

Antibiotic resistance genes that occur in Biosynthetic gene clusters

DATA

CARD database contains antibiotic resistance (AR) genes.
MIBiG database is composed by Biosynthetic gene clusters.
antiSMASHDB contains predicted BGCs.

Methodology

CARD gene families

Genes in CARD were classified into gene families using FastORTHO with default parameters.

Homology searches between CARD and MIBiG

Homology search was conducted using blast (e-value) in an all vs all comparison with query antibiotic resistance genes from The Comprehensive Antibiotic Resistance Database (CARD) against genes in BGCs from MIBiG database.

In the following sections we will refer to the following questions: How many BGCs contain an AR gene? How many families of AR genes are present in BGCs? Which families are over represented? How many BGCs per AR gene/Family?

Protein families in CARD

The following figure shows the most populated families in CARD database.

The total number of antibiotic gene resistance families is 652 counting singletons. Without singletons there are 311 families.

These first five families, in a preliminar search corresponds to:

Family	Number of genes	Annotation
Family 0	209	class A broad spectrum β -lactamase
Family 1	166	class A extended-spectrum β -lactamase
Family 2	159	cephalosporin-hydrolyzing class C β -lactamase
Family 3	138	class A extended-spectrum β -lactamase
Family 4	112	class D β -lactamase

BGC classes in MiBIG

The following figure shows the most populated classes in MiBIG database.

There are 1811 different BGCs reported at MIBiG distributed in several classes as indicated in the following

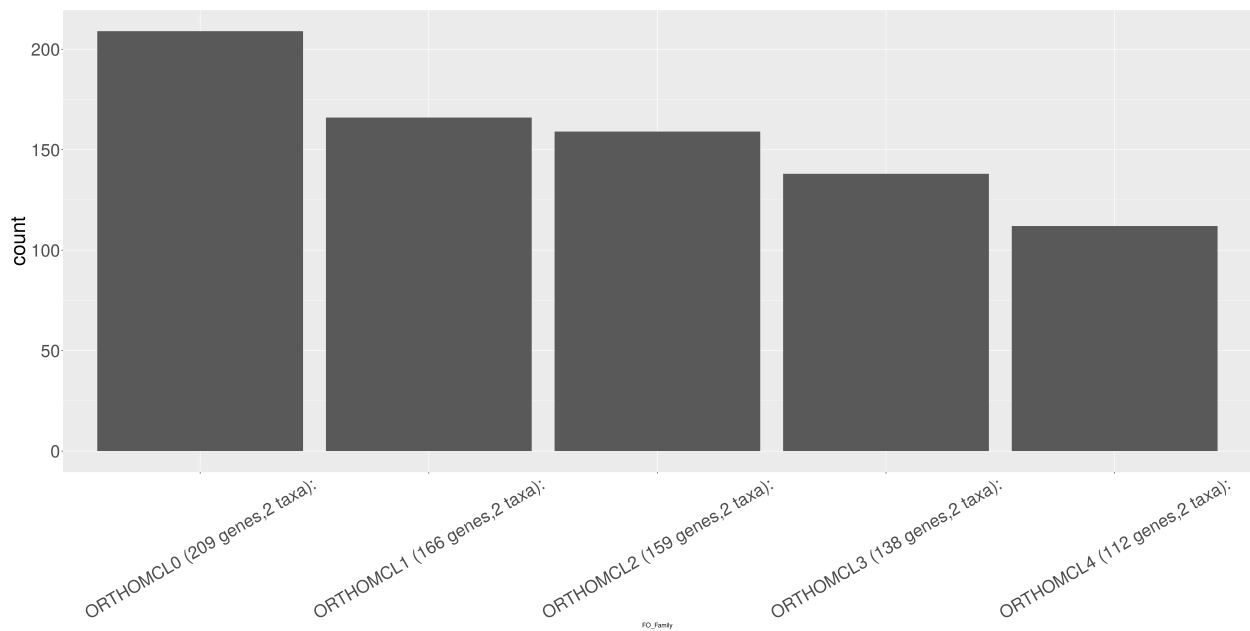


Figure 1: CARD families

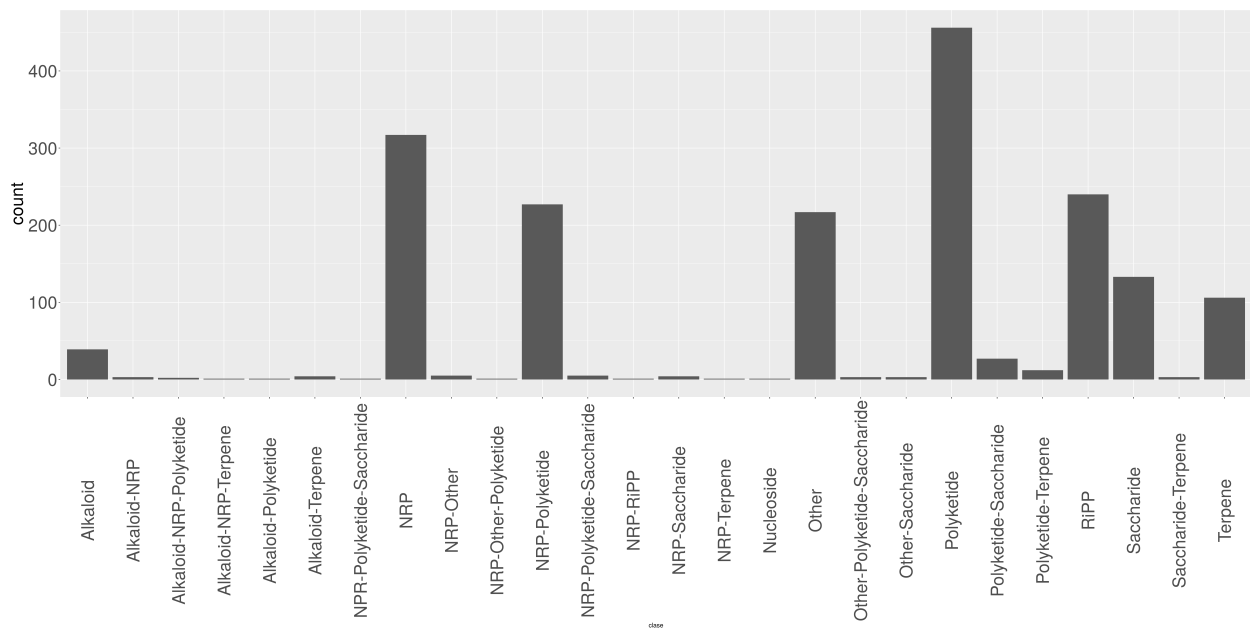


Figure 2: MIBiG classes

table.

Elements	MiBiG class
39	Alkaloid
3	Alkaloid-NRP
2	Alkaloid-NRP-Polyketide
1	Alkaloid-NRP-Terpene
1	Alkaloid-Polyketide
4	Alkaloid-Terpene
1	NRP-Polyketide-Saccharide
317	NRP
5	NRP-Other
1	NRP-Other-Polyketide
227	NRP-Polyketide
5	NRP-Polyketide-Saccharide
1	NRP-RiPP
4	NRP-Saccharide
1	NRP-Terpene
1	Nucleoside
217	Other
3	Other-Polyketide-Saccharide
3	Other-Saccharide
456	Polyketide
27	Polyketide-Saccharide
12	Polyketide-Terpene
240	RiPP
133	Saccharide
3	Saccharide-Terpene
106	Terpene

CARD families and their interaction with BGCs in MiBiG

The following figure shows the average of CARD families by MiBiG classes

From 1811 different BGCs at MiBiG 1240 contains a CARD hit

The following figure shows which CARD families are the most populated classes in MiBiG database. and now an histogram by class with the number of BGCs with different frequencies of CARD by color

which CARD families are over represented in MiBiG which MiBiG classes are over represented as CARD hits

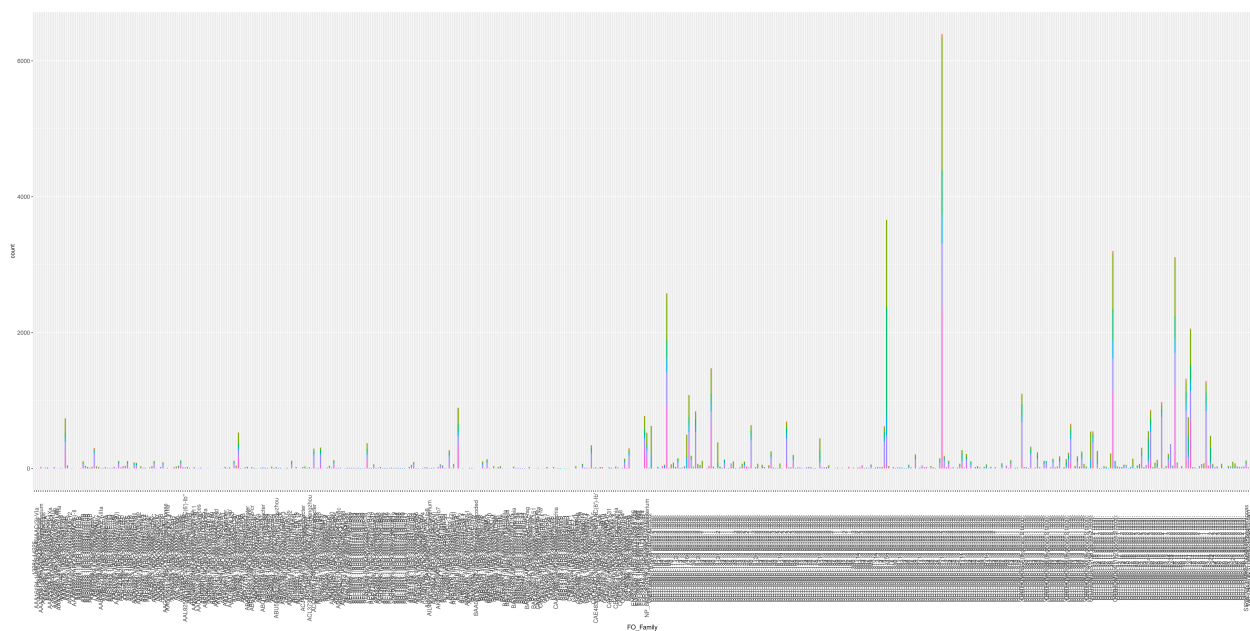


Figure 3: CARD families in MIBiG classes

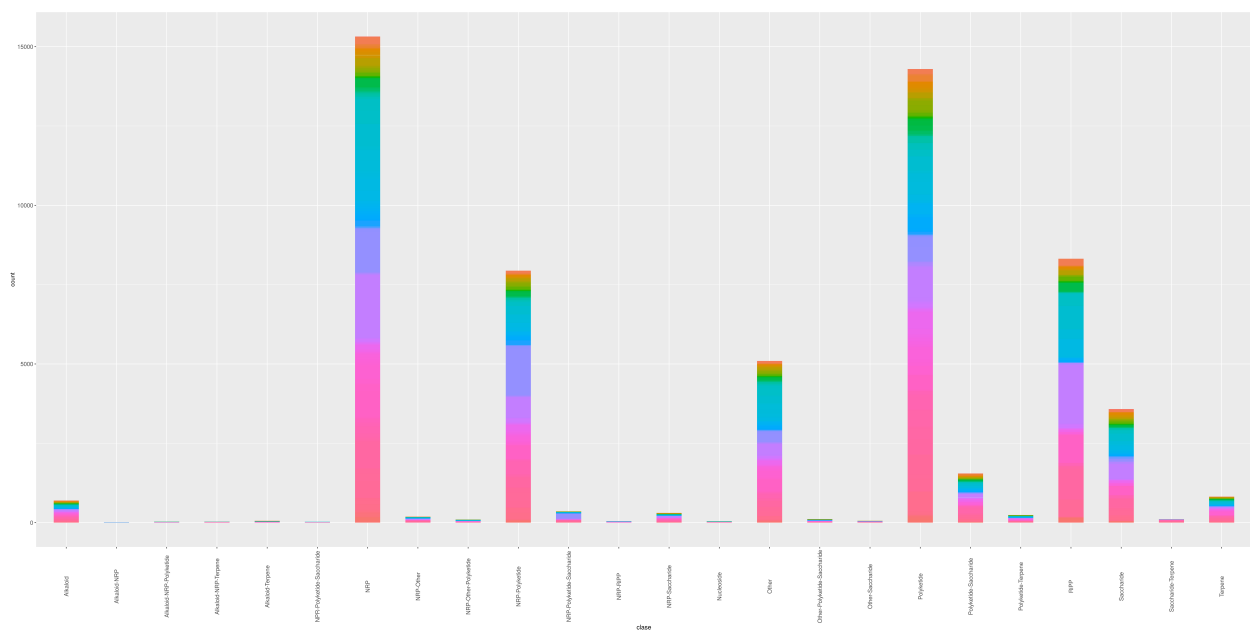


Figure 4: CARD families in MIBiG classes

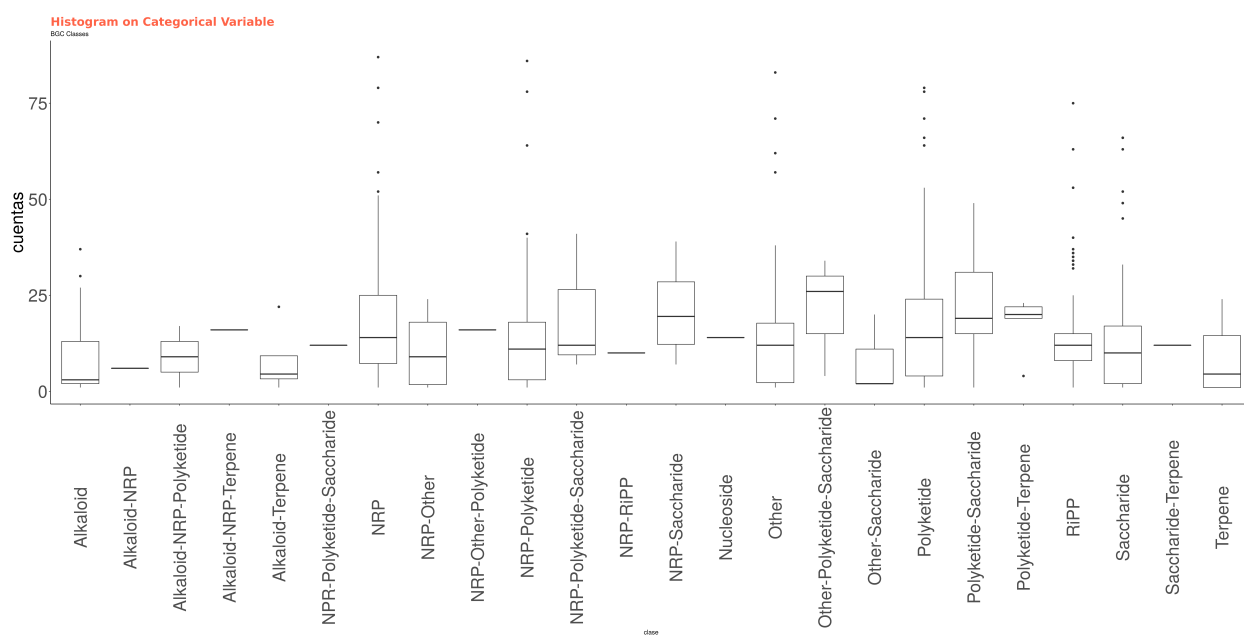


Figure 5: CARD families in MIBiG classes