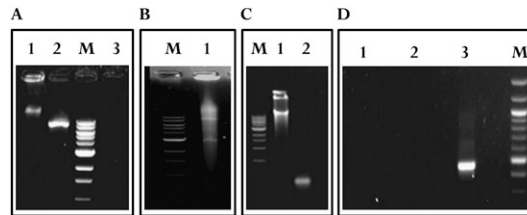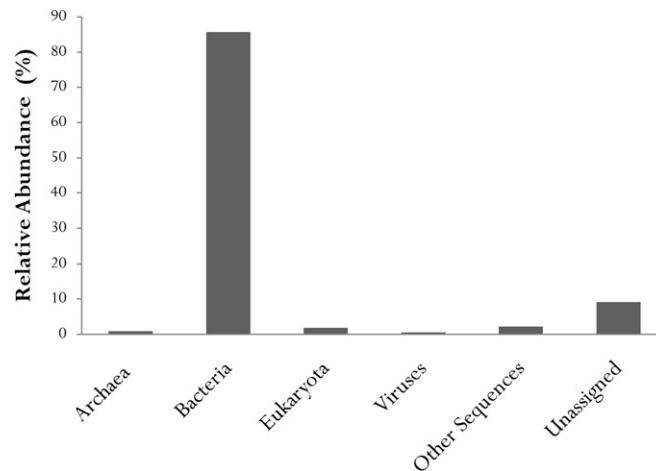# Supporting Information

Brown Kav et al. 10.1073/pnas.1116410109

Fig. S1. Phi29 polymerase amplification of plasmid DNA. (*A*) Phi29 polymerase amplifies circular DNA in a selective manner. 1, Phi29 polymerase amplification of plasmid pME6031 extracted from *Escherichia coli* TG1 (*E. coli* vector, ~8,300 bp) treated with Plasmid Safe DNase; 2, digestion of the phi29 amplification product of pME6031 plasmid with BamHI, a unique site in plasmid pME6031; M, 1-kb ladder; 3, Phi29 polymerase amplification of *E. coli* TG1 genomic DNA treated with Plasmid Safe DNase. (*B*) Purification of the plasmid population from rumen-isolated bacteria (*Materials and Methods*). M, 1-kb ladder; 1, rumen-purified plasmids. (*C*) Phi29 polymerase amplification of rumen-purified plasmids. M, 1-kb ladder; 1, Phi29 polymerase-amplified rumen plasmid DNA; 2, negative control, water as template. (*D*) 16S rRNA PCR analysis. 1, Phi29 polymerase-amplified plasmid DNA; 2, negative control (water as template); 3, rumen plasmid DNA untreated by DNase or Phi29.

Fig. S2. Domain analysis of the rumen plasmidome by MG-RAST using similarity to the SEED database with a maximum E-value of $<10^{-5}$.

Fig. S3. Taxonomic assignation of the rumen plasmidome vs. the plasmid databases. The relative abundance of each bacterial phylum is shown for the rumen plasmidome and the ACLAME (A Classification of Mobile Genetic Elements database) and IMG (Integrated Microbial Genome database) plasmid genes, as revealed by MG-RAST analysis using similarity to the SEED database with a maximum E-value of $\leq 10^{-5}$. The difference in phylum distribution was tested using χ^2 statistics and the distributions were found significantly different from one another (rumen plasmidome vs. ACLAME, $P = 9.88E\text{-}09$; rumen plasmidome vs. IMG, $P = 2.49E\text{-}05$; ACLAME vs. IMG, $P = 0.001141$).

Fig. S4. Taxonomic assignation of the rumen plasmidome. The relative abundance of each bacterial class (*A*) and order (*B*) (established for the same 16 cows) is shown for the bovine rumen plasmidome, as revealed by MG-RAST analysis using similarity to the SEED database with a maximum E-value of ≤10⁻⁵, and for the bacterial populations in the same samples as revealed by amplicon pyrosequencing of the V2 and V3 regions of 16S rRNA gene sequences. Asterisk (*) indicates significant difference at *P* < 0.05.

Fig. S5. A general scheme of BRPA (Blast Report Plasmid Aggregator). The computational analysis pipeline was designed for analysis of the rumen plasmi-dome. (*A*) First, the generated sequences are compared with five different databases using different BLAST algorithms. Next, the different hits from each of the databases are mapped on each contig, using a best-hit selection mechanism. After all hits have been mapped to contigs, the program summons contigs according to their classifications via string search for several purposes. (*B*) Here, we used BRPA to select contigs for a downstream application of contigs that carry both a plasmid backbone gene and an accessory gene in a certain functional category.

Table S1. Summary of the assembly data obtained from two different assembly methods using the total 33,874,912 reads generated

Assembly data	Velvet	SOAP de novo
Total number of contigs	13,018	5,771
Mean contig length (bp)	182	469
Longest contig length (bp)	4661	7779
Number of reads in contigs	12,344,320	12,457,888
Percentage of reads in contigs (%)	36.4	36.8

Table S2. χ^2 statistic P values for the comparative functional analysis of the rumen plasmidome vs. rumen metagenomes and plasmid databases (Fig. 2)

SEED subsystem	Brulc et al. (1)	Hess et al. (2)	ACLAME	IMG
Amino acids	NS	NS	1.79E-05	4.51E-05
Carbohydrates	NS	NS	NS	0.016049
Cell division and cell cycle	NS	NS	NS	NS
Cell wall and capsule	NS	NS	0.00017	0.000212
Cofactors, vitamins, etc.	NS	NS	0.000583	0.000212
DNA metabolism	0.046556	NS	0.002345	0.000342
Fatty acids and lipids	NS	NS	NS	NS
Membrane transport	0.040827	NS	NS	3.64E-05
Aromatic compounds	NS	NS	NS	NS
Motility and chemotaxis	NS	NS	NS	0.018521
Nitrogen metabolism	NS	NS	NS	NS
Nucleosides and nucleotides	NS	NS	NS	NS
Mobile genetic elements	NS	NS	1.48E-14	0.019629
Protein metabolism	NS	NS	3.98E-08	1.59E-06
RNA metabolism	NS	NS	NS	NS
Regulation and cell signaling	NS	NS	NS	NS
Respiration	NS	NS	NS	0.035336
Stress response	NS	NS	NS	NS
Sulfur metabolism	NS	NS	NS	NS
Virulence	NS	NS	NS	NS

Differences in the abundances of the each SEED subsystem compared with the rumen plasmidome were tested using χ^2 statistics. The P values for differences of each subsystem between the rumen plasmidome and the other datasets are mentioned. NS, not significant.

1. Brulc JM, et al. (2009) Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc Natl Acad Sci USA* 106:1948–1953.
2. Hess M, et al. (2011) Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* 331:463–467.

Table S3. Proportion of annotated contigs

Database	CDD	Protein NR	Nucleotide NR	ACLAME (only genes)	IMG (only plasmid genes)
Percentage of contigs with annotation	22.3%	28.7%	3.8%	23.3%	26%

The percentage of SOAP de novo contigs with annotations in each of the five databases: ACLAME, A Classification of Mobile Genetic Elements; CDD, Conserved Domain database; IMG, Integrated Microbial Genome database; Nucleotide NR, National Center for Biotechnology Information (NCBI) nucleotide nonredundant; Protein NR, NCBI protein nonredundant.

Table S4. Previously isolated plasmids from terrestrial aquatic and rumen environment

Source	Bacterial species	Plasmid	Reference
Rumen-isolated bacteria	*Selenomonas ruminantium*	pONE430	GI32455406
	Selenomonas ruminantium	pONE429	GI32455408
	Selenomonas ruminantium	pJW1	GI4809199
	Selenomonas ruminantium	pSR1	GI 5739173
	Selenomonas ruminantium	pSRD191	GI50659892
	Selenomonas ruminantium	pJDB23	Attwood and Brooker (1)
	Prevotella ruminicola	pRAM4	GI10956058
	Lactobacillus plantarum	pLD1	GI225166856
	Butyrivibrio fibrisolvens	pRJF1	GI209808771
	Butyrivibrio fibrisolvens	pRJF2	GI209808768
	Butyrivibrio fibrisolvens	pOM1	GI468363
	Ruminococcus flavefaciens	pBAW301	GI:726294
Terrestrial-isolated bacteria	*Ralstonia solanacearum*	pJTPS1	GI10954629
	Ralstonia solanacearum	pRSI13	GI300698618
	Burkholderia cenocepacia	pBCJ2315	GI206479926
	Pseudomonas putida	pDTG1	GI32469878
	Pseudomonas putida	pYQ39	GI32469969
	Pseudomonas putida	pPP81	GI10954624
	Bacillus subtilis	p1414	GI10956512
	Bacillus subtilis	pTA1015	GI10956497
	Bacillus subtilis	pTA1040	GI10956489
	Bacillus subtilis	pTA1060	GI10956503
	Bacillus licheniformis	pFL7	GI40445271
	Bacillus licheniformis	pFL5	GI40804652
	Corynebacterium diphtheriae	pNG2	GI32470491
	Staphylococcus aureus	pUB110	GI150521
Aquatic-isolated bacteria	*Synechococcus* sp.	pAQ1	GI170079460
	Synechococcus sp.	pAQ3	GI170079464
	Synechococcus elongatus	pUH24	GI32455412
	Eromonas salmonicida	pAsa3	GI32186818
	Eromonas salmonicida	pAsa2	GI32186830
	Eromonas salmonicida	pAsa1	GI32186809
	Vibrio fischeri	pES213	GI42794283
	Vibrio sp.	pPS41	GI5817769
	Methylophaga thalassica	pMTS1	GI10956807
	Vibrio vulnificus	pMP1	GI82791436
	Synechococcus sp.	pSY10	GI216847
	Methylophaga thalassica	pMTS1	GI10956807
	Shewanella sp.	pSFKW33	GI222692192
	Leptolyngbya sp.	pRF1	GI32455760

The sequences of plasmids previously isolated from three different environments were downloaded from the NCBI nucleotide database. Three different datasets were generated: rumen-isolated plasmids, terrestrial-isolated plasmids and aquatic-isolated plasmids (Table S5). Each of the datasets were compared with the rumen plasmidome using TBLASTX with a maximum E-value of $\leq 10^{-3}$. When hits overlapped, the best hit (lowest E-value) was selected. Finally, the number of contigs of the rumen plasmidome which were hit by the different plasmids in the dataset was counted.

1. Attwood GT, Brooker JD (1992) Complete nucleotide sequence of a Selenomonas ruminantium plasmid and definition of a region necessary for its replication in Escherichia coli. *Plasmid* 28(2):123–129.

Table S5. Number of rumen plasmidome contigs with hits in the different plasmid datasets

	Aquatic	Terrestrial	Rumen
Number of contigs with hits from each dataset	20*	59*	338*

See legend for Table S4. The difference in the number of hits for each contig was measured using χ^2 statistics and was found to be significantly different. Asterisk (*) indicates significant difference at $P < 0.05$.

Table S6. Mosaic nature of contigs and the phylogenetic association of their ORFs

Contig	Mosaic nature	ORF 1	ORF 2	ORF 3	ORF 4
C48430	Nonmosaic	Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; *Bacteroides*	Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; *Bacteroides*		
Scaffold883	Nonmosaic	Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; *Bacteroides*	Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; *Bacteroides*		
C46542	Nonmosaic	Bacteroidetes; Bacteroidia; Bacteroidales; Prevotellaceae; Prevotella	Bacteroidetes; Bacteroidia; Bacteroidales; Prevotellaceae; Prevotella		
C48422	Nonmosaic	Bacteroidetes; Bacteroidia; Flavobacteriia; Flavobacteriales; Flavobacteriaceae	Bacteroidetes; Bacteroidia; Flavobacteriia; Flavobacteriales; Flavobacteriaceae		
Scaffold737	Nonmosaic	Firmicutes; Bacilli; Lactobacillales; Enterococcaceae; *Enterococcus*	Firmicutes; Bacilli; Lactobacillales; Enterococcaceae; *Enterococcus*		
Scaffold19	Nonmosaic	Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; Butyrivibrio	Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; Butyrivibrio		
C48342	Nonmosaic	Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Ruminococcus	Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Ruminococcus		
Scaffold 44	Nonmosaic	Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Ruminococcus	Firmicutes; Clostridia; Clostridiales; Ruminococcaceae; Ruminococcus		
Scaffold755	Nonmosaic	Proteobacteria; Gammaproteobacteria; Pseudomonadales; Moraxellaceae; *Acinetobacter*	Proteobacteria; Gammaproteobacteria; Pseudomonadales; Moraxellaceae; *Acinetobacter*		
Scaffold521	Nonmosaic	Spirochaetes; Spirochaetia; Spirochaetales; Spirochaetaceae; *Treponema*	Spirochaetes; Spirochaetia; Spirochaetales; Spirochaetaceae; *Treponema*		
Scaffold326	Genus	Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; Oribacterium	Firmicutes; Clostridia; Clostridiales; Lachnospiraceae; Butyrivibrio		
C47816	Family	Bacteroidetes; Bacteroidia; Bacteroidales; Prevotellaceae	Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae		
C48410	Family	Firmicutes; Bacilli; Bacillales; Bacillaceae	Firmicutes; Bacilli; Bacillales; Paenibacillaceae		
C48346	Family	Firmicutes; Clostridia; Clostridiales; Clostridiaceae	Firmicutes; Clostridia; Clostridiales; Lachnospiraceae		
Scaffold352	Family	Firmicutes; Clostridia; Clostridiales; Clostridiaceae	Firmicutes; Clostridia; Clostridiales; Lachnospiraceae	Firmicutes; Clostridia; Clostridiales; Lachnospiraceae	
Scaffold654	Family	Firmicutes; Clostridia; Clostridiales; Clostridiaceae	Firmicutes; Clostridia; Clostridiales; Ruminococcaceae		
C48428	Family	Firmicutes; Clostridia; Clostridiales; Eubacteriaceae	Firmicutes; Clostridia; Clostridiales; Clostridiaceae		
Scaffold593	Family	Firmicutes; Clostridia; Clostridiales; Ruminococcaceae	Firmicutes; Clostridia; Clostridiales; Lachnospiraceae		
C48424	Family	Firmicutes; Clostridia; Clostridiales; Ruminococcaceae	Firmicutes; Clostridia; Clostridiales; Lachnospiraceae		
C47544	Family	Firmicutes; Clostridia; Clostridiales; Ruminococcaceae	Firmicutes; Clostridia; Clostridiales; Lachnospiraceae		
Scaffold792	Order	Bacteroidetes; Bacteroidia; Bacteroidales	Bacteroidetes; Bacteroidia; Flavobacteriia		
Scaffold1084	Order	Bacteroidetes; Bacteroidia; Bacteroidales	Bacteroidetes; Bacteroidia; Flavobacteriia		
C48408	Order	Bacteroidetes; Bacteroidia; Bacteroidales	Bacteroidetes; Bacteroidia; Flavobacteriia		
Scaffold628	Order	Bacteroidetes; Bacteroidia; Bacteroidales	Bacteroidetes; Bacteroidia; Flavobacteriia		
C48448	Order	Bacteroidetes; Bacteroidia; Bacteroidales	Bacteroidetes; Bacteroidia; Flavobacteriia		
Scaffold252	Order	Bacteroidetes; Bacteroidia; Bacteroidales	Bacteroidetes; Bacteroidia; Bacteroidales	Bacteroidetes; Bacteroidia; Cytophagia	
Scaffold814	Order	Bacteroidetes; Bacteroidia; Bacteroidales	Bacteroidetes; Bacteroidia; Flavobacteriia		

Contig	Mosaic nature	ORF 1	ORF 2	ORF 3	ORF 4
Scaffold651	Order	Bacteroidetes; Bacteroidia; Bacteroidales	Bacteroidetes; Bacteroidia; Flavobacteriia		
C48060	Order	Bacteroidetes; Bacteroidia; Bacteroidales	Bacteroidetes; Bacteroidia; Flavobacteriia		
Scaffold657	Order	Bacteroidetes; Bacteroidia; Flavobacteriia	Bacteroidetes; Bacteroidia; Bacteroidales		
Scaffold565	Order	Bacteroidetes; Bacteroidia; Flavobacteriia	Bacteroidetes; Bacteroidia; Bacteroidales		
Scaffold1048	Order	Bacteroidetes; Bacteroidia; Flavobacteriia	Bacteroidetes; Bacteroidia; Bacteroidales		
Scaffold306	Order	Firmicutes; Bacilli; Lactobacillales	Firmicutes; Clostridia; Clostridiales		
Scaffold584	Order	Firmicutes; Bacilli; Lactobacillales	Firmicutes; Bacilli; Bacillales		
caffold504	Order	Firmicutes; Bacilli; Lactobacillales	Firmicutes; Bacilli; Bacillales		
Scaffold279	Order	Firmicutes; Bacilli; Lactobacillales	Firmicutes; Bacilli; Bacillales		
Scaffold617	Order	Firmicutes; Clostridia; Clostridiales	Firmicutes; Clostridia; Thermoanaerobacterales		
Scaffold415	Order	Proteobacteria; Epsilonproteobacteria; Campylobacterales	Proteobacteria; Epsilonproteobacteria; Nautiliales		
C48432	Class	Firmicutes; Bacilli	Firmicutes; Clostridia		
Scaffold121	Class	Firmicutes; Bacilli	Firmicutes; Clostridia		
Scaffold799	Class	Firmicutes; Bacilli	Firmicutes; Clostridia		
C48416	Class	Firmicutes; Bacilli	Firmicutes; Clostridia		
C48434	Class	Firmicutes; Bacilli	Firmicutes; Clostridia		
Scaffold66	Class	Firmicutes; Bacilli	Firmicutes; Clostridia		
Scaffold155	Class	Firmicutes; Bacilli	Firmicutes; Clostridia		
Scaffold128	Class	Firmicutes; Bacilli	Firmicutes; Clostridia		
C48366	Class	Firmicutes; Clostridia	Firmicutes; Bacilli		
Scaffold559	Class	Firmicutes; Clostridia	Firmicutes; Erysipelotrichi		
Scaffold1082	Class	Firmicutes; Clostridia	Firmicutes; Erysipelotrichi		
Scaffold381	Class	Firmicutes; Clostridia	Firmicutes; Bacilli		
C48234	Class	Firmicutes; Clostridia	Firmicutes; Bacilli;		
Scaffold294	Class	Firmicutes; Clostridia	Firmicutes; Bacilli		
C48384	Class	Firmicutes; Clostridia	Firmicutes; Negativicutes		
C48404	Class	Firmicutes; Clostridia	Firmicutes; Bacilli		
Scaffold562	Class	Firmicutes; Clostridia	Firmicutes; Bacilli	Firmicutes; Clostridia	Firmicutes; Clostridia
Scaffold713	Class	Firmicutes; Clostridia	Firmicutes; Negativicutes		
Scaffold695	Class	Firmicutes; Clostridia	Firmicutes; Bacilli		
C48344	Class	Firmicutes; Clostridia	Firmicutes; Bacilli		
C48286	Class	Firmicutes; Clostridia	Firmicutes; Bacilli		
Scaffold406	Class	Firmicutes; Clostridia	Firmicutes; Bacilli		
C48438	Class	Firmicutes; Clostridia	Firmicutes; Bacilli	Firmicutes; Bacilli	
Scaffold64	Class	Firmicutes; Clostridia	Firmicutes; Clostridia	Firmicutes; Erysipelotrichi	
Scaffold269	Class	Firmicutes; Clostridia	Firmicutes; Clostridia	Firmicutes; Bacilli	
Scaffold782	Class	Firmicutes; Clostridia	Firmicutes; Bacilli		
Scaffold980	Class	Firmicutes; Clostridia	Firmicutes; Bacilli		
Scaffold181	Class	Firmicutes; Clostridia	Firmicutes; Bacilli		
Scaffold335	Class	Firmicutes; Negativicutes	Firmicutes; Bacilli		
caffold1034	Class	Firmicutes; Negativicutes	Firmicutes; Bacilli		
Scaffold206	Class	Firmicutes; Negativicutes	Firmicutes; Clostridia		
Scaffold453	Class	Firmicutes; Negativicutes	Firmicutes; Clostridia		
C48394	Class	Firmicutes; Negativicutes	Firmicutes; Bacilli		
Scaffold413	Class	Proteobacteria; Alphaproteobacteria	Proteobacteria; Gammaproteobacteria		
C48392	Phylum	Actinobacteria	Firmicutes		
caffold1076	Phylum	Actinobacteria	Firmicutes		
C48360	Phylum	Actinobacteria	Firmicutes	Firmicutes	
Scaffold900	Phylum	Actinobacteria	Proteobacteria		
C48340	Phylum	Actinobacteria	Firmicutes		
C48316	Phylum	Actinobacteria	Firmicutes		
Scaffold166	Phylum	Actinobacteria	Firmicutes		

Table S6.   Cont.

Contig	Mosaic nature	ORF 1	ORF 2	ORF 3	ORF 4
Scaffold463	Phylum	Firmicutes	Actinobacteria		
Scaffold533	Phylum	Firmicutes	Actinobacteria	Firmicutes	
Scaffold433	Phylum	Firmicutes	Bacteroidetes		
Scaffold600	Phylum	Proteobacteria	Firmicutes		
Scaffold721	Phylum	Spirochaetes	Proteobacteria		

The mosaic nature of each contig was determined by the agreement in taxonomic assignment between its ORFs. The lowest level at which the ORFs on each contig did not exhibit the same phylogenetic affinity, and the taxonomic lineage up to this level, are shown for each of the contigs.

Table S7.   Interactions between five major phyla that were detected in mosaic contigs

Phyla	Firmicutes	Proteobacteria	Spirochaetes	Actinobacteria	Bacteroidetes
Firmicutes		*		*	*
Proteobacteria			*	*	*
Spirochaetes					*
Actinobacteria					
Bacteroidetes					

Boxes with asterisks indicate the existence of ORFs from these phyla residing on the same contig. Shading indicates the same type of interaction as described in the white boxes.