

System	phyletic distribution	comments
I. Systems with two primary components		
A. Systems with a nucleotide-synthesizing enzyme		
1. Systems with a nucleotide-synthesizing enzyme coupled to enzymatic effector domains		
SMODS[+AGS-C], [unk.2+]HNH+SAVED (Figure 2C-D)	$\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria, bacteroidetes, actinobacteria, firmicutes, <i>Coleofasciculus</i> , <i>Rhodopirellula</i>	Basic system often couple to Ub conjugation system (see below). The unk.2 domain of unknown homology is primarily restricted to $\beta$ -proteobacteria.
SMODS[+AGS-C], Rease+SAVED (Figure 2C-D)	$\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria, <i>Zavarzinella</i>	Very few total number of systems with this arrangement, most Rease+SAVED-including systems co-occur with Ub or HORMA-TRIP13/Pch2 systems (see below).
SMODS+AGS-C, Saf-2TM+SAVED (Figure 1C, 2D)	<i>Acholeplasma</i> , $\alpha$ -, $\gamma$ -proteobacteria, <i>Cupriavidus</i> , firmicutes, <i>Polaribacter</i>	AGS-C fusion nearly universal. Saf-2TM domain contains several conserved secondary structure elements in regions between TMs (Supplementary Material).
SMODS[+AGS-C], Caspase-like peptidase+SAVED (Figure 2C-D)	<i>Myxococcus</i> , <i>Hyalangium</i> , <i>Oscillatoria</i>	Limited phyletic spread. The cyanobacterium <i>Oscillatoria</i> contains further fusion to P-loop NTPase between the Caspase-like peptidase and SAVED domains.
SMODS, metallopeptidase (MPTase)+SAVED (Figure 2C-D)	$\delta$ -proteobacteria	Limited phyletic spread.
SMODS, REase (Figure 2C)	<i>Mycoplasma</i> , $\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria, firmicutes, <i>Tolypothrix</i>	Lack of clear receptor domain indicates secondary messenger likely directly activates REase domain.
SMODS[+AGS-C, insert], Patatin (Figure 2C)	<i>Verrucomicrobia</i> , $\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Paenibacillus</i> , cyanobacteria, bacteroidetes, actinobacteria	AGS-C fusion present in diverse lineages. SMODS with characteristic helical insert found in $\gamma$ -proteobacteria. Patatin effector likely directly activated by secondary messenger. Some $\gamma$ -proteobacteria are adjacent to a DisA_N domain.
CARF+HEPN+mCpol (Figure 1G)	cyanobacteria, <i>Candidatus Magnetobacterium bavaricum</i>	Minimal conflict system encoded on a single polypeptide, with mCpol generating secondary messenger which likely binds CARF, potentially activating HEPN RNase effector.
2. Systems with a nucleotide-synthesizing enzyme coupled to potential pore-forming effectors		
SMODS[+AGS-C], SLATT.4 (Figure 1C)	<i>Mycoplasma</i> , $\alpha$ -, $\beta$ -, $\delta$ -, $\epsilon$ -, $\gamma$ -proteobacteria, firmicutes, deinococci, <i>Prochloron</i> , bacteroidetes, actinobacteria, euryarchaeota	AGS-C fusion restricted mostly to $\gamma$ -proteobacteria and actinobacteria. SLATT likely functioning as membrane-perforating effector, SMODS as generates secondary messenger which could bind to variable SLATT C-terminal helical tail region, effectively "gating" the pore.
SMODS+AGS-C, SLATT.6+TM (Figure 1C)	<i>Leptospira</i> , $\alpha$ -, $\beta$ -, $\delta$ -, $\epsilon$ -, $\gamma$ -proteobacteria, firmicutes, bacteroidetes	AGS-C fusion nearly absolutely conserved in system. AGS-C potentially competes with SLATT C-terminal tail for secondary messenger binding. SLATT domain rapidly diverging, fused to additional TM at C-terminus.

RelA/SpoT, SLATT.1, NA37/YejK (Figure 1E)	$\gamma$ -proteobacteria	Found across a range of enteric $\gamma$ -proteobacteria, core RelA/SpoT-NA37/YejK association also found in bacteroidetes, firmicutes. DNA-binding role of NA37/Yejk likely acts as recognition component, activating RelA/SpoT and secondary messenger production. Secondary messenger binds to SLATT domain, activating pore-forming effector activity.
SMODS[+AGS-C], S-4TM (Figure 1C)	<i>Brachyspira</i> , <i>Pedosphaera</i> , $\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Candidatus Brocadia</i> , <i>Leptotrichia</i> , firmicutes, <i>Stanieria</i> , bacteroidetes, actinobacteria	AGS-C fusion rarely observed. One rare occasions, systems with two or more SMODS domains are observed in actinobacteria and $\beta$ -proteobacteria. SMODS domains typically flank the S-4TM-encoding gene on the genome.
SMODS[+AGS-C], S-2TM $\beta$ (Figure 1C)	<i>Mollicutes</i> , <i>Treponema</i> , $\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Fusobacterium</i> , firmicutes, <i>Scytonema</i> , <i>Nitrolancea</i> , bacteroidetes, actinobacteria, euryarchaeota	S-2TM $\beta$ domain contains 2 N-terminal TMs with C-terminal $\beta$ -strand rich region possibly related to lipocalin-like fold which could bind secondary messenger. Very few AGS-C fusions observed.
DisA_N, S-2TM $\beta$ (Figure 1F)	$\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria, <i>Streptomyces</i>	DisA_N likely displacing the SMODS domain in more widely distributed system immediately above. On rare occasions (including <i>Streptomyces</i> ), SMODS is adjacent in the genome.
mCpol, S-2TM $\beta$ (Figure 1G)	$\alpha$ -, $\gamma$ -proteobacteria, <i>Clostridium</i>	mCpol displaces SMODS domain. This and linkage to DisA_N implies S-2TM $\beta$ capable of recognizing broad spectrum of secondary messengers. System tends to co-occur on the genome with Ub conjugation systems.
mCpol, CARF+2TM (Figure 1G)	cyanobacteria, bacteroidetes, <i>Thermomonospora</i>	Secondary messenger-binding CARF domain with C-terminal, potential pore-forming TM helices. One version is directly fused.
SMODS+AGS-C, 2TM+NUDIX (Figure 1C)	<i>Mycoplasma</i> , firmicutes, <i>Flavobacterium</i>	AGS-C fusion universally present. Highly divergent NUDIX domain contains unusual extended “outflow” region and 2 N-terminal TM helices; however, all active site residues conserved. NUDIX likely functions as effector.
[GrpB, DUF2204-like], S-4TM (Figure 1D)	<i>Porphyromonas</i> , $\gamma$ -proteobacteria	GrpB N-terminally fused, DUF2204-like C-terminally fused pol $\beta$ superfamily members. Likely generating a secondary messenger which interacts with S-4TM.
B. Systems with a nucleotide-binding or processing protein		
TIR, Sir2+SLOG.STALD (Figure 3D)	<i>Mycoplasma</i> , $\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Sebaldella</i> , firmicutes, <i>Chloroflexi</i> , bacteroidetes	Often multiple TIR domains present. In three firmicutes, all three components are fused in the same polypeptide (Figure 3D). Sir2 likely functions as effector, TIR domain appears inactive, could function in nucleotide recognition.

SLOG.YspA, TIR, Caspase (Figure 3E)	$\alpha$ -, $\delta$ -proteobacteria, <i>Bacillus</i> , <i>Tolypothrix</i> , actinobacteria	Caspase likely functions as effector. TIR domain in this group likely inactive.
SLOG.YpsA, SLATT.3+SLATT.1, [TIR] (Figure 1H)	<i>Micromonas</i> , verrucomicrobia, $\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Pelodictyon</i> , actinobacteria, euryarchaeota, cyanobacteria	<i>Micromonas</i> chlorophyte acquisition suggests very late horizontal transfer into eukaryotes. Three domains are typically fused on the same polypeptide. TIR and/or TPR repeats also observed.
TIR, [SLATT.3+]SLATT.1 (Figure 1I)	$\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, bacteroidetes, actinobacteria, <i>Staphylococcus</i> phage, firmicutes, euryarchaeota	TIR domain appears to be inactive, could bind nucleotide or nucleotide-derived ligand, gating the SLATT domain(s). Multiple TIR domains sometimes present, these could multimerize.
SLOG.LDcluster2, TIR (Figure 3F)	$\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Verrucomicrobium</i> , <i>Rhodopirellula</i> , firmicutes, cyanobacteria, <i>Chlorobium</i> , bacteroidetes, <i>Persephonella</i> , actinobacteria, <i>Ixodes</i>	2 copies of TIR can be fused to LDcluster2, with up to 4 total copies of TIR in the systems, including neighborhood associations. At least one TIR domain in the system appears enzymatically active, potentially processing nucleotide or derived substrate. Alternatively, it could act as the effector in this system.
SLOG.LDcluster2, TIR, Patatin (Figure 3G)	$\beta$ -, $\delta$ -, $\gamma$ -proteobacteria	Patatin domain often directly fused to SLOG domain. Typically multiple TIR domains present in systems, at least one of which appears to be enzymatically active. Active TIR domain could process nucleotide substrate or act as co-effector with Patatin.
SLOG.LDcluster3, TIR (Figure 3F)	$\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria, bacteroidetes, actinobacteria, <i>Candidatus Koribacter</i>	Between 1 and 5 copies of the TIR domain found in these systems, typically one appears to be enzymatically active.
SLOG, TIR, [PNP, CASPASE] (Figure 3G)	$\beta$ -, $\delta$ -, $\gamma$ -proteobacteria	SLOG domain belongs to the SIR2/TIR-associating clade, but family membership within the clade remains undetermined. The three domains are fused in a single polypeptide, in <i>Cystobacter</i> PNP displaced by caspase.
II. Multi-component systems combinatorically derived from core systems with two components		
A. Systems with Ubl-conjugation or HORMA-TRIP13/Pch2 components		
1. Complete Ub systems		
SMODS, [unk.1,unk.1+]HNH+SAVED, E2+E1, JAB (Figure 4B,D)	$\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, firmicutes, actinobacteria	Ubl conjugation system either acts as co-effector (with HNH) targeting proteins of invasive element for degradation and/or regulates the activity of core SMODS-HNH+SAVED by targeting them for destruction. 3'-5' RNaseH fold exonuclease sometimes embedded in operon, could function as a secondary effector. Unk.1 and unk.2 domains are widespread, could assist HNH domain in target recognition/discrimination. SMODS, HNH+SAVED often first two gene of neighborhood; however, SMODS is sometimes found as last with

		HNH+SAVED as first, demonstrating circularity in neighborhood construction and implying the protein products form a large complex.
SMODS, Rease+SAVED, E2+E1, JAB (Figure 4B)	$\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Rhodospirellula</i> , firmicutes, <i>Deinococcus</i> , cyanobacteria, bacteroidetes, actinobacteria, <i>Methanobacterium</i>	Same as above, with Rease displacing HNH as effector in the core system. Both SMODS and Rease+SAVED appear as first gene in neighborhood.
SMODS, Saf-2TM+SAVED, E2+E1, JAB (Figure 4B)	$\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria, firmicutes	Same as above 2 systems, with Saf-2TM acting as likely pore-forming effector.
SMODS, TIR+SAVED, E2+E1, JAB (Figure 4B)	actinobacteria, euryarchaeota	Actinobacterial JAB sometimes fused to TIR+SAVED. TIR is potentially active, consistent with syntactical effector position.
SMODS, calcineurin+SAVED, E2+E1, JAB (Figure 4B)	firmicutes	Limited to <i>Clostridium</i> and <i>Thermoanaerobacter</i> , calcineurin phosphoesterase acts as primary effector in the system.
SMODS[AGS-C, insert], Patatin, E2+E1, JAB (Figure 4B)	$\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Lysinibacillus</i> , <i>Gloeocapsa</i> , bacteroidetes, actinobacteria	Insert-containing SMODS, including the DncV protein, belong to this conserved system. Patatin likely acting as primary effector, possibly directly activated by SMODS-generated secondary messenger given absence of candidate receptor domain. Patatin is always first in the gene neighborhood.
SMODS, [JAB+]Rease[+SAVED], E2+E1, JAB (Figure 4B,E)	$\alpha$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Dehalococcoides</i> , bacteroidetes, actinobacteria	In some bacteroidetes, JAB domain N-terminally fused to REase domain, supporting a role for the Ub component in direct regulation of effector. Both SMODS and REase can appear first in gene neighborhood.
SMODS, S-2TM $\beta$ , E2+E1, [JAB] (Figure 4B)	$\alpha$ -, $\gamma$ -proteobacteria, <i>Firmicutes</i> , <i>Bacteroides</i> , <i>Cryptobacterium</i>	Unique to this system, S-2TM $\beta$ often positioned in opposite genome direction from SMODS. JAB sometimes absent.
SMODS, SUa-2TM, E2+E1, JAB (Figure 4B)	<i>Fusobacterium</i> , firmicutes, <i>Coriobacteriaceae</i>	SUa-2TM (SMODS, Ub-associating 2 TM containing) domain is a distinct N-terminal 2TM domain found only in these systems, likely acting as pore-forming effector domain.
SMODS, E2+E1, JAB (Figure 4B)	$\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, firmicutes, bacteroidetes, actinobacteria	Several of these systems are likely to represent inactive fragments of genuine systems.
2. Partial Ub systems		
SMODS, HNH+SAVED, E2 (Figure 4C)	$\alpha$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Rhodospirellula</i> , firmicutes, <i>Leptolyngbya</i> , bacteroidetes, actinobacteria, <i>Acidobacterium</i>	E2 domain is only Ub conjugation component found in system. As all of the above Ub systems likely apprehend a Ub-like (Ubl) domain from elsewhere in the genome, this and the following systems likely similarly apprehend E1/JAB domains. HNH+SAVED consistently the first gene in the neighborhood.

SMODS, REase+SAVED, E2 (Figure 4C)	<i>Anaeromyxobacter</i> , cyanobacteria	Limited distribution. SMODS consistently the first gene in the neighborhood.
SMODS, Saf-2TM, E2 (Figure 4C,D)	<i>Sulfurospirillum</i> , $\gamma$ -proteobacteria	Limited distribution. <i>Serratia</i> has a 3'-5' RNaseH exonuclease in neighborhood, potentially functioning as secondary effector.
SMODS, SAVED+MPTase, E2 (Figure 4C)	<i>Leptospira</i> , <i>Niastella</i> , <i>Fibrisoma</i>	Limited distribution. Only systems where SAVED domain is N-terminal to the effector. <i>Niastella</i> contains complete Ub conjugation system (includes E1 and JAB).
SMODS, Patatin, E2 (Figure 4C)	$\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, firmicutes, <i>Rhodopirellula</i> , bacteroidetes, <i>Calothrix</i> , <i>Streptomyces</i>	While SMODS is typically the first gene in the neighborhood, occasionally Patatin is observed as the first gene.
SMODS, S-2TM $\beta$ , E2 (Figure 4C)	<i>Ponticoccus</i> , bacteroidetes	E2 is generally the last gene in the neighborhood.
SMODS, S-4TM, E2 (Figure 4C)	$\beta$ -proteobacteria, firmicutes, <i>Fusobacterium</i> , bacteroidetes	E2 is usually sandwiched between the SMODS and S-4TM domains, two SMODS domains occasionally present.
SMODS, 2TM+NUDIX, E2 (Figure 4C,E)	bacteroidetes, <i>Blautia</i>	E2 is directly fused to SMODS in bacteroidetes <i>Spirosoma</i> . Same NUDIX family as in above two-component system.
3. HORMA-TRIP13/Pch2 systems		
SMODS, REase+SAVED, HORMA, TRIP13/Pch2, [HORMA] (Figure 4G)	$\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Sediminibacterium</i> sp., actinobacteria	Some systems, including this one, contain second, distinct HORMA domain. Like their eukaryotic counterparts, prokaryotic HORMA likely multimerizes into a scaffold or captures unstructured peptide tails of invasive proteins. TRIP13/Pch2 acts as switch, catalyzing resolution of HORMA captured substrates.
SMODS, SAVED+MPTase, HORMA, TRIP13/Pch2 (Figure 4G)	bacteroidetes	Limited distribution, same C-terminal SAVED MPTase fusion as observed in systems above.
SMODS, 2TM+NUDIX, HORMA, TRIP13/Pch2 (Figure 4G)	bacteroidetes	Limited distribution, same NUDIX family as in above systems.
SMODS, S-4TM, HORMA, HORMA, TRIP13/Pch2 (Figure 4H)	actinobacteria	Limited distribution.
SMODS, REase, HORMA, TRIP13/Pch2, [HORMA] (Figure 4G)	$\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria, <i>Schlesneria</i>	Roughly half of the systems contain the second HORMA domain. Systems lack candidate receptor domain, suggesting SMODS-generated secondary messenger directly interacts with REase.
SMODS, HORMA, TRIP13/Pch2 (Figure 4G)	$\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Bacillus</i> , bacteroidetes, actinobacteria	System lacks effector and receptor domains. Likely fragmented systems akin to SMODS-Ub systems above.
SMODS, RNaseH 3'-5' exonuclease, HORMA, TRIP13/Pch2 (Figure 4I)	$\beta$ -, $\gamma$ -proteobacteria	Limited distribution, same family of 3'-5' RNase exonucleases as above.

SMODS, SLOG.TPALS+[TIR,PNP], HORMA, TRIP13/Pch2, RNaseH 3'-5' exonuclease (Figure 4J)	$\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria	SLOG.TPALS likely functioning as receptor domain, with TIR/PNP acting as the effectors, 3'-5' RNaseH exonuclease (same family as above) potentially acting as secondary effector.
B. Combinations with CRISPR/Cas systems		
CRISPR/Cas coupled to SMODS, SLATT.4 (Figure 2E)	$\beta$ -, $\gamma$ -proteobacteria	Limited distribution, positioning of the SMADS, SLATT.4 pair within and the content of the CRISPR/Cas system variable
CRISPR/Cas coupled to [Lon_C, 2TM]+SAVED (Figure 2E)	thermotogae, deferribacteres, <i>Thermodesulfobacterium</i> , cyanobacteria, aquificae	Cyanobacteria representation limited to <i>Microcystis</i> , 2TM module only found in cyanobacteria and aquificae <i>Persephonella</i> . <i>Persephonella</i> system also includes SMODS. SAVED fusion protein sometimes inserted into CRISPR/Cas system, also found at the genome edge of system.
C. Combinations with diverse R-M and DNA-modification systems		
Classical R-M systems coupled to SLATT- SMODS systems (Figure 2F)	$\beta$ -, $\gamma$ -, $\epsilon$ -proteobacteria, firmicutes, bacteroidetes, actinobacteria, <i>Lentisphaera</i>	Coupled systems predominantly contain SMODS with SLATT.4 or SLATT.6. The nucleotide cyclase or a comparable recognition component is often absent, leaving only the SLATT domain.
PreQ0-based R-M system coupled to SMODS and SLATT.4 (Figure 1M)	$\alpha$ -, $\gamma$ -proteobacteria, firmicutes, <i>Methanococoides</i>	PreQ0 R-M system includes restriction component (HhH DNA glycosylase/nuclease) and modification components (tRNA: guanine transglycosylase (QTG) and QueC-like PP-loop ATPase).
PreQ0-based R-M system coupled to TIR and [SLATT.3+]SLATT.1 (Figure 1M)	$\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria	Inactive TIR displaces SMODS in above system, strongly suggesting functional equivalency, potentially through nucleotide binding. Second TIR domain N-terminally fused to QTG, potentially active.
SLOG.LDcluster2, TIR, TIR+TIR, MazG+MazG-C, Nmad2, $\alpha$ - glutamyl/putresciny l thymine pyrophosphorylase (aGPT-PPlase1), [Uracil DNA glycosylase (UDG)] (Figure 3H)	Leptospira, $\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria, bacteroidetes, Candidatus <i>Blastococcus</i>	Between 1 and 3 copies of the TIR domain, up to two directly fused to SLOG domain. Second TIR domain of TIR+TIR protein potentially active. MazG is a pyrophosphatase, the MazG-C domain likely assists in transfer of a pyrophosphate moiety to substrates. Function of Nmad2 is unknown, but often found in predicted, obscure nucleotide modification pathways. Presence of UDG suggests base transfer involved in modification.
Sir2+SLOG.STALD, TIR, MazG+MazG-C, aGPT-PPlase1, Nmad2, [UDG]	$\alpha$ -, $\gamma$ -proteobacteria, <i>Pedobacter</i>	Same as above, with Sir2 appearing to displace the active TIR domain of the above system.
III. Disparate systems sharing certain components with the above systems		
A. Novel retroelement with diversity-generating potential		
SLATT.5+TM, Reverse Transcriptase (RT) (Figure 1J)	$\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Psychrilyobacter</i> , firmicutes, bacteroidetes	RT domain contains RVT_N N-terminal, RdRP fold RT, and "domain X", known as GIIIM in pfam. "Hyper-mobile", self-replicating retroelement system. SLATT domain likely mediates transport of the RT-generated DNA out of the host cell.

B. Systems linked to NAD utilization and ADP-ribosylation		
SLOG.YspA/SLOG.cpYsA/SLOG.YAcAr associating with diverse NAD+/NADH processing enzymes (Figure 5B-D)	<i>Ktedonobacter</i> (NUDIX), <i>Deinococcus</i> (NADAR), <i>Bacillus</i> and cyanobacteria (MACRO, NADAR), <i>Chryseobacterium</i> (NAD-processing “mega-operons”)	Various associations point to a role for SLOG.YspA in binding a NAD or NAD-derived substrate. Large cyanobacterial proteins display extensive domain polymorphism. SLOG.YAcAr in <i>Chryseobacterium</i> is part of large, previously-described NAD processing “mega-operons”.
IV. Nucleotide-centric systems in non-conflict systems		
A. Prokaryotic and eukaryotic systems regulating ion flux and membrane transport		
SLOG.YpsA, Ion channel+TrkA_N+TrkA_N+RyR+RyR (Figure 5F)	<i>Nitrosomonas</i> , actinobacteria <i>Cyclobacterium</i>	Ion channel distantly-related VIC superfamily 2 TM channels. TrkA_N domain binds NAD. System echoes eukaryotic TRPMs. In <i>Cyclobacterium</i> , SLOG.YpsA domain fused directly to channel.
SLOG.LSDAT.prokaryotic, [SLATT.3]+SLATT.1 (Figure 5E)	cyanobacteria, $\gamma$ -, $\delta$ -proteobacteria, actinobacteria	Can be fused or operonic. SLOG.LSDAT.prok appears active (see Table 1): enzymatic reaction could contribute to channel gating.
SLOG.LSDAT.eukaryotic +Ank+Ank+Ank+Ion_channel+TRP-box+[Pkinase/NUDIX] (Figure 5H)	choanoflagellids, Guillardia, Emiliana, animals, ciliates	Encompasses TRPM1-8 paralogues. Most ciliate versions are standalone, one version fused to distinct ion channel. Ankyrin repeats are highly divergent. Some invertebrates display extensive, distinctive domain accretion (see Supplementary material).
B. Miscellaneous signaling systems		
SLOG.YpsA, cNMP cyclase, TPR repeats (Figure 5I)	$\alpha$ -, $\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, <i>Zavarzinella</i> , <i>Oscillatoriothycideae</i>	Domains fused in a single polypeptide. SLOG.YpsA domain is likely recognizes the cyclic nucleotide generated by the cNMP cyclase. One instance further N-terminally fused to RyR domain.
SLOG.YpsA, TIR, TPR repeats (Figure 5I)	$\beta$ -, $\delta$ -, $\gamma$ -proteobacteria, cyanobacteria, <i>Chlorobium</i> , bacteroidetes, actinobacteria	Appears analogous to the above system, with TIR replacing the cNMP cyclase: likely processes a nucleotide substrate subsequently bound by the SLOG.YpsA domain.
SLOG.YspA, Patatin, [SLATT.3+SLATT.1], [TPR repeats]	<i>Haloferula</i> , $\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria, <i>Microlunatus</i>	SLATT and/or TPR repeats sometimes present in the systems. In <i>Microlunatus</i> , Patatin is directly fused to SLOG.YspA.
cNMP cyclase+AGS-C, pJV1.spdB3 3TM domain (Figure 5J)	$\beta$ -, $\gamma$ -proteobacteria, firmicutes, bacteroidetes	AGS-C fused C-terminally to both SMODS and cNMP cyclase. cNMP cyclase+AGS-C associates operonically with pJV1.spdB3, a 3TM domain contributing to plasmid spreading during <i>Streptomyces</i> conjugation. pJV1.spdB3 likely involved in DNA/nucleotide recognition in signal transduction pathways.
cNMP_cyclase, pJV1.spdB3 3TM domain, cNMPBD+[HTH,TIR] (Figure 5J)	$\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria, actinobacteria	cNMPBD+HTH association could directly enact transcriptional change via binding of cNMP cyclase product. Role of TIR fusion remains unclear; potentially an enzymatically active TIR domain.
HD+pJV1.spdB3 3TM domain, [cNMP_cyclase], proteins with	bacteroidetes, chlorobi	While the cNMP_cyclase association is scattered across only a few bacteria, the fusion to HD domain is strongly reminiscent of signaling systems which utilize HD to hydrolyze cNMPs. Associating domains

predicted roles in eDNA recognition, processing (Figure 5J)		include calcineurin, MBB, and distinct $\beta$ -propeller families with enzymatic activity.
SLOG.YAcAr[+HTH,+KilA-C], PRTase (Figure 5K)	<i>Treponema</i> , $\alpha$ -, $\beta$ -, $\delta$ -, $\epsilon$ -, $\gamma$ -proteobacteria, firmicutes, <i>Fusobacterium</i> , <i>Flexistipes</i> , <i>Pyrinomonas</i>	SLOG.YAcAr domain sometimes fused to either HTH or KilA-C DNA-binding domains. This SLOG family is potentially active, suggesting the PRTase coupling forms a switch acting on a nucleotide-derived messenger substrate: the SLOG domain removes and PRTase removes a phosphoryl-based group. Switch directionality could be mediated by the DNA-binding domain.
SLOG. cpYspA system with several poorly-understood domains	$\alpha$ -, $\beta$ -, $\gamma$ -proteobacteria	SLOG.cpYspA combines with uncharacterized DUF4755 domain, an uncharacterized domain with a single conserved cysteine residue fused to an N-terminal signal peptide, an uncharacterized domain with two conserved cysteine residues and two N-terminal TMs, a small, uncharacterized domain, and a Lon-type AAA+ P-loop NTPase domain N-terminally fused to a signal peptide and Zinc Finger and ferritin-like domains. Presence of several conserved cysteine residues could indicate a transfer cascade involving a phosphodiester intermediate.