# **Electronic Supplementary Information**

# Smartphone-based image analysis and chemometric pattern recognition of the thin-layer chromatographic fingerprints of herbal materials

Sarah May Sibug-Torres<sup>a</sup>, Isagani D. Padolina<sup>c</sup>, Philip Cruz<sup>d</sup>, Felan Carlo Garcia<sup>b</sup>, Mark Joseph Garrovillas<sup>a</sup>, Maria Regina Yabillo<sup>a</sup>, Erwin P. Enriquez<sup>a‡</sup>

<sup>a</sup> Department of Chemistry, Ateneo de Manila University, Katipunan Ave., Loyola Heights, Quezon City, Philippines 1108

<sup>b</sup> Department of Electronics, Computer, and Communications Engineering, Ateneo de Manila University, Katipunan Ave., Loyola Heights, Quezon City, Philippines 1108

<sup>c</sup> Pascual Pharma Corp., 23rd Floor, The Taipan Place, F. Ortigas Jr. Road, Ortigas Center, Brgy. San Antonio, Pasig City, Metro Manila, Philippines 1605

<sup>d</sup> Herbanext Laboratories, Inc., Negros South Road, Bago City, Negros Occidental, Philippines

**Corresponding Author** <sup>‡</sup>epenriquez@ateneo.edu

## CONTENTS

#### **Figures**

Figure ESI-1. Cropping algorithm for extracting the TLC fingerprint region-of-interest (ROI)

Figure ESI-2. Application of white reference strip in the 3D-printed photo-box to check for consistent lighting conditions during image capture.

Figure ESI-3. Schematic of nested 5-fold cross-validation with one-class model training using only within-specifications samples.

**Figure ESI-4.** Classification predictions for *B. balsamifera* samples based on the outer loops in nested 5-fold cross-validation and aggregated SIMCA models

Figure ESI-5. Classification predictions for *V. negundo* samples based on the outer loops in nested 5-fold cross-validation and aggregated SIMCA models

### **Tables**

**Table ESI-1**. Sample details and TLC profiles of *B. balsamifera* within-specifications (WS), off-specifications (OS) samples and mixtures.

**Table ESI-2.** Sample details and TLC profiles of *V. negundo* within-specifications (WS), off-specifications (OS) samples and mixtures.

Table ESI-3. Results for tuning the B. balsamifera SIMCA models during inner cross-validation

Table ESI-4. Results for tuning the V. negundo SIMCA models during inner cross-validation



**Figure ESI-1.** Cropping algorithm for extracting the TLC fingerprint region-of-interest (ROI) implemented in Python (Python Software Foundation, https://www.python.org/) with SciPy's multi-dimensional image processing package



**Figure ESI-2.** The white reference strip in the 3D-printed photo-box is used to check for consistent lighting conditions during image capture. The mean RGB intensity values of the white reference strip is compared to the range of RGB intensity values of calibration images. If lighting is within range, further image processing proceeds. If not, the user is prompted to repeat image capture.

**Table ESI-1**. Sample details and TLC profiles of *B. balsamifera* within-specifications (WS), off-specifications (OS) samples and mixtures. Note that TLC profiles (Full Pharmacopeia vs. adapted TLC methods) are not to scale. Sample BB21 is shown with *B. balsamifera* standards BL (blumeatin), DQDE (dihydroquercetin-7,4'-dimethyl ether), DQME (dihydroquercetin-4'-methyl ether), THFE (5,7,3',5'-tetrahydroxyflavanone), and QN (quercetin)

Class	ID	Source	Date Collected	Processing History	Pharmacopeia Full TLC	Adapted TLC
WS	BB01	Farm 1, Luzon, Philippines	June 2017	Air Dried		
WS	BB02	Farm 1, Luzon, Philippines	June 2017	Oven Dried		
WS	BB03	Farm 2, Visayas, Philippines	April 2016	Air Dried		
WS	BB04	Farm 2, Visayas, Philippines	April 2016	Oven Dried		
WS	BB05	Farm 2, Visayas, Philippines	July 2016	Oven Dried		
WS	BB06	Farm 3, Luzon, Philippines	Feb 2017	Air Dried	0.0.0	
WS	BB07	Farm 3, Luzon, Philippines	Feb 2017	Oven Dried		
WS	BB08	Farm 4, Visayas, Philippines	June 2017	Air Dried		
WS	BB09	Farm 5, Visayas, Philippines	June 2017	Air Dried		
WS	BB10	Farm 6, Visayas, Philippines	June 2017	Air Dried		
WS	BB11	Farm 7, Visayas, Philippines	June 2017	Air Dried		
WS	BB12	Farm 8, Visayas, Philippines	June 2017	Air Dried		
WS	BB13	Farm 9, Luzon, Philippines	June 2016	Air Dried		

Class	ID	Source	Date Collected	Processing History	Pharmacopeia Full TLC	Adapted TLC
WS	BB14	Farm 9, Luzon, Philippines	June 2016	Oven Dried		
WS	BB15	Farm 10, Luzon, Philippines	Feb 2017	Air Drying	001111	
WS	BB16	Farm 11, Luzon, Philippines	Feb 2017	Oven Dried		
WS	BB17	Farm 12, Luzon, Philippines	Aug 2017	Oven Dried		
WS	BB18	Farm 12, Luzon, Philippines	April 2017	Oven Dried		
WS	BB19	Farm 13, Luzon, Philippines	June 2016	Oven Dried		
WS	BB20	Farm 13, Luzon, Philippines	July 2016	Dehydrator		
WS	BB21	Farm 13, Luzon, Philippines	Nov 2017	Oven Dried		
BB std	BL				- 0	
BB std	DQDE				0	
BB std	DQME					
BB std	QN					
BB std	THFE				C	
WS	BB22	Farm 14, Luzon, Philippines	Nov 2017	Oven Dried		-
OS	BB_OS1	Farm 4, Visayas, Philippines	Oct 2017	Stored in humid conditions, 24 hrs		
OS	BB_OS2	Farm 4, Visayas, Philippines	Oct 2017	Stored in humid conditions, 48 hrs		

Class	ID	Source	Date Collected	Processing History	Pharmacopeia Full TLC	Adapted TLC
OS	BB_OS3	Farm 4, Visayas, Philippines	Feb 2016	Steam Sterilized	010 1111	
OS	BB_OS4	Farm 4, Visayas, Philippines	July 2016	Fermented	2011	
OS	BB_OS5	Farm 4, Visayas, Philippines	July 2016	Incomplete drying followed by storage	000000	
OS	BB_OS6	Farm 4, Visayas, Philippines	July 2016	High Heat Drying, 100°C		
OS	BB_OS7	Farm 13, Visayas, Philippines	Nov 2017	Stored in humid conditions, 48 hrs		
OS	BB_OS8	Farm 13, Visayas, Philippines	Nov 2017	High Heat Drying, 100°C		
MIX	BB_BL05	Farm 4, Visayas, Philippines	Feb 2016	Oven Dried, 5% <i>B. lacera</i>		
MIX	BB_BL10	Farm 4, Visayas, Philippines	Feb 2016	Oven Dried, 10% <i>B. lacera</i>		
MIX	BB_BL20	Farm 4, Visayas, Philippines	Feb 2016	Oven Dried, 20% <i>B. lacera</i>		
MIX	BB_BL30	Farm 4, Visayas, Philippines	Feb 2016	Oven Dried, 30% <i>B. lacera</i>		
MIX	BB_BL40	Farm 4, Visayas, Philippines	Feb 2016	Oven Dried, 50% <i>B. lacera</i>		
MIX	BB_BL50	Farm 4, Visayas, Philippines	Feb 2016	Oven Dried, 50% <i>B. lacera</i>		
MIX	BB_BL60	Farm 4, Visayas, Philippines	Feb 2016	Oven Dried, 60% <i>B. lacera</i>		
MIX	BB_BL70	Farm 4, Visayas, Philippines	Feb 2016	Oven Dried, 70% <i>B. lacera</i>		
MIX	BB_BL80	Farm 4, Visayas, Philippines	Feb 2016	Oven Dried, 80% <i>B. lacera</i>	0 00 01 11 11	

Class	ID	Source	Date Collected	Processing History	Pharmacopeia Full TLC	Adapted TLC
MIX	BB_BL90	Farm 4, Visayas, Philippines	Feb 2016	Oven Dried, 90% <i>B. lacera</i>		
MIX	BB_BL	Farm 4, Visayas, Philippines	Feb 2016	Oven Dried, 100% <i>B. lacera</i>		
MIX	BB_VN05	Farm 11, Luzon, Philippines	Nov 2017	Oven Dried, 5% <i>V. negundo</i>		
MIX	BB_VN10	Farm 11, Luzon, Philippines	Nov 2017	Oven Dried, 10% <i>V. negundo</i>		
MIX	BB_VN20	Farm 11, Luzon, Philippines	Nov 2017	Oven Dried, 20% <i>V. negundo</i>		
MIX	BB_VN30	Farm 11, Luzon, Philippines	Nov 2017	Oven Dried, 30% <i>V. negundo</i>		
MIX	BB_VN40	Farm 11, Luzon, Philippines	Nov 2017	Oven Dried, 40% <i>V. negundo</i>		
MIX	BB_VN50	Farm 11, Luzon, Philippines	Nov 2017	Oven Dried, 50% <i>V. negundo</i>		
MIX	BB_VN60	Farm 11, Luzon, Philippines	Nov 2017	Oven Dried, 60% <i>V. negundo</i>		
MIX	BB_VN70	Farm 11, Luzon, Philippines	Nov 2017	Oven Dried, 70% <i>V. negundo</i>		
MIX	BB_VN80	Farm 11, Luzon, Philippines	Nov 2017	Oven Dried, 80% <i>V. negundo</i>		
MIX	BB_VN90	Farm 11, Luzon, Philippines	Nov 2017	Oven Dried, 90% <i>V. negundo</i>		
MIX	BB_VN	Farm 11, Luzon, Philippines	Nov 2017	Oven Dried, 100% <i>V. negund</i> o		

**Table ESI-2.** Sample details and TLC profiles of *V. negundo* within-specifications (WS), off-specifications (OS) samples and mixtures. Note that TLC profiles (Full Pharmacopeia vs. adapted TLC methods) are not to scale. Sample VN15 is shown with *V. negundo* standard, AGN (agnuside)

Class	ID	Source	Date Collected	Processing History	Pharmacopeia Full TLC	Adapted TLC
WS	VN01	Farm 1, Luzon, Philippines	Feb 2017	Oven Dried		
WS	VN02	Farm 2, Luzon, Philippines	Feb 2017	Air Dried	0000000000	
WS	VN03	Farm 2, Luzon, Philippines	Feb 2017	Oven Dried	00000	
WS	VN04	Farm 3, Luzon, Philippines	May 2017	Air Dried		
WS	VN05	Farm 4, Visayas, Philippines	June 2017	Air Dried	00000	
WS	VN06	Farm 5, Visayas, Philippines	June 2017	Air Dried		
WS	VN07	Farm 6, Luzon, Philippines	Feb 2017	Oven Dried	00000	
WS	VN08	Farm 7, Luzon, Philippines	Feb 2017	Oven Dried		
WS	VN09	Farm 8, Luzon, Philippines	Nov 2017	Air Dried		
WS	VN10	Variety A - Farm 9, Luzon, Phil.	Nov 2017	Air Dried		
WS	VN11	Variety B - Farm 9, Luzon, Phil.	Nov 2017	Air Dried		
WS	VN12	Farm 10, Luzon, Philippines	Feb 2017	Oven Dried		
WS	VN13	Farm 11, Luzon, Philippines	June 2017	Air Dried		
WS	VN14	Farm 12, Luzon, Philippines	June 2017	Air Dried	000 000	

Class	ID	Source	Date Collected	Processing History	Pharmacopeia Full TLC	Adapted TLC
std	AGN					
WS	VN15	Farm 13, Luzon, Philippines	Nov 2017	Oven Dried		
OS	VN_OS1	Farm 2, Luzon, Philippines	Feb 2017	High Heat Drying, 100°C		
OS	VN_OS2	Farm 2, Luzon, Philippines	Feb 2017	Stored in humid conditions		
OS	VN_OS3	Farm 2, Luzon, Philippines	Feb 2017	Fermented	••	
OS	VN_OS4	Farm 2, Luzon, Philippines	Feb 2017	Incomplete drying followed by storage		-
OS	VN_OS5	Farm 4, Visayas, Philippines	Nov 2017	Stored in humid conditions, 24 hrs		
OS	VN_OS6	Farm 5, Visayas, Philippines	Nov 2017	Stored in humid conditions, 48 hrs		
OS	VN_OS7	Farm 9, Luzon, Philippines	Nov 2017	Stored in humid conditions, 24 hrs		
OS	VN_OS8	Farm 9, Luzon, Philippines	Nov 2017	Stored in humid conditions, 48 hrs	000000000	
MIX	VN_BB05	Farm 13, Luzon, Philippines	June 2017	Oven Dried, 5% <i>B. balsamifera</i>		
MIX	VN_BB10	Farm 13, Luzon, Philippines	June 2017	Oven Dried, 10% <i>B. balsamifera</i>		
MIX	VN_BB20	Farm 13, Luzon, Philippines	June 2017	Oven Dried, 20% <i>B. balsamifera</i>		
MIX	VN_BB30	Farm 13, Luzon, Philippines	June 2017	Oven Dried, 30% <i>B. balsamifera</i>		
MIX	VN_BB40	Farm 13, Luzon, Philippines	June 2017	Oven Dried, 40% <i>B. balsamifera</i>		

Class	ID	Source	Date Collected	Processing History	Pharmacopeia Full TLC	Adapted TLC
MIX	VN_BB50	Farm 13, Luzon, Philippines	June 2017	Oven Dried, 50% <i>B. balsamifera</i>		
MIX	VN_BB60	Farm 13, Luzon, Philippines	June 2017	Oven Dried, 60% <i>B. balsamifera</i>		
MIX	VN_BB70	Farm 13, Luzon, Philippines	June 2017	Oven Dried, 70% <i>B. balsamifera</i>	00 000 1 I III	
MIX	VN_BB80	Farm 13, Luzon, Philippines	June 2017	Oven Dried, 80% <i>B. balsamifera</i>	000000000000000000000000000000000000000	
MIX	VN_BB90	Farm 13, Luzon, Philippines	June 2017	Oven Dried, 90% <i>B. balsamifera</i>	000011	•
MIX	VN_BB	Farm 13, Luzon, Philippines	June 2017	Oven Dried, 100% <i>B.</i> balsamifera	11 11	



**Figure ESI-3**. Schematic of nested 5-fold cross-validation with one-class model training using only within-specifications samples. Off-specifications samples are used only as validation or test samples for the corresponding inner and outer folds for the within-specifications sample set. Inner cross-validation loops are repeated for all folds in the outer cross-validation loop. The number of samples in the within-specifications and off-specifications classes are set as n=15 for illustrative purposes. Each sample has three technical replicates A, B, and C.

Table	ESI-3.	Results for	1	90.2%	87.5%	1	90.2%	87.5%	 2	92.6%	83.3%	3	78.4%	91.7%
tuning	the R	halsamifara	1	92.6%	79.2%	1	88.9%	83.3%	2	90.2%	91.7%	3	82.4%	95.8%
cuning			1	92.6%	95.8%	1	87.0%	91.7%	 2	76.5%	87.5%	3	79.6%	95.8%
SIMC	A models	during inner	1	92.6%	87.5%	1	88.2%	87.5%	2	92.2%	87.5%	3	79.6%	100.0%
cross-v	alidation		1	92.6%	83.3%	1	96.3%	87.5%	 2	88.9%	87.5%	3	74.5%	95.8%
	0/	0/	1	94.4%	87.5%	1	90.7%	95.8%	 2	79.6%	87.5%	3	81.5%	91.7%
# OT PCs	% Sonsitivity	% Specificity	1	92.6%	79.2%	1	86.3%	83.3%	 2	83.3%	95.8%	3	82.4%	95.8%
103	Genativity	opeenieity	1	92.6%	95.8%	1	88.9%	91.7%	 2	88.2%	100.0%	3	78.4%	100.0%
1	98.1%	87.5%	1	86.3%	91.7%	1	86.3%	87.5%	 2	83.3%	100.0%	3	87.0%	91.7%
1	90.7%	100.0%	1	88.9%	87.5%	1	90.2%	87.5%	 2	84.3%	91.7%	3	78.4%	100.0%
1	86.3%	87.5%	1	90.7%	100.0%	1	96.3%	91.7%	 2	90.7%	87.5%	3	85.2%	95.8%
1	87.0%	79.2%	1	88.2%	79.2%	1	88.2%	87.5%	 2	90.2%	91.7%	3	85.2%	100.0%
1	78.4%	91.7%	1	92.6%	91.7%	1	92.6%	87.5%	 2	85.2%	91.7%	3	84.3%	95.8%
1	86.3%	100.0%	1	92.2%	91.7%	1	88.2%	79.2%	 2	84.3%	95.8%	3	81.5%	95.8%
1	94.1%	91.7%	1	81.5%	87.5%	1	85.2%	95.8%	 2	75.9%	100.0%	3	86.3%	91.7%
1	92.2%	91.7%	1	90.2%	87.5%	1	90.7%	95.8%	 2	85.2%	79.2%	3	85.2%	87.5%
1	90.7%	91.7%	1	92.6%	79.2%	1	88.9%	83.3%	 2	86.3%	95.8%	3	80.4%	100.0%
1	86.3%	87.5%	1	90.7%	91.7%	1	94.1%	87.5%	 2	85.2%	91.7%	3	77.8%	100.0%
1	90.2%	95.8%	1	92.6%	100.0%	1	88.2%	79.2%	 2	85.2%	87.5%	3	86.3%	95.8%
1	88.9%	91.7%	1	96.1%	87.5%	1	90.7%	91.7%	 2	90.2%	95.8%	3	75.9%	91.7%
1	92.2%	79.2%	1	90.7%	95.8%	1	92.6%	95.8%	 2	87.0%	87.5%	3	87.0%	95.8%
1	94.4%	79.2%	1	88.9%	79.2%	1	92.2%	83.3%	 2	79.6%	83.3%	3	75.9%	95.8%
1	92.6%	100.0%	1	92.2%	87.5%	1	90.7%	87.5%	 2	83.3%	91.7%	3	80.4%	91.7%
1	90.7%	83.3%	1	88.9%	95.8%	1	90.7%	95.8%	 2	90.2%	83.3%	3	83.3%	100.0%
1	90.2%	87.5%	1	94.1%	83.3%	1	85.2%	87.5%	 2	77.8%	95.8%	3	84.3%	91.7%
1	90.7%	91.7%	1	90.7%	87.5%	1	92.2%	83.3%	 2	83.3%	95.8%	3	88.2%	91.7%
1	94.4%	87.5%	1	85.2%	87.5%	1	94.4%	79.2%	 2	82.4%	100.0%	3	81.5%	91.7%
1	86.3%	95.8%	1	86.3%	87.5%	1	84.3%	87.5%	 2	90.7%	100.0%	3	84.3%	100.0%
1	88.9%	91.7%	1	92.2%	95.8%	1	94.1%	83.3%	 2	88.9%	83.3%	4	79.6%	100.0%
1	88.2%	83.3%	1	92.6%	95.8%	1	85.2%	87.5%	 2	92.2%	95.8%	4	74.5%	100.0%
1	92.2%	95.8%	1	88.9%	95.8%	1	92.2%	95.8%	 2	90.7%	83.3%	4	77.8%	100.0%
1	94.4%	87.5%	1	84.3%	83.3%	2	88.9%	91.7%	 2	88.9%	91.7%	4	83.3%	100.0%
1	94.4%	100.0%	1	94.4%	87.5%	2	87.0%	87.5%	 2	81.5%	83.3%	mean	87.9%	90.8%
1	90.2%	83.3%	1	94.1%	83.3%	2	81.5%	95.8%	 2	90.7%	91.7%	s.d.	5.1%	6.1%
1	90.7%	100.0%	1	96.1%	95.8%	2	90.7%	91.7%	 2	92.2%	95.8%			
1	88.9%	83.3%	1	88.9%	91.7%	2	92.6%	100.0%	 2	92.6%	100.0%			
1	94.1%	83.3%	1	90.7%	87.5%	2	87.0%	91.7%	 2	88.9%	95.8%			
1	88.2%	83.3%	1	98.0%	95.8%	2	80.4%	83.3%	 2	88.9%	87.5%			
1	92.6%	91.7%	1	94.4%	83.3%	2	85.2%	95.8%	 3	87.0%	95.8%			
1	88.9%	91.7%	1	100.0%	100.0%	2	87.0%	91.7%	 3	80.4%	95.8%			
1	92.2%	87.5%	1	86.3%	83.3%	2	87.0%	83.3%	 3	79.6%	95.8%			
1	94.4%	87.5%	1	92.2%	79.2%	2	87.0%	91.7%	 3	83.3%	87.5%			
1	87.0%	83.3%	1	90.2%	87.5%	2	88.9%	83.3%	 3	74.1%	100.0%			
1	90.2%	95.8%	1	92.6%	83.3%	2	88.2%	100.0%	 3	84.3%	95.8%			



**Figure ESI-4.** Classification predictions for *B. balsamifera* samples based on the outer loops in nested 5-fold cross-validation and aggregated SIMCA models (1 = within-specifications, -1 = off-specifications, 0 = abstain)

Table	ESI-4. F	Results for	1	77.8%	90.5%	1	69.4%	100.0%	_	1	75.0%	87.5%	1	75.0%	90.5%
tuning	tha V	nagundo	1	63.9%	90.5%	1	75.0%	90.5%		1	72.2%	100.0%	1	61.1%	85.7%
tunnig	, the v.	negunao	1	80.6%	100.0%	1	63.9%	95.2%		1	72.2%	95.2%	1	61.1%	91.7%
SIMC	A model	s during	1	72.2%	100.0%	1	77.8%	90.5%		1	63.9%	90.5%	1	77.8%	95.2%
inner	prose-valid	ation	1	77.8%	100.0%	1	72.2%	100.0%		1	72.2%	91.7%	1	77.8%	95.2%
	21055-vanu		1	72.2%	90.5%	1	75.0%	100.0%		1	69.4%	100.0%	1	63.9%	90.5%
# of	%	%	1	83.3%	85.7%	1	75.0%	100.0%		1	69.4%	100.0%	1	72.2%	90.5%
PCs	Sensitivity	Specificity	1	72.2%	95.2%	1	63.9%	95.8%		1	77.8%	95.2%	1	72.2%	91.7%
- 05	Selisierreg	sprenning	1	77.8%	91.7%	1	77.8%	90.5%		1	69.4%	95.2%	1	83.3%	81.0%
1	58.3%	91.7%	1	80.6%	90.5%	1	66.7%	90.5%		1	69.4%	100.0%	1	63.9%	81.0%
1	66.7%	90.5%	1	66.7%	95.2%	1	80.6%	95.2%		1	69.4%	90.5%	1	75.0%	100.0%
1	66.7%	100.0%	1	83.3%	90.5%	1	72.2%	95.2%		1	72.2%	95.2%	1	66.7%	95.2%
1	86.1%	90.5%	1	66.7%	100.0%	1	86.1%	95.8%		1	77.8%	95.2%	2	63.9%	95.2%
1	66.7%	95.8%	1	80.6%	90.5%	1	66.7%	85.7%		1	69.4%	83.3%	2	52.8%	100.0%
1	77.8%	90.5%	1	63.9%	95.2%	1	75.0%	90.5%		1	72.2%	90.5%	2	63.9%	100.0%
1	80.6%	100.0%	1	61.1%	87.5%	1	66.7%	100.0%		1	75.0%	100.0%	2	75.0%	95.2%
1	52.8%	95.2%	1	66.7%	100.0%	1	66.7%	90.5%		1	72.2%	90.5%	2	50.0%	95.2%
1	83.3%	100.0%	1	80.6%	90.5%	1	63.9%	91.7%		1	66.7%	95.2%	2	52.8%	95.2%
1	69.4%	87.5%	1	80.6%	100.0%	1	77.8%	100.0%		1	75.0%	91.7%	2	58.3%	95.2%
1	69.4%	90.5%	1	72.2%	90.5%	1	75.0%	90.5%		1	77.8%	100.0%	2	61.1%	100.0%
1	72.2%	100.0%	1	69.4%	91.7%	1	69.4%	95.2%		1	77.8%	95.2%	2	58.3%	90.5%
1	63.9%	95.2%	1	69.4%	81.0%	1	72.2%	91.7%		1	80.6%	90.5%	2	69.4%	95.2%
1	72.2%	90.5%	1	72.2%	90.5%	1	80.6%	100.0%		1	72.2%	95.2%	2	52.8%	100.0%
1	66.7%	91.7%	1	75.0%	95.2%	1	80.6%	100.0%		1	75.0%	91.7%	2	61.1%	100.0%
1	69.4%	90.5%	1	61.1%	95.2%	1	69.4%	100.0%		1	66.7%	100.0%	2	69.4%	95.2%
1	72.2%	90.5%	1	75.0%	100.0%	1	69.4%	95.8%		1	77.8%	95.2%	2	66.7%	95.2%
1	75.0%	83.3%	1	72.2%	85.7%	1	69.4%	100.0%		1	86.1%	90.5%	2	66.7%	100.0%
1	77.8%	90.5%	1	75.0%	81.0%	1	77.8%	100.0%		1	75.0%	91.7%	2	61.1%	95.2%
1	69.4%	90.5%	1	66.7%	100.0%	1	66.7%	95.2%		1	75.0%	95.2%	2	63.9%	100.0%
1	72.2%	95.2%	1	66.7%	85.7%	1	75.0%	90.5%		1	77.8%	100.0%	2	63.9%	95.2%
1	72.2%	100.0%	1	66.7%	87.5%	1	75.0%	91.7%		1	58.3%	81.0%	2	55.6%	95.2%
1	69.4%	100.0%	1	77.8%	95.2%	1	77.8%	100.0%		1	72.2%	100.0%	2	75.0%	100.0%
1	72.2%	100.0%	1	75.0%	100.0%	1	61.1%	100.0%		1	63.9%	95.8%	2	63.9%	100.0%
1	80.6%	81.0%	1	80.6%	90.5%	1	77.8%	85.7%		1	75.0%	90.5%	Mean	71.4%	93.9%
1	63.9%	91.7%	1	80.6%	95.2%	1	72.2%	100.0%		1	80.6%	85.7%	s.d.	7.1%	5.3%
1	63.9%	90.5%	1	72.2%	100.0%	1	75.0%	100.0%		1	77.8%	90.5%			
1	86.1%	90.5%	1	66.7%	85.7%	1	66.7%	100.0%		1	69.4%	100.0%			
1	69.4%	90.5%	1	61.1%	100.0%	1	77.8%	90.5%		1	72.2%	90.5%			
1	69.4%	100.0%	1	72.2%	95.2%	1	75.0%	90.5%		1	77.8%	90.5%			
1	66.7%	81.0%	1	69.4%	95.2%	1	63.9%	87.5%		1	77.8%	100.0%			
1	77.8%	81.0%	1	80.6%	100.0%	1	63.9%	81.0%		1	75.0%	85.7%			
1	77.8%	90.5%	1	77.8%	90.5%	1	66.7%	100.0%		1	75.0%	95.8%			
1	77.8%	95.2%	1	75.0%	95.2%	1	69.4%	95.2%		1	80.6%	100.0%			
1	61.1%	87.5%	1	77.8%	91.7%	1	77.8%	100.0%		1	80.6%	100.0%			
			-												

																	W	ith	in-	Spe	cif	ica	tio	ns	Sar	npl	es																						0	ff-S	peo	cifi	cat	ion	IS S	Sam	nple	s									
	VN01-A	VN01-B	VND1-C	VN02-A	VN02-B	VN02-C	VN03-A	ANU3-B	D-EDNA	VNU4-A	VNU4-B	VN04-C	A 105 P	VN05-B	VN05-C	A-90M	AND6-B	VNUD-C	VAND7-B	VN07-C	VN08-A	VN08-B	VN08-C	A-90MV	8-60N/	VNID-C	VN10-B	VN10-C	WII-A	WII-B	0-UNA	VN12-A	VN12-C	VN13-A	VN13-B	VN13-C	W14-A	VN14-B	VN14-C	VN15-B	VN15-C	W_0S1-A	W_0S1-B	VN_0S1-C	VN_0S2-A	VN_052-C	A-CSO_NV	AN_0S3-B	VN_0S3-C	VN_0S4A	VN_054.8	VN_054-C	AN DS6.B	VN_055-C	VN 056-A	VN_056-B	VN_056-C	VN_057.A	VN 057-C	VN_0S8-A	VN_0S8-B	VN_0S8-C	88-A	88 8 88 C	% Selectivity	y s	% pecificity
	1	-1	1	-1	-1	-1	1	-1	1	1	1	-1	1	1	1	1	1	1	1 '	1 1	-1	1	-1	1	1	-1	1	1	1	1	1	1 1	1 1	1	1	1	1	-1	-1	1 -1	1 -1	-1	-1	-1 -	-1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	71.1%		88.9%
	1	1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1	1 1	1	1	1	1	1	-1 1	1	1	1	1	1	1	1 1	1	1	1	1	1	-1	1 -1	1 1	-1	-1	-1	1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	71.1%		88.9%
	1	1	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	-1	1	1	1	1 1	-1	1	1	1	1	-1 :	1	1	1	1	1	1	1 1	1	1	1	1	1	-1	1 1	1	-1	-1	-1 -	-1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	84.4%		85.2%
	1	-1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1	1 1	-1	-1	-1	1	1	-1 '	1	1	1	1	1	1	1 1	1	1	1	1	-1	-1	1 -1	1 -1	-1	-1	-1	1 -	-1 1	1 -1	1 -1	-1	-1	-1 -	-1 -	-1 -	-1 -1	1 -1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	//.8%		88.9%
	1	1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1	1 1	-1	1	1	1	1	-1 1	1	1	1	1	1	1	1 1	1	1	1	-1	-1	-1	1 1	1 1	-1	-1	-1	1 -	-1 1	1 -1	1 -1	-1	-1	-1 -	-1 -	-1 -	-1 -1	1 1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	/1.1%		92.6%
		1	1	-1	-1	-1	1	-1	1	1	1	-1	1	1	1	1	1	1	1	1 1	-1	-1	-1	1	1	-1	1	1	1	1	1	1	1 1	1	1	1	1	-1	-1	1 -1	1 -1	-1	-1	-1	1 -	-1 1	-1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 -1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	80.0%		88.9%
		1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	-1	1	1	1	1 1	-1	1	-1	1	1	-1	1	1	1	1	1	1	1 1	1	1	1	1	-1	-1	] -]	11	-1	-1	-1 -	-1 -	-] ]	-	-1	-1	-1	-1 -	-] -		1 -1		-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	71.1%		92.6%
		-1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1	1 1	1	1	1	1	1	-1	1	1	1	1	1	1	11		1	1	1	-1	-1	] -]	11	1-1	-1	-1	] -	-] ]	-	-1	-1	-1	-1 -	-1 -	-1 -	1 -1		-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	/5.6%		88.9%
			-1	-1	-1	-1	1	-1	1	-		1	1	1		1	1				-1	-1	-1	1	1			1	4	1	-	¦ .						-1	-1				-1	-1	1 -		-		-1	-1	-1 -	-! -	1 -	1 -		-1	1	-1 -	-1 -		-1	늰	-1	-1 -1	80.0%		80.2%
		-1	1	-1	-1	-1	-	-1	1	1	1	1	1	-	1	1	1	1			-1		-	-	-				4	-	1	1					-	-	-1		1 I 1 1		-1	-1	-		-		-1	-1	-1 -	-1 -				-1		-1 -		1 -1	-1	-1	-1 -	-1 -1	00.0%		05.2% 05.2%
		1	÷	-1	1	-1	-		-		1	4	-	-	-	-	÷	-					- 1	4	4	1				-	-						4	-	-					1	11	1			-	-			1.7					1.	1 -		-1	4			00.0%		00.270
		1	1	_1	_1	_1	1	-1	-	1	1	1	1	1	1	-	-			, , , ,	1	1	1	-	-		. 1		-	-			1 1	. 1	. 1	1	-1	-1	1		1 1		_1	1	-1 -		1.	-1	_1	_1	-1 -	_1 _		1 1		_1	-	-1 -1		1 _1	_1	-1	-1	-1 -1	71 1%		00.970 00.970
		÷	÷	_1	_1	_1	-1	-1	-1	÷	i .	-1	÷	÷		÷	÷	÷	÷		-1	i.	÷	4	4			- 1	÷	÷	÷	÷ .				1	1	_1	_1	1 1		1	_1	_1	1				_1	_1	_1 .	_1 _		1 .1		_1	1	_1		1 _1	_1	_1	_1	_1 _1	80.0%		88.9%
		-1	i	-1	-1	-1	1	-1	1	i	i	1	i	i	i	i.	i.	i		1 1	-1	i	÷	÷	i.		1	i	÷	i.	i	÷	1 1	1	1	1	1	1	-1	ii	ii	-1	-1	-1 -	-1 -	-1 1	-1	-1	-1	-1	-1 -	-1 -	1 -	1 -1	1 1	-1	-1	-1 -	1 -	1 -1	-1	-1	-1 -	-1 -1	75.6%		85.2%
-		1	i	-1	-1	-1	-1	-1	-1	i	i .	-1	i	i	i	i.	i	i	1 ·	 1 1	1	i	i	÷	i.	-1 -	. 1	1	i.	1	i	1 ·	1 1	1	. 1	1	i	1	-1	1 -1	1 1	-1	-1	-1 -	-1 -	4 -	1 -1	1	-1	-1	-1 -	-1 -	1 -	-1 -1	1 1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	82.2%		92.6%
tion 1		-1	1	-1	-1	-1	1	1	1	1	1	1	1	1	1	1	1	1	1 ·	1 1	-1	1	1	1	i	-1 -	1 1	1	1	1	i	1 .	1 1	1	1	1	1	1	-1	1 -1	1 -1	-1	-1	-1 -	-1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	77.8%		96.3%
La	1	1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1 .	1 1	-1	1	1	1	1	-1 -	1	1	1	1	1	1 .	1 1	1	1	1	-1	-1	-1	1 1	1 1	-1	-1	-1 -	-1 -	-1 -	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 -1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	80.0%		88.9%
Ite	1	-1	-1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1 .	1 1	-1	1	1	1	1	-1 -	1	1	1	1	1	1 .	1 1	1	1	1	1	-1	-1	1 -1	1 1	-1	-1	-1 -	-1 -	-1 1	i -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	80.0%		100.0%
S	-	-1	-1	-1	-1	-1	-1	-1	-1	1	1	-1	1	1	1	1	1	1	1 .	1 1	-1	1	-1	1	1	-1 -	1	1	1	1	1	1 .	1 1	1	1	1	1	1	-1	1 -1	1 -1	-1	-1	-1 -	-1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	75.6%		88.9%
믿	-	-1	-1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	-1	1	1	1 .	1 1	-1	1	1	1	1	-1 -	1	1	1	1	1	1	1 1	1	1	1	1	1	-1	1 -1	1 -1	-1	-1	-1 -	-1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	64.4%		92.6%
ę	1	1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1 .	-1	1	1	1 .	1 1	-1	1	1	1	1	-1 -	1	1	1	1	1	1	1 1	1	1	1	1	1	-1	1 -1	1 1	-1	-1	-1 -	-1 -	-1 1	1 -1	I -1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 -1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	71.1%		92.6%
5 .	1	1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1	1 1	-1	-1	-1	-1	1	-1 :	1	1	1	1	1	1 1	1 1	1	1	1	-1	-1	-1	1 -1	1 -1	-1	-1	-1 -	-1 -	-1 -	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	80.0%		96.3%
ute	1	1	1	-1	-1	-1	1	-1	1	1	1	-1	1	1	1	-1	1	1	1 '	1 1	-1	1	1	1	1	-1 1	1	1	1	1	1	1	1 1	1	1	1	1	1	-1	1 -1	1 -1	-1	-1	-1 -	-1 -	-1 -1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	68.9%		92.6%
ō	1	-1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1 '	1 1	1	1	1	1	1	-1 1	1	1	1	1	1	1	1 1	1	1	1	-1	-1	-1	1 -1	1 1	-1	-1	-1 -	-1 -	-1 -1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	75.6%		92.6%
	1	1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1 '	1 1	-1	1	1	1	1	-1 -	1	1	1	1	1	1 .	1 1	1	1	1	1	-1	-1	1 -1	1 1	-1	-1	-1	1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	77.8%		92.6%
	1	1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1 '	1 1	-1	1	1	1	1	-1 '	1	1	1	1	1	1	1 1	1	1	1	1	-1	-1	1 -1	1 1	-1	-1	-1	1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	80.0%		88.9%
	1	-1	-1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1 '	1 1	-1	1	1	1	1	-1 :	1	1	1	1	1	1 .	1 1	1	1	1	1	1	-1	1 -1	1 1	-1	-1	-1 -	-1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	80.0%		85.2%
	1	-1	1	-1	-1	-1	1	-1	1	1	1	1	-1	1	1	1	1	1	1 '	1 1	1	1	1	1	1	-1 1	1	1	1	1	1	1 1	1 1	1	1	1	1	-1	-1	1 1	1	-1	-1	-1 -	-1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 -1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1	-1 -1	77.8%		88.9%
	1	1	-1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1	1 1	-1	1	1	1	1	-1 1	1	1	1	1	1	1	1 1	1	1	1	1	1	-1	1 1	1 1	-1	-1	-1 -	-1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 -1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	80.0%		96.3%
	1	1	1	-1	-1	-1	1	1	1	1	1	1	1	1	1	1	1	1	1	1 1	-1	1	1	1	1	-1 '	1	1	1	1	1	1	1 1	1	1	1	1	-1	-1	1 -1	1 1	-1	-1	-1 -	-1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 -1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	82.2%		96.3%
	1	1	1	-1	-1	-1	1	1	1	1	1	1	1	1	1	1	1	1	1	1 1	-1	1	1	1	1	-1 '	1	1	1	1	1	1	1 1	1	1	1	1	-1	-1	1 1	1	-1	-1	-1 -	-1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	-1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	82.2%		96.3%
	1	1	1	-1	-1	-1	1	-1	1	1	1	-1	1	1	1 .	-1	1	1	1 '	1 1	-1	1	1	1	1	-1 1	1	1	1	1	1	1	1 1	1	1	1	1	-1	-1	1 -1	1 1	-1	-1	-1 -	-1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 -1	-1	-1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	84.4%		88.9%
	1	1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	-1	1	1	1	1 -	-1	-1	-1	1	1	-1 1	1	1	1	1	1	1 1	1 1	1	1	1	1	1	-1	1 -1	1 -1	-1	-1	-1 -	-1 -	-1 1	-1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	75.6%		96.3%
		1	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1	1 1	-1	1	1	1	1	-1	1	1	1	1	1	1	1 1	1	1	1	-1	-1	-1	1 -1	1 1	-1	-1	-1	1 -	-1 1	1 -1	-1	-1	-1	-1 -	-1 -	-1 -	1 -1	1 1	-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	/1.1%		88.9%
		-1	-1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1	1 1	-1	1	-1	1	1	-1		1	1	1	1	1	1 1	1	1	1	1	-1	-1	1 -1	11	1-1	-1	-1 -	-1 -	1 -	1 -1	-1	-1	-1	-1 -	-1 -	.] -	1 -1		-1	1	-1 -	-1 -	1 -1	-1	-1	-1 -	-1 -1	77.8%		85.2%
		1	1	-1	-1	-1	1	-1	1	1	1	-1	1	-	1	1	1	1		1	-1	1	1	1	1	1		1	1	1	-		1		1	1	1	-1	-1		1	-1	-1	-1		-	U -1	-1	-1	-1	-1 -	-1 -		1 -1		-1	1	-1 -		1 -1	-1	-1	-1 -	-1 -1	/3.3%		92.0%
		1	1	-1	-1	-1	-1	-1	-1	1	1	1	-	-	1	1	1	1		1 1	-1	1	1	1	1	-		1	1	1	-		1 1	1	1	1	-1	-1	-1	1		-1	-1	-1 -	-1 -		1 -1 1 -	-1	-1	-1	-1 -	-1 -		1 -1 4 -		-1	1	-1 -		1 -1	-1	-1	-1 -	-1 -1	80.0%		00.2% 00.6%
		-1	1	-1	-1	-1	-	-	-1	-	4	1	-	-	1	1	1	-		1 1	-1	-	-	-	-	-			-	1	-		1 1		1		1	1	-				-1				-	-1	-1	-1	-1 -	-1 -		1 -1		-1	-	-1 -		-1	-1	-1	-1 -	-1 -1	75.6%		92.0% 02.6%
		- 1		-1	-1	-1	-	-1	1	1	1	1		1	1	1	1	1	1		-1	-		1	-	-1			1	1		1					-1	-1	-1	1 -1	1 1	-1	-1	-1		-1			-1	-1	-1 -	-1 -	- 1 -		- 1	-1	-	-1 -	- 1 -		-1	-1	-1 -	-1 -1	15.0%		92.070
Aggregated	1	0	1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	1	1	1	1	-1	1	1	1	1	-1 1	1	1	1	1	1	1	1	1	1	1	1	0	-1	1 -1	1 1	-1	-1	-1	0 -	-1 1	-1	-1	-1	-1	-1	-1	-1 -	-1 -1	1 1	-1	0	-1	-1 -	-1	-1	-1	-1	-1 -1	81.4%		92.0%
% WS Votes	1.0	0.6	0.8	0.0	0.0	0.0	0.9	0.1 0	).9 -	.0 1	.0 C	0.8 1	1.0 1.	1.0 1	1.0 0	).8 1	1.0 1	.0 1	0 1.	0 1.0	0.2	0.9	0.8	1.0	1.0 0	.0 1.	0 1.0	0 1.0	1.0	1.0 1	1.0 1	.0 1.	0 1.0	0 1.0	0 1.0	1.0	0.8	0.4	0.0 1	0 0.3	3 0.8	8 0.0	0.0	0.0 0	0.4 0.	.0 0.	8 0.0	0.0	0.0	0.0 0	0.0.0	0.0 0	.0 0.	.0 0.1	0 0.8	3 0.0	0.6	0.0 0	.0 0.	0.0	0.0 (	0.0	0.0 C	0.0 0.0	D		

**Figure ESI-5**. Classification predictions for *V. negundo* samples based on the outer loops in nested 5-fold cross-validation and aggregated SIMCA models (1 = within-specifications, -1 = off-specifications, 0 = abstain)