Supplemental Information

Methods

Fermentation and sampling

Fermentation length was ~8 days (12 to 20 Sept) and temperature was maintained at 28°C (Miller et al. 2019). Pump-overs were performed every day, starting with six volumes per day the first two to three days, then decreasing by one volume each following day until completion of fermentation (total soluble solids [Brix] ≤ 0). Grapes were sampled from the top and bottom of the fermenter each day during fermentation, rinsed with water, and processed for imaging. Pulp and skin from the grapes were scraped off with a razor blade, then the skins were placed on a slide, hydrated with water, and sealed via coverslip with the outside of the skin facing up. Seeds were removed from the berry, rinsed with water, cleaned with a paper towel, sectioned with a razor blade, placed on a slide, hydrated with water, and sealed via cover slip.

The imaging of grape seeds was from a previous study (Miller et al. 2019), where Cabernet Sauvignon grapes (Vitis vinifera L.) were harvested in 2018, brought to the UC Davis Teaching and Research Winery, fermented in duplicate 2000-L pilot fermenters, then sampled for imaging. Two sets of grape samples were collected during this fermentation. The first set was used for epifluorescence microscopy (Miller et al. 2019). The other set was used for x-ray microcomputed tomography (x-ray µ-CT) of seeds only that was not reported. Data for this study was from analyses of the collected x-ray µ-CT images.

Model parameters

Ten models were trained for 500 epochs using a binary cross-entropy loss function and an Adam optimizer for stochastic optimization with the learning rate set to 0.001. Eighty percent of the image/annotation pairs were used for validation of the model during training (Rippner et al. 2022). The batch size was set to 1 for training due to graphics processing unit VRAM limitations when working with large images (2000 × 2000 pixels). Initially, models were trained based on four image feature classes (background, integument, endosperm, and seed pore space). Subsequently, a fifth feature class was added to the annotations (endosperm pore space) and used to train a new set of models. Results for the background, integument, endosperm, and seed pore space come from the best-performing model trained on four image feature classes. Results for the endosperm pore space come from the best-performing model trained on five image feature classes.

Model performance was evaluated based on the accuracy, precision, recall, and f1 score calculated after evaluating one image from each seed (nine total) that was not included in the training or validation data sets (Rippner et al. 2022). In our work, the accuracy, precision, recall, and f1 scores are defined as:

\[
\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \\
\text{Precision} = \frac{TP+1E^{-9}}{TP+FP+1E^{-9}} \\
\text{Recall} = \frac{TP+1E^{-9}}{TP+FN+1E^{-9}} \\
F1 = \frac{(TP+1E^{-9})}{\frac{1}{2}(TP+1E^{-9})+(\frac{1}{2}(FP+FN))}
\]

Where TP = true positive, FP = false positive, TN = true negative, and FN = false negative.

From the binary segmentations generated by the fully convolutional network, area and perimeter data for each grape seed component were calculated using the region properties function in scikit-image (van der Walt et al. 2014). The region properties function calculates area by summing the total pixels of a certain value in an image. Perimeter is calculated by the region properties function using four neighborhood connectivity for border pixel determination, which is then used to compute the contour of the pixels, generating the area. Total surface area for each fermentation stage replicate was calculated by summing the perimeter data. Similarly, area data from each slice was summed to determine the volume of the seed component for each fermentation stage and replicate.

The surface area and volume of each unconnected endosperm pore was calculated using the Porespy package in Python (Oliphant 2007, Gostick et al. 2019). To accomplish this, all individual binary slices of the endosperm pore space for a specific seed were sorted by slice number, stacked into a 3-D array, and saved as a 3-D .tif file using the Pil and Numpy packages in Python (van der Walt et al. 2014, van Kemenade et al. 2022). Individual, unconnected pores were then identified and labeled using the label function in scikit image (van der Walt et al. 2014). Individual volumetric data was calculated in Porespy using the region_volumes function with the mode set to ‘voxel’ (Gostick et al. 2019). This calculation was extremely RAM-intensive, requiring 117 Gb of RAM to compute. To reduce the RAM load, endosperm pores with volumes <1000 voxels were ignored. This calculation was checked against the volume calculated by the scikit image region properties function for the same data and was found to be identical (van der Walt et al. 2014). Agreement with the original total endosperm pores space calculation was >99%, with the newly calculated volumes being slightly less (<1%) due to the exclusion of pore volumes <1000 voxels.

Surface areas of each unconnected pore were calculated using the region_surface_area function in Porespy. This approach computes a mesh over each unconnected pore 3-D volume, which is then smoothed using a spherical element.
with a radius of 1 (Gostick et al. 2019). This approach leads to slightly different values than the original endosperm pore space calculation, but the magnitude of the differences remains the same, leading to the same statistical result outcomes as the original calculations.

Results

3-D pore network model

A link to the 3-D model is available on YouTube: https://www.youtube.com/watch?v=KkAzzw4aQQ0

Additional Discussion

In a modeled system where seeds and skins were extracted separately, there was more release of seed tannin, a flavonoid, than of skin tannin (Abi-Habib et al. 2021, 2022, 2023). When extracted simultaneously, higher levels of skin tannins were found in solution despite seeds having higher tannin content and extractability (Thorngate and Singleton 1994, Abi-Habib et al. 2023). Interestingly, in the same experiment, total polyphenol extraction was much lower than in the seed extraction. It is likely that when extracted alone, seed tannins were released more easily due to the absence of interactions with the grape skin's cell wall. This phenomenon could also be applied to the mass transfer theory of flavonoids from the inner to outer layer of a seed proposed previously (Miller et al. 2019), where a decrease in integument thickness might allow for the diffusion of endosperm tannins. Nonetheless, it appears that multiple diffusion pathways may play a significant factor in overall tannin release to wine. Additionally, other mechanisms have been identified as influencing flavonoid extraction, including re-adsorption of seed flavonoids by grape skin cell walls and reactions and precipitation with proteins, polysaccharides, and other polyphenols in solution (Abi-Habib et al. 2023).

References


Miller KV, Noguera R, Beaver J, Medina-Plaza C, Oberholster A and Block DE. 2019. A mechanistic model for the extraction of phenolics from grapes during red wine fermentation. Molecules 24:1275. DOI: 10.3390/molecules24071275


Supplemental Data for:

**Supplemental Table 1** Calculated surface area (SA) and volume (V) measurements from annotated x-ray microcomputed tomography images. SA and V were calculated using the Porespy package in Python for each seed replicate at three sampling time points: Field, Fermentation Start (FS), and Fermentation End (FE). Thickness (T) of the integument layer was calculated by dividing V over SA.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Rep</th>
<th>Integument</th>
<th></th>
<th></th>
<th>Endosperm</th>
<th></th>
<th></th>
<th>Endosperm pore space</th>
<th></th>
<th></th>
<th>Total pore space</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SA (mm²)</td>
<td>V (mm³)</td>
<td>T (m)</td>
<td>SA (mm²)</td>
<td>V (mm³)</td>
<td>SA (mm²)</td>
<td>V (mm³)</td>
<td>SA (mm²)</td>
<td>V (mm³)</td>
<td>SA (mm²)</td>
<td>V (mm³)</td>
</tr>
<tr>
<td>Field 1</td>
<td></td>
<td>60.5</td>
<td>4.0</td>
<td>0.65</td>
<td>28.1</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field 2</td>
<td></td>
<td>70.0</td>
<td>5.3</td>
<td>0.076</td>
<td>21.1</td>
<td>4.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field 3</td>
<td></td>
<td>68.2</td>
<td>5.0</td>
<td>0.073</td>
<td>21.9</td>
<td>4.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. ± Std. Dev.</td>
<td></td>
<td>66 ± 5 a</td>
<td>4.8 ± 0.7 a</td>
<td>0.27 ± 0.3 a</td>
<td>24 ± 4 ab</td>
<td>4.1 ± 1 a</td>
<td>13 ± 3 b</td>
<td>0.74 ± 0.3 ab</td>
<td>49 ± 2 a</td>
<td>2.1 ± 0.3 a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS 1</td>
<td></td>
<td>54.8</td>
<td>4.8</td>
<td>0.088</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS 2</td>
<td></td>
<td>71.3</td>
<td>6.7</td>
<td>0.094</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS 3b</td>
<td></td>
<td>66.5</td>
<td>6.0</td>
<td>0.090</td>
<td>0.045</td>
<td>0.00038</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. ± Std. Dev.</td>
<td></td>
<td>64 ± 9 a</td>
<td>5.8 ± 0.9 a</td>
<td>0.09 ± 0.003 a</td>
<td>21.0 ± 2 b</td>
<td>4.8 ± 1 a b</td>
<td>6.9 ± 8 b</td>
<td>0.10 ± 0.1 b b</td>
<td>37 ± 4 a b</td>
<td>1.3 ± 0.3 a b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FE 1</td>
<td></td>
<td>47.4</td>
<td>4.1</td>
<td>0.086</td>
<td>34.4</td>
<td>4.2</td>
<td>25.5</td>
<td>0.80</td>
<td>42.2</td>
<td>1.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FE 2</td>
<td></td>
<td>63.7</td>
<td>5.4</td>
<td>0.084</td>
<td>33.6</td>
<td>4.7</td>
<td>25.2</td>
<td>0.73</td>
<td>53.2</td>
<td>2.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FE 3</td>
<td></td>
<td>75.2</td>
<td>4.7</td>
<td>0.062</td>
<td>27.5</td>
<td>4.7</td>
<td>26.8</td>
<td>0.83</td>
<td>67.2</td>
<td>2.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. ± Std. Dev.</td>
<td></td>
<td>62 ± 1 a</td>
<td>4.7 ± 0.6 a</td>
<td>0.08 ± 0.01 a</td>
<td>32 ± 4 a</td>
<td>4.5 ± 0.3 a</td>
<td>26 ± 0.8 a</td>
<td>0.79 ± 0.1 a</td>
<td>54 ± 13 a</td>
<td>2.2 ± 0.6 a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| | | | | | | | | | | | | |

aLetters associated with average SA, V, and T data represent significant differences of these averages between sampling time points.
bSA and V measurements of Replicate 3 FS seeds were measured and reported but results were not incorporated into the average SA and V data for the endosperm, endosperm pore space, and total pore space, due to the seed missing its endosperm.
Supplemental Data for:

Supplemental Figure 1 Annotation of an x-ray microcomputed tomography (x-ray µ-CT) grape seed image using a fully convolutional network (FCN) model. Comparison of original x-ray µ-CT grape seed images to human and computer-generated annotations: A) individual image slice from an x-ray µ-CT scan of a grape seed from Fermentation End samples; B) human annotation of the seed components; C) FCN model prediction of grape seed components; and D) FCN model prediction of the individual seed components with endosperms, connected in 3-D space, uniquely colored.
Supplemental Figure 2  Representative slices from seed scans. Representative slices from each of the individual seed scans for the Field, Fermentation Start, and Fermentation End treatments.
Supplemental Data for:

Supplemental Figure 3  Endosperm porosity density plots. Surface area (SA) and volume (V) of individual pores located within the endosperm layer of each seed was used to assess pore size distribution across sampling time points using density plots created in RStudio software. Results for Replicate 3 of Fermentation Start seeds were removed for statistical analyses due to a hollow endosperm layer. Levene's test for homogeneity of variance and the Kruskal-Wallis test of analysis of variance were used to assess the normality of the data. Both tests confirmed the data was non-parametric (p < 0.001). A Dunn test was then used to compare the differences between sampling time points for SA and V of the endosperm pores. Significant differences are represented in letters within each density plot, along with a dotted line indicating the median SA and/or V for Field, Fermentation Start, and Fermentation End samples.