Supporting information

Machine learning powered detection of biological toxins in association with confined lateral flow immunoassay (c-LFA)

Seoyeon Choi [‡,a,b], Seongmin Ha [‡,a], Chanmi Kim [b], Cheng Nie [a], Ju-Hong Jang [c], Jieun Jang [c,d], Do Hyung Kwon [c,d], Nam-Kyung Lee [c,d], Jangwook Lee [c,d], Ju Hwan Jeong [e], Wonjun Yang [c,*] and Hyo-Il Jung [a,b,*]

[a] School of Mechanical Engineering, Yonsei University, Seoul, 03722, Republic of Korea

[b] TheDABOM Inc., Seoul, 03722, Republic of Korea

[c] Biotherapeutics Translational Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, 34141, Republic of Korea.

[d] Department of Biomolecular Science, Korea Research Institute of Bioscience and Biotechnology, School of Bioscience, Korea University of Science and Technology, Daejeon, 34113, Republic of Korea

[e] Chem-Bio Technology Center, Agency for Defense Development, Daejeon, 34186, Republic of Korea.

Reagents and materials

Monoclonal anti-human serum albumin IgG (11800; Bore Da Biotech, Korea), monoclonal anti-SEB IgG antibody (KRIBB, Korea, 2G8), monoclonal anti-RICIN IgG antibody (Kerafast, UK, EMT012), and monoclonal anti-BoNT-A IgG antibody (KRIBB, Korea, 3C5) were used as capture antibodies for the test spots on each chip. Monoclonal anti-human serum albumin IgG (11803; Bore Da Biotech, Korea), monoclonal anti-SEB IgG antibody (Fitzgerald, USA, 10-7825), monoclonal anti-RICIN antibody (PhD9, AntoXa Corp., CA), and monoclonal anti-BoNT-A IgG antibody (KRIBB, Korea, 4A4) were used as detection antibodies. An anti-mouse IgG antibody (M8642; Sigma-Aldrich) was used as the control antibody. Human serum albumin (Sigma-Aldrich, USA, A3782), Staphylococcal enterotoxin B (SEB, KRIBB, Korea), Ricinus communis agglutinin I (Vector Laboratories, USA, RL-1082), and heavy chain binding domain of Botulinum neurotoxin type A (BoNT-A, KRIBB, Korea) were used as the target analytes.

Colloidal gold nanoparticles (AuNPs, BBCG.20) with a diameter of 20 nm were purchased from Bore Da Biotech (Seoul, Korea). Sodium hydroxide (NaOH), sodium chloride (NaCl), and hydrochloric acid (HCl) were obtained from Duksan General Science (DGS, Seoul, Korea). Skim milk powder (LP0033B) was purchased from OXOID (Hampshire, UK). Sodium azide (NaN3), boric acid, phosphate-buffered saline (PBS), bovine serum albumin (BSA), Tween 20, and HEPES were purchased from Sigma-Aldrich (Missouri, USA). DABOM and the dilution buffers were obtained from DABOM (Seoul, Korea).

The chip was composed of a sample pad (Ahlstrom, 8980), a nitrocellulose membrane (Sartorius Unisart, Germany, CN140), an absorbent pad (Ahlstrom, Finland, Grade 222), and a backing card (PJ Co., Ltd., Korea, H-30). Cartridges (K030) were purchased from Shengfeng (Hebei, China).

Preparation of detection antibody-gold Nanoparticles (AuNPs) conjugate

For the conjugation of gold nanoparticles and the detection antibody, the AuNPs solution (Optical density, O.D. = 1) of 3mL was adjusted with 0.2 M borate buffer of 150 μL. To the solution, 1 μg/mL detection antibody was added to the AuNPs and gently mixed for 1 h using the multi mixer (Seoulin bioscience, SLRM-3, Korea). Then, the AuNPs-detection antibody conjugates were blocked with 10% BSA of 30 μL for 30 minutes at room temperature and centrifuged at 12,000 g for 30 minutes at 4 ℃ (Labogene, Denmark). The supernatant was discarded, and the AuNPs-detection antibody pellets were re-dispersed by adding storage buffer (0.1% BSA, 0.1% Tween 20, 0.01% NaN_3) to a concentration of O.D 3 and stored at 4 ℃ until used. Antibody-AuNP conjugation was verified by monitoring the absorbance in the range of 260–760 nm (Figure S4) using a UV-Vis spectrophotometer (TECAN, Switzerland). The same procedure was used for all the antibody-AuNP conjugates.

Table S1. The number of datasets used in the TOCBoost algorithm. A total of 70% of the entire training dataset was allocated for algorithm training, with the remaining 30% reserved for validation. The ground truth (GT) images had dimensions of 1533 × 719 × 3 pixels, and preprocessing was conducted using the ImageJ program, specifically focusing on the detection area measuring 724 × 242 × 3 pixels. The size of the images in the blind test data set was 1533 x 719 x 3 pixels and consisted of 270 images.

Class	Concentration (ng/mL)	Ground truth (GT) (1533 × 719 × 3)		Training data set & Validation data set (724 × 242 × 3)		Blind test data set (1533 × 719 × 3)	
		Parameter	image	Parameter	image	Parameter	image
SEB	0	50	210	50	120	-	90
	10	50		50		-	
	17.5	50		-		50	
	25	50		50		-	

	37.5	50		-		50	
	50	50		50		-	
	75	50		-		50	
	100	50		50		-	
Ricin	0	50		50		-	
	10	50		50		-	
	17.5	50		-		50	
	25	50	182	50	92	-	90
	37.5	50		-		50	
	50	50		50		-	
	75	50		-		50	
	100	50		50		-	
BoNT-A	0	50		50		-	
	10	50		50		-	
	17.5	50		-		50	
	25	50	208	50	118	-	90
	37.5	50		-		50	
	50	50		50		-	
	75	50		-		50	
	100	50		50		-	
3-type	24-class	1200	600	750	330	450	270

Table S2. Composition of tap water in Korea.

Components	**Tap water**
Mineral	45.00
Calcium (Ca, mg/L)	14.90
Sodium (Na, mg/L)	5.30
Magnesium (Mg, mg/L)	2.60
Potassium (K, mg/L)	1.60
Silicon (Si, mg/L)	3.60

pH	7.10
Residual Chlorine (RC, mg/L)	0.15
Fishy taste	x
Bitter taste	x

Table S3. The sensitivity and specificity for 24 classes in the test set.

Class-(Concentration)	**Performance measure**	
	Sensitivity (%)	Specificity (%)
SEB-0	92.9	94.5
SEB-10	94.8	95.2
SEB-17.5	100.0	100.0
SEB-25	96.4	97.3
SEB-37.5	100.0	100.0
SEB-50	96.4	97.3
SEB-75	95.2	96.1
SEB-100	96.4	98.8
Ricin-0	95.2	96.4
Ricin-10	95.2	96.4
Ricin-17.5	90.5	92.8
Ricin-25	94.7	95.1
Ricin-37.5	90.5	91.9
Ricin-50	95.2	96.4
Ricin-75	81.0	84.5
Ricin-100	95.2	96.4
BoNT-A-0	95.2	96.7
BoNT-A-10	95.2	96.7
BoNT-A-17.5	81.0	83.6
BoNT-A-25	94.8	95.9
BoNT-A-37.5	85.7	87.2
BoNT-A-50	95.2	96.7

BoNT-A-75	90.5	91.8
BoNT-A-100	100	99.3

Table S4. RGB index values of the TOCBoost hybrid algorithm for extracting morphological characteristics of three toxin group markers at different concentration values (0–100 ng/mL).

Concentration (ng/mL)	SEB (Sn)	Ricin (Rn)	BoNT-A (Bn)
	Pixel size (10 × 10)		
	RGB index* (<10)		
0	0.71	0.86	0.69
10	0.83	0.91	0.74
17.5	0.92	0.94	0.88
25	1.63	0.98	1.34
37.5	3.78	1.29	3.86
50	6.91	1.46	7.92
75	8.98	4.67	8.25
100	9.34	8.85	9.96

* RGB index = $(R,G,B)_{max} – (R,G,B)_{min} / 10^{n}$

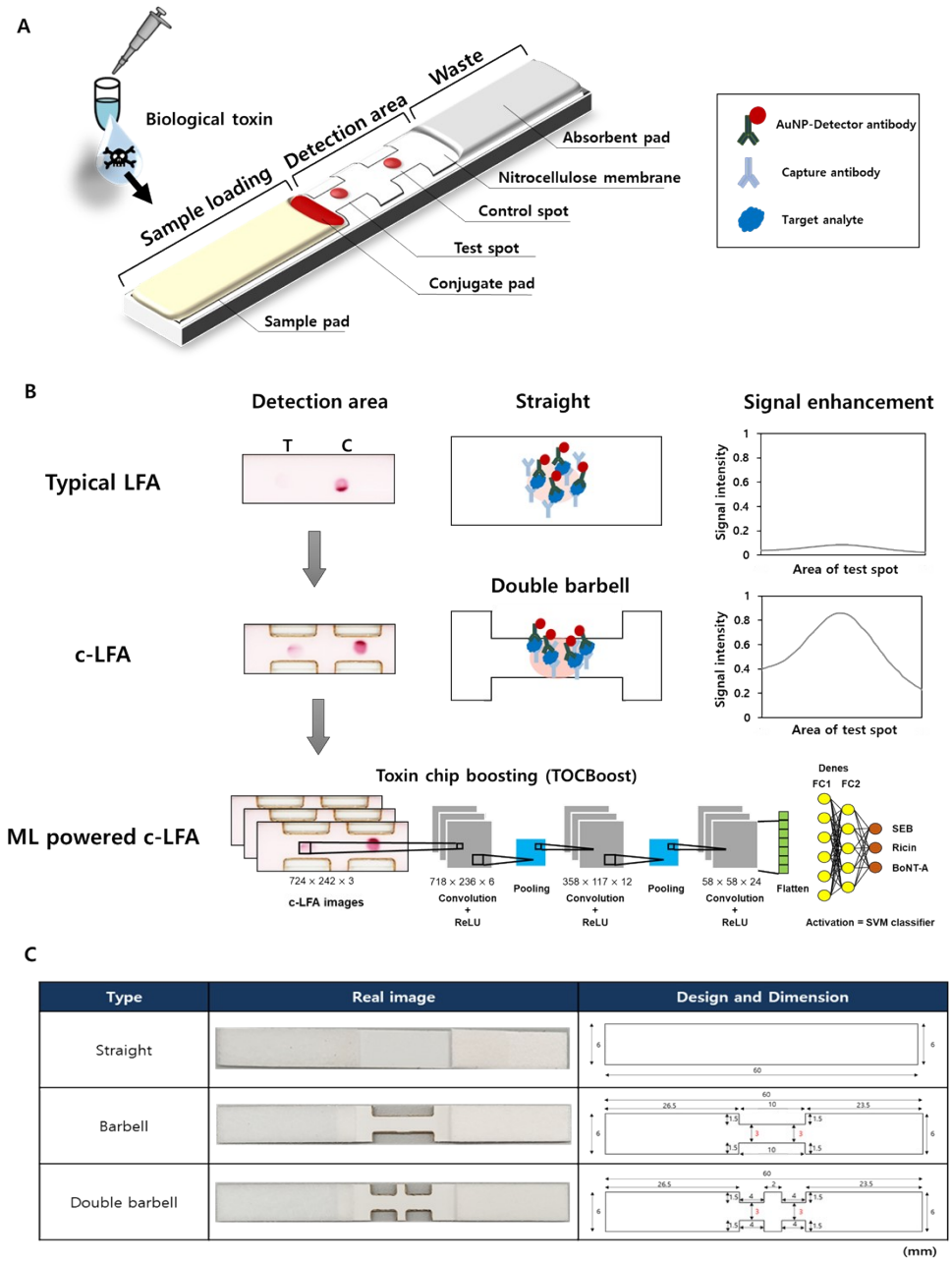

Figure S1. (A) Composition of the lateral flow immunoassay strip. (B) Schematic illustration of strategies to improve sensitivity, specificity, and accuracy of lateral flow immunoassay (LFA). LFA and c-LFA used the sandwich immunoassay using gold nanoparticles for the detection of toxins such as SEB, Ricin, and BoNT-A. Machine learning (ML) was used with c-LFA to enhance the classification accuracy of toxin classes, through a toxin chip boosting (TOCBoost). (C) Detailed design and dimension of the straight, barbell, and double barbell sensor.

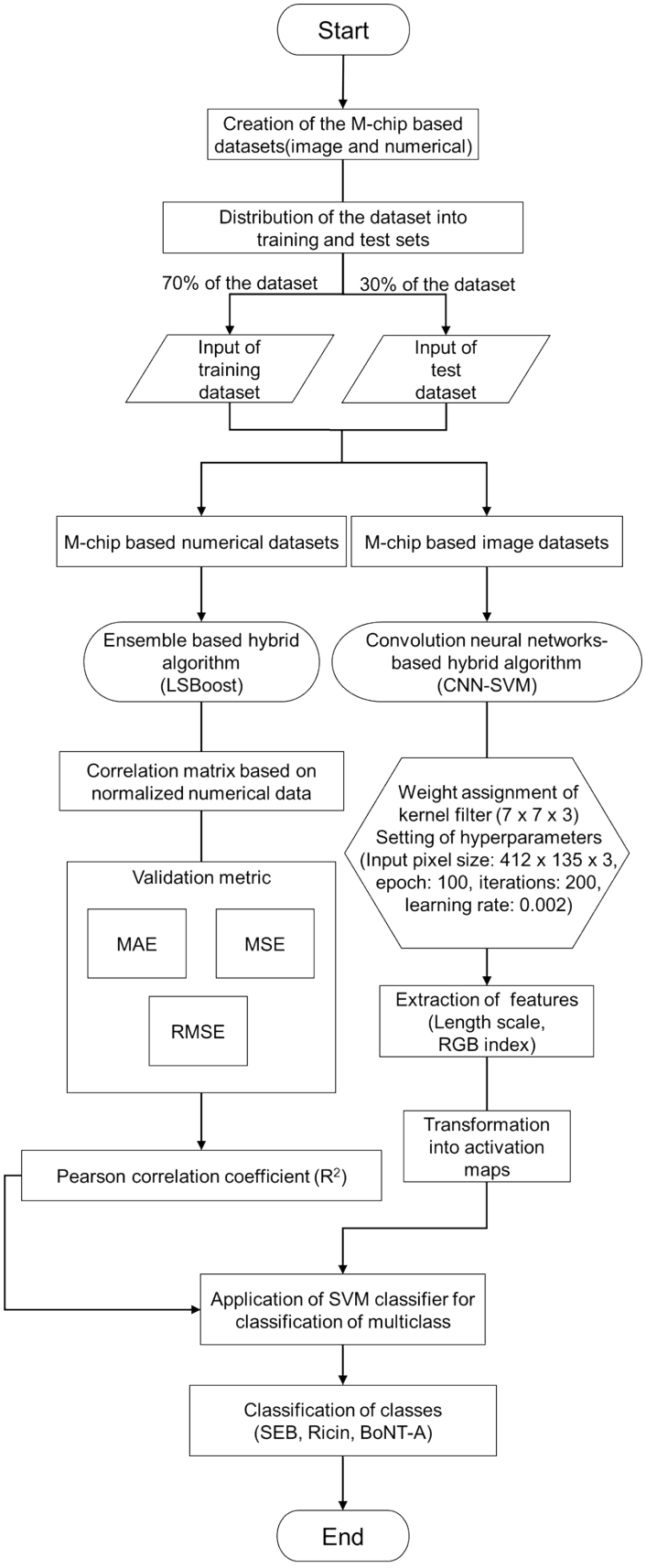

Figure S2. Flowchart of M-chip based Image dataset processing and TOCBoost hybrid algorithm.

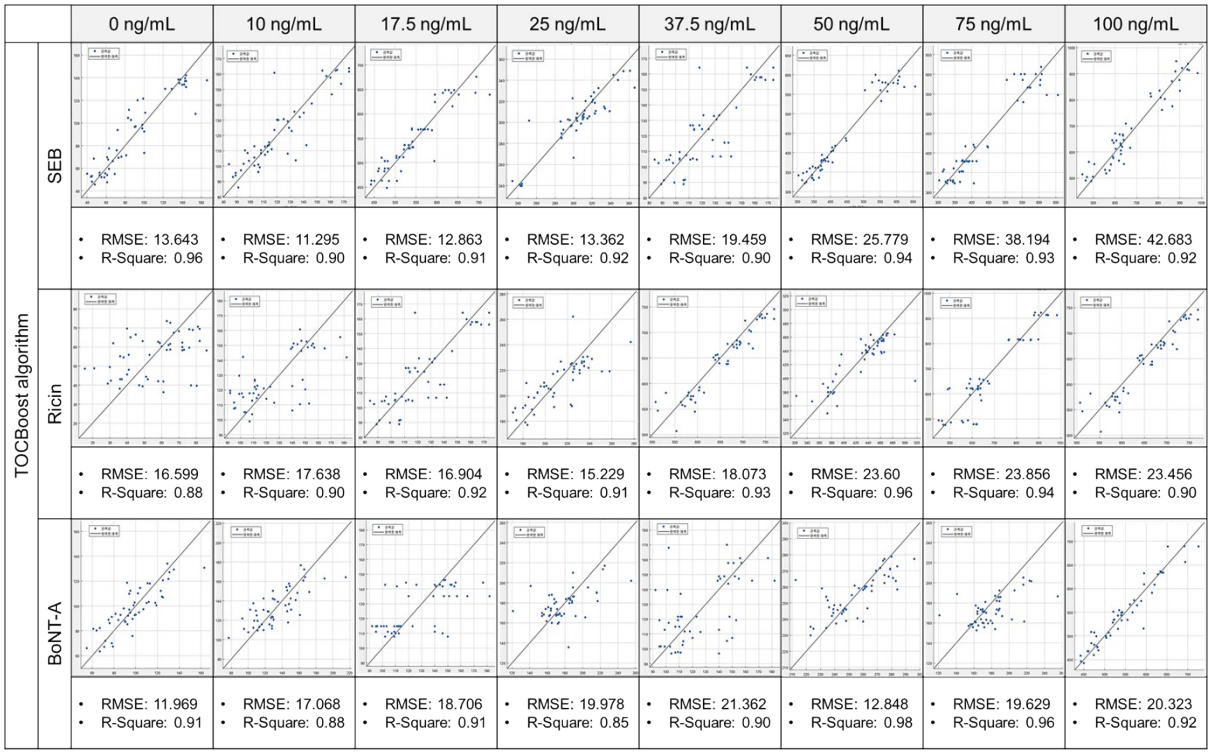

Figure S3. Pearson correlation coefficients for each concentration of three toxins (SEB, Ricin, BoNT-A) based on regression learning using the TOCBoost hybrid algorithm.

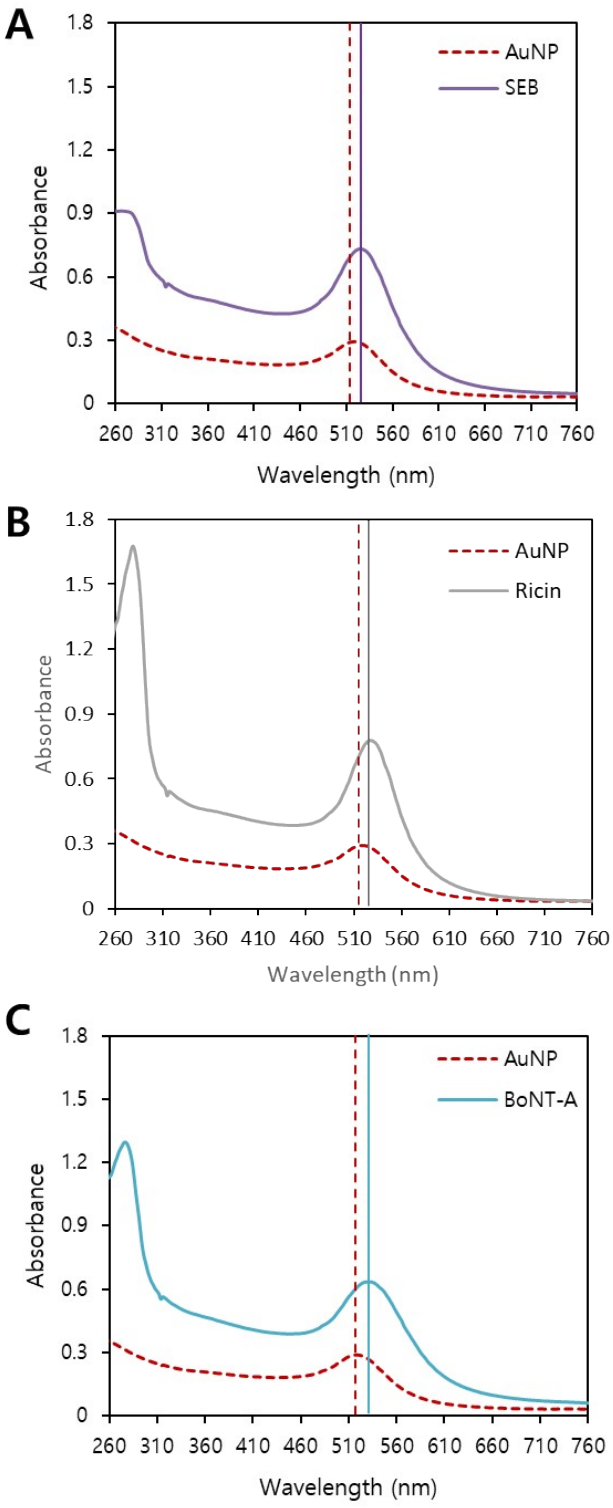

Figure S4. (A-C) Verification of antibody-AuNP conjugation: The bare AuNPs do not exhibit a signal peak at 280 nm. However, upon conjugation with antibodies, a peak can be observed at 280 nm, confirming the red shift from 520 nm to 525 nm.

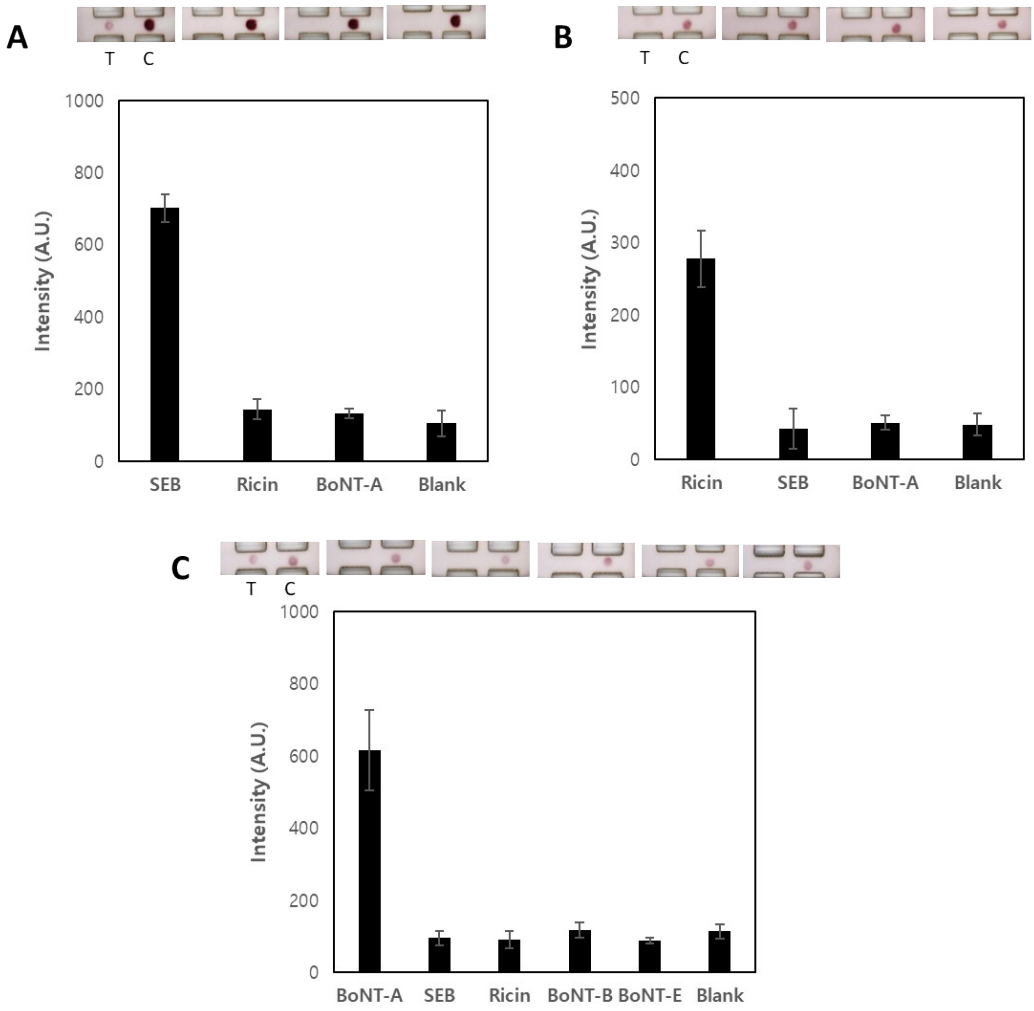

Figure S5. Specificity confirmation of the c-LFA for each toxin, such as (A) SEB, (B) Ricin, and (C) BoNT-A. The concentration of each analyte used was 50 ng/mL (n=3).