

# Gas Chromatographic Fingerprint Analysis for the Comparison of Seized Cannabis Samples

Amorn Slosse <sup>1,2</sup>, Filip Van Durme <sup>1</sup>, Nele Samyn <sup>1</sup>, Debby Mangelings <sup>2</sup> and Yvan Vander Heyden <sup>2,\*</sup>

<sup>1</sup> Drugs and Toxicology Department, National Institute for Criminalistics and Criminology (NICC), Vilvoordsesteenweg 100, B-1120 Brussels, Belgium; amorn.slosse@just.fgov.be (A.S.); Filip.VanDurme@just.fgov.be (F.V.D.); Nele.samyn@just.fgov.be (N.S.)

<sup>2</sup> Department of Analytical Chemistry, Applied Chemometrics and Molecular Modelling, Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, B-1090 Brussels, Belgium; Debby.Mangelings@vub.be

\* Correspondence: yvanvdh@vub.be

**Abstract:** *Cannabis sativa* L. is widely applied as recreational illegal drugs. Illicit cannabis profiling, comparing seized samples, is challenging because of the natural Cannabis heterogeneity. The aim of this study is to use GC-FID and GC-MS herbal fingerprints for intra (within)- and inter (between)-location variabilities evaluation. This study focuses on finding an acceptable threshold to link seized samples. Through Pearson correlation-coefficient calculations between intra-location samples, 'linked' thresholds were derived using 95% and 99% confidence limits. False negative (FN) and false positive (FP) error rate calculations, aiming at getting the lowest possible FP value, were performed for different data pretreatments. Fingerprint-alignment parameters were optimized using Automated Correlation Optimized Warping (ACOW) or Design of Experiments (DoE), which resulted in similar results. Hence, ACOW data, as reference, showed 54% and 65% FP values (95/99% confidence). Additional fourth root normalization pretreatment provided the best results for both the GC-FID and GC-MS data sets. For GC-FID, which showed the best improved FP error rate, 54/65% FP for the reference data decreased to 24 and 32% after fourth root transformation. Cross validation showed similar FP values as the entire calibration set, indicating the representativeness of the thresholds. A noteworthy improvement in discrimination between seized Cannabis samples could be concluded.

**Keywords:** Chromatographic fingerprint; alignment optimization; design of experiments; data preprocessing; comparison intra- and inter-location samples

**Citation:** Slosse, A.; Van Durme, F.; Samyn, N.; Mangelings, D.; Vander Heyden, Y. Gas chromatographic fingerprint analysis for the comparison of seized cannabis samples *Molecules* **2021**, *26*, 6643. <https://doi.org/10.3390/molecules26216643>

Academic Editor: Michal Daszykowski, Joaquim Jaumot, Hadi Parastar

Received: 30 September 2021

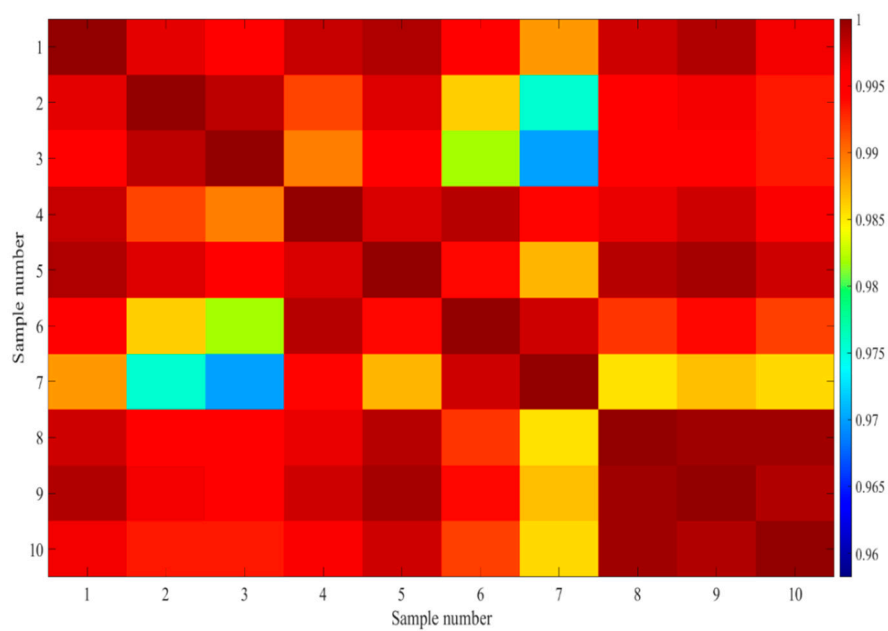
Accepted: 26 October 2021

Published: 2 November 2021

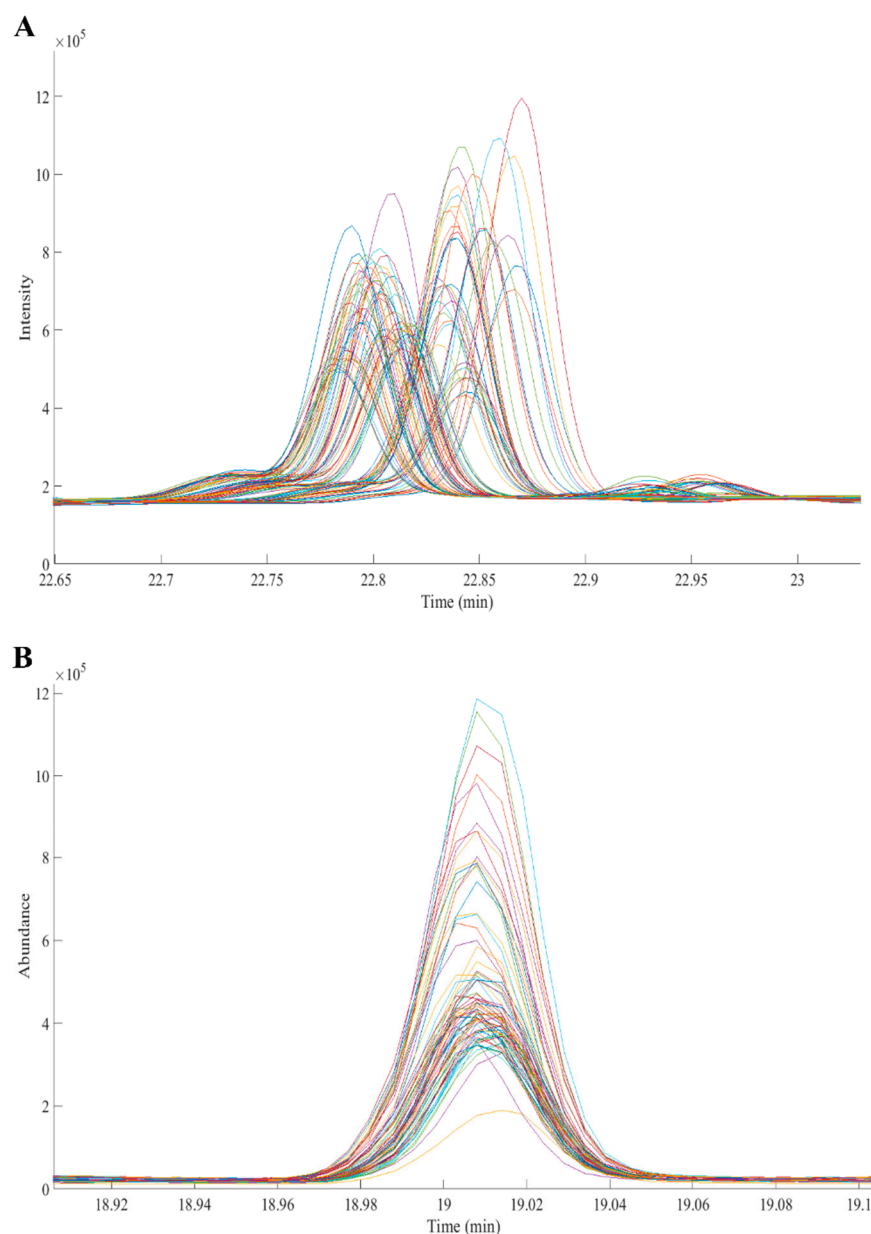
**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



**Figure S1.** Intra-correlation coefficients colour map from the 8<sup>th</sup> cultivation site. GC-FID data were used.



**Figure S2.** A detail plot of one peak for (A) GC-FID raw data, and (B) GC-MS TIC raw data for the 10 different cultivation sites.

**Table S1.** Summary of the FN and FP values of GC-FID Cannabis fingerprints for the ACOW and DoE approach. The effect of multiple alignments are also shown. The best FP error rates are denoted in bold.

	95% CL		99% CL	
	FN (%)	FP (%)	FN (%)	FP (%)
<b>Raw Data</b>	6	34	3	39
<b>Range SL 15-100 &amp; SS 1-10</b>				
1 x ACOW	6	57	4	65
2 x ACOW	6	<b>47</b>	4	<b>55</b>
3 x ACOW	5	49	4	58
1 x DoE	5	<b>47</b>	3	<b>55</b>
2 x DoE	4	58	3	65
3 x DoE	6	50	4	59

<b>Range SL 25-200 &amp; SS 1-10</b>				
1 x ACOW	6	58	4	63
2 x ACOW	5	54	4	62
3 x ACOW	5	66	2	75
1 x DoE	6	58	4	60
2 x DoE	5	57	3	62
3 x DoE	4	<b>49</b>	3	<b>55</b>

**Table S2.** Representation of the alignment optimization on the GC-MS raw data. The warping effect is expressed as the obtained FN and FP error rates.

	<b>95% CL</b>		<b>99% CL</b>	
	FN (%)	FP (%)	FN (%)	FP (%)
<b>Raw Data</b>	4	53	3	58
<b>Range SL 15-100 &amp; SS 1-10</b>				
1 x ACOW	7	53	3	58
1 x DoE	7	53	3	58
<b>Range SL 25-200 &amp; SS 1-10</b>				
1 x ACOW	6	54	2	57
1 x DoE	6	54	3	57

**Table S3.** The obtained total % FN and FP for all pretreatment techniques after aligning the GC-FID raw data twice using ACOW.

<b>Pre-treatment method</b>	<b>95% CL</b>		<b>99% CL</b>	
	FN (%)	FP (%)	FN (%)	FP (%)
<b>Aligned data (2xACOW)</b>	6	47	3	55
<b>Column centering (CC)</b>	9	55	6	63
<b>Normalization + CC (N+CC)</b>	7	54	4	67
<b>Standard normal variate + CC (SNV+CC)</b>	7	52	4	66
<b>Square root normalization</b>	7	29	3	38
<b>Fourth root normalization</b>	7	24	4	32
<b>Auto-scaling</b>	6	19	4	27

**Table S4.** Total % FN and FP for all pretreatment techniques after removal of THC of the aligned GC-FID data matrix.

<b>Pre-treatment method</b>	<b>95% CL</b>		<b>99% CL</b>	
	FN (%)	FP (%)	FN (%)	FP (%)
<b>Aligned data (1xACOW)</b>	6	57	4	65
<b>Column centering (CC)</b>	7	31	3	42
<b>Normalization + CC (N+CC)</b>	4	29	3	40
<b>Standard normal variate + CC (SNV+CC)</b>	7	29	3	38
<b>Square root normalization</b>	5	30	4	40
<b>Fourth root normalization</b>	5	28	4	37
<b>Auto-scaling</b>	6	19	4	27