

Supporting Information

Wang et al. 10.1073/pnas.1401734111

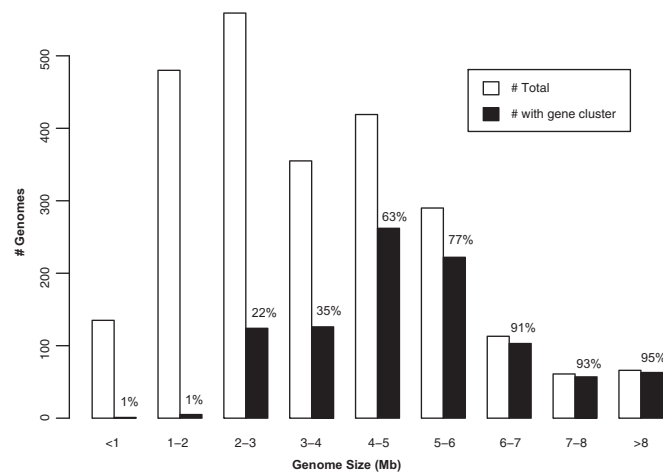


Fig. S1. Distribution of bacterial strains found with gene clusters among different genome-size ranges. The gene clusters are nearly absent from those strains with a genome size less than 2 Mb.

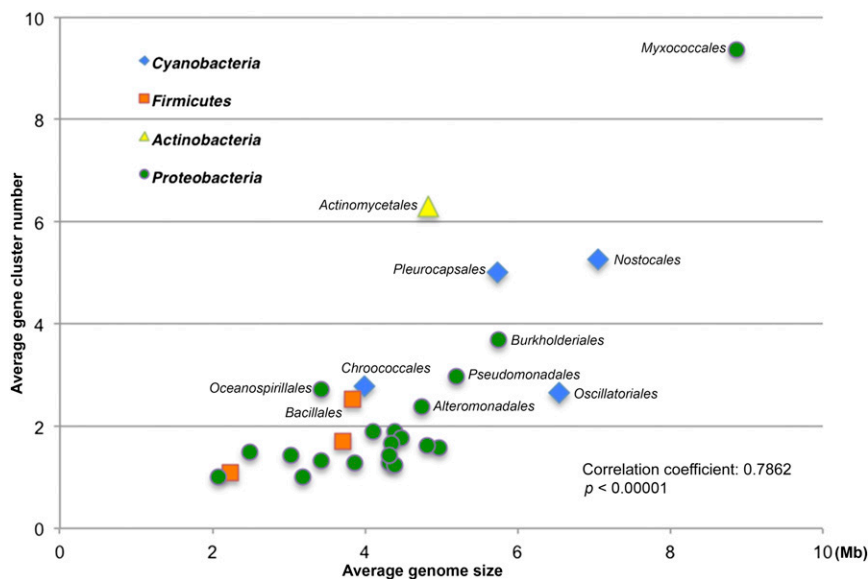
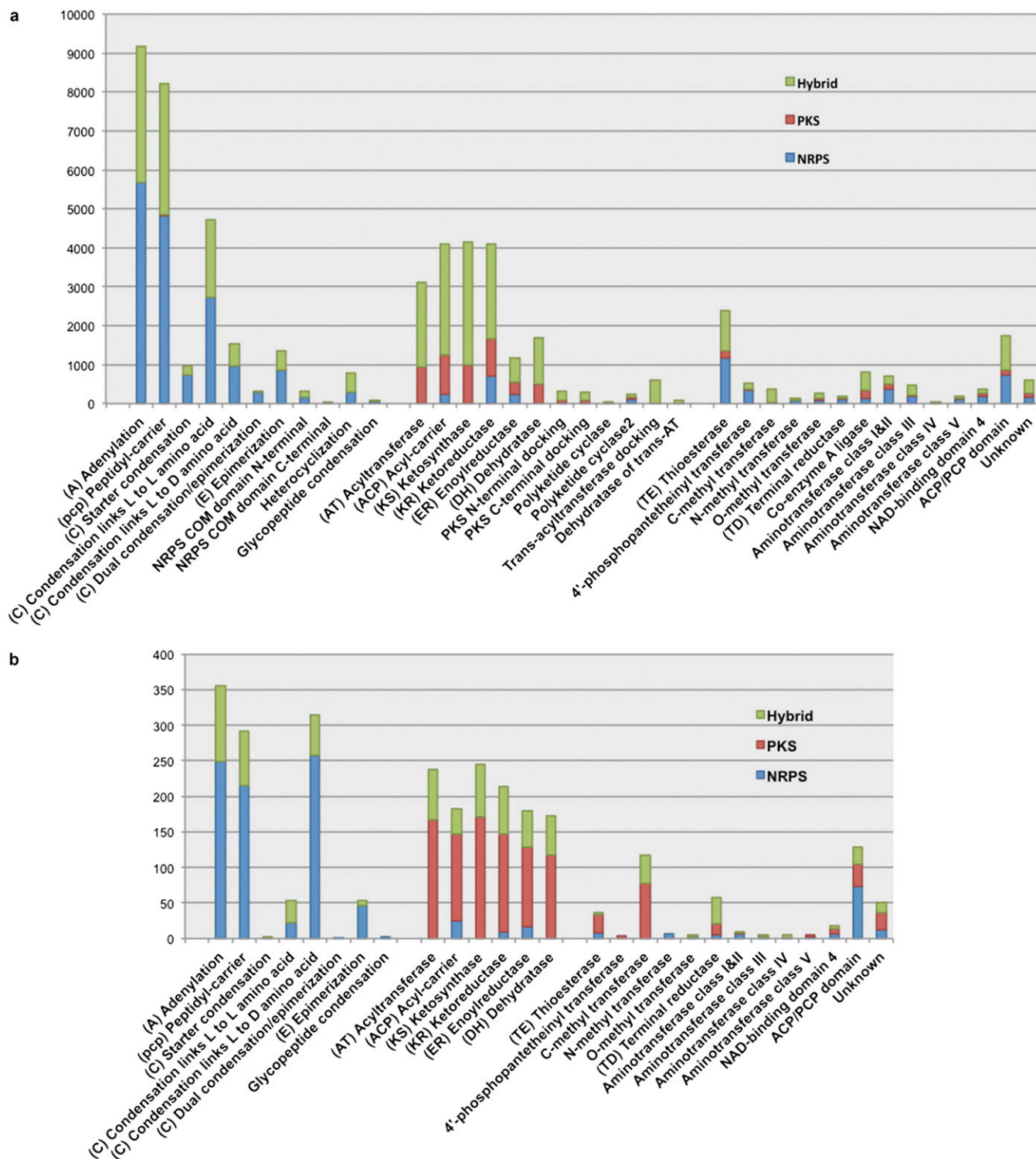


Fig. S2. Significant correlation ($P < 0.00001$) between average cluster number and average genome size in bacterial orders with enriched NRPSs and PKSs. The names are displayed for the orders whose average cluster number is over two per genome.



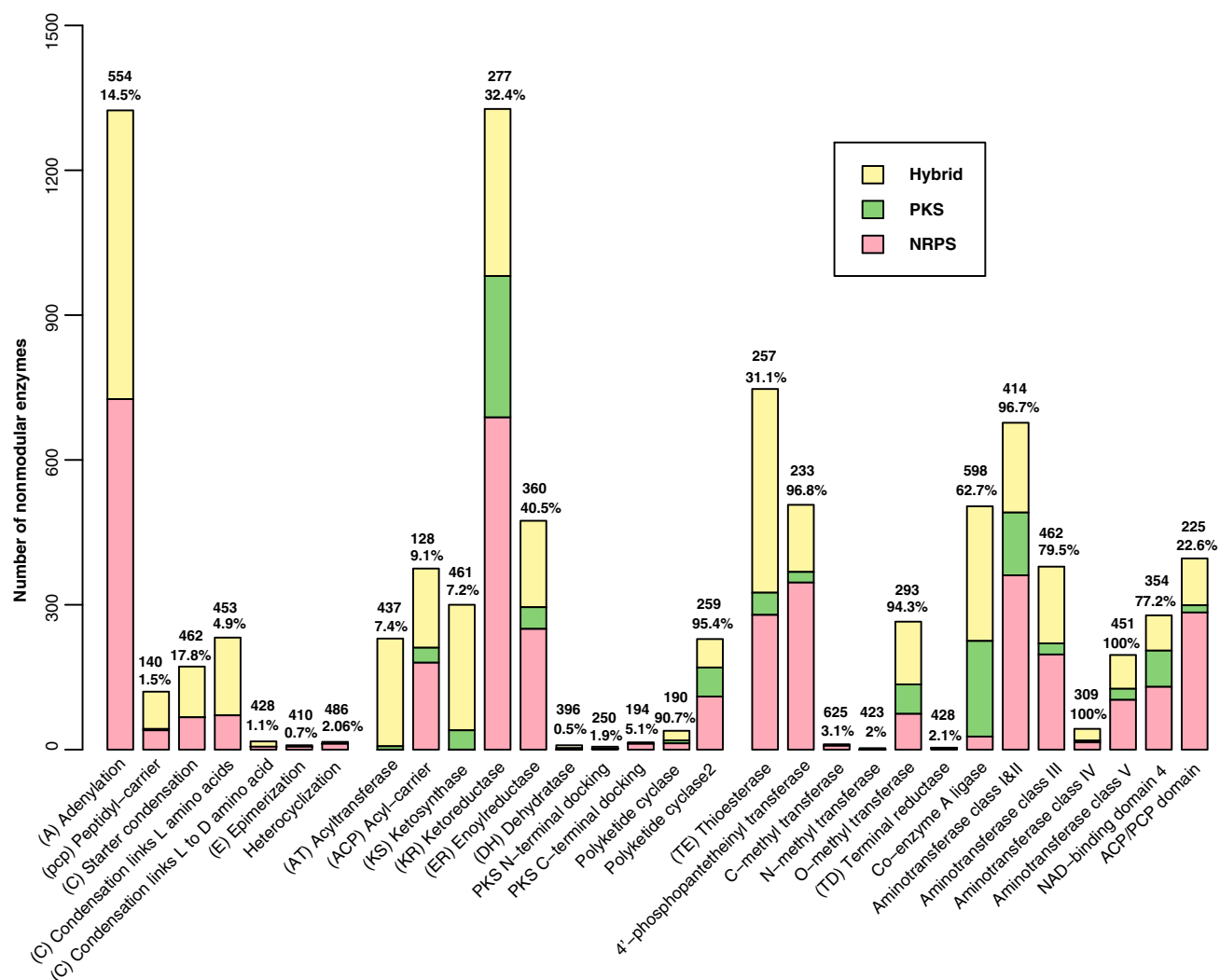


Table S1. Summary of bacterial NRPS and PKS gene clusters

Phylum	No. of strains		Avg genome size, Mb	No. of gene clusters				No. of proteins	No. of domains
	Total	With clusters (%)		Total (plasmid*)	NRPS	PKS	Hybrid		
<i>Proteobacteria</i>	1,098	519 (47%)	3.93	1,167 (59)	701	103	363	5,905	19,769
<i>Firmicutes</i>	573	178 (31%)	3.11	396 (14)	279	5	112	2,202	9,293
<i>Actinobacteria</i>	271	197 (73%)	4.44	1,225 (14)	381	319	525	6,742	23,627
<i>Tenericutes</i>	77	0 (0)	0.90	0	0	0	0	0	0
<i>Cyanobacteria</i>	72	37 (51%)	4.55	136 (14)	51	32	53	763	2,447
<i>Bacteroidetes</i>	96	13 (14%)	3.60	15 (1)	4	4	7	95	432
<i>Chlamydiae</i>	103	1 (1%)	1.23	1	1	0	0	3	14
<i>Spirochaetes</i>	57	0 (0)	2.31	0	0	0	0	0	0
<i>Deinococcus-Thermus</i>	19	1 (5%)	2.97	1	1	0	0	6	6
<i>Chloroflexi</i>	18	1 (6%)	3.37	9	3	1	5	67	284
<i>Thermotogae</i>	15	0 (0)	2.08	0	0	0	0	0	0
<i>Chlorobi</i>	11	1 (9%)	2.62	1	1	0	0	3	8
<i>Aquificae</i>	11	0 (0)	1.66	0	0	0	0	0	0
<i>Acidobacteria</i>	8	4 (50%)	5.69	7	2	2	3	33	126
<i>Fusobacteria</i>	8	0 (0)	2.31	0	0	0	0	0	0
<i>Planctomycetes</i>	7	5 (71%)	6.28	10	0	5	5	29	107
<i>Synergistetes</i>	5	0 (0)	2.14	0	0	0	0	0	0
<i>Verrucomicrobia</i>	4	1 (25%)	3.66	1	0	0	1	14	67
<i>Nitrospirae</i>	4	1 (25%)	2.82	2	1	0	1	12	51
<i>Deferribacteres</i>	4	0 (0)	2.63	0	0	0	0	0	0
<i>Ignavibacteria</i>	2	1 (50%)	3.48	1	0	0	1	4	12
<i>Elusimicrobia</i>	2	1 (50%)	1.97	2	1	1	0	6	16
<i>Thermodesulfobacteria</i>	2	0 (0)	1.98	0	0	0	0	0	0
<i>Dictyoglomi</i>	2	0 (0)	1.91	0	0	0	0	0	0
<i>Gemmatimonadetes</i>	1	1 (100%)	4.64	1	1	0	0	4	7
<i>Fibrobacteres</i>	1	1 (100%)	3.84	1	1	0	0	1	3
<i>Armatimonadetes</i>	1	0 (0)	3.44	0	0	0	0	0	0
<i>Thermobaculum</i>	1	0 (0)	3.10	0	0	0	0	0	0
<i>Chrysiogenetes</i>	1	0 (0)	2.93	0	0	0	0	0	0
<i>Caldiserica</i>	1	0 (0)	1.56	0	0	0	0	0	0
<i>Saccharobacteria</i>	1	0 (0)	1.01	0	0	0	0	0	0

*Numbers of gene clusters found in plasmids.

Table S2. Distribution of NRPS and PKS gene clusters among classes within the phylum of *Ascomycota* in fungi

Class	No. of strains		Avg genome size, Mb	No. of clusters	No. of gene clusters			No. of proteins	No. of domains
	Total	With clusters (%)			NRPS	PKS	Hybrid		
<i>Saccharomycetes</i>	18	0 (0)	11.38	0	0	0	0	0	0
<i>Sordariomycetes</i>	6	6 (100%)	38.07	127	33	64	30	215	1,143
<i>Eurotiomycetes</i>	5	5 (100%)	32.48	179	57	86	36	386	1,601
<i>Schizosaccharomycetes</i>	1	1 (100%)	12.59	1	1	0	0	1	12

Table S3. Organisms used for construction of the rRNA gene tree in Fig. 1

Domain	Phylum/Class	Organism
Bacteria	Proteobacteria	<i>Proteus mirabilis</i> HI4320
Bacteria	Firmicutes	<i>Acetobacterium woodii</i> DSM 1030
Bacteria	Actinobacteria	<i>Actinoplanes missouriensis</i> 431
Bacteria	Tenericutes	Aster yellows witches'-broom phytoplasma AYWB
Bacteria	Cyanobacteria	<i>Anabaena</i> sp. 90
Bacteria	Bacteroidetes	<i>Flavobacterium johnsoniae</i> UW101
Bacteria	Chlamydiae	<i>Parachlamydia acanthamoebae</i> UV-7
Bacteria	Spirochaetes	<i>Spirochaeta africana</i> DSM 8902
Bacteria	Deinococcus-Thermus	<i>Deinococcus maricopensis</i> DSM 21211
Bacteria	Chloroflexi	<i>Herpetosiphon aurantiacus</i> DSM 785
Bacteria	Thermotogae	<i>Marinitoga piezophila</i> KA3
Bacteria	Chlorobi	<i>Chlorobium phaeobacteroides</i> DSM 266
Bacteria	Aquificae	<i>Aquifex aeolicus</i> VF5
Bacteria	Acidobacteria	<i>Acidobacterium capsulatum</i> ATCC 51196
Bacteria	Fusobacteria	<i>Fusobacterium nucleatum</i> subsp. nucleatum ATCC 25586
Bacteria	Planctomycetes	<i>Planctomyces brasiliensis</i> DSM 5305
Bacteria	Synergistetes	<i>Thermanaerovibrio acidaminovorans</i> DSM 6589
Bacteria	Verrucomicrobia	<i>Opitutus terrae</i> PB90-1
Bacteria	Nitrospirae	<i>Candidatus Nitrospira defluvii</i>
Bacteria	Deferribacteres	<i>Deferribacter desulfuricans</i> SSM1
Bacteria	Ignavibacteria	<i>Ignavibacterium album</i> JCM 16511
Bacteria	Elusimicrobia	<i>Elusimicrobium minutum</i> Pei191
Bacteria	Thermodesulfobacteria	<i>Thermodesulfobacterium</i> sp. OPB45
Bacteria	Dictyoglomi	<i>Dictyoglomus turgidum</i> DSM 6724
Bacteria	Gemmatimonadetes	<i>Gemmatimonas aurantiaca</i> T-27
Bacteria	Fibrobacteres	<i>Fibrobacter succinogenes</i> subsp. succinogenes S85
Archaea	Thermoprotei	<i>Pyrolobus fumarii</i> 1A
Archaea	Archaeoglobi	<i>Archaeoglobus veneficus</i> SNP6
Archaea	Halobacteria	<i>Halobacterium salinarum</i> R1
Archaea	Methanobacteria	<i>Methanobrevibacter ruminantium</i> M1
Archaea	Methanococci	<i>Methanococcus aeolicus</i> Nankai-3
Archaea	Methanomicrobia	<i>Methanocella paludicola</i> SANAE
Archaea	Methanopyri	<i>Methanopyrus kandleri</i> AV19
Archaea	Thermococci	<i>Thermococcus onnurineus</i> NA1
Archaea	Thermoplasmata	<i>Thermoplasma volcanium</i> GSS1
Eukarya	Chlorophyta	<i>Chlorella variabilis</i>
Eukarya	Mollusca	<i>Lottia gigantea</i>
Eukarya	Rotifera	<i>Adineta vaga</i>
Eukarya	Ascomycota	<i>Aspergillus oryzae</i> RIB40
Eukarya	Chordata	<i>Mus musculus</i>
Eukarya	Arthropoda	<i>Drosophila melanogaster</i>
Eukarya	Nematoda	<i>Caenorhabditis elegans</i> Bristol N2
Eukarya	Annelida	<i>Capitella teleta</i>
Eukarya	Streptophyta	<i>Arabidopsis thaliana</i>
Eukarya	Apicomplexa	<i>Cryptosporidium parvum</i> Iowa II
Eukarya	Slime-molds	<i>Dictyostelium discoideum</i> AX4
Eukarya	Microsporidia	<i>Encephalitozoon cuniculi</i> GB-M1
Eukarya	Basidiomycota	<i>Cryptococcus gattii</i> WM276
Eukarya	Amoebozoa	<i>Acanthamoeba castellanii</i> Neff ATCC 50373
Eukarya	Dinoflagellata	<i>Karenia brevis</i> isolate CCMP718
Eukarya	Choanoflagellida	<i>Salpingoeca</i> sp. ATCC 50818
Eukarya	Stramenopiles	<i>Phytophthora infestans</i> T30-4

Other Supporting Information Files

Datasets S1–S4 (XLS)