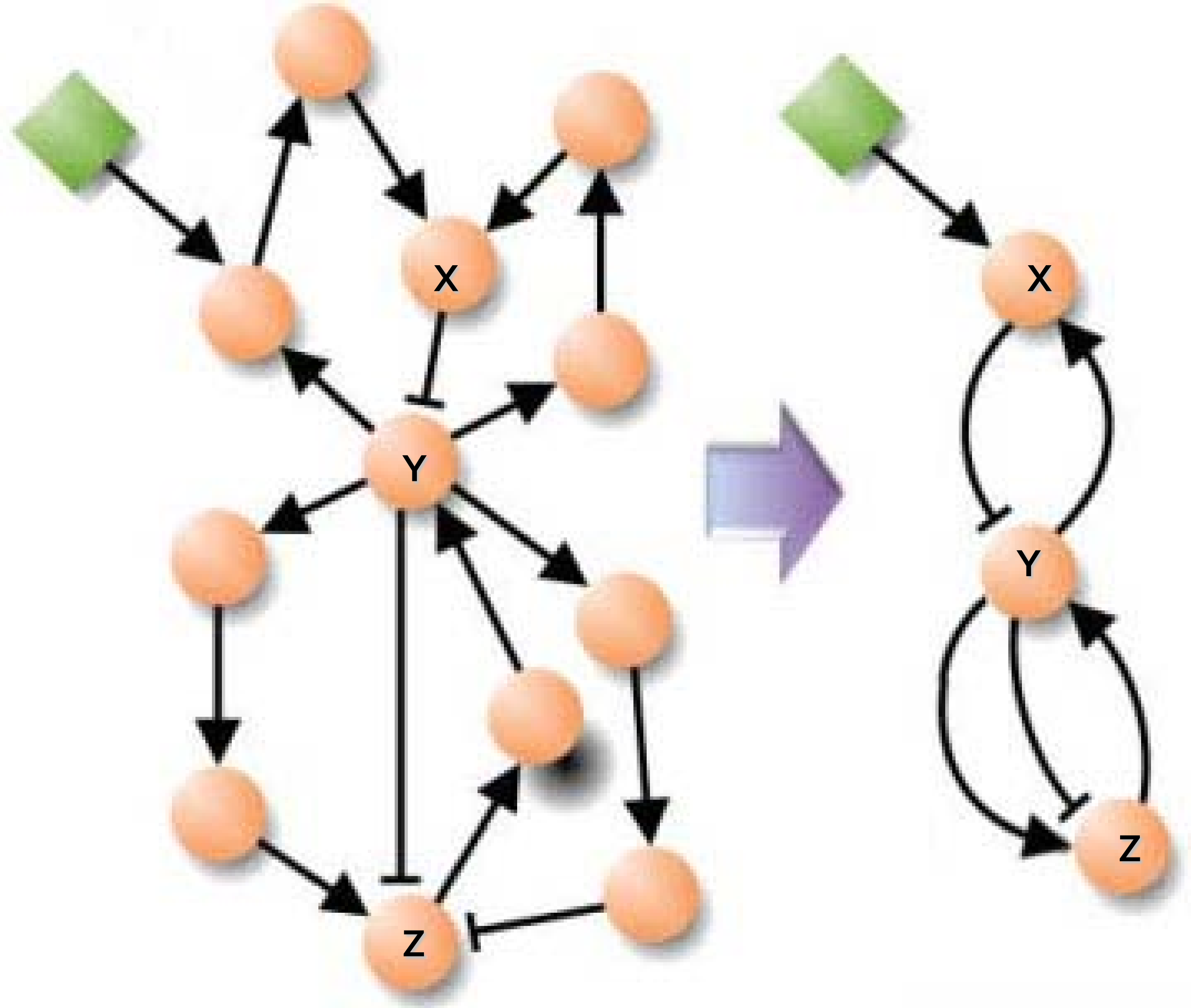
euchromatic arrays are close to, genes or within introns

Array size could be selectively constrained by a role as gene regulators

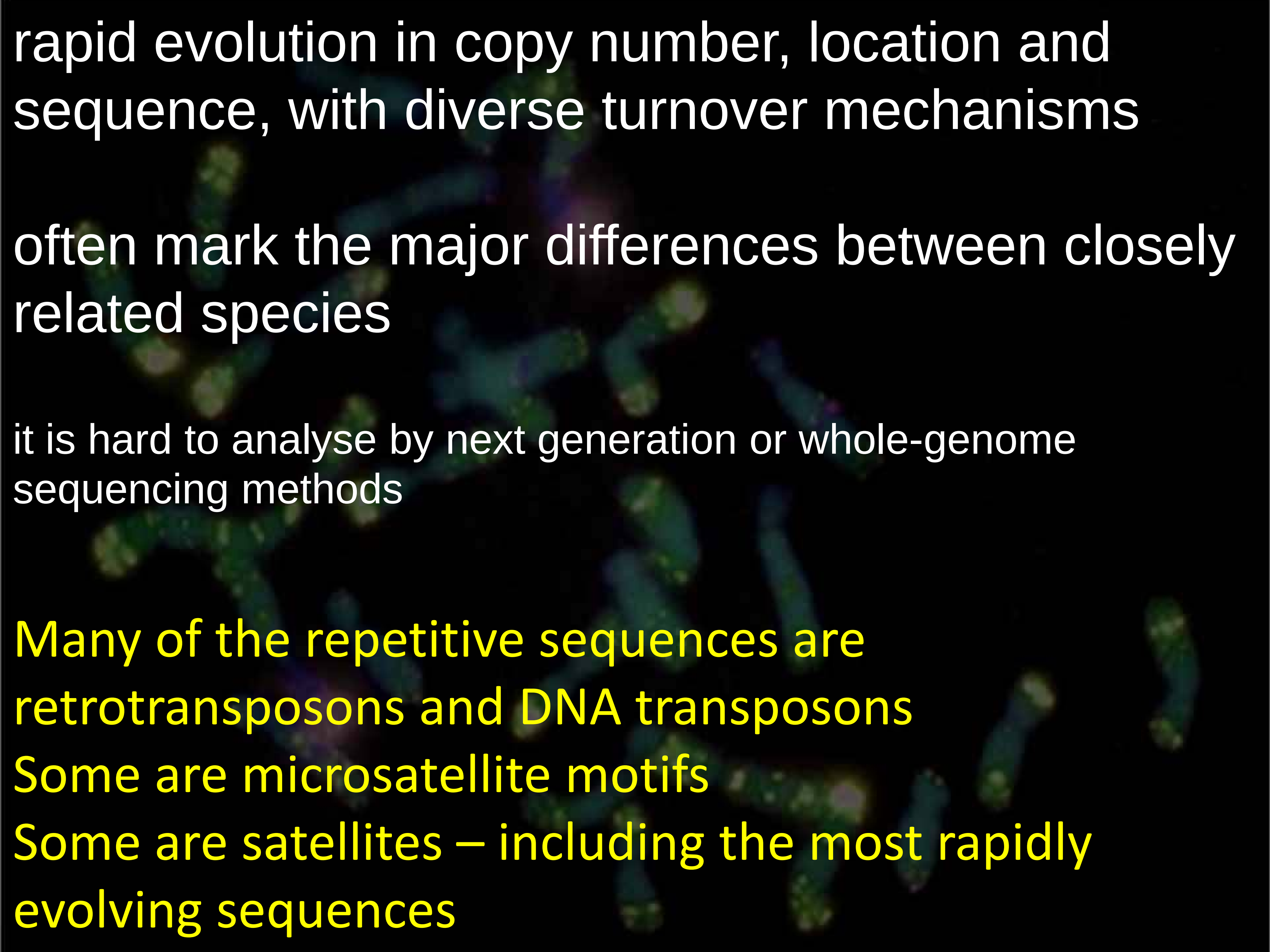
Kuhn et al. 2011 in press Mol Biol Evol

Network reduction

Circadian Clock regulation
after Leloup & Goldbeter
cf Andrew Millar
in Arabidopsis



Kim, HH, Cho et al. 2011 Science Signaling

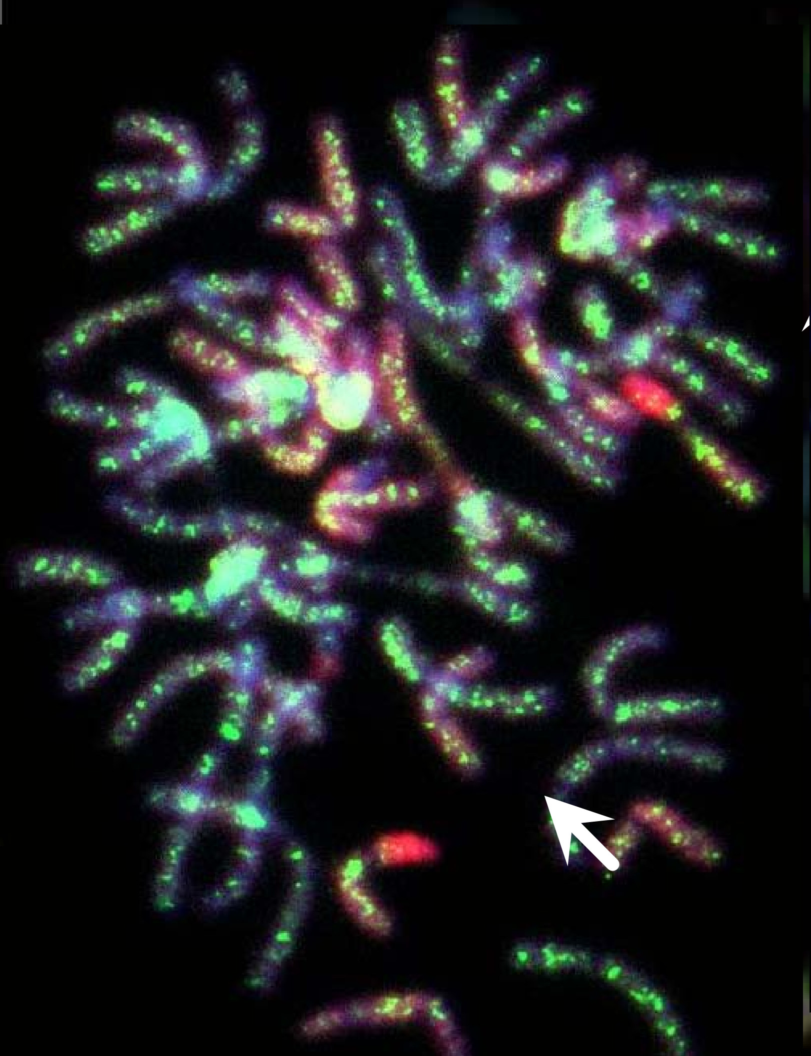


rapid evolution in copy number, location and sequence, with diverse turnover mechanisms

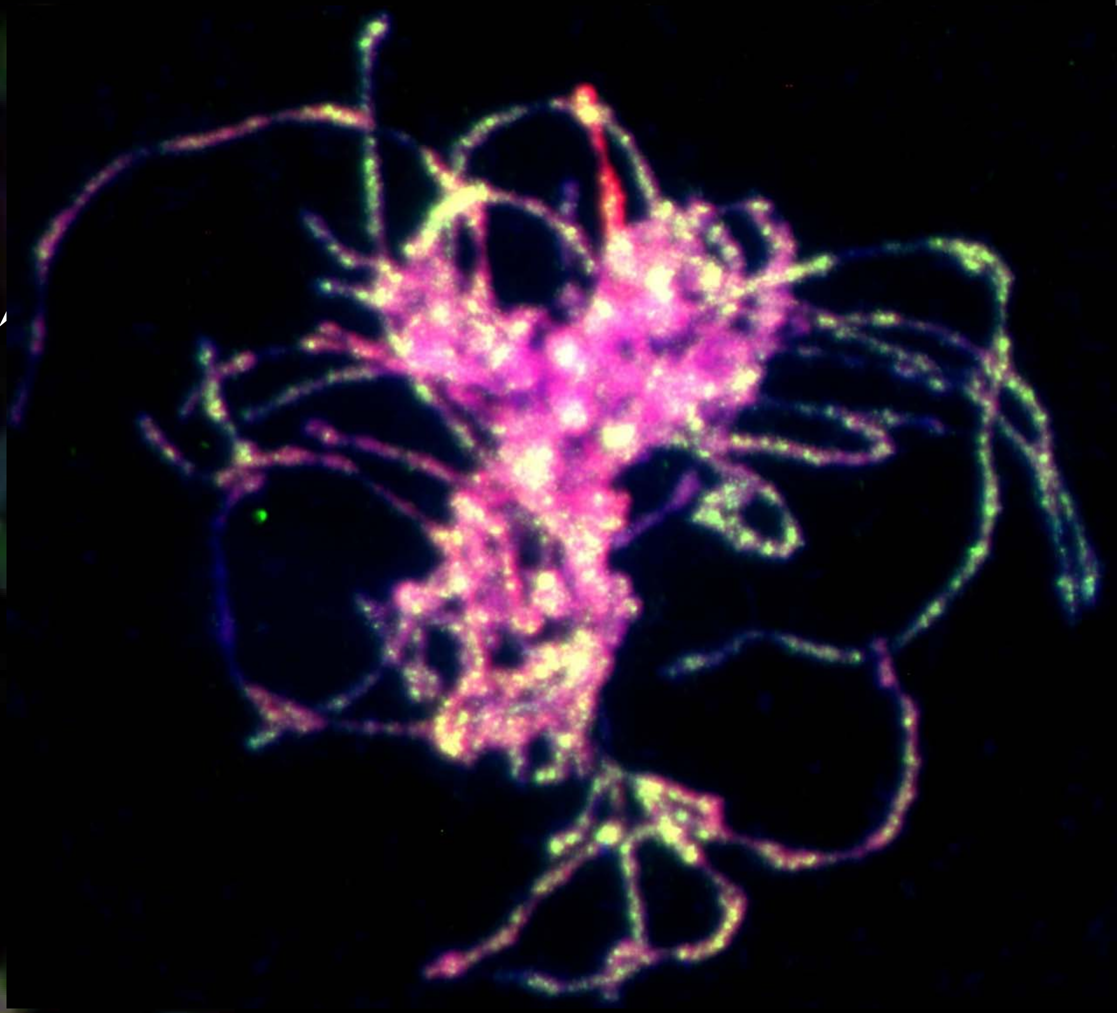
often mark the major differences between closely related species

it is hard to analyse by next generation or whole-genome sequencing methods

Many of the repetitive sequences are retrotransposons and DNA transposons
Some are microsatellite motifs
Some are satellites – including the most rapidly evolving sequences



Th. intermedium **Anti 5*meC***



Th. intermedium **DNA Anti 5*meC***

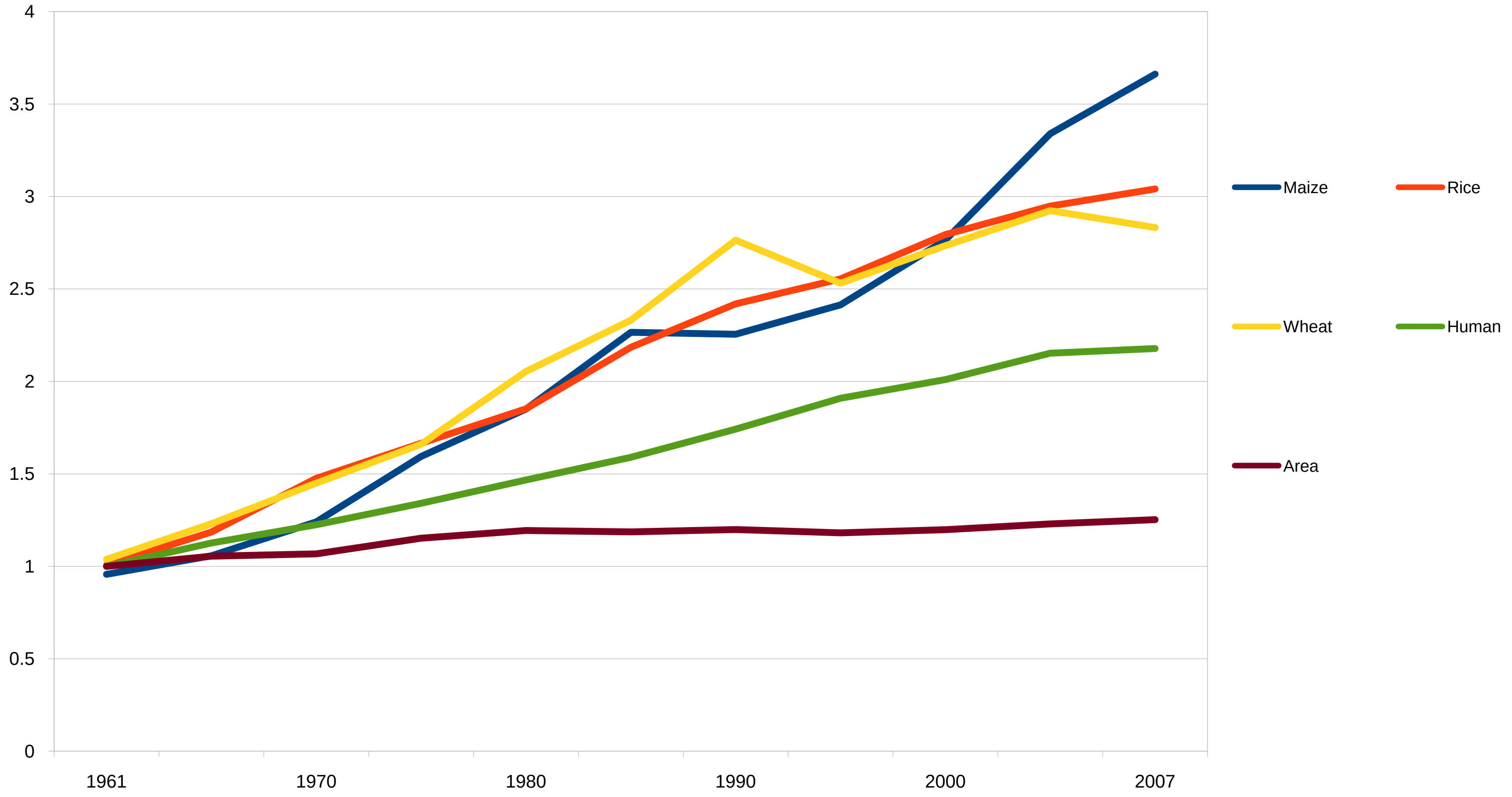
Niaz Ali, Trude Schwarzacher – Poster 59

Wsm-1: only highly effective source of resistance to WSMV

Greybosch et al. 2009


ank You For
Shopping

50 years of plant breeding progress



United Nations
Millennium Development Goals- MDGs

Goal 1 – Eradicate extreme poverty and hunger

Goal 2 – Achieve universal primary education

Goal 3 – Promote gender equity and empower women

Goal 4 – Reduce child mortality

Goal 5 – Improve maternal health

Goal 6- Combat HIV/AIDS, malaria and other diseases

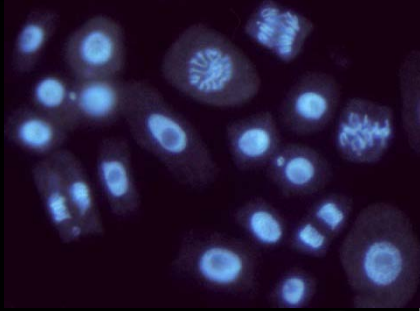
Goal 7 - Ensure environmental sustainability

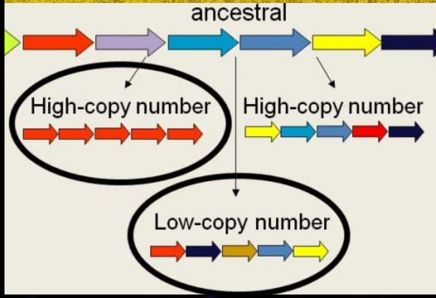
Goal 8 - Develop a global partnership for development

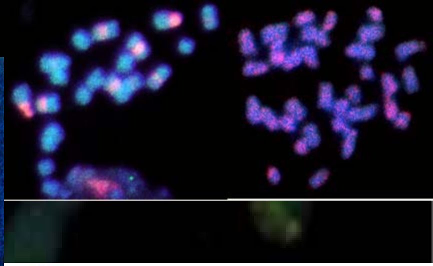
Rapid evolution of repetitive DNA: Chromosome evolution and speciation

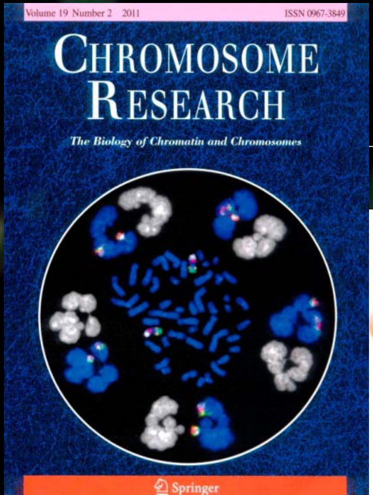
Pat Heslop-Harrison
phh4@le.ac.uk
www.molcyt.com pw/user: ‘visitor’

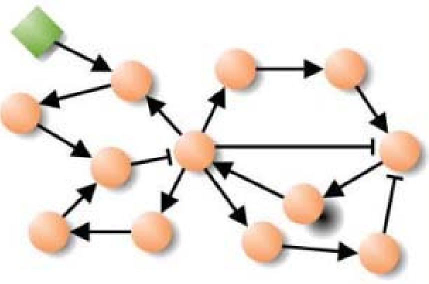
Social media: #ICC18 and
Pathh1 on Twitter
Reports: AoBBlog.com and
Storify.com/pathh1

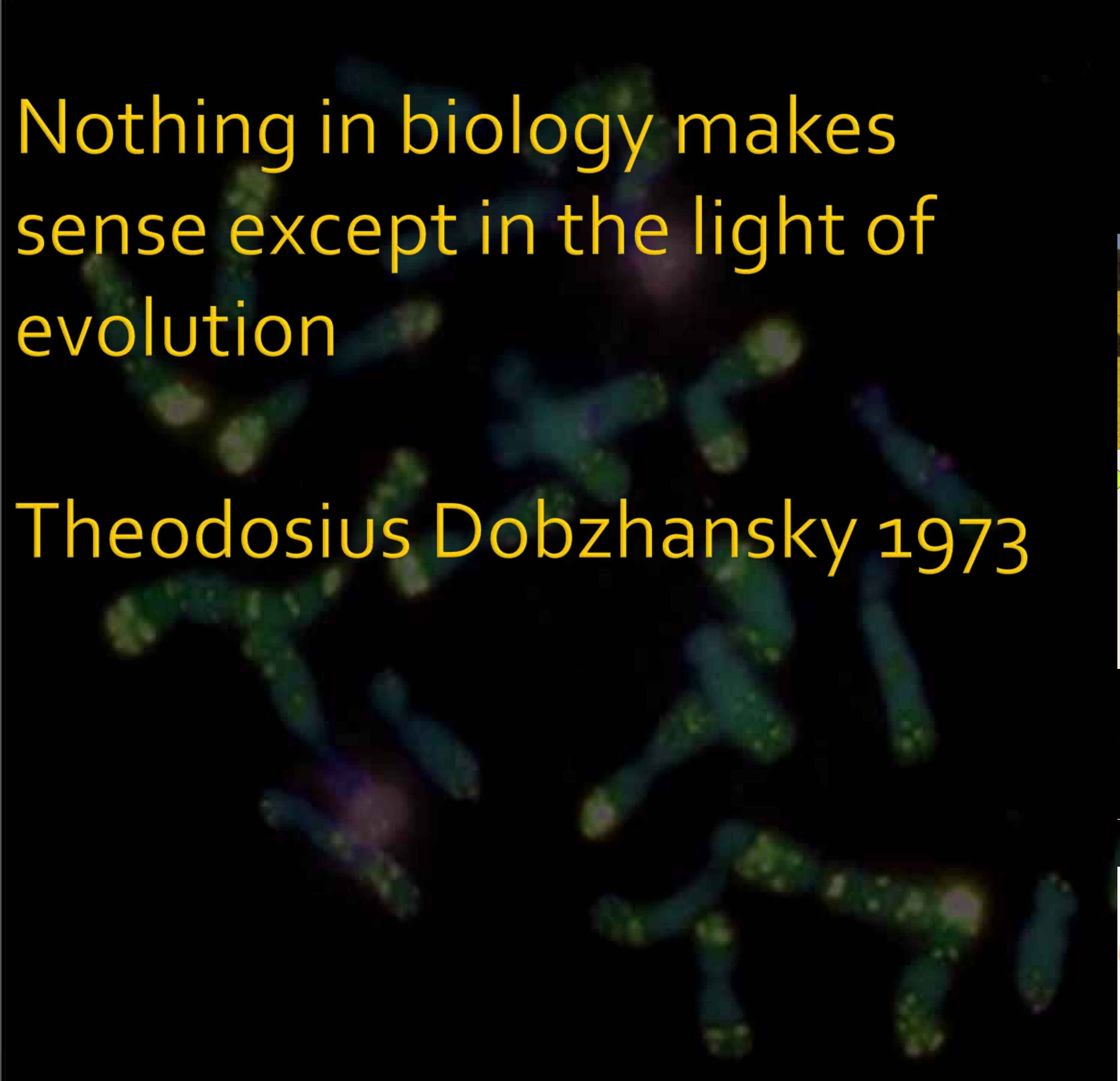







Volume 19 Number 2 2011 ISSN 0967-3849
CHROMOSOME RESEARCH
The Biology of Chromatin and Chromosomes
Springer





Nothing in biology makes sense except in the light of evolution

Theodosius Dobzhansky 1973