MetaNetX Chemicals Summary

Moritz E. Beber

2021-06-04

Contents

Intro

A summary of the raw tables downloaded from MetaNetX.

Properties

Table 1: The number of chemicals with properties and sources for them in MetaNetX.

Number of Chemicals	with formula	with charge	with SMILES	with mass	with InChI	with InChIKey	Number of Namespaces
1,043,606	1,000,166	1,000,166	1,000,017	814,413	803,265	803,265	17

Table 2: The percentage of chemical information in MetaNetX.

Percent of Chemicals	with formula	with charge	with SMILES	with mass	with InChI	with InChIKey
100.00%	95.84%	95.84%	95.82%	78.04%	76.97%	76.97%

Table 3: The source namespaces from which chemicals are included in MetaNetX.

Namespace	Frequency
bigg.metabolite	6,911
chebi	104,207
envipath	8,468
hmdb	77,611
kegg.compound	2,680
kegg.drug	6,567
kegg.environ	563
kegg.glycan	10,737

Namespace	Frequency
lipidmaps	27,397
metacyc.compound	9,469
metanetx.chemical	3
reactome	2,721
rhea.generic	1,473
rhea.polymer	228
sabiork.compound	5,436
seed.compound	6,243
slm	772,892

Table 4: Chemical formulae that are not fully determined.

Number of Formulae	with `*`	with `R`	with `Z[z]`
1,000,166	11,063	0	0

Table 5: InChIs that are not fully determined.

Number of InChIs	with `*`	with `R`	with `Z[z]`
803,265	1,943	0	0

Table 6: SMILES that are not fully determined.

Number of SMILES	with `*`	with `R`	with `Z[z]`
1,000,017	196,889	0	0

Table 7: Duplicated InChIs that are not fully determined.

Number of InChIs	with `*`	with `R`	with `Z[z]`
181	10	0	0

Cross-References

Table 8: Chemicals with cross-references to other databases.

Number of Cross-References	Number of Namespaces
2,437,660	17

Table 9: The namespaces to which cross-references from MetaNetX are established.

Namespace	Frequency
bigg.metabolite	27,304
chebi	250,829

Namespace	Frequency
envipath	24,612
hmdb	195,008
kegg.compound	56,019
kegg.drug	33,441
kegg.environ	2,592
kegg.glycan	33,126
lipidmaps	86,170
metacyc.compound	40,592
metanetx.chemical	6
reactome	11,052
rhea.generic	1,494
rhea.polymer	228
sabiork.compound	17,888
seed.compound	101,985
slm	1,555,314

Deprecated MetaNetX Identifiers

Table 10: The total number of chemicals and how many deprecated identifiers there are.

Number of Chemicals	Number of Deprecated Identifiers
42512	40600

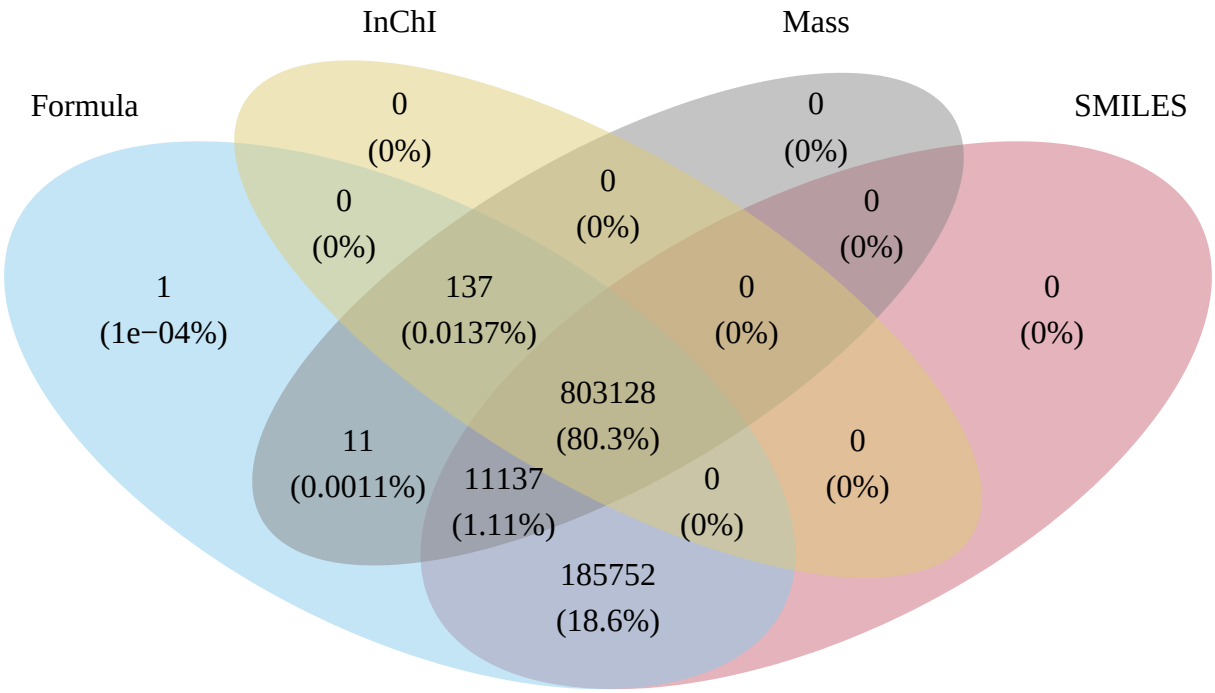

Figure 1: Venn diagram of structural annotation and mass.

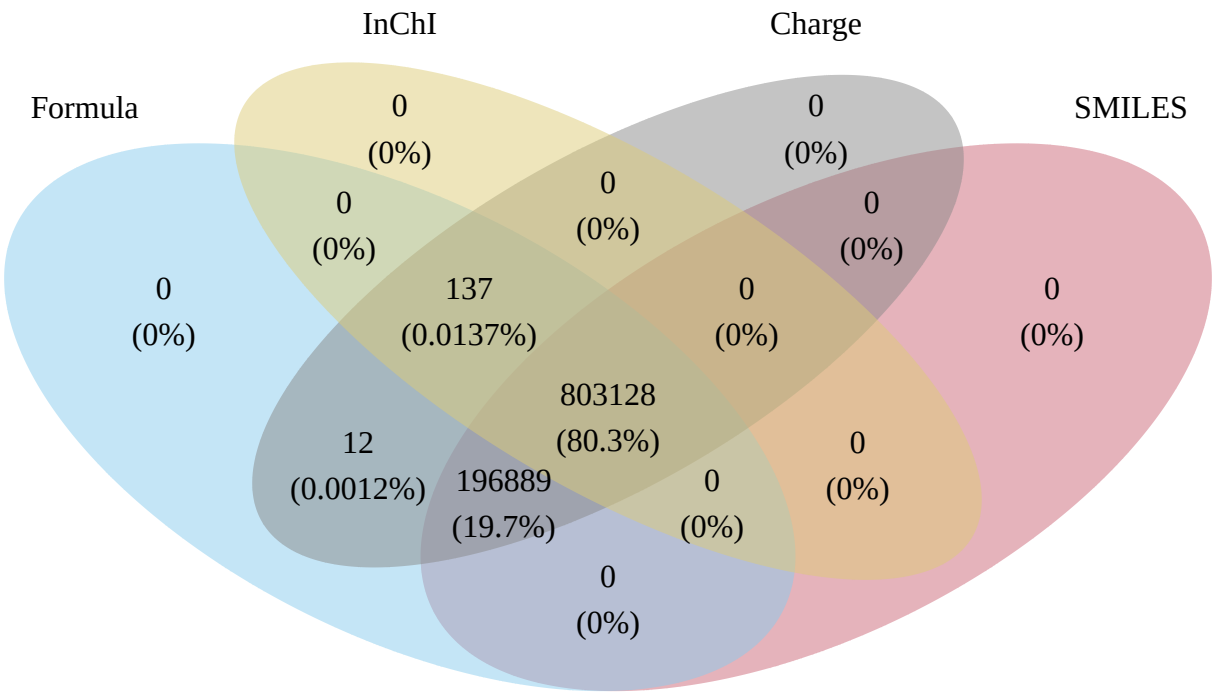

Figure 2: Venn diagram of structural annotation and electric charge.

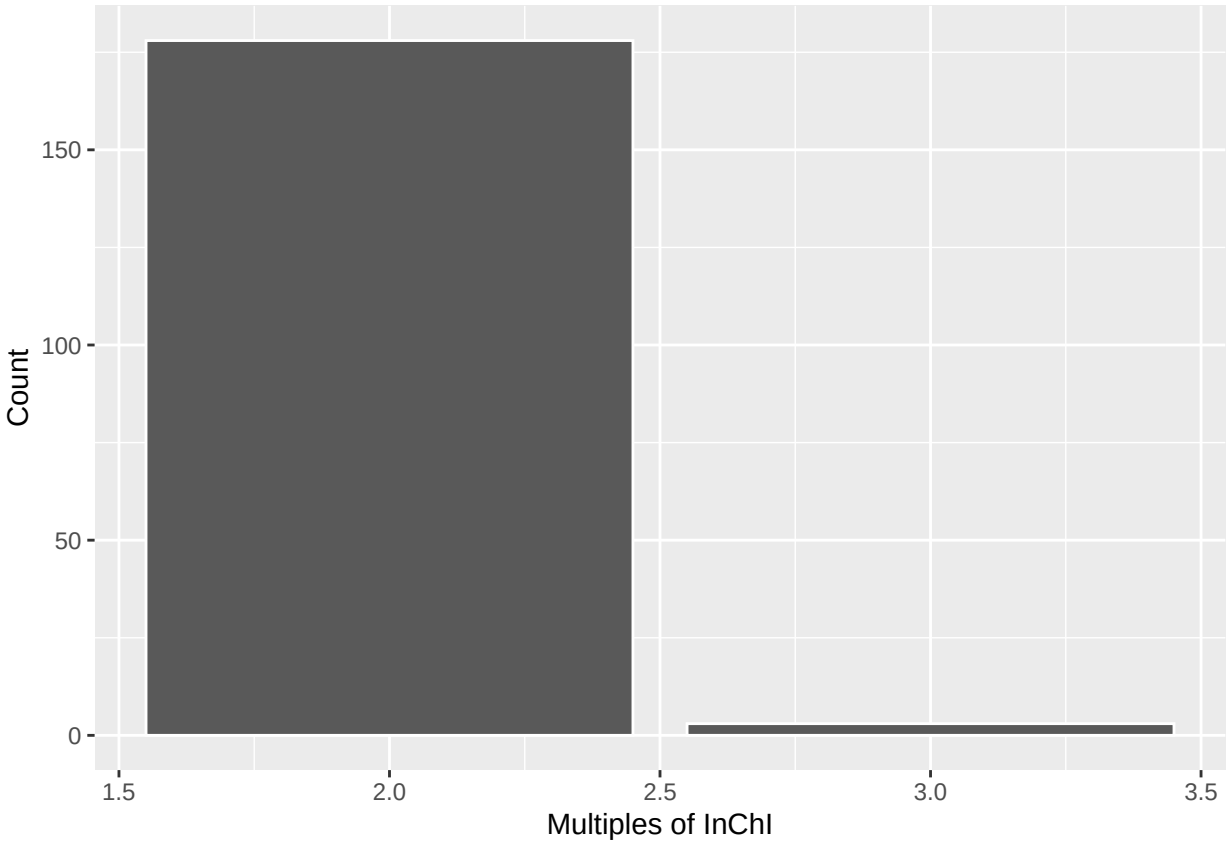

Figure 3: Multiple occurrences of the same InChI.

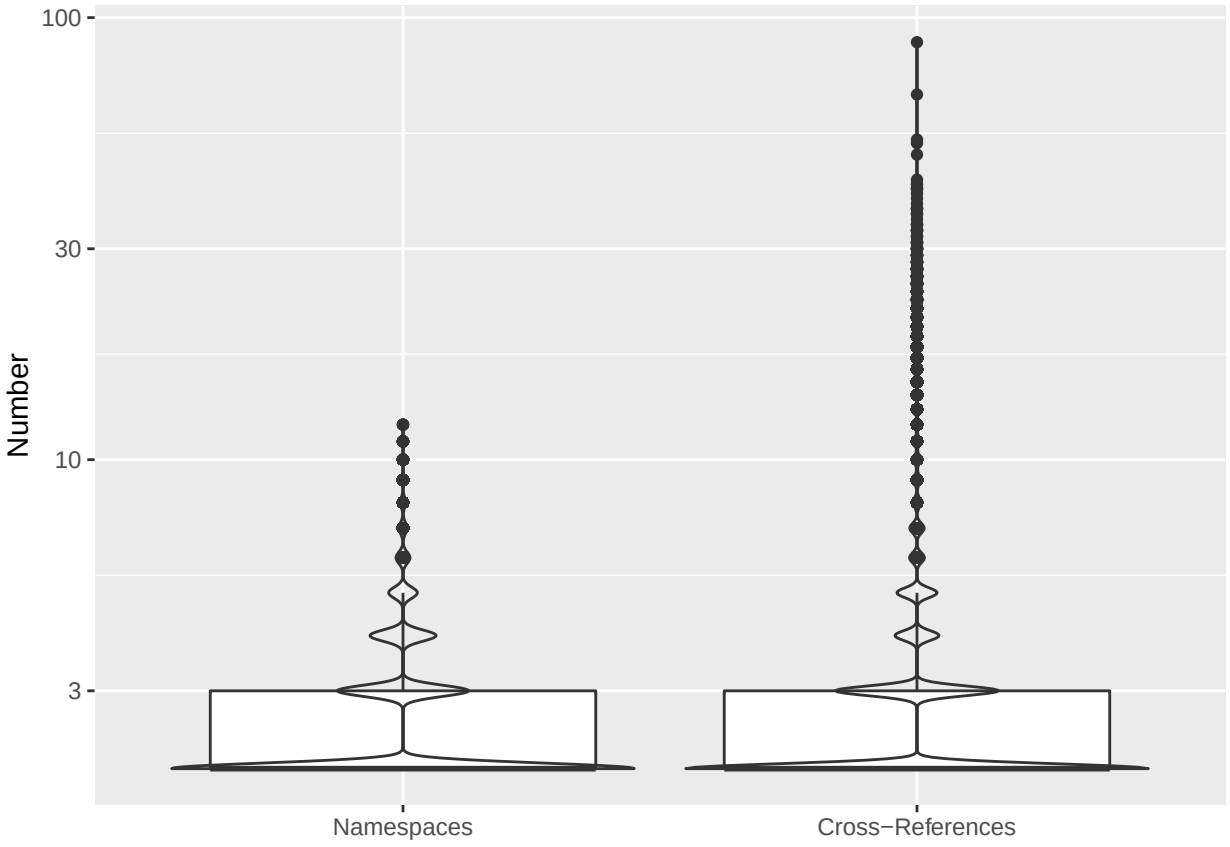

Figure 4: Number of namespaces and cross-references per chemical in MetaNetX. Only shown for chemicals where this number is greater than one.

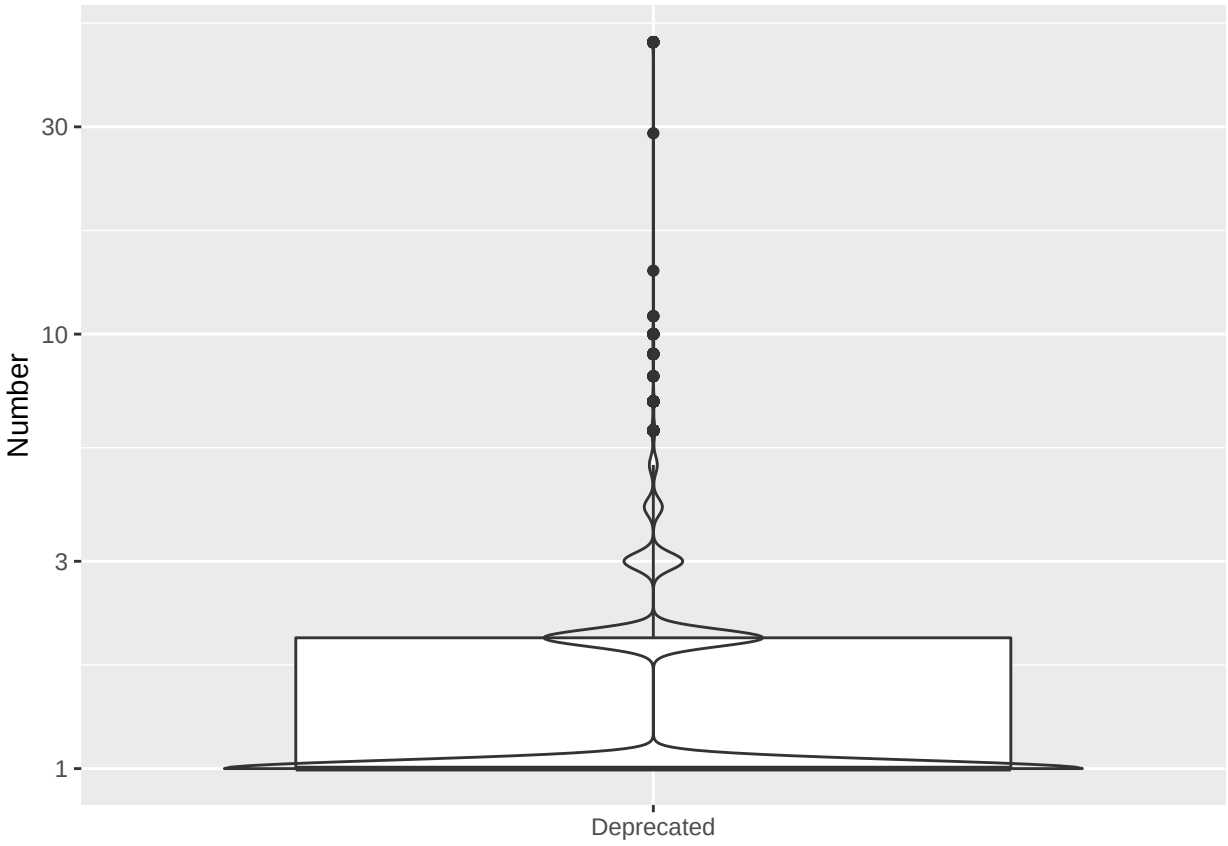

Figure 5: Number of deprecated MNX identifiers per chemical