Supplementary Information

"Shedding Light on Carob Seeds: A Non-Destructive Approach to Assess Dehusking Efficiency using Diffuse Reflectance Spectroscopy and Kubelka-Munk theory"

As explained in the article, the second derivative spectra were presented in Figure 6 because they provide a more intuitive visualization of absorbance features, with peaks appearing as valleys at the same positions. However, for principal component analysis we used the first derivative, since its loadings are less complex and therefore easier to interpret than those from second derivative spectra, which contain more peaks and valleys. Importantly, both approaches yield similar results in the PCA scores plot (Figure 7), so the choice of first derivative data for PCA was made primarily for clarity of interpretation.

In this supplementary file, we present the complementary plots relatively to those shown in the article, that is,

Figure S1: the first derivative spectra (equivalent to Figure 6, but with the first derivative)

Figure S2: the PCA scores plot from the second derivative data (equivalent to Figure 7, but with the second derivative)

Figure S3: the PCA loadings plot from the second derivative data (equivalent to Figure 8, but with the second derivative and without the simultaneous plot of loadings for the raw reflectance data)

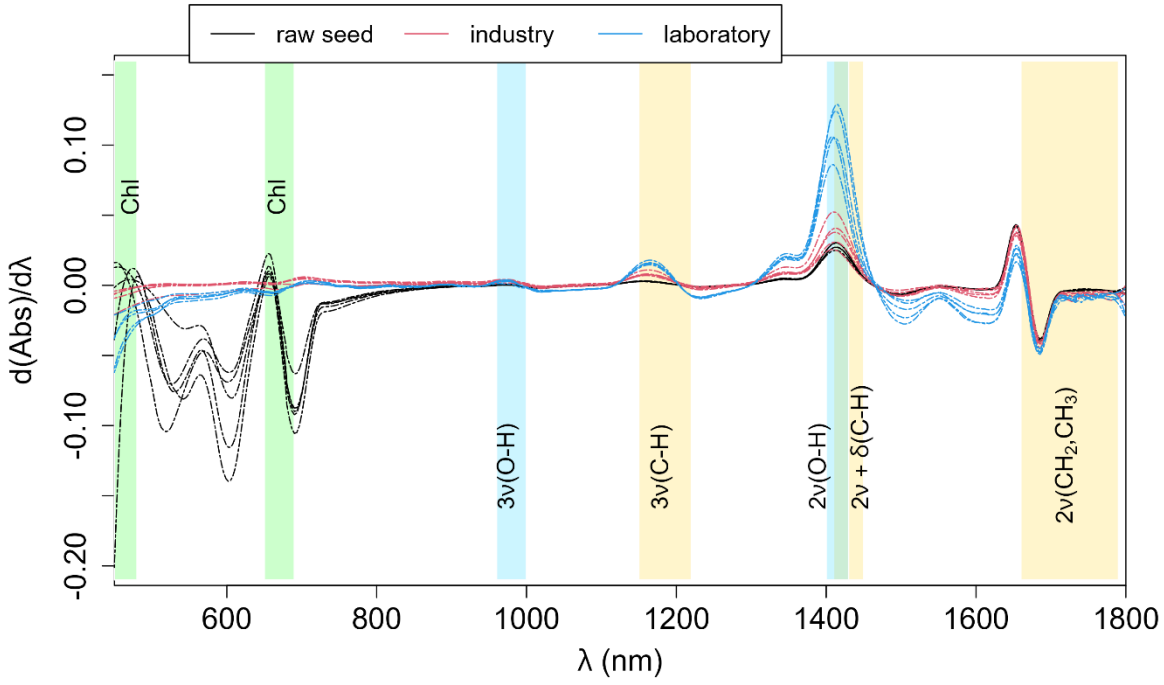


Figure S1. First derivative of the absorbance spectra for native seeds (black curves), commercially dehusked seeds obtained through the industrial process (red curves), and seeds dehusked using the method developed by us

(green curves). Five spectra per seed type are shown. The vertical bands show the principal absorption bands (see figure 1 for details).

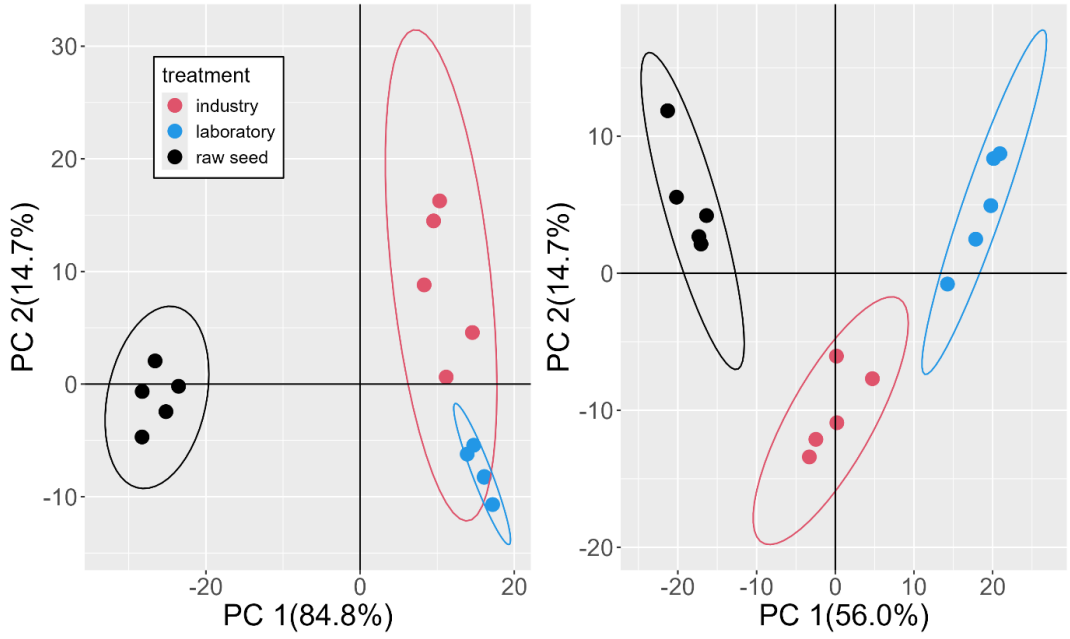


Figure S2. PCA analysis of the reflectance data. Left: from raw reflectance. Right: from the second derivative of absorbance data. The three clusters highlighted correspond to the native seeds (black cluster), commercial dehusked seeds obtained with the industrial process (red cluster) and dehusked seeds with our developed method (blue cluster). Each spectral variable was mean-centered and normalized to the standard deviation.

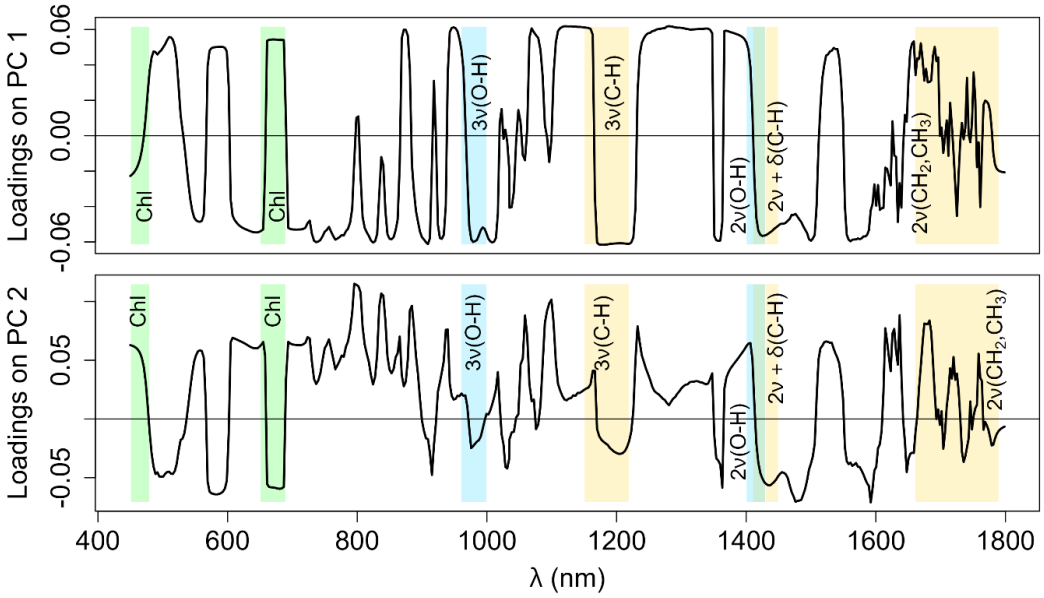


Figure S3. Loadings of the first (top) and second (bottom) principal components for the PCA decomposition of the second derivative of absorbance.