Supplementary Table 1. Assessment of improvement in the number of annotations of small molecules (m/z < 600 Da) in three public datasets: Global FoodOmics (MSV000085226), Evaluation dataset (MSV000083306) and cheetah fecal metabolome dataset (MSV000082969).

Dataset	**# features (m/z < 600)**	**# spectral annotations; cosine score > 0.7 (%)**	**# fingerprints; Zodiac score > 0.98 (%)**	% **fingerprints with structures (%)**
Global FoodOmics	1067	210 (19.68)	663 (62.13)	90.64 (56.32)
Evaluation dataset	5062	953 (18.8)	4399 (86.90)	91.67 (79.67)
Cheetah Fecal Metabolome	903	317 (35.10)	684 (75.74)	80.40 (60.90)

Supplementary Table 2. An assessment of the quality of the CSI:FingerID fingerprint-based chemical classifications in three public datasets: Global FoodOmics (MSV000085226), Evaluation dataset (MSV000083306) and cheetah fecal metabolome dataset (MSV000082969). We report the percent agreement in the chemical taxonomy for the features that had both MS/MS library matches and CSI:FingerID structural annotations.

Dataset	% Kingdom	% Superclass	% Class	% Subclass	% Direct Parent
Global FoodOmics	100	78	65	66	58
Evaluation dataset	100	86	70	41	26
Cheetah Fecal Metabolome	100	70	65	58	36

Supplementary Table 3. The composition of samples in the evaluation dataset.

Sample Name	Mixture type	% Fecal_1	% Fecal_2	% Tomato	% NIST_1950_SRM
F1	Unmixed	100	0	0	0
F1-3_To-1	Binary mixture	75	0	25	0
F1-1_To-1	Binary mixture	50	0	50	0
F1-1_To-3	Binary mixture	25	0	75	0
To	Unmixed	0	0	100	0
To-3_NIST-1	Binary mixture	0	0	75	25
To-1_NIST-1	Binary mixture	0	0	50	50
To-1_NIST-3	Binary mixture	0	0	25	75
NIST	Unmixed	0	0	0	100
F2-1_NIST-3	Binary mixture	0	25	0	75
F2-1_NIST-1	Binary mixture	0	50	0	50
F2-3_NIST-1	Binary mixture	0	75	0	25
F2	Unmixed	0	100	0	0
F1-1_F2-3	Binary mixture	25	75	0	0
F1-1_F2-1	Binary mixture	50	50	0	0
F1-3_F2-1	Binary mixture	75	25	0	0
F1-3_NIST-1	Binary mixture	75	0	0	25
F1-1_NIST-1	Binary mixture	50	0	0	50
F2-1_To-3	Binary mixture	0	25	75	0
F1-1_NIST-3	Binary mixture	25	0	0	75
F2-1_To-1	Binary mixture	0	50	50	0
F2-3_To-1	Binary mixture	0	75	25	0
F1-2_F2-1_NIST-1_To-1	Quaternary mixture	40	20	20	20
F1-1_F2-1_NIST-1_To-2	Quaternary mixture	20	20	40	20
F1-1_F2-1_NIST-2_To-1	Quaternary mixture	20	20	20	40
F1-1_F2-2_NIST-1_To-1	Quaternary mixture	20	40	20	20
F1-1_F2-1_NIST-1_To-1	Quaternary mixture	25	25	25	25

Supplementary Table 4. A comparison of Bray-Curtis, CSCS, Cosine-UniFrac, Fingerprint-UniFrac for clustering samples in Global Foodomics dataset (N=126). The distances within- and between-sample groupings were compared using a one-sided permutational ANOVA (PERMANOVA) test; pseudo F-statistic are shown below.

Metadata Column	**Bray-Curtis**	CSCS	Cosine-UniFrac	Fingerprint-UniFrac
Food ontology level 1	4.33612	4.63819	4.82414	5.34768
Food ontology level 2	4.22387	4.53509	4.75484	5.15483
Food ontology level 3	4.28973	4.74771	4.73052	4.85554
Food ontology level 4	4.30789	5.76645	5.038	5.35642