Reactions Summary

Moritz E. Beber

2021-06-03

Contents

Intro

A summary of the final database content.

Transformation Steps:

- Deduplicated identifiers per MNX reaction and namespace
- Deduplicated names per MNX reaction and namespace
- Added deprecated MNX identifiers
- Added reaction names from KEGG, ModelSEED, and BiGG
- Added further EC numbers from ExPASy

Annotation

Table 1: Overall number of identifiers and of unique source namespaces.

Identifiers	Unique Namespaces
348,883	8

Table 2: Number of identifiers per source namespace. Identifiers are deduplicated compared to raw tables.

Namespace	Frequency
bigg.reaction	56,340
ec-code	37,311
kegg.reaction	11,160
metacyc.reaction	17,198
metanetx.reaction	124,859
rhea	50,042
sabiork.reaction	8,118
seed.reaction	43,855

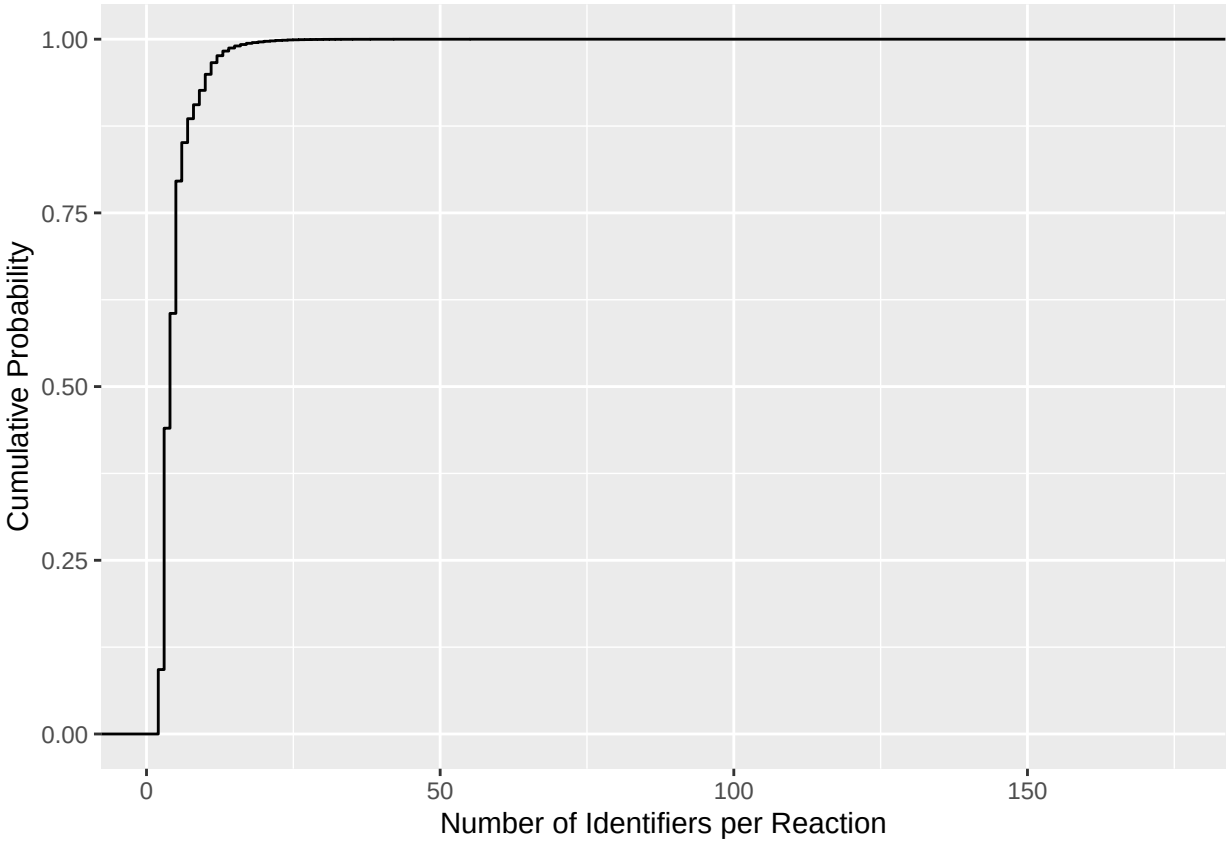

Figure 1: The empirical cumulative distribution function (eCDF) of the number of distinct identifiers per reaction.

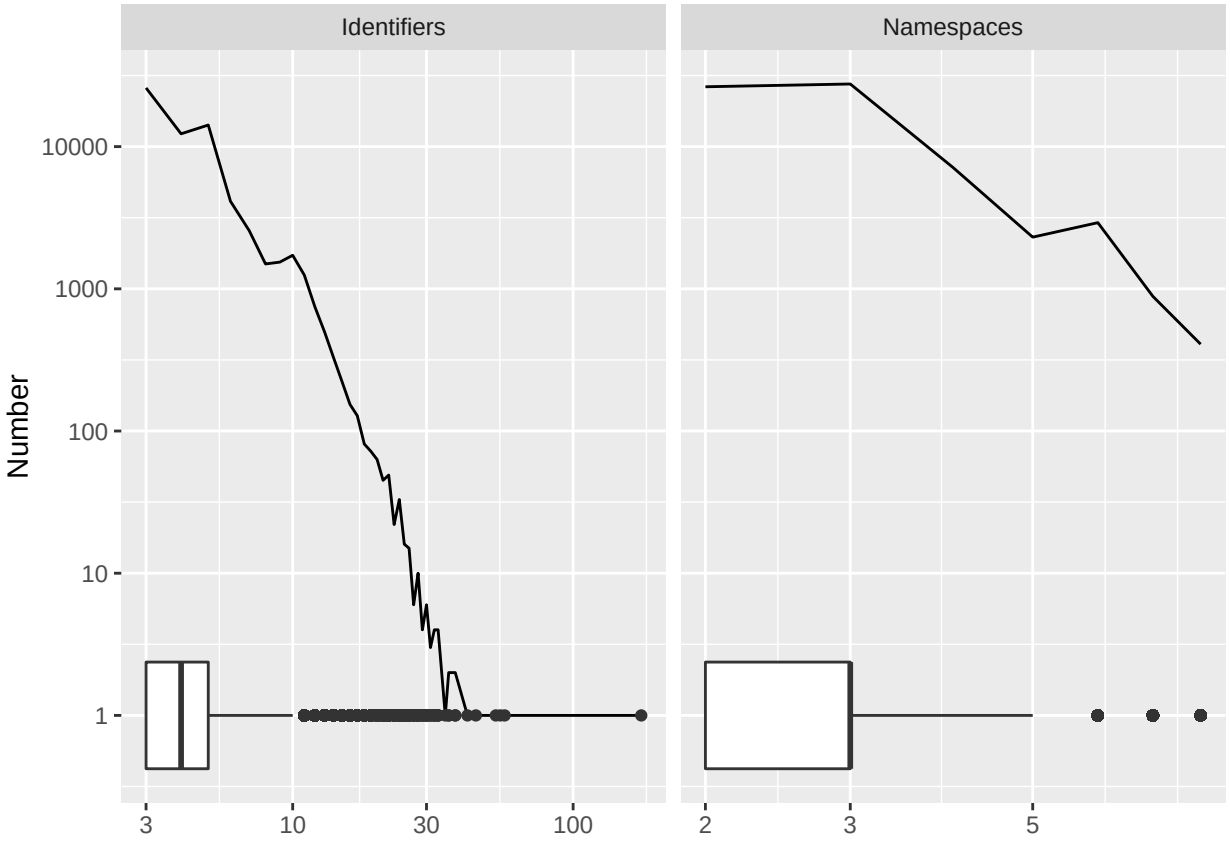

Figure 2: The number of identifiers per reaction and the number of unique source namespaces per reaction. Only reactions that have more than two identifiers are included.

Names

Table 3: Overall number of names and of unique source namespaces.

Names	Unique Namespaces
370,468	6

Table 4: Number of names per source namespace. Names are deduplicated compared to raw tables.

Namespace	Frequency
	1
bigg.reaction	78,374
kegg.reaction	31,190
metacyc.reaction	34,904
rhea	50,042
sabiork.reaction	16,236
seed.reaction	159,721

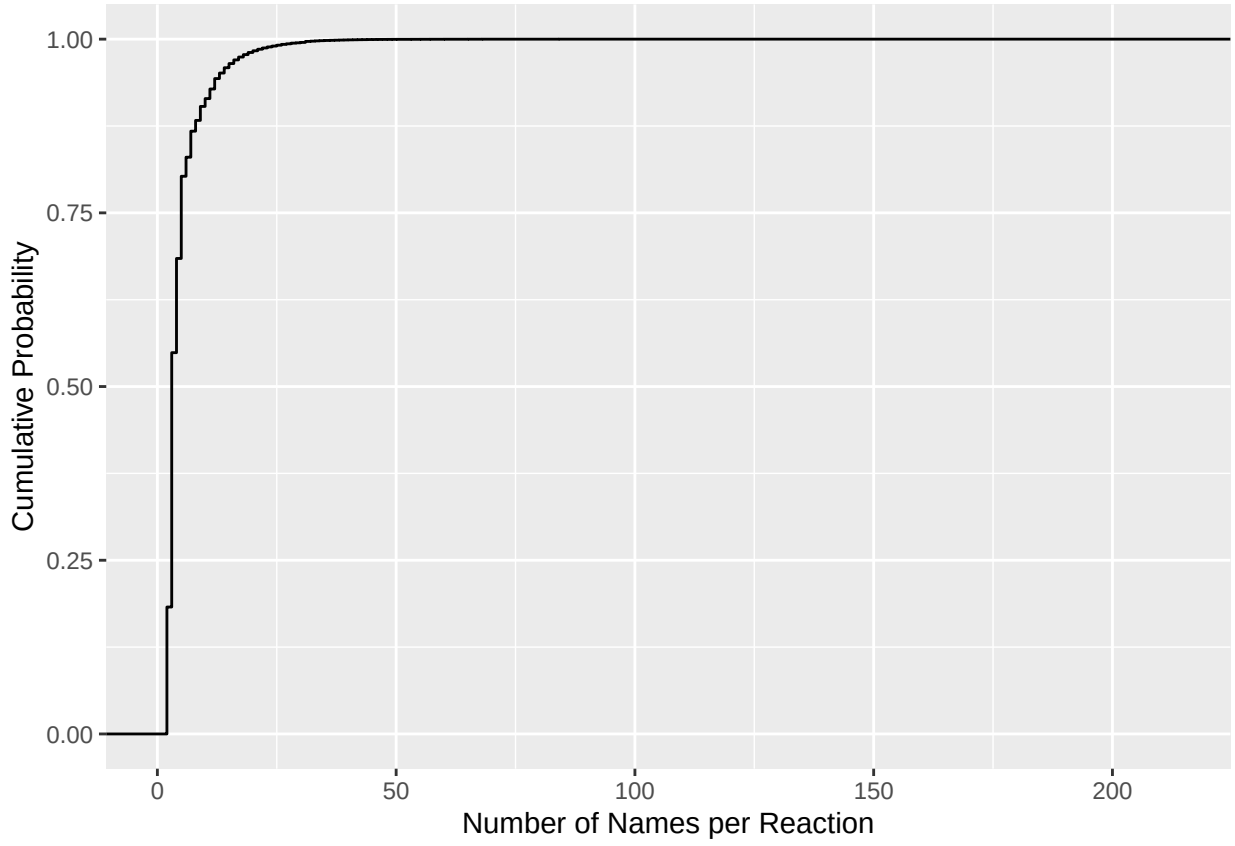

Figure 3: The empirical cumulative distribution function (eCDF) of the number of distinct names per reaction.

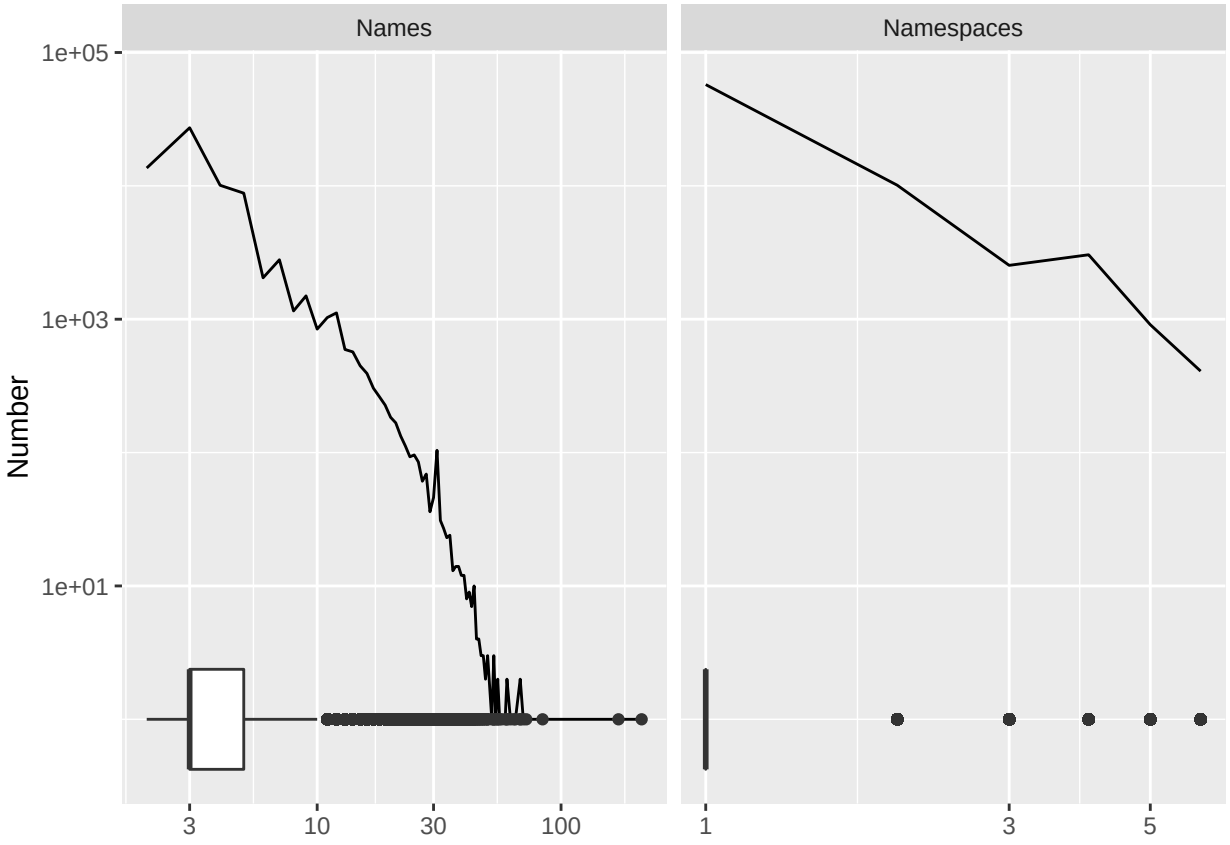

Figure 4: The number of names per reaction and the number of unique source namespaces per reaction. Only reactions that have more than one name are included.