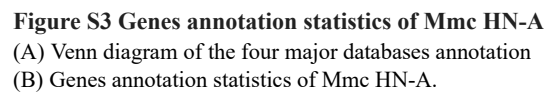


A

Venn diagram illustrating the overlap of protein families identified by four databases: COG (blue), KEGG (orange), NR (green), and Swiss-Prot (red). The counts for each region are as follows:

Region	Count
COG only	0
KEGG only	0
NR only	252
Swiss-Prot only	0
COG & KEGG	0
COG & NR	15
COG & Swiss-Prot	40
KEGG & NR	52
KEGG & Swiss-Prot	10
NR & Swiss-Prot	27
COG & KEGG & NR	29
COG & KEGG & Swiss-Prot	0
COG & NR & Swiss-Prot	0
KEGG & NR & Swiss-Prot	0
All four databases	418



Database	Number of annotated genes
GO	731
KEGG	509
COG	502
NR	843
Secretory_Protein	57
Swiss-Prot	495
PHI	130
VFDB	15
CARD	38
T3SS	1
CAZy	92

Database	Number of annotated genes
GO	731
KEGG	509
COG	502
NR	843
Secretory_Protein	57
Swiss-Prot	495
PHI	130
VFDB	15
CARD	38
T3SS	1
CAZy	92