

Widespread Gene-Scale Structural Variants Revealed by Long-Range Sequencing



Harmeet Singh Chawla

 @GiessenSingh

Department of Plant Breeding, Justus Liebig University, Giessen, Germany

Outline

Introduction



**Genome-wide identification of
small intra-genic SV**



**SV associated with eco-
geographical adaptations**



**SV contributing to disease
resistance**

Outline

Introduction



**Genome-wide identification of
small intra-genic SV**

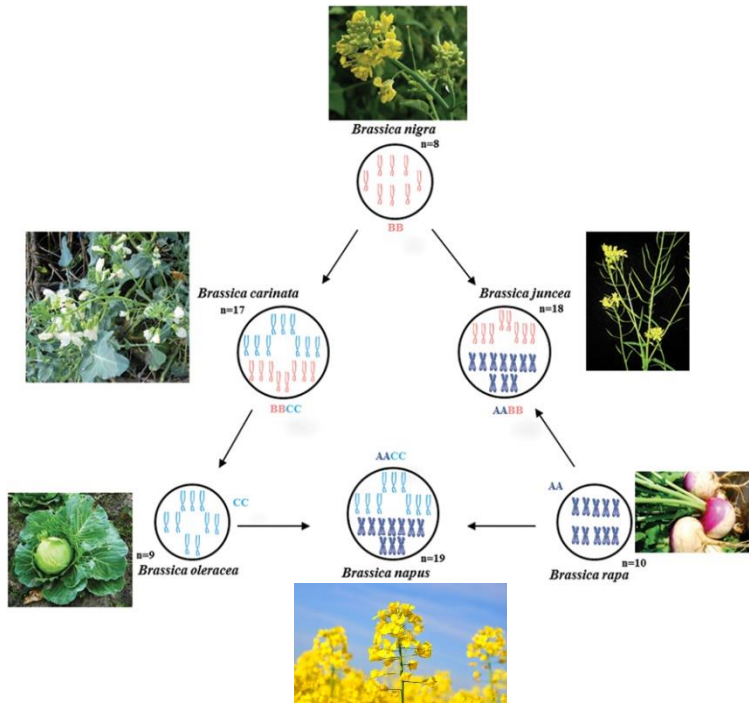


**SV associated with eco-
geographical adaptations**

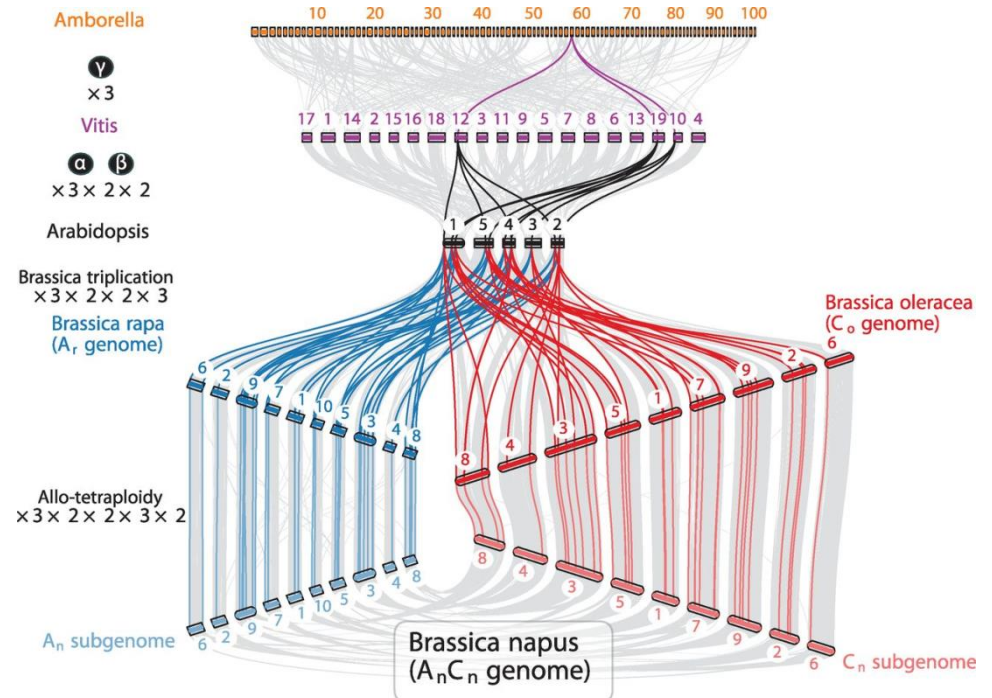


**SV contributing to disease
resistance**

Brassica napus



Extensive genome structural variations both at large chromosomal scale and small gene scale



Chalhoub et al., *Science* (2014)

Genome Structural Variations (SV)

Reference



Genome Structural Variations (SV)

Reference

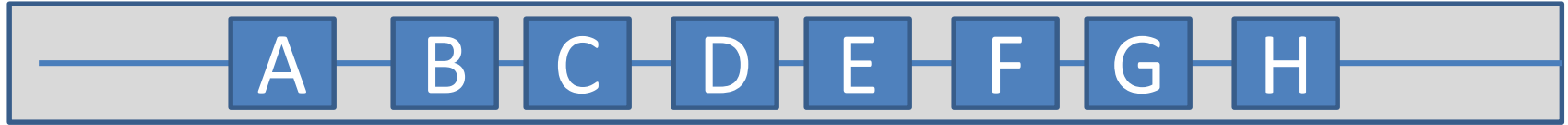


Novel-Insertion



Genome Structural Variations (SV)

Reference



Novel-Insertion



Deletion



Genome Structural Variations (SV)

Reference



Novel-Insertion



Deletion



Copy Number Variations



Genome Structural Variations (SV)

Reference



Novel-Insertion



Deletion



Copy Number Variations



Why study SV?

SV underlining important traits in rapeseed

Flowering time

ORIGINAL RESEARCH ARTICLE
published: 25 August 2014
doi: 10.3389/fpls.2014.00404



frontiers in
PLANT SCIENCE

Capturing sequence variation among flowering-time regulatory gene homologs in the allopolyploid crop species *Brassica napus*

Sarah Schiessl^{1*}, Birgit Samans¹, Bruno Hüttel², Richard B. Smith³

Plant Biotechnology Journal (2018) 16, pp. 2102–2112

doi: 10.1111/pbi.12942

Disease resistance

Finding invisible quantitative trait loci with missing data

Julian Gabur^{1,#}, Harmeet S. Chawla^{1,#}, Xiwei Liu¹, Vinod Kumar², Sébastien Fauriol³, Christophe Jestin⁵, Emmanuelle Dancourt⁶

Plant Biotechnology Journal (2018)

doi: 10.1111/pbi.12134

Seed quality traits

Mapping of homoeologous chromosome exchanges influencing quantitative trait variation in *Brassica napus*

Anna Stein^{1,*}, Olivier Coriton², Mathieu Rousseau-Gueutin², Birgit Samans¹, Sarah V. Schiessl¹, Christian Obermeier¹, Isobel A.P. Parkin³, Anne-Marie Chèvre² and Rod J. Snowdon¹

¹Department of Plant Breeding, IFZ Research Centre for Biosystems, Land Use and Nutrition, Justus Liebig University, Giessen, Germany

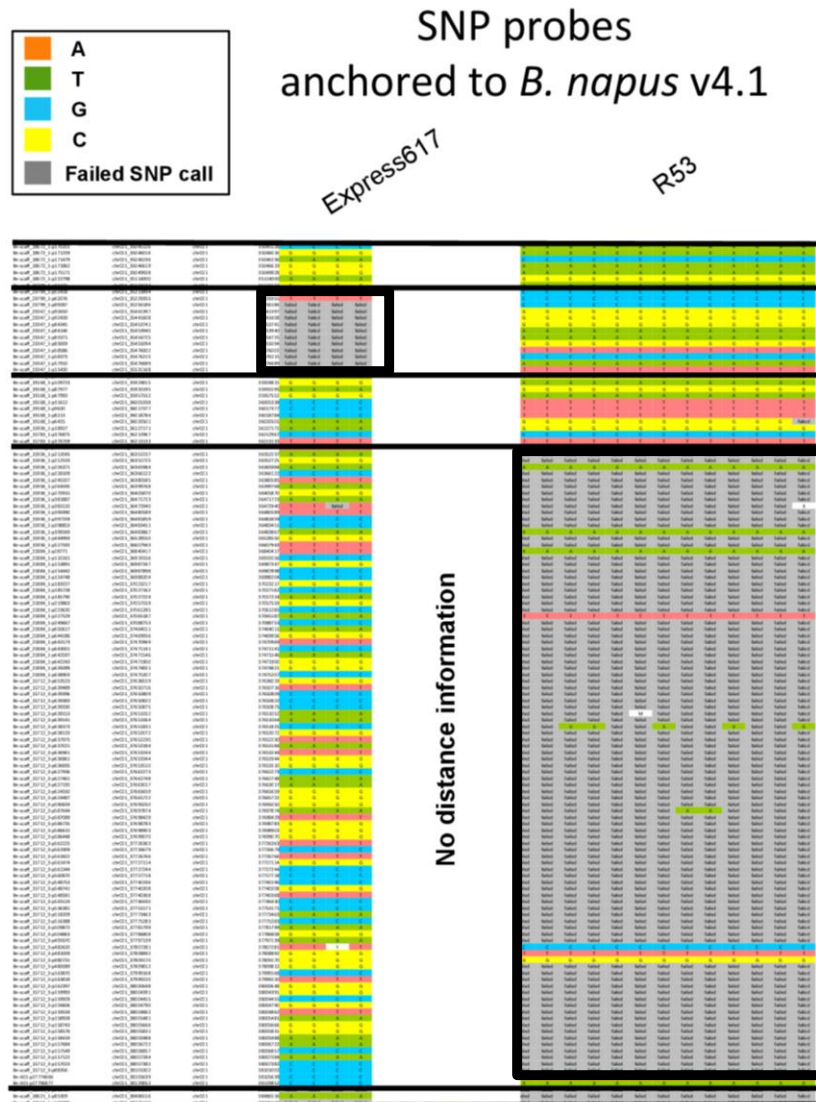
²IGEPP, INRA, Agrocampus Ouest, Université de Rennes 1, Le Rheu, France

³Agriculture and Agri-Food Canada, Saskatoon, Canada

Summary

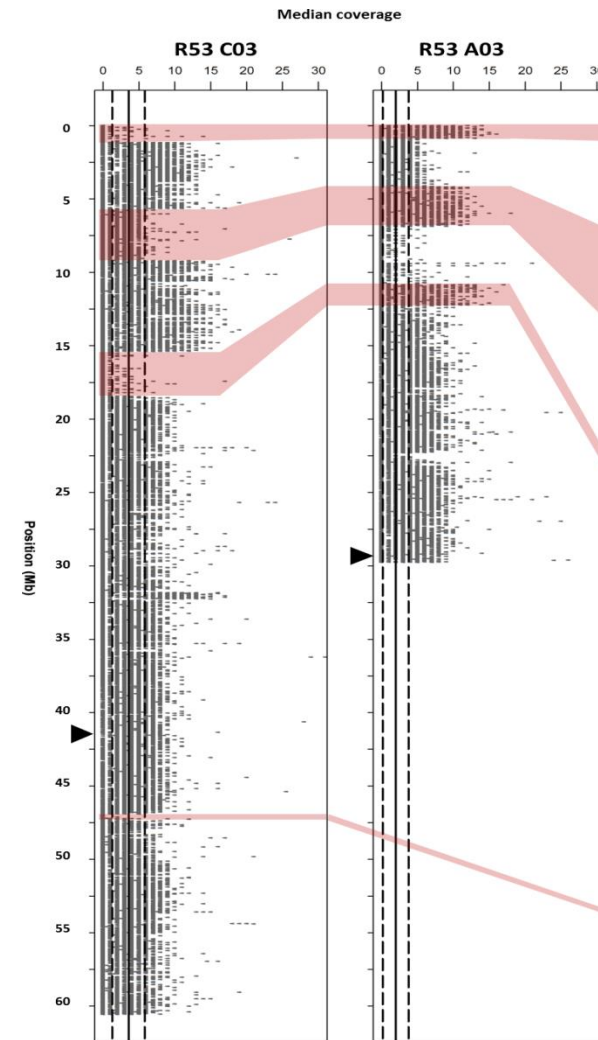
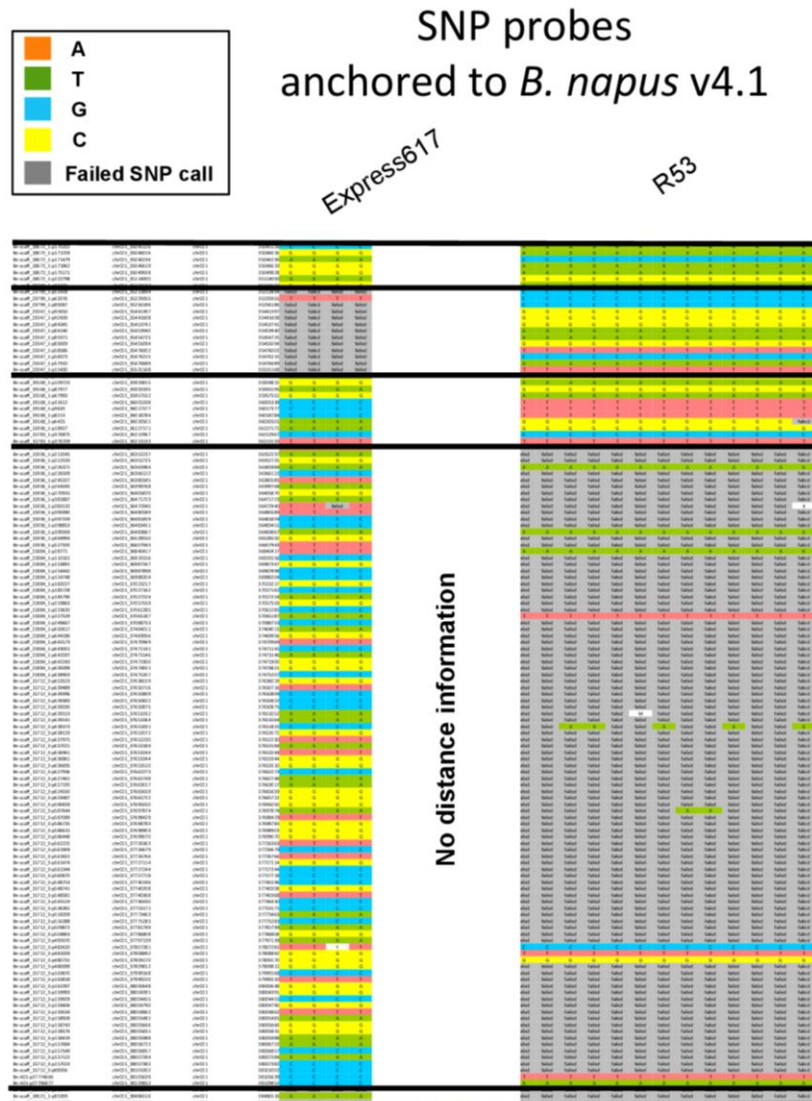
Chromosomal rearrangements arising during polyploidization are an important source of genetic variation in the recent allopolyploid crop *Brassica napus*. Exchanges among homoeologous chromosomes, interhomoeologue pairing, and deletions without reciprocal translocations in both natural *B. napus* and synthetic varieties may induce gene copy number changes.

SNP array or Illumina sequencing-based SV detection approaches

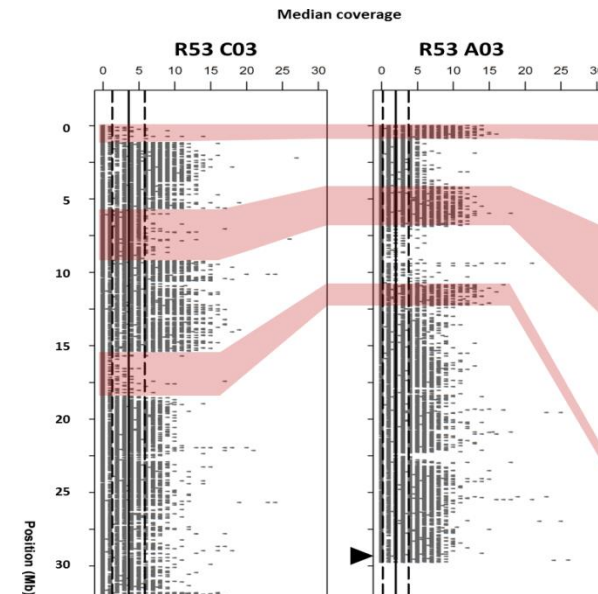
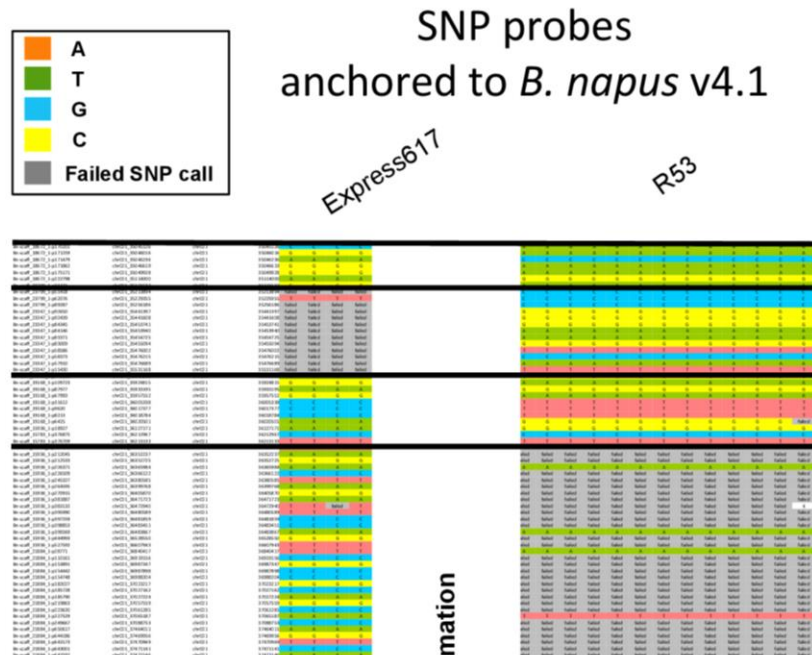


SNP arrays can detect large deletions but fail to pick up smaller ones due to large physical distance of SNP probes

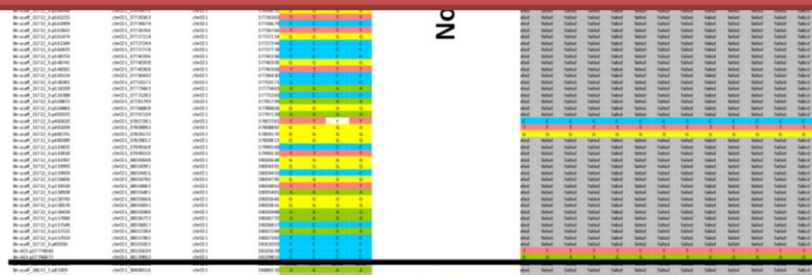
SNP array or Illumina sequencing-based SV detection approaches



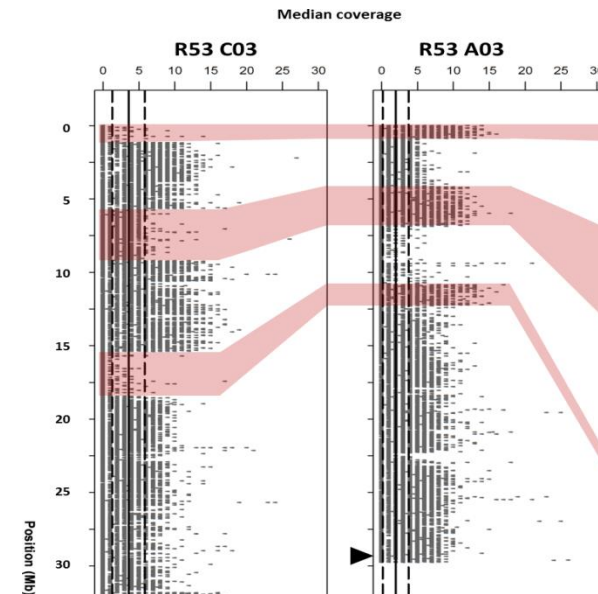
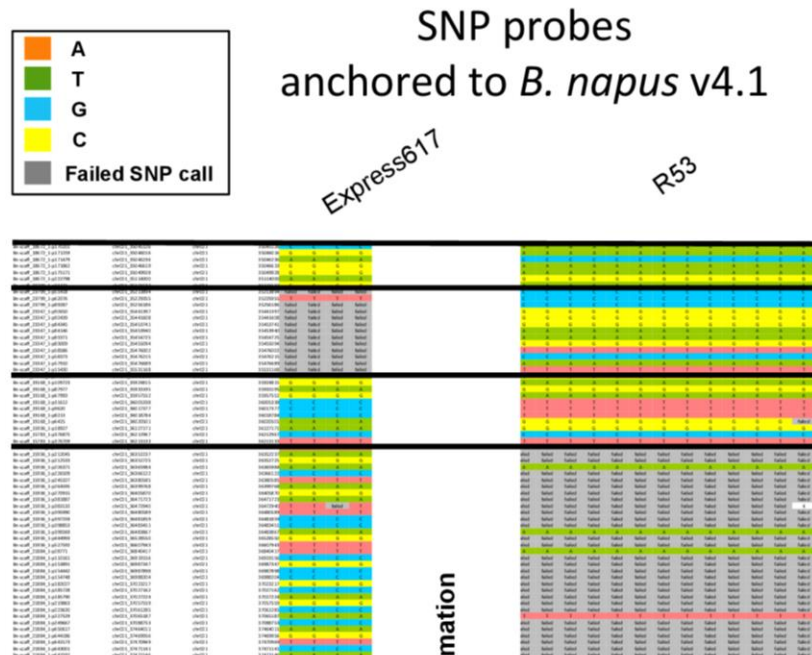
SNP array or Illumina sequencing-based SV detection approaches



Can detect large SV but not enough resolution to detect gene size SV due to short read length



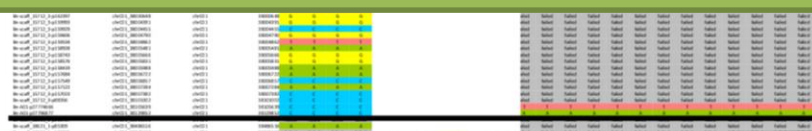
SNP array or Illumina sequencing-based SV detection approaches



Can detect large SV but not enough resolution to detect gene size SV due to short read length



Solution ?? Generate longer reads



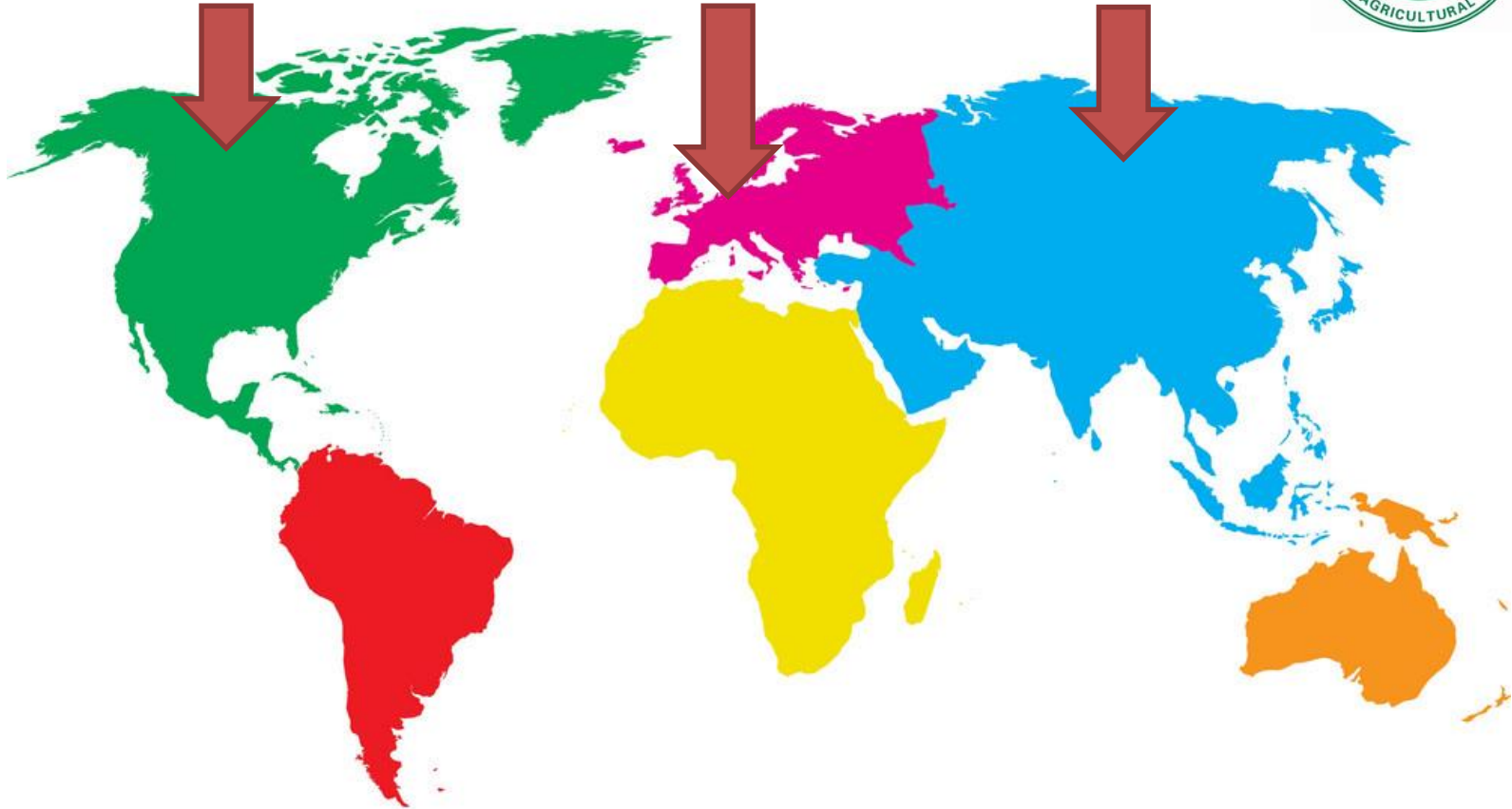
Long read sequencing for *Brassica napus*



Spring

Winter

Semi-winter



PACBIO®



Oxford
NANOPORE
Technologies

Outline

Introduction

Genome-wide identification of
small intra-genic SV

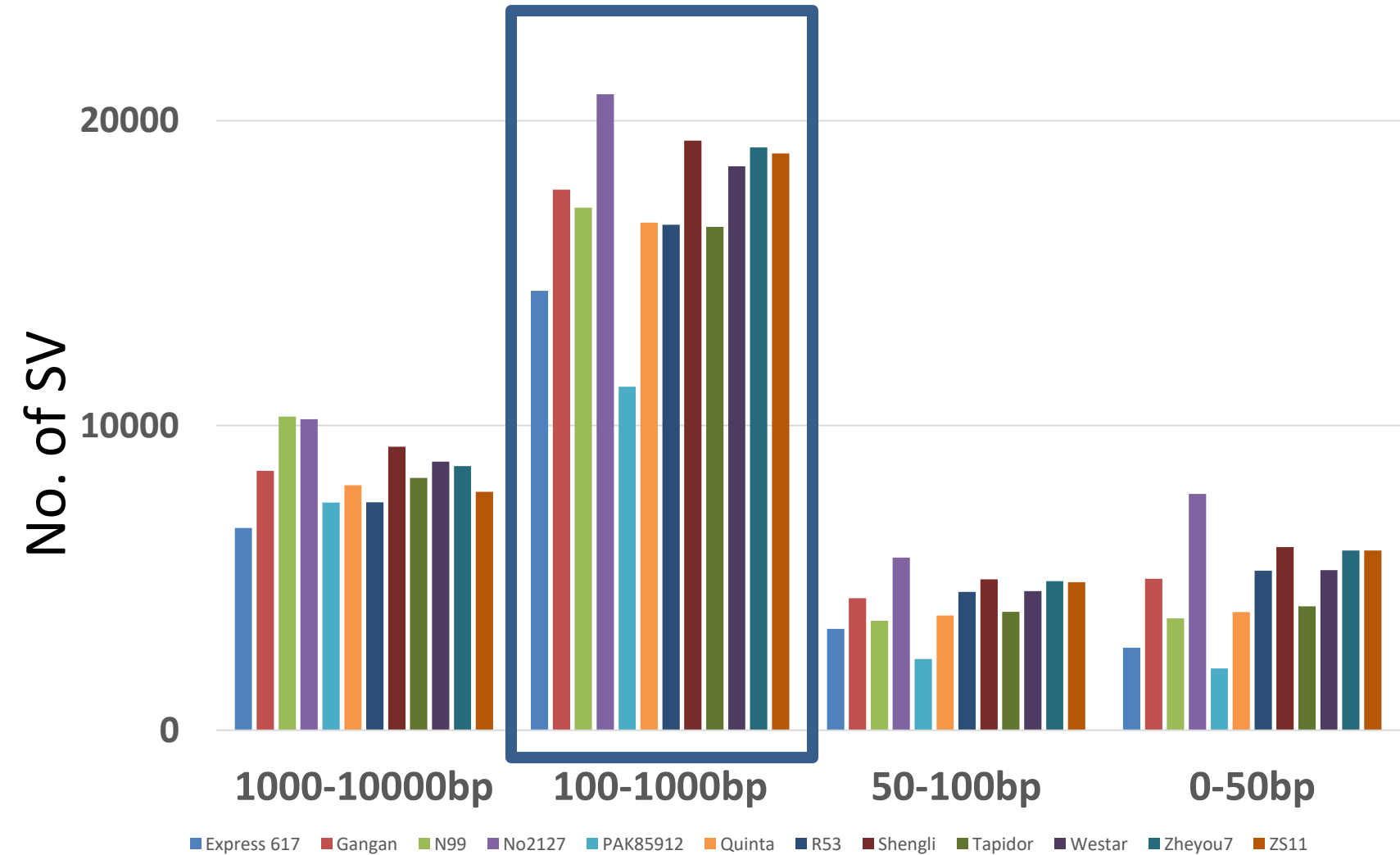
SV associated with eco-geographical
adaptations

SV contributing to disease
resistance

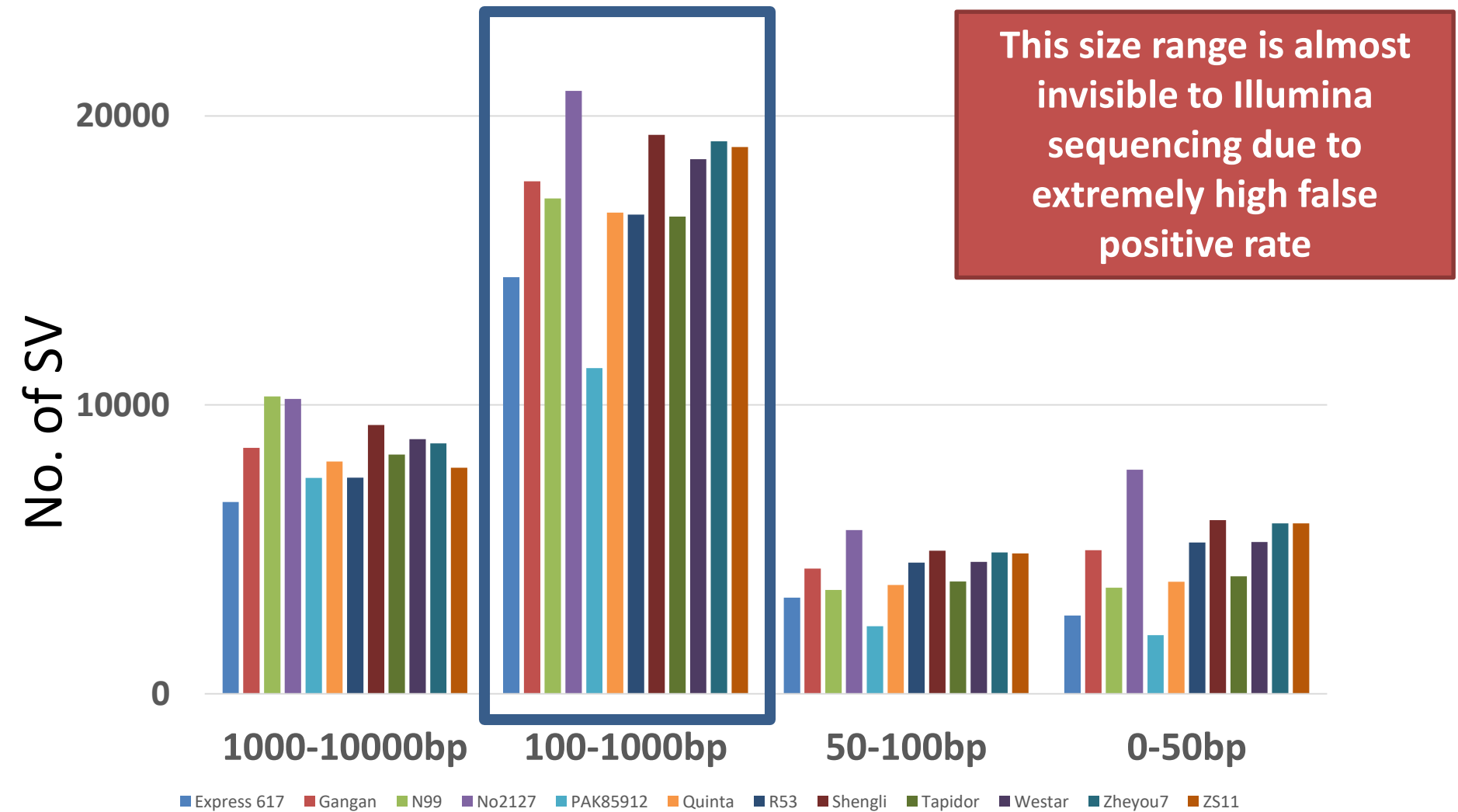
Long read sequencing for *Brassica napus*

Genotype	Data type	Type	N50 for raw reads	Quality filtered SV
Express 617	ONT	Winter	10,756	27,107
Quinta	Pacbio		14,192	32,349
Tapidor	Pacbio		14,448	32,757
ZS11	Pacbio	Semi-winter	10,552	37,496
Zheyu7	Pacbio		12,370	38,590
Gangan	Pacbio		14,064	35,560
Shengli	Pacbio		13,828	39,622
PAK85912	ONT	Spring	28,916	23,177
N99	ONT		27,139	34,848
Westar	Pacbio		13,810	37,138
R53	ONT	Winter Synthetic	11,253	33,851
No2127	Pacbio	Spring Synthetic	15,369	44,516

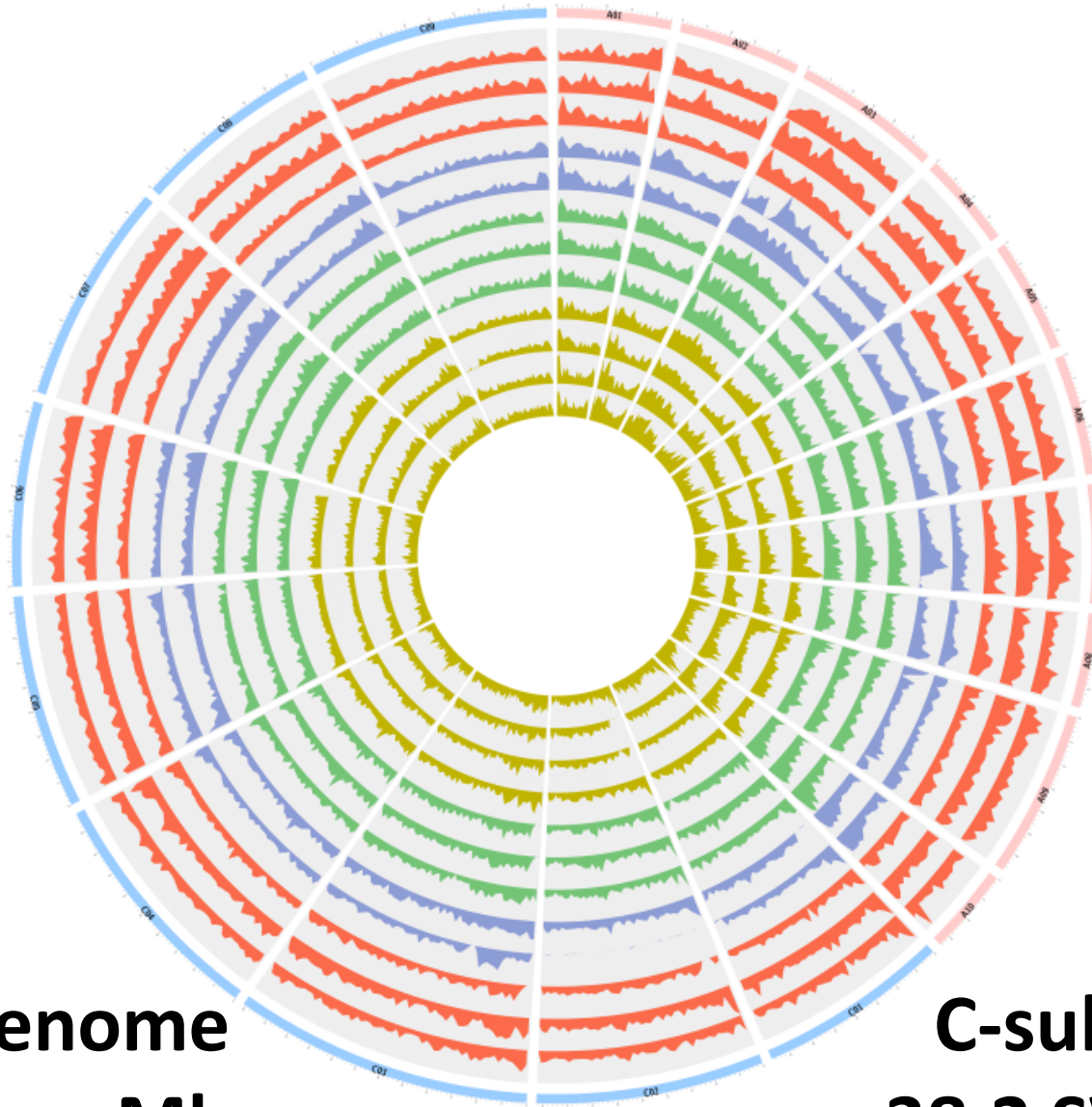
Most SV are between 100 to 1000 bp



Most SV are between 100 to 1000 bp



Size and distribution of SV in rapeseed genome



Winter

Synthetic

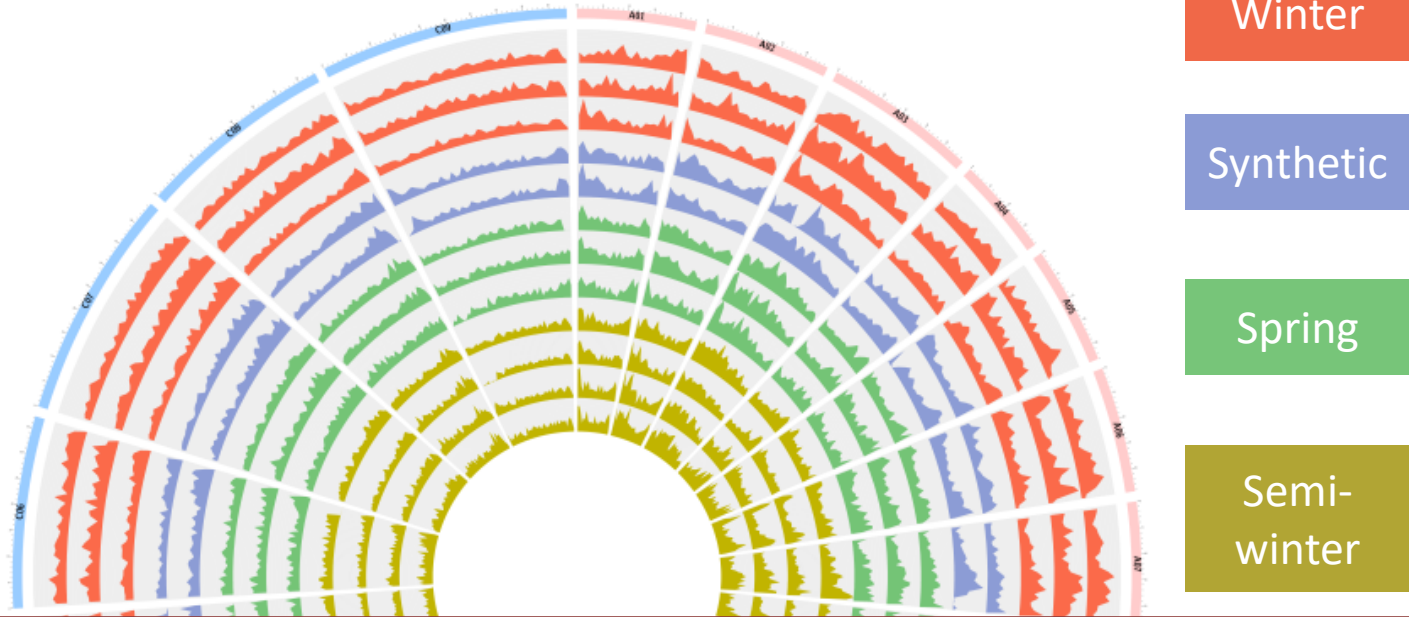
Spring

Semi-
winter

A-subgenome
57.5 SV per Mbp

C-subgenome
38.2 SV per Mbp

Size and distribution of SV in rapeseed genome

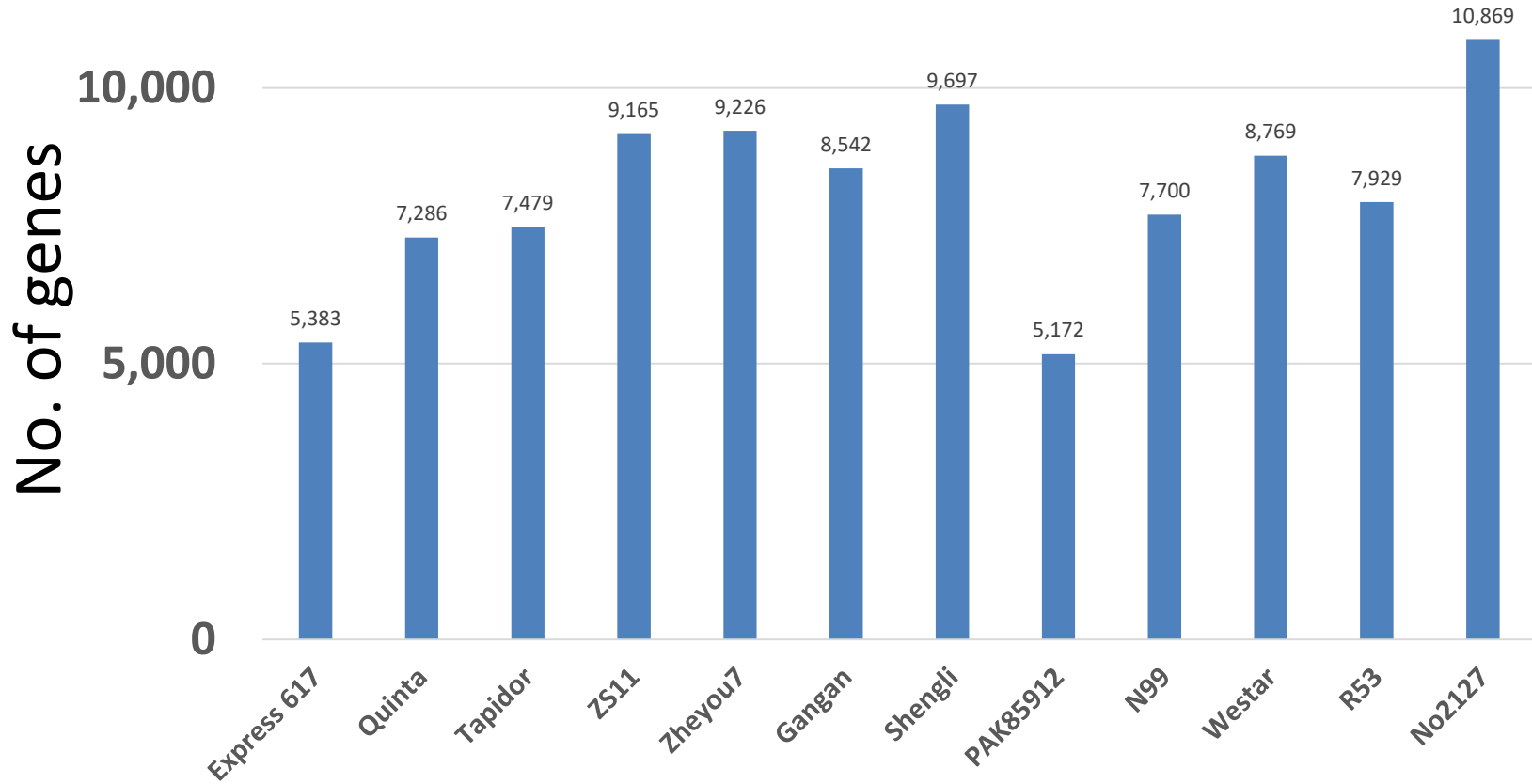


Lower SV frequency in the C-subgenome compared to A-subgenome

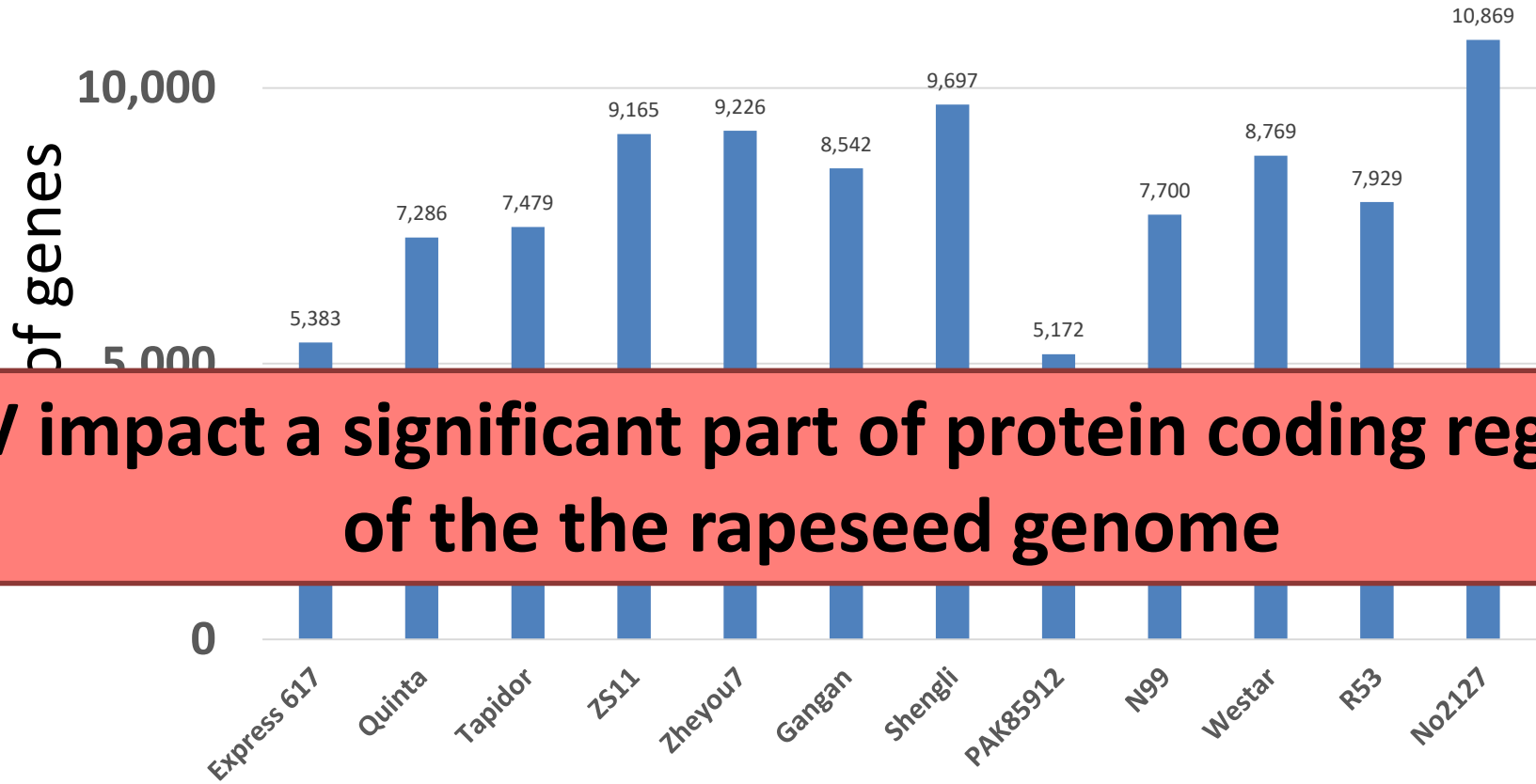
**A-subgenome
57.5 SV per Mbp**

**C-subgenome
38.2 SV per Mbp**

10 percent genes effected by small scale SV



10 percent genes effected by small scale SV



SV impact a significant part of protein coding regions of the the rapeseed genome

Outline

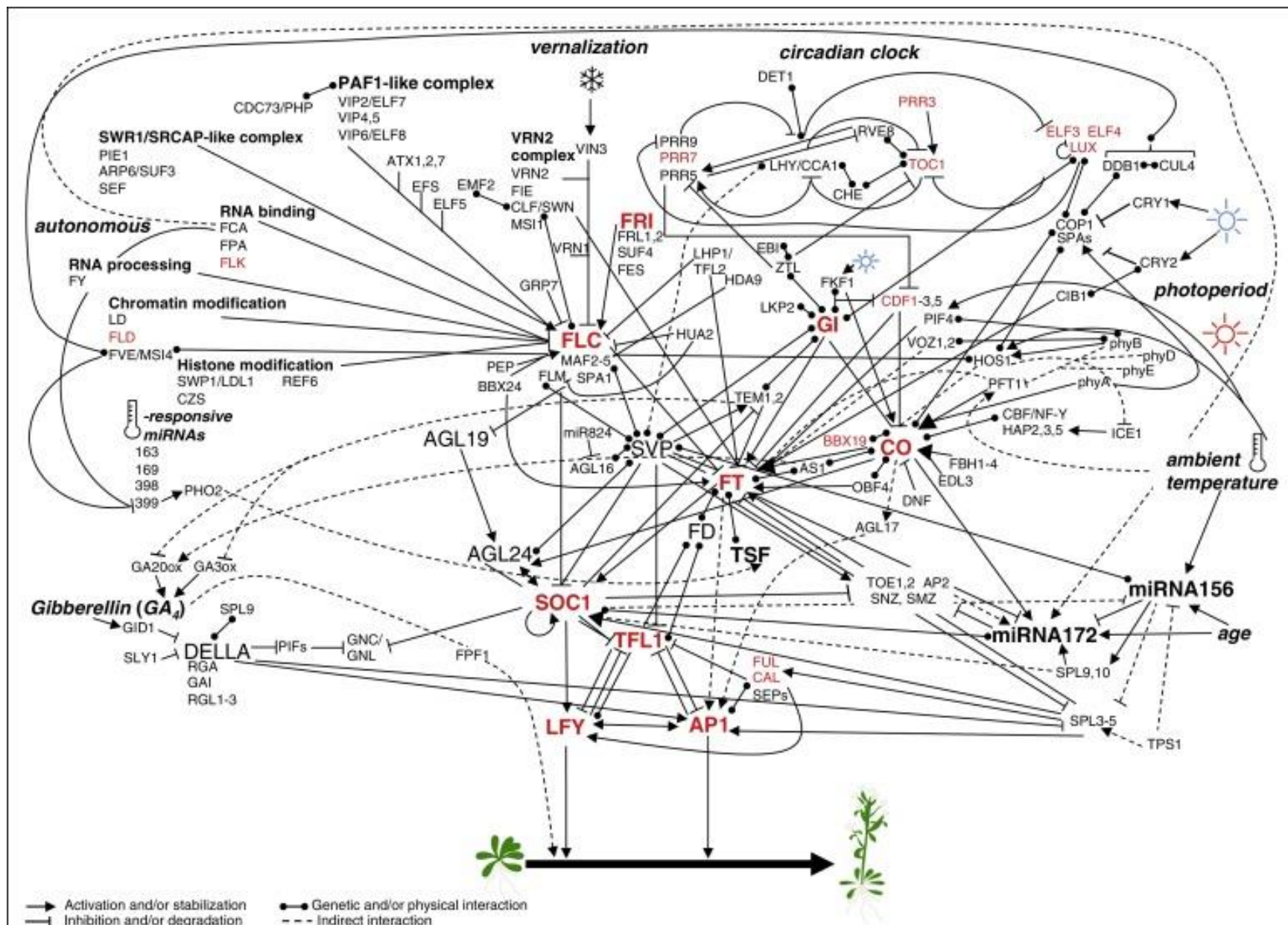
Introduction

Genome-wide identification of
small intra-genic SV

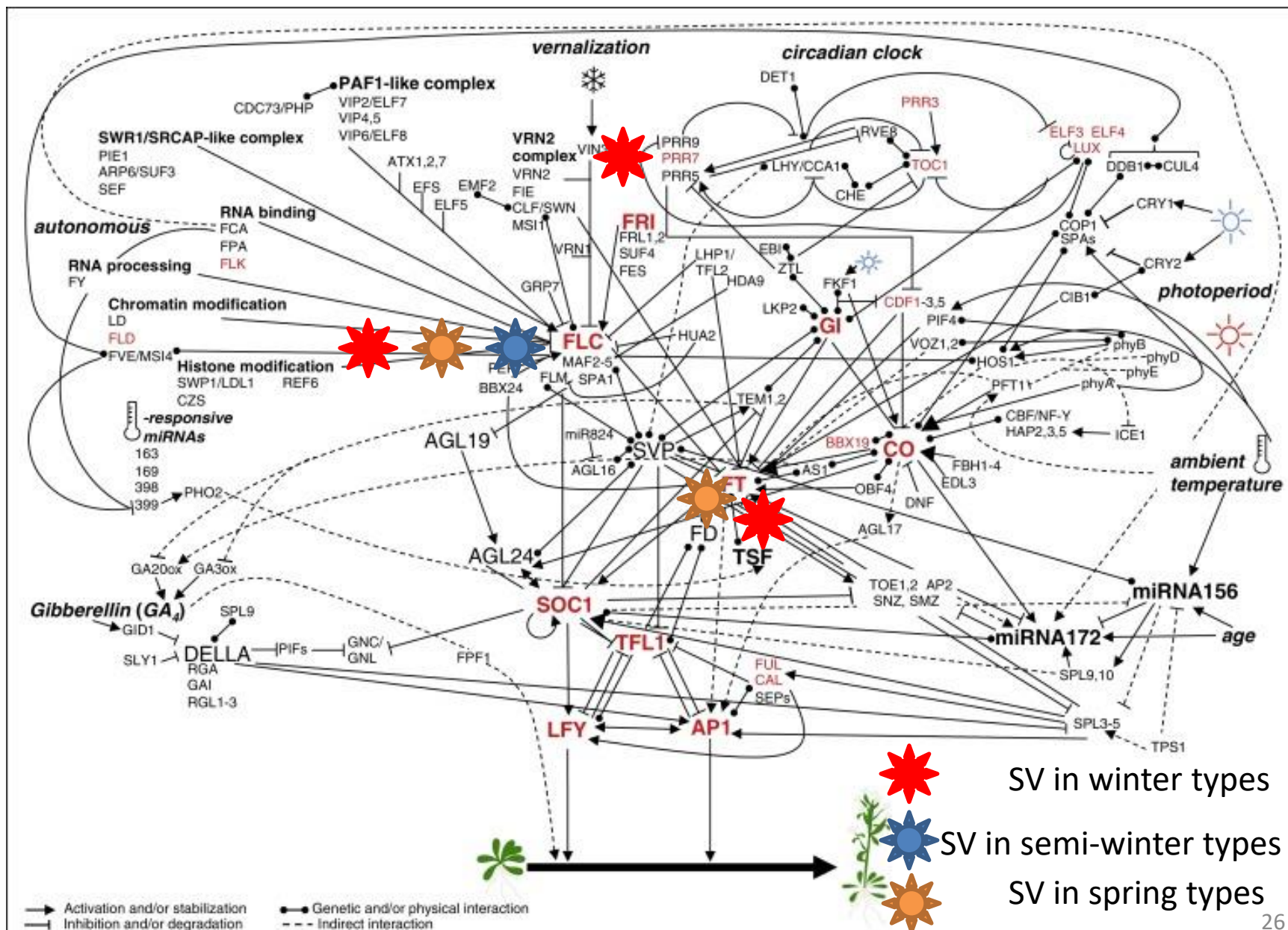
SV associated with eco-geographical
adaptations

SV contributing to disease
resistance

SV in flowering time pathway



SV in flowering time pathway



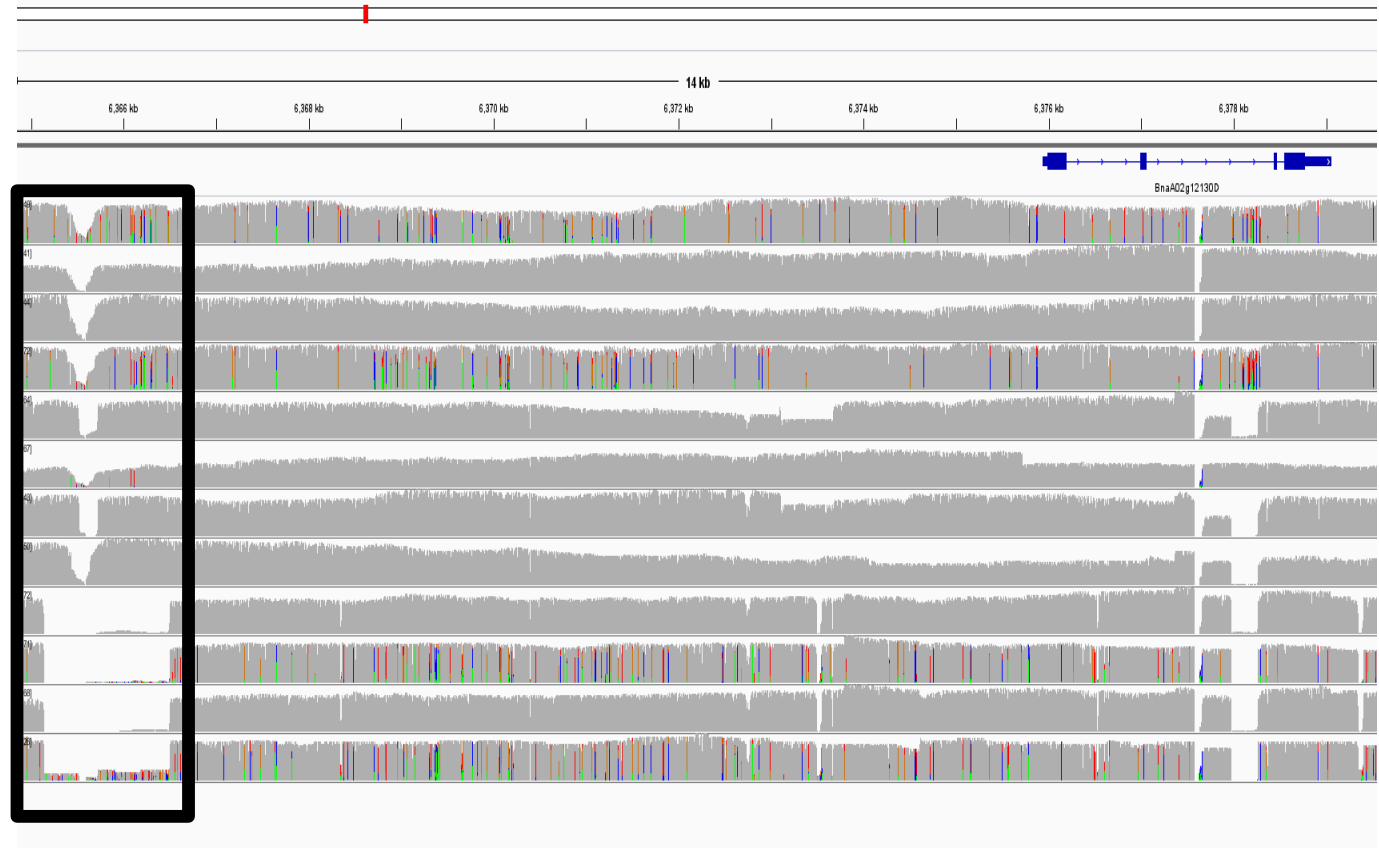
Deletions in Flowering Time Locus T

Winter

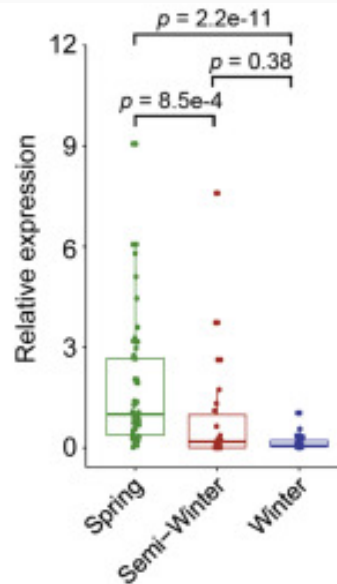
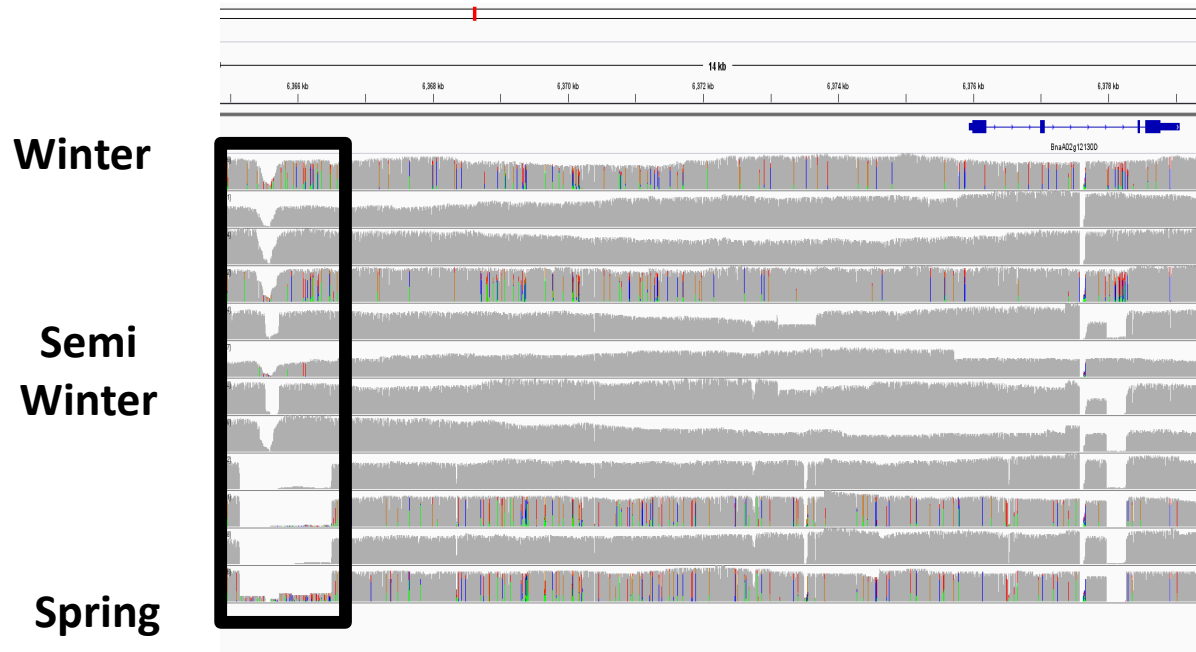
Semi

Winter

Spring



Deletions in Flowering Time Locus T



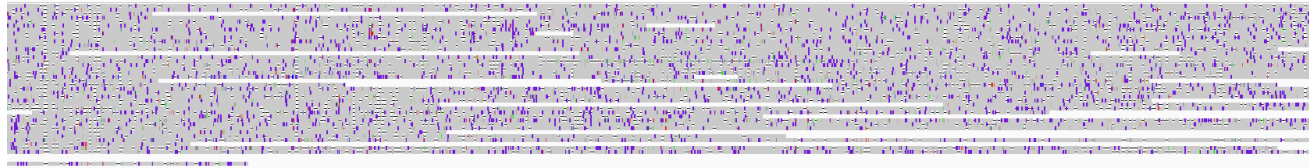
90bp insertion in *BnVIN3*

Express617

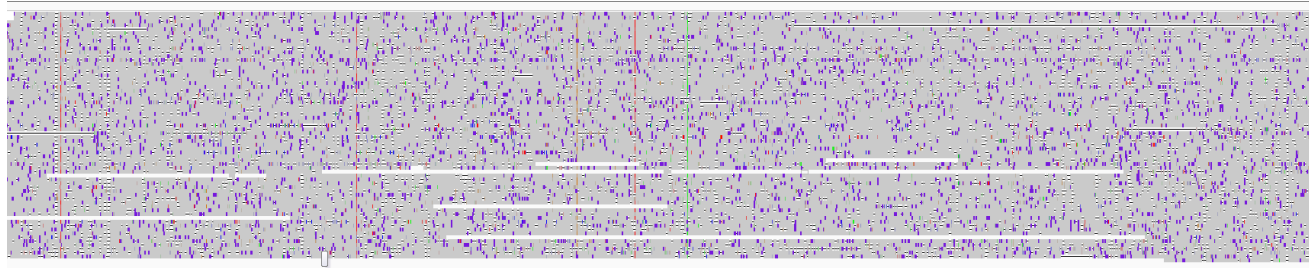


90bp insertion
BnVIN3 exclusively
present in one of
the
winter line

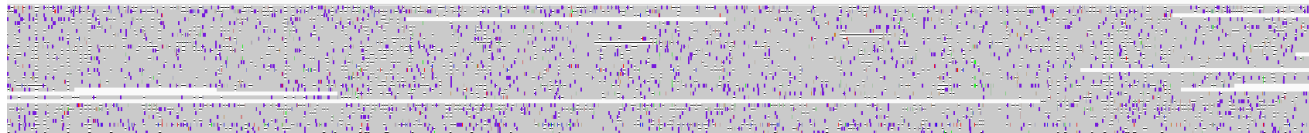
R53



N99

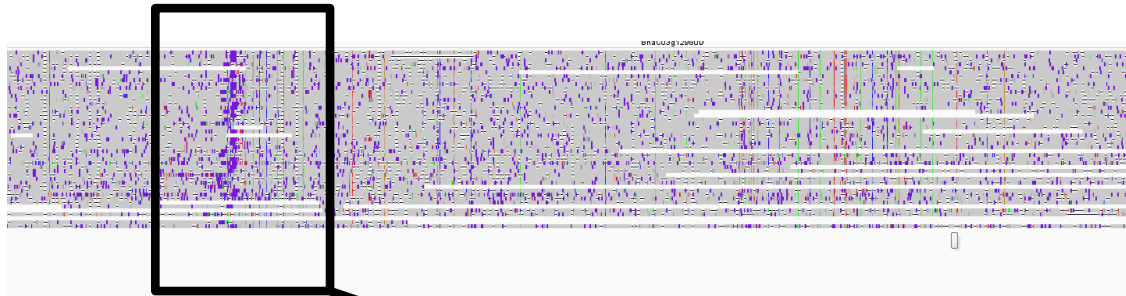


PAK85912



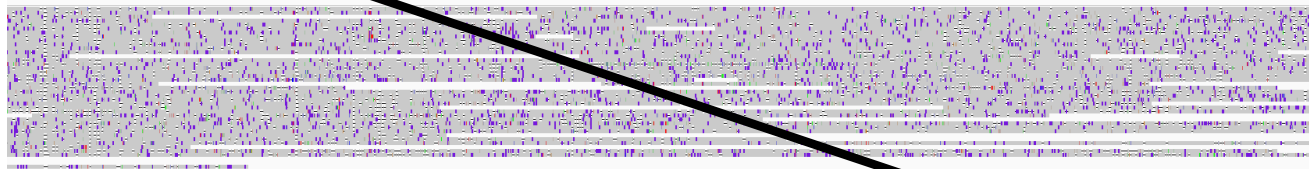
90bp insertion in *BnVIN3*

Express617

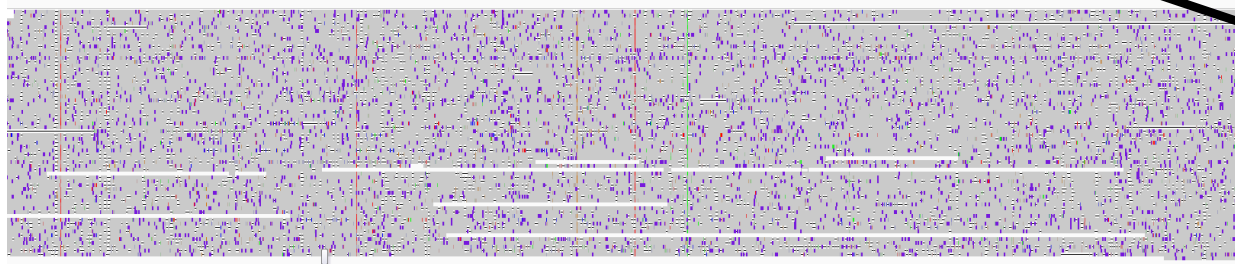


90bp insertion
BnVIN3 exclusively
present in one of
the
winter line

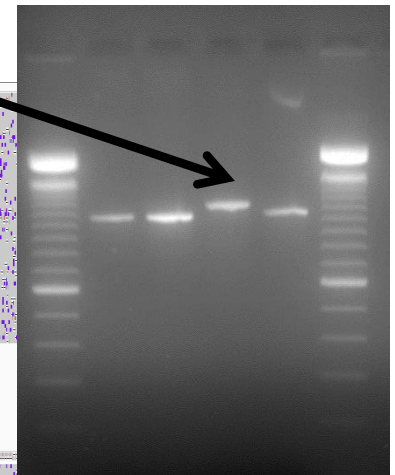
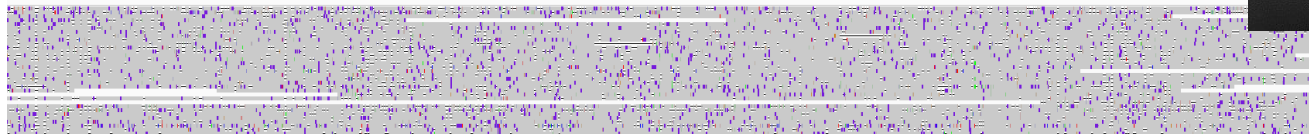
R53



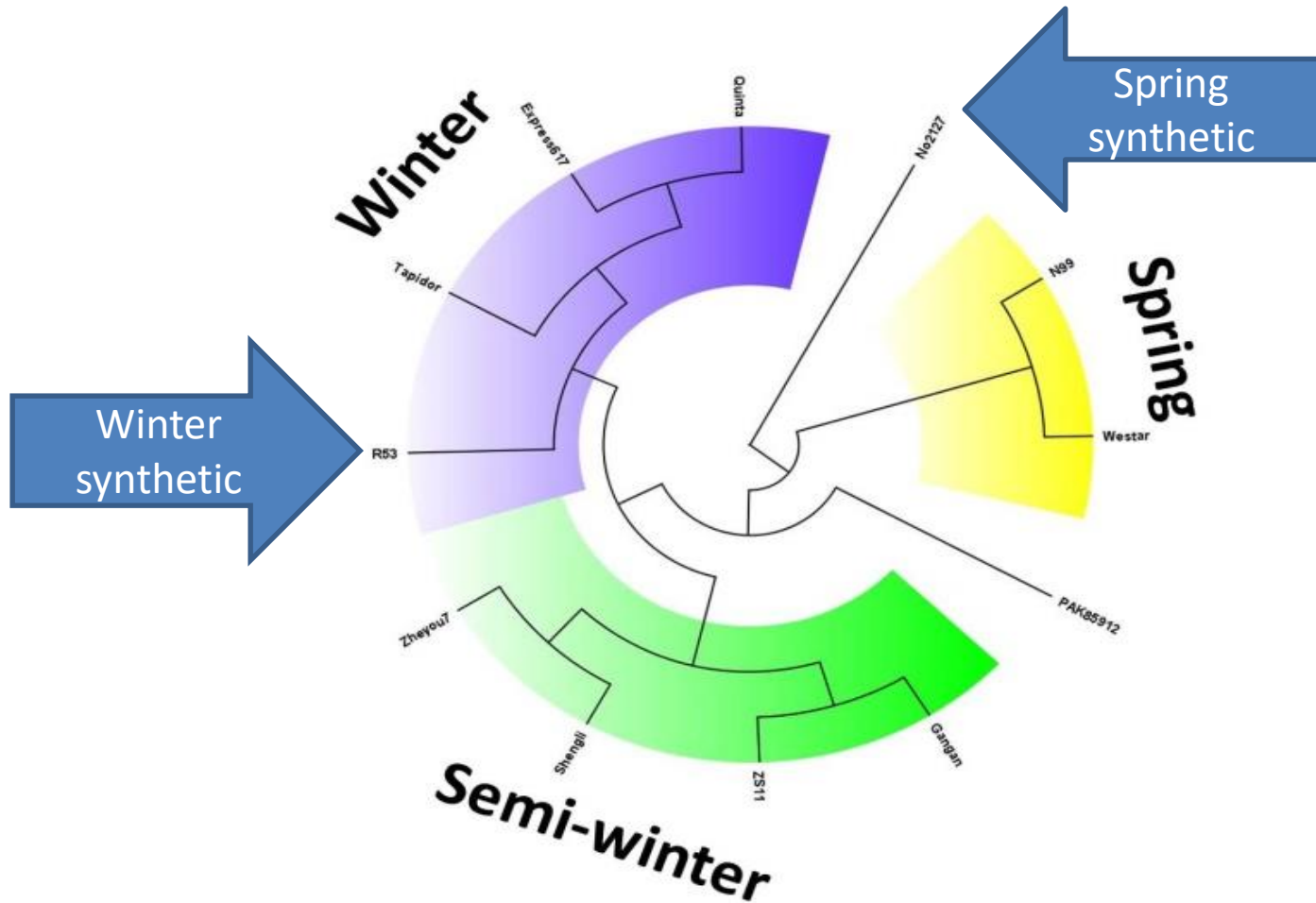
N99



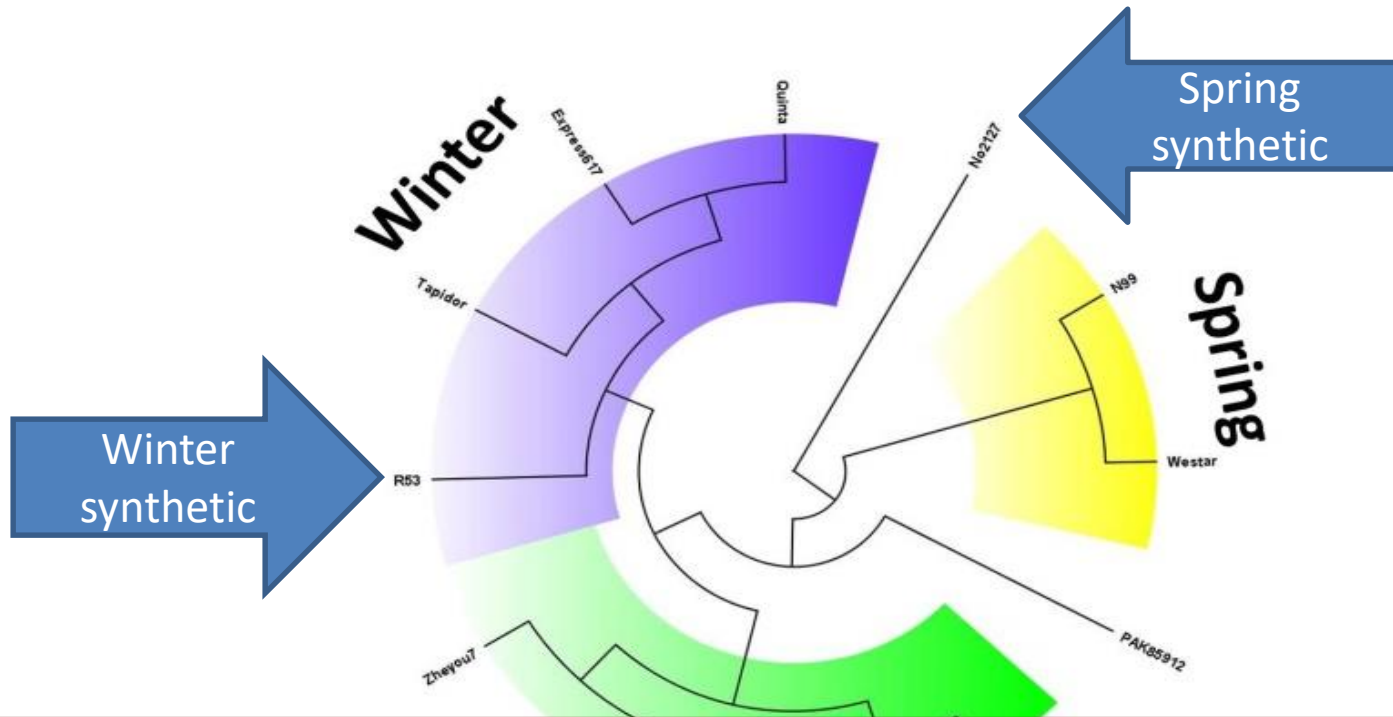
PAK85912



Clustering of gene scale SV



Clustering of gene scale SV



Gene scale SV events originating during or immediately after allo-polyploidisation might rapidly confer ecogeographical adaptation

Outline

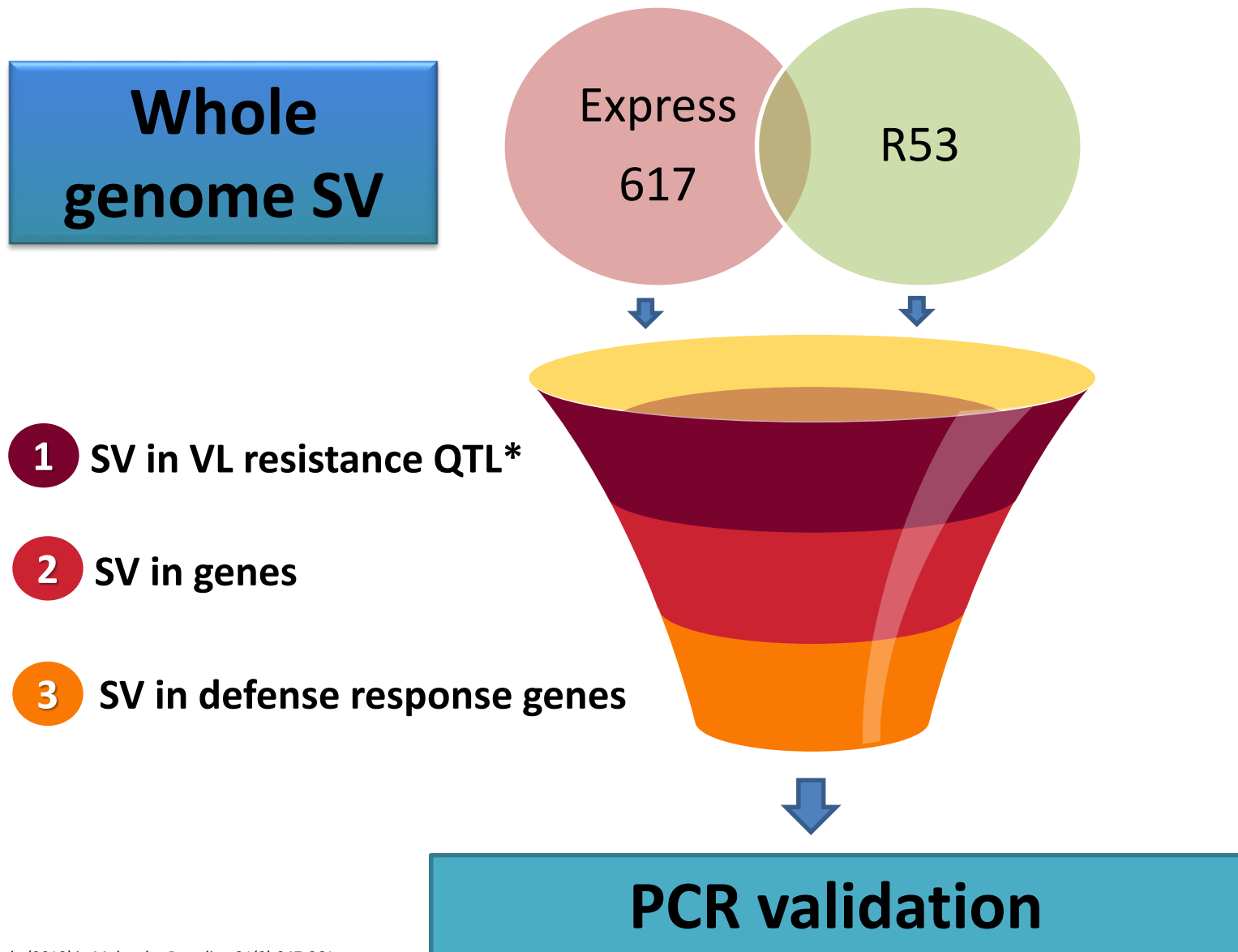
Introduction

Genome-wide identification of
small intra-genic SV

SV associated with eco-geographical
adaptations

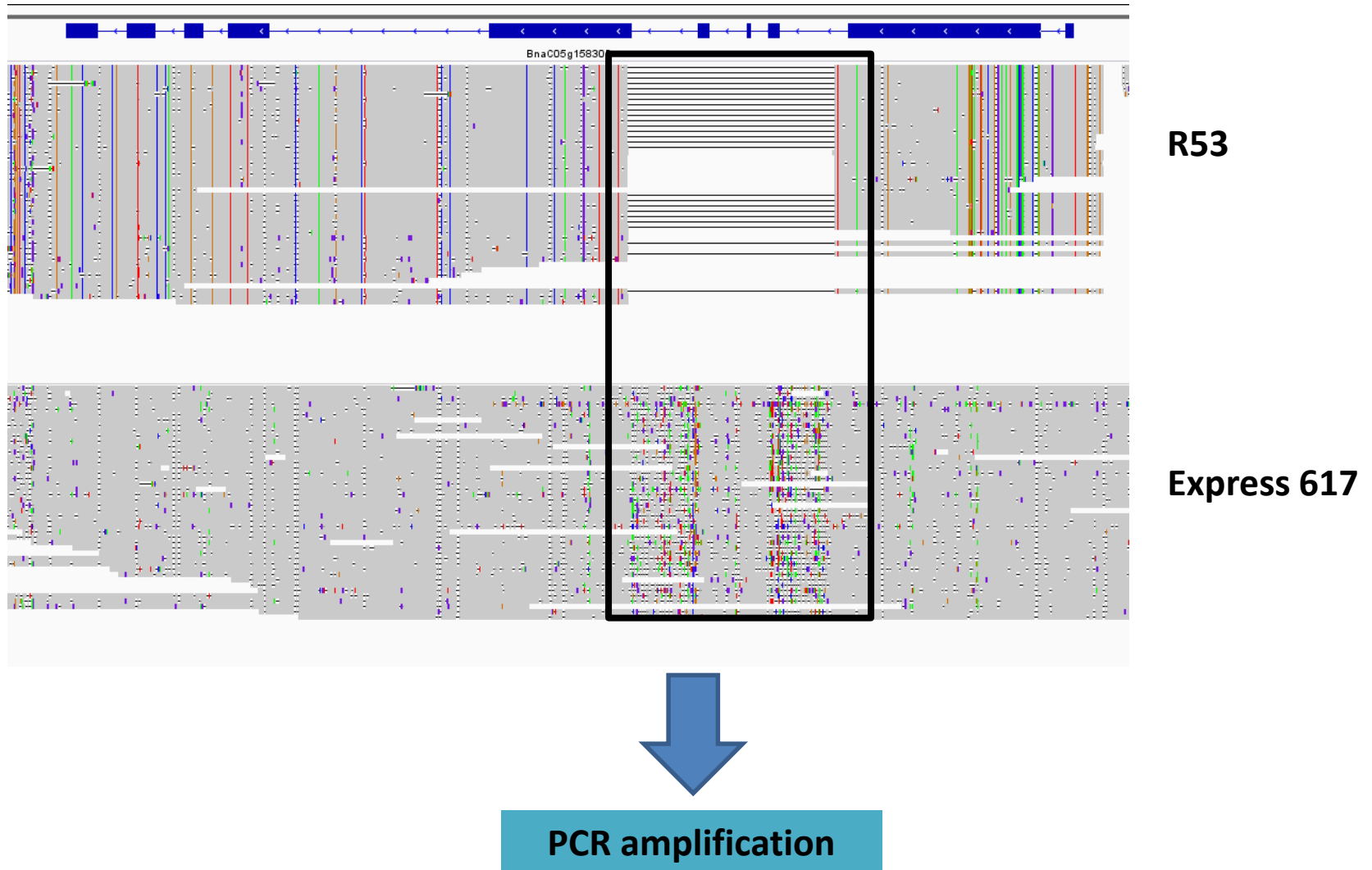
SV contributing to disease
resistance

SV contributing to disease resistance: Example *Verticillium* resistance



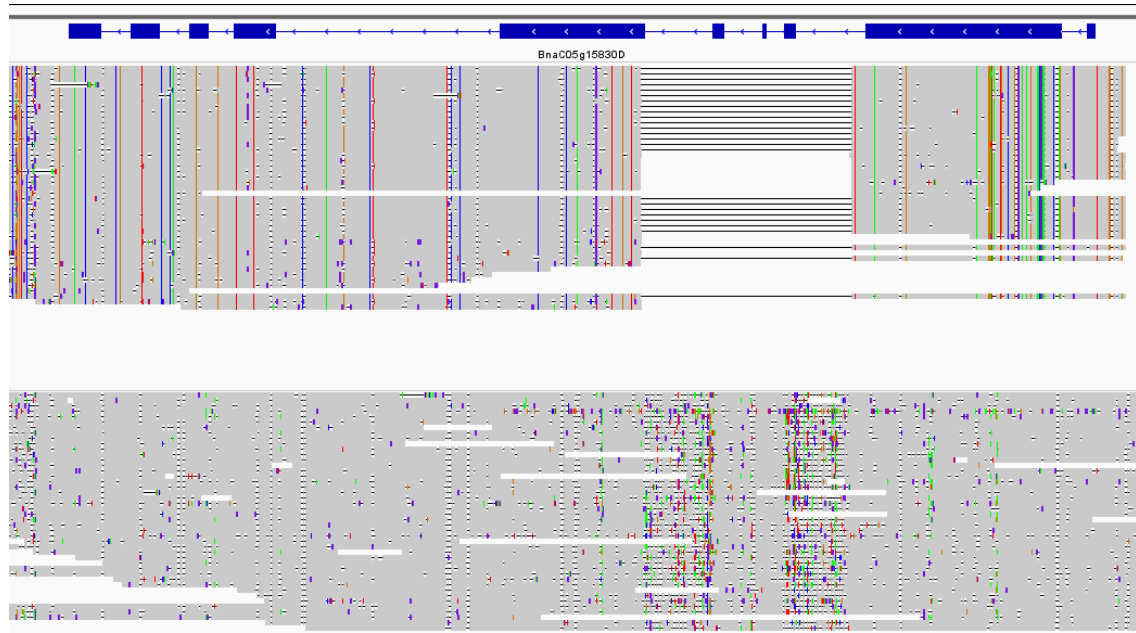
Deletion in VL resistance QTL

725bp deletion in 4-Coumarate:CoA Ligase gene



Deletion in VL resistance QTL

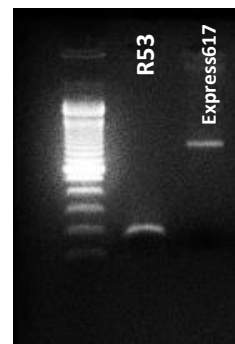
725bp deletion in 4-Coumarate:CoA Ligase gene



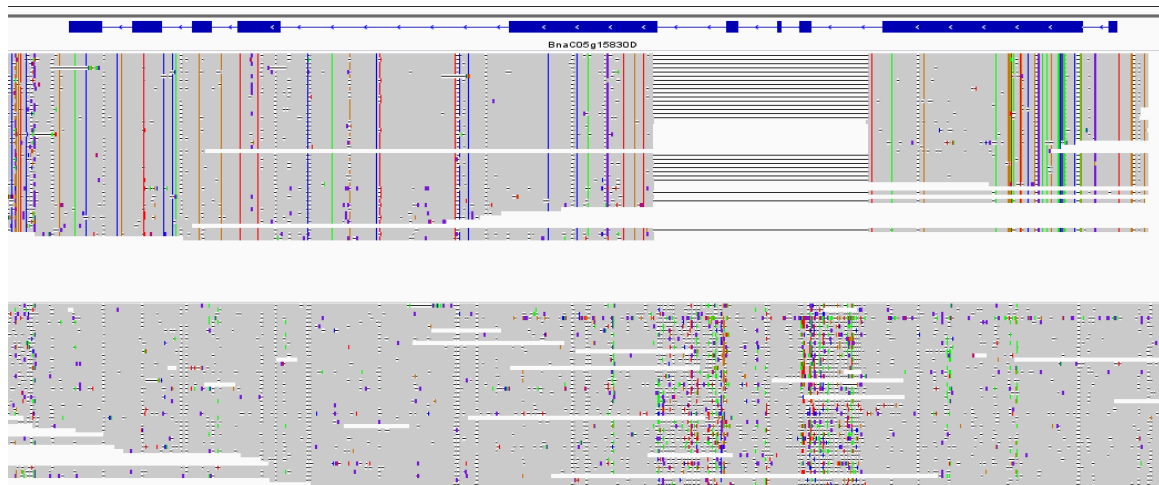
R53

Express 617

PCR amplification



Deletion in VL resistance QTL



R53

Express 617

PCR amplification for 250 lines of ExR53 population

Single marker regression analysis

Experiment no	No. of individuals phenotyped	No. of genotypes sequenced	Adjusted R^2 (%)
1	100	250	19.4
2	32		NS
3	82		17.3
4	98		13.5

Genomic prediction with SV

BMC Genomics

[Home](#) [About](#) [Articles](#) [Submission Guidelines](#)

Research Article | [Open Access](#) | Published: 29 October 2019

Transcriptomic and presence/absence variation in the barley genome assessed from multi-tissue mRNA sequencing and their power to predict phenotypic traits

[Marius Weisweiler](#), [Amaury de Montaigu](#), [David Ries](#), [Mara Pfeifer](#) & [Benjamin Stich](#) 

[BMC Genomics](#) 20, Article number: 787 (2019) | [Cite this article](#)

658 Accesses | 1 Altmetric | [Metrics](#)

Abstract

Background

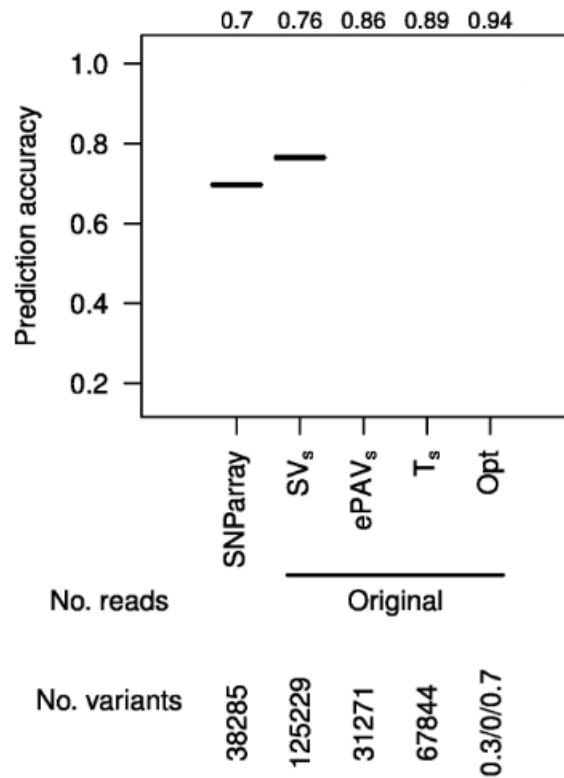
Barley is the world's fourth most cultivated cereal and is an important crop model for genetic studies. One layer of genomic information that remains poorly explored in barley is presence/absence variation (PAV), which has been suggested to contribute to phenotypic variation of agronomic importance in various crops.

Genomic prediction with SV

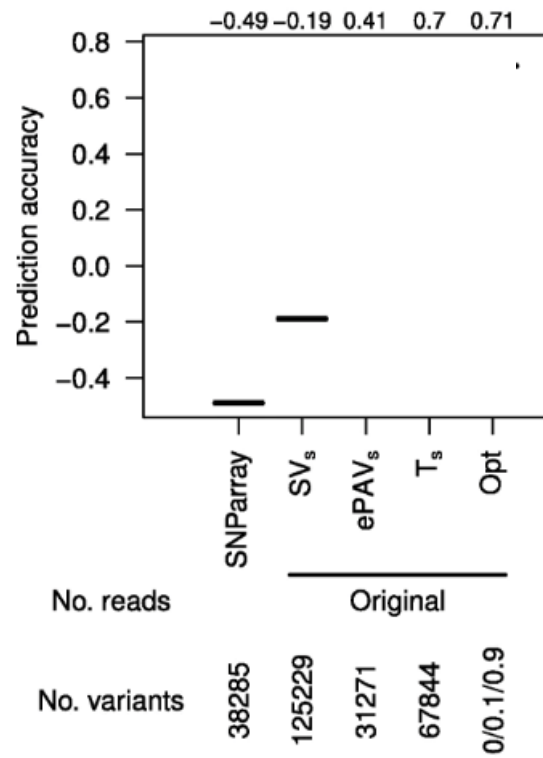
Research Article | Open Access | Published: 29 October 2019

Transcriptomic and presence/absence variation in the barley genome assessed from multi-tissue mRNA sequencing and their power to predict phenotypic traits

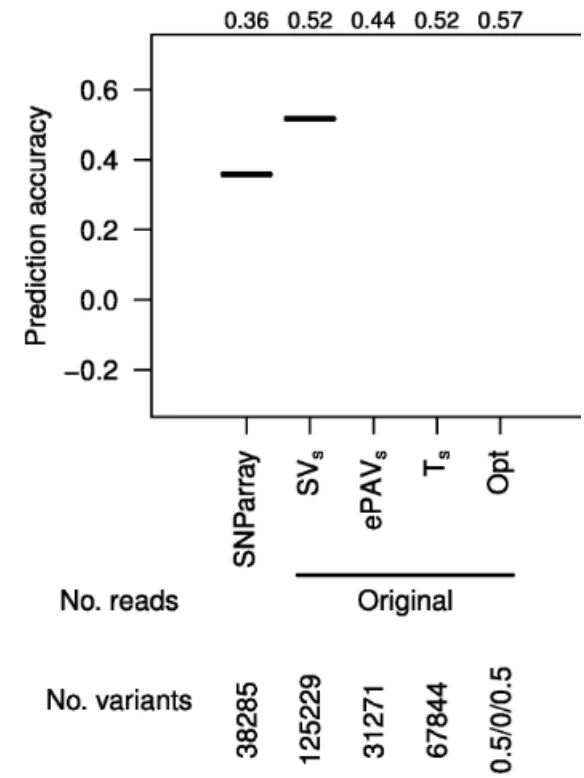
Leaf angle



Heading date



Plant Height

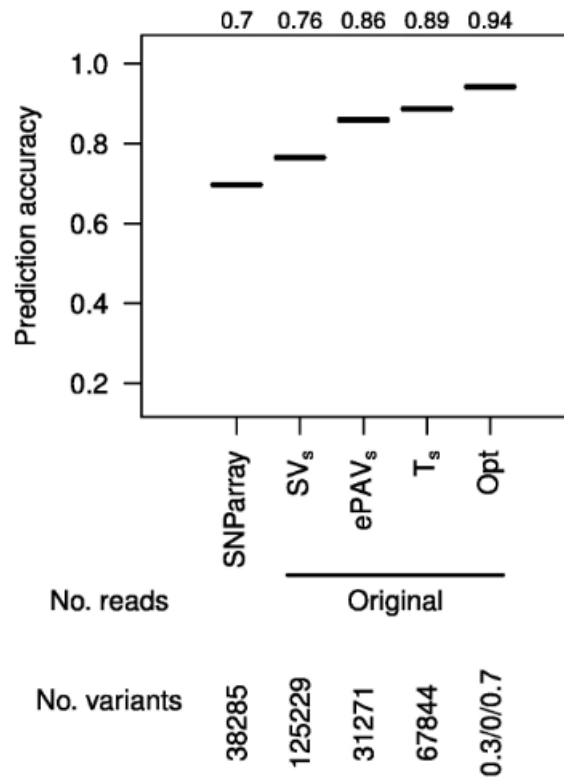


Genomic prediction with SV

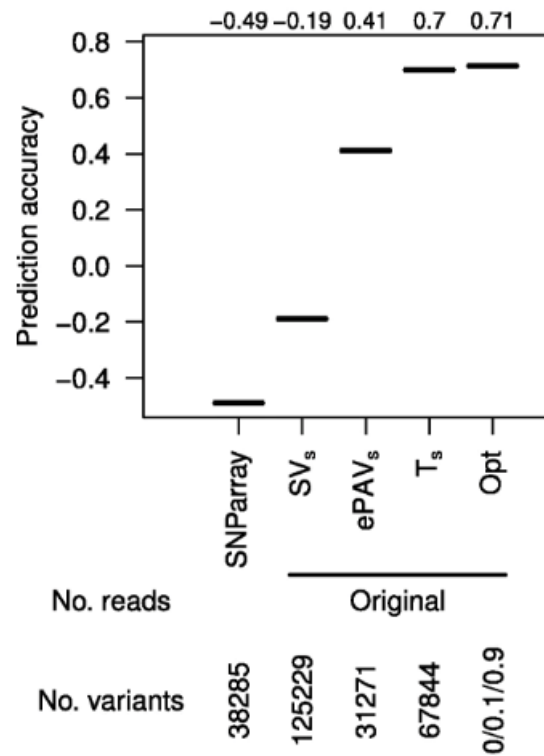
Research Article | Open Access | Published: 29 October 2019

Transcriptomic and presence/absence variation in the barley genome assessed from multi-tissue mRNA sequencing and their power to predict phenotypic traits

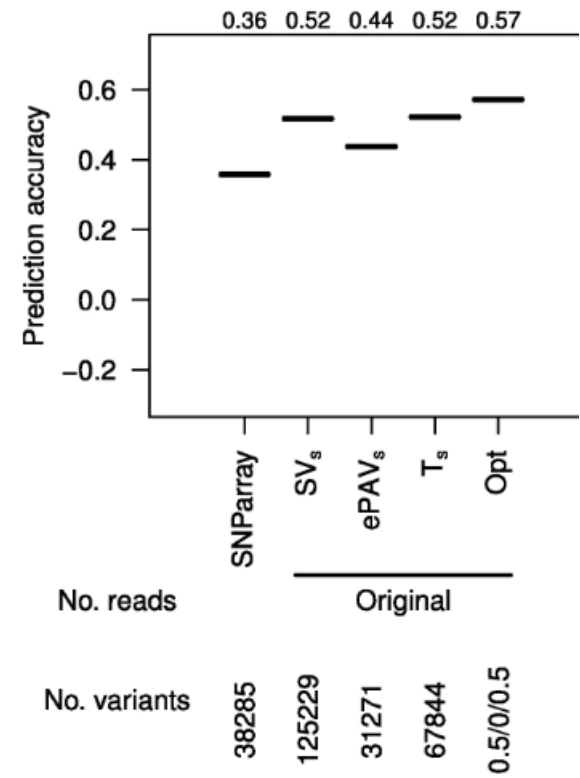
Leaf angle



Heading date



Plant Height

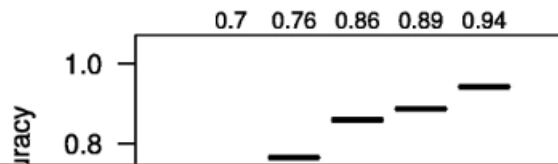


Genomic prediction with SV

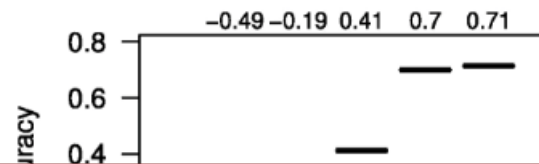
Research Article | Open Access | Published: 29 October 2019

Transcriptomic and presence/absence variation in the barley genome assessed from multi-tissue mRNA sequencing and their power to predict phenotypic traits

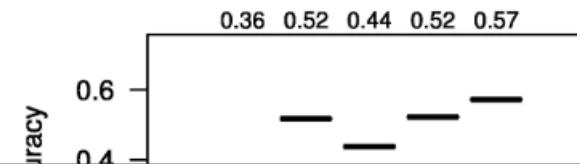
Leaf angle



Heading date



Plant Height



“Expression PAV were not in strong linkage disequilibrium with neighbouring sequence variants (SV), and therefore provided an additional layer of genetic information.”

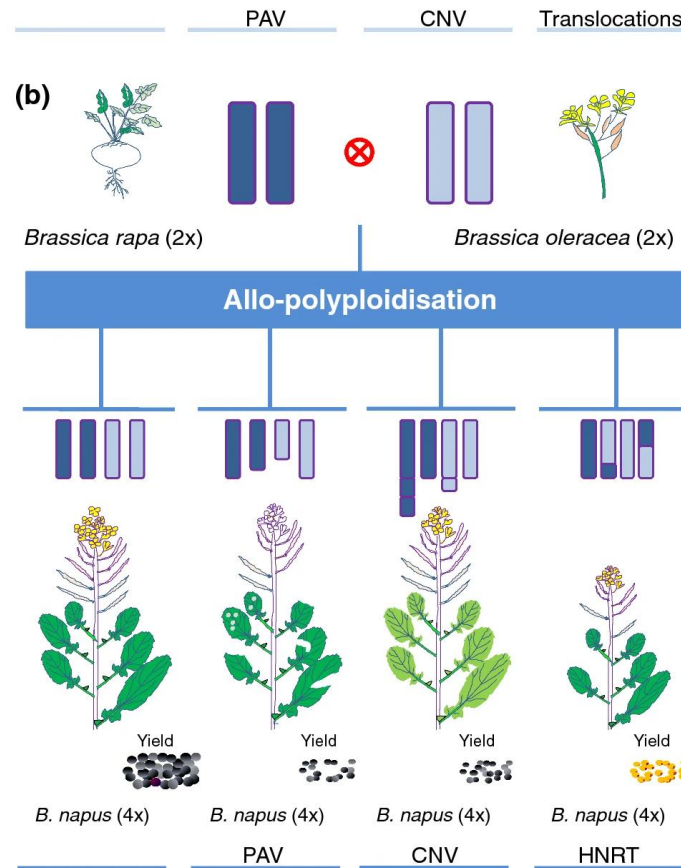
	SNParray	SV	ePAV	T	Op
No. reads		Original			
No. variants	38285	125229	31271	67844	0.3/0/0.7

	SNParray	SV	ePAV	T	Op
No. reads		Original			
No. variants	38285	125229	31271	67844	0/0.1/0.9

	SNParray	SV	ePAV	T	Op
No. reads		Original			
No. variants	38285	125229	31271	67844	0.5/0/0.5

SV were discovered before. Then what's new?

Previous studies describe SV ranging from from few 100's Kbps to Mbps

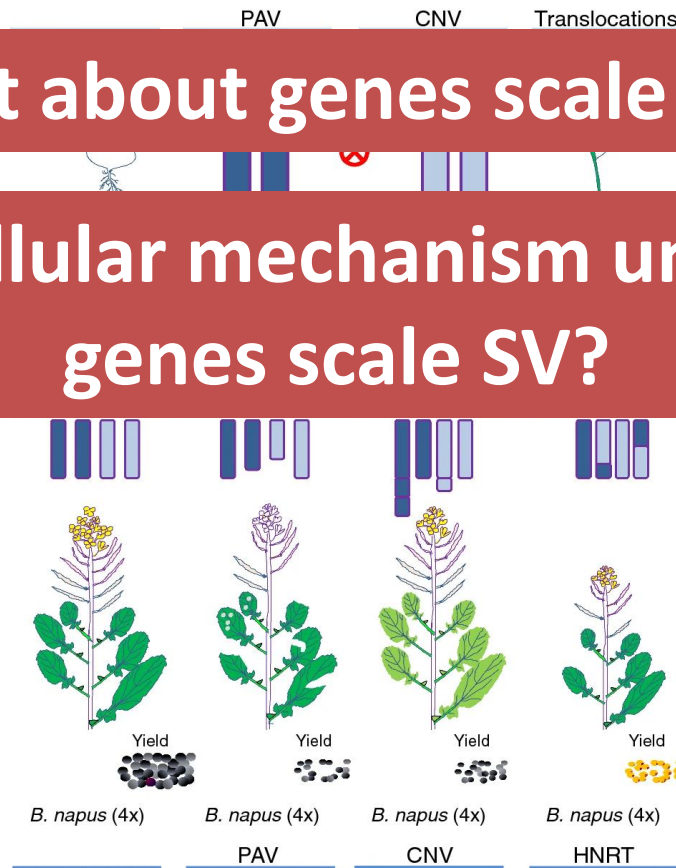


SV were discovered before. Then what's new?

Previous studies describe SV ranging from from few
100's Kbps to Mbps

What about genes scale SV?

What cellular mechanism underlines
genes scale SV?



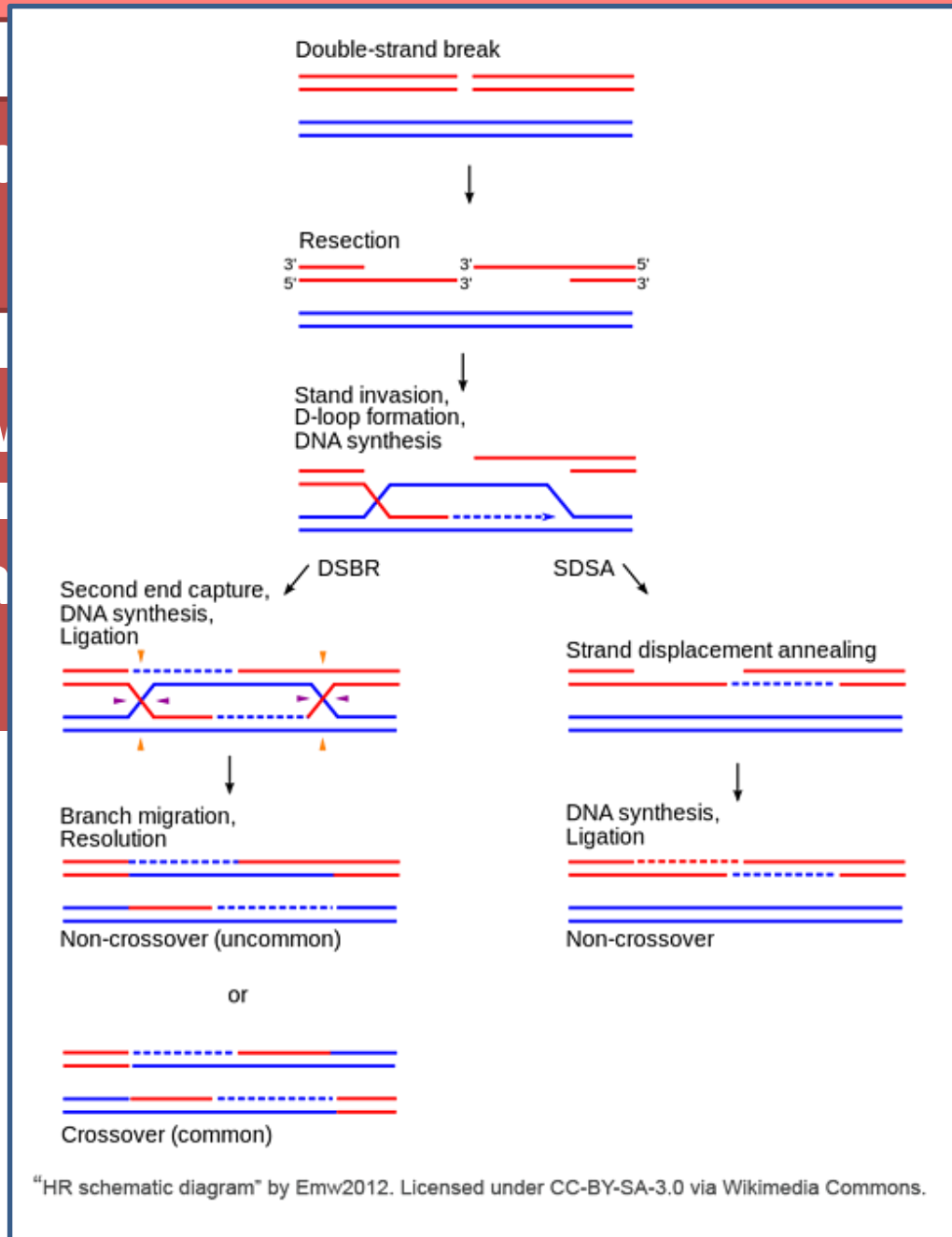
SV were discovered before. Then what's new?

Previous studies

from few

What

nes



Take home messages

- Identified gene scale SV with precise break-points using long range sequencing in 12 diverse *B. napus* genotypes
- More than 50 percent of SV were in between 100 to 1000 bp
- Higher frequency of SV in A subgenome than C subgenome
- 5-10 percent of the total genes contain SV in oilseed rape
- Gene scale SV might contribute to eco-geographical adaptation and disease resistance in *B. napus*
- Including SV in GS models might increase prediction accuracies

Acknowledgements

Huey Tyng Lee
Iulian Gabur
Suriya Tamilselvan-Nattar-Amutha
Christian Obermeier
Sarah V. Schiessl
Jia-Ming Song
Kede Liu
Liang Guo
Isobel A. P. Parkin
Rod J. Snowdon



Agriculture and
Agri-Food Canada



JUSTUS-LIEBIG-



UNIVERSITÄT
GIESSEN











@GiessenSingh

Interested in learning more about SV?

New Results

[Comment on this paper](#)

Long-read sequencing reveals widespread intragenic structural variants in a recent allopolyploid crop plant

 Harmeet Singh Chawla,  HueyTyng Lee, Iulian Gabur, Suriya Tamilselvan-Nattar-Amutha,  Christian Obermeier,  Sarah V. Schiessl, Jia-Ming Song,  Kede Liu,  Liang Guo,  Isobel A. P. Parkin,  Rod J. Snowdon

doi: <https://doi.org/10.1101/2020.01.27.915470>

This article is a preprint and has not been certified by peer review [what does this mean?].

