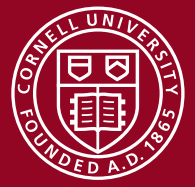
Cornell AgriTech

New York State Agricultural Experiment Station



Cornell CALS

College of Agriculture
and Life Sciences

Robot-Enabled Field Phenotyping of Grape Berry Cluster Color Using Multimodal Vision Foundation Models for Genetic Mapping

Yiyuan Lin[1], Madan Pandey[2], Lance Cadle-Davidson[3,4], Matthew Clark[2], Soon Li Teh[2], Yu Jiang[3]

1 School of Electrical and Computer Engineering, College of Engineering, Cornell University, Ithaca, NY, USA
2 Department of Horticultural Science, University of Minnesota, Minneapolis, MN, USA
3 School of Integrative Plant Science, Cornell University, Geneva, NY, USA
4 Grape Genetics Research Unit, United States Department of Agriculture-Agricultural Research Service, Geneva, NY, USA

ASABE

Introduction

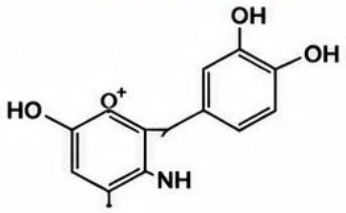
Why Grape Berry Color Matters

Fruit Quality & Wine Style



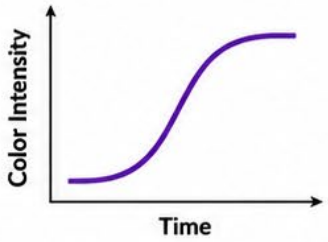
Berry color influences wine quality, flavor, and market value.

Pigment Accumulation (Anthocyanin)



Anthocyanin accumulation determines red / black coloration in berries.

Ripening Dynamics



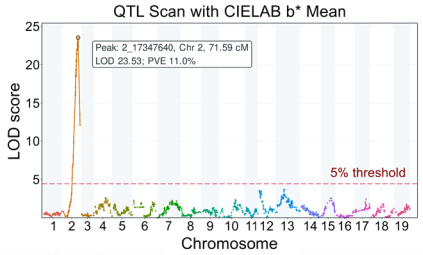
Berry color changes reflect ripening progression.

Genetic Variation



Substantial genetic variation underlies observable differences in berry color.

Breeding & QTL Mapping



Accurate color phenotypes enable QTL mapping and support breeding selection.

Challenges in Field Phenotyping



Occlusion

Leaves and canopies partially block clusters.



Varying Illumination

Changing sunlight and shadows cause large appearance variation.



Biological Variability

Substantial within- and between-genotype color variation.



Manual Scoring

- Labor intensive
- Discrete classes (white / red / noir)

Compresses variation into discrete score

Lab Measurements (Controlled Environment)

1) Imaging in Controlled Setup



- Controlled lighting
- Standardized background
- High-throughput imaging

OR

2) Instrumental Measurements



- Destructive sampling
- Time-consuming
- Post-processing

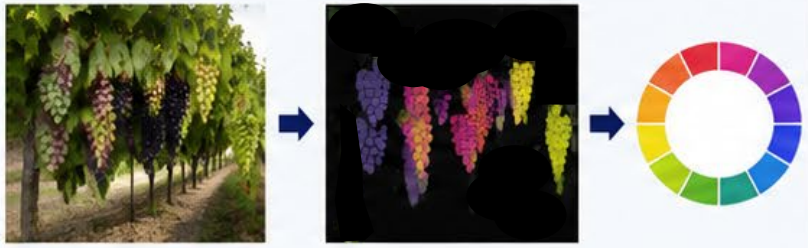
Can capture Substantial variation, but are not scalable

Overarching Goal and Specific Objectives

Our Overarching Goal

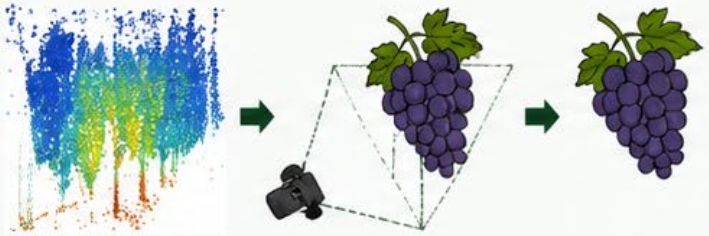
Develop a **scalable field-based phenotyping workflow** that transforms vineyard imagery into **biologically meaningful vine-level color phenotypes** for genetic analysis.

1 Image-based Vine-level Phenotype Extraction



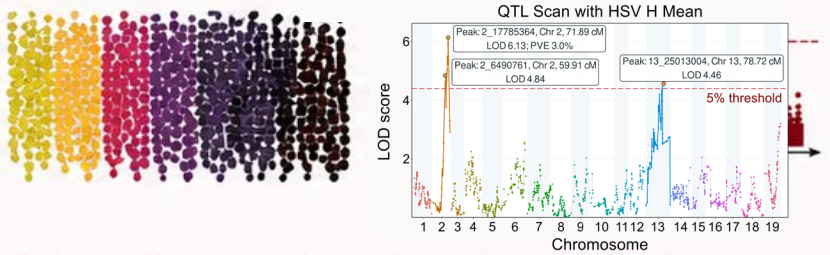
Develop a robust image-based workflow to segment berry clusters and quantify vine-level grape berry color phenotypes under realistic vineyard conditions.

2 Depth-aware Phenotype Stabilization



Investigate how depth-aware filtering influences phenotype stability and ensures sufficient biological sampling support across vines.

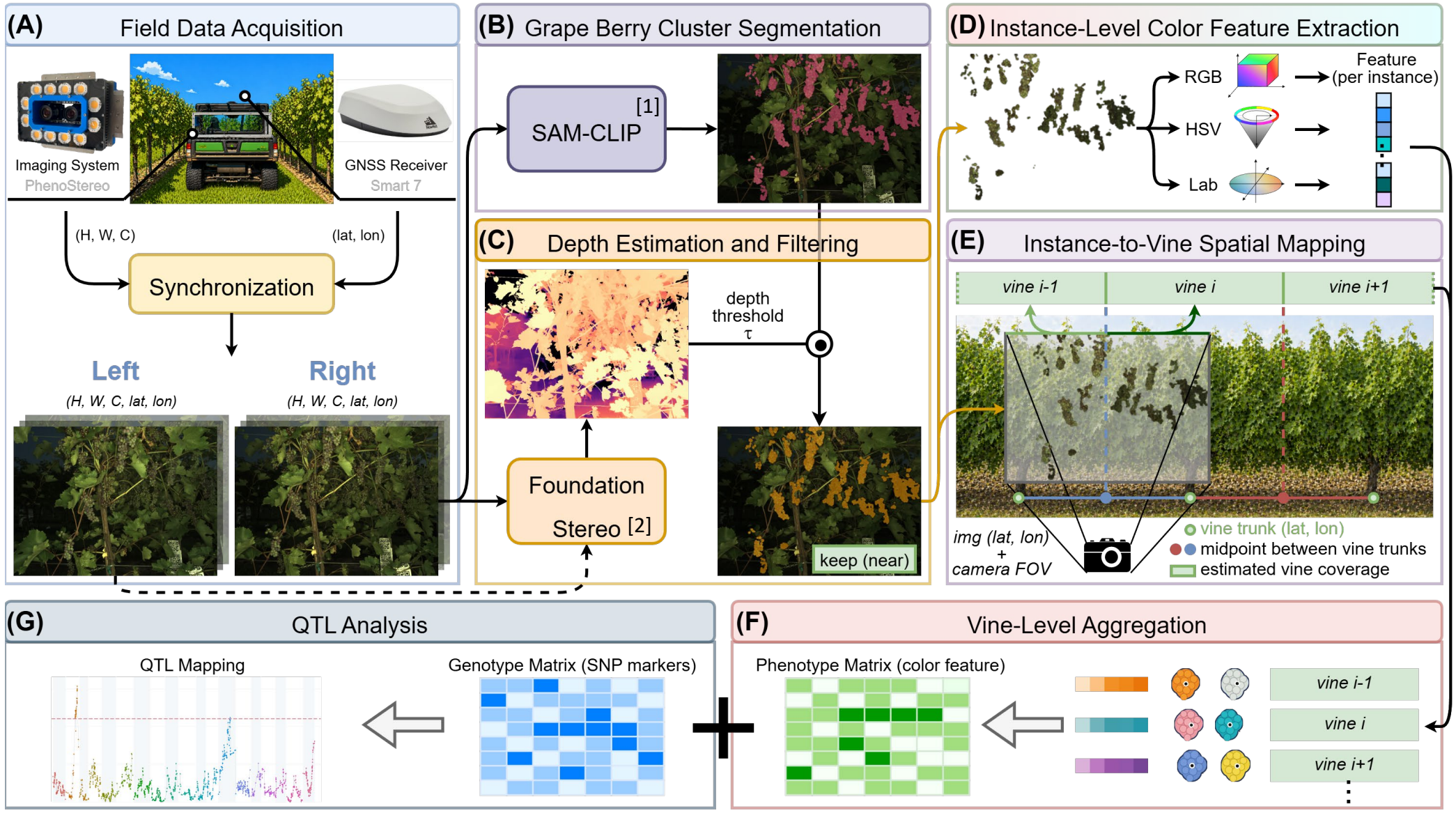
3 Continuous Phenotypes for Genetic Analysis



Evaluate whether continuous image-derived color phenotypes can improve the detection of QTL and capture biologically meaningful genetic variation.

Together, these objectives support scalable integration of field imaging, quantitative phenotyping, and grape genetics for future vineyard breeding and genetic improvement.

Overall Pipeline

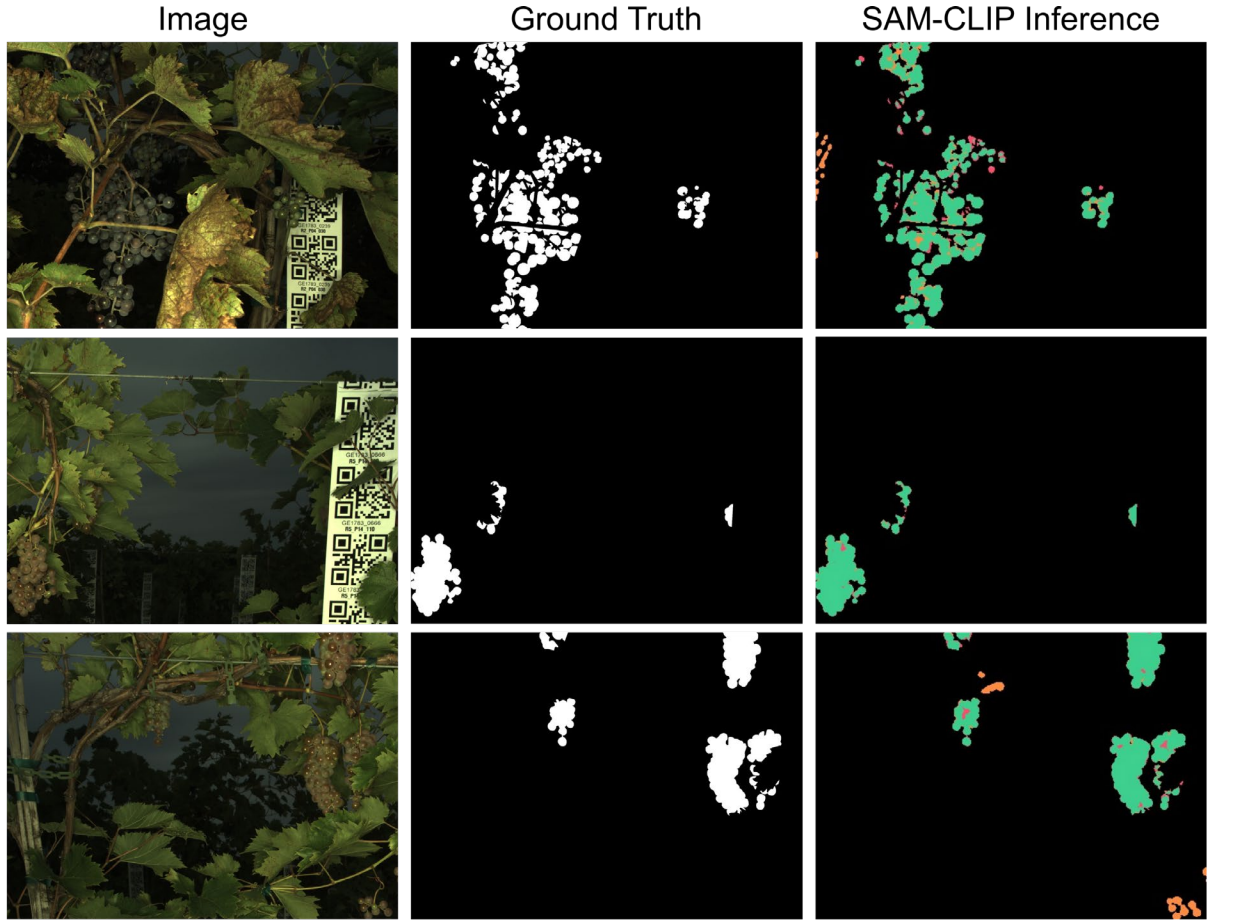


[1] Lin et al., *Effective integration of vision foundational models for semantic segmentation to quantify grape foliage powdery mildew infection*. ASABE AIM (2024)
[2] Wen et al., FoundationStereo: Zero-Shot Stereo Matching. CVPR (2025)

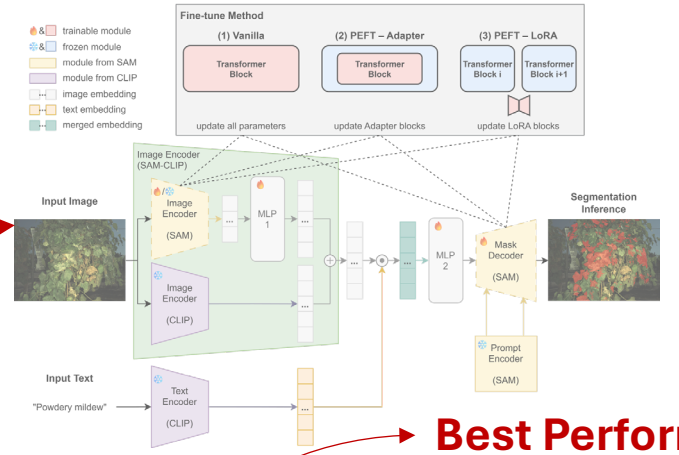
Segmentation Performance

Table 1. Details of datasets used in this study.

Dataset	Train	Val	Test	Total	Total Instances	Usage
UMN-Berry-2024	164	21	21	206	3064	Segmentation model development
UMN-Berry-2025	/	/	/	5853	/	Color extraction and QTL analysis



SAM-CLIP[1,2] Framework



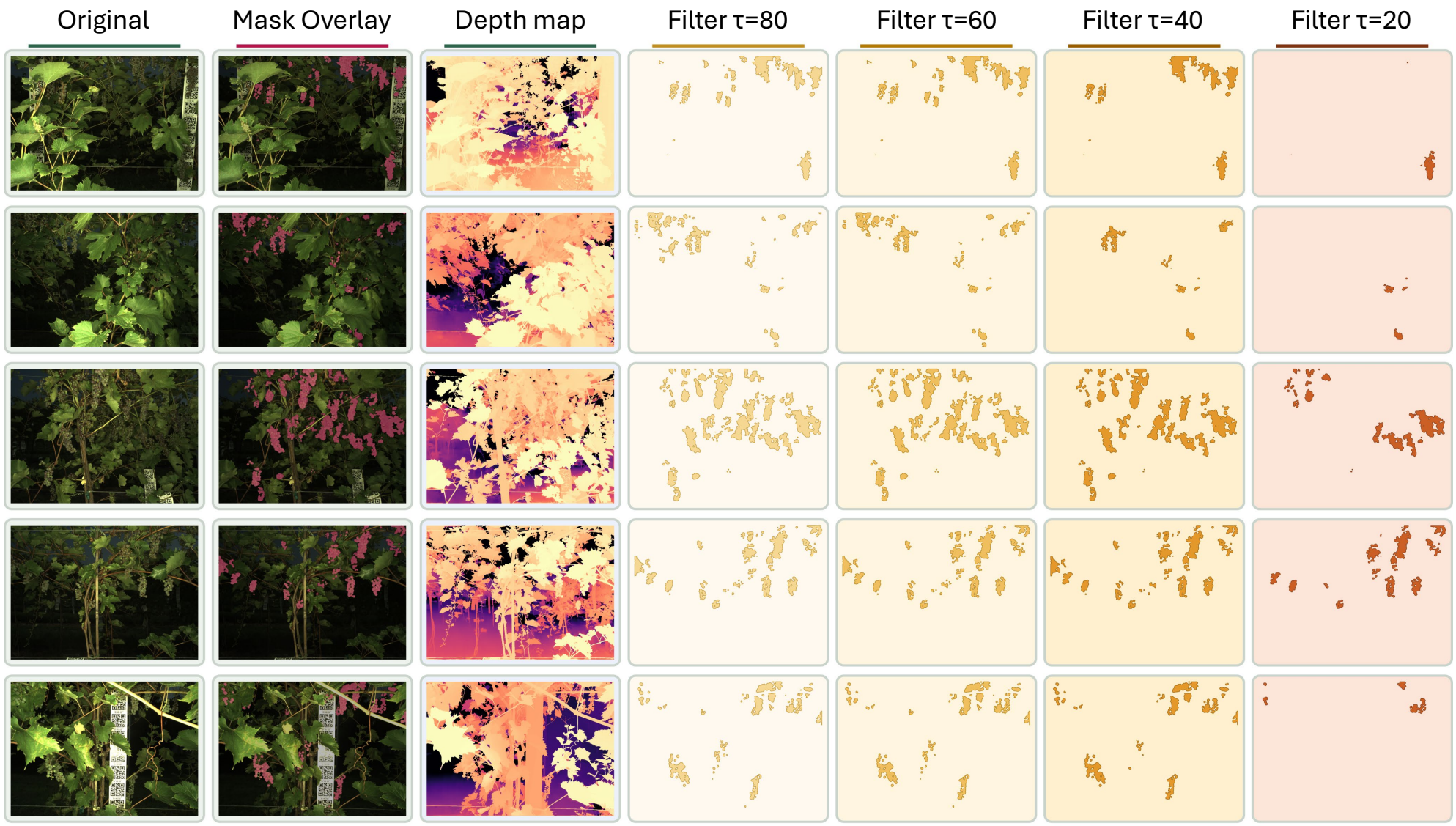
Best Performance!

Table 2. Model segmentation performance on UMN-berry-2024 test set.

Model	$mIoU^I$	$mIoU^D$	IoU^{C5}	IoU^{C10}	IoU^{C15}
SAM-CLIP	0.8526	0.8820	0.5049	0.5324	0.5749
SAM	0.7175	0.7644	0.3152	0.3436	0.4915
KNet	0.7117	0.7669	0.0660	0.1088	0.1930
Mask2Former	0.7097	0.7762	0.0000	0.0207	0.0999
DeepLabv3+	0.6832	0.7674	0.0310	0.1110	0.1911
SegFormer	0.6728	0.7019	0.0000	0.0430	0.1359

[1] Lin et al., *Effective integration of vision foundational models for semantic segmentation to quantify grape foliage powdery mildew infection*. ASABE AIM (2024)
[2] Lin et al., *Integrating Large Multi-Modal Models for Automated Powdery Mildew Phenotyping in Grapevines*. Artificial Intelligence in Agriculture (under review, 2026)

Depth Filtering and Vine-Level Feature Stability

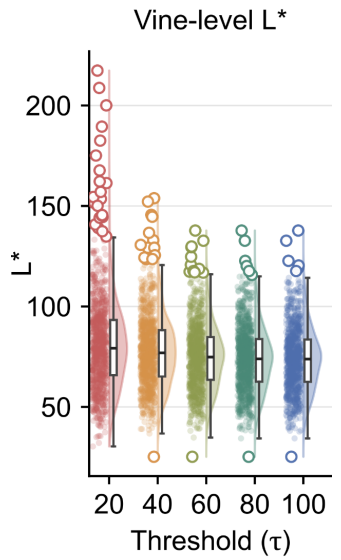


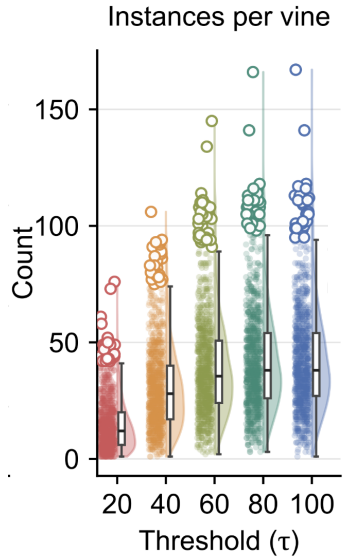
τ: retained depth percentile threshold

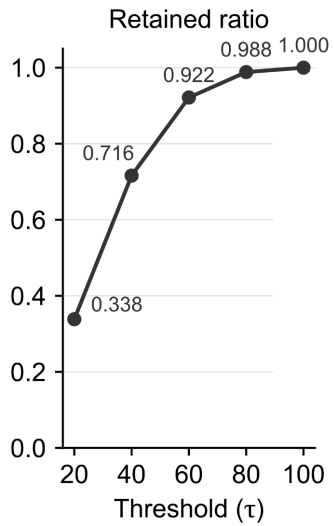
Evaluating Depth Filtering for Vine-Level Phenotyping

Depth filtering influences vine-level phenotype representation through observation reliability and biological sampling support.

Depth filtering is used to remove deeper canopy clusters that often receive suboptimal illumination and introduce additional noise.







Moderate filtering **preserves overall phenotype structure** across vines.

Aggressive filtering **reduces the number of retained berry** clusters per vine.

Lower τ substantially **decreases the overall retained instance ratio.**

These results show a trade-off between phenotype quality and biological sampling support.
τ = 60 is selected for downstream phenotype extraction and subsequent analyses.

Key interpretation

Large τ
- Retains deeper canopy clusters
- May introduce illumination-related noise

Small τ
- Reduces biological sampling density
- Fewer clusters contribute to each vine phenotype

Moderate τ
- Balances observation reliability and biological sampling support

Selected operating point

τ = 60

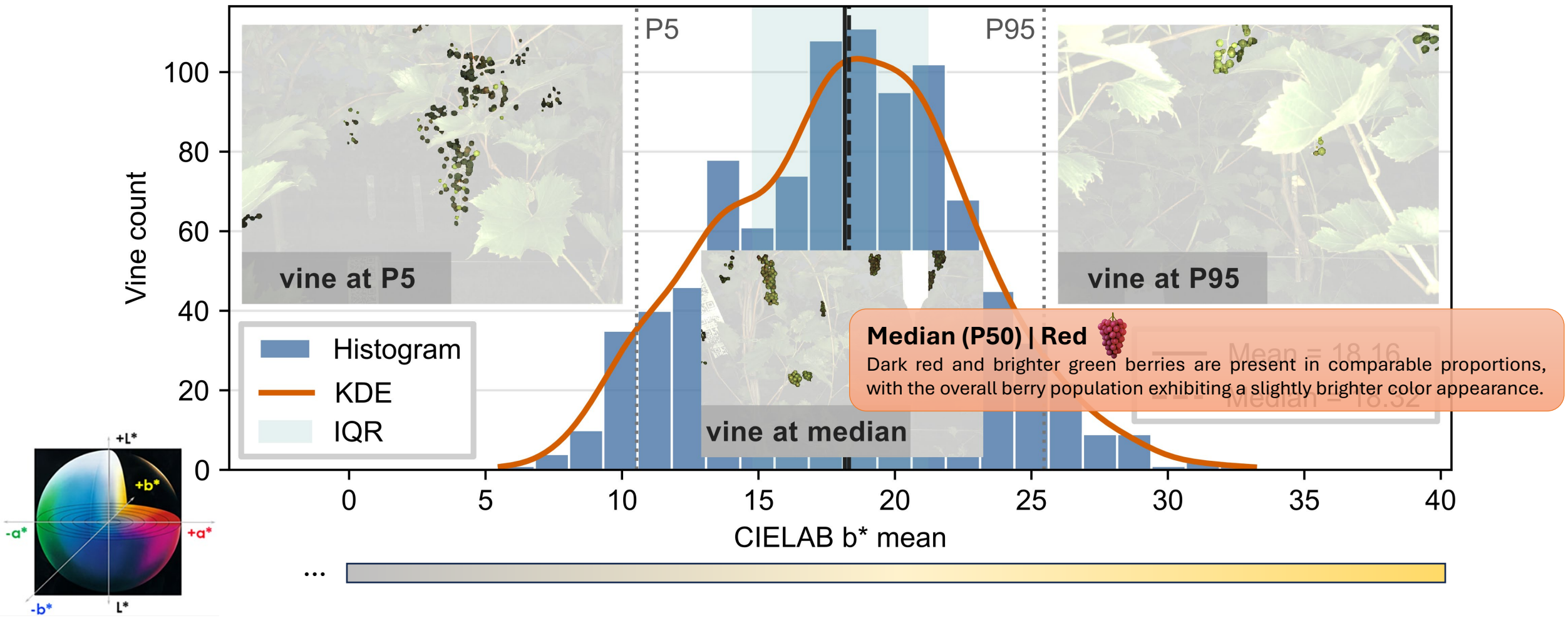
Provides a balanced compromise between observation reliability and biological sampling support.

We will evaluate how this choice affects downstream genetic analysis later in the presentation.

Vine-Level Color Features

P5 | Noir
Most berries show darker purple coloration, with only a small fraction showing brighter green coloration.

P95 | White
Most berries show brighter green coloration.



Median (P50) | Red
Dark red and brighter green berries are present in comparable proportions, with the overall berry population exhibiting a slightly brighter color appearance.

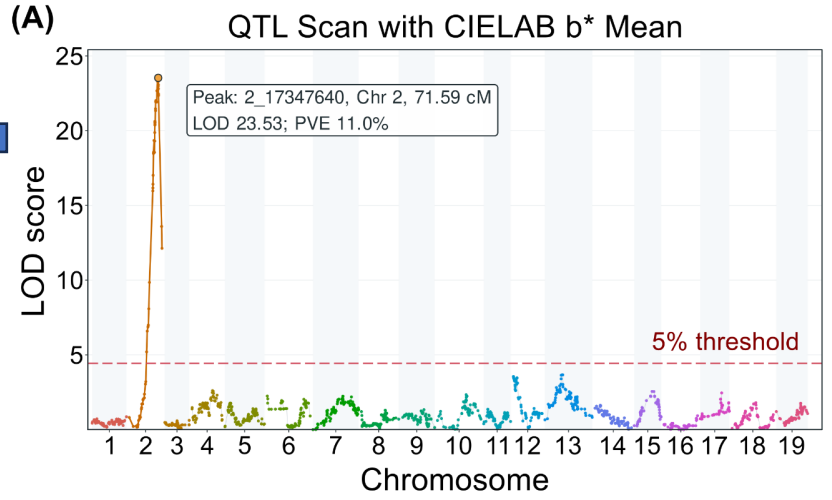
QTL Analysis Results

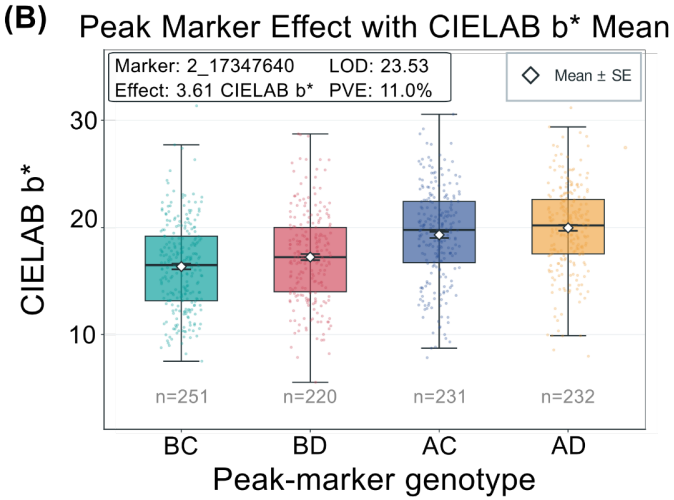
- Continuous image-derived phenotypes recover known QTL signals

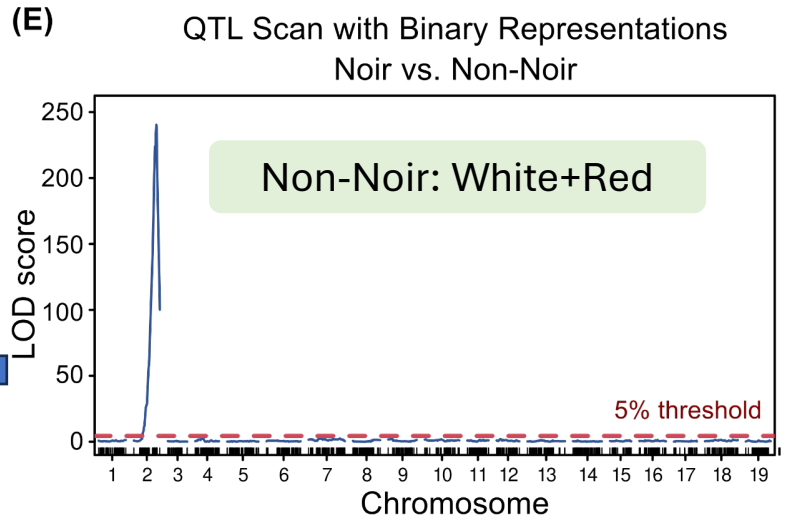
Continuous image-derived phenotypes
→ **chromosome 2** locus

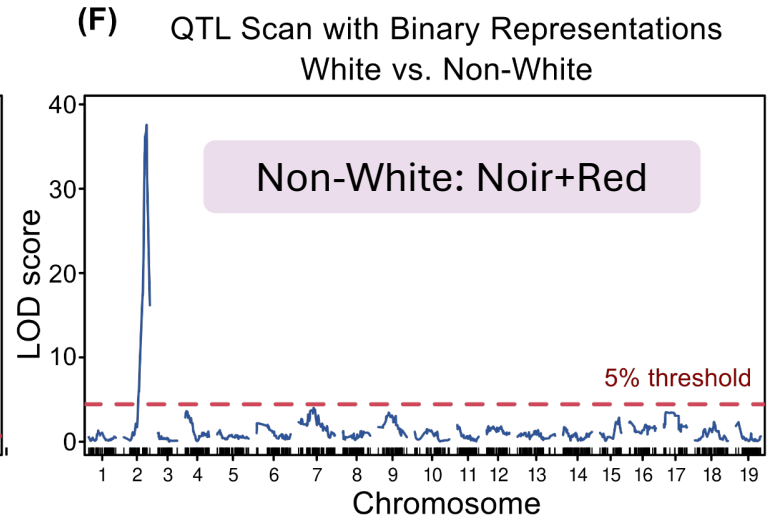
Known berry color locus[3]:
chromosome 2 (VvMYBA1 / VvMYBA2)

Human scouting binary phenotypes:
→ **chromosome 2** locus









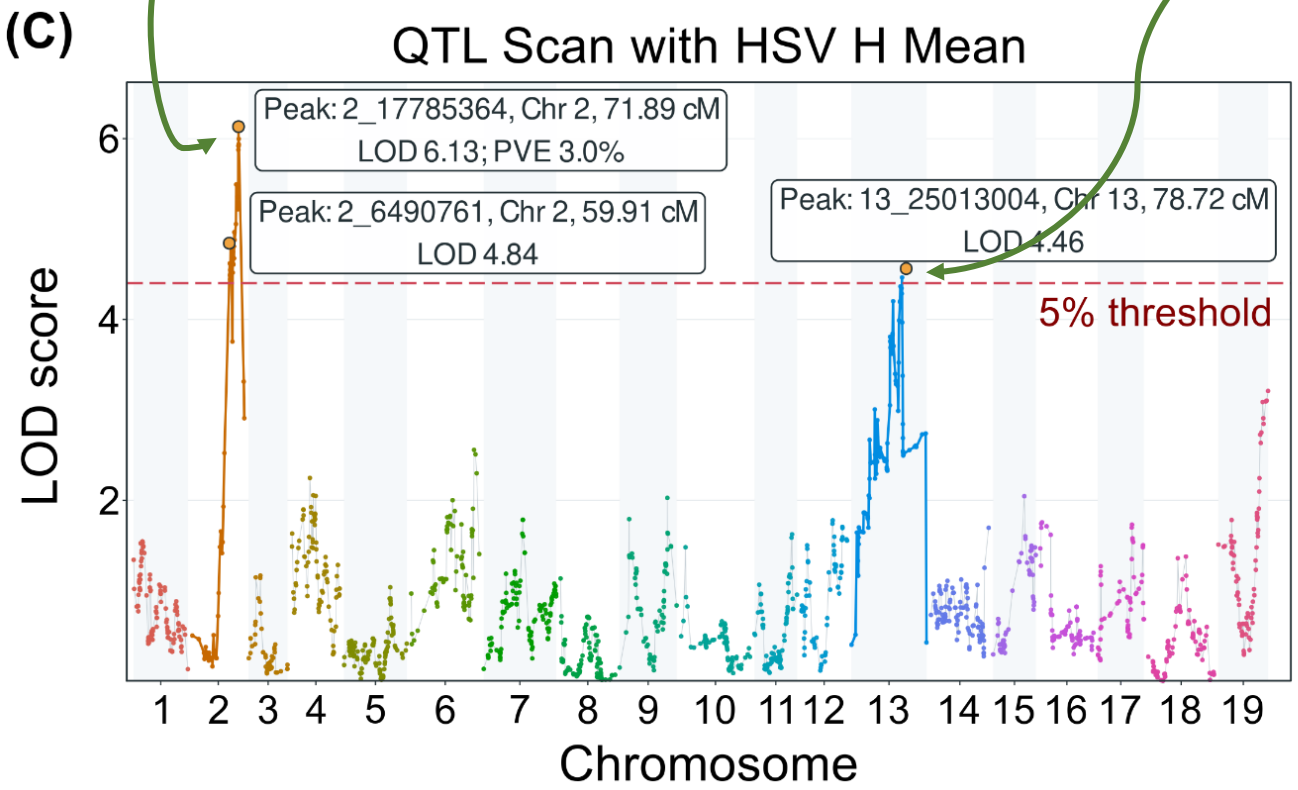
[3] Walker et al., *White grapes arose through the mutation of two similar and adjacent regulatory genes*. The Plant Journal (2007)

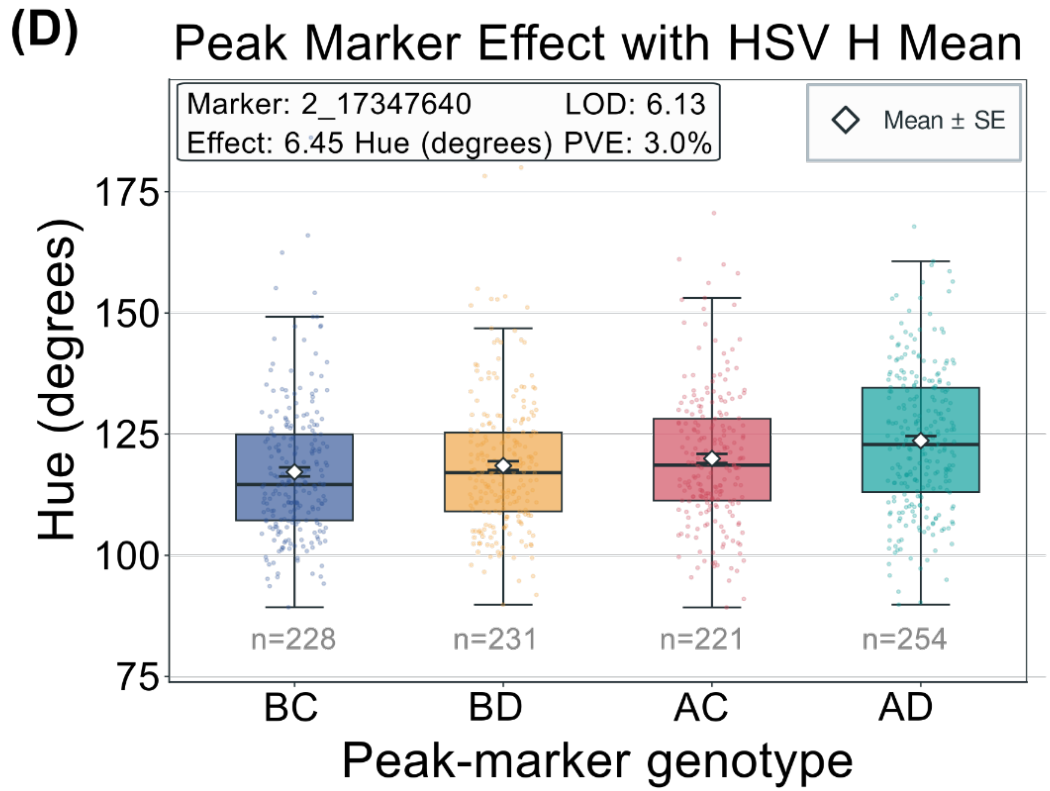
QTL Analysis Results

- Image-derived continuous phenotypes reveal additional genetic signals

Major QTL: chromosome 2

Additional Signal: chromosome 13





Beyond Segmentation: Toward Quantitative Phenotypes

Robust segmentation is an **essential foundation**, but our goal is to generate **biologically meaningful, genetics-ready phenotypes** from field imagery.

1. Segmentation (Enabling Step)

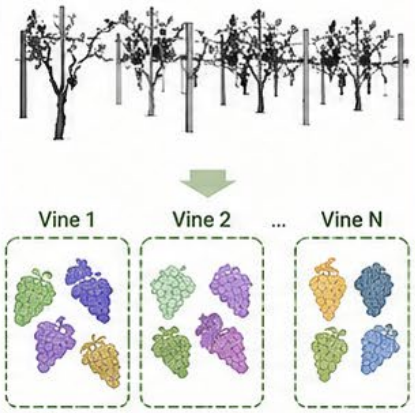
Detect berry clusters in complex field conditions



Robust mask quality (e.g., mIoU) is necessary.

2. Vine-Level Aggregation

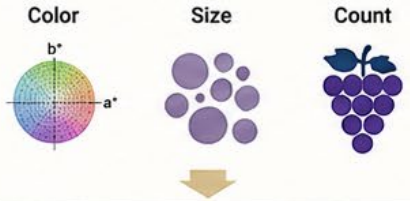
Aggregate detected clusters by vine and by visit



Move from object level to vine and visit level.

3. Trait Extraction

Extract biologically meaningful traits (e.g., color, size, count)

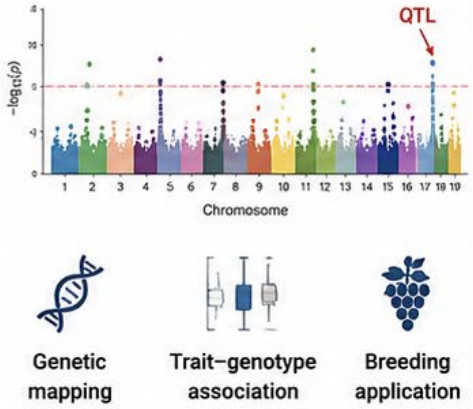


Vine	Visit	Lab	L	a	b	Size	Count	...
1	1	...	28.1	3.2	-5.6	15.3	42	...
1	2	...	27.4	2.8	-6.1	14.7	38	...
...	...	...	...	...	...	...	...	...
N	K	...	29.3	3.7	-4.8	16.1	45	...

Generate reproducible, quantitative phenotypes.

4. Downstream Genetics

Link phenotypes to genetics and plant performance



Produce genetics-ready phenotypes and biological insights.

Outcome

From field imagery to biological and genetic insights



Reliable, biologically meaningful, and genetics-ready phenotypes for grape improvement.

Our focus: not only better masks, but biologically meaningful, genetics-ready phenotypes.

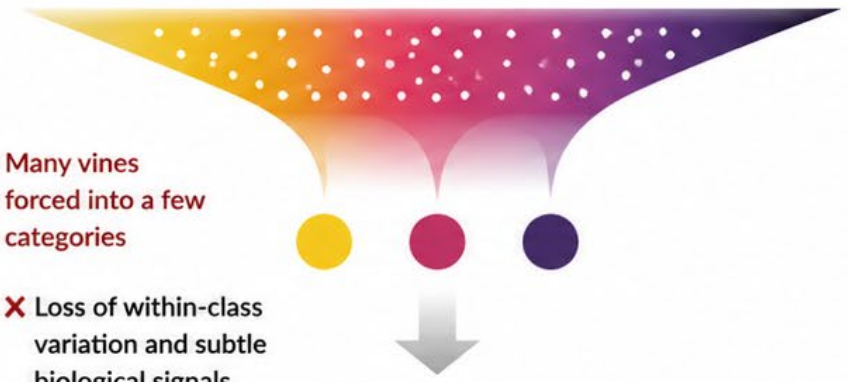
Why Continuous Phenotypes Matter

Traditional Categorical Phenotypes

Information Compression and Loss

Discrete labels (e.g., white / red / noir)

White | Red | Noir | ...



Many vines forced into a few categories

✗ Loss of within-class variation and subtle biological signals

Continuous phenotypes preserve more information and improve genetic sensitivity

Continuous Image-derived Phenotypes

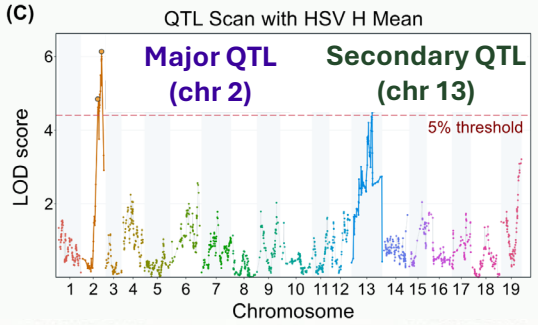
Preserves Gradual Variation and Biological Signals

Continuous color phenotypes (e.g., CIELAB b*, HSV hue)

Low — Captures gradual variation across the population — High

- Retains within- and between-genotype variation
- Preserves intermediate color states with biological relevance
- Improves detection power for QTL mapping

Greater Genetic Sensitivity



Continuous descriptors can reveal additional loci and capture complementary genetic effects.

Continuous image-derived phenotypes capture richer biological variation, preserve subtle signals, and improve the power and resolution of genetic analyses.

More informative phenotypes lead to more insightful genetics.

Acknowledgement

- This study was supported by the Specialty Crop Research Initiative Competitive Grant, Award No. 2022-51181-38240, of the USDA National Institute of Food and Agriculture.



USDA

Generative AI (ChatGPT, OpenAI) was used to assist in preparing presentation graphics and illustrative figures.
All scientific content, data analysis, interpretations, and conclusions are the author's own.

Thanks for your attention!



Yiyuan Lin

Cyber-Agricultural Intelligence and Robotics Laboratory
Email: yl3663@cornell.edu
Personal Website: https://yiyuanlinxx.github.io/
CAIR Lab Website: https://www.cair.cals.cornell.edu/

Yiyuan Lin

CAIR Lab

ASABE