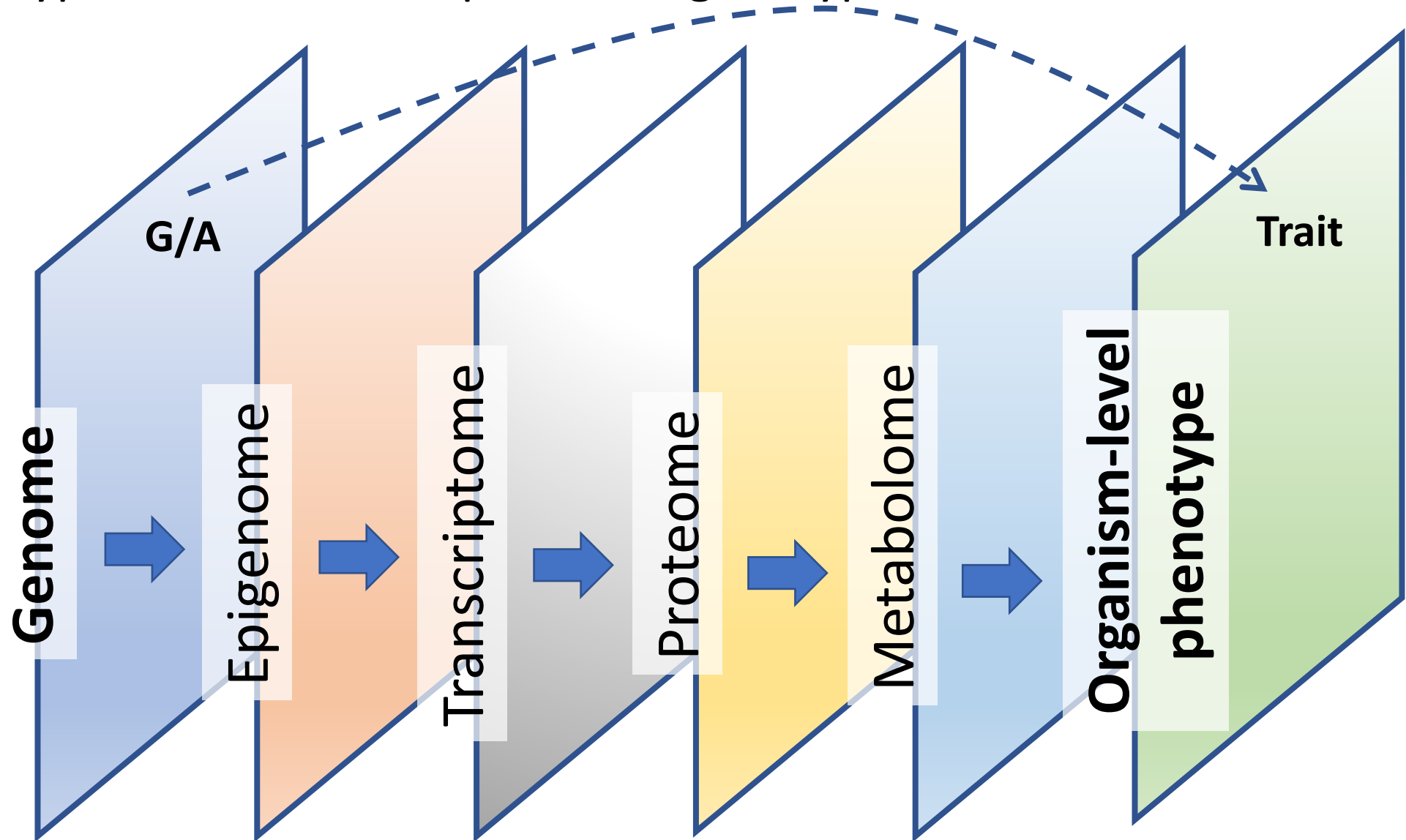


# Genomic resources

## WheatCAP

Eduard Akhunov  
Kansas State University

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



# Develop eQTL maps for wheat

200 genetically diversity wheat lines (worldwide sample)  
~ 3 million SNPs mapped to CS RefSeq genome

## **2-week seedlings**

8568 cis-eQTL -> 8837 transcripts

14,645 trans-eQTL -> 8500 transcripts

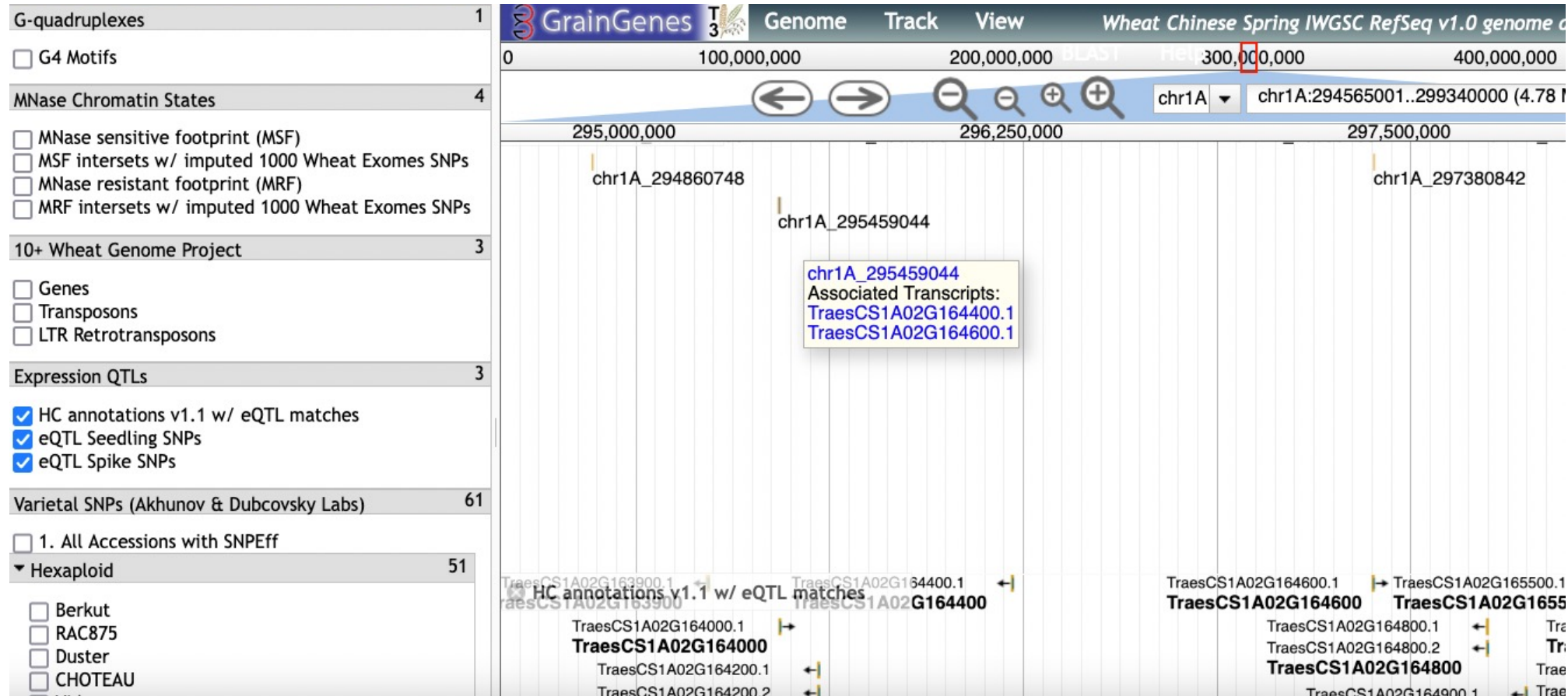
## **Post-anthesis spikes**

3172 cis-eQTL -> 3476 transcripts

9891 trans- eQTL -> 7250 transcripts



## eQTL maps for wheat seedlings and spikes in GrainGenes



<https://graingenes.org/jb/?data=%2Fggds%2Fwhe-iwgsc2018>

# Genomic resources for characterizing regulatory diversity

- ***200 accessions for studying regulatory diversity in the wheat genome***
  - US wheat breeding programs
  - Founders of NAM populations
  - Primary synthetics and wild wheat from introgression mapping projects
- ***Characterization of regulatory diversity in the wheat genome***
  - ***Tiered sequencing strategy***
    - *Few genomes at whole genome level*
    - *Regulatory sequence capture*
    - *Low-pass sequencing (Skim-Seq)*
    - *Updating PHG and Imputation*
- ***Tissues and developmental stages selected for Quant-seq analyses***
  - 10 tissues; developmental stages related to biomass production and grain yield
- **eQTL mapping and integration with other OMICs data (epigenome, chromatin acc.)**
- **Joint analysis to identify drivers of trait variation**

# Expanding Wheat PanGenome for mapping regulatory variants

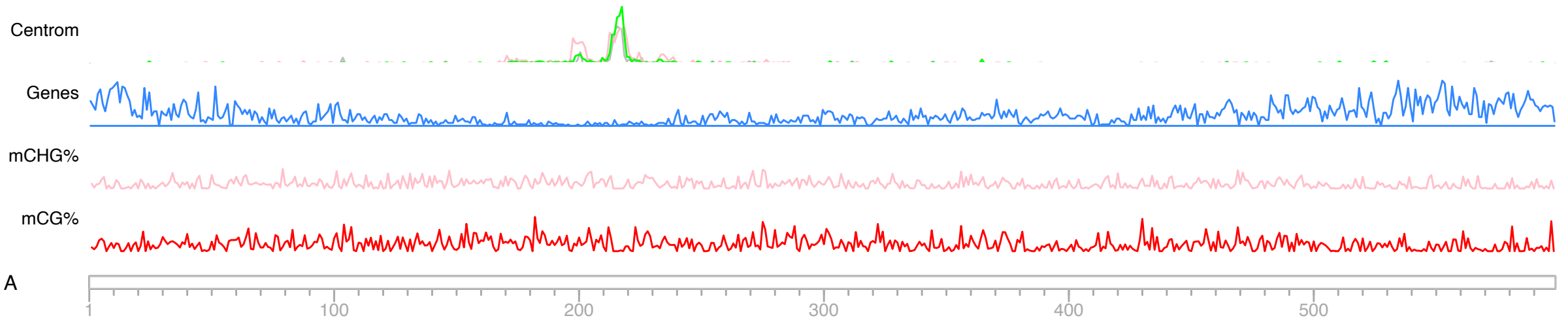
## Preliminary assembly based on PacBio data

Accessions	Total, bp	Contigs	N50
GF361	14395750022	9698	3219032
GF371	14477383491	4174	9462587
GF313	14902901758	23614	2885859
GF150	14485612752	4249	14547818
GF197	14441884101	4747	9831389
Kariega*	14689034153	5309	27611474
Fielder*	14676177301	4498	18988619

\*previously published PacBio data from public databases

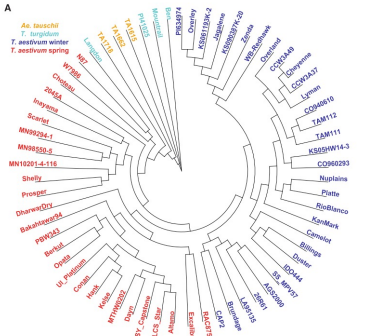
# Epigenomic variation in wheat

- 96 winter wheat cultivars from the US Great Planes breeding programs
- 6.8 Tb bisulfite-treated sequencing data
- 1.5x coverage of each genome
- Call methylated positions in the CHH, CHG or CpG contexts



# Updating Practical Haplotype Graph

## Version 1 (G3, 2022)



Whole-exome capture sequencing data  
from a diverse set of 65 wheat accessions

Imputation accuracy: 92-97%

## Version 2 (developed and will be shared with T3 for database-wide imputations)

Whole exome capture data (Genotyping Labs + KSU group)

Alignment of additional 250 wheat exomes to CS RefSeq v.2. (U. Yunusbaev, KSU)

~470 wheat lines representing majority of the US breeding programs

Construction of PHG v.2 (K. Jordan, USDA ARS)



# PHG version 2

- Reference ranges – Coordinates based on genes from RefSeq v2.1
- 472 taxa sequenced using Exome Capture
  - 90 Hard Winter Wheat – Southern Plains
  - 94 Northern Great Plains – both spring and winter habit
  - 95 Southern Region – soft winter wheat
  - 193 Pacific Northwest region
- Database footprint 146Gb
- In process of getting this to T3 database to use for imputation

# PHGv2 > 5 million segregating variants

