Investigating the Impact of Replicates on Non-Targeted Analysis Results in Different Food Matrices

Karen E. Butler,[1] Erica L. Bakota,[2] Christine M. Fisher,[1] Ann M. Knolhoff[1]
[1]Center for Food Safety and Applied Nutrition, Food and Drug Administration, 5001 Campus Dr, College Park, MD 20740;
[2]Kansas City Human and Animal Food Laboratory, Food and Drug Administration, 10749 W. 84th Terrace, Lenexa, KS 66214

Introduction

- Non-targeted analysis (NTA) is incredibly useful for the detection and identification of unknown compounds.
- Thousands of compounds can be detected within a single food sample with liquid chromatography coupled to high resolution mass spectrometry (LC/HR-MS); however, ensuring compounds are reliably detected and extracted from each data file is a challenge.
- This study investigates the impact of extraction and injection replicates on the quality of the data output.
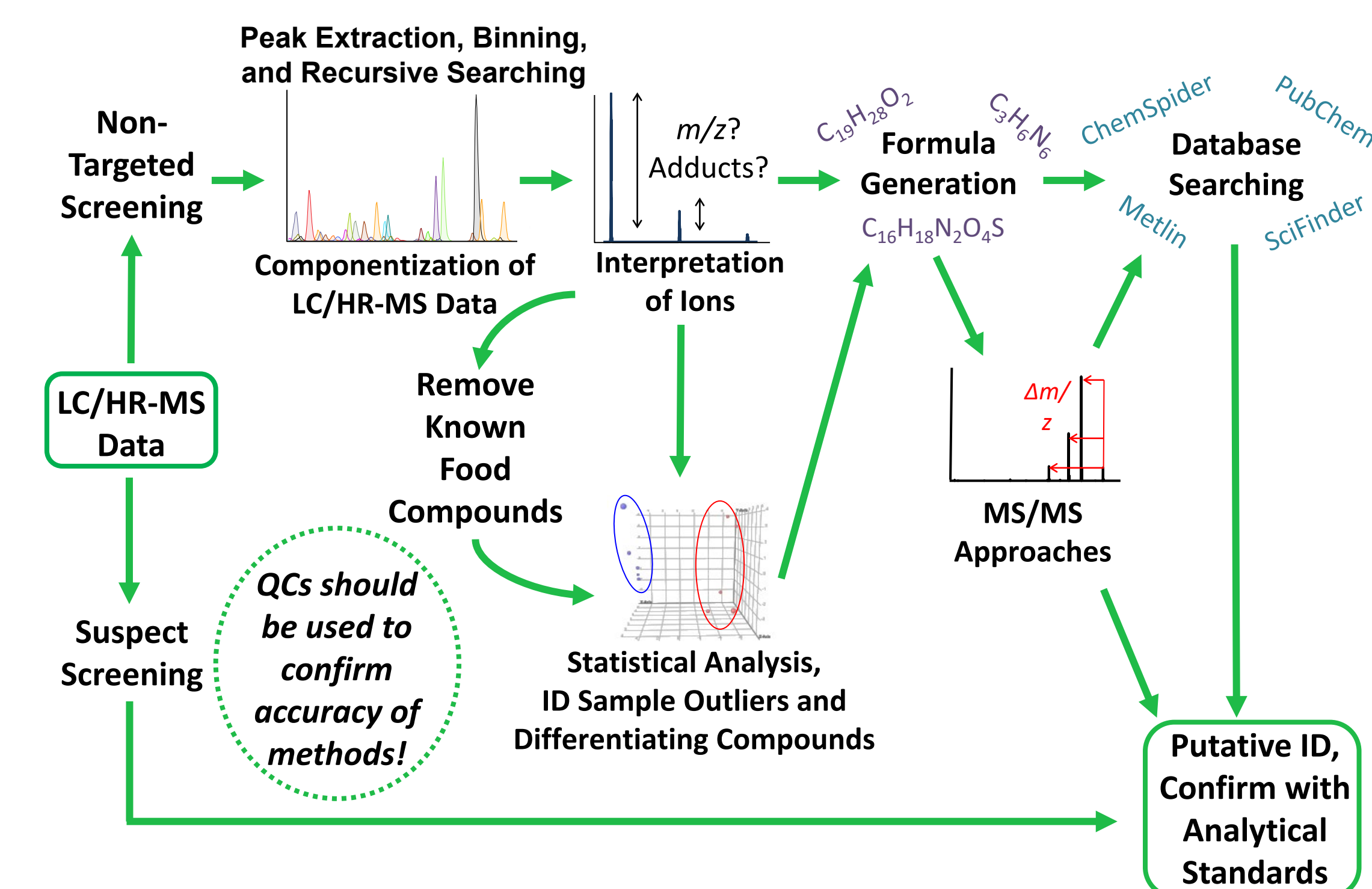

Figure 1. Data Analysis Workflow. Figure modified from Knolhoff, A. M. & Croley, T. R. *J. Chrom. A.* **2016**, *1428*, 86-96.

Materials and Methods

Analytical Strategy for Non-Targeted Analysis of Food Samples
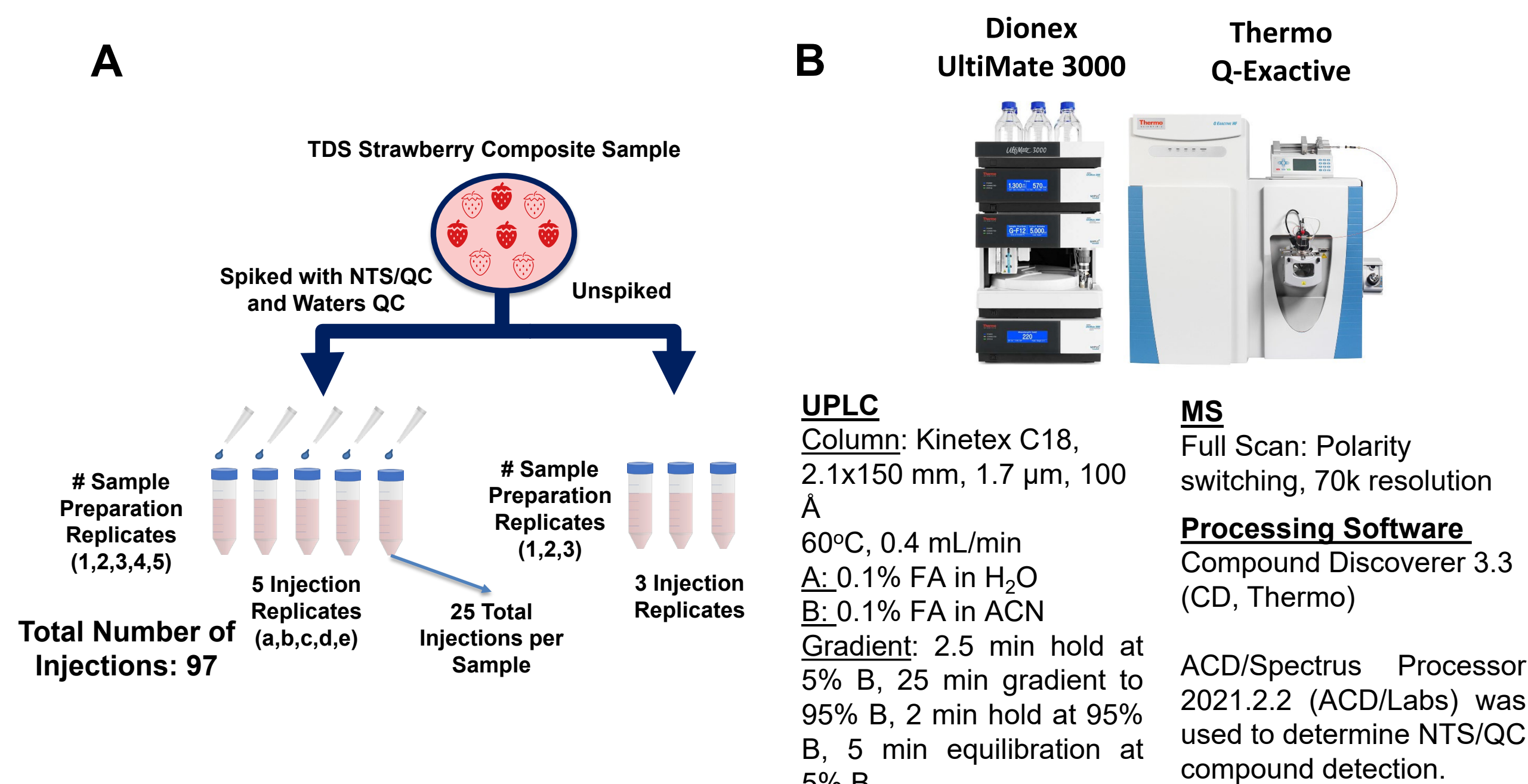

Figure 2. Experimental design (A) and instrumental parameters (B) used to analyze a pooled strawberry sample collected as part of the FDA's Total Diet Study. Each sample was spiked with a previously developed quality control standard mixture (86 compounds covering a broad range of chemical properties: NTS/QC, Knolhoff, A.M. et al, *Anal. Chem.* **2021**, 93(3), 1596-1603) and the LCMS QC Reference Standard (Waters: 9 compounds). More details for method parameters can be found here: Knolhoff, A.M. et al, *Anal. Chem.* **2021**, 93(3), 1596-1603.

Results and Discussion

Total Diet Study Matrix Selection from AOAC Foods Triangle

Image for AOAC food triangle: https://www.nist.gov/mml/csd/organic-chemical-measurement-science/primary-focus-areas/food-nutrition-and-safety/food
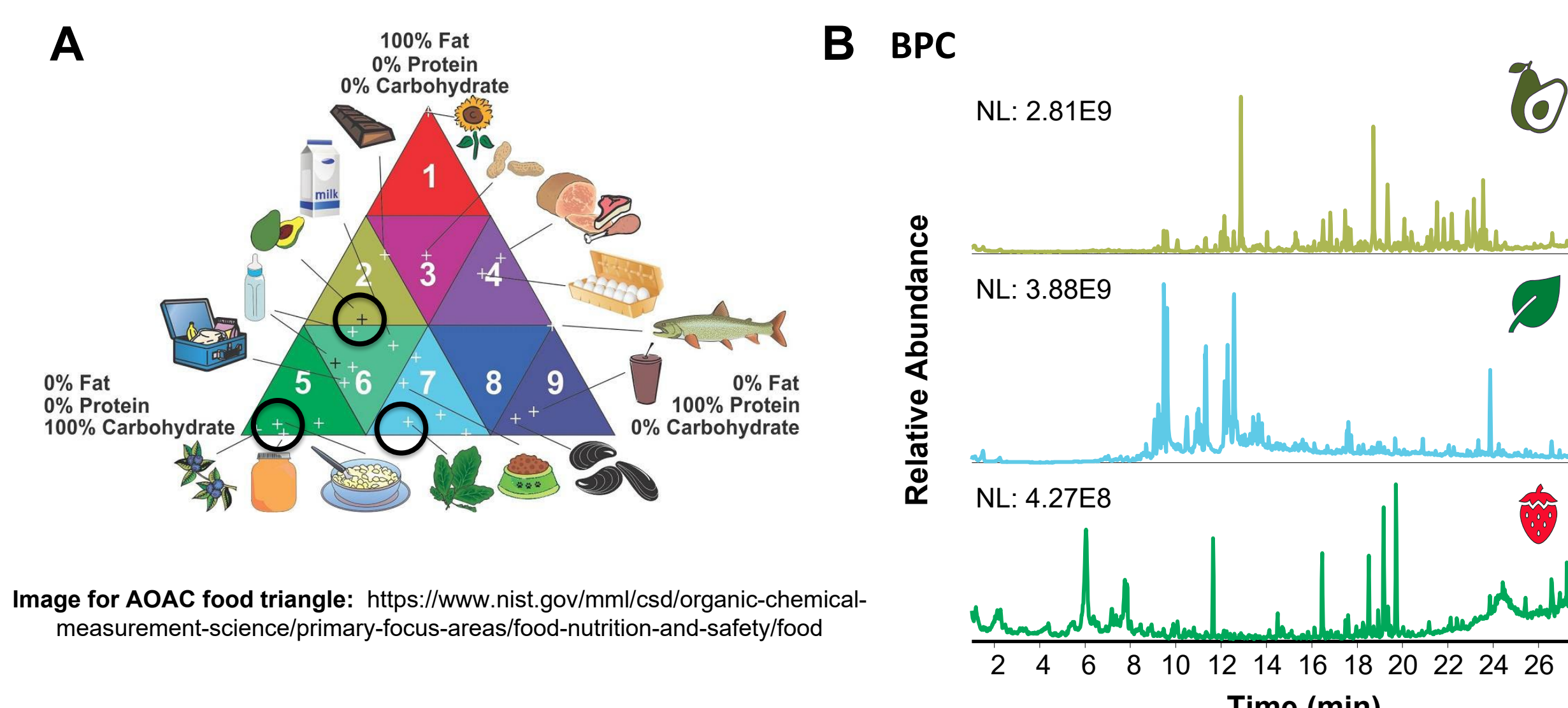
Figure 3. Matrix selection was done based on the AOAC Foods Triangle (A), which separates foods based on their fat, protein, and carbohydrate content. Initial investigations were done using strawberry (5), spinach (7), and avocado (2). (B) Qualitative differences between the three matrices were readily apparent from looking at their respective base peak chromatograms (BPC).

Monitoring Spiked QC Compounds

Matrix	Skyline (SSA) Average	Skyline (SSA) %CV	CD (NTA) Average	CD (NTA) %CV
Neat	5.23	9.85	3.71	7.63
Strawberry	7.44	8.53	6.37	8.02
Spinach	11.46	8.82	9.96	9.09
Avocado	9.44	7.74	9.08	7.04

Spiked QC compounds indicate the method/data is reproducible for all three matrices.
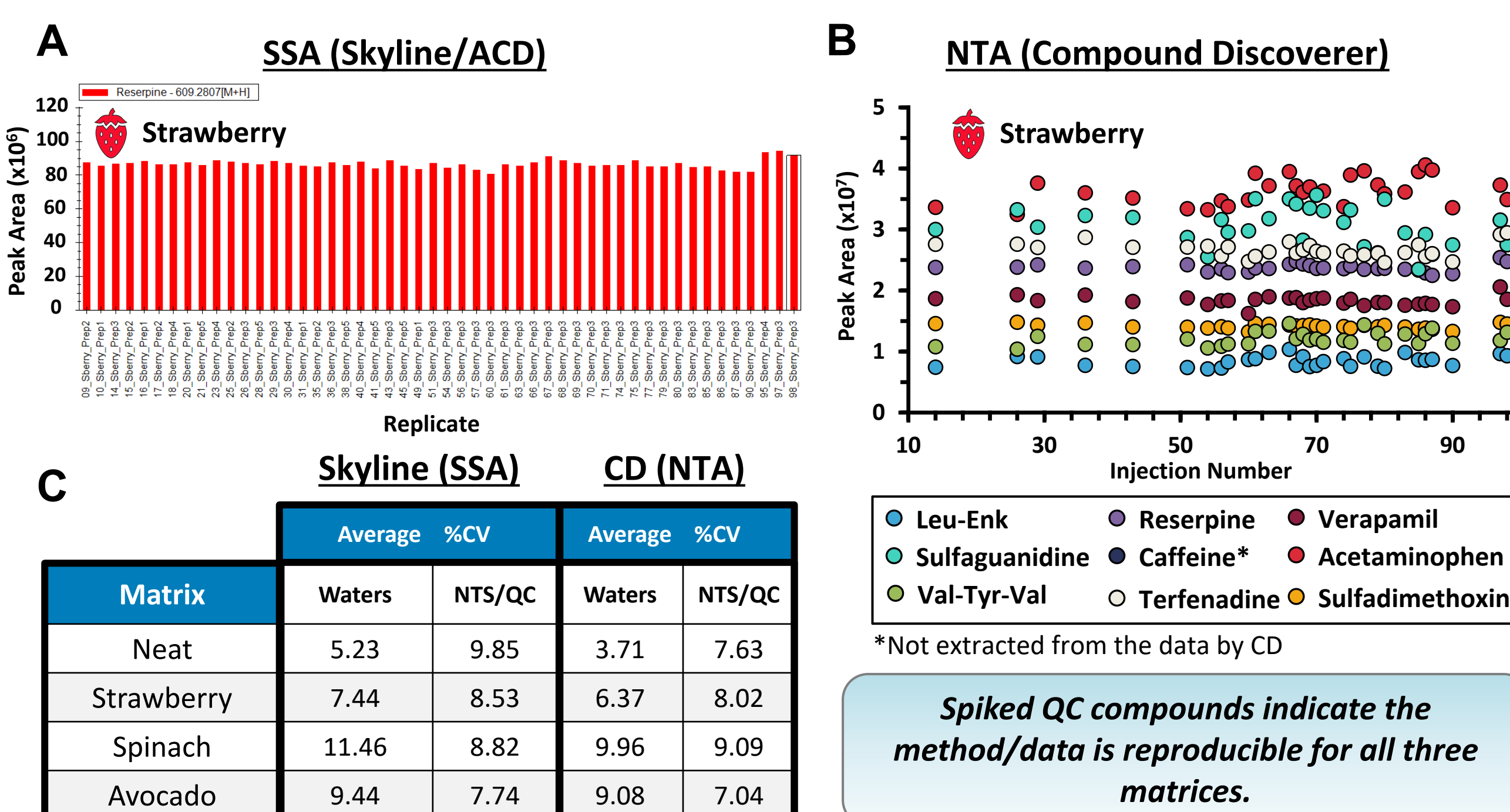
Figure 4. Spiked QC compounds were monitored to ensure instrument performance was maintained for the duration of data collection. Peak areas were assessed using both targeted (Skyline, A) and non-targeted (Compound Discoverer, B) data analysis software. Overall %CVs for peak area for QC compounds in both neat and matrix (C) were comparable between the different matrices assessed.

Table 1. Quality control compounds detected and extracted from the data by Compound Discoverer (CD).

Results File	Replicates Processed Together	NTS/QC Compounds Found (Detectable*: 81) 🍓	🥬	🥑	Waters QC Compounds Found (Detectable*: 9) 🍓	🥬	🥑
Single injection of triplicate preparations	1-3 a	68	70	73	8	8	8
	1-3 b	66	69	70	8	8	8
	1-3 c	71	72	72	8	8	8
Triplicate injections of individual preparations	1 a-c	68	72	73	8	8	8
	2 a-c	68	67	71	8	8	8
	3 a-c	67	71	73	8	8	8
Duplicate injections of duplicate preparations	1a-b, 2a-b	68	71	71	8	8	8
	1a-b, 3a-b	67	71	73	8	7	8
	2a-b, 3a-b	69	72	72	8	8	8

*Detected using ACD/Labs Spectrus Processor
1,2,3,4,5: extraction replicates
a,b,c: injection replicates

Investigating the Use of Replicates for Feature Filtering Strategies

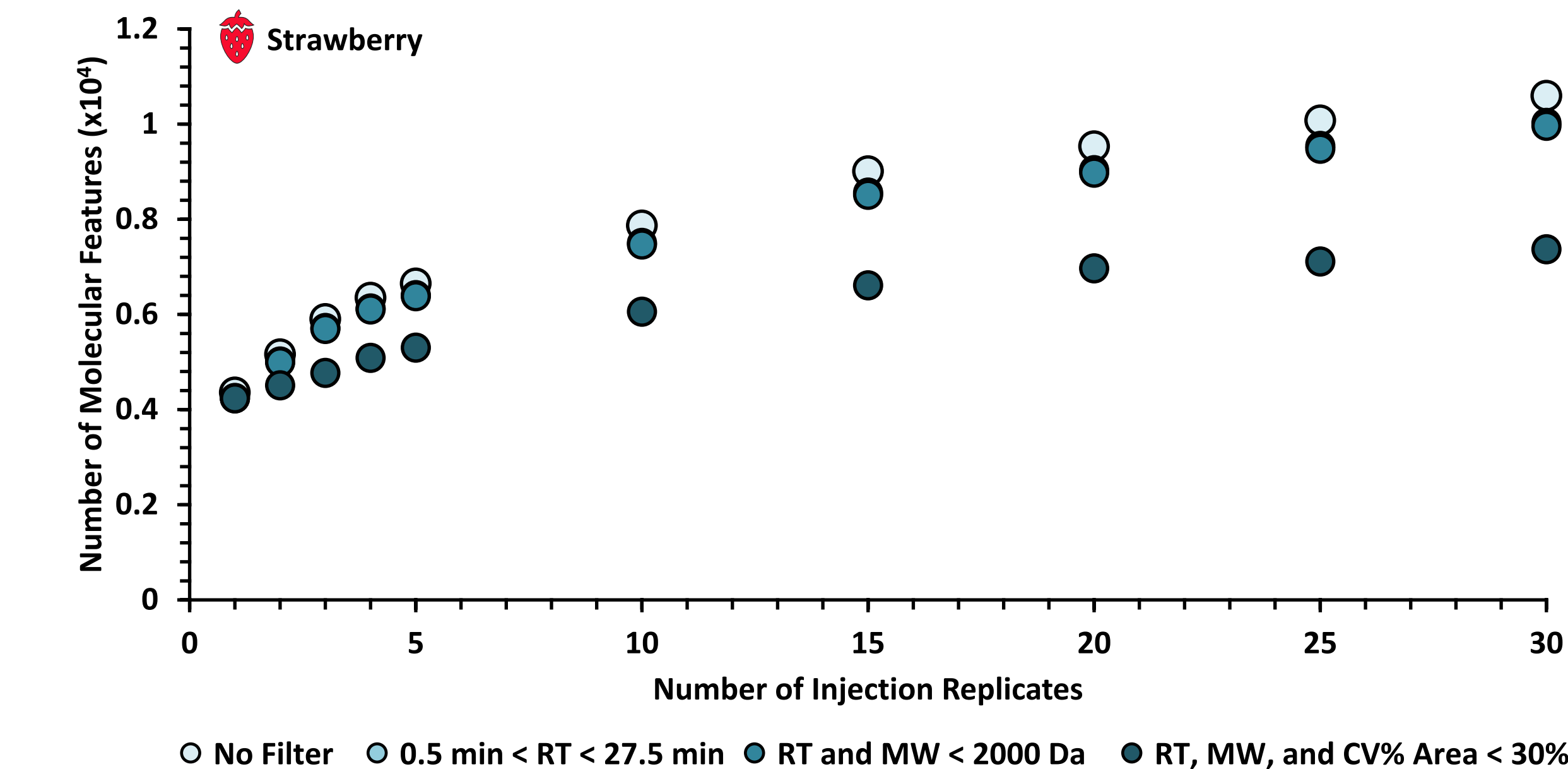○ No Filter ● 0.5 min < RT < 27.5 min ● RT and MW < 2000 Da ● RT, MW, and CV% Area < 30%

Figure 5. While metrics like retention time (RT) and molecular weight (MW) cut-off do not require replicates, having replicates can allow for the use of additional metrics, such as %CV for peak area, to be used to prioritize features for further analysis.

3 Injection Replicates	4 Injection Replicates	5 Injection Replicates
🥑 14435 Features (+ 17%)	🥑 15356 Features (+ 25%)	🥑 16163 Features (+ 31%)
🥬 10477 Features (+ 8%)	🥬 10591 Features (+ 9%)	🥬 11435 Features (+ 18%)
🍓 4771 Features (+ 13%)	🍓 5087 Features (+ 20%)	🍓 5302 Features (+ 25%)

Filtered by retention time (0.5 min < RT < 27.5 min), calculated MW < 2000 Da, and ≤ 30% CV for peak area
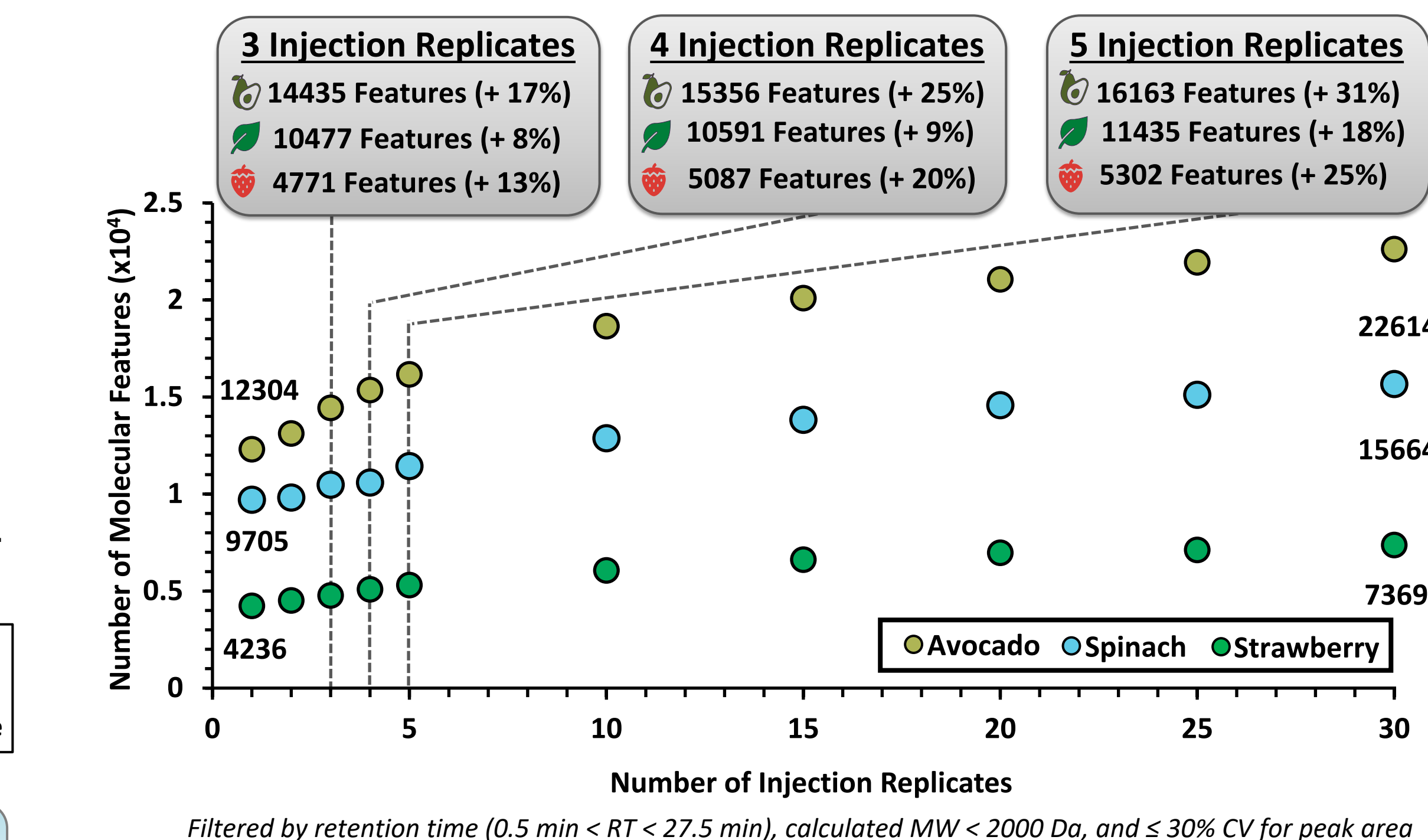
Figure 6. Molecular features (total) extracted from the data by CD increases with increasing numbers of injection replicates processed together, regardless of matrix. This is due to recursive searching of features from other samples in the analysis queue.

Injection vs. Preparation Replicates for NTA Assessments

🍓 Strawberry

Injection Replicate Molecular Features (n=3 files processed together)

Prep 1 a-c	Prep 2 a-c	Prep 3 a-c	Average	%RSD
5293	5336	4771	5133	6.13

Preparation Replicate Molecular Features (n=3 files processed together)

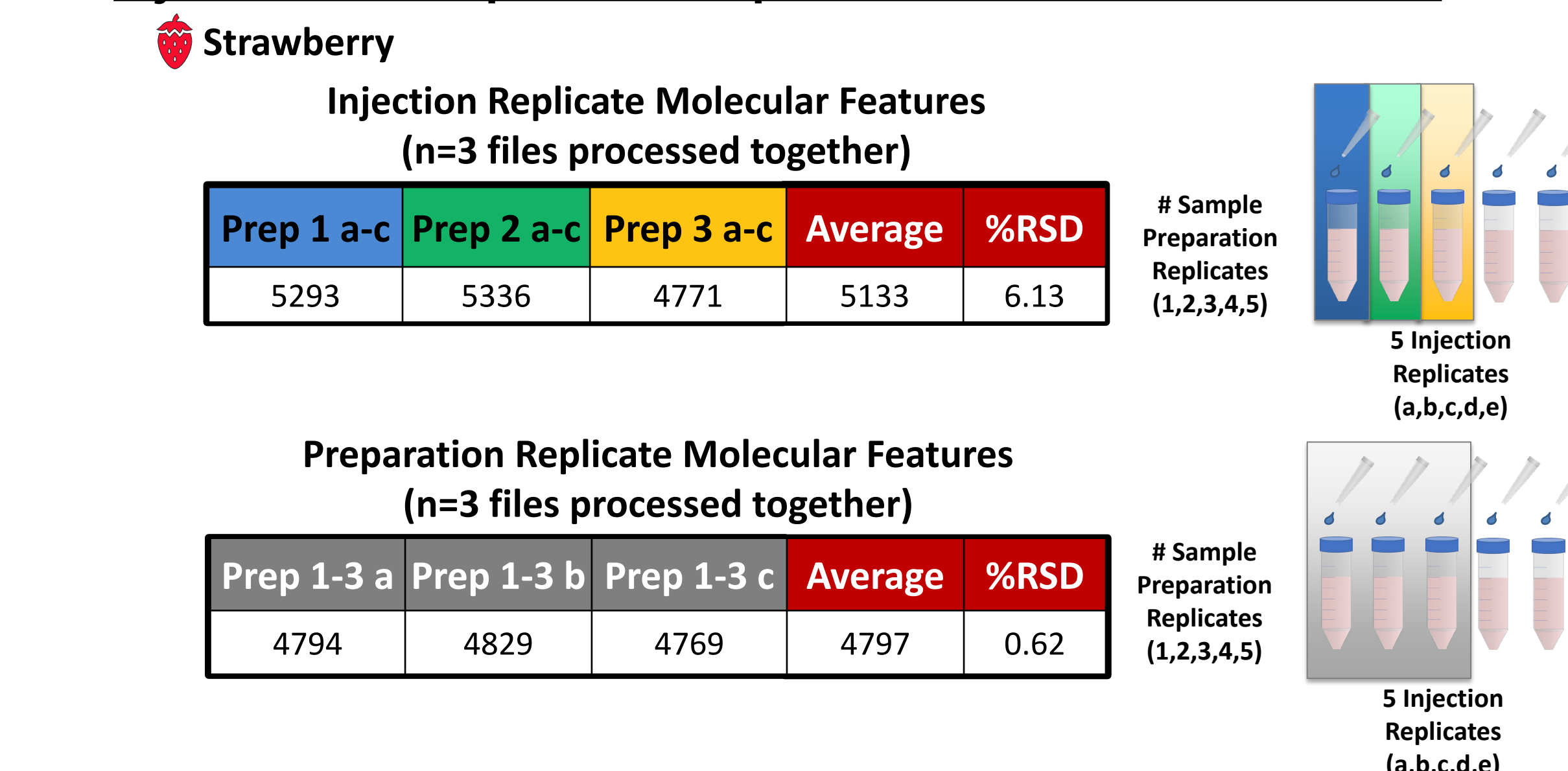| Prep 1-3 a	Prep 1-3 b	Prep 1-3 c	Average	%RSD
4794	4829	4769	4797	0.62

Figure 7. A comparable number of molecular features were observed between injection and preparation replicates processed together in CD.

Table 2. Extracted molecular features from the data by CD for injection vs. preparation replicate combinations

Results File	Replicates Processed Together	Average 🍓	🥬	🥑	%RSD 🍓	🥬	🥑
Single injection of triplicate preparations	1-3 a, 1-3 b, 1-3 c	4794	10933	4268	0.63	1.64	13.80
Triplicate injections of individual preparations	1 a-c, 2 a-c, 3 a-c	5133	11180	12611	6.1	5.49	32.98
Duplicate injections of duplicate preparations	1a-b, 2a-b, 1a-b, 3a-b, 2a-b, 3a-b	5022	11688	6927	1.8	4.50	89.60

1,2,3,4,5: extraction replicates
a,b,c: injection replicates

The number of molecular features extracted from the data is similar between the three conditions for both strawberries and spinach, but not for avocado.

Preparation Replicate Comparison by Matrix

● Prep 1 ● Prep 2 ● Prep 3 ● Prep 4 ● Prep 5

D	Strawberry	Spinach	Avocado
Min. Sample Mass	2.50 g	2.50 g	2.49 g
Max. Sample Mass	2.52 g	2.53 g	2.51 g
Avg. Sample Mass	2.51 g	2.52 g	2.50 g
Stdev	0.01 g	0.01 g	0.01 g
%RSD	0.32	0.48	0.50

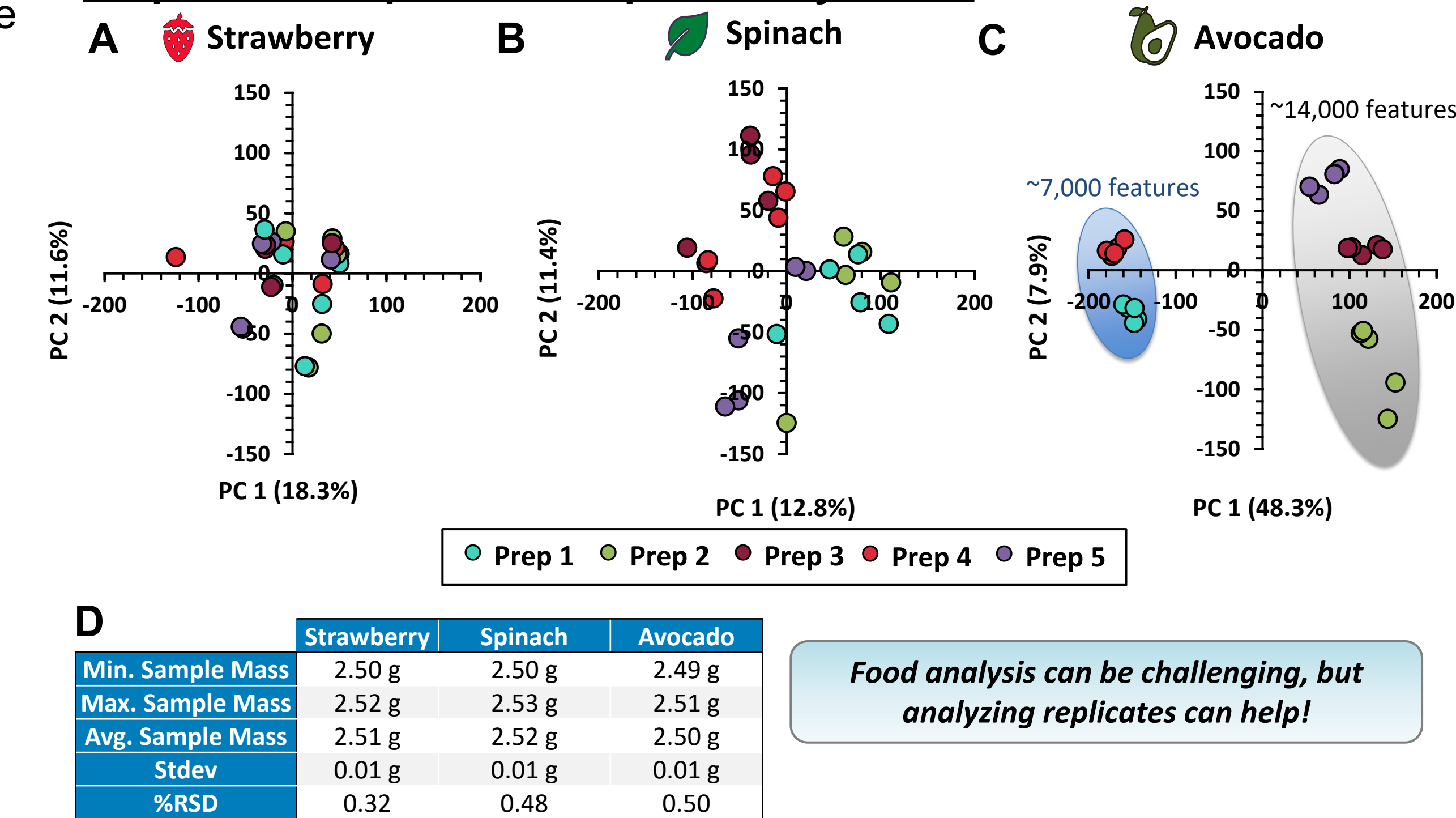*Food analysis can be challenging, but analyzing replicates can help!*

Figure 8. PCA plots for five injection replicates of five preparation replicates for strawberry (A), spinach (B), and avocado (C). While there was little differentiation between preparation replicates in the strawberry and spinach samples, there was clustering of the individual preparation replicates for the composite avocado sample, with preparation replicates 1 and 4 observed to have half the overall number of molecular features extracted from the data by CD than the other preparations. Examination of the sample amounts prepared for each matrix (D) combined with a comparable number of QC extracted from the data by CD (Table 1) did not reveal obvious differences between the three matrices, which points to the need for using replicates for NTA assessments for novel food matrices.

Conclusion

How many and what type of replicates should be run in NTA?

- The answer to this question is complicated and may be situational!
- We have observed an increase in features with increased number of replicates processed together in CD.
- There are many sources of variability, but replicates combined with standards can help address this issue.
- Replicates may be especially beneficial when analyzing a new matrix type, especially one that is challenging to homogenize.