Molecular Composition of Soybean Seeds from Drought Susceptible and Tolerant Cultivars

Final Progress Report: 31 Jan 2024

This grant investigated the nutritional molecular composition of soybean seeds. In our study the drought susceptible cultivar Pana (drought susceptible) was investigated. There were two goals. The first goal was to build a metabolite database for each seed type using the combination of liquid chromatography and tandem mass spectrometry (MS/MS). The second goal was to compare and contrast the metabolite databases of the Pana cultivar with the PI 567731 (drought tolerant) cultivar to determine how different their molecular compositions are.

While ultimately we were unable to develop a working LC-MS/MS method to separate the components extracted from the seeds, and therefore unable to achieve the second goal, we were able to glean significant information on the molecular composition of Pana seeds even in the absence of chromatography.

Using a mortar and pestle, it was possible to mill Pana cultivar seeds into powders. No singular liquid extraction method is ideal to obtain the known rich diversity in molecular composition of soybean seeds. Thus, the seed powder was divided into two subsets. Subset A was subjected liquid extraction using ethyl acetate to extract *nonpolar* analytes from the seed powder. A 250 mg sample of powder was extracted in 50 mL of pure ethyl acetate solvent with ultrasonication at room temperature. Subset B was subjected to liquid extraction by methanol/water (9:1) spiked with 0.1% acetic acid in order to obtain the *polar* analytes. A 250 mg sample of powder was extracted in 50 mL of the extraction solvent with ultrasonication at room temperature. In both cases, ultrasonication was conducted for 20 minutes. The soluble liquid portion was decanted off into separate containers, leaving any insoluble material behind. The soluble fractions were then vacuum-dried to remove solvent, leaving two powders. The material from subset A will be hereafter referred to as the nonpolar powder, and the material from subset B will be referred to as the polar powder.

For mass spectrometry analysis, 30 ppm solutions of the polar and nonpolar powders were prepared in 9:1 methanol/water spiked with 0.1% formic acid for analysis by high-resolution Orbitrap mass spectrometry using electrospray ionization (ESI) analysis. Analysis was performed in the positive ion mode and the negative ion mode for both powder extracts in order to provide a broader view of the molecular composition of each power. In the positive ion mode, the polar powder extract showed the detection of 161 individual mass-to-charge values, all as singly-charged ions. The nonpolar powder extract showed the detection of a slightly larger set of 205 individual mass-to-charge values. The mass spectra themselves are displayed as Figure 1a and Figure 1b.

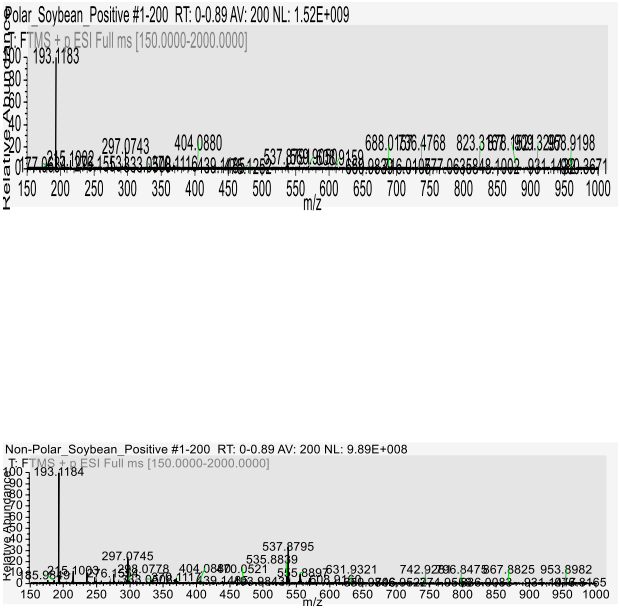
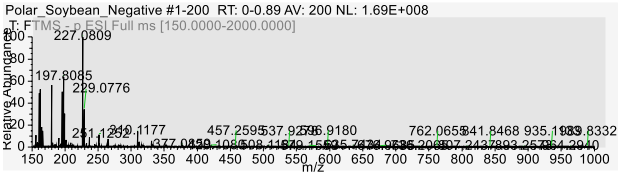


Figure 1. *Positive ion ESI mass spectra of a) polar extract of Pana soy seed powder and b) nonpolar extract of Pana soy seed powder.*

While the report by Jervis et al. identified a number of distinct *m/z* values in reversed-phase and hydrophobic interaction liquid chromatography (HILIC) when detected by MS, very few of these species were correlated with any molecular formula or structure [1]. The base (or most intense) peak in each spectrum corresponds to $C_7H_{16}N_2O_4$. Other peaks in the spectra are attributed to $C_7H_{14}N_2O_3$ (*m/z* 175.108, N-acetylornithine), $C_{13}H_{22}OS_2$ (*m/z* 297.074), $C_{11}H_{21}N_3O_5$ (*m/z* 276.155, likely a tripeptide), $C_6H_7N_5O_2S$ (*m/z* 236.021), $C_{15}H_{17}N_5O_5$ (*m/z* 370.112, an adenosine), $C_{19}H_{23}NOS_2$ (*m/z* 368.111), $C_{18}H_{22}O_2S_2$ (*m/z* 357.095), and $C_7H_{10}N_2O_2$ (*m/z* 177.063). However, without a chromatographic method that successfully isolates these components from one another, tandem mass spectrometry at this stage provides little additional information on structures.

Analysis performed in the negative ion mode for both powder extracts are displayed as Figure 2a and Figure 2b. For the polar powder extract, a slightly greater number of mass-to-charge ratios was detected in negative ion (193) than in positive ion mode (161); of these, two were doubly-charged, the others were all singly-charged. For the nonpolar powder extract, 202 individual mass-to-charge ratios were detected, close to the 205 identified in the positive ion mode mass spectrum. This includes one quadruply-charged species and one doubly-charged species.



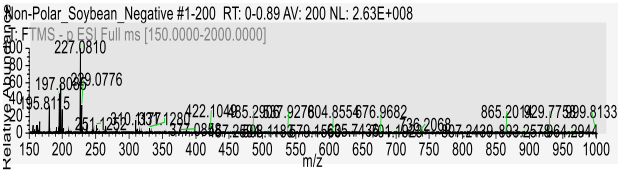


Figure 2. *Negative ion ESI mass spectra of a) polar extract of Pana soy seed powder and b) nonpolar extract of Pana soy seed powder.*

In both negative ion mass spectra, the base peak is detected at *m/z* 227.081, with a substantial peak shifted 2 Da away and in approximately 1/3 of the ratio of the base peak. This demonstrates that the species is ionized with the addition of Cl-, again pointing to a substance with a molecular formula $C_7H_{16}N_2O_4$. That both positive and negative ions corresponding to the same molecular formula suggests that this substance is both abundant and easily ionized with cations or anions. A few other molecular formulas are probable, including $C_6H_7N_5O$ (*m/z* 166.073, methylguanine), $C_7H_{17}NS_2$ (*m/z* 180.088, a dithiol), $C_9H_{18}N_2O_6$ (m/z 251.125, legionaminic acid), $C_{12}H_{24}O_4S$ (*m/z* 265.148), and $C_{11}H_{21}N_3O_5$ (*m/z* 310.118, likely a tripeptide like that in the positive ion mass spectra but with a chloride ion adduct as indicated by peak shifted higher by 2 Da).

One disappointment during the research was the inability of reversed-phase LC to separate the components. The columns used were all of the C18 capped particle type. This leads to an interesting hypothesis that perhaps more polar columns are needed to separate the components. However, because these columns are expensive, a better approach in the near-term will be to explore more mobile phases over a range of polarities but continue to employ the reversed-phase columns.