

Supplementary Material

Global genome mining reveals the distribution of diverse thioamidated RiPP biosynthetic gene clusters

Jessie James Limlingan Malit¹, Chuanhai Wu¹, Ling-Li Liu^{1, 2*}, Pei-Yuan Qian^{1,*}

¹ Department of Ocean Science, and Hong Kong Branch of Southern Marine Science and Engineering Guangdong Laboratory, The Hong Kong University of Science and Technology, Hong Kong, China

² Shaanxi Key Laboratory of Natural Products & Chemical Biology, College of Chemistry & Pharmacy, Northwest A&F University, 22 Xinong Road, Yangling 712100, Shaanxi, People' s Republic of China

*Correspondence: boqianpy@ust.hk; leonie@nwsuaf.edu.cn; Tel.: +852-23587331

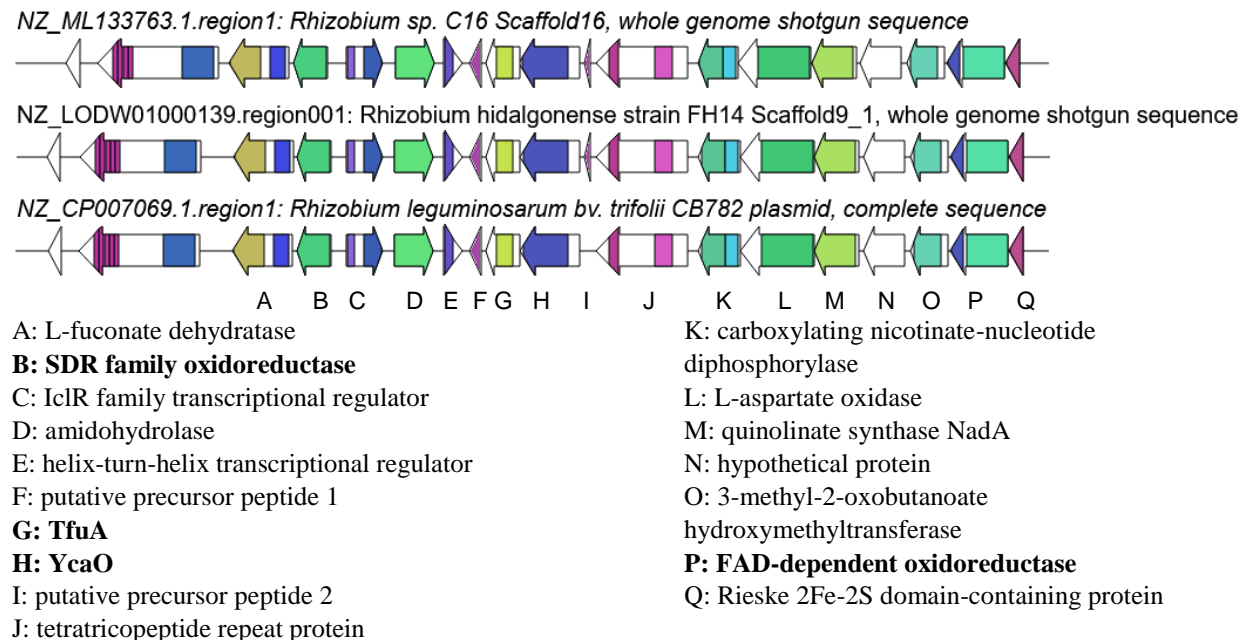
Index

Figure S1. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Proteobacteria <i>Rhizobium sp.</i> Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.....	5
Figure S2. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Proteobacteria, <i>Rhizobium sp.</i> Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.....	6
Figure S3. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Proteobacteria, <i>Agrobacterium sp.</i> Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.....	7
Figure S4. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Proteobacteria, <i>Rhizobium sp.</i> Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.....	8
Figure S5. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Proteobacteria, <i>Coralloccoccus sp.</i> Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.....	9
Figure S6. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Euryarchaeota, <i>Methanosarcina sp.</i> Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.....	10
Figure S7. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Euryarchaeota, <i>Methanobrevibacter sp.</i> Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.....	11
Figure S8. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Euryarchaeota, <i>Methanothermobacter sp.</i> Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.....	12
Figure S9. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Euryarchaeota, <i>Methanobrevibacter sp.</i> Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.....	13
Figure S10. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Euryarchaeota, <i>Methanoculleus sp.</i> Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.....	14

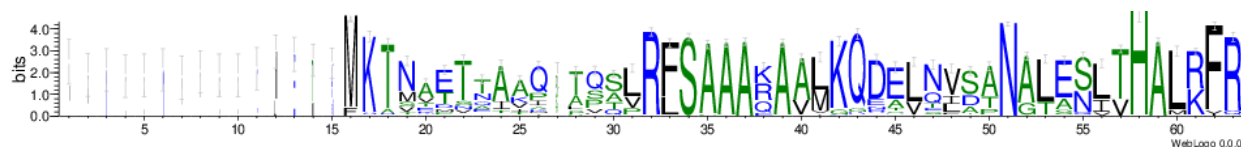
Figure S11. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Thaumarchaeota <i>Nitrosocosmicus</i> sp.. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.	15
Figure S12. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Cyanobacteria. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.	16
Figure S13. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Cyanobacteria <i>Nostoc</i> sp.. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.	17
Figure S14. Sequence similarity network of TfuA protein sequences generated using EFI-EST.	18
Figure S15. Sequence similarity network of precursor peptides identified through RiPPER within TfuA-related BGCs.	19
Figure S16. TfuA-related GCFs containing an CCRG-2 family-like precursor peptide as predicted by RiPPER. GCF from <i>Nostoc</i> sp. share the same predicted precursor peptide network. Sequence logo of predicted precursor peptide using RiPPER is indicated.	20
Figure S16 (cont). TfuA-related GCFs containing an CCRG-2 family-like precursor peptide as predicted by RiPPER. GCFs from <i>Nostoc</i> share the same predicted precursor peptide network. Sequence logo of predicted precursor peptide using RiPPER is indicated.	21
Figure S17. TfuA-related GCFs containing an albusnodin-like precursor peptide as predicted by RiPPER. Sequence logo of predicted precursor peptide using RiPPER is indicated.	22
Figure S18. TfuA-related GCFs containing berninamycin-like precursor peptide as predicted by RiPPER. GCFs from <i>Rhizobium</i> share the same predicted precursor peptide network. Sequence logo of predicted precursor peptide using RiPPER is indicated.	23
Figure S19. TfuA-related GCF containing with the precursor peptide annotated as a bacteriocin containing thiopeptide-type modifications as predicted by RiPPER. Sequence logo of predicted precursor peptide using RiPPER is indicated.	24
Figure S20. TfuA-related GCFs containing Nif-like precursor peptides as predicted by RiPPER.	25
Figure S21. Frequencies of additional modification genes (A) and other genes (B) as annotated by antiSMASH found within TfuA-related BGCs.	26
Figure S22. TfuA-related GCF containing an iron-containing redox enzyme family protein. Sequence logo of predicted precursor peptide using RiPPER is indicated.	27
Figure S23. TfuA-related GCF containing a fused <i>tfuA</i> and <i>ycalO</i> genes. Sequence logo of predicted precursor peptide using RiPPER is indicated.	28
Figure S24. TfuA-related GCF annotated to also produce ranthipeptides as annotated by antiSMASH. Both GCFs share the same precursor peptide network. Sequence logo of predicted precursor peptide using RiPPER is indicated.	29
Figure S25. TfuA-related GCF annotated to also produce linear azole-containing peptides as annotated by antiSMASH. Sequence logo of predicted precursor peptide using RiPPER is indicated.	30

Figure S26. TfuA-related GCFs annotated to also produce thiopeptides as annotated by antiSMASH. Sequence logo of predicted precursor peptide using RiPPER is indicated.	31
Figure S26 (cont.). TfuA-related GCFs annotated to also produce thiopeptides as annotated by antiSMASH. Sequence logo of predicted precursor peptide using RiPPER is indicated.	32
Figure S27. TfuA-related GCFs containing multiple <i>ycaO</i> genes without further annotation. Sequence logo of predicted precursor peptide using RiPPER is indicated.	33
Figure S27 (cont.). TfuA-related GCFs containing multiple <i>ycaO</i> genes without further annotation. Sequence logo of predicted precursor peptide using RiPPER is indicated.	34
Figure S28. Sequence logo of the ATP-binding sequence motifs identified from TfuA-associated YcaO proteins.....	35

Figure S1. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Proteobacteria *Rhizobium* sp. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.



Putative precursor peptide 1 (Cluster 5)



Putative precursor peptide 2 (Cluster 6)

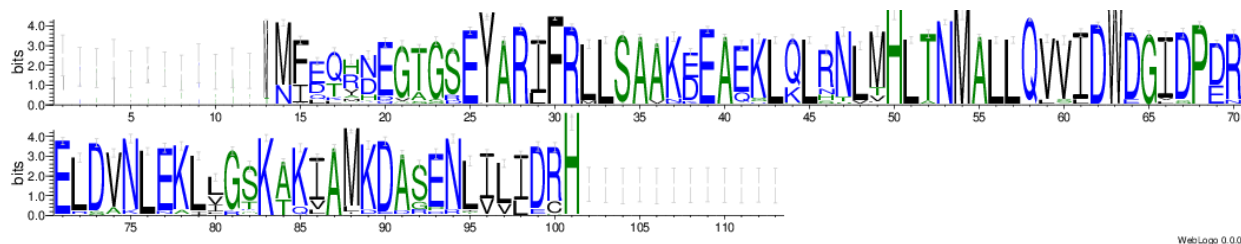
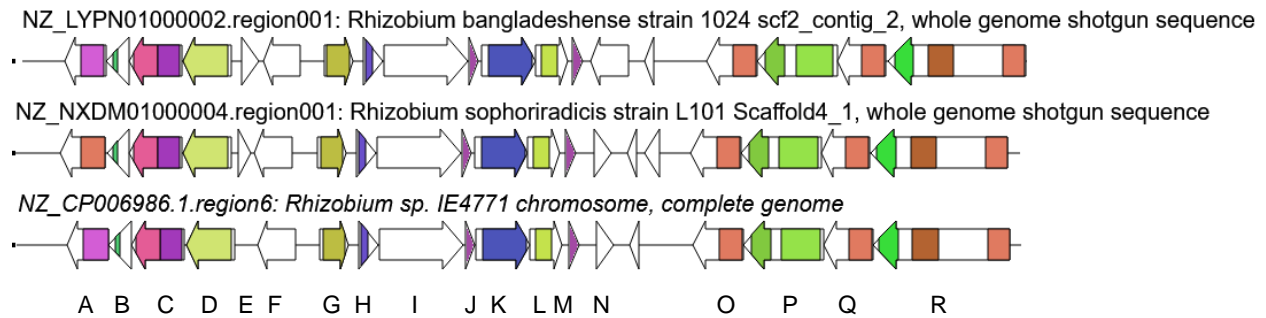


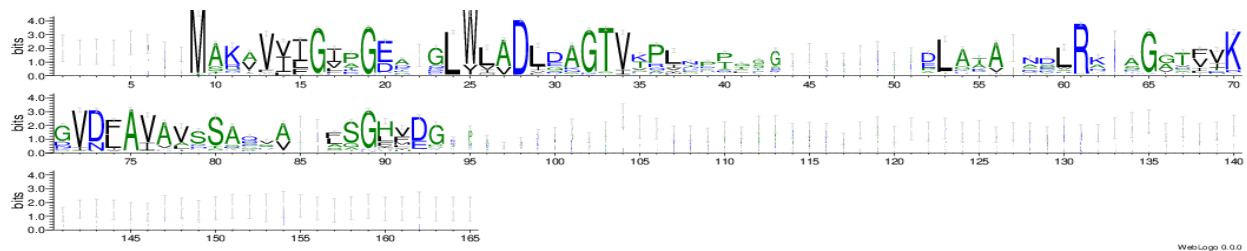
Figure S2. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Proteobacteria, *Rhizobium* sp. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.



A: polysaccharide pyruvyl transferase family protein
 B: acyltransferase
C: glycosyltransferase
 D: endo-1,4-beta-xylanase
 E: hypothetical protein
 F: exopolysaccharide biosynthesis protein
G: class I SAM-dependent methyltransferase
 H: helix-turn-helix domain-containing protein
 I: peptide antibiotic resistance protein

J: putative precursor peptide 1
K: YcaO
L: TfuA
 M: putative precursor peptide 2
 N: hypothetical protein
O: glycosyltransferase family 2 protein
P: ABC transporter ATP-binding protein
Q: glycosyltransferase family 2 protein
R: glycosyltransferase

Putative precursor peptide 1 (Cluster 1)



Putative precursor peptide 2 (Cluster 4)

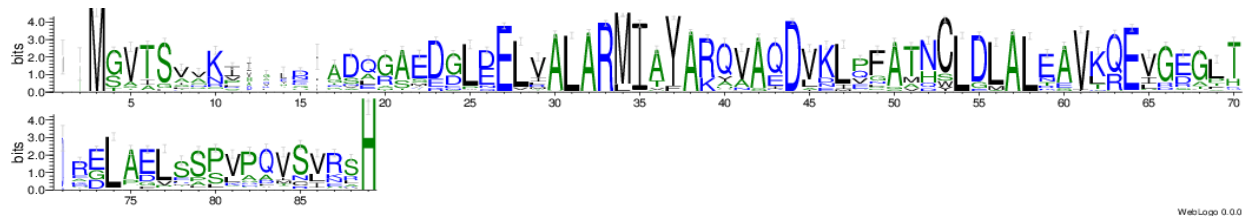
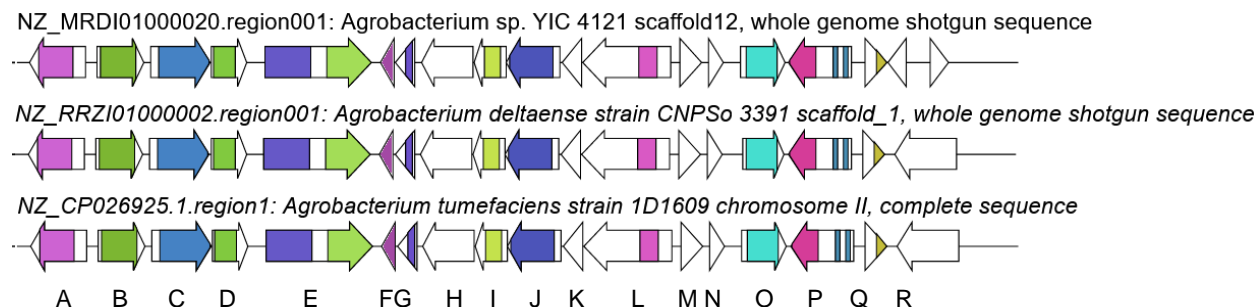


Figure S3. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Proteobacteria, *Agrobacterium* sp. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.



A: ROK family protein

B: D-xylose ABC transporter substrate-binding protein

C: sugar ABC transporter permease

D: sugar ABC transporter ATP-binding protein

E: 5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase

F: putative precursor peptide

G: helix-turn-helix transcriptional regulator

H: hypothetical protein

I: TfuA

J: YcaO

K: hypothetical protein

L: SARP family transcriptional regulator

M: hypothetical protein

N: hypothetical protein

O: TIM barrel protein

P: family 16 glycosylhydrolase

Q: transcription elongation factor GreA

R: SGNH/GDSL hydrolase family protein

Putative precursor peptide 2 (Cluster 3)

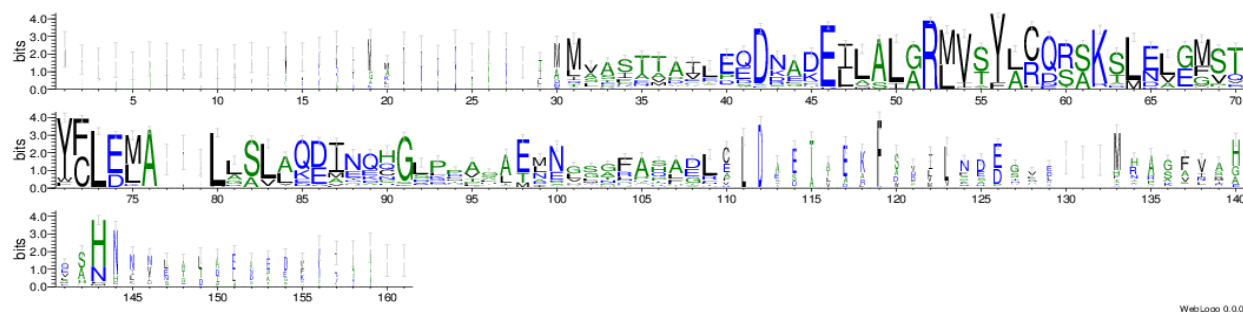
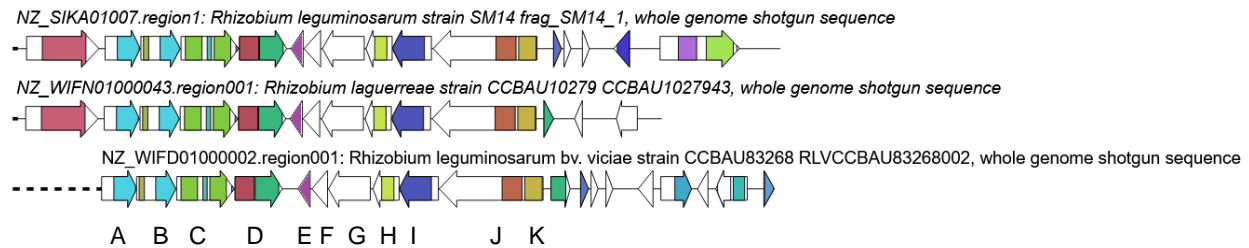


Figure S4. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Proteobacteria, *Rhizobium* sp. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.



A: ABC transporter ATP-binding protein
B: ABC transporter permease
C: ABC transporter ATP-binding protein
D: alpha-glucosidase/alpha-galactosidase
E: precursor peptide
F: hypothetical protein

G: hypothetical protein
H: TfuA
I: YcaO
J: AAA family ATPase
K: alpha-glucosidase/alpha-galactosidase

Putative precursor peptide (Cluster 23)

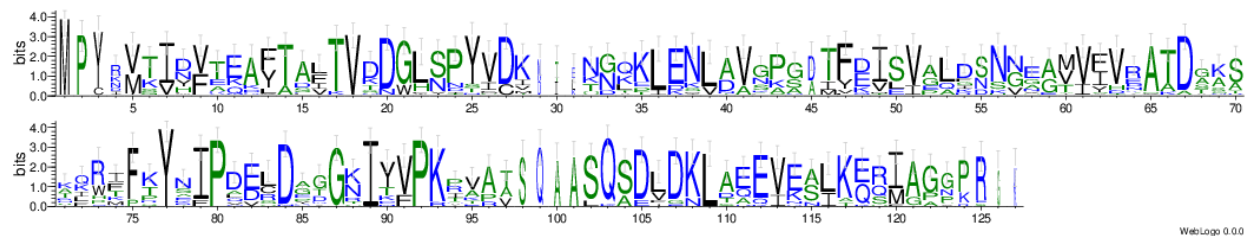
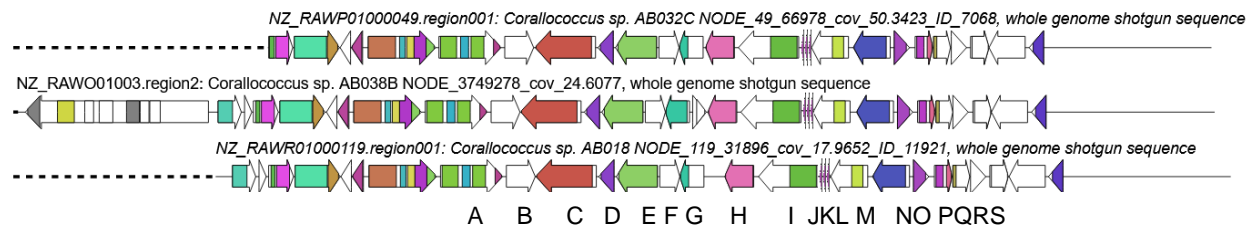


Figure S5. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Proteobacteria, *Corallococcus* sp. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.



A: ABC-F family ATP-binding cassette domain-containing protein

B: flagellar motor protein MotB

C: OPT/YSL family transporter

D: peptidylprolyl isomerase

E: dicarboxylate/amino acid:cation symporter

F: hypothetical protein

G: site-specific DNA-methyltransferase

H: DUF2381 family protein

I: serine/threonine protein kinase

J: putative precursor peptide 1

K: putative precursor peptide 2

L: putative precursor peptide 3

M: TfuA

N: YcaO

O: peroxiredoxin

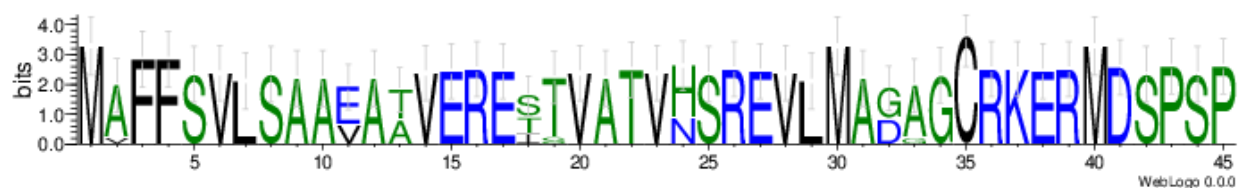
P: sigma-70 family RNA polymerase sigma factor

Q: zf-HC2 domain-containing protein

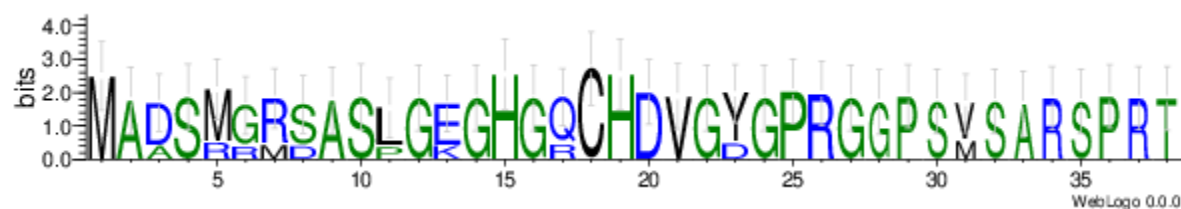
R: hypothetical protein

S: DUF4136 domain-containing protein

Putative precursor peptide 1 (Cluster 100)



Putative precursor peptide 2 (Cluster 218)



Putative precursor peptide 3 (Cluster 612)

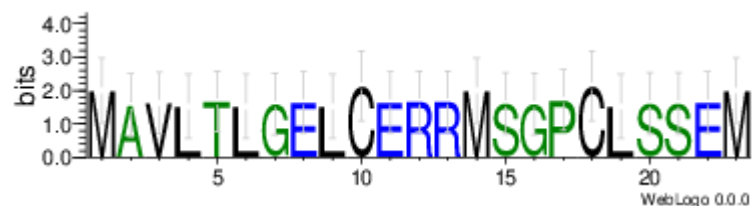
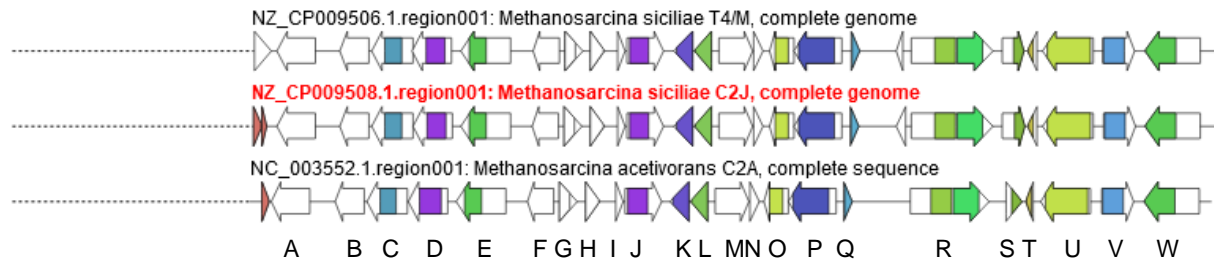


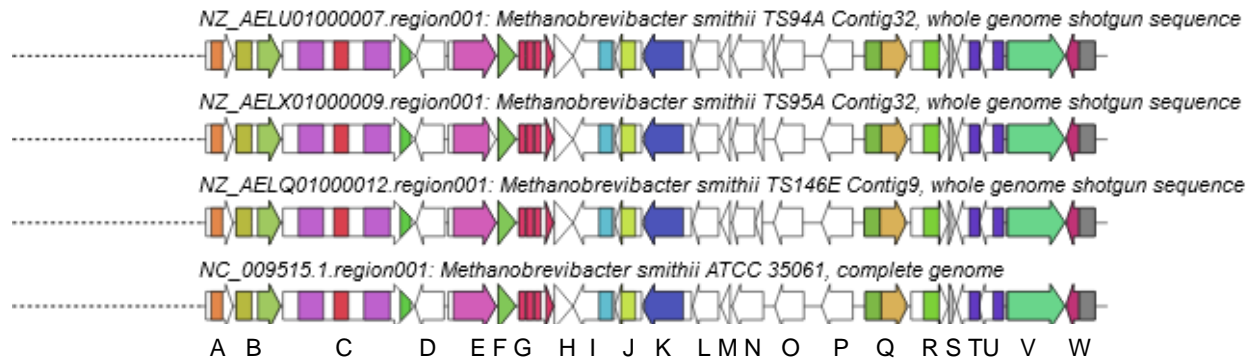
Figure S6. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Euryarchaeota, *Methanosarcina* sp. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.



A: polysaccharide deacetylase family protein
 B: 3-methylornithyl-N6-L-lysine dehydrogenase
 PylD
 C: 3-methylornithine--L-lysine ligase PylC
 D: methylornithine synthase PylB
 E: pyrrolysine--tRNA(Pyl) ligase
 F: hypothetical protein
 G: DUF134 domain-containing protein
 H: DUF5320 domain-containing protein
 I: hypothetical protein
J: radical SAM protein
 K: ferritin

L: universal stress protein
 M: hypothetical protein
 N: hypothetical protein
O: TfuA
P: YcaO
 Q: DUF2892 domain-containing protein
 R: hypothetical protein
 S: copper-translocating P-type ATPase
 T: cytochrome c
 U: PLDc N-terminal domain-containing protein
 V: MFS transporter
W: class I SAM-dependent methyltransferase

Figure S7. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Euryarchaeota, *Methanobrevibacter* sp. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.



A: class I SAM-dependent methyltransferase

B: glutamine-hydrolyzing carbamoyl-phosphate synthase small subunit

C: carbamoyl-phosphate synthase large subunit

D: zinc ribbon domain-containing protein

E: amidohydrolase family protein

F: universal stress protein

G: CBS domain-containing protein

H: hypothetical protein

I: ATP-dependent sacrificial sulfur transferase LarE

J: TfuA

K: YcaO

L: site-2 protease family protein

M: hypothetical protein

N: hypothetical protein

O: hypothetical protein

P: hypothetical protein

Q: aminopeptidase P family protein

R: type II secretion system F family protein

S: class III signal peptide-containing protein

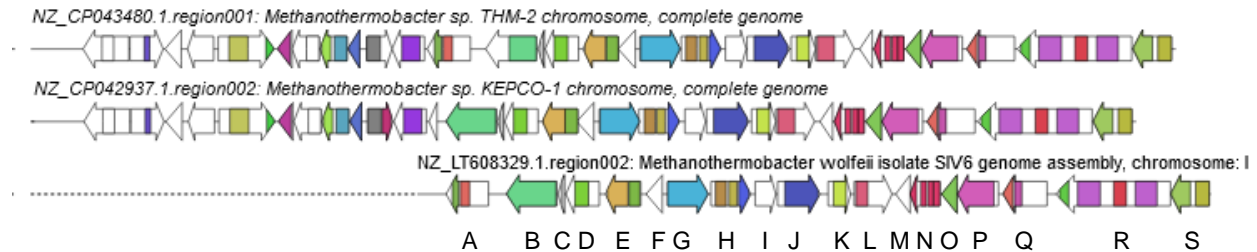
T: DUF1211 domain-containing protein

U: DUF1211 domain-containing protein

V: aldehyde dehydrogenase family protein

W: tRNA-binding protein

Figure S8. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Euryarchaeota, *Methanothermobacter* sp. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.



A: chitinase/beta-hexosaminidase C-terminal domain-containing protein

B: aldehyde dehydrogenase family protein

C: class III signal peptide-containing protein

D: type II secretion system F family protein

E: aminopeptidase P family protein

F: HIT family protein

G: iron-containing alcohol dehydrogenase

H: PAS domain S-box protein

I: site-2 protease family protein

J: YcaO

K: TfuA

L: ATP-dependent sacrificial sulfur transferase LarE

M: hypothetical protein

N: CBS domain-containing protein

O: universal stress protein

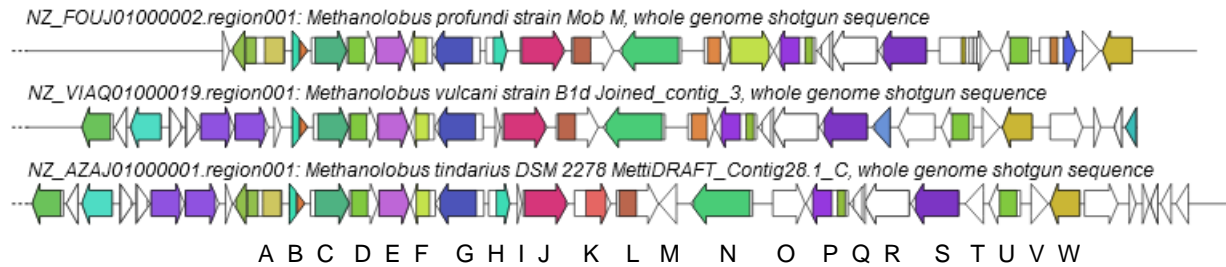
P: amidohydrolase family protein

Q: PKD domain-containing protein

R: carbamoyl-phosphate synthase large subunit

S: glutamine-hydrolyzing carbamoyl-phosphate synthase small subunit

Figure S9. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Euryarchaeota, *Methanobolus* sp. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.



A: MBL fold metallo-hydrolase

B: metal-dependent transcriptional regulator

C: zinc ABC transporter substrate-binding protein

D: ABC transporter ATP-binding protein

E: metal ABC transporter permease

F: TfuA

G: YcaO

H: cytochrome P460 family protein

I: hypothetical protein

J: cysteine desulfurase

K: IS5 family transposase

L: ORC1-type DNA replication protein

M: hypothetical protein

N: thermosome subunit

O: YkgJ family cysteine cluster protein

P: TIGR04013 family B12-binding domain/radical SAM domain-containing protein

Q: DUF126 domain-containing protein

R: DUF521 domain-containing protein

S: UbiD family decarboxylase

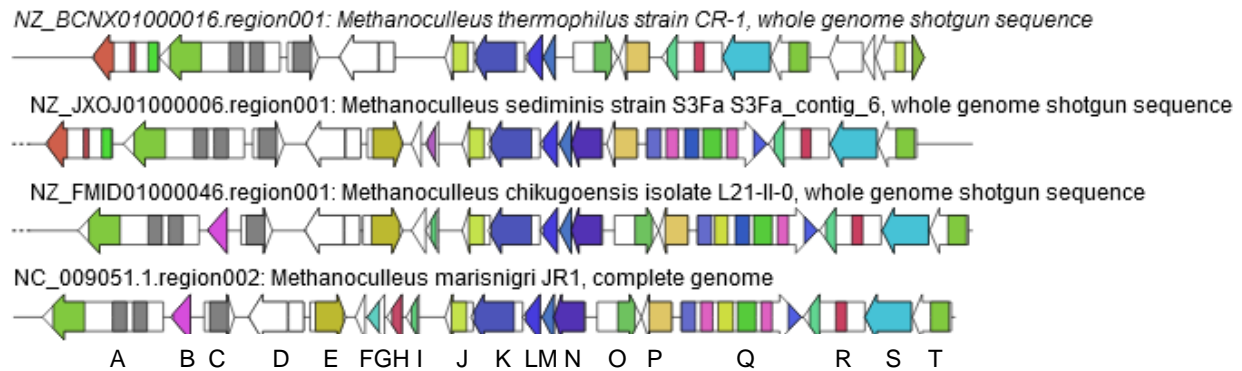
T: hypothetical protein

U: ATP-binding cassette domain-containing protein

V: hypothetical protein

W: 4,5-DOPA dioxygenase extradiol

Figure S10. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Euryarchaeota, *Methanoculleus* sp. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.



A: excinuclease ABC subunit UvrA

B: ferritin family protein

C: DUF5591 domain-containing protein

D: type III ribulose-bisphosphate carboxylase

E: dienelactone hydrolase family protein

F: hypothetical protein

G: winged helix-turn-helix transcriptional regulator

H: hypothetical protein

I: hypothetical protein

J: TfuA

K: YcaO

L: molybdenum cofactor biosynthesis protein MoaE

M: MoaD/ThiS family protein

N: HesA/MoeB/ThiF family protein

O: GNAT family N-acetyltransferase

P: DUF4935 domain-containing protein

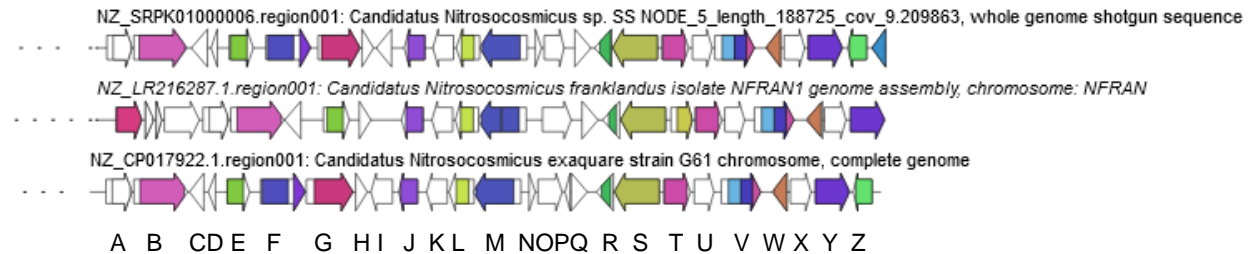
Q: PAS domain S-box protein

R: carboxypeptidase regulatory-like domain-containing protein

S: ABC transporter permease

T: ABC transporter ATP-binding protein

Figure S11. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Thaumarchaeota *Nitrosocosmicus* sp.. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH.



A: ribose 5-phosphate isomerase A

B: amidohydrolase family protein

C: hypothetical protein

D: hypothetical protein

E: ABC transporter ATP-binding protein

F: ABC transporter permease

G: PQQ-dependent sugar dehydrogenase

H: hypothetical protein

I: hypothetical protein

J: YfcE family phosphodiesterase

K: hypothetical protein

L: TfuA

M: YcaO

N: hypothetical protein

O: hypothetical protein

P: hypothetical protein

Q: hypothetical protein

R: 3-isopropylmalate dehydratase small subunit

S: 3-isopropylmalate dehydratase large subunit

T: coenzyme F420-0:L-glutamate ligase

U: hypothetical protein

V: GTP-binding protein

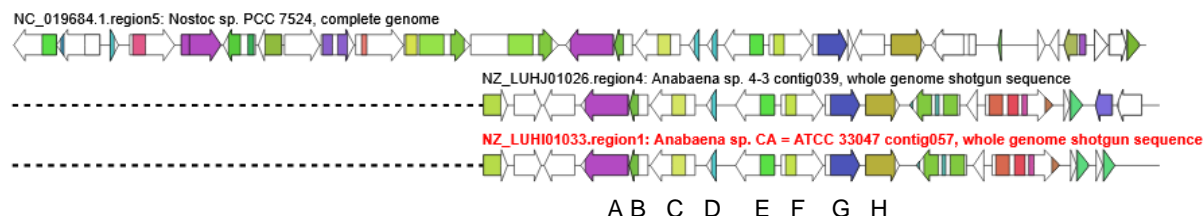
W: reverse transcriptase-like protein

X: hypothetical protein

Y: isocitrate/isopropylmalate dehydrogenase family protein

Z: methyltransferase

Figure S12. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Cyanobacteria. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.



A: tryptophan 7-halogenase

B: 2OG-Fe(II) oxygenase family protein

C: SPASM domain-containing protein

D: precursor peptide (Nif11 family protein)

E: MBL fold metallo-hydrolase

F: TfuA

G: YcaO

H: NAD(P)/FAD-dependent oxidoreductase

Putative precursor peptide 2 (Cluster 275)

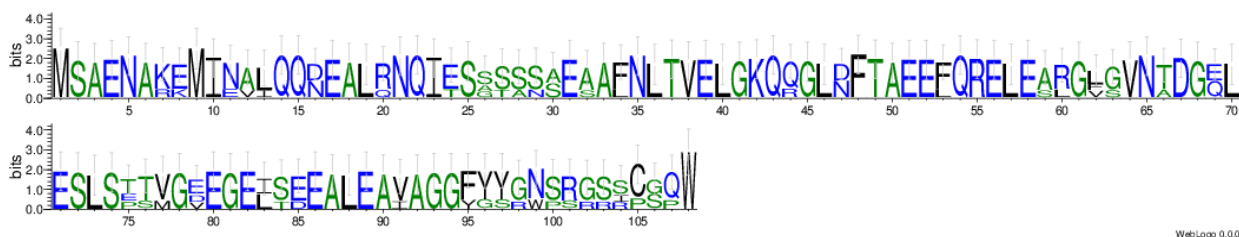
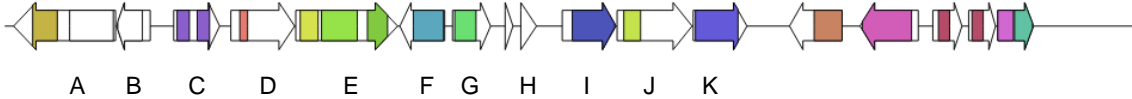


Figure S13. Conserved gene cluster architecture of TfuA - YcaO gene cluster families with most members in Cyanobacteria *Nostoc sp.*. Genes in bold indicate additional biosynthetic or transport genes as annotated by antiSMASH. Sequence logos of predicted precursor peptides using RiPPER are indicated.

NZ_MRBZ01116.region1: *Nostoc calcicola* FACHB-389 FACHB-389_Scaffold_116, whole genome shotgun sequence



NZ_AP018223.1.region3: *Nostoc linckia* NIES-25 plasmid plasmid1 DNA, nearly complete genome



A: adenylate/guanylate cyclase domain-containing protein

B: DUF928 domain-containing protein

C: cyclic nucleotide-binding domain-containing protein

D: NHLP bacteriocin system secretion protein

E: NHLP family bacteriocin export ABC

transporter peptidase/permease/ATPase subunit

F: cupin-like domain-containing protein

G: thiopeptide-type bacteriocin biosynthesis protein

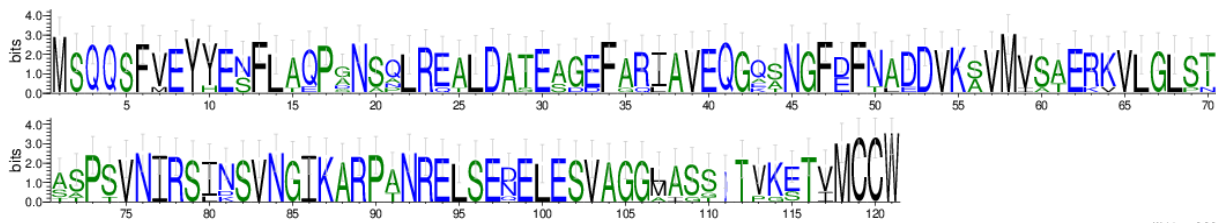
H: precursor peptide 1

I: YcaO

J: TfuA

K: cupin-like domain-containing protein

Putative precursor peptide (Cluster 136)



WebLogo 0.0.0

Figure S14. Sequence similarity network of TfuA protein sequences generated using EFI-EST.

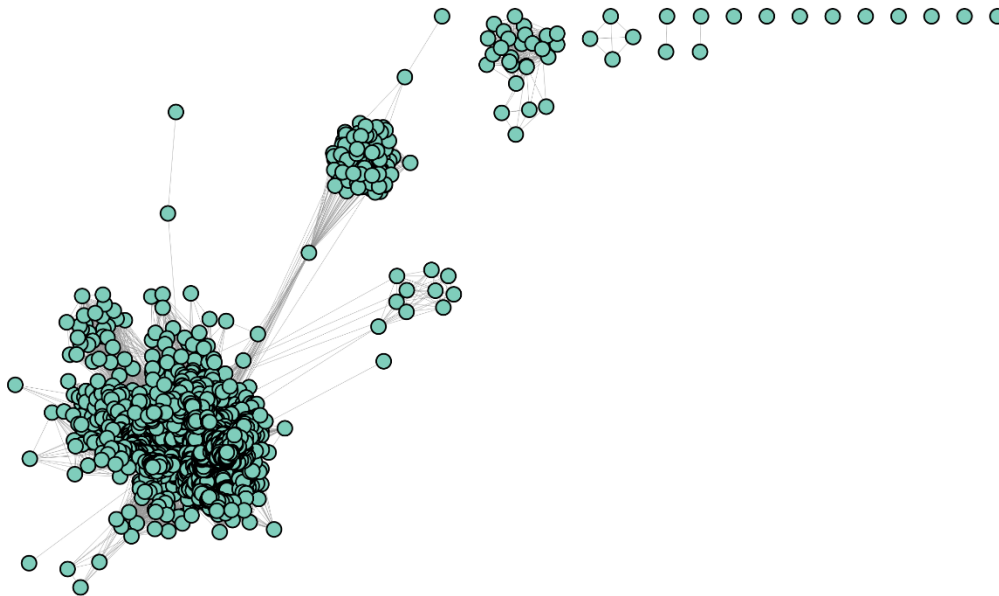
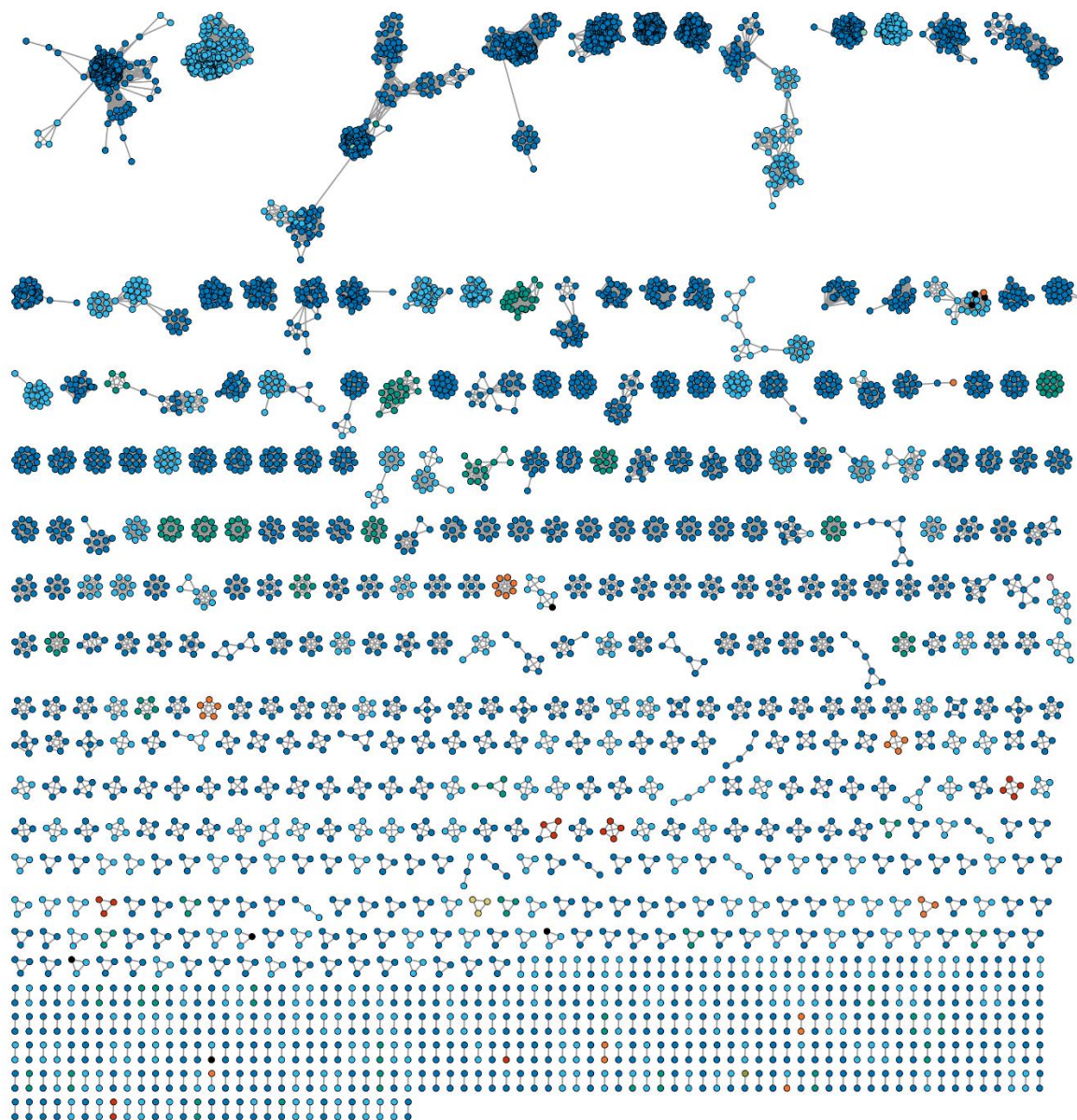


Figure S15. Sequence similarity network of precursor peptides identified through RiPPER within TfuA-related BGCs.



- | | | |
|---|--|---|
| ■ Proteobacteria | ■ Acidobacteria | ■ Firmicutes |
| ■ Actinobacteria | ■ Bacteroidetes | ■ Nitrospirae/Tectomicrobia group |
| ■ Euryarchaeota | ■ Balneolaeota | ■ Nitrospirae |
| ■ Cyanobacteria | ■ Candidatus Korarchaeota | ■ Unidentified |
| ■ Thaumarchaeota | ■ Chlamydiae | |

Figure S16. TfuA-related GCFs containing an CCRG-2 family-like precursor peptide as predicted by RiPPER. GCF from *Nostoc sp.* share the same predicted precursor peptide network. Sequence logo of predicted precursor peptide using RiPPER is indicated.

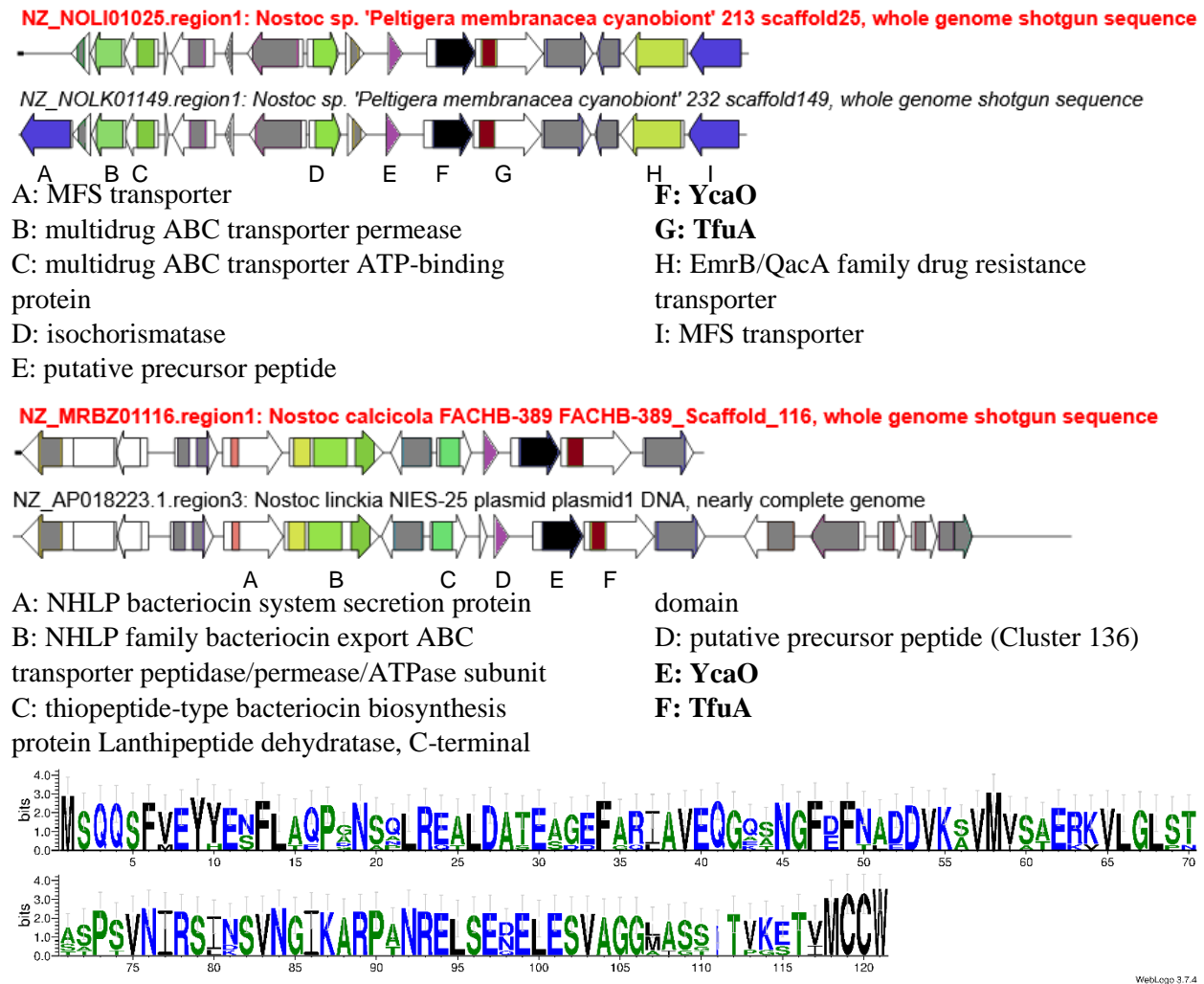


Figure S16 (cont). TfuA-related GCFs containing an CCRG-2 family-like precursor peptide as predicted by RiPPER. GCFs from *Nostoc* share the same predicted precursor peptide network. Sequence logo of predicted precursor peptide using RiPPER is indicated.

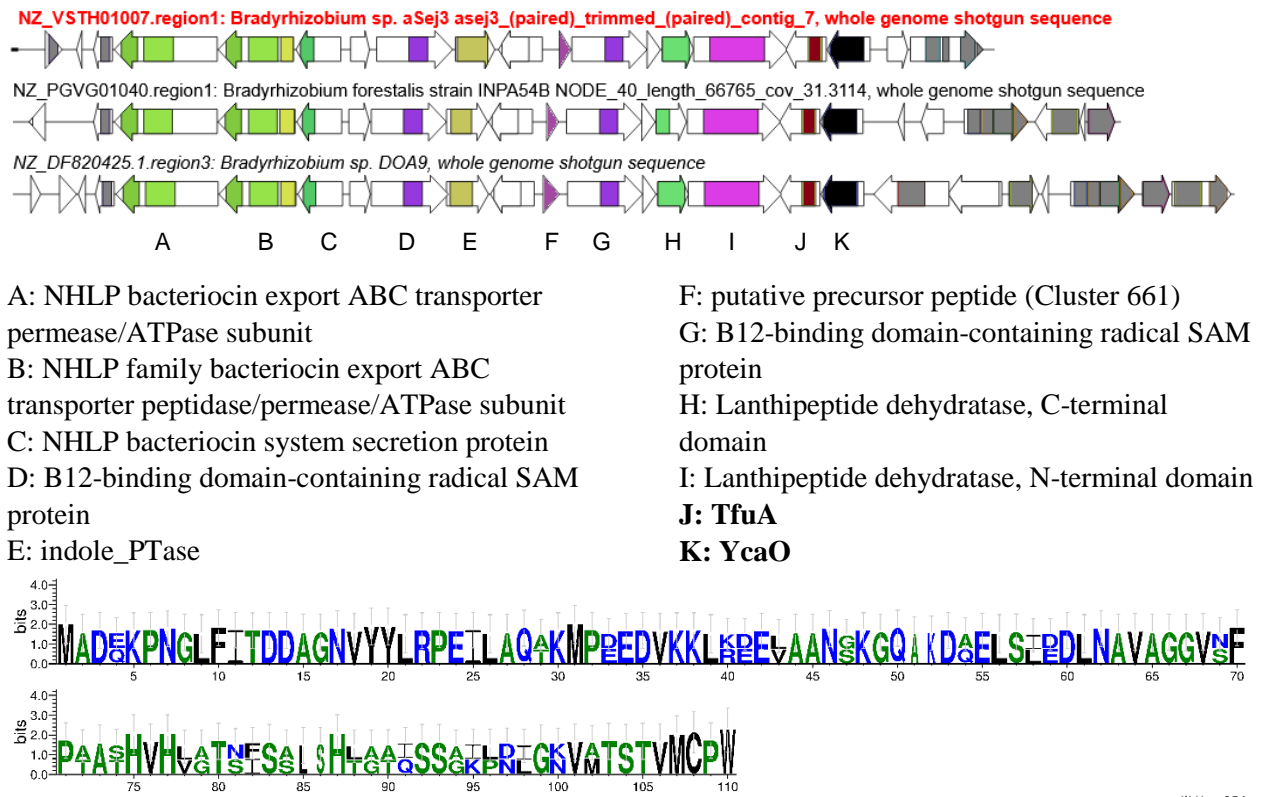
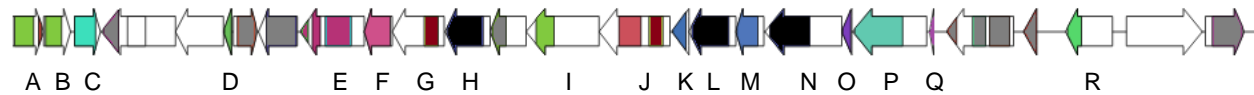


Figure S17. TfuA-related GCFs containing an albusnodin-like precursor peptide as predicted by RiPPER. Sequence logo of predicted precursor peptide using RiPPER is indicated.

NZ_FNVT01015.region3: *Nonomuraea solani* strain CGMCC 4.7037, whole genome shotgun sequence



A: branched-chain amino acid transport system
ATP-binding protein
B: branched-chain amino acid transport system
ATP-binding protein
C: HAD-superfamily hydrolase
D: acyl carrier protein
E: AMP-binding protein
F: thioesterase
G: TfuA
H: YcaO
I: ABC transporter ATP-binding protein
J: protein-L-isoaspartate(D-aspartate) O-
methyltransferase with TfuA domain
K: Transglutaminase-like superfamily protein

L: YcaO (ribosomal protein S12
methylthiotransferase accessory factor)
M: SagB-type dehydrogenase domain-
containing protein
N: YcaO (thiazole/oxazole-forming peptide
maturase, SagD family component)
O: Coenzyme PQQ synthesis protein D (PqqD)
P: albusnodin/ikarugamycin family macrolactam
cyclase
Q: putative precursor peptide (Cluster 249)
R: Cytochrome P450

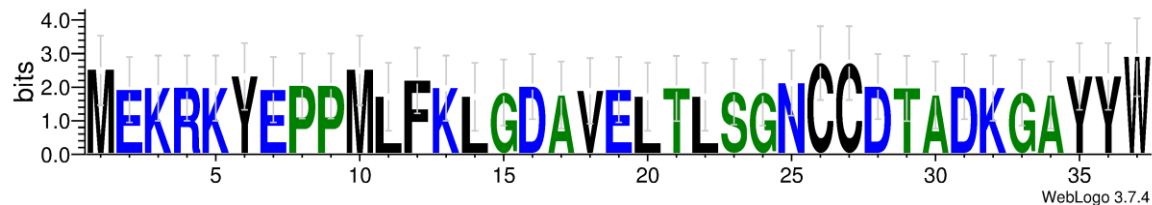
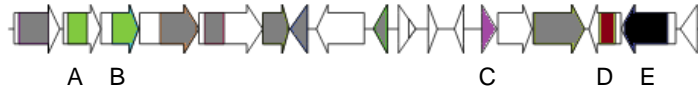


Figure S18. TfuA-related GCFs containing berninamycin-like precursor peptide as predicted by RiPPER. GCFs from *Rhizobium* share the same predicted precursor peptide network. Sequence logo of predicted precursor peptide using RiPPER is indicated.

NZ_SJNQ01000005.region002: *Rhizobium laguerreae* strain SPF2A11 ESF2A1105, whole genome shotgun sequence



A: ABC transporter permease

B: ABC transporter ATP-binding protein

C: putative precursor peptide (Cluster 288)

D: TfuA

E: YcaO

NZ_KI912103.1.region2: *Rhizobium leguminosarum* bv. *trifolii* CC278f Rleg6DRAFT_RLE.7, whole genome shotgun sequence



A: ABC transporter substrate-binding protein

B: ABC transporter permease

C: ABC transporter permease

D: ABC transporter ATP-binding protein

E: ABC transporter ATP-binding protein

F: EAL domain-containing protein (diguanylate cyclase)

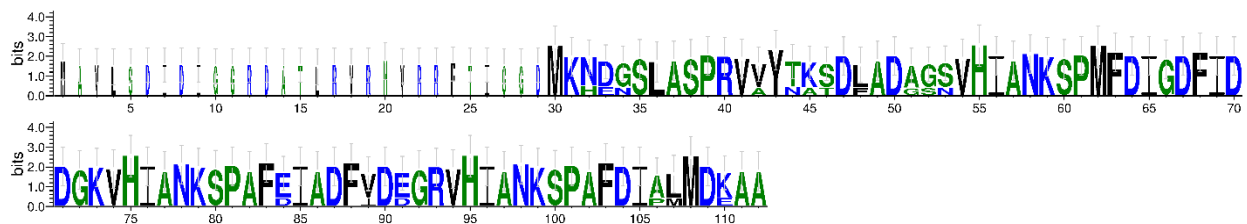
G: TfuA

H: YcaO

I: putative precursor peptide (Cluster 288)

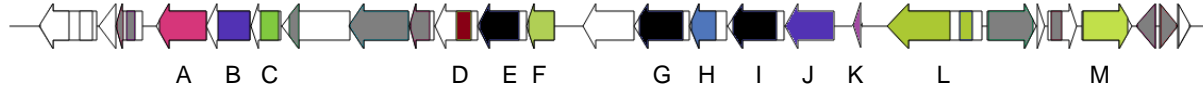
J: GGDEF domain-containing protein (diguanylate cyclase)

K: SDR family oxidoreductase



WebLogo 3.7.4

NZ_VFNH01004.region1: *Herbaspirillum* sp. SJZ107 Ga0314257_104, whole genome shotgun sequence



A: aminotransferase class V-fold PLP-dependent enzyme

B: ThiF family adenylyltransferase

C: ABC transporter ATP-binding protein

D: TfuA

E: YcaO

F: ATP-binding cassette domain-containing protein

G: YcaO

H: SagB/ThcOx family dehydrogenase

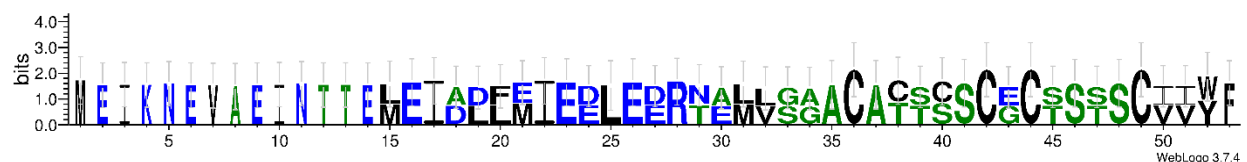
I: YcaO

J: TOMM precursor leader peptide-binding protein (ThiF)

K: putative precursor peptide (Cluster 486)

L: TonB-dependent receptor

M: multidrug efflux MFS transporter



WebLogo 3.7.4

Figure S19. TfuA-related GCF containing with the precursor peptide annotated as a bacteriocin containing thiopeptide-type modifications as predicted by RiPPER. Sequence logo of predicted precursor peptide using RiPPER is indicated.

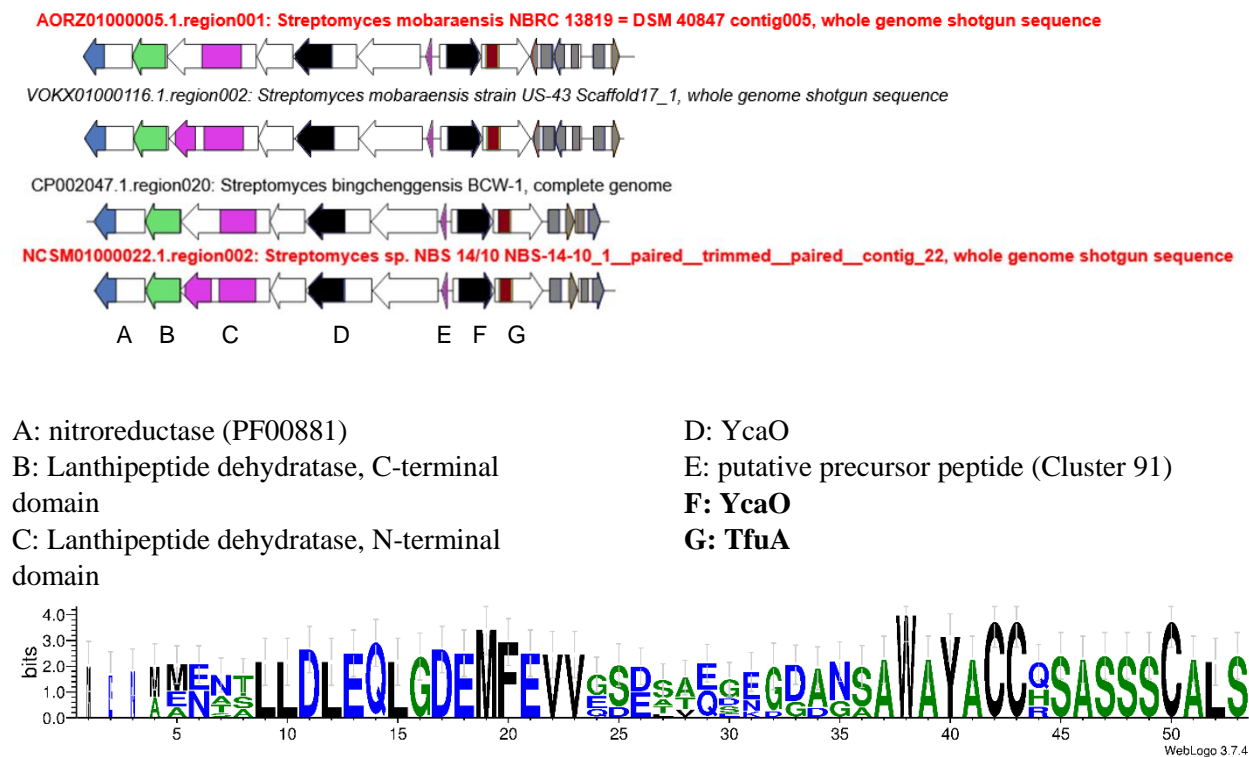
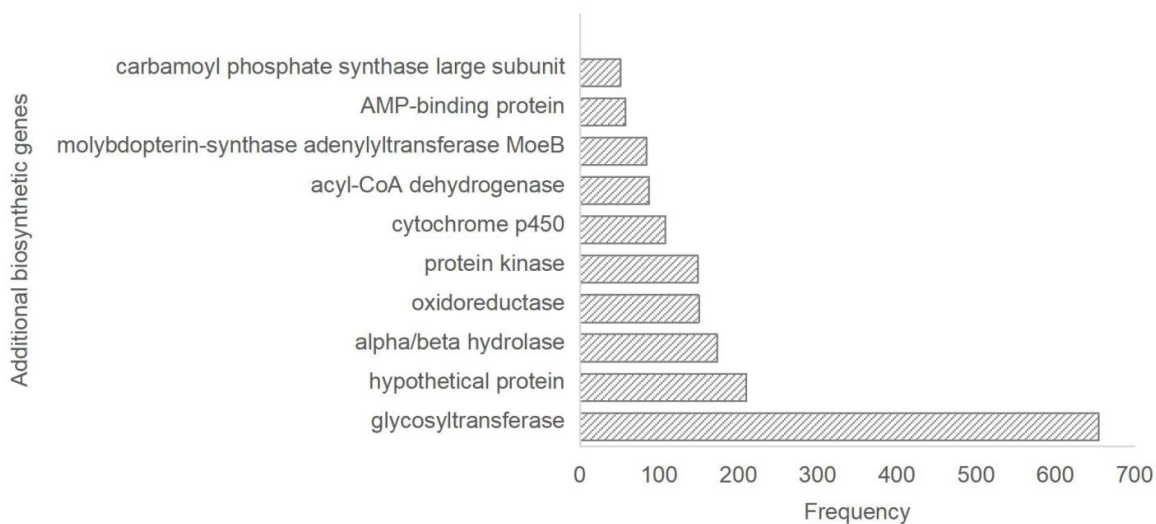


Figure S20. TfuA-related GCFs containing Nif-like precursor peptides as predicted by RiPPER.



Figure S21. Frequencies of additional modification genes (A) and other genes (B) as annotated by antiSMASH found within TfuA-related BGCs.

A



B

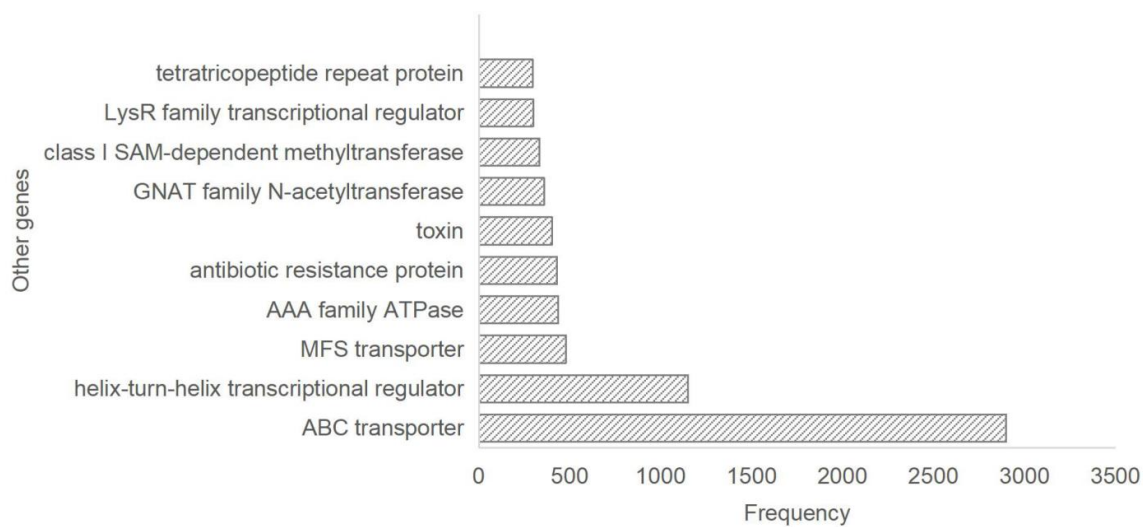


Figure S22. TfuA-related GCF containing an iron-containing redox enzyme family protein. Sequence logo of predicted precursor peptide using RiPPER is indicated.

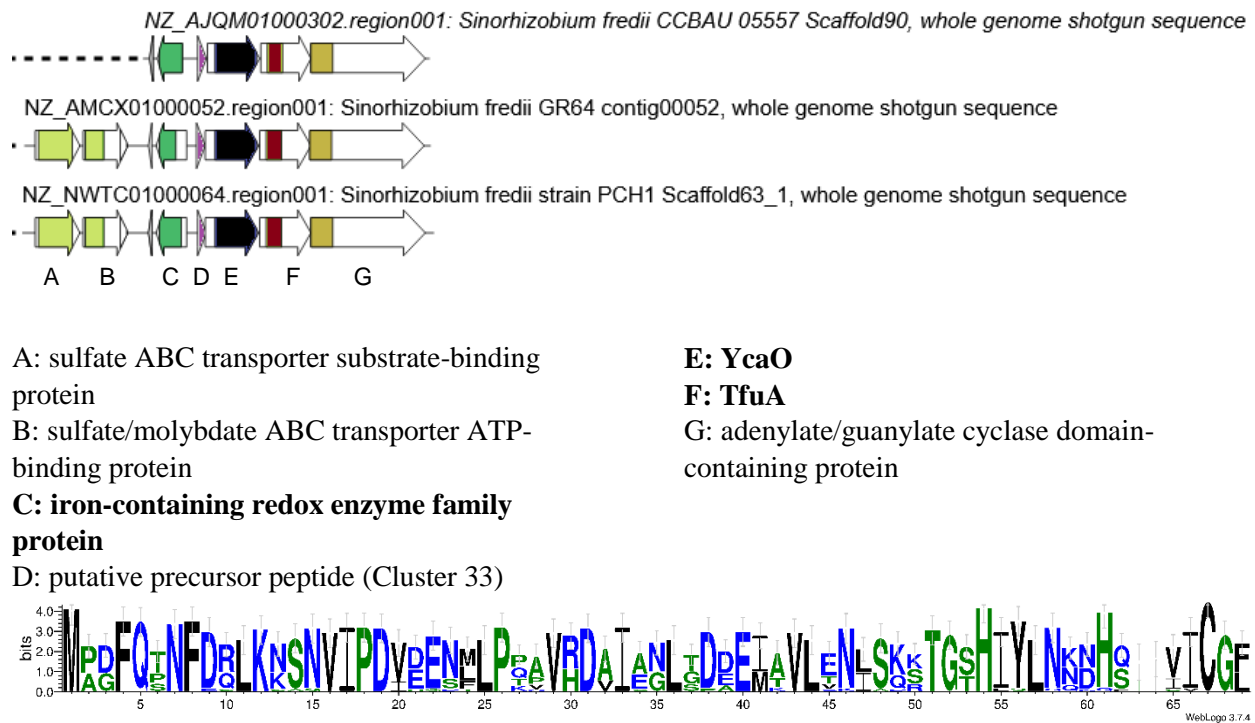


Figure S23. TfuA-related GCF containing a fused *tfuA* and *ycaO* genes. Sequence logo of predicted precursor peptide using RiPPER is indicated.

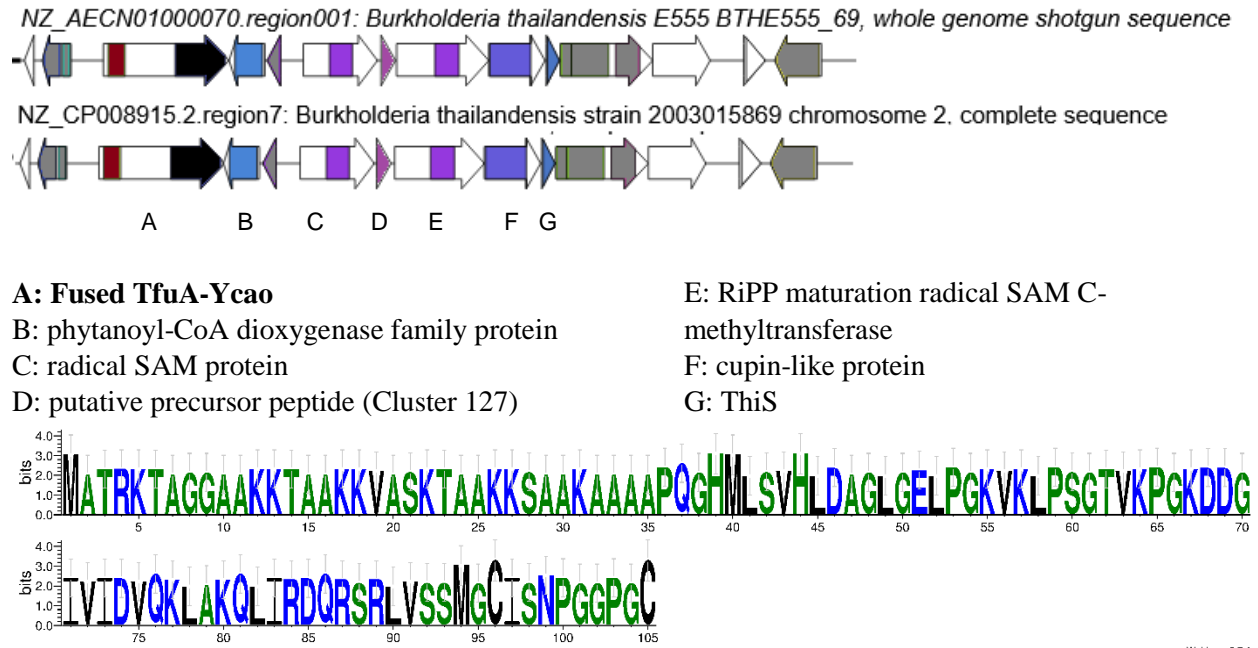
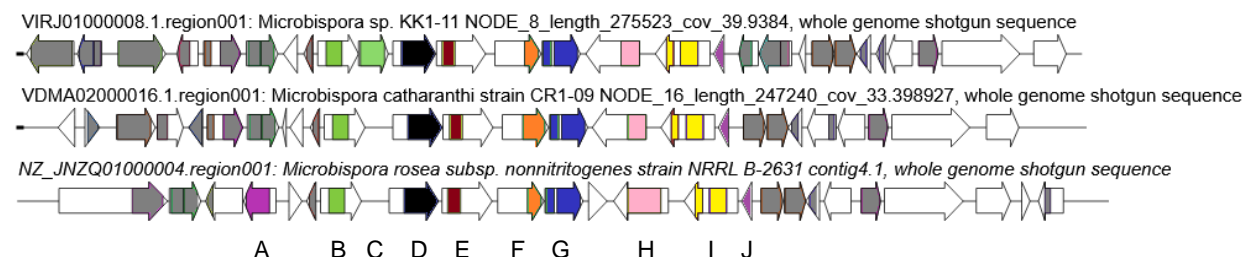
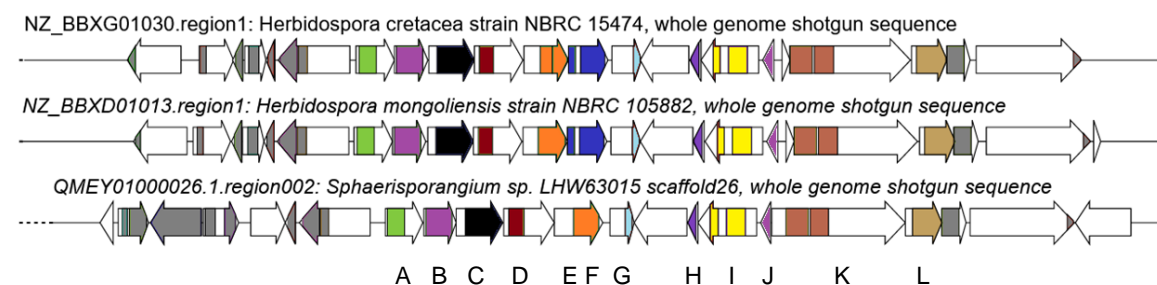


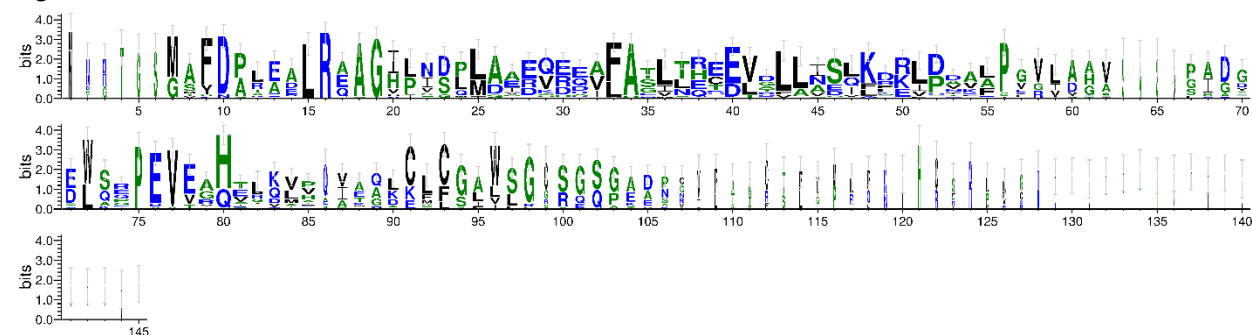
Figure S24. TfuA-related GCF annotated to also produce ranthipeptides as annotated by antiSMASH. Both GCFs share the same precursor peptide network. Sequence logo of predicted precursor peptide using RiPPER is indicated.



- A: alpha/beta hydrolase
 B: ABC transporter ATP-binding protein
 C: ABC transporter permease
D: YcaO
E: TfuA
 F: cytochrome P450
 G: O-methyltransferase
 H: FAD-dependent monooxygenase
 I: radical SAM protein (ranthipeptide)
 J: putative precursor peptide (Cluster 14)



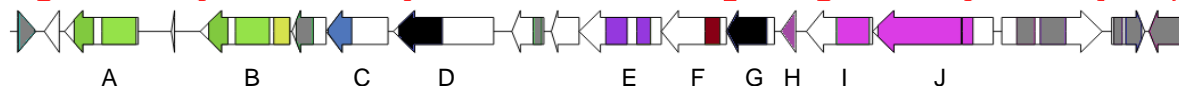
- A: ABC transporter ATP-binding protein
 B: ABC transporter permease
C: YcaO
D: TfuA
 E: cytochrome P450
 F: O-methyltransferase
 G: LuxR family DNA-binding response regulator
 H: PqqD family protein (Stand_Alone_Lasso_RRE PF05402)
 I: radical SAM protein (ranthipeptide)
 J: putative precursor peptide (Cluster 14)
 K: adenylate/guanylate cyclase domain-containing protein
 L: S8 family serine peptidase



WebLogo 3.7.4

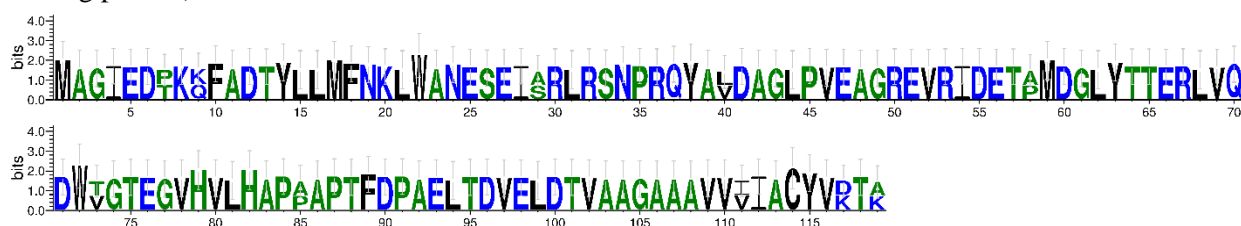
Figure S25. TfuA-related GCF annotated to also produce linear azole-containing peptides as annotated by antiSMASH. Sequence logo of predicted precursor peptide using RiPPER is indicated.

NZ_KB893721.1.region1: Frankia elaeagni strain BMG5.12 B207DRAFT_scaffold_23.24, whole genome shotgun sequence



A: ATP-binding cassette domain-containing protein
 B: NHP family bacteriocin export ABC transporter peptidase/permease/ATPase subunit
 C: SagB/ThcOx family dehydrogenase
 D: YcaO (TOMM precursor leader peptide-binding protein)

E: RiPP maturation radical SAM protein 1
 F: TfuA
 G: YcaO
 H: putative precursor peptide (Cluster 609)
 I: Lanthipeptide dehydratase, C-terminal domain
 J: Lanthipeptide dehydratase, N-terminal domain



WebLogo 3.7.4

Figure S26. TfuA-related GCFs annotated to also produce thiopeptides as annotated by antiSMASH. Sequence logo of predicted precursor peptide using RiPPER is indicated.

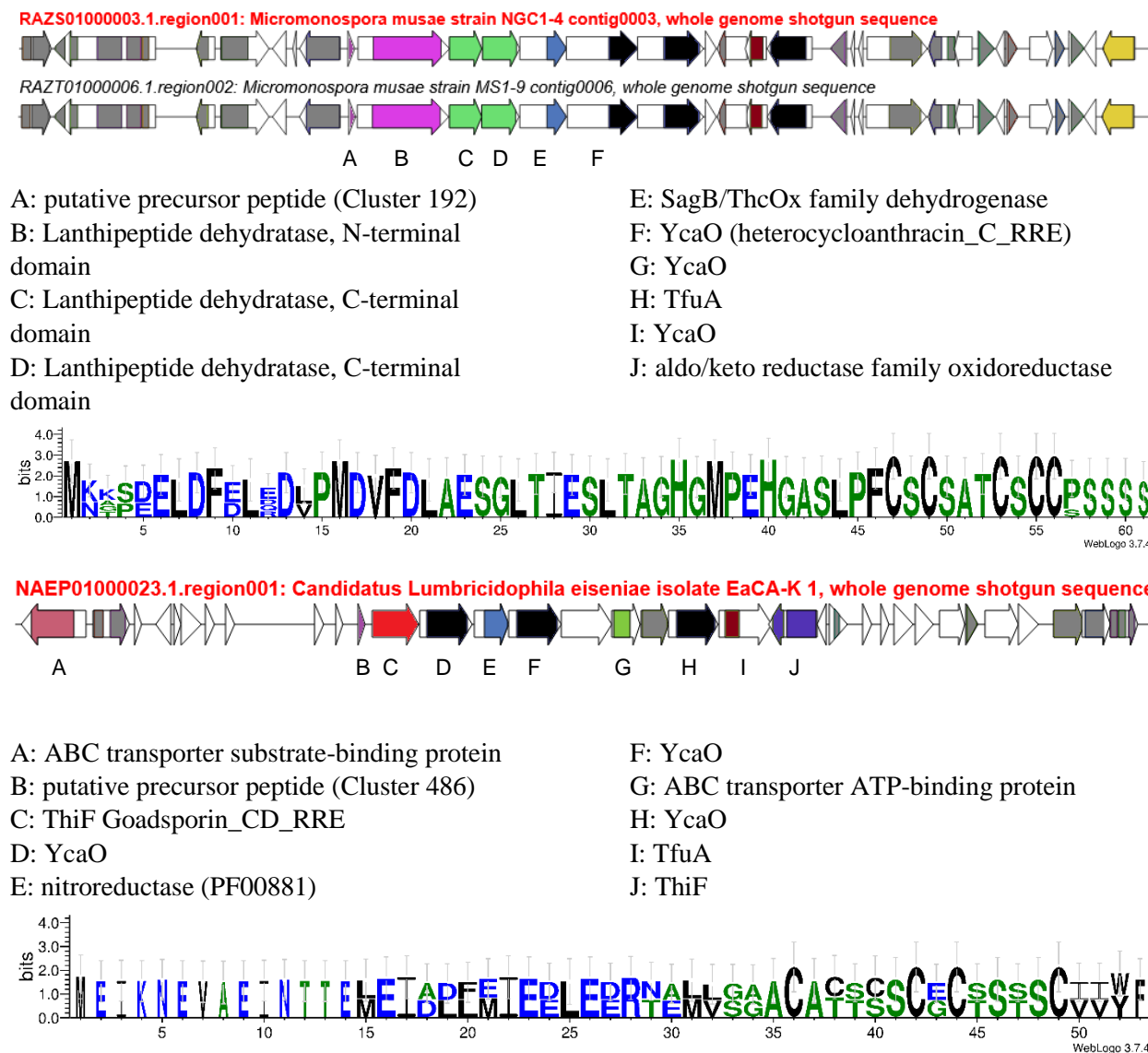
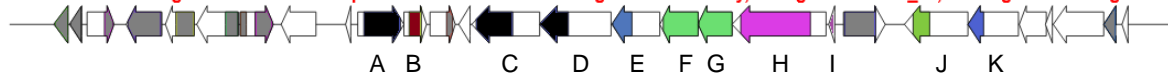
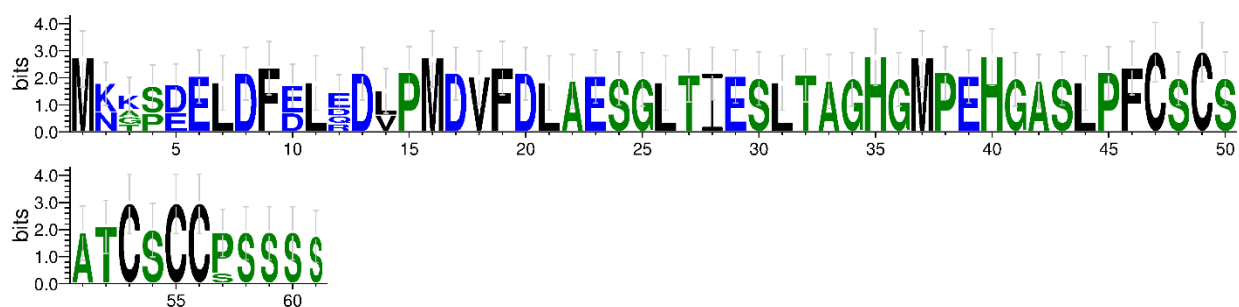


Figure S26 (cont.). TfuA-related GCFs annotated to also produce thiopeptides as annotated by antiSMASH. Sequence logo of predicted precursor peptide using RiPPER is indicated.

FMHZ01000002.1.region004: *Micromonospora citrea* strain DSM 43903 genome assembly, contig: Ga0070606_12, whole genome shotgun sequence



- | | |
|--|---|
| A: YcaO | G: Lantibiotic biosynthesis dehydratase C-term |
| B: TfuA | H: Lanthipeptide dehydratase, N-terminal domain |
| C: YcaO (ribosomal protein S12 methylthiotransferase accessory factor) | I: precursor peptide (Cluster 192) |
| D: YcaO | J: Alpha/beta hydrolase family protein (Peptidase_S9) |
| E: SagB-type dehydrogenase domain-containing protein | K: dehydrogenase |
| F: Lantibiotic biosynthesis dehydratase C-term | |



WebLogo 3.7.4

Figure S27. TfuA-related GCFs containing multiple *ycaO* genes without further annotation. Sequence logo of predicted precursor peptide using RiPPER is indicated.

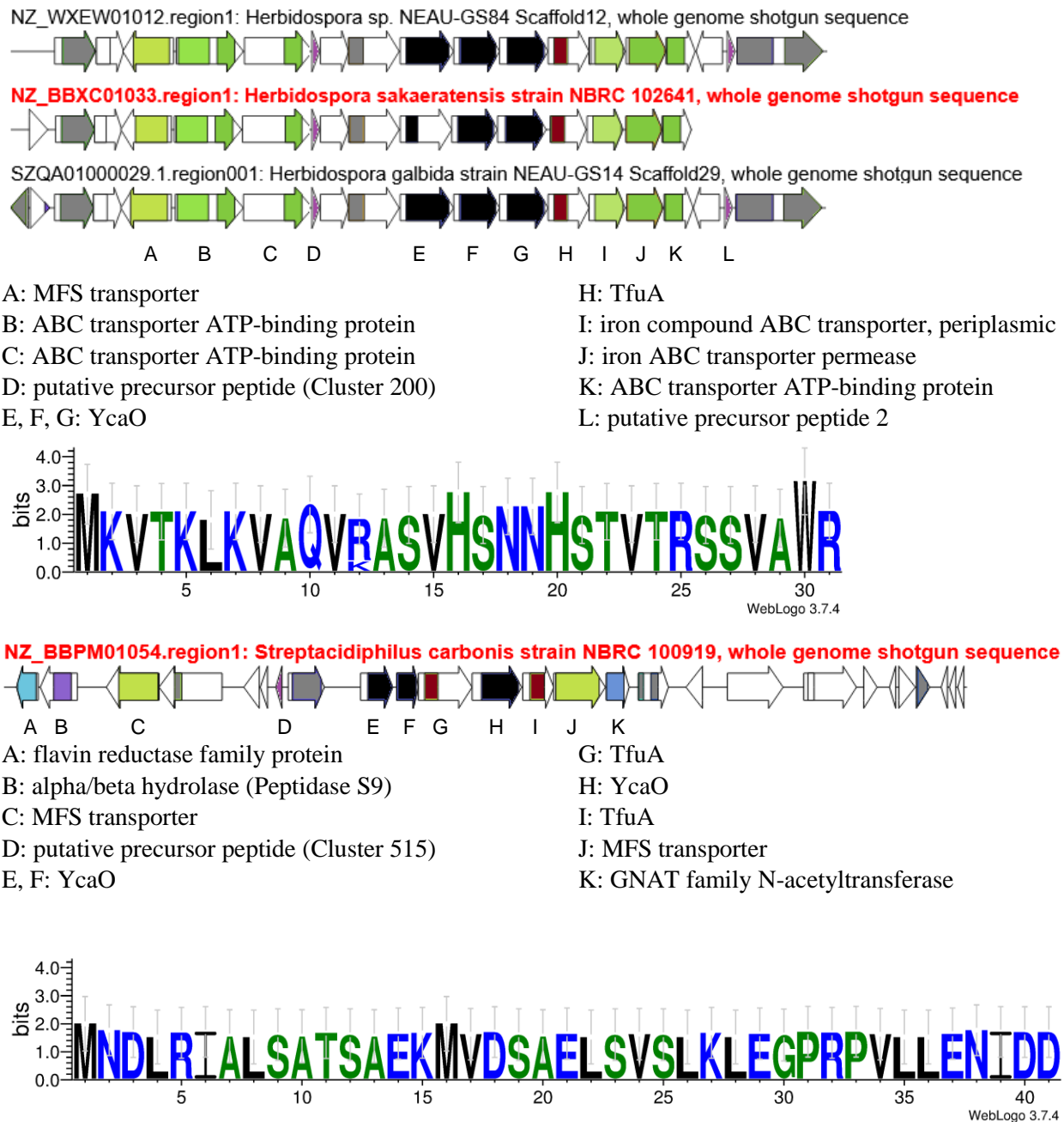


Figure S27 (cont.). TfuA-related GCFs containing multiple *ycaO* genes without further annotation. Sequence logo of predicted precursor peptide using RiPPER is indicated.

AP012204.1.region007: Microlunatus phosphovorus NM-1 DNA, complete genome



- | | |
|---|--|
| A: ABC transporter permease/ATP-binding protein | G: TfuA |
| B: ABC transporter permease/ATP-binding protein | H: prolyl endopeptidase (Peptidase S9) |
| C: putative precursor peptide (Cluster 224) | I: MatE efflux family protein |
| D, E, F: YcaO | J: ABC transporter substrate-binding protein |
| | K: ABC transporter ATP-binding protein |
| | L: ABC transporter permease protein |

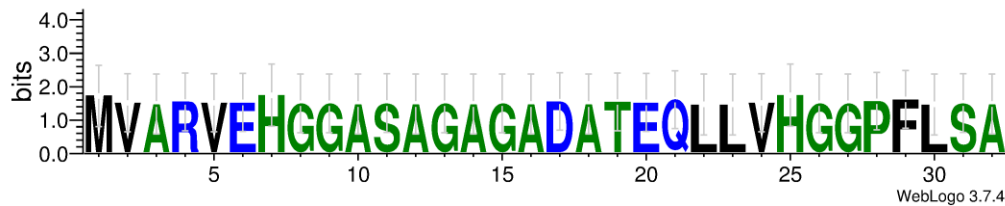


Figure S28. Sequence logo of the ATP-binding sequence motifs identified from TfuA-associated YcaO proteins.

