Supplementary material

Supplementary Table 1

List of systems, rules and HMM profiles used in DefenseFinder

Supplementary Table 2

Anti-phage systems detected.

Supplementary Table 3

Genes involved in anti-phage systems

Supplementary Table 4

Anti-phage systems per genome

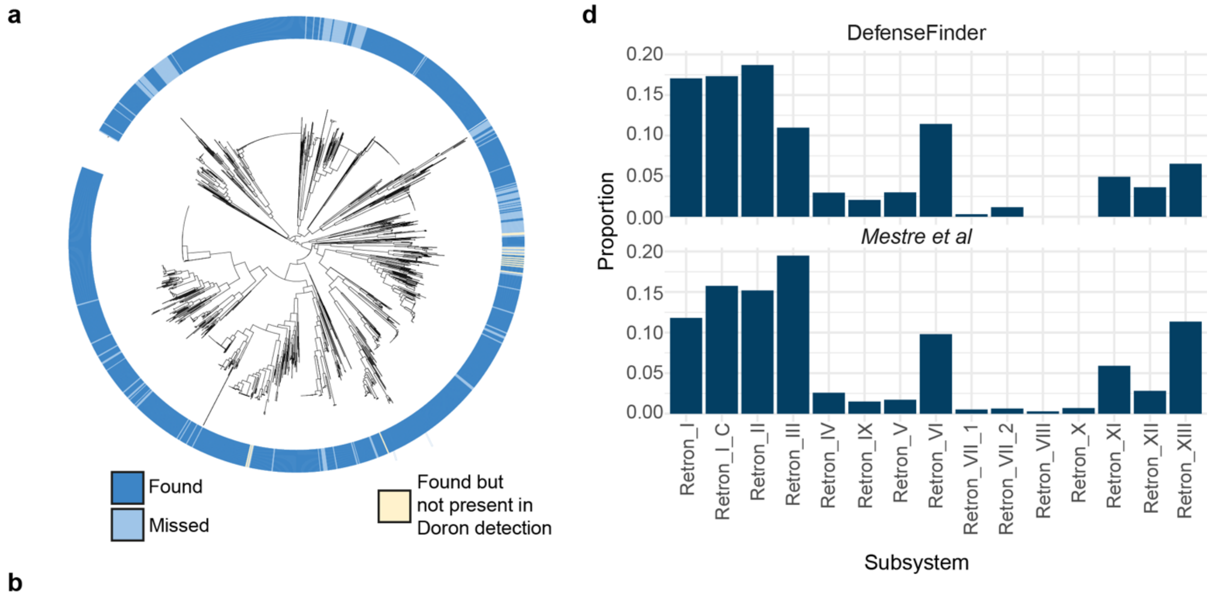
Supplementary Table 5

Detection of prophages

Supplementary Figures

Supplementary figure 1: Validation of DefenseFinder models

a. Phylogenetic tree of SduA. Dark blue was found by an initial round of the DefenseFinder models, light blue was missed (compared to detection from Doron et al) **b.** Table of sensitivity and sensibility for systems from Doron et al. **c.** Validation of the detection of RM **d.** Comparison of distribution of retron subtypes.



b

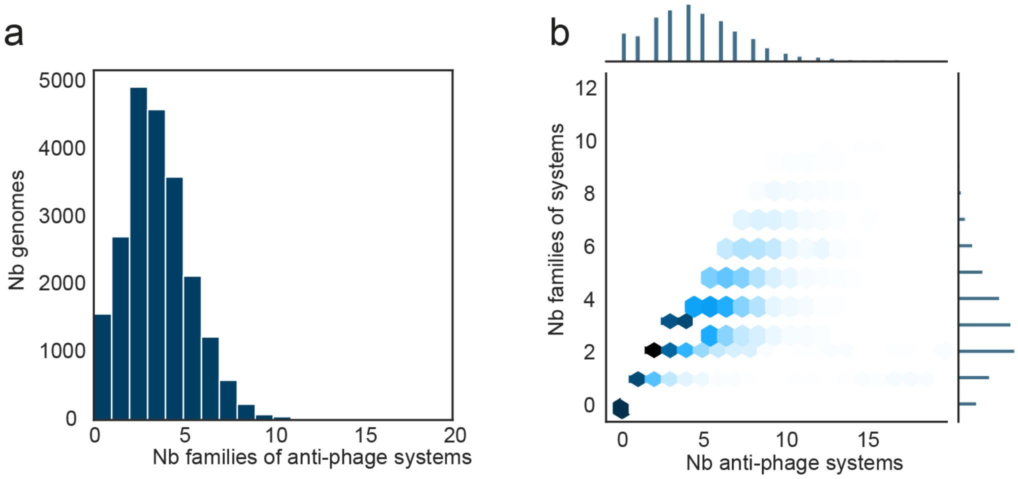
System	Druantia	Gabija	Hachiman	Kiwa	Lamassu	Septu	Shedu	Thoeris	Wadjet	Zorya
Sensitivity	99,0 %	99,1 %	98,8 %	98,9 %	99,2 %	99,4 %	98,7 %	98,5 %	99,3 %	99,0 %
Specificity	98,5 %	99,2 %	99,9 %	99,5 %	99,9 %	99,3 %	99,9 %	99,9 %	99,8 %	99,9 %

c

R-M Protein	Type I M	Type I R	Type II M	Type II R	Type III M	Type III R	Type IV
Sensitivity	98,3 %	98,9 %	91,1 %	89,1 %	97,6 %	83,7 %	77,4 %

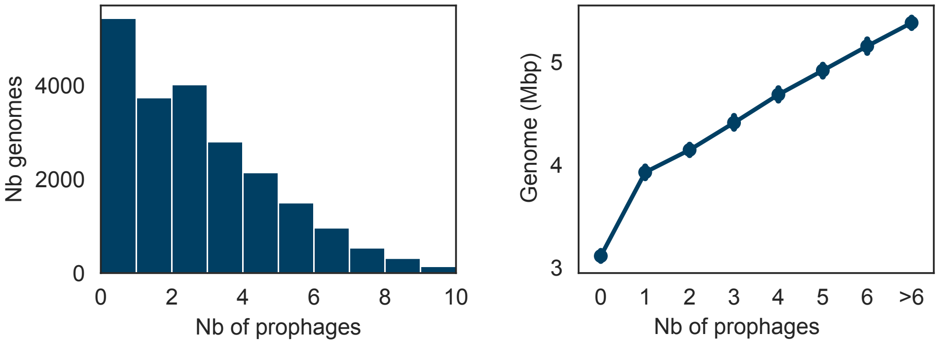
Supplementary figure 2: Families of anti-phage systems are correlated with the total number of anti-phage systems

a. Distribution of the number of families anti-phage systems per genome. The x-axis was cut at 20 for data visualization purposes. **b.** Correlation between the families of anti-phage systems and the total number of anti-phage systems (Spearman ρ=0.79 P-value<0.0001).



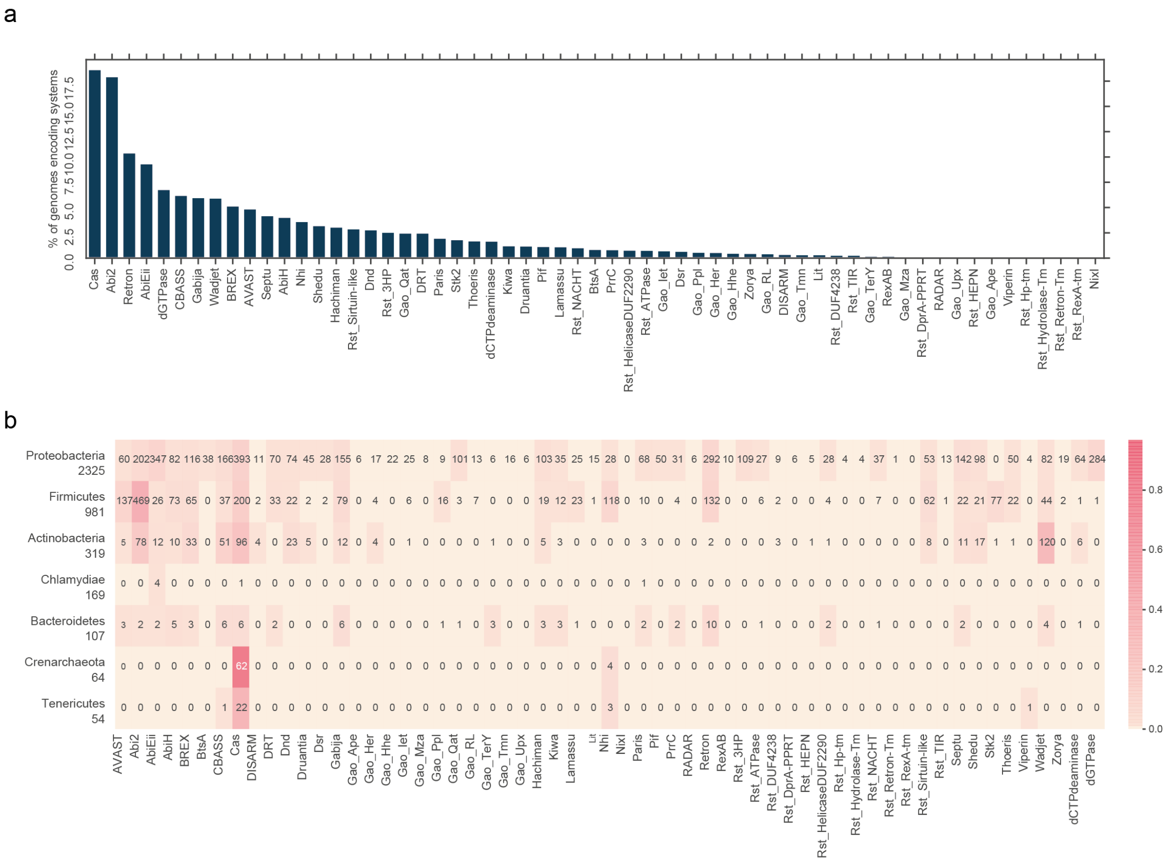
Supplementary figure 3: Prophage detection

a. Distribution of the number of prophages per genome. **b.** Correlation between number of prophages and genome size.



Supplementary figure 4: Distribution of anti-phage systems in genomes without RM

a. Distribution of anti-viral systems. **b.** Distribution per phyla



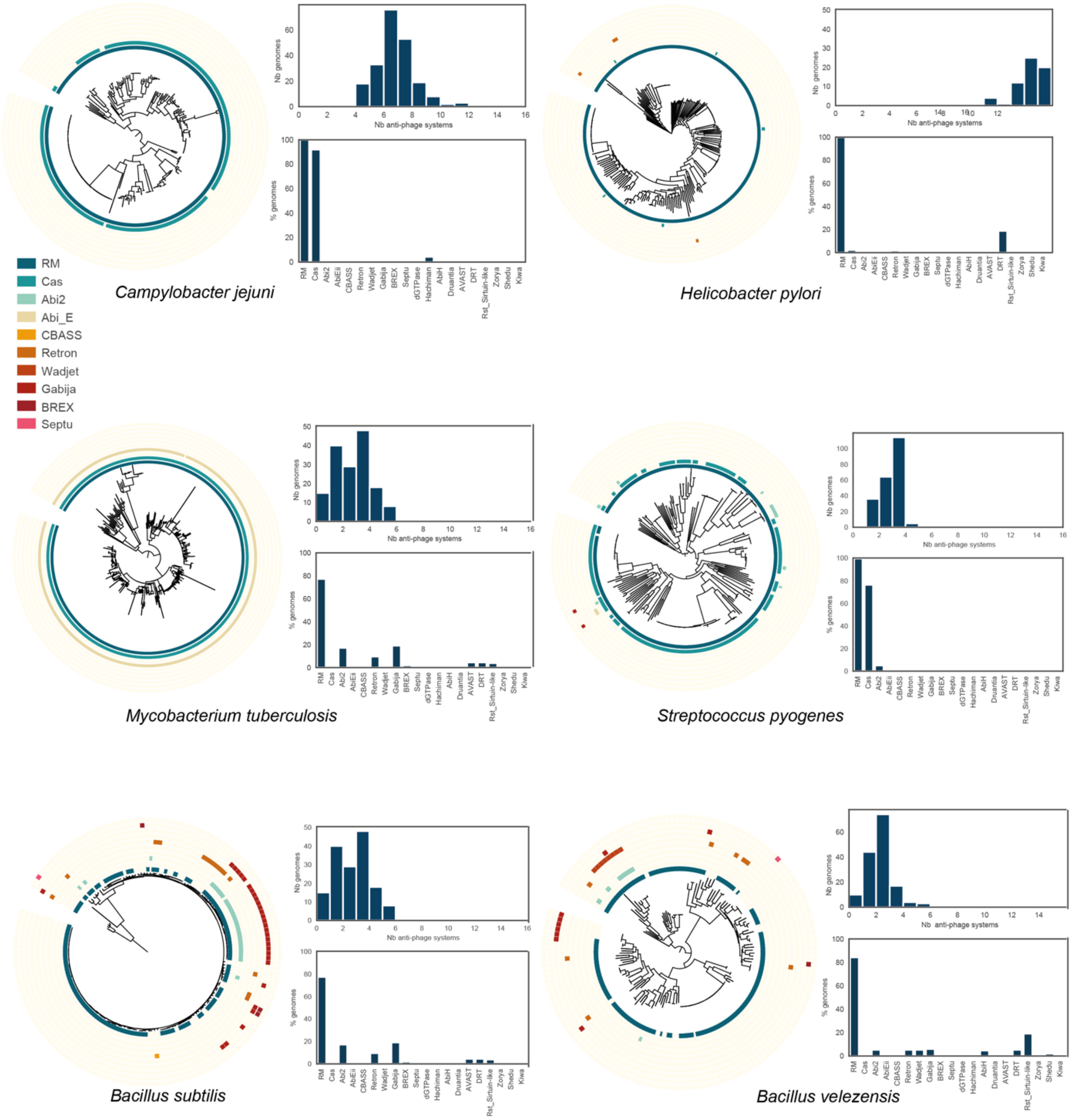
	AVAST	Abi2	AbiEii	AbiH	BREX	BtsA	CBASS	Cas	DISARM	DRT	Dnd	Druantia	Dsr	Gabija	Gao_Ape	Gao_Her	Gao_Hhe	Gao_Iet	Gao_Mza	Gao_Ppl
Proteobacteria 2325	60	202	347	82	116	38	166	393	11	70	74	45	28	155	6	17	22	25	8	9
Firmicutes 981	137	469	26	73	65	0	37	200	2	33	22	2	2	79	0	4	0	6	0	16
Actinobacteria 319	5	78	12	10	33	0	51	96	4	0	23	5	0	12	0	4	0	1	0	0
Chlamydiae 169	0	0	4	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Bacteroidetes 107	3	2	2	5	3	0	6	6	0	2	0	0	0	6	0	0	0	0	0	1
Crenarchaeota 64	0	0	0	0	0	0	0	62	0	0	0	0	0	0	0	0	0	0	0	0
Tenericutes 54	0	0	0	0	0	0	1	22	0	0	0	0	0	0	0	0	0	0	0	0

	Gao_Qat	Gao_RL	Gao_TerY	Gao_Tmn	Gao_Upx	Hachiman	Kiwa	Lamassu	Lit	Nhi	NixI	Paris	Pif	PrrC	RADAR	Retron	RexAB	Rst_3HP	Rst_ATPase	Rst_DUF4238
Proteobacteria 2325	101	13	6	16	6	103	35	25	15	28	0	68	50	31	6	292	10	109	27	9
Firmicutes 981	3	7	0	0	0	19	12	23	1	118	0	10	0	4	0	132	0	0	6	2
Actinobacteria 319	0	0	1	0	0	5	3	0	0	0	0	3	0	0	0	2	0	0	0	3
Chlamydiae 169	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Bacteroidetes 107	1	0	3	0	0	3	3	1	0	0	0	2	0	2	0	10	0	0	1	0
Crenarchaeota 64	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0
Tenericutes 54	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0

	Rst_DprA-PPRT	Rst_HEPN	Rst_HelicaseDUF2290	Rst_Hp-tm	Rst_Hydrolase-Tm	Rst_NACHT	Rst_Retron-Tm	Rst_RexA-tm	Rst_Sirtuin-like	Rst_TIR	Septu	Shedu	Stk2	Thoeris	Viperin	Wadjet	Zorya	dCTPdeaminase	dGTPase
Proteobacteria 2325	6	5	28	4	4	37	1	0	53	13	142	98	0	50	4	82	19	64	284
Firmicutes 981	0	0	4	0	0	7	0	0	62	1	22	21	77	22	0	44	2	1	1
Actinobacteria 319	0	1	1	0	0	0	0	0	8	0	11	17	1	1	0	120	0	6	0
Chlamydiae 169	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bacteroidetes 107	0	0	2	0	0	1	0	0	0	0	2	0	0	0	0	4	0	1	0
Crenarchaeota 64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tenericutes 54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0

Supplementary figure 5: Anti-viral arsenal of diverse bacterial species

Each panel shows the distribution of the total number of systems in the species (top panel), the frequency of the 20 most common anti-phage systems in prokaryotes in this species (bottom panel) and a phylogenetic tree of the species with the presence/absence of the 10 most common anti-phage systems in prokaryotes.



Campylobacter jejuni

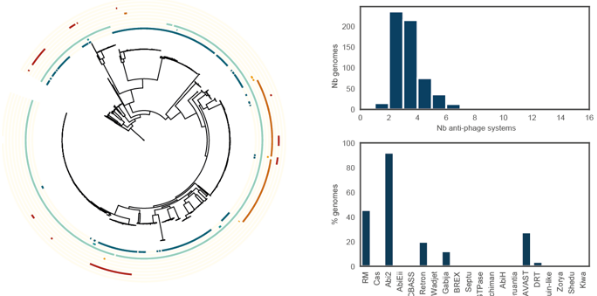
Helicobacter pylori

Mycobacterium tuberculosis

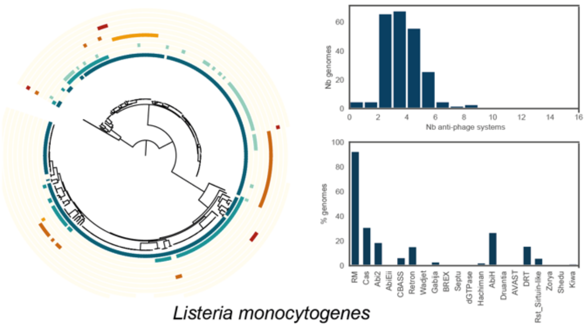
Streptococcus pyogenes

Bacillus subtilis

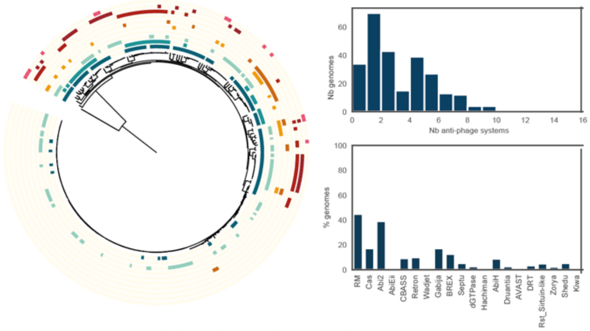
Bacillus velezensis



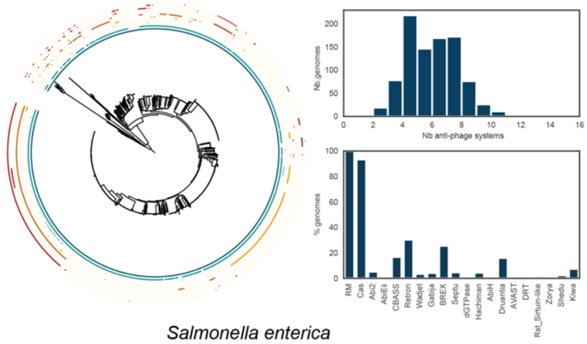
Staphylococcus aureus



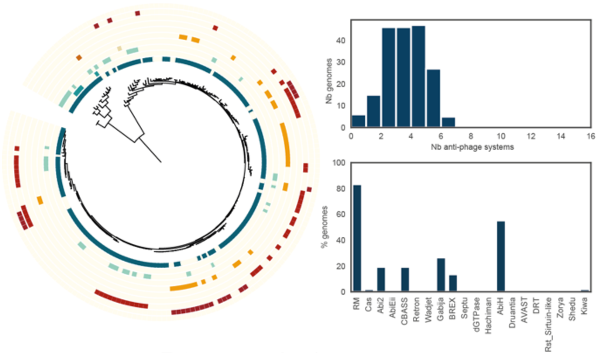
Listeria monocytogenes



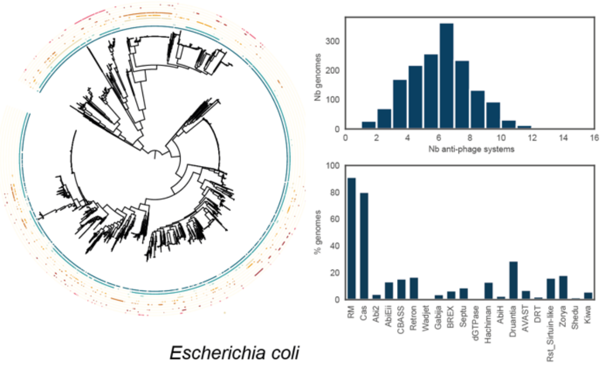
Acinetobacter baumanii



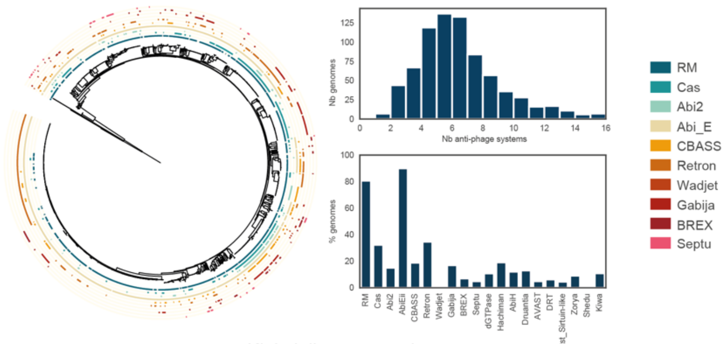
Salmonella enterica



Enterococcus faecium



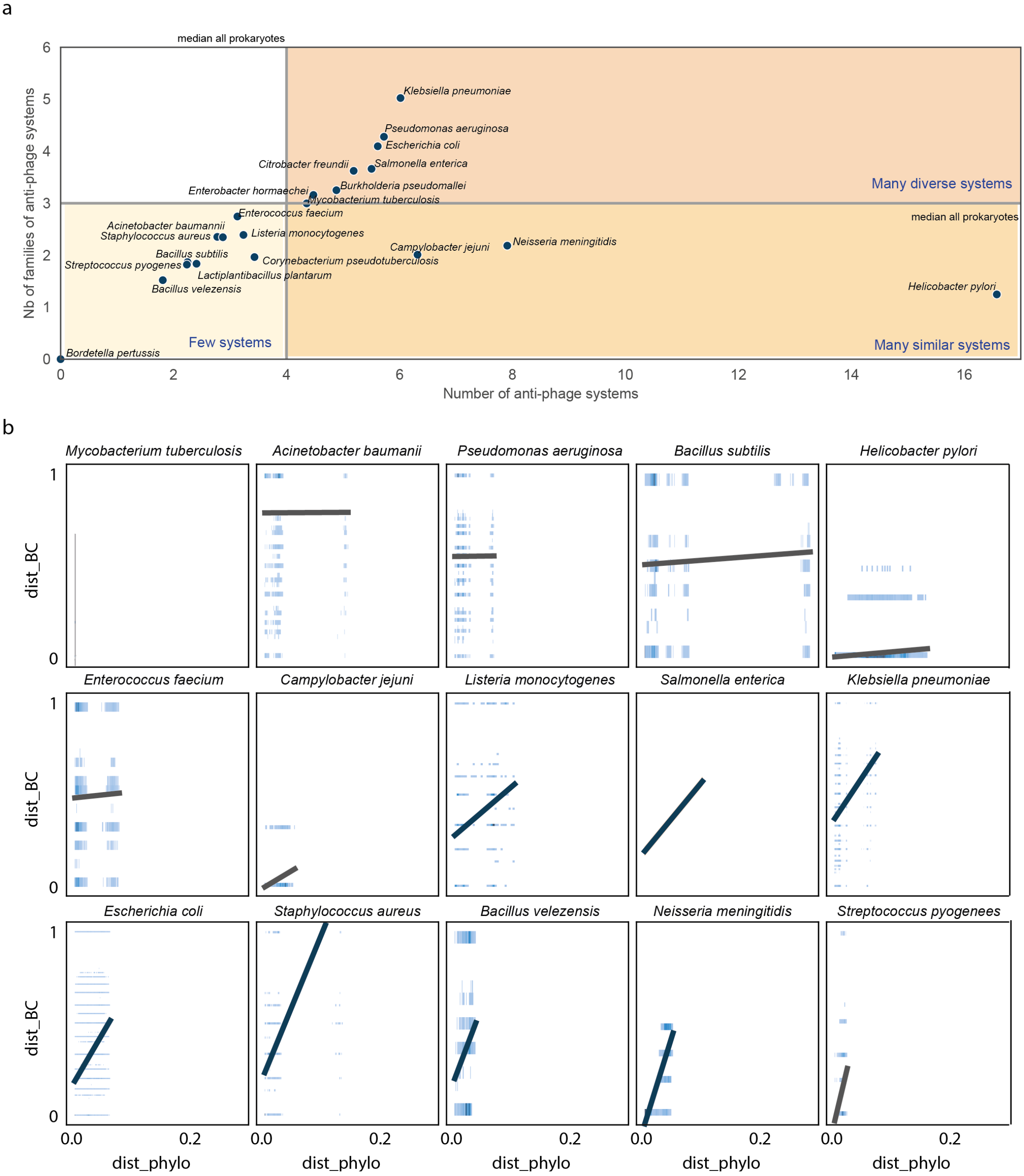
Escherichia coli



Klebsiella pneumoniae

Supplementary Figure 6: Anti-viral arsenal of bacterial species are diverse

a. Scatter plot for the anti-viral arsenal of prokaryotes. **b.** Correlation between the phylogenetic distance and the Bray-Curtis distance of the anti-viral arsenal of diverse species. Each plot corresponds to one species. For each species, the Bray-Curtis distance (dist_BC) of all pairs of anti-viral arsenals was computed as well as the phylogenetic distance (dist_phylo) between all strains. Each line corresponds to a pearson fit of the data.



Supplementary figure 7: Determinants of the anti-viral arsenal of bacterial species

a. Genome size of bacterial species. **b.** Correlation between the families of anti-phage systems and the genome size (Linear regression pearson r=0.65, p-value=0.0014). **c.** Correlation between the number of anti-phage systems and the number of prophages (linear regression r=-0.2 p-value=0.4). **d.** Correlation between the number of anti-phage systems and the genome size (linear regression r=-0.16 p-value=0.5).