

Module output: report

March 5, 2018

Module Analysis

Data set and modules summary

- 82 data sets
- 0 datasets without any module fraction match
- 8 modules analysed initially

Module fraction presence: heatmap and variance

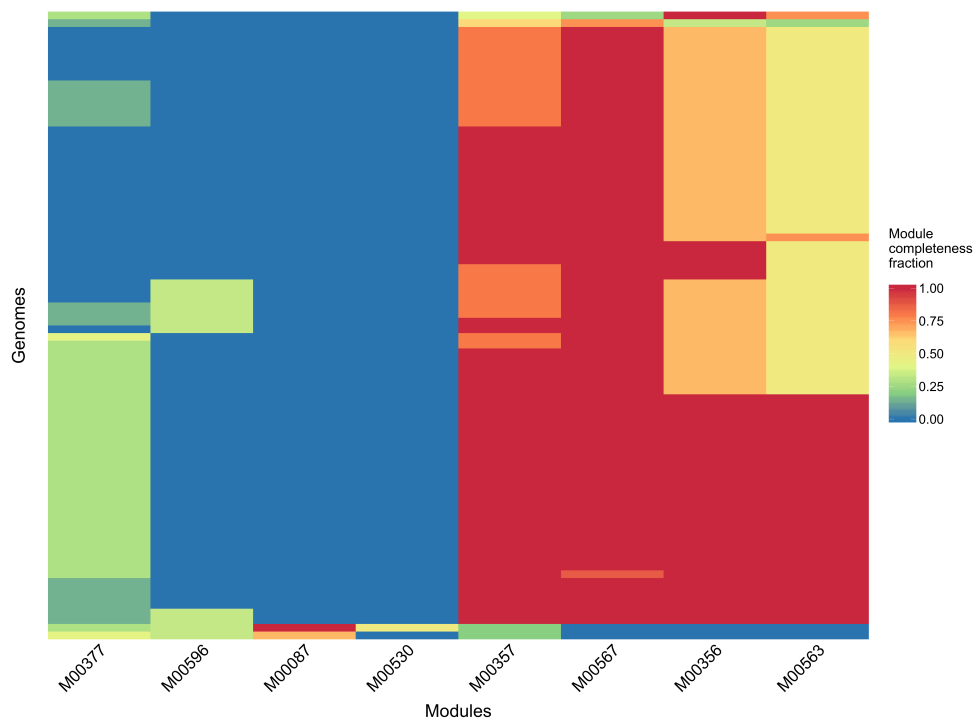


Figure 1: Module fraction presence across data sets

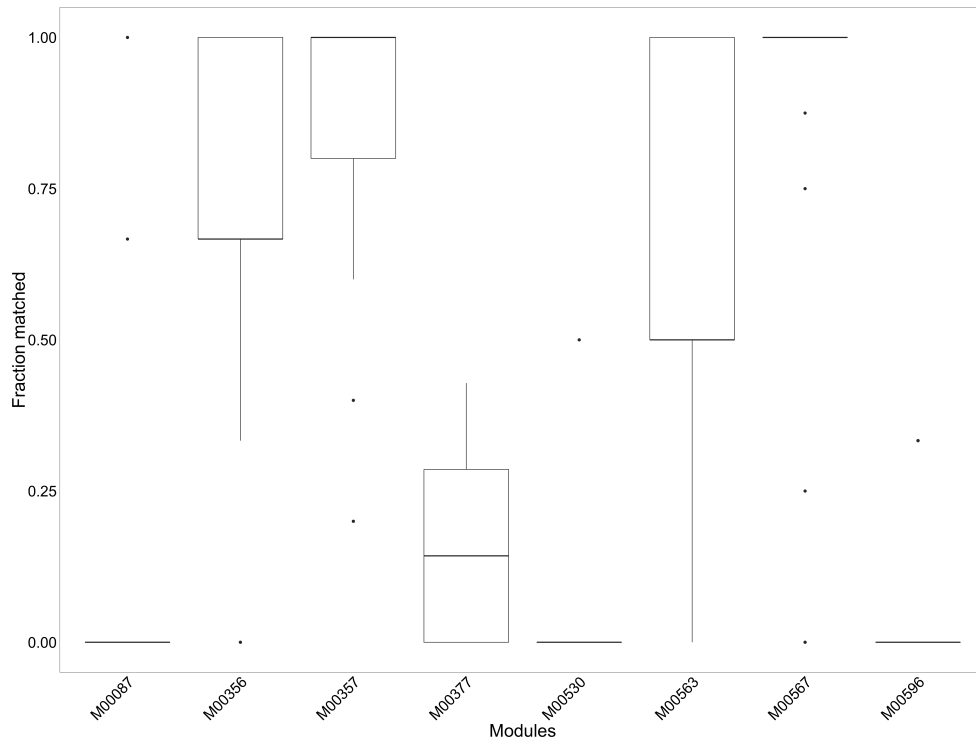


Figure 2: Module fraction presence variance - boxplot

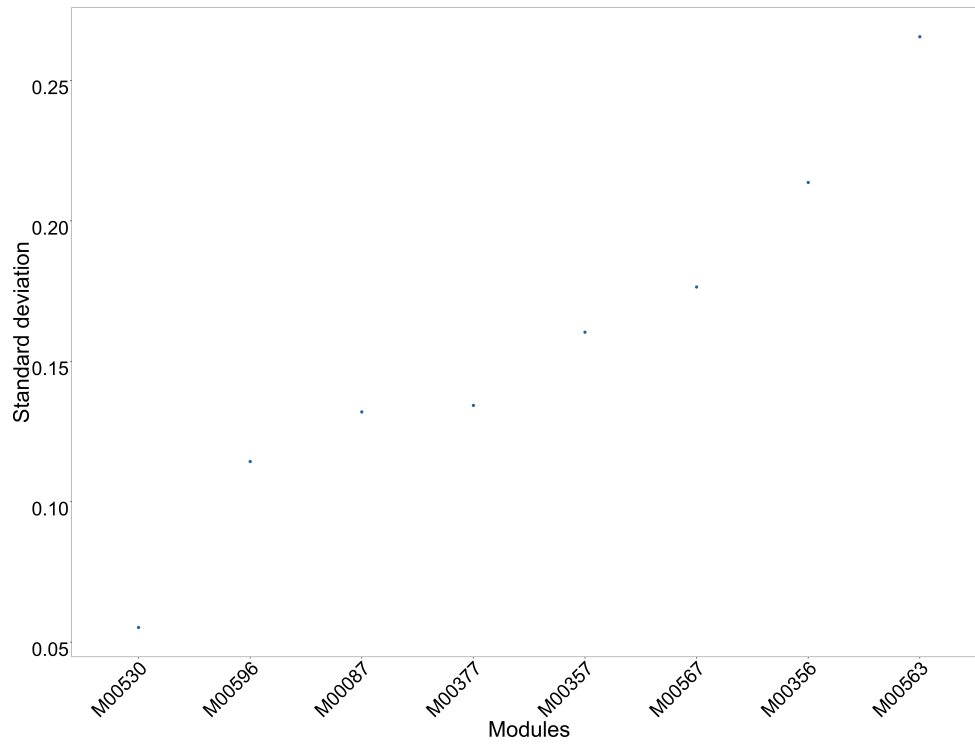


Figure 3: Module fraction presence variance - scatter plot

Module fraction presence: mean and variance by grouping factors

No modules with zero variance.

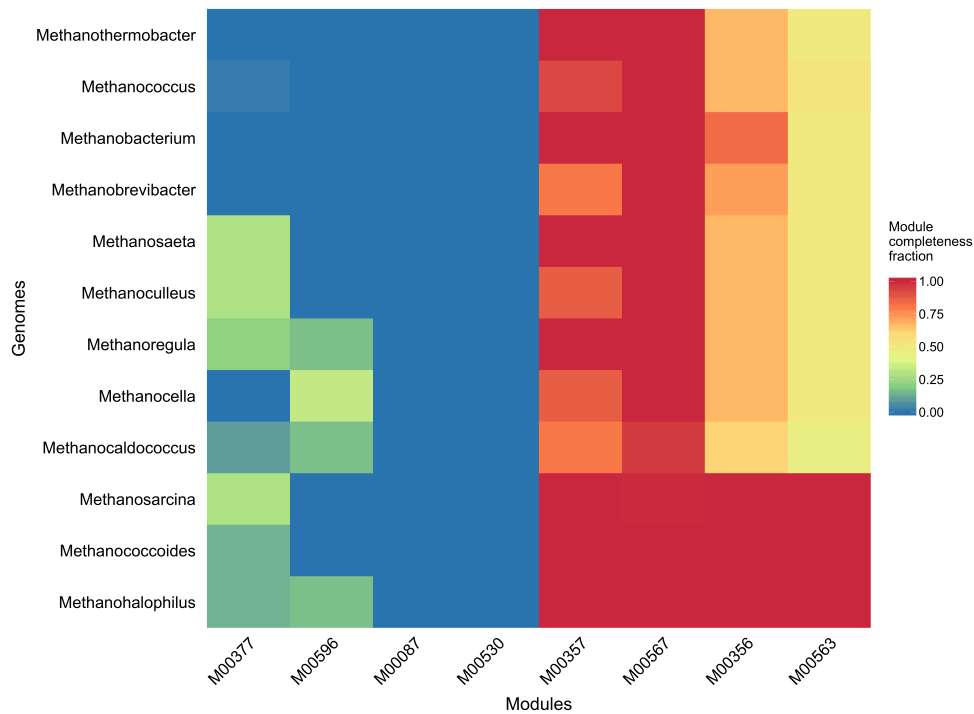


Figure 4: genusFactor: mcf - mean

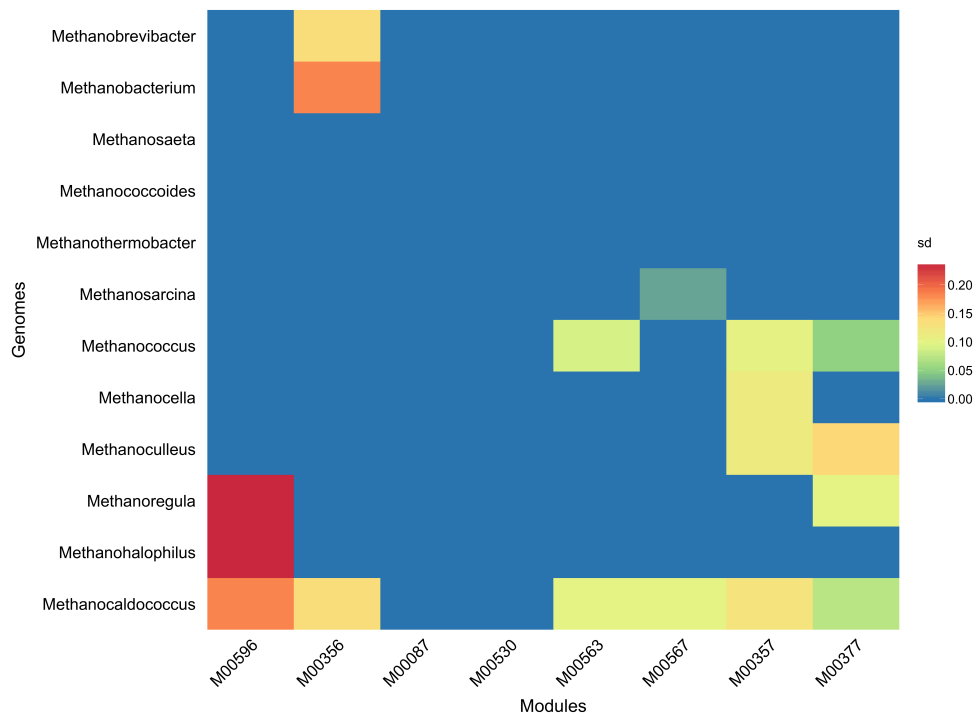


Figure 5: genusFactor: mcf - standard deviation

PCA

PCA was done with the module analysis output.

Grouping “factors” were used to look at the data. The mean Euclidean distance (m) across the groups generated by factoring was calculated according to the following equation:

$$m = \frac{\sum_i^N d_i}{N} \quad (1)$$

where N is the total number of distances between p points, given by

$$N = \frac{p * (p - 1)}{2} \quad (2)$$

and are contained in the (module_mean_dist_output_FACTOR.rda object)

Factors	n.Groups
genusFactor	27.00

Table 1: Number of groups by factors

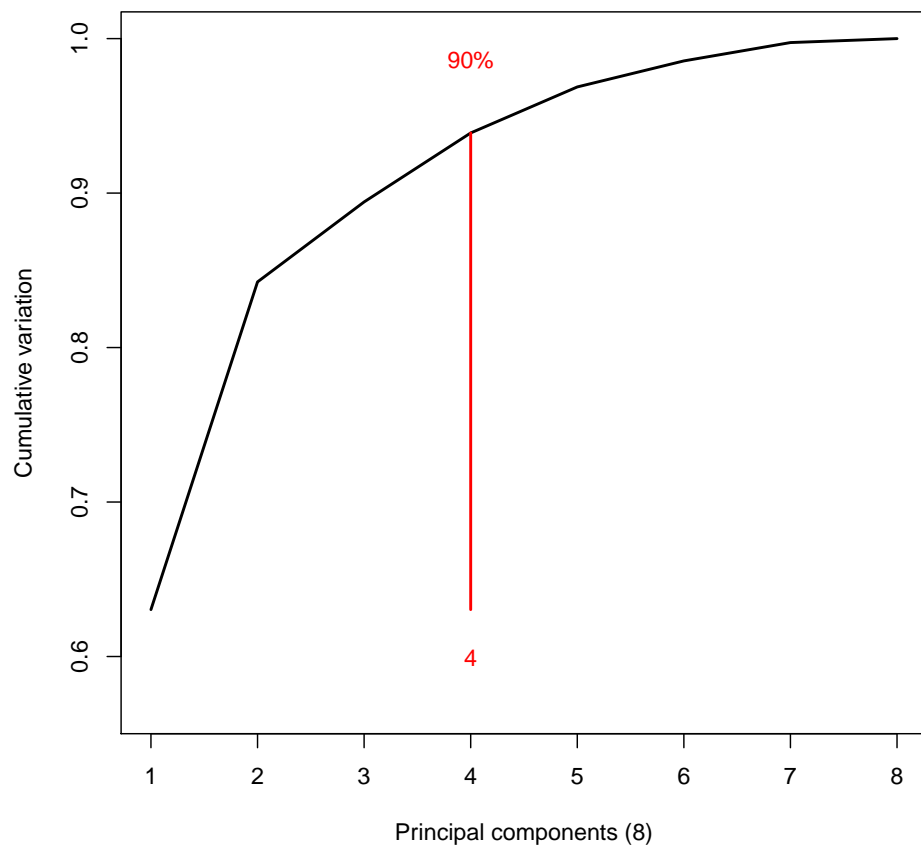


Figure 6: Cumulative variance captured by principal components

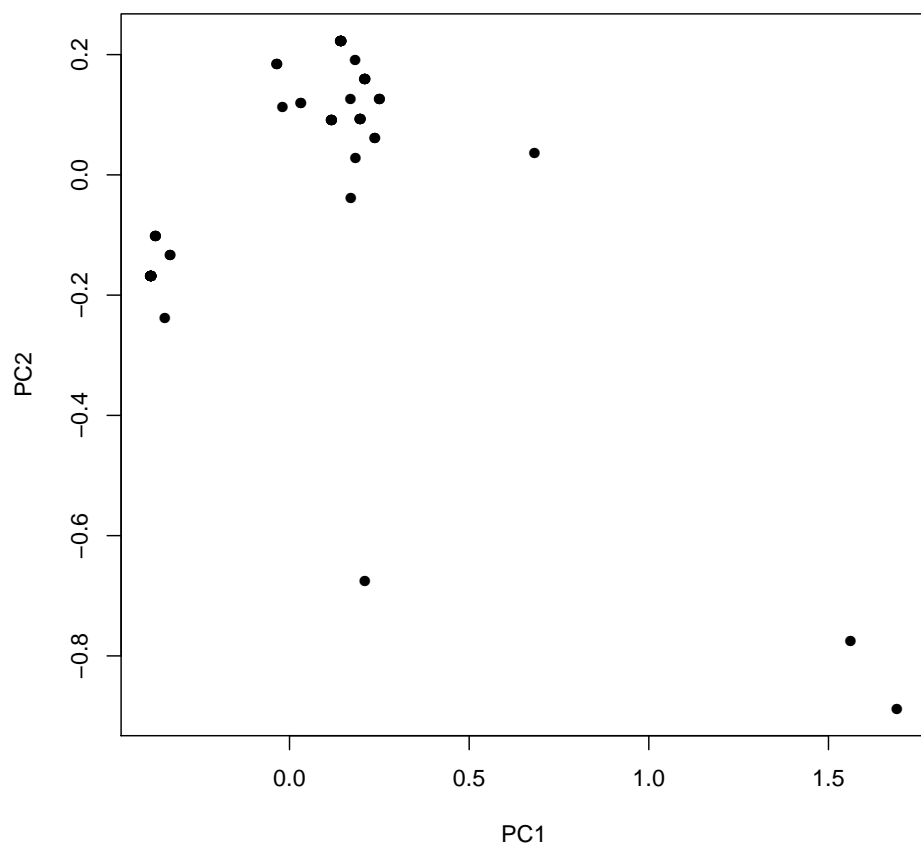


Figure 7: Princial component scatter plot

genusFactor

Groups	group_counts
Amycolatopsis	1.00
Bacillus	1.00
Methanobacterium	6.00
Methanobrevibacter	6.00
Methanocaldococcus	6.00
Methanocella	3.00
Methanococcoides	2.00
Methanococcus	8.00
Methanocorpusculum	1.00
Methanoculleus	3.00
Methanogenic	1.00
Methanohalobium	1.00
Methanohalophilus	2.00
Methanolacinia	1.00
Methanolobus	1.00
Methanomethylovorans	1.00
Methanopyrus	1.00
Methanoregula	2.00
Methanosaeta	3.00
Methanosalsum	1.00
Methanosarcina	22.00
Methanosphaera	1.00
Methanospirillum	1.00
Methanothermobacter	4.00
Methanothermococcus	1.00
Methanothermus	1.00
Methanotorris	1.00

Table 2: genusFactor

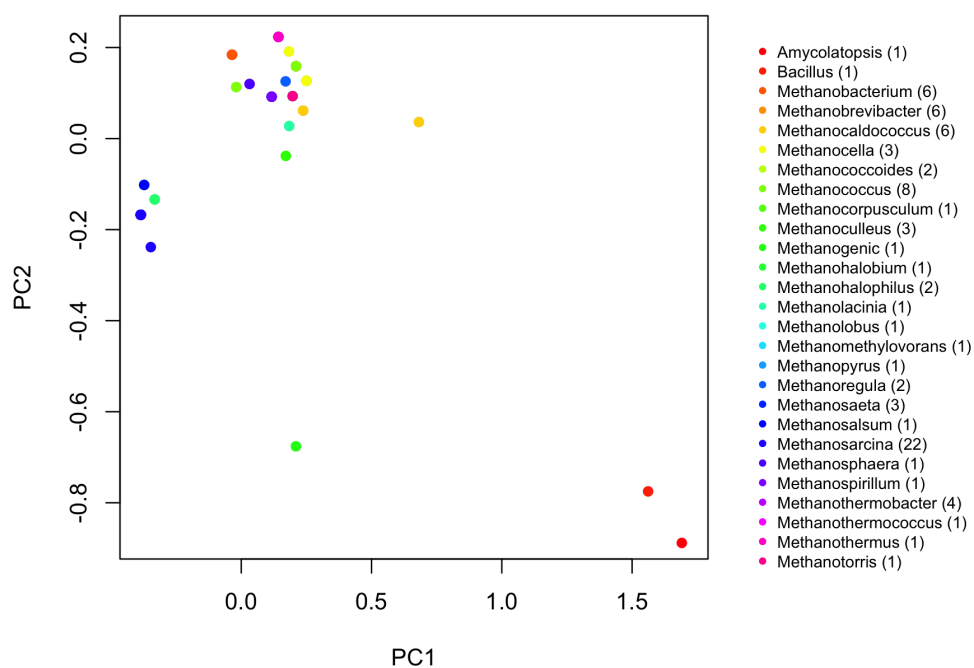


Figure 8: genusFactor: PC plot

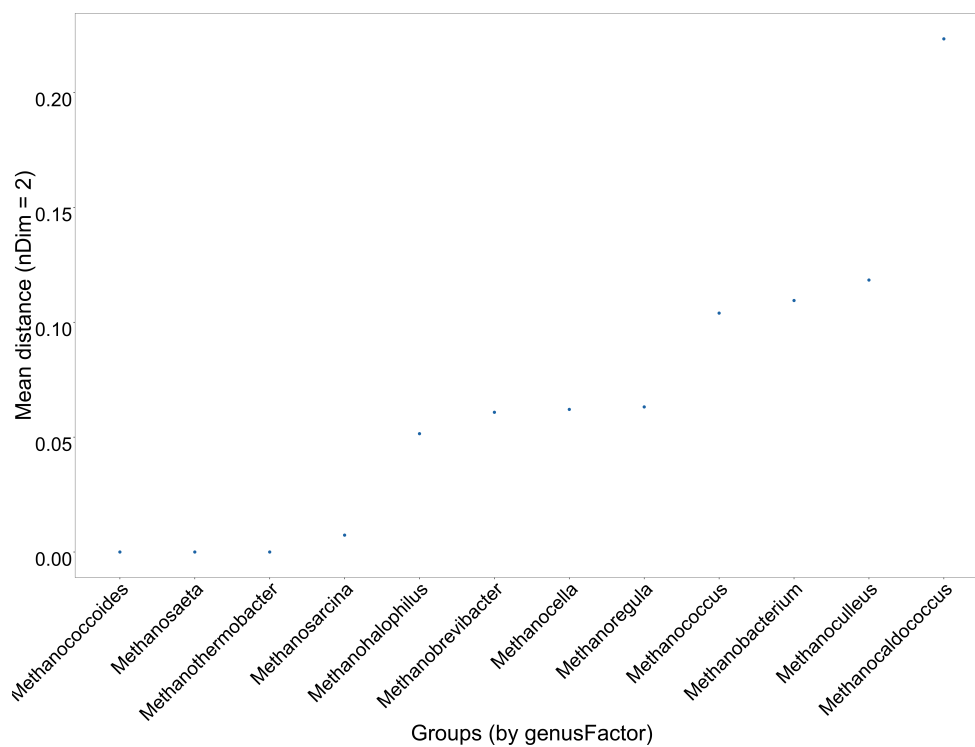


Figure 9: genusFactor: mean Euclidean distance