Differential chromatin accessibility map as a new resource for studying wheat genome function and genotype-to-trait relationships

Eduard Akhunov

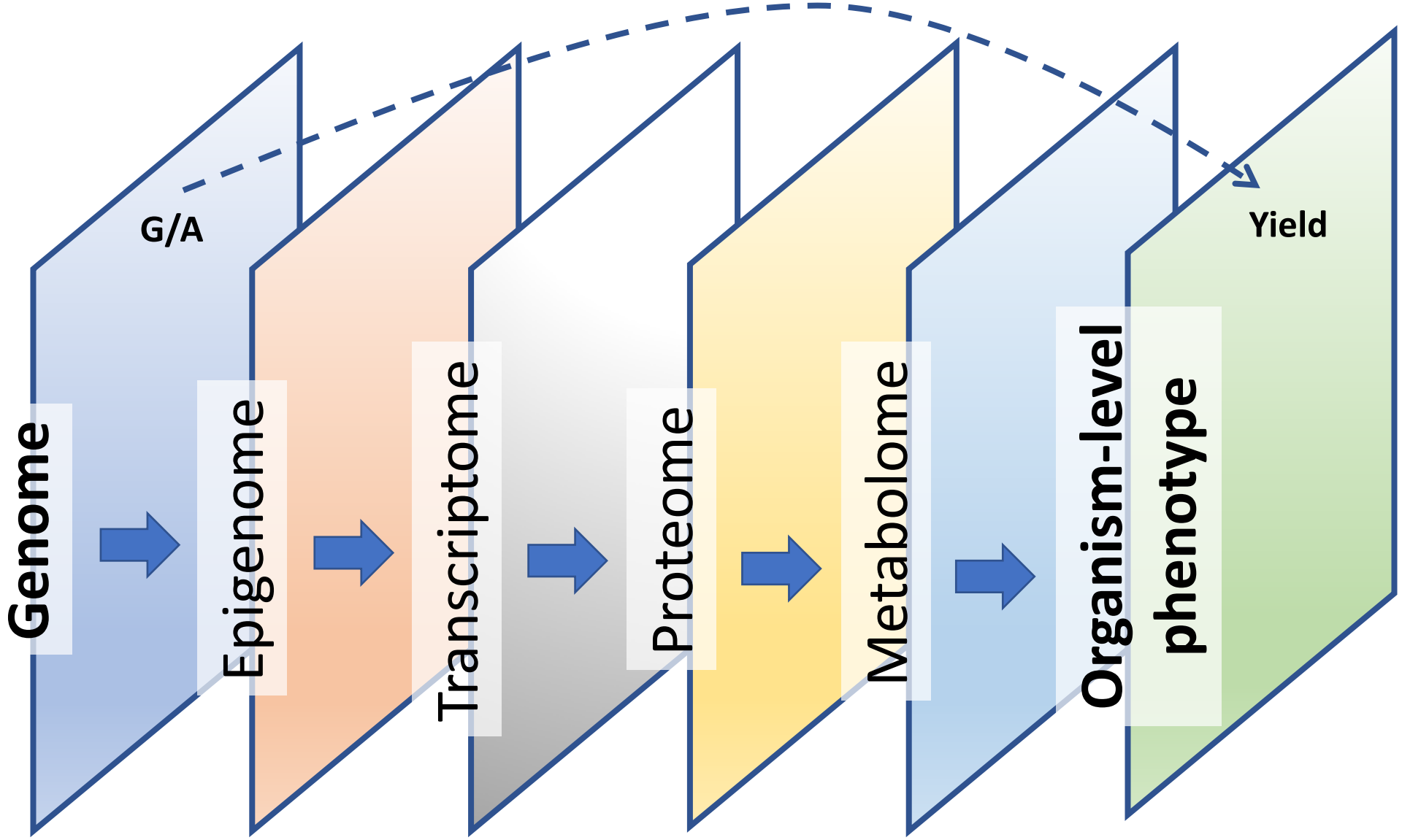
Throckmorton Plant Sciences Center
Department of Plant Sciences
Kansas State University

WHEAT CAP
Coordinated Agricultural Project

KANSAS STATE UNIVERSITY

IWGSC Fall 2020 Webinar Series, Oct. 7

Constructing genome-to-phenome maps is critical for predicting phenotypic outcomes of a particular genotype



Reference genomes and diverse populations of re-sequenced and phenotyped wheat accessions provide a powerful resource for connecting genomic and phenotypic variation

Reference genomes

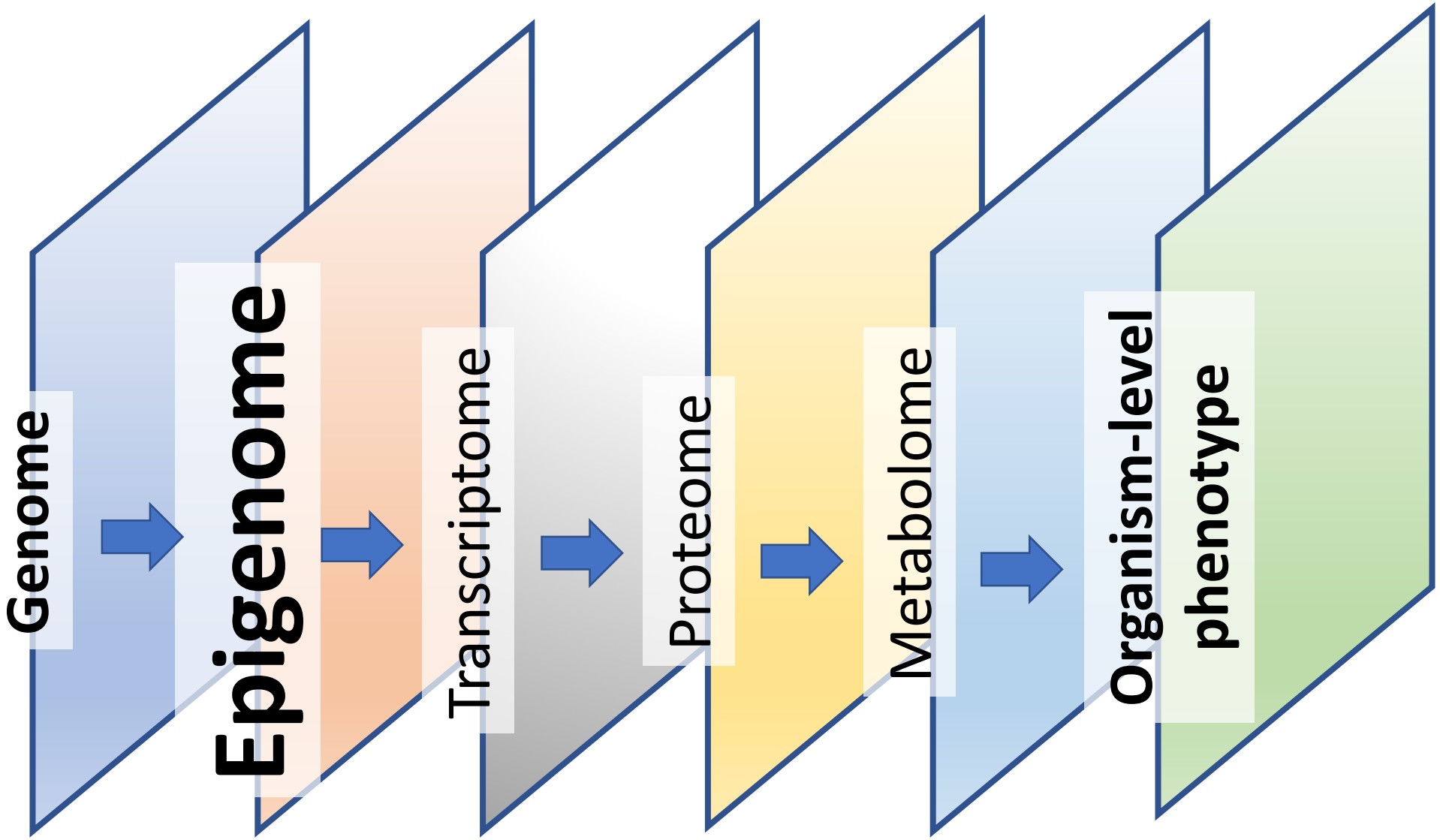
- Bread wheat cv. Chinese Spring (IWGSC, 2018)
- Wild emmer wheat (Avni et al., 2017)
- *Aegilops tauschii* (Luo et al., 2017)
- Durum wheat (Maccaferri et al., 2019)
- Wheat Pangenome (10+ Wheat Genome Project)

Densely genotyped diversity panels

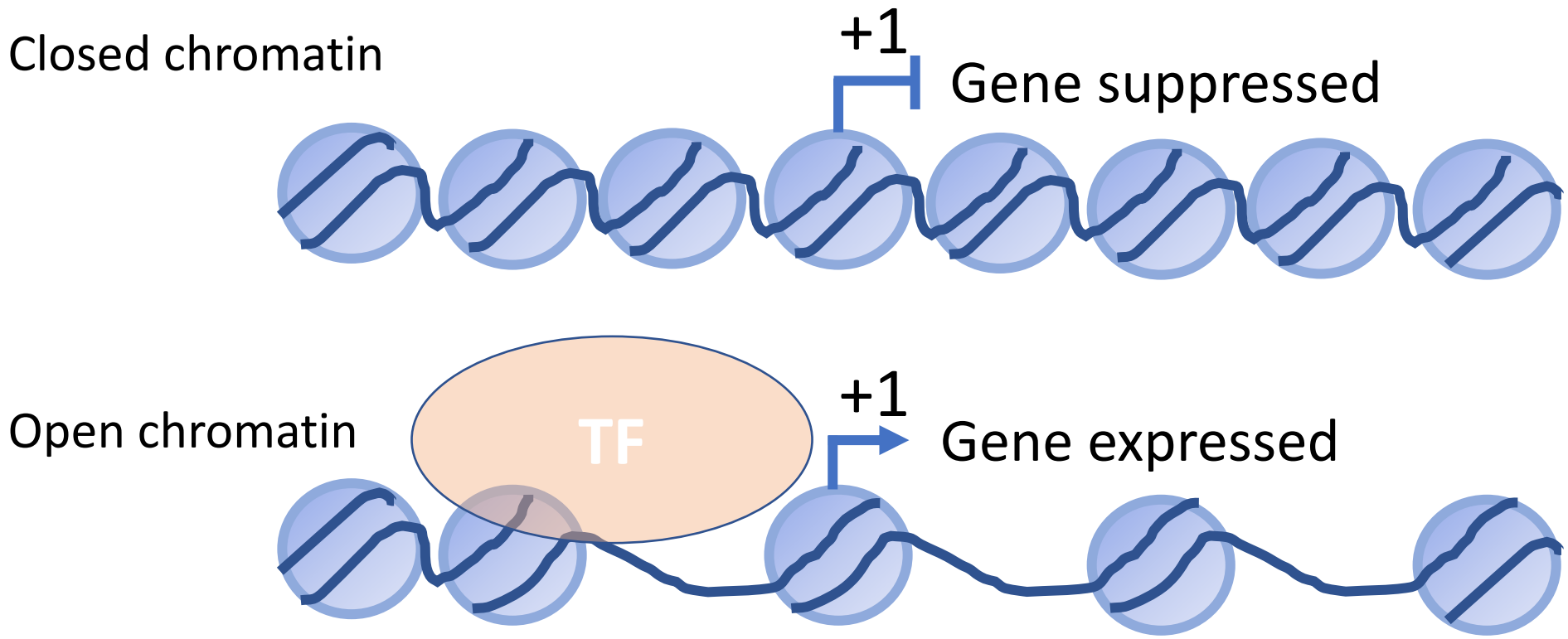
- 1,000 wheat lines (whole exome capture) (He et al., 2019)
- 2100 RILs from NAM population (whole exome capture + GBS + 90K iSelect) (Jordan et al., 2018)
- 44K wheat breeding lines (GBS) (Juliana et al., 2019)

Mechanistic understanding of the genetic effects of individual mutations requires integration of other types of genomics data

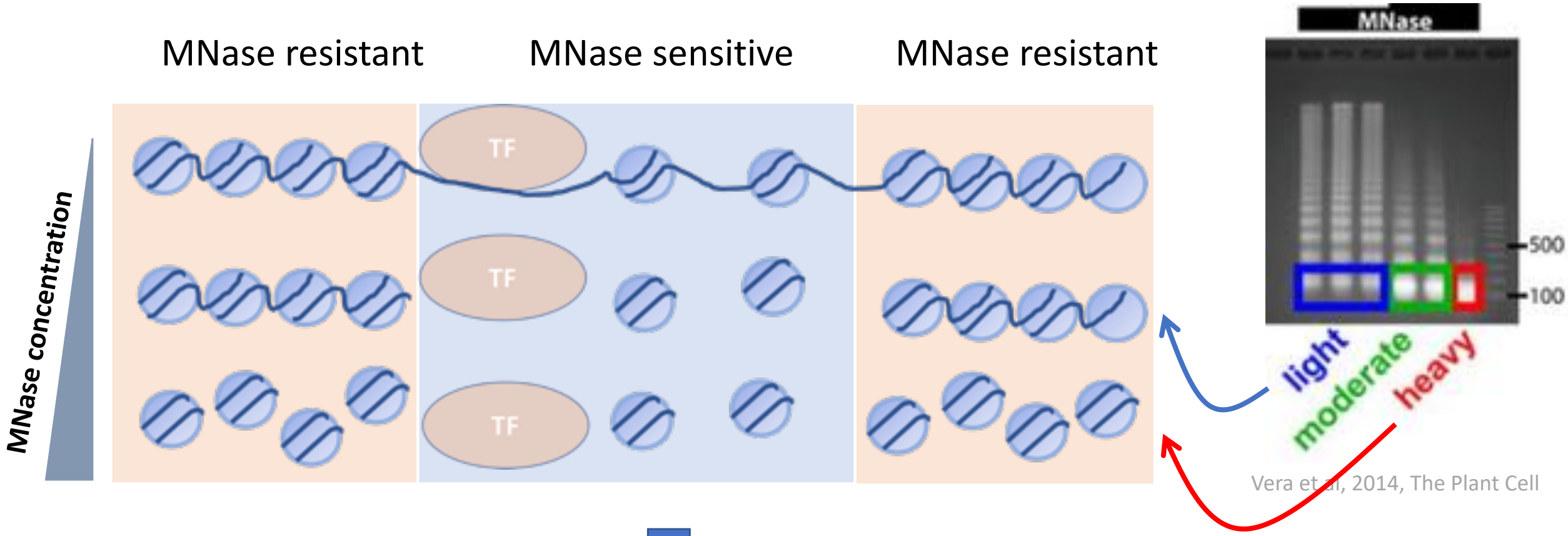
Genome-to-phenome map – epigenome dimension



Chromatin accessibility assays are critical for detecting epigenetically modified regions of genome involved in regulation of gene expression



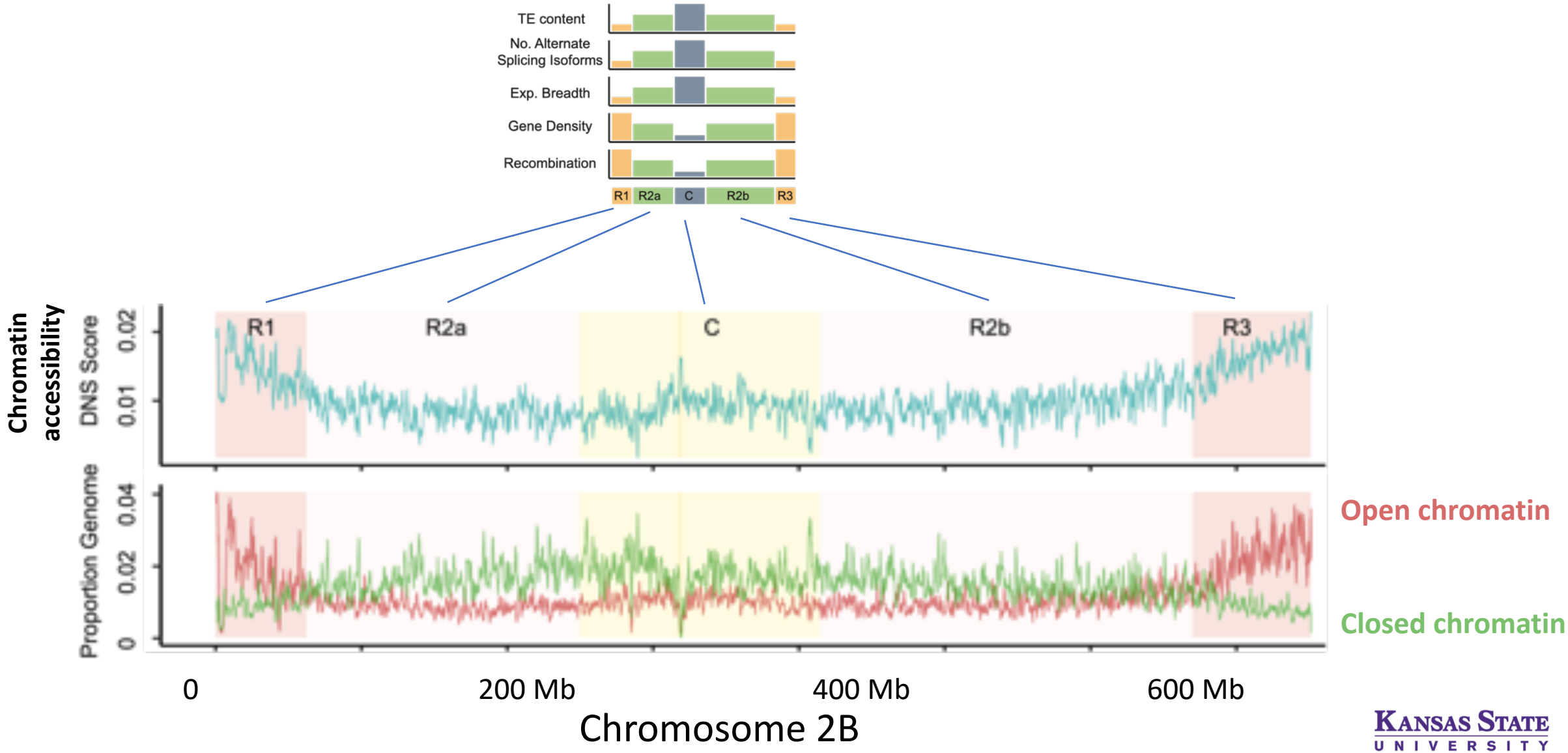
Difference in the NGS read coverage between MNase-sensitive and MNase-resistant treatments reflects open or closed chromatin states



Vera et al, 2014, The Plant Cell

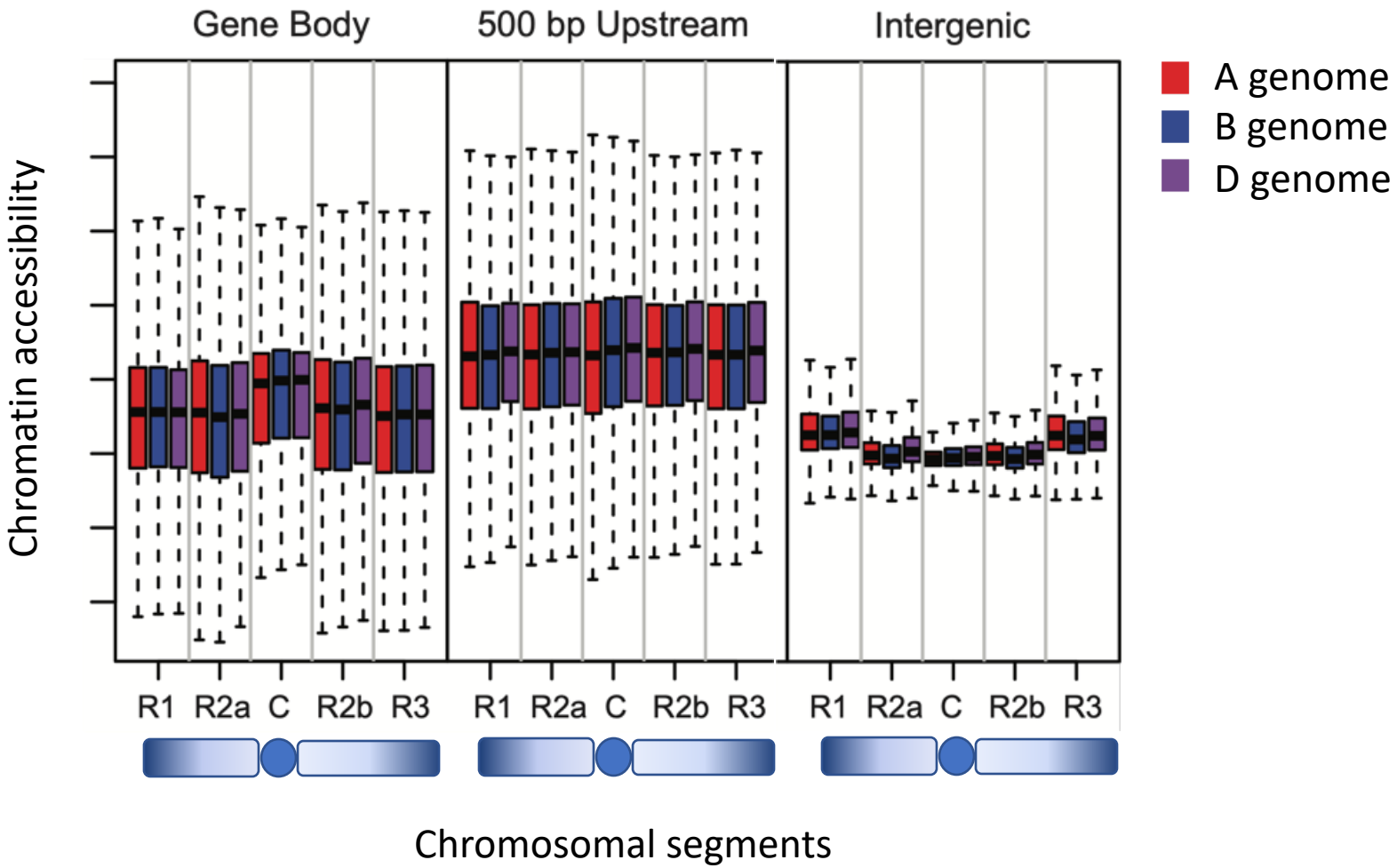
NGS with Illumina and mapping reads to the wheat reference genome

Chromosome-level chromatin accessibility follows the patterns of gene density, gene expression, recombination, and genetic diversity

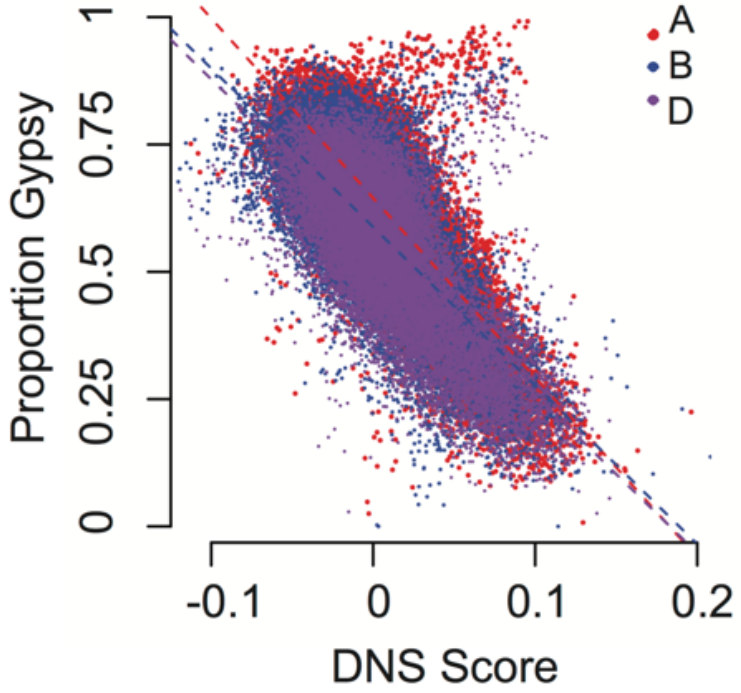


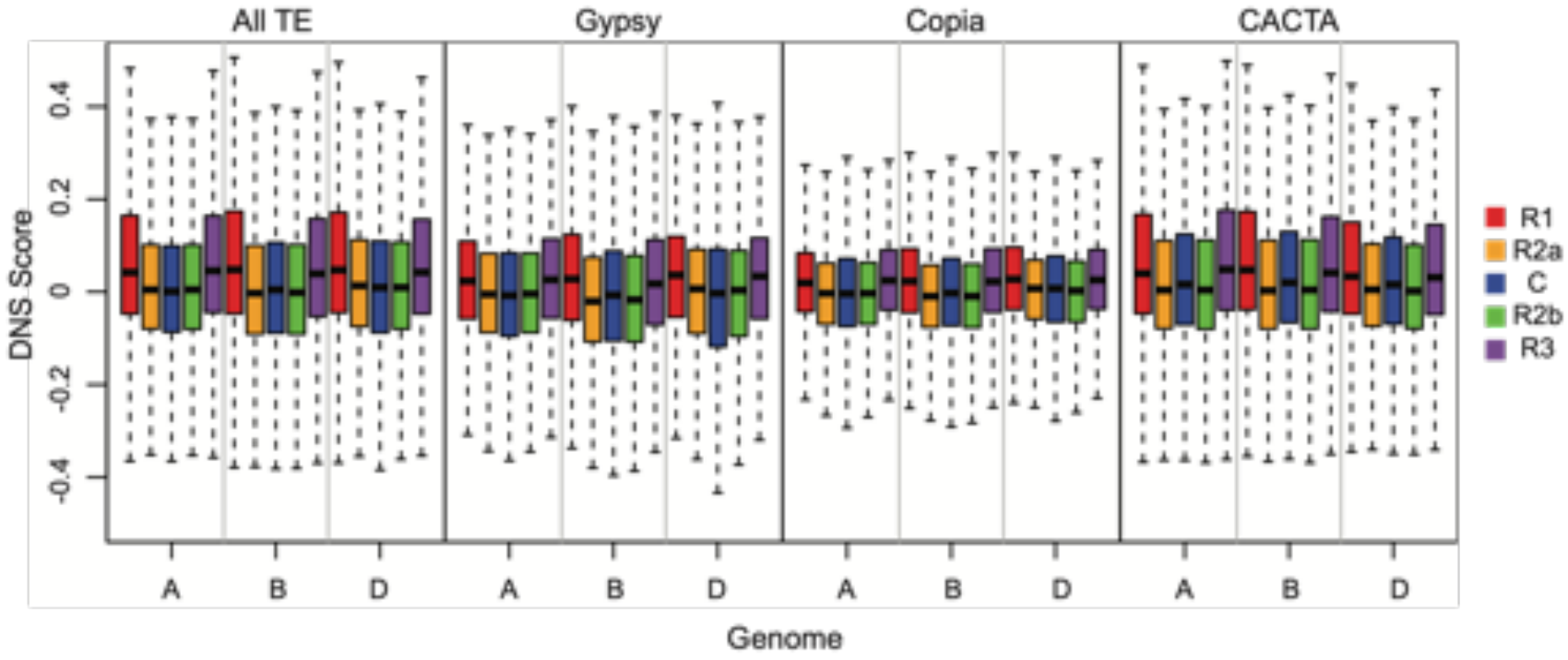
KANSAS STATE UNIVERSITY

While chromatin states in **intergenic regions change** from open to more closed from telomere to centromere, chromatin states in **genic regions remain mostly uniform** along the chromosomes

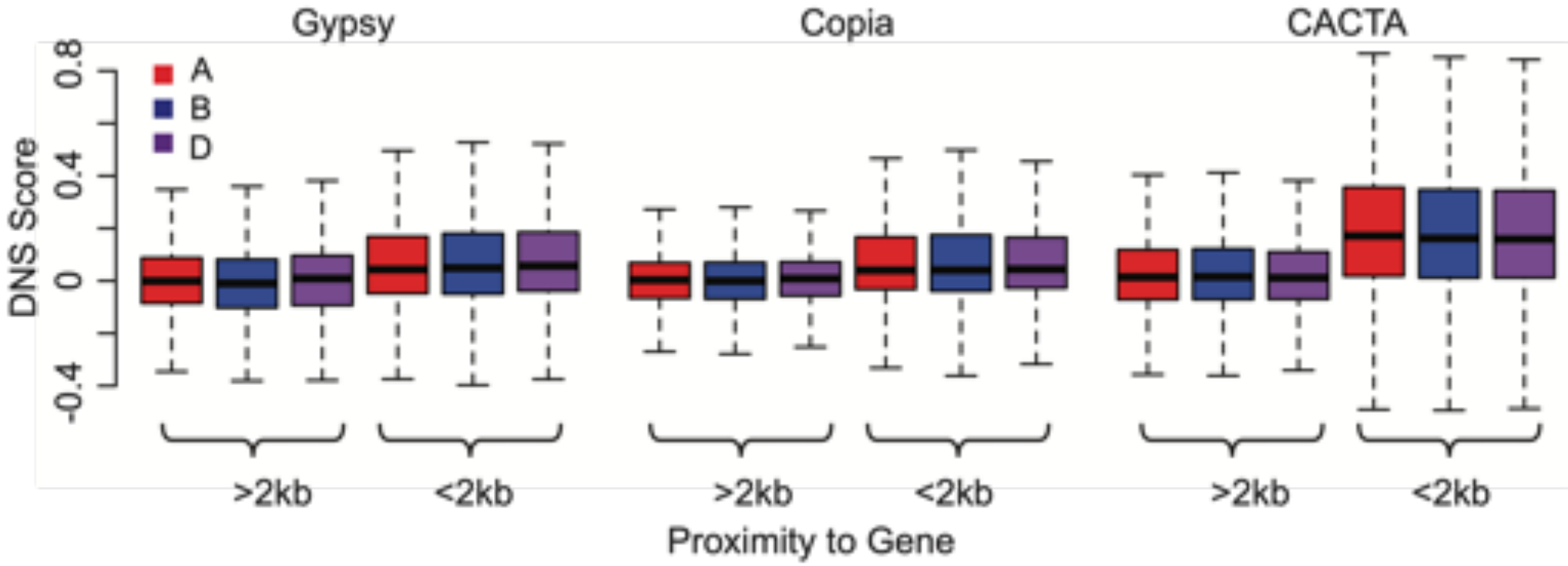


Even though chromatin accessibility of a region depended on %TEs, TEs in centrometic regions still showed lower chromatin accessibility than TEs in the pericentromeric regions

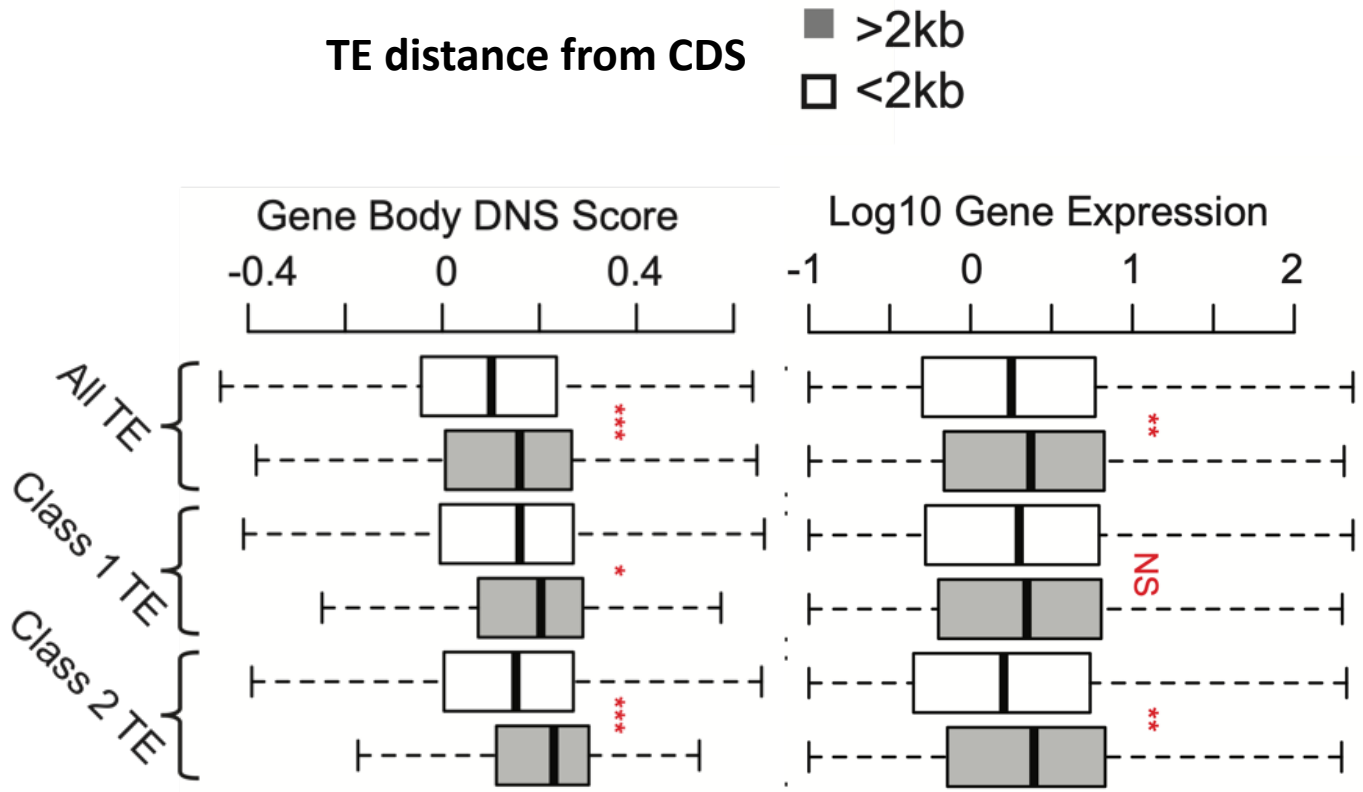




Chromatin accessibility of TEs, independent of TE type was associated with their proximity to genes

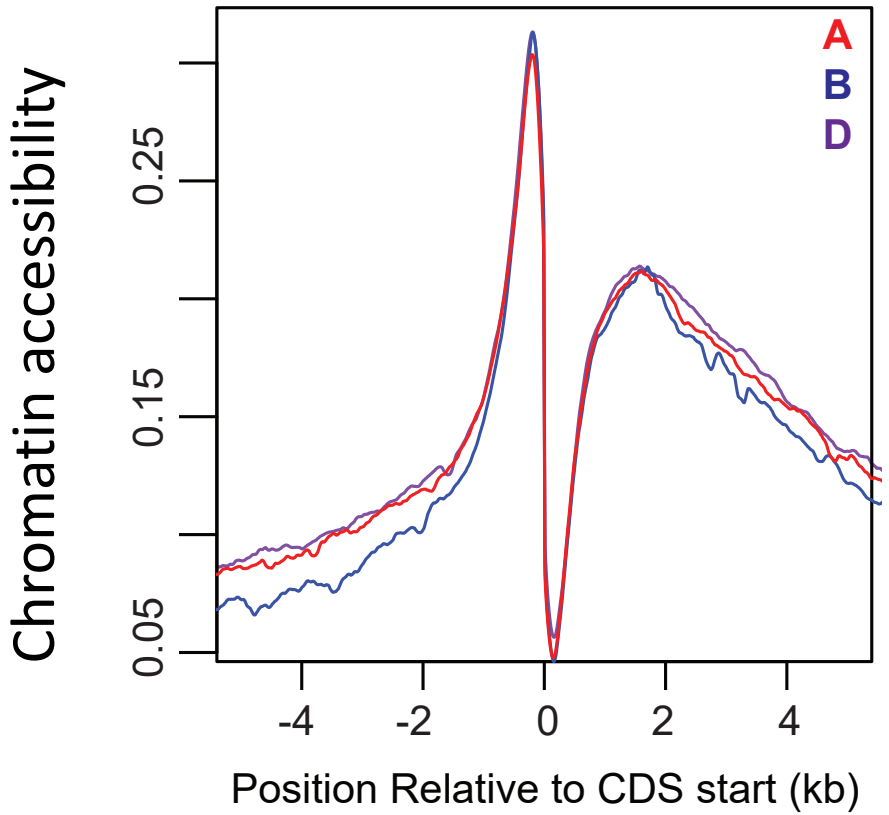


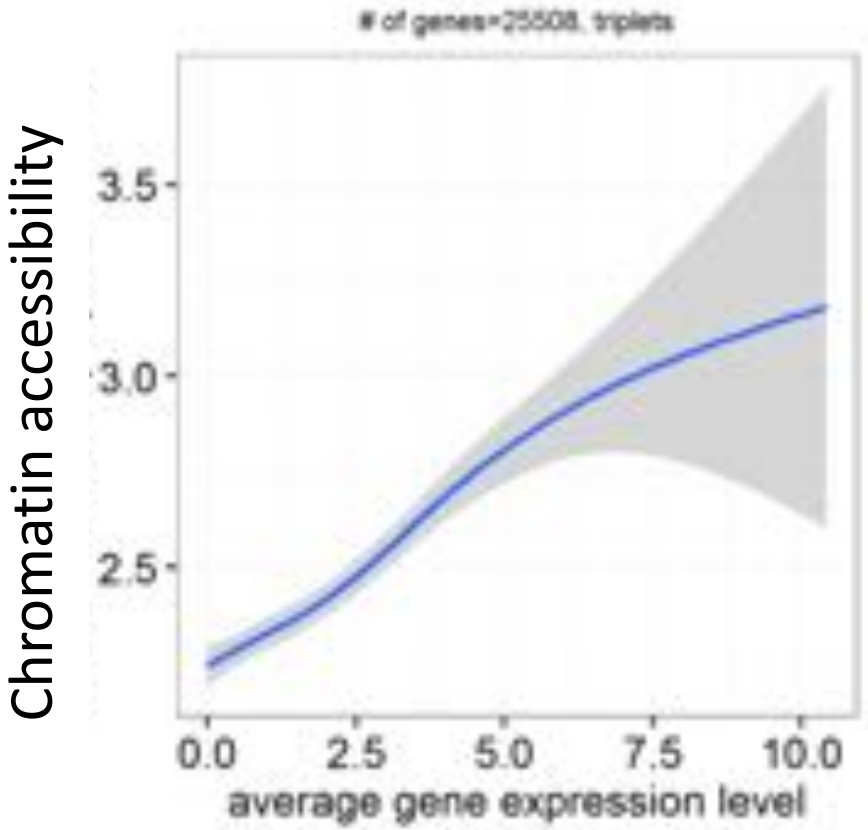
Chromatin accessibility of genes and their expression levels were associated with the presence of TEs in the promoter regions, suggesting that TEs might affect gene function by modifying chromatin states of regulatory regions



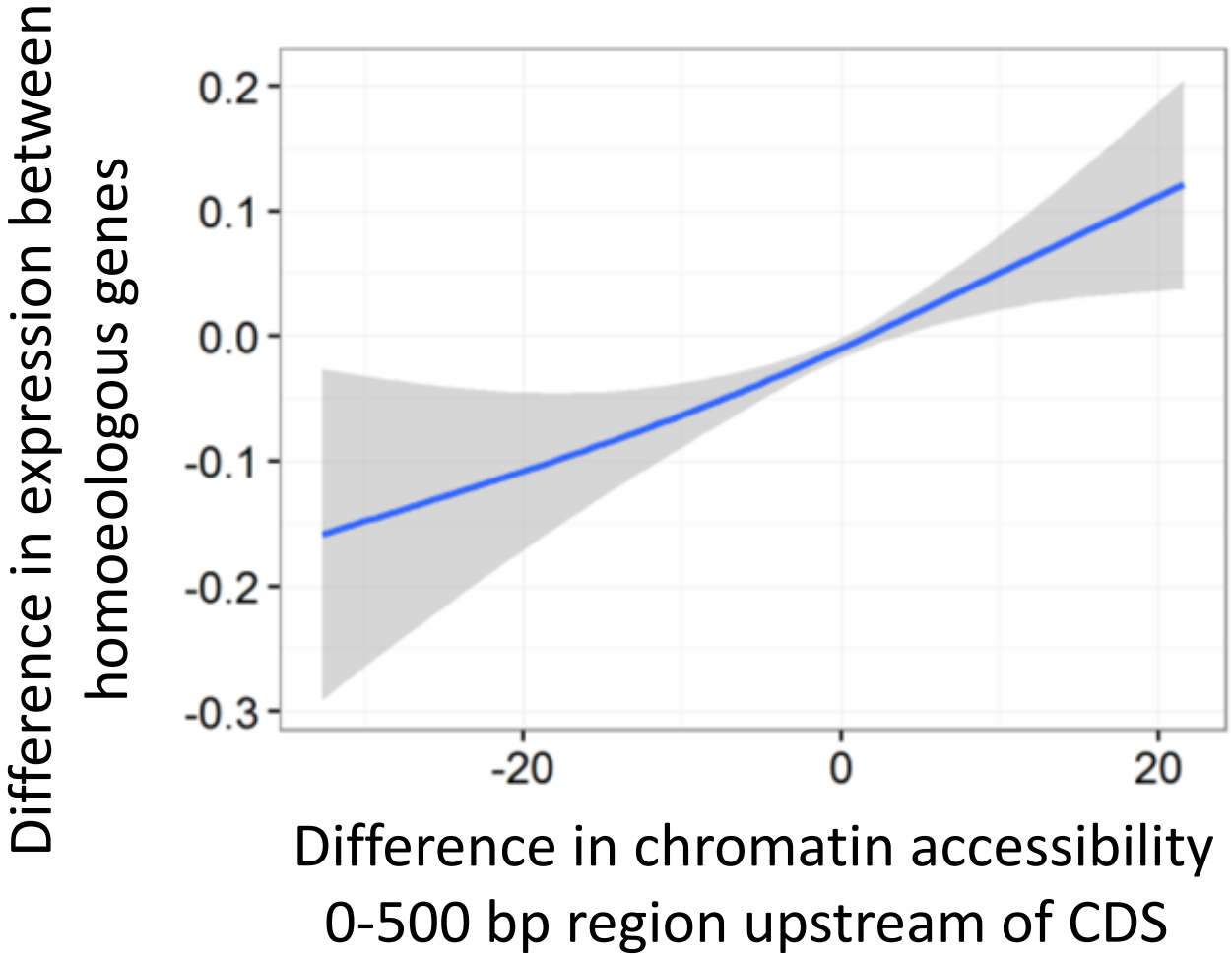
Gene expression data from
Ramirez-Gonzalez, et. al. 2018

Promoter regions showed highest levels of chromatin accessibility, which correlates positively with gene expression levels

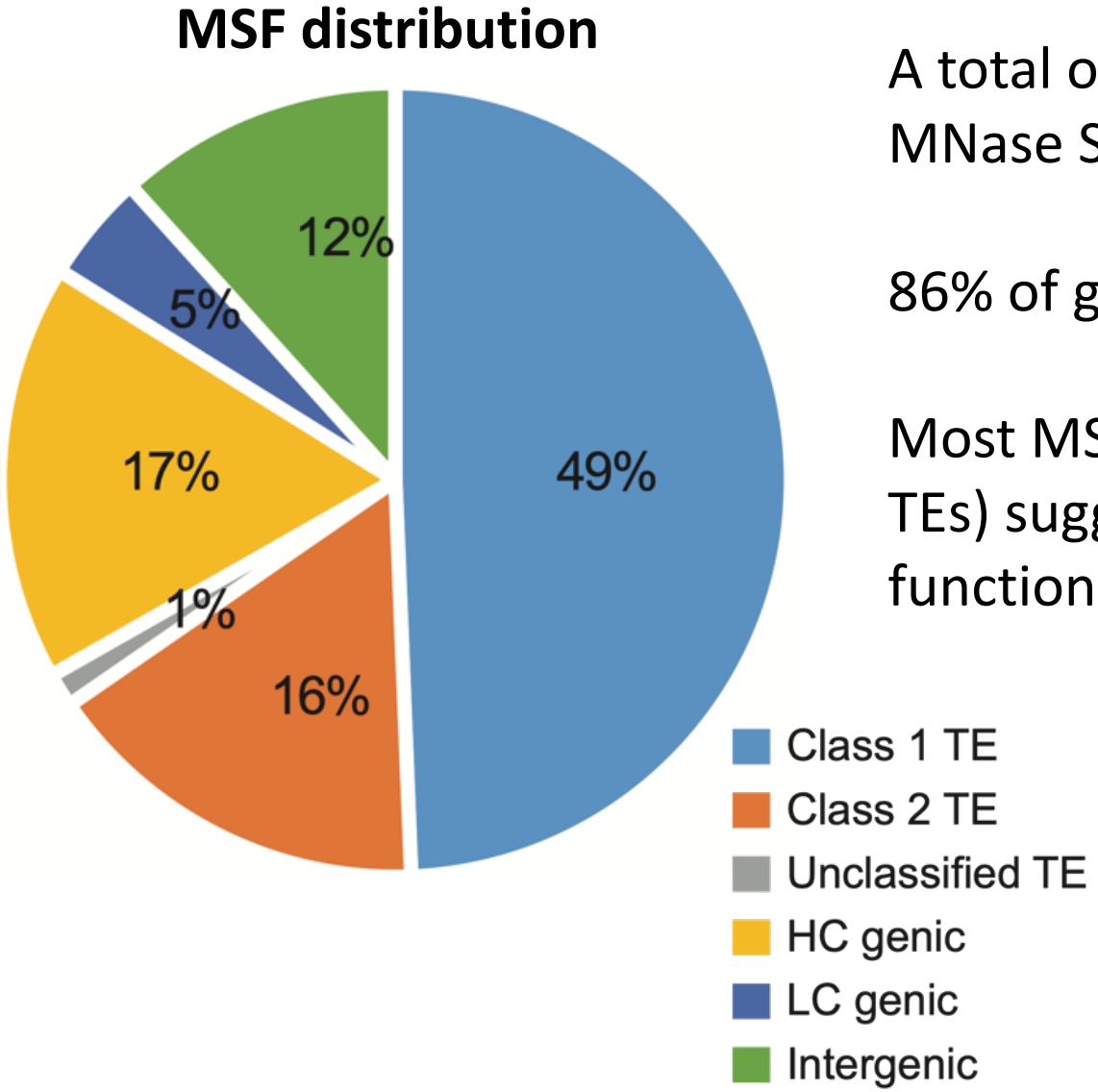




Gene expression bias between homoeologous genes from the wheat genomes is associated with differences in chromatin states of their promoter regions



Though MNase Sensitive Footprints (MSF - outliers of chromatin accessibility) are enriched around genes, majority are found within TEs.



A total of 177 Mb (1.26%) of the genome were classified as MNase Sensitive Footprints (MSF)

86% of genes were located within 2 kb of at least one MSF

Most MSF are located in the intergenic regions (67% within TEs) suggesting importance of TEs in regulation of gene function

To assess functional relevance of chromatin accessibility data we partitioning genetic variation for agronomic traits using phenotyping and genotyping data from the 1000 wheat exome project

http://wheatgenomics.plantpath.ksu.edu/1000EC/

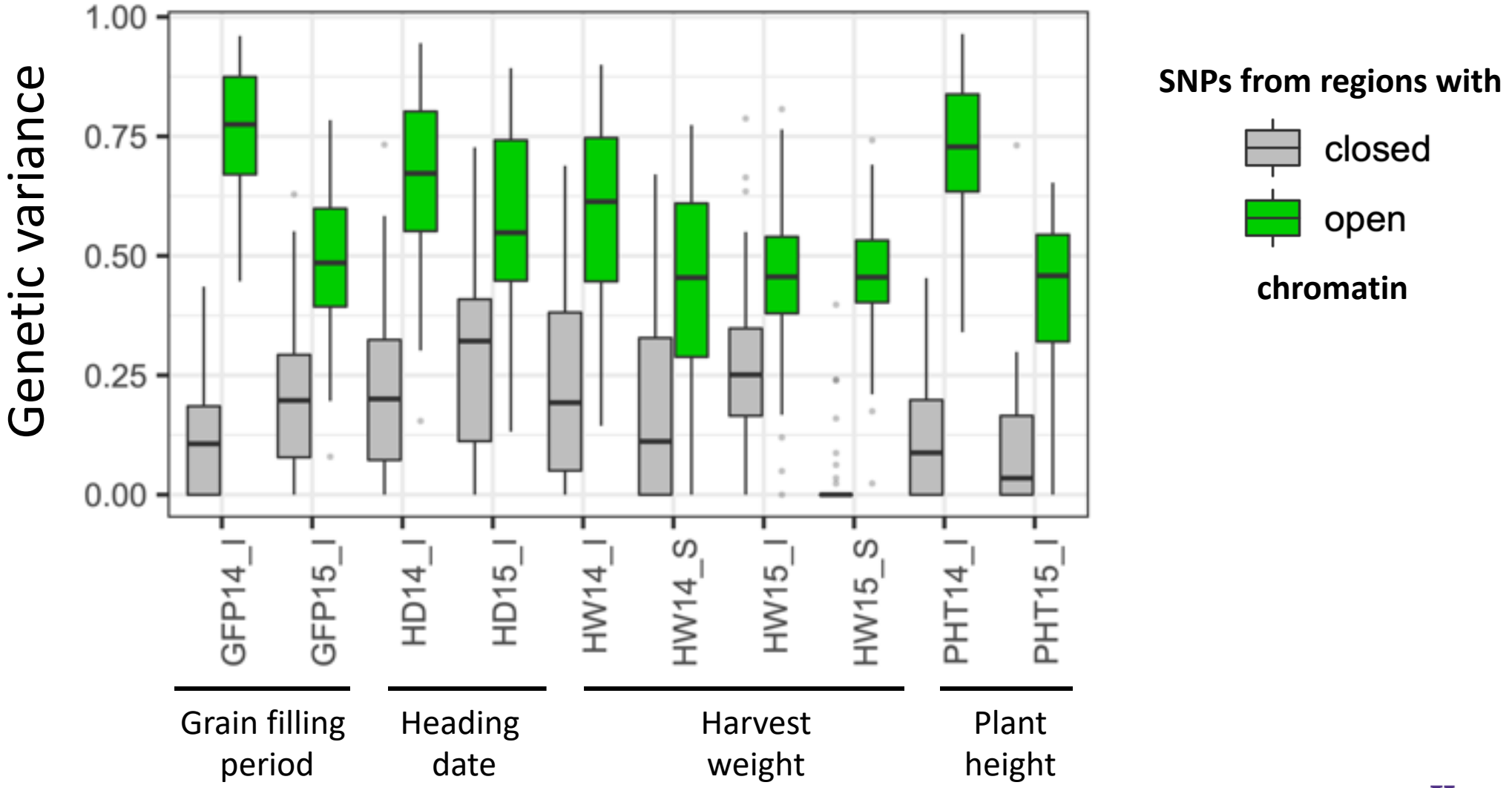
He at al., 2019

1,000 genetically and geographically diverse wheat accessions were re-sequenced using exome capture resulting in discovery > 7 million SNPs

Phenotyping was performed for a number of agronomic traits

Partitioning genetic variation (GCTA-GREML)

Increase in chromatin accessibility was associated with an increase in the proportion of phenotypic variation explained

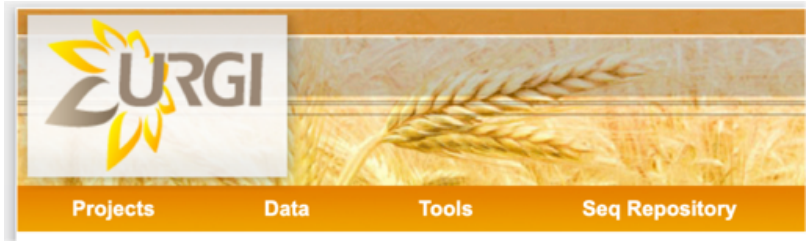


Chromatin accessibility is a good predictor of the effect of SNP variation on phenotype

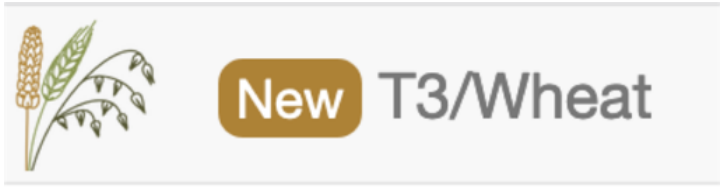
Developed map of chromatin states across the wheat genome could be useful for prioritizing SNPs in genomic selection or detecting causal SNPs in gene mapping studies or GWAS

Differential chromatin accessibility resources and tools

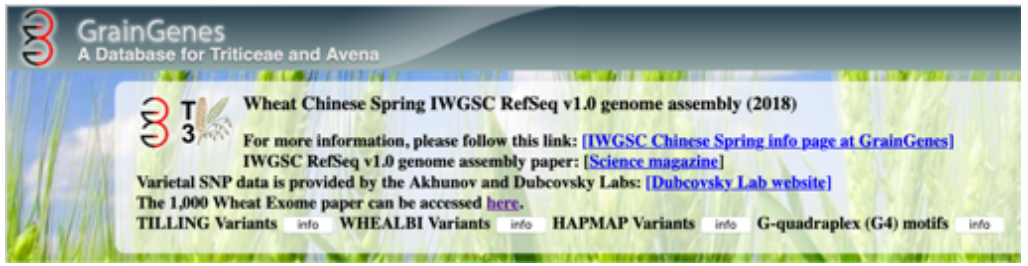
Chromatin accessibility data
https://wheat-urgi.versailles.inra.fr/Seq-Repository/Chromatin-accessibility



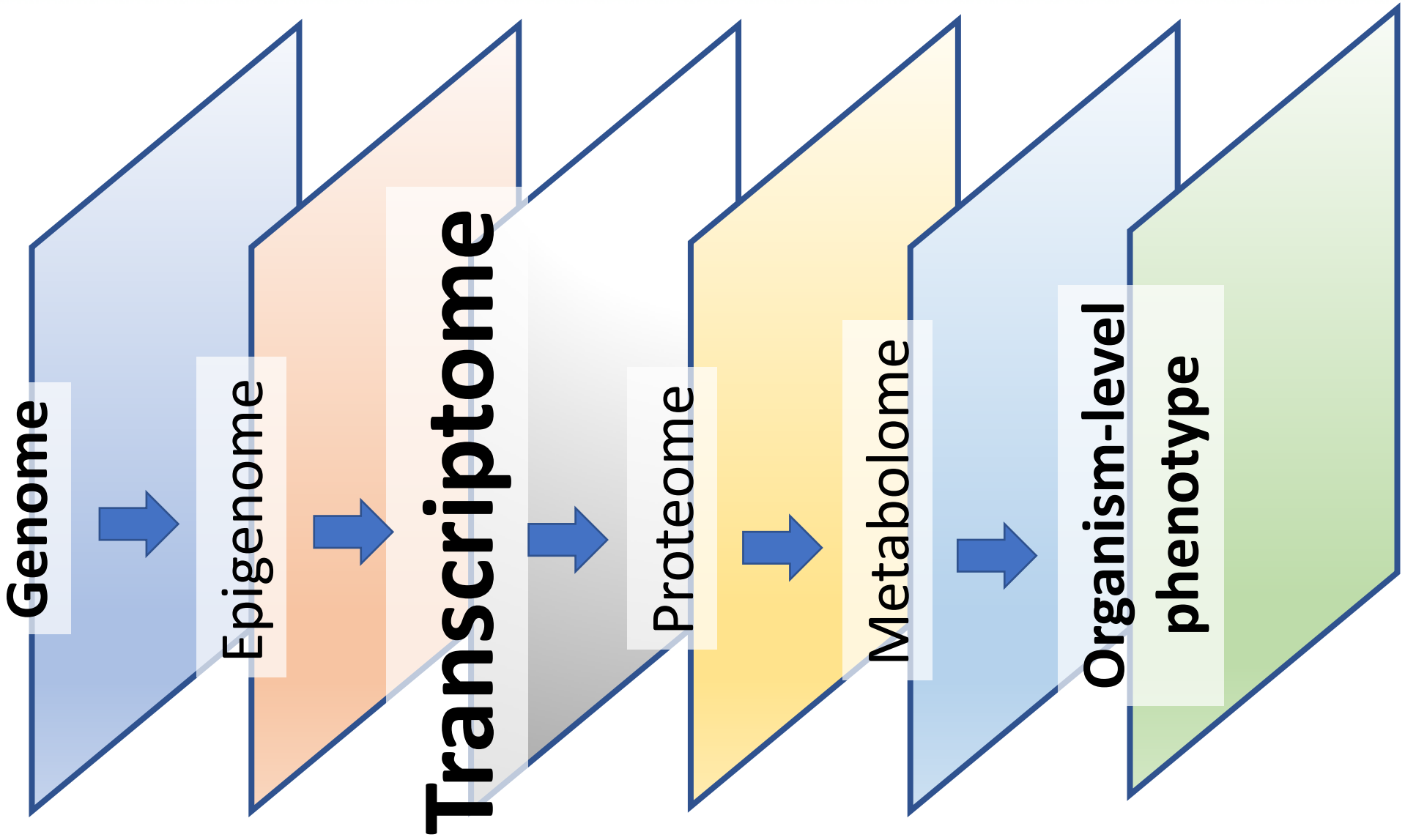
Filtering SNPs using differential chromatin accessibility profiles:
T3 database (C. Birkett, J. Jannink)
https://wheat.triticeaetoolbox.org/genome/mnase.pl



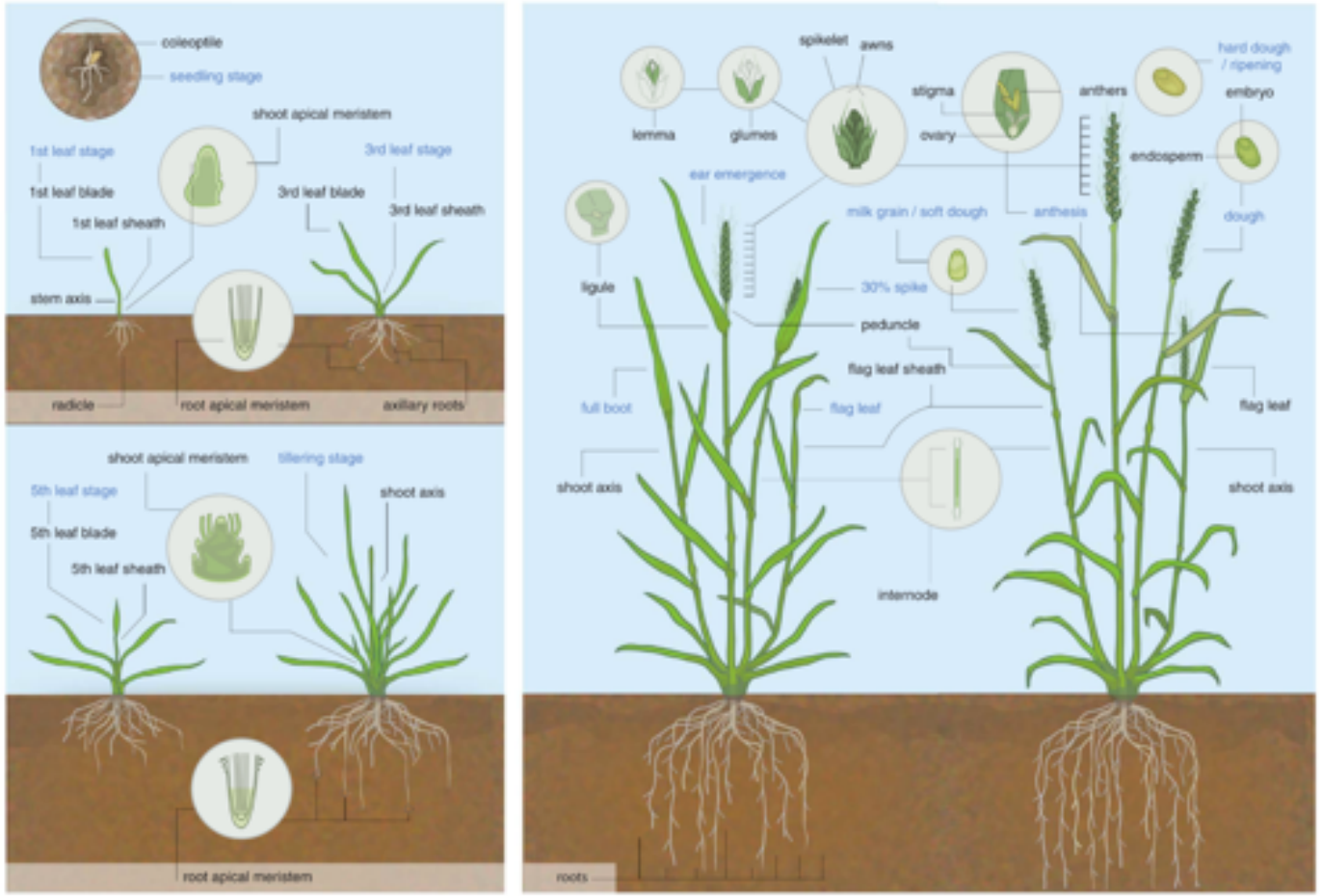
GrainGenes (T. Sen)
https://wheat.pw.usda.gov/GG3/content/september-2020-mnase-chromatin-states-tracks-iwgsc-chinese-spring-genome-browser



Genome-to-phenome map – transcriptome dimension



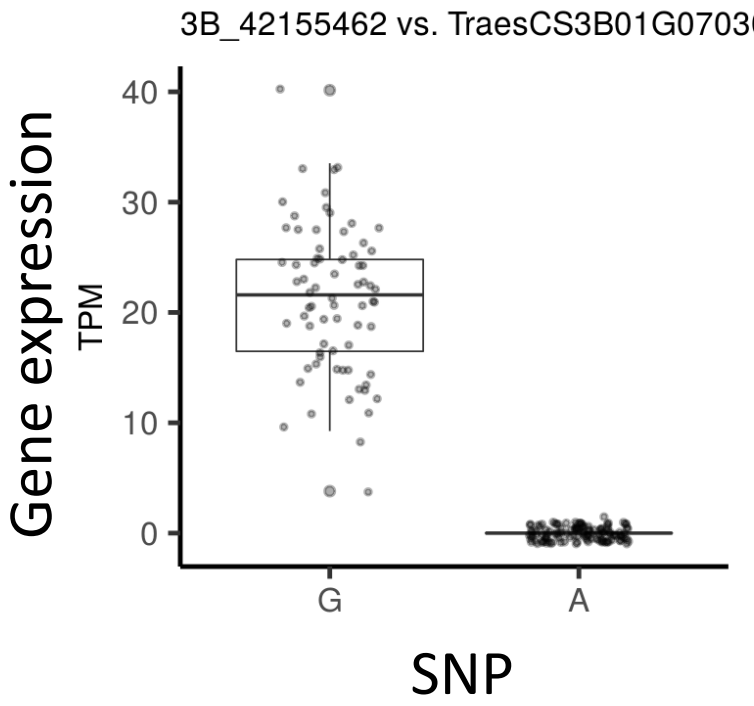
Transcriptional atlas of wheat provides description of gene expression changes across tissues at differential stages, yet we still need to understand genetic control of expression in wheat

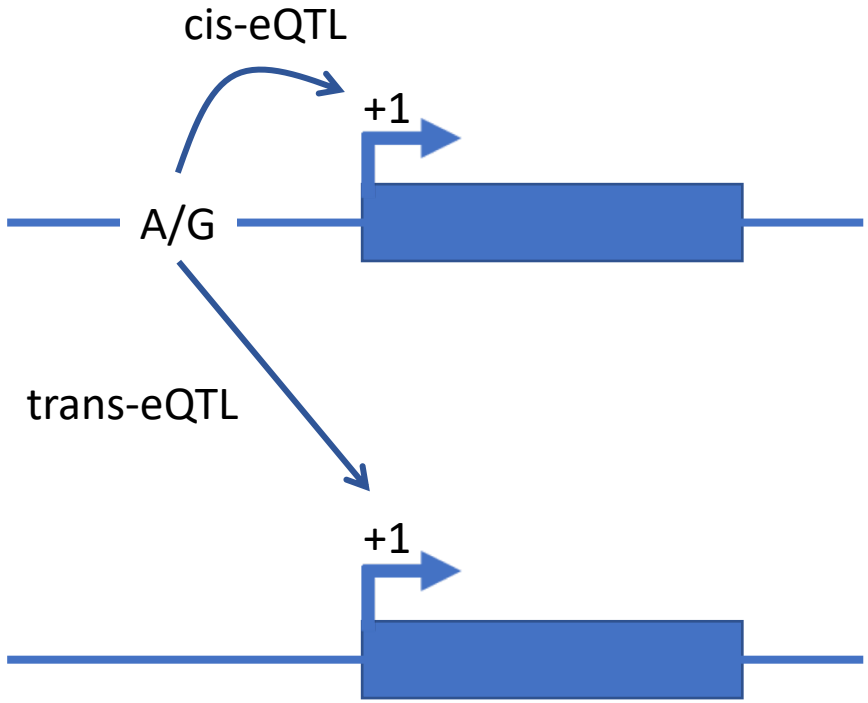


Ramírez-González et al. Science. 2018;361:eaar6089

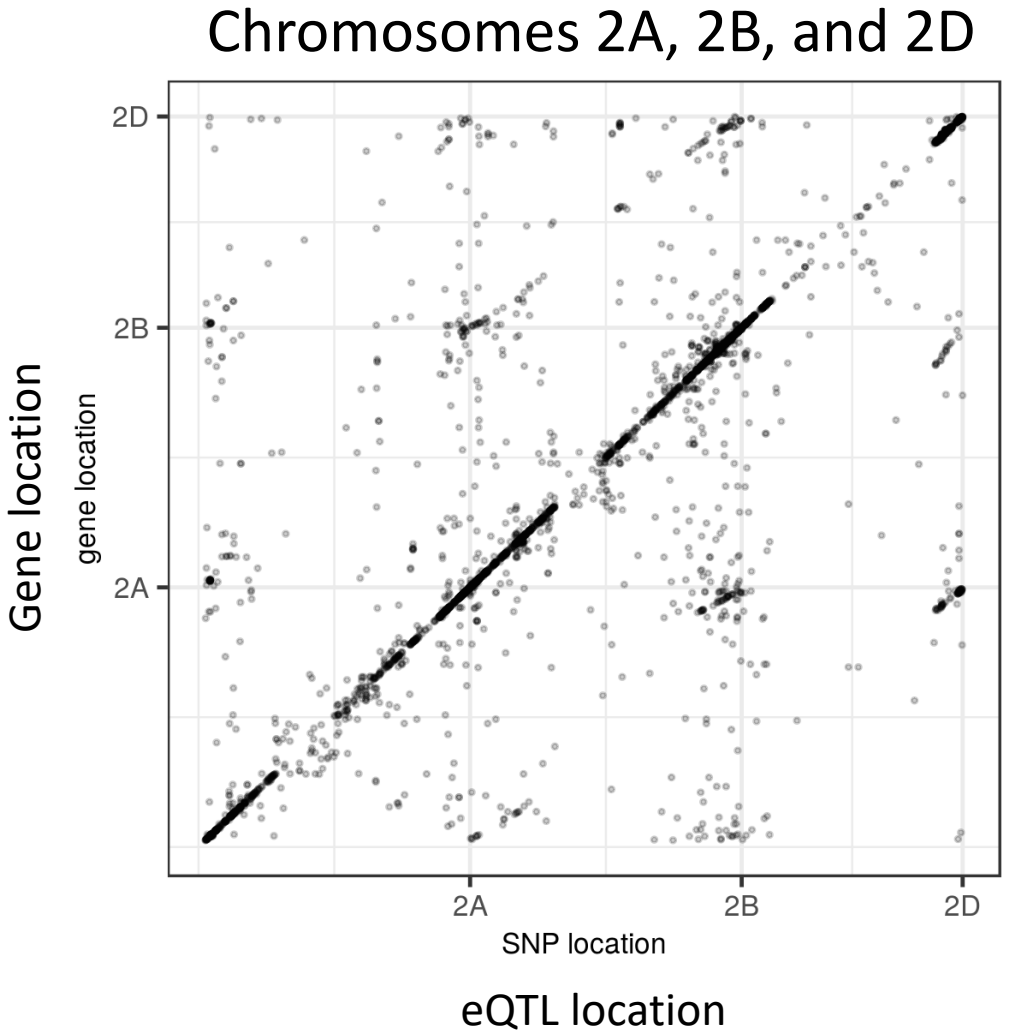
Understanding the genetic control of gene expression variation (eQTL) can improve our ability to detect causal genes and pathways

RNA-seq data from diverse panel of wheat accessions
GWAS of gene expression traits

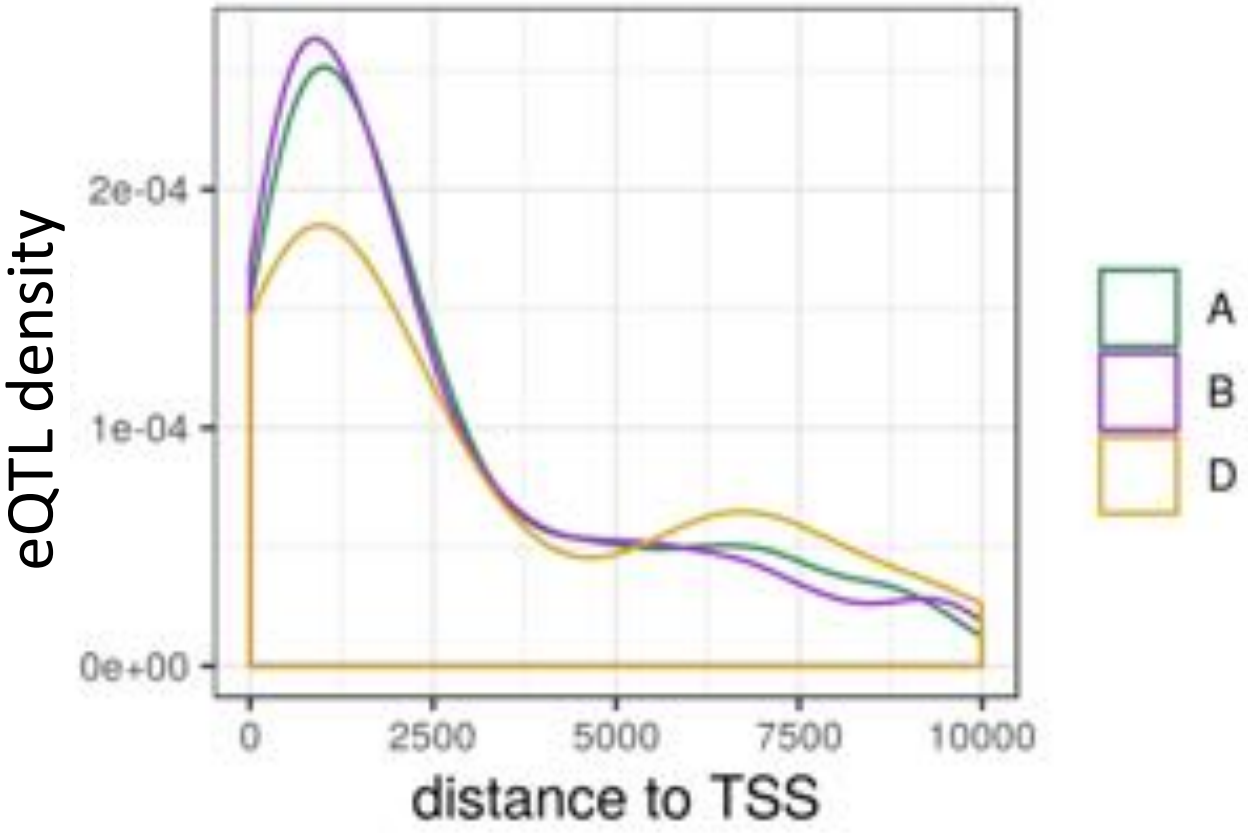


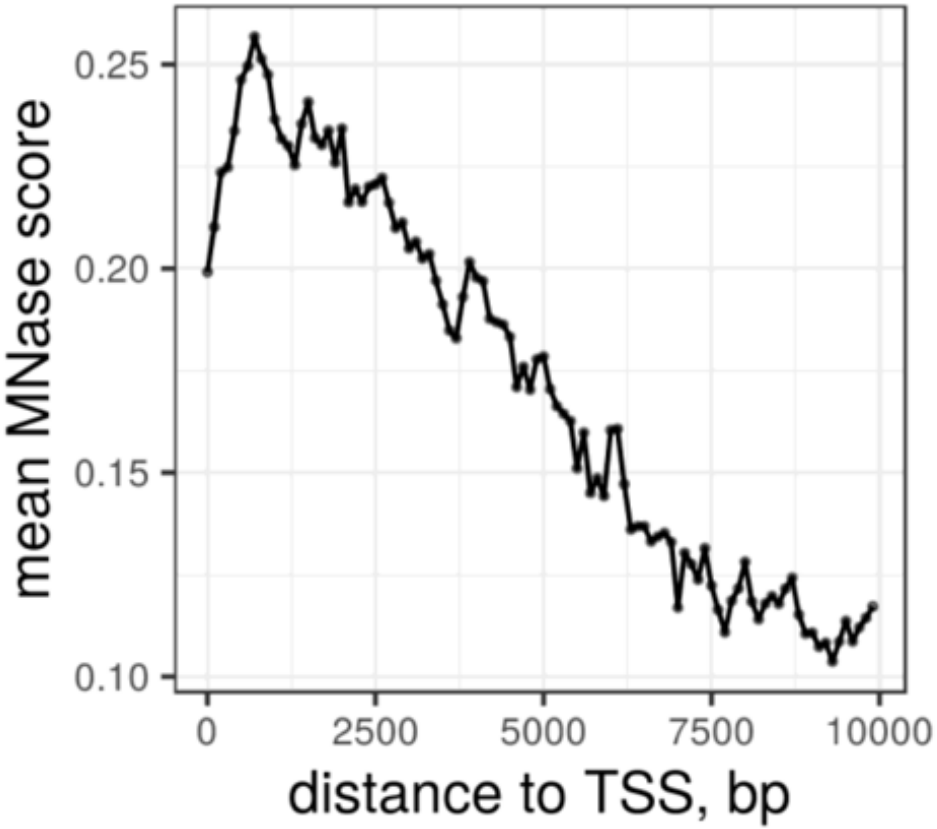


GWAS identified ~37,000 eQTL regulating expression of ~8,000 genes

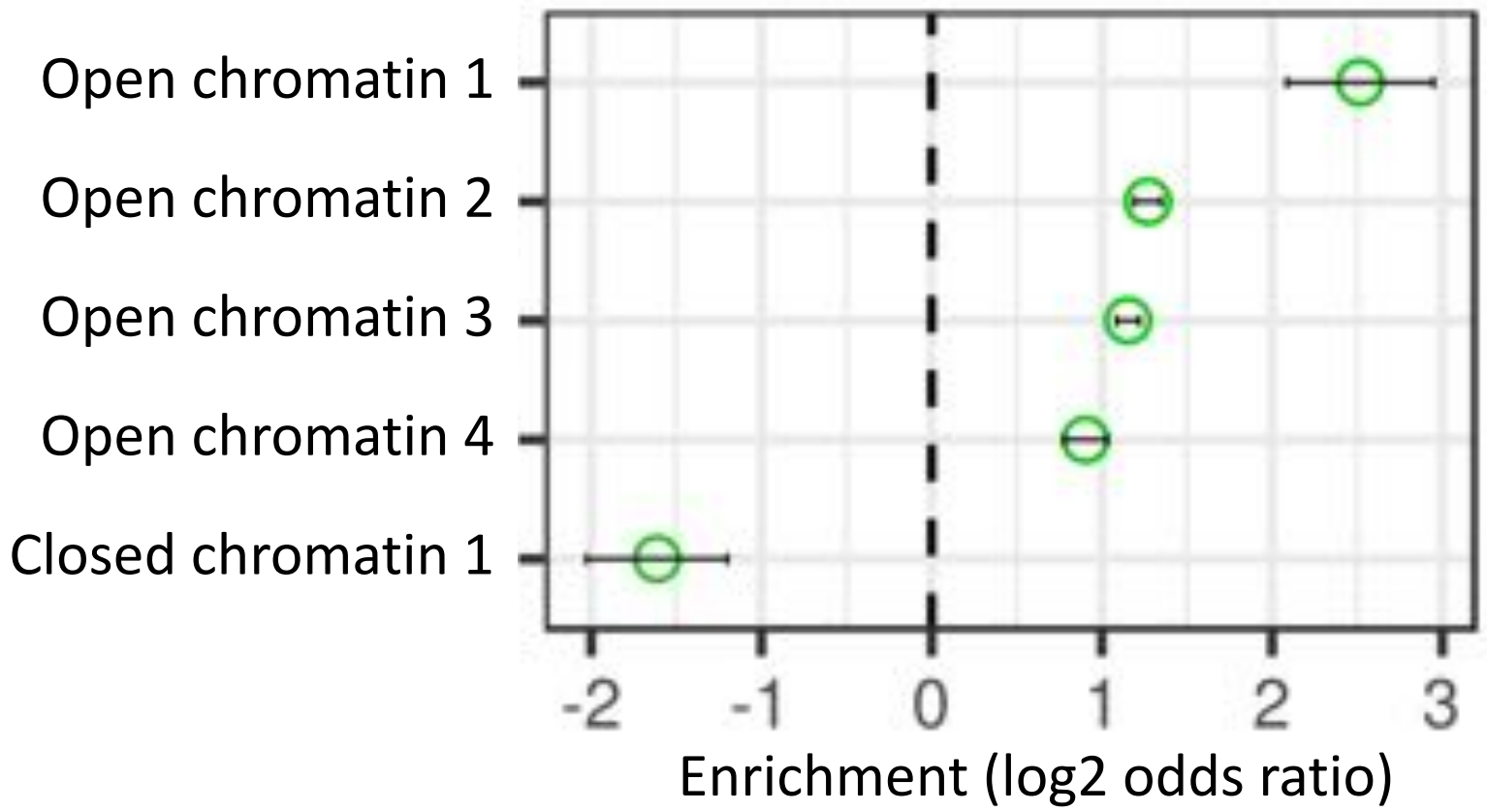


cis-eQTL density was highest in 2 kb regions upstream of genes in all three wheat genomes mirroring the patterns of chromatin accessibility

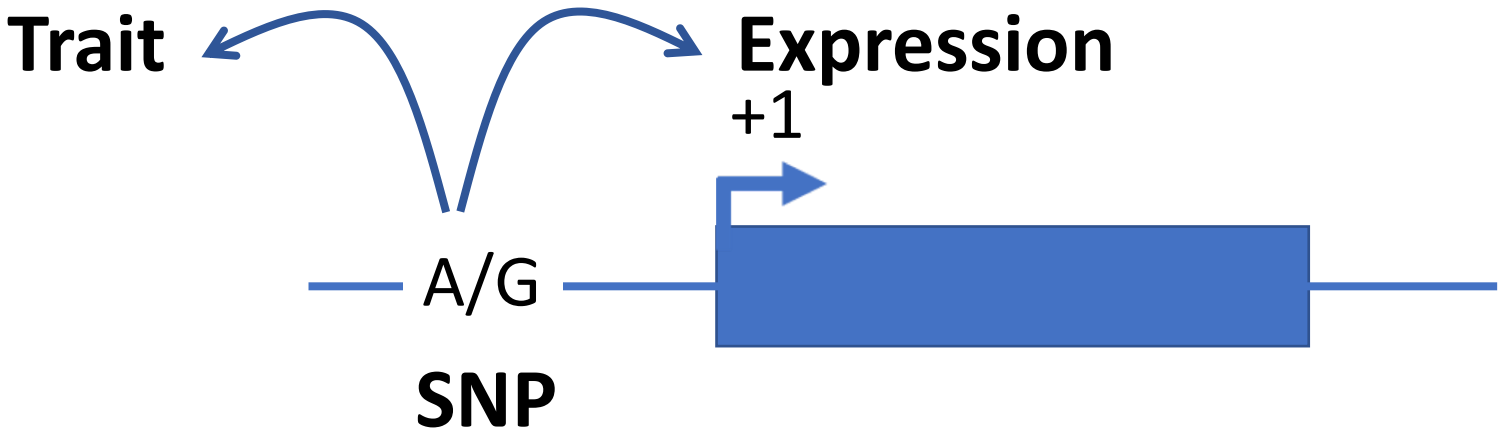




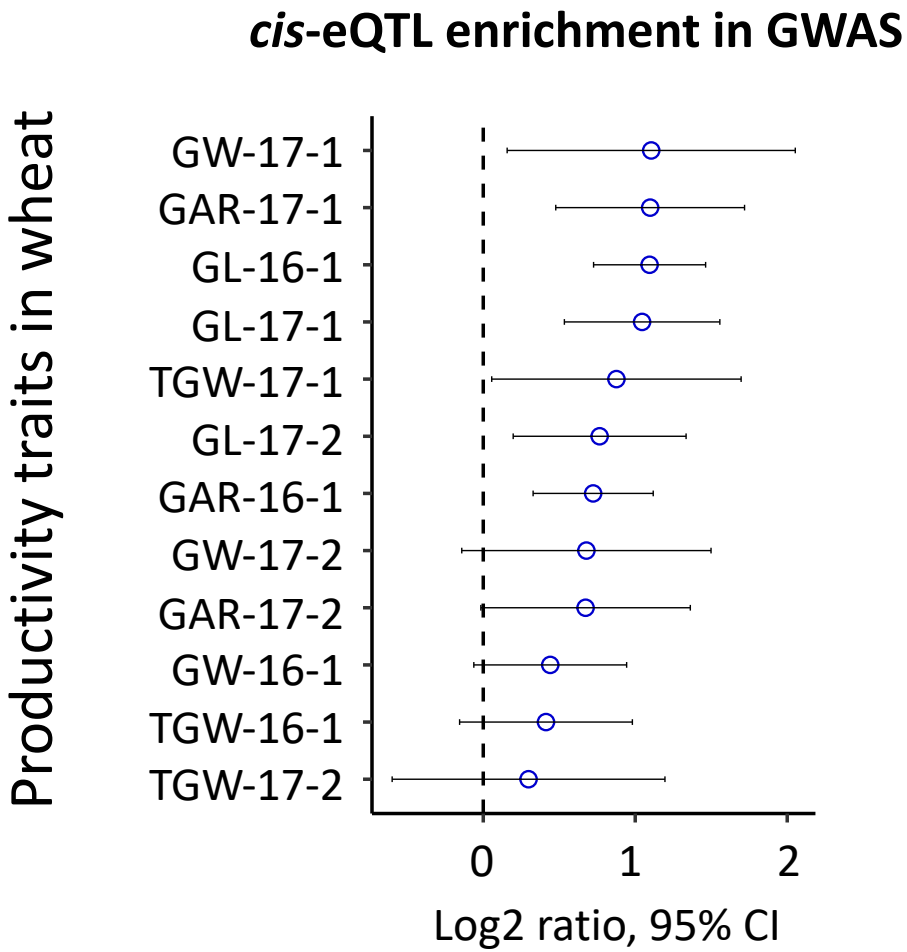
eQTL are enriched in the regions of open chromatin defined using different approaches including regions located distantly from genes



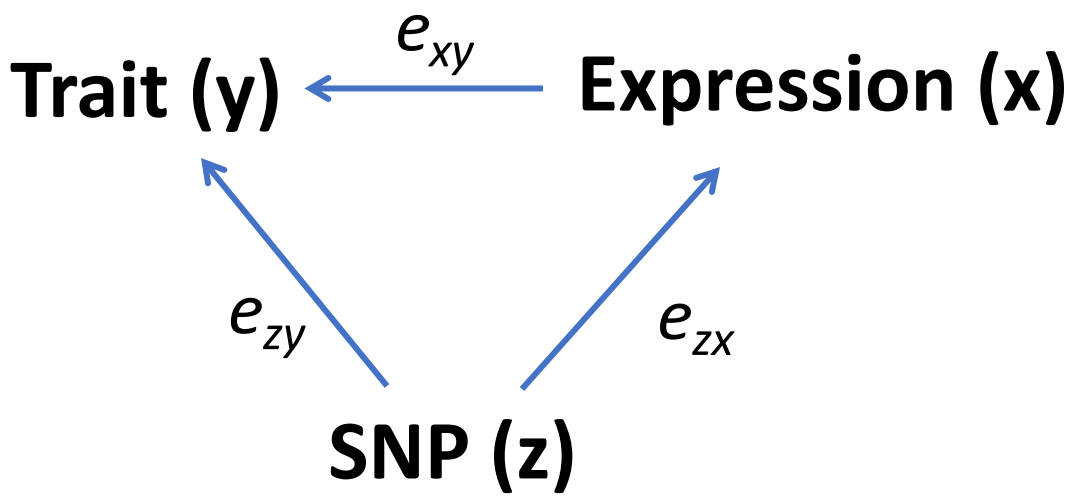
Joint analysis of eQTL and trait-associated SNPs provides opportunity to identify genes that are part of pathways controlling trait variation



Enrichment of eQTL among trait-associated SNPs suggest their contribution to trait variation



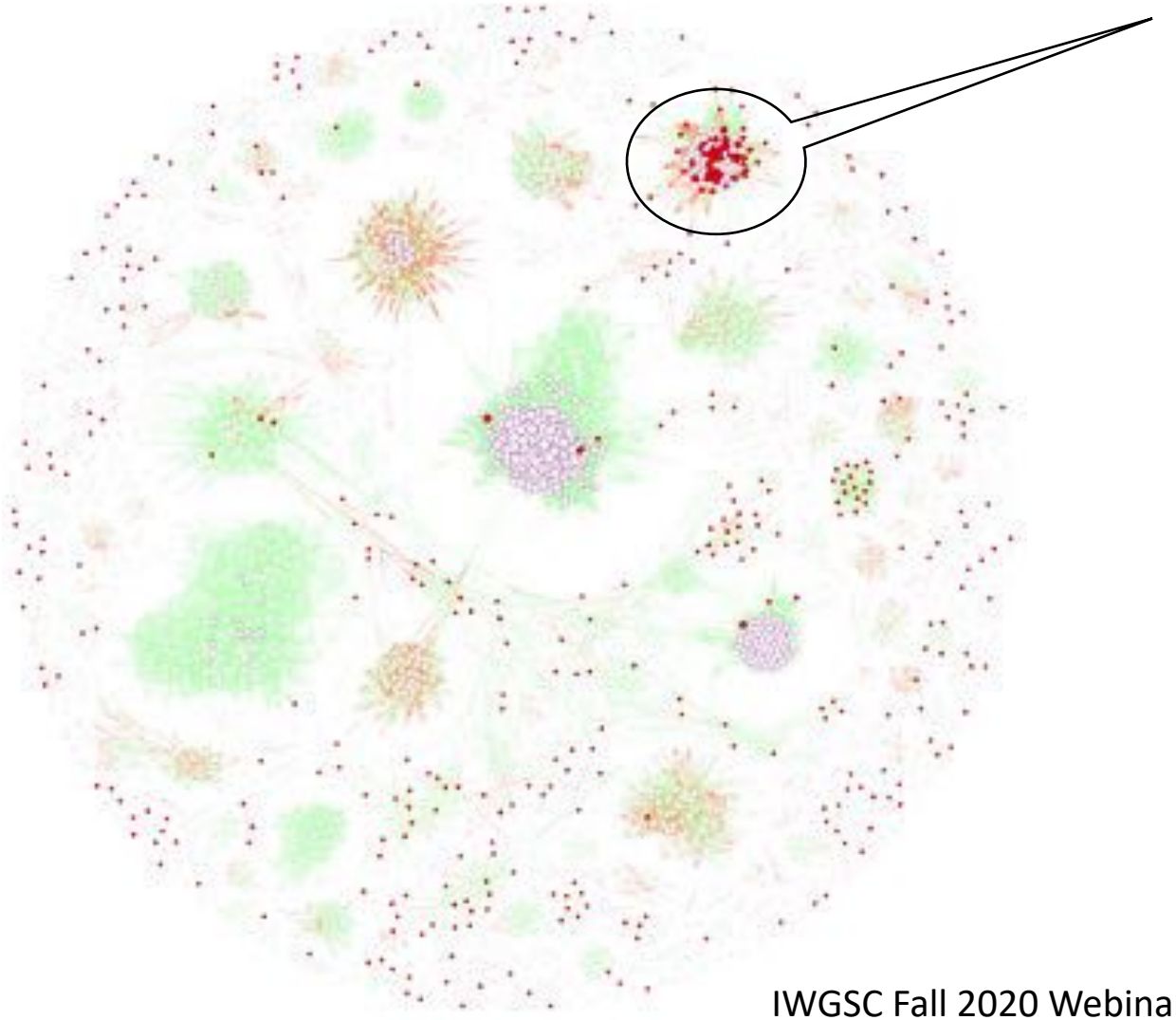
Joint modeling of eQTL and trait-associated SNPs provides opportunity to identify genes whose expression is associated with trait variation



$$e_{xy} = e_{zy}/e_{zx}$$

Effect of expression of gene expression **x** on trait **y**

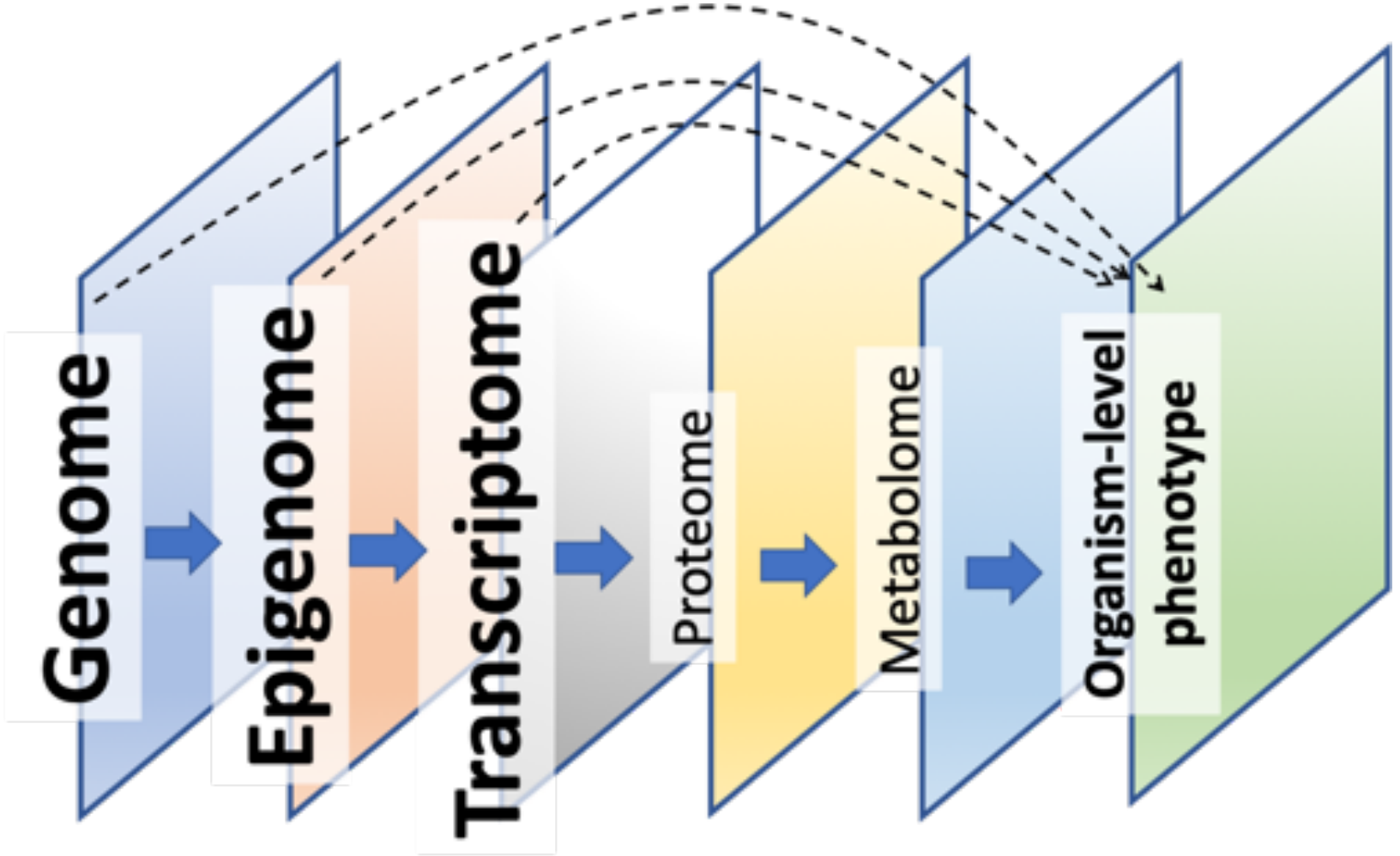
By incorporating genes associated with agronomic traits and eQTL into the gene co-expression networks it is possible to identify pathways controlling processes underlying variation in a trait



Genes involved in regulation of heading date and number of spikelets per spike (includes EARLY FLOWERING 3 - *Elf3*).

Co-expressed gene modules including genes associated with eQTL and trait (● red)

Integration of genomic data from multiple dimensions of G2P map could help to identify regulatory regions, genes and pathways underlying variation in traits



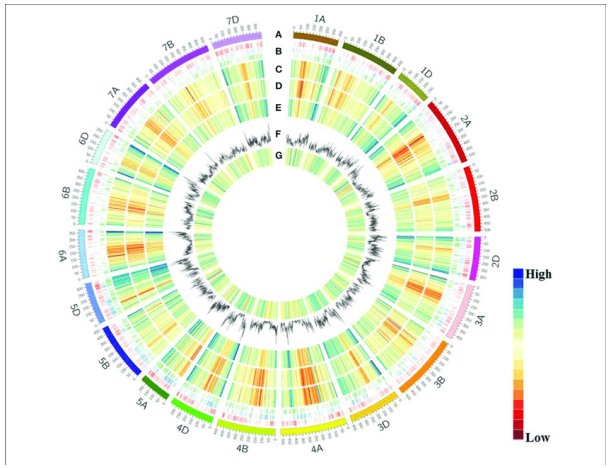
IWGSC RefSeq allows to connect identified genes and pathways associated with variation in traits with functional data from other crops and provide targets for modifying genome using genome editing.

Genes from other crops

Multi-OMICs datasets

Gene mapping

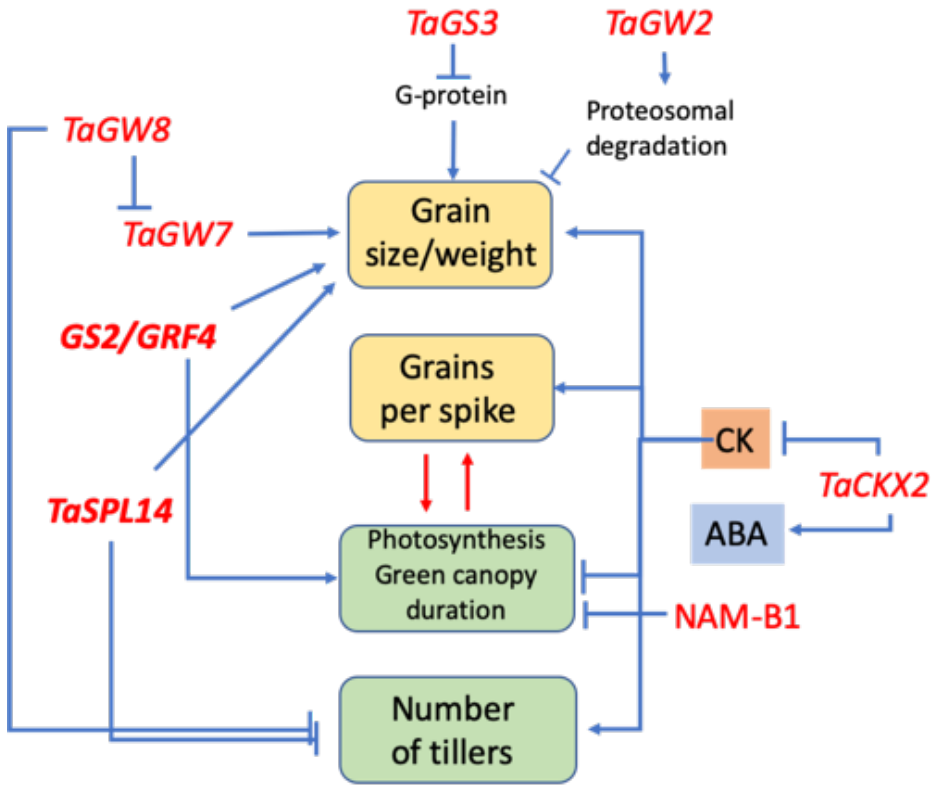
IWGSC RefSeq

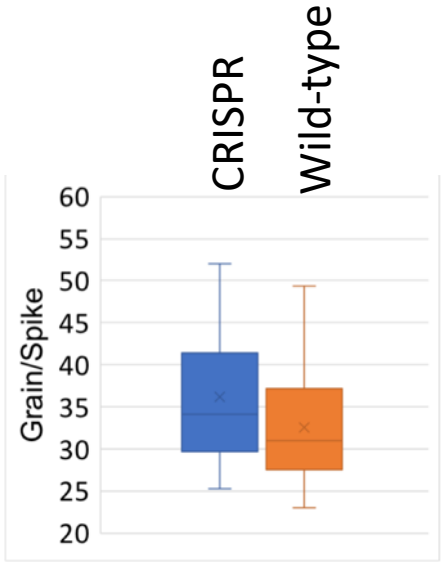


- Genes
- Regulatory regions
- Pathways

- Targeted CRISPR/Cas9-induced variation
- EMS mutagenesis

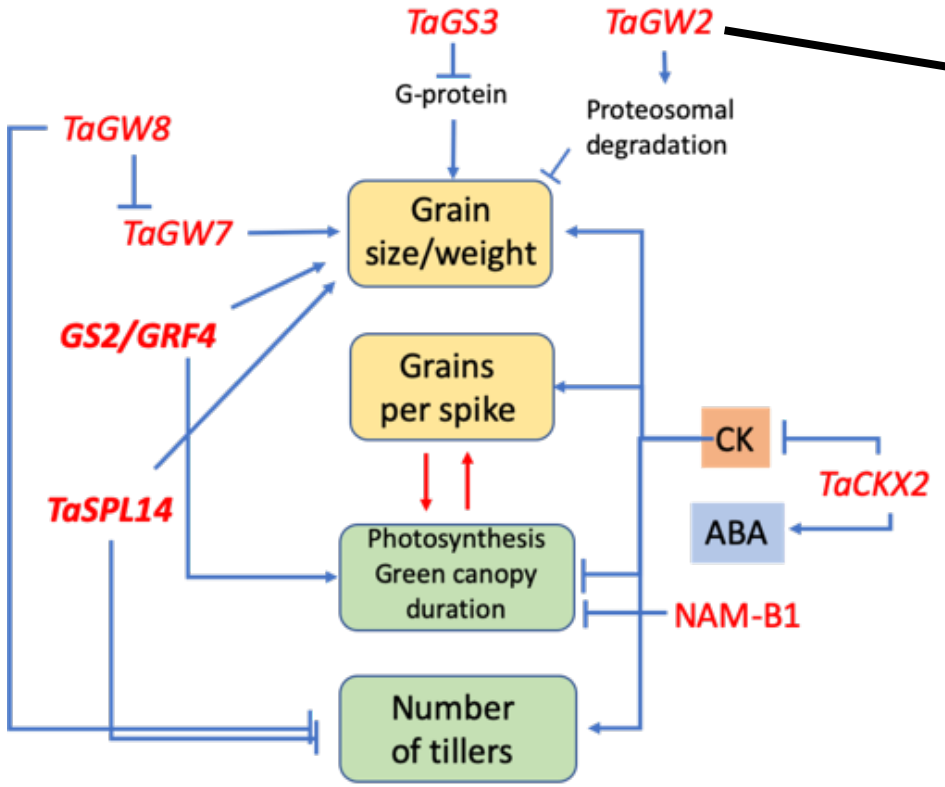
Knowledge of critical elements of biological pathways and their modes of interaction opens possibilities for engineering traits by CRISPR-based editing and creating alleles with specific effect on network

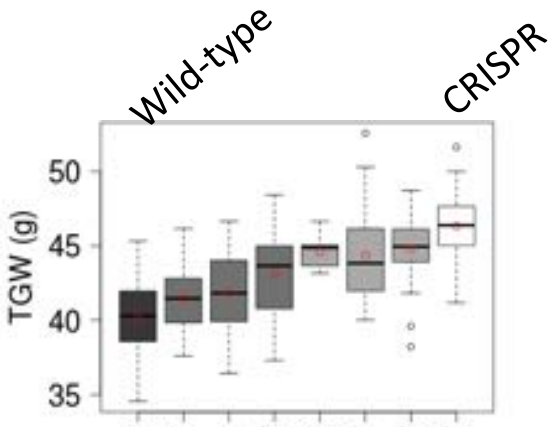




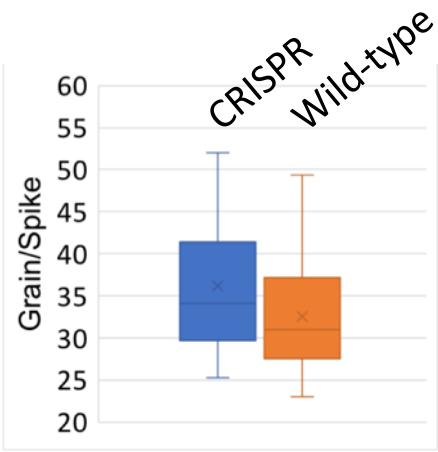
Increase # grains/head

Knowledge of critical elements of biological pathways and their modes of interaction opens possibilities for engineering traits by CRISPR-based editing and creating alleles with specific effect on network



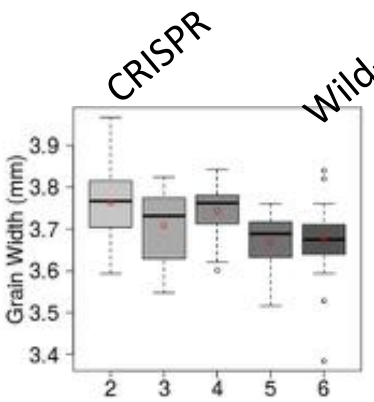


Increase grain size/weight

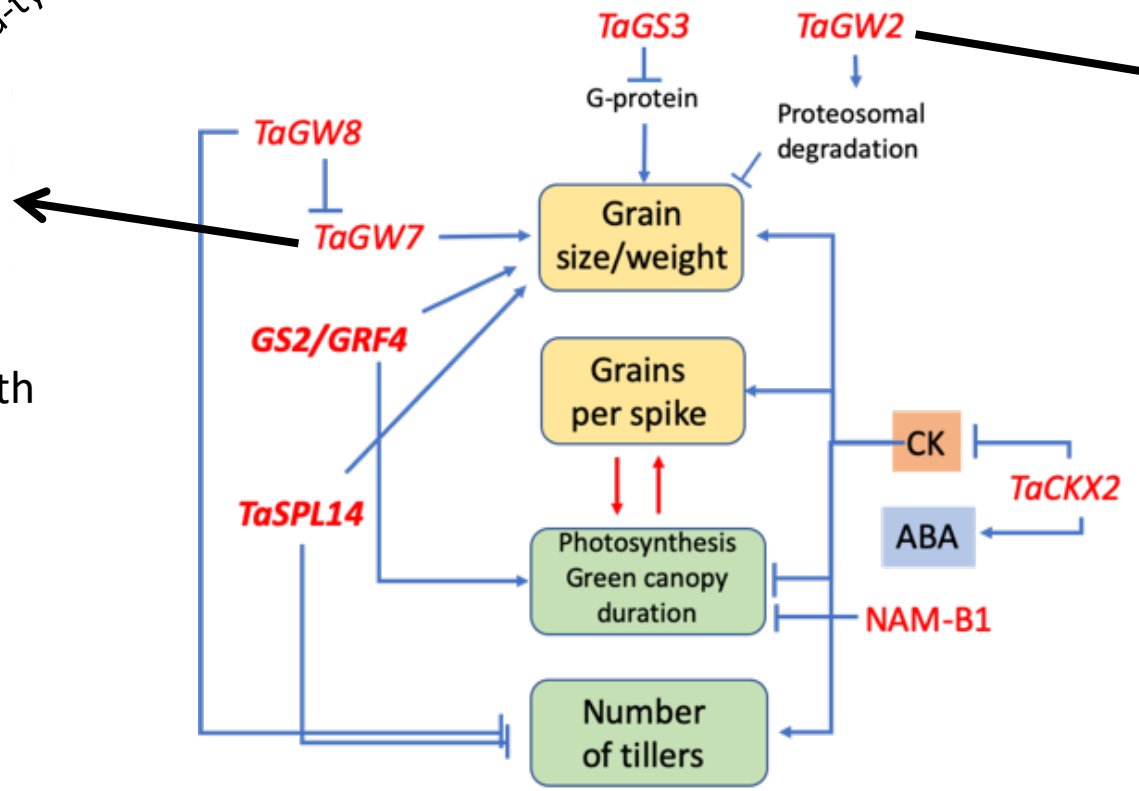


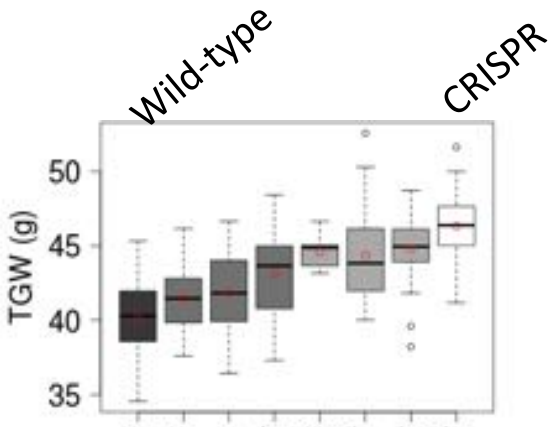
Increase # grains/head

Knowledge of critical elements of biological pathways and their modes of interaction opens possibilities for engineering traits by CRISPR-based editing and creating alleles with specific effect on network

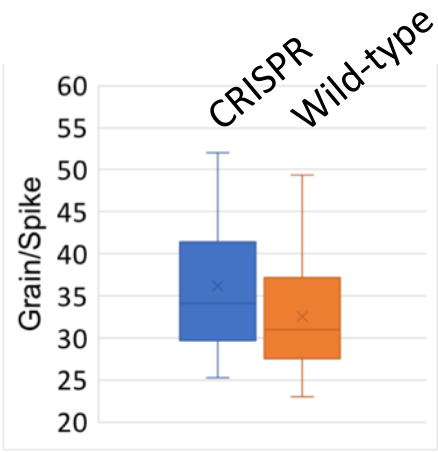


Increase grain width





Increase grain size/weight



Increase # grains/head

Conclusions

Joint analysis of eQTL and chromatin accessibility data provides a powerful tool for detecting regulatory regions of the large wheat genome

Enrichment of trait-associated variants in the regions of open chromatin or eQTL indicates relevance of these genome-to-phenome map's dimensions for understanding the genetic basis of major agronomic traits

Integrating additional 'omics' datasets into phenotypic trait analysis could help to identify critical genes and pathways controlling biological processes underlying these traits

This information could improve our ability to predict phenotypic outcomes of any particular genome, and select genomic targets for engineering desired traits

Acknowledgements

Akhunov Lab
Katherine Jordan
Fei He
Wei Wang
Andres Salcedo
Will Rutter
Shichen Wang
Dwight Davidson
Bliss Betzen
Alyssa Dunnivan
Zitong Yu
Moses Nyine

KSU Beocat
Dan Andresen

USDA CDL (MN)
Matt Rouse
Les Szabo
Yu Jin

USDA (ND)
Shiaoman Chao

KSU IGF
Alina Akhunova
Monica Fernandez de Soto
Jie Ren
Yanni Lun

UC Davis
Jorge Dubcovsky

University of Saskatchewan
Curtis Pozniak
Pierre Hucl
Krystalee Wiebe

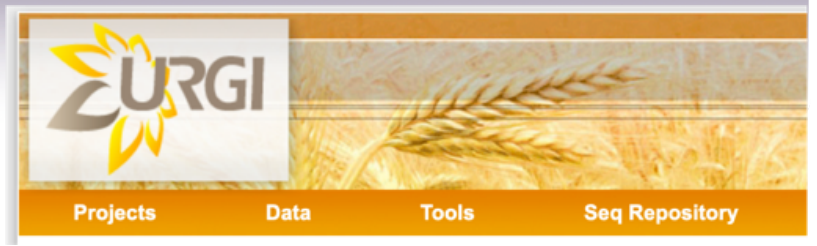
Agriculture Victoria, and, La Trobe University
Matthew Hayden
Raj Pasam
Fan Shi
Surya Kant
Gabriel Keeble-Gagnere
Pippa Kay
Kerrie Forrest
German Spangenberg
Ben Hayes
Hans Daetwyler
Josquin Tibbits

United States Department of Agriculture

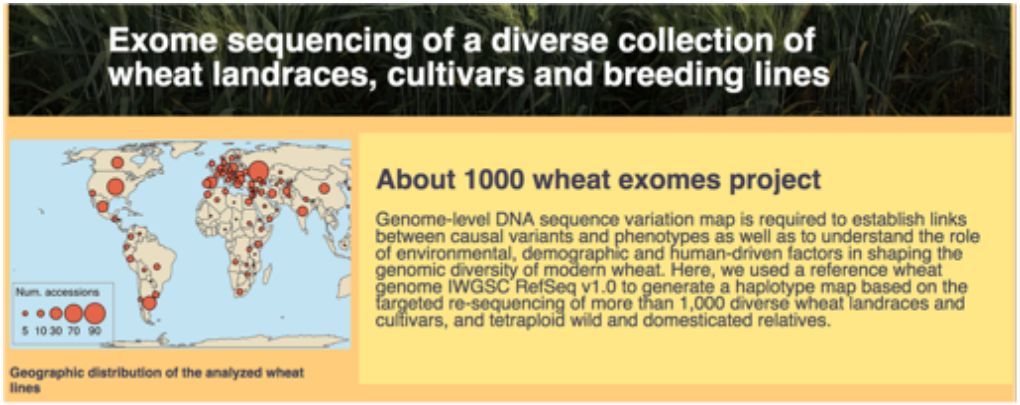
National Institute of Food and Agriculture

Resources

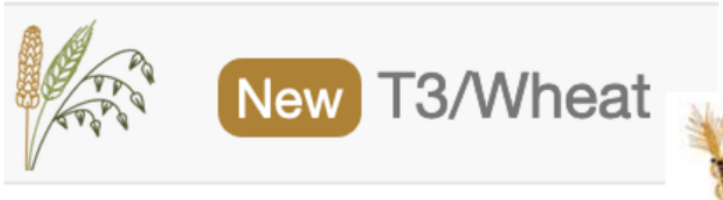
Chromatin accessibility data:
https://wheat-urgi.versailles.inra.fr/Seq-Repository/Chromatin-accessibility



1000 exome project: http://wheatgenomics.plantpath.ksu.edu/



Filtering SNPs using differential chromatin accessibility profiles:
T3 database (C. Birkett, J. Jannink)
https://wheat.triticeaetoolbox.org/genome/mnase.pl



GrainGenes (T. Sen)
https://wheat.pw.usda.gov/GG3/content/september-2020-mnase-chromatin-states-tracks-iwgsc-chinese-spring-genome-browser