Structural regulation of the grapevine genome at global and local levels

Rachel Schwope[1,2], Aldo Tocci[2,3], Fabio Marroni[1,2], Mara Miculan[1,2], Michele Morgante[1,2]

1. University of Udine, Udine, Italy. 2. Institute of Applied Genomics, Udine, Italy. 3. Scuola Internazionale Superiore di Studi Avanzati, Trieste, Italy.

Grapevine is a valuable model for plant epigenomes

Grapevine (*Vitis vinifera*) is an economically and culturally valuable crop worldwide. A desire to produce better grapes led to the sequencing and assembly of a reference genome, isolated from a Pinot Noir cultivar (Jaillon et al., 2007). Its 19 diploid chromosomes contain approximately 482 Mb of DNA, comprising large structural variants and over 30,000 genes.
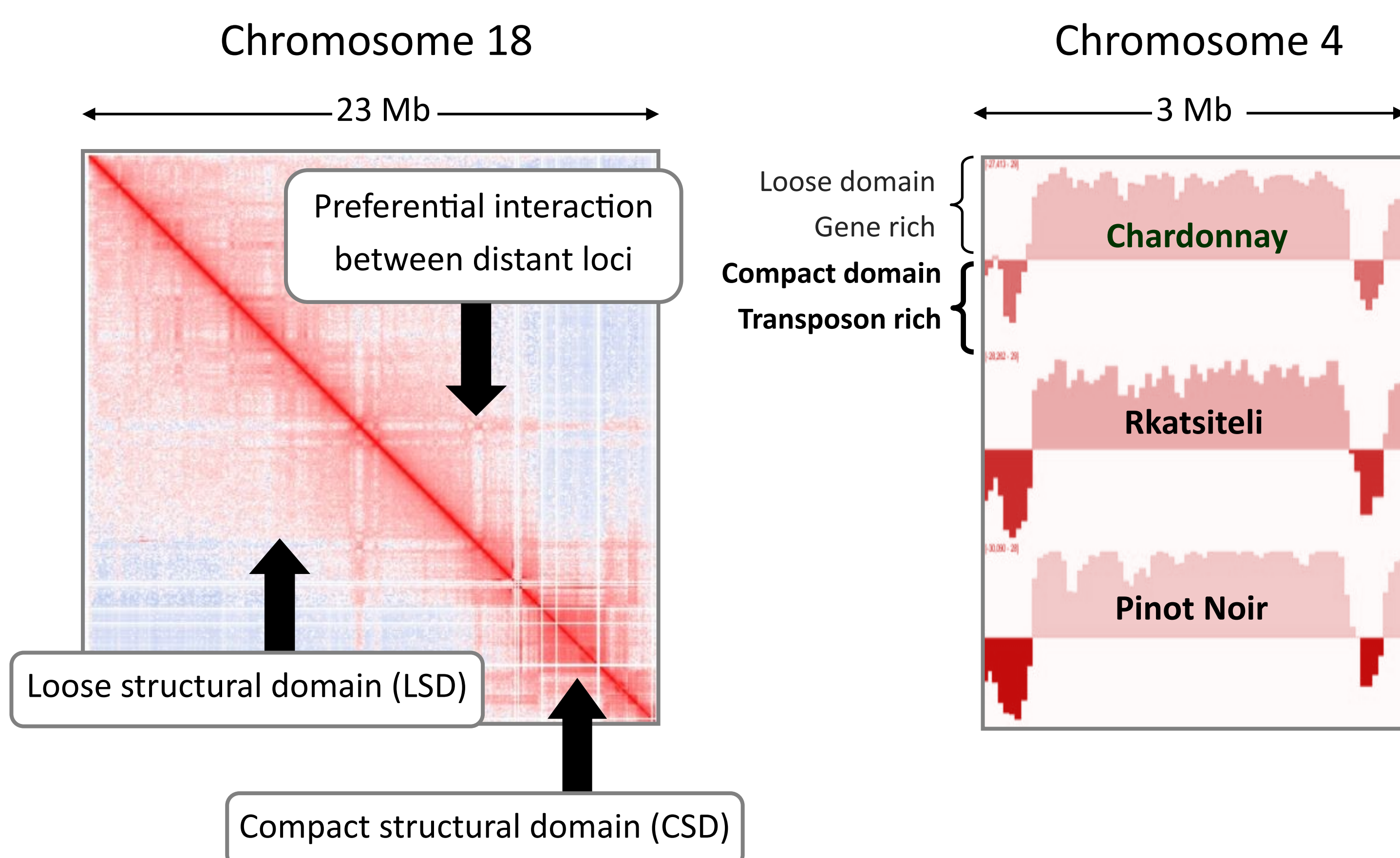
Given the presence of both gene-rich euchromatin and repeat-rich heterochromatin, a grapevine nucleus must organize its genome to maximize efficient gene transcription while minimizing deleterious effects of transposons. Our goals are to understand the structural factors that support such extensive genome regulation and to identify the epigenetic factors that maintain the nuclear architecture required for this task.

Experimental approach: investigate the grapevine genome structure at both global and local scale
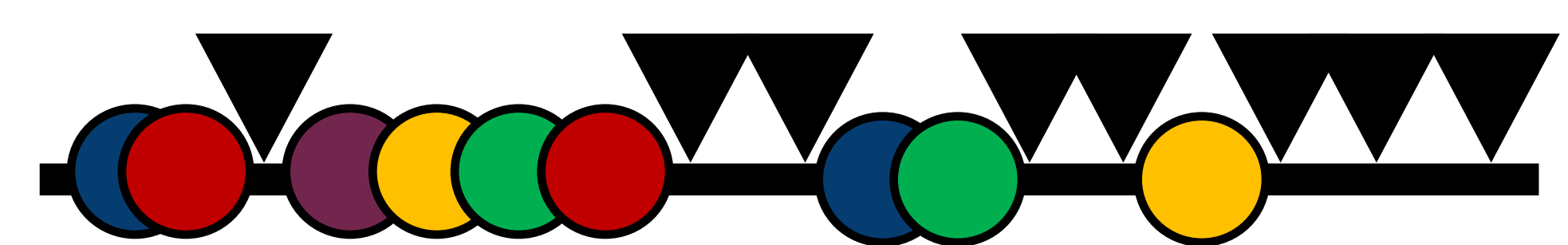
1. Nuclear space: Hi-C provides a three-dimensional map of the genome

Two dimensional genome

Interaction captured by Hi-C

Three dimensional genome
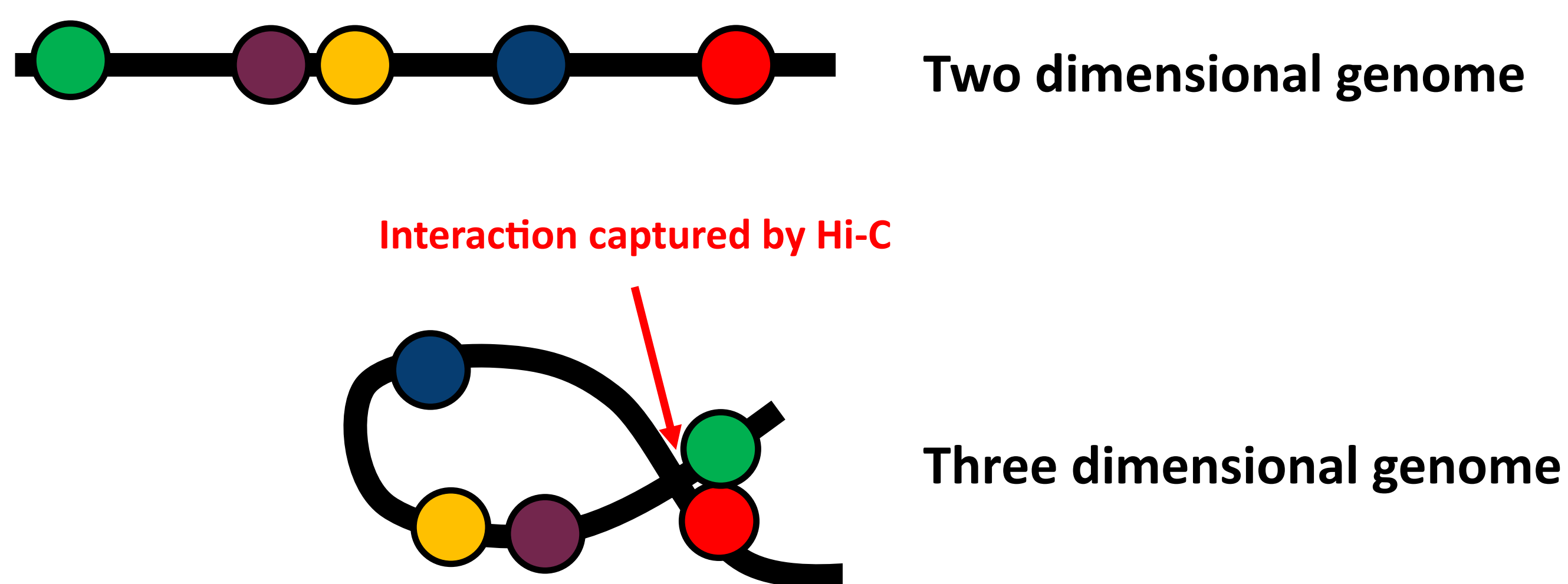
Hi-C principal components analysis reveals two types of chromatin, which are conserved across grapevine varieties

Chromosome 18

23 Mb

Preferential interaction between distant loci

Loose structural domain (LSD)

Compact structural domain (CSD)

Chromosome 4

3 Mb

Loose domain
Gene rich
Compact domain
Transposon rich

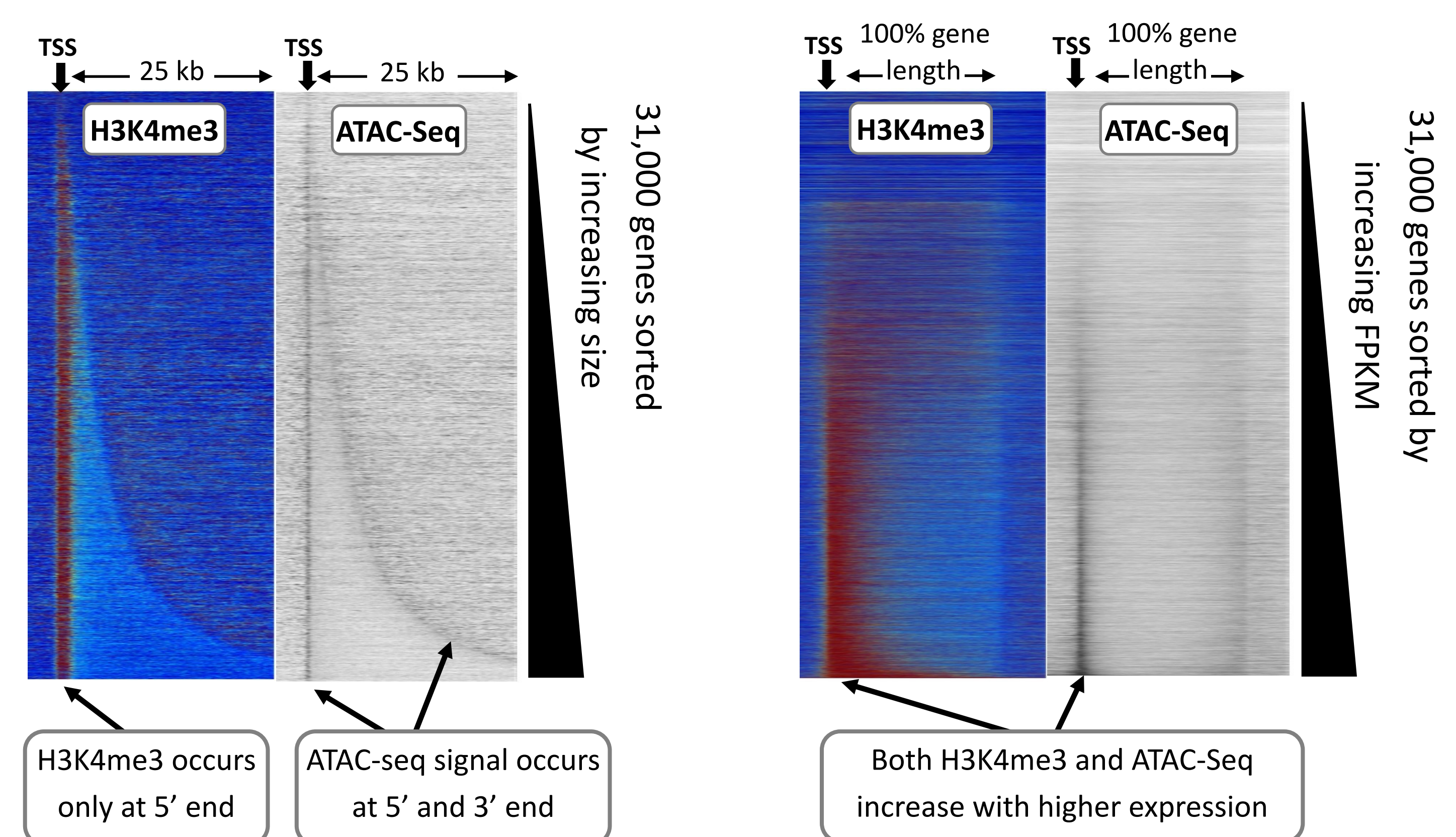**Chardonnay**

Rkatsiteli

Pinot Noir

2. Local DNA packaging: ATAC-Seq measures accessibility of chromatin
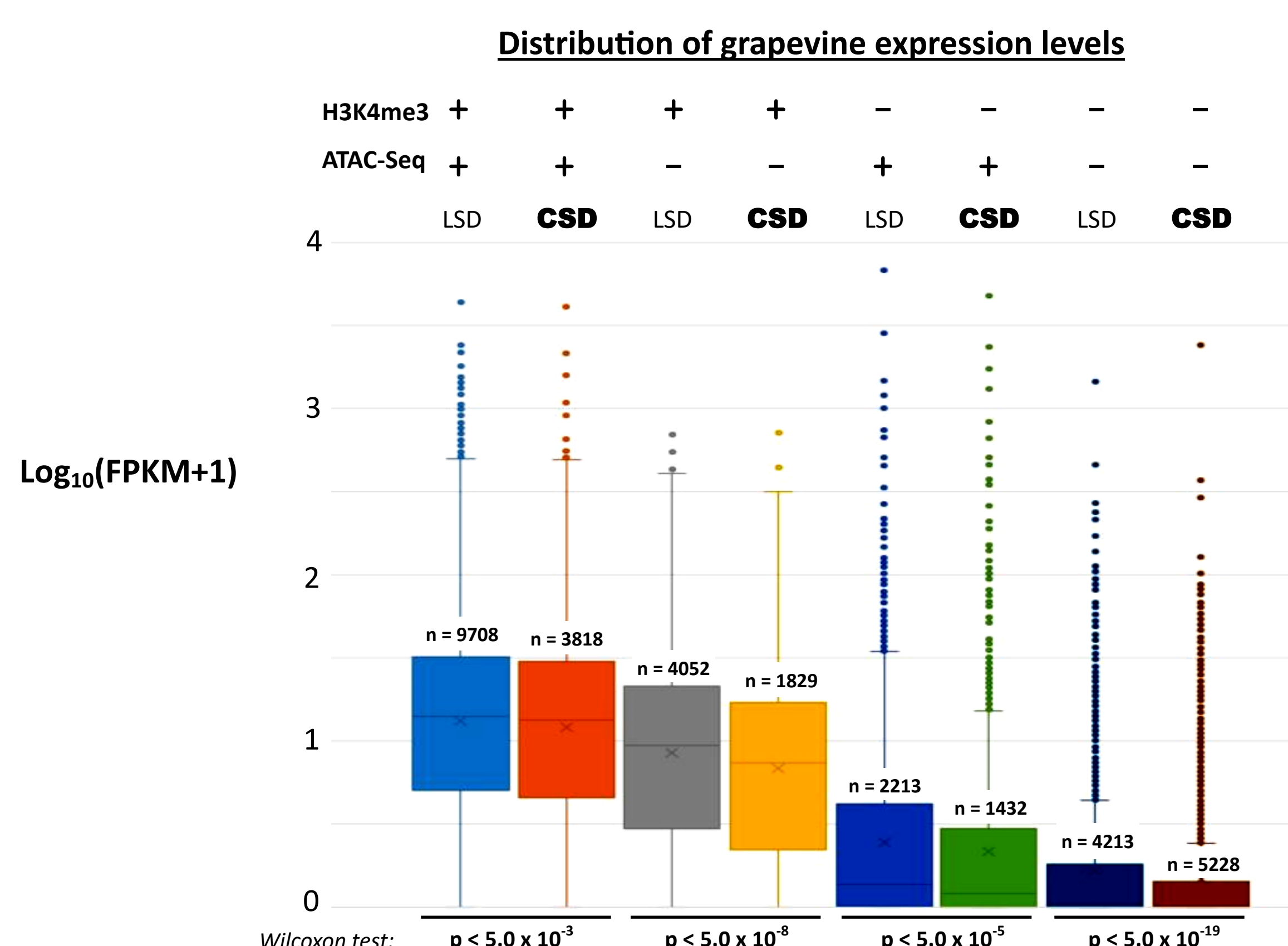
Lower DNA availability
Less Tn5 insertion

Higher DNA availability
More Tn5 insertion

ATAC-Seq shows that open chromatin occurs frequently at 5' and 3' ends of genes and is correlated with higher gene expression

TSS 25 kb TSS 25 kb
H3K4me3   ATAC-Seq
31,000 genes sorted by increasing size

TSS 100% gene length TSS 100% gene length
H3K4me3   ATAC-Seq
31,000 genes sorted by increasing FPKM

H3K4me3 occurs only at 5' end

ATAC-seq signal occurs at 5' and 3' end

Both H3K4me3 and ATAC-Seq increase with higher expression

Global and local chromatin characteristics produce a combined effect on gene expression

Distribution of grapevine expression levels

H3K4me3	+	+	+	+	−	−	−	−
ATAC-Seq	+	+	−	−	+	+	−	−
	LSD	**CSD**	LSD	**CSD**	LSD	**CSD**	LSD	**CSD**

$Log_{10}(FPKM+1)$

n = 9708 n = 3818 n = 4052 n = 1829 n = 2213 n = 1432 n = 4213 n = 5228

Wilcoxon test: $p < 5.0 \times 10^{-3}$ $p < 5.0 \times 10^{-8}$ $p < 5.0 \times 10^{-5}$ $p < 5.0 \times 10^{-19}$

Conclusions: Gene activity is influenced by multiple levels of structural regulation

The grapevine genome forms both loose (LSD) and compact (CSD) structural domains (Grob et al., 2014) that are conserved across varieties. Genes in the CSD are generally expressed at lower levels than genes in the LSD, although genes from either domain featuring both H3K4me3 and an open conformation are the most highly expressed groups. The absence of one or both of these features is correlated with a reduction in transcription, with the most significant domain-specific differences observed in groups lacking an open conformation. Our grapevine genome studies continue with 1) allele-specific analyses to examine structural variant roles, 2) tissue-specific studies to determine how cell differentiation affects genome organization, and 3) an enhancer search to identify regulatory elements in grapevine.