Supporting Information

Pipeline for High-Throughput Modeling of Marijuana and Hemp Extracts

Zewei Chen and Peter Harrington*

Center for Intelligent Chemical Instrumentation, Clippinger Laboratories, Department of Chemistry and Biochemistry, Ohio University, Athens, OH 45701, USA

*Author to whom correspondence should be addressed; E-Mail: peter.harrington@ohio.edu; Tel: +1-740-994-0265; Fax: +1-740-593-0148.

This file includes additional PC scores plots for the marijuana and hemp extracts from NMR, MS, and UV datasets. The results from a paired *t*-test are also included to compare different classifiers with respect to the classification accuracies in chemotyping process.

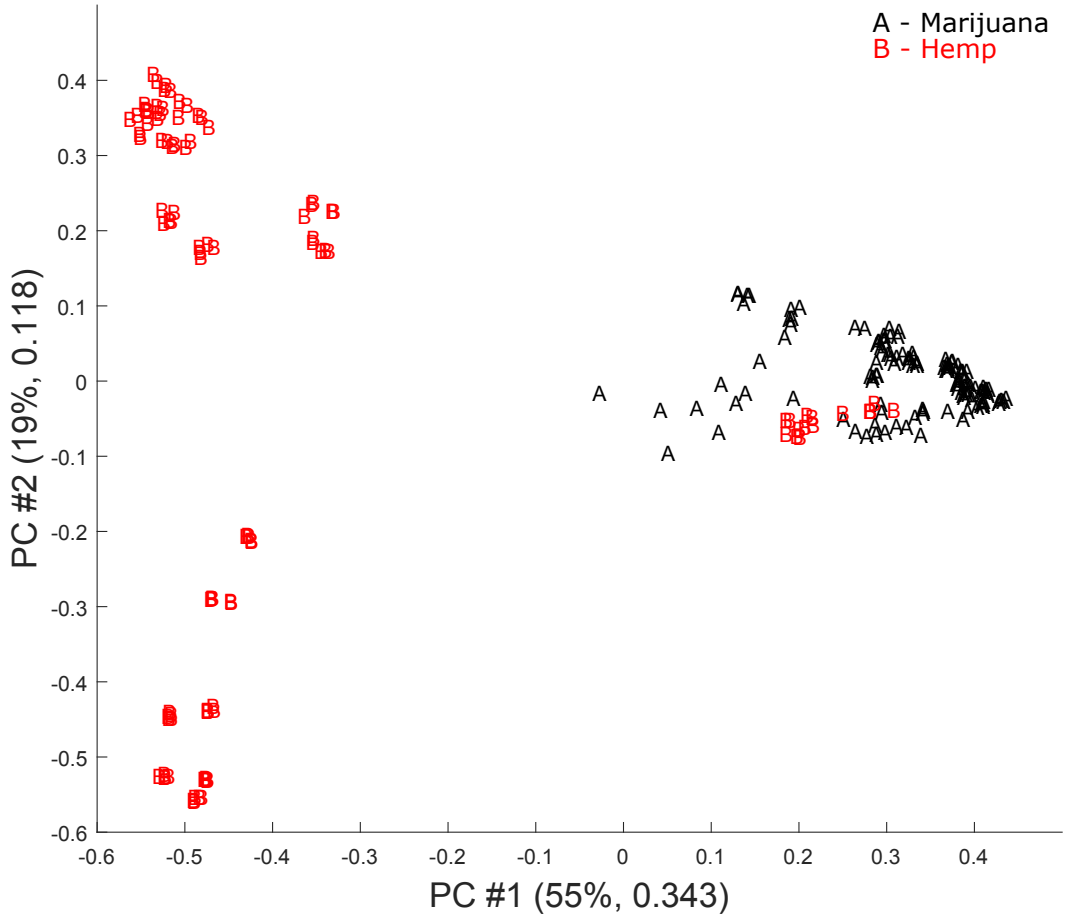

Figure S1. PC scores plots for marijuana and hemp with the normalized NMR real spectra.

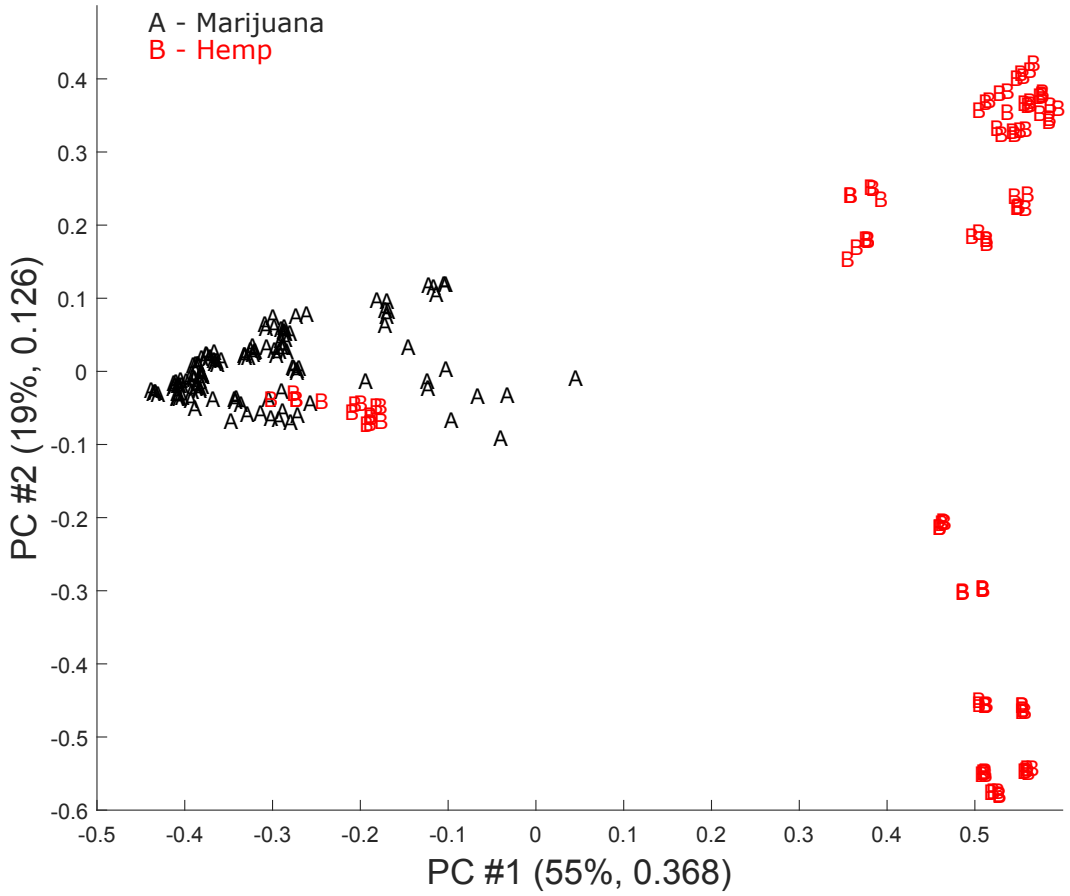

Figure S2. PC scores plots for marijuana and hemp with the normalized NMR imaginary spectra.

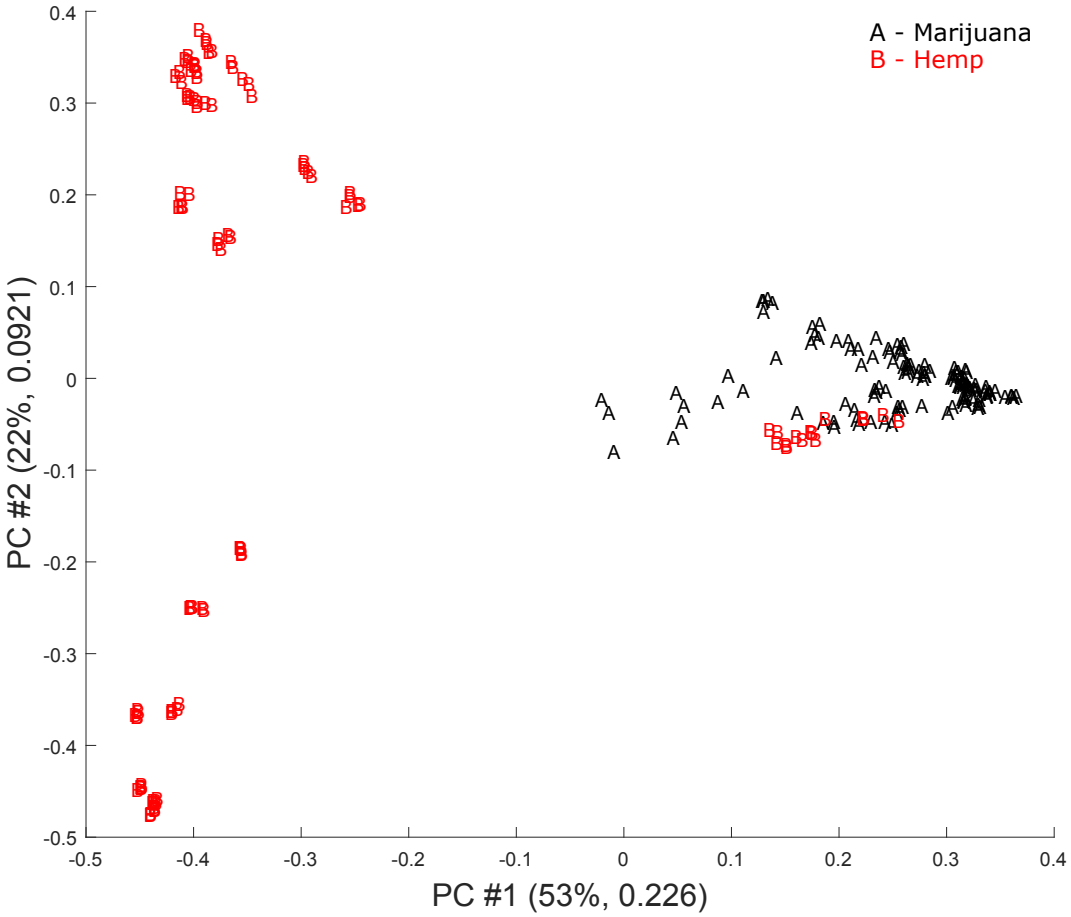

Figure S3. PC scores plots for marijuana and hemp with the normalized NMR absolute spectra.

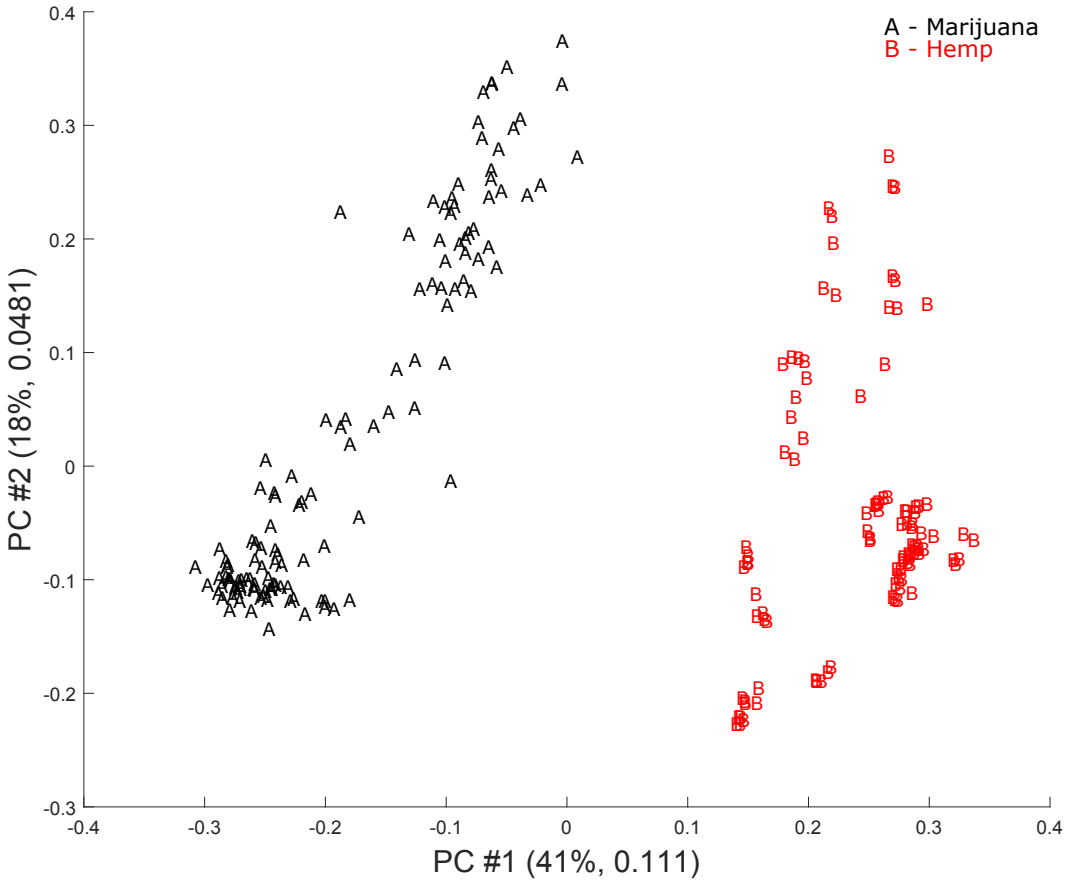

Figure S4. PC scores plots for marijuana and hemp with the normalized mass spectra after binning to integral mass-to-charge ratio and square root transformation.

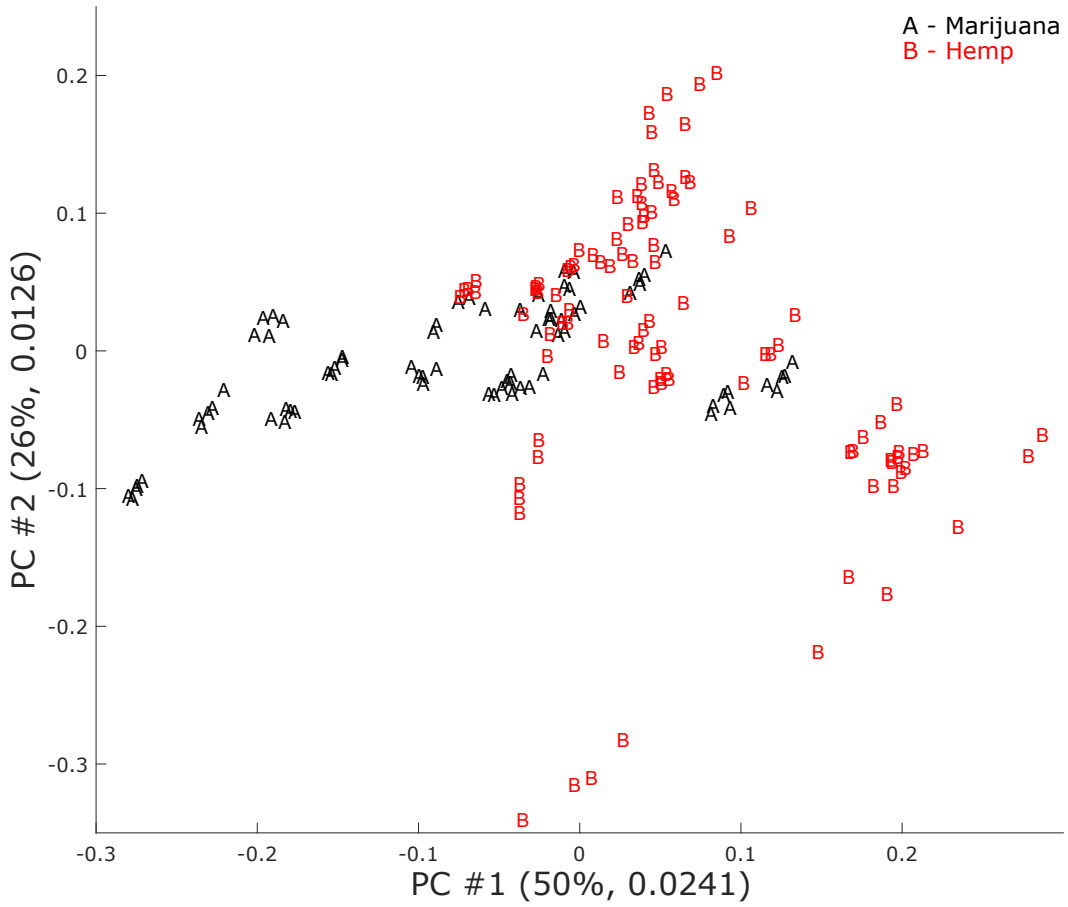

Figure S5. PC scores plots for marijuana and hemp with the normalized UV spectra.

Table S1. Results of the paired *t*-test to compare the three classifiers with respect to classification accuracy of the chemotyping process with the real NMR spectra.

Real NMR spectra				
	Marijuana		Hemp	
	t score	*p* value	*t* score	*p* value
FuRES vs sPLS-DA	-7.5	«0.01	0	1
FuRES vs SVMtreeH	-6.1	<0.01	0	1
sPLS-DA vs SVMtreeH	1.8	0.11	0	1

Table S2. Results of the paired *t*-test to compare the three classifiers with respect to classification accuracy of the chemotyping process with the imaginary NMR spectra.

Imaginary NMR spectra				
	Marijuana		Hemp	
	t score	*p* value	*t* score	*p* value
FuRES vs sPLS-DA	-9.4	«0.01	0	1
FuRES vs SVMtreeH	-3	0.01	0	1
sPLS-DA vs SVMtreeH	4.1	<0.01	0	1

Table S3. Results of the paired *t*-test to compare the three classifiers with respect to classification accuracy of the chemotyping process with the absolute NMR spectra of the marijuana and hemp extracts.

Absolute NMR spectra				
	Marijuana		Hemp	
	t score	*p* value	*t* score	*p* value
FuRES vs sPLS-DA	-5.1	<0.01	0	1
FuRES vs SVMtreeH	3.5	<0.01	0	1
sPLS-DA vs SVMtreeH	11.1	«0.01	0	1

Table S4. Results of the paired *t*-test to compare the three classifiers with respect to classification accuracy of the chemotyping process with the mass spectra of the marijuana and hemp extracts.

	Mass spectra			
	Marijuana		Hemp	
	t score	*p* value	*t* score	*p* value
FuRES vs sPLS-DA	-24.5	«0.01	0	1
FuRES vs SVMtreeH	-16.8	«0.01	0	1
sPLS-DA vs SVMtreeH	6	<0.01	0	1

Table S5. Results of the paired *t*-test to compare the three classifiers with respect to classification accuracy of the chemotyping process with the UV spectra of the marijuana and hemp extracts.

	UV spectra			
	Marijuana		Hemp	
	t score	*p* value	*t* score	*p* value
FuRES vs sPLS-DA	12.8	«0.01	11.7	«0.01
FuRES vs SVMtreeH	0	1	1.2	0.28
sPLS-DA vs SVMtreeH	-12.8	«0.01	-15	«0.01