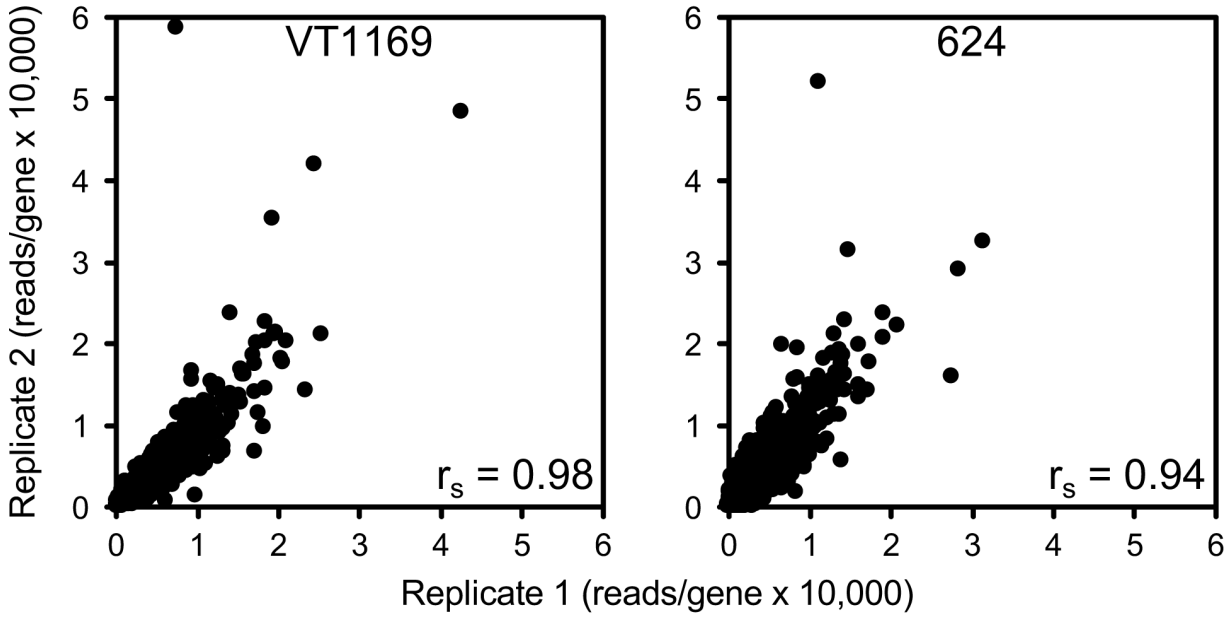

Figure S1. Correlation between Tn-seq replicates for determining the *A. actinomycetemcomitans* essential genome. Each point corresponds to a gene in VT1169 (left) or 624 (right). The normalized reads/gene are plotted for replicates 1 and 2 on the X and Y axis, respectively. r_s, Spearman's rank correlation coefficient.

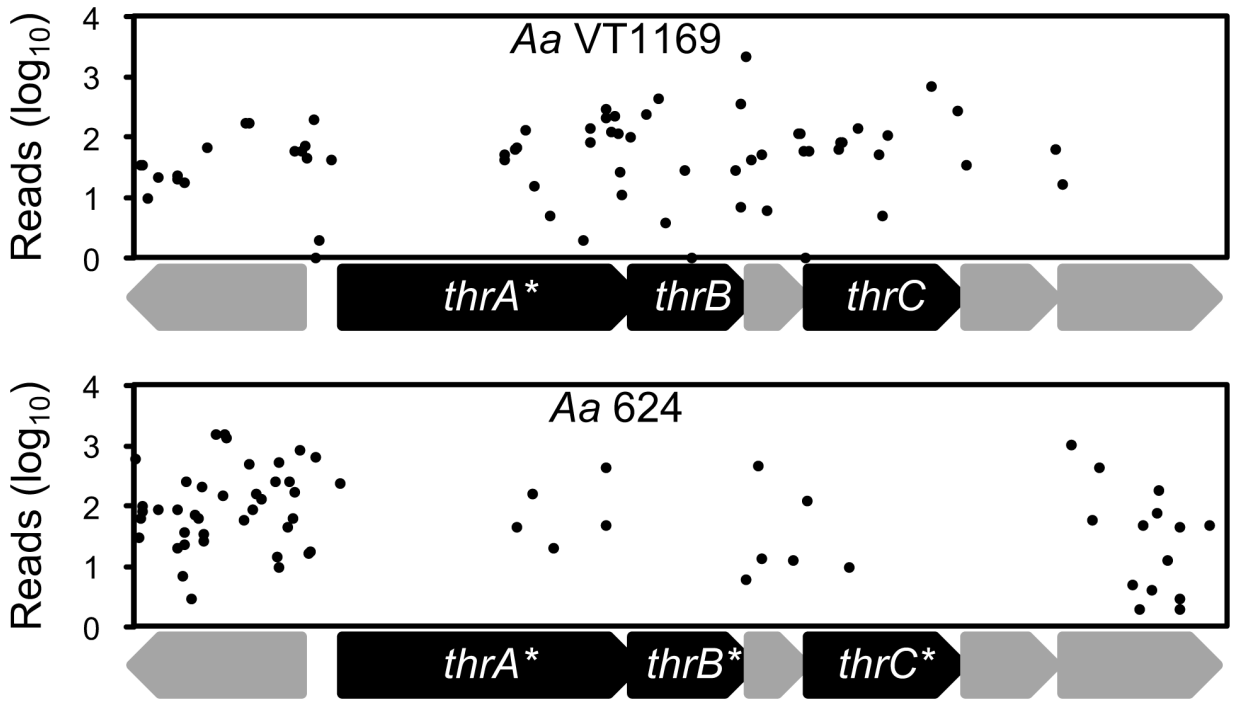

Figure S2. Example of *A. actinomycetemcomitans* strain-specific gene essentiality. Each point corresponds to a transposon insertion in VT1169 (top) or 624 (bottom). X axis, genome position. Y axis, $\log_{10}$ read count. Black, threonine biosynthesis genes. * indicates less than expected (DESeq2 adjusted $p < 0.01$).

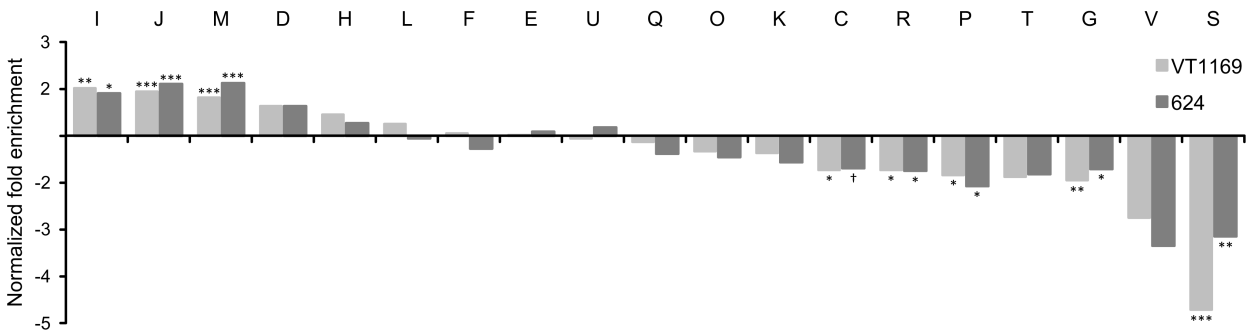

Figure S3. Functional enrichment in the *A. actinomycetemcomitans* essential genome. Overlying data labels indicate clusters of orthologous groups (COGs). Normalized fold enrichment (Y axis) is the percent of the essential genome made up by a COG divided by the percent of the whole genome made up by the same COG. The Y axis is adjusted such that 2-fold higher than the genome is +2, and 2-fold lower than the genome is -2. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$ (two-tailed Fisher's exact test). I, Lipid transport and metabolism. J, Translation, ribosomal structure, and biogenesis. M, Cell wall/membrane/envelope biogenesis. D, Cell cycle control, cell division, and chromosome partitioning. H, Coenzyme transport and metabolism. L, Replication, recombination, and repair. F, Nucleotide transport and metabolism. E, Amino acid transport and metabolism. U, Intracellular trafficking, secretion, and vesicular transport. Q, Secondary metabolite biosynthesis, transport, and catabolism. O Post-translational modification, protein turnover, and chaperones. K, Transcription. C, Energy production and conversion. R, General function prediction only. P, Inorganic ion transport and metabolism. S, Signal transduction mechanisms. G, Carbohydrate transport and metabolism. V, Defense mechanisms. S, Function unknown.

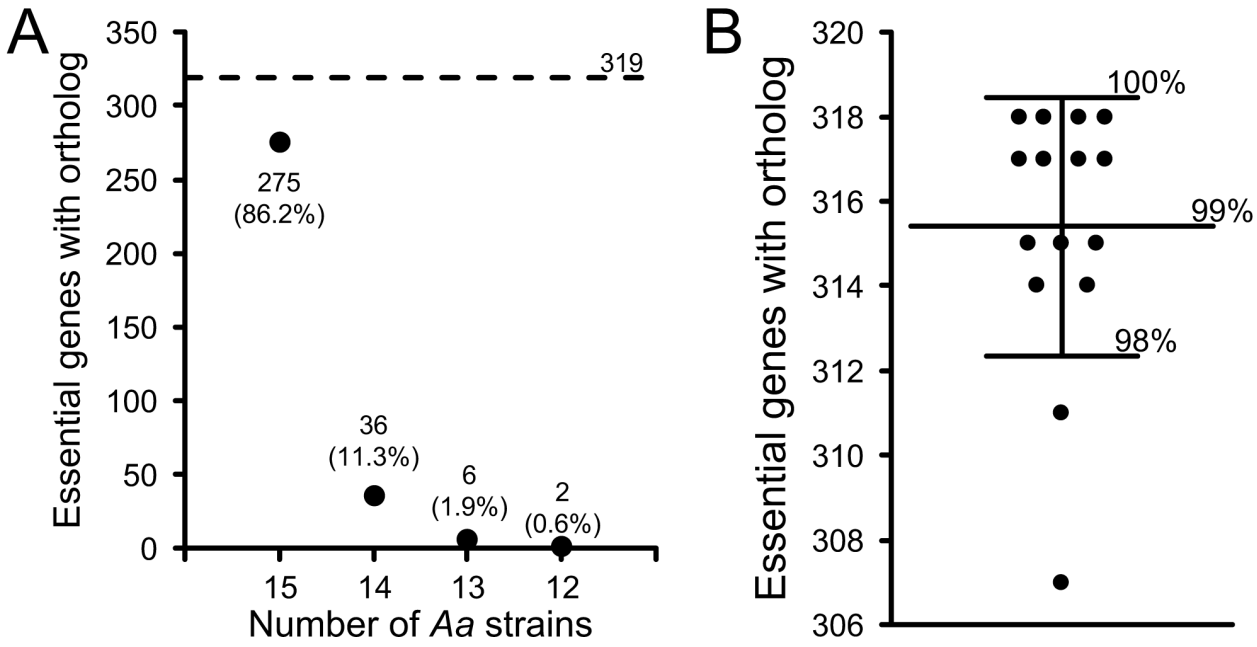

Figure S4. Conservation of the *A. actinomycetemcomitans* essential genome among strains. (A) Y axis, number of core essential genes in VT1169 with an ortholog in the indicated number of other *A. actinomycetemcomitans* strains (X axis). Dotted line, total core essential genes in VT1169. Data labels, number of orthologs (percentages are relative to 319). (B) Y axis, number of core essential genes in VT1169 with an ortholog in another individual strain. Each point represents an individual strain of *A. actinomycetemcomitans*. The percentages (relative to 319) indicate the average ± standard deviation. (A) and (B) were essentially the same for 624 (see Dataset S1).

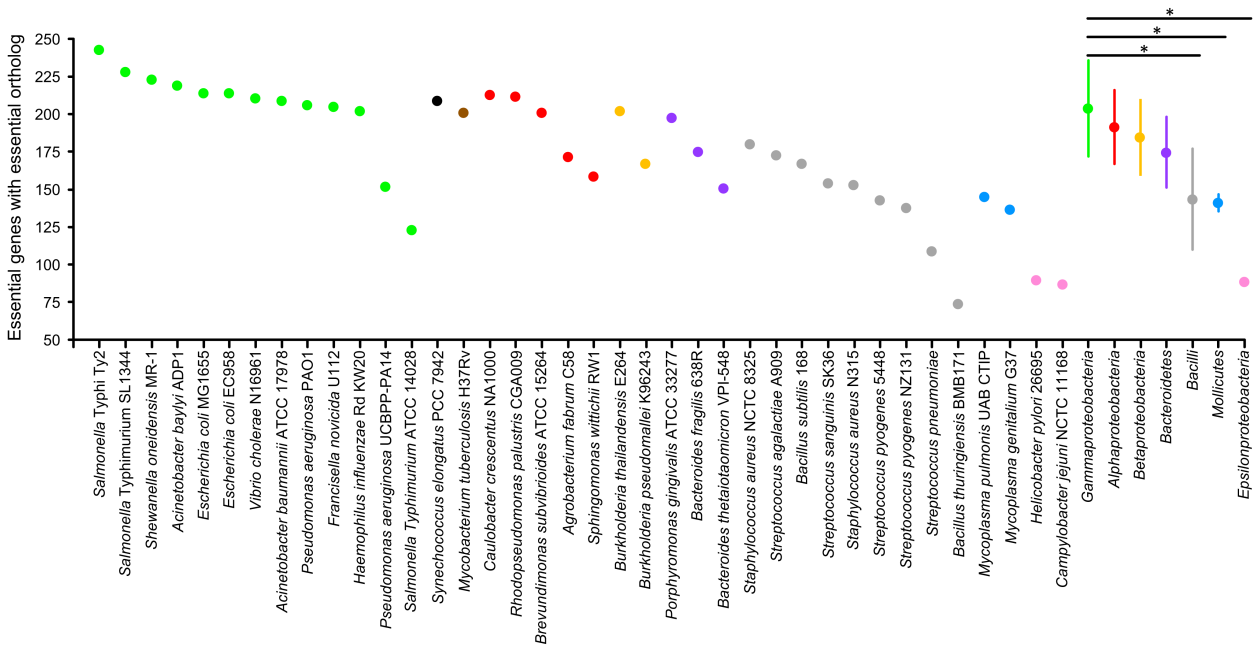

Figure S5. Conservation of the core *A. actinomycetemcomitans* essential genome among other species. Y axis, number of core essential genes in VT1169 with an ortholog in the essential genome of the indicated species (X axis). The corresponding numbers in 624 varied at most from VT1169 by 1 (see Dataset S1). Each species is colored according to its phylogenetic class. The average for each class is shown on the right. The *Cyanobacteria* and *Actinobacteria* are not shown since they are represented by single species (*S. elongatus* and *M. tuberculosis*, respectively). Error bars represent standard deviation (n = 2-13). *, $p < 0.001$ (two-tailed Student's *t* test).

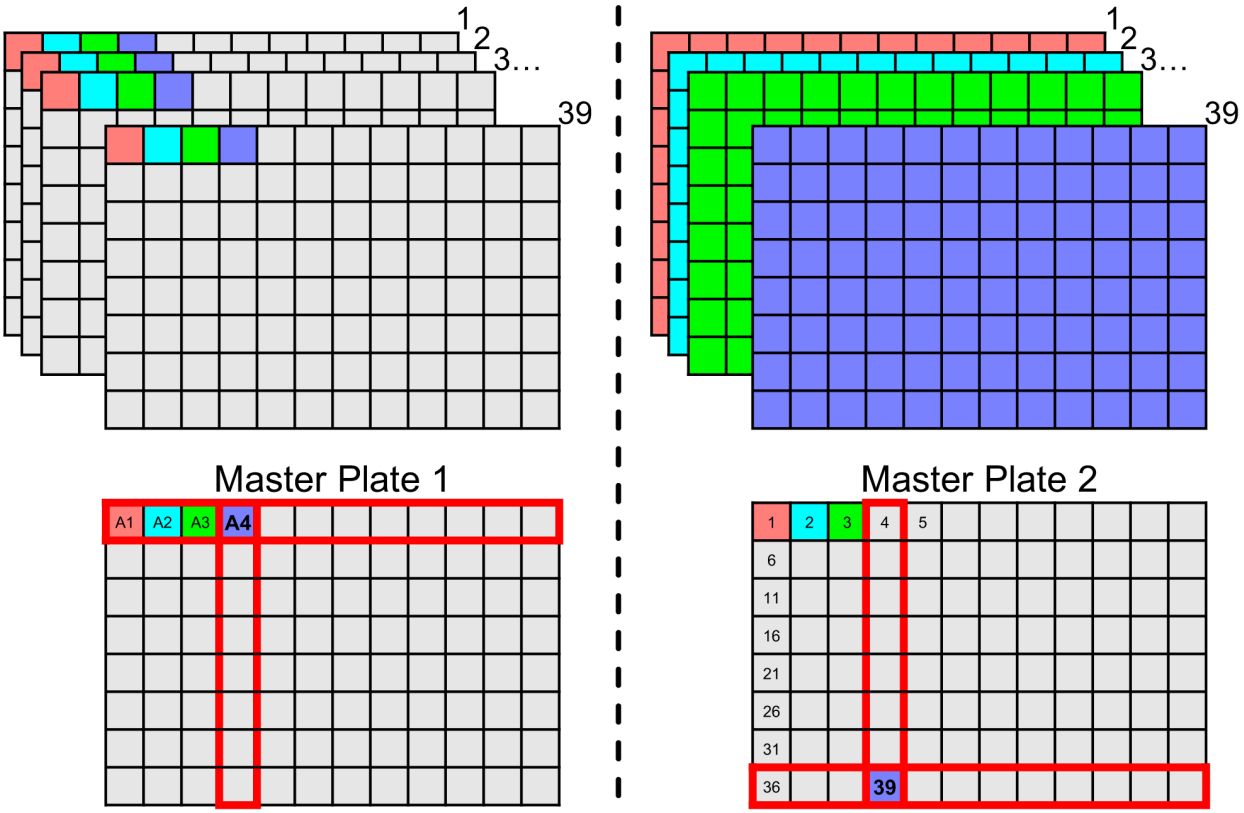

Figure S6. The Cartesian Pooling-Coordinate Sequencing method. In the CP-CSeq method, 2 master plates (MP) are made. Each well of each plate goes into the corresponding well in MP1 (left), whereas all of the wells per plate each go into a separate well in MP2 (right). Then, Tn-seq is done on each row and column of each MP. The *A. actinomycetemcomitans* ordered library comprises 39 plates, so it required 33 Tn-seq reactions (8 MP1 rows + 12 MP1 columns + 8 MP2 rows + 5 MP2 columns). In the example shown, the mutant in well A4 of plate 39 can be identified by finding the intersection between row A and column 4 in MP1 and row H and column 4 in MP2.

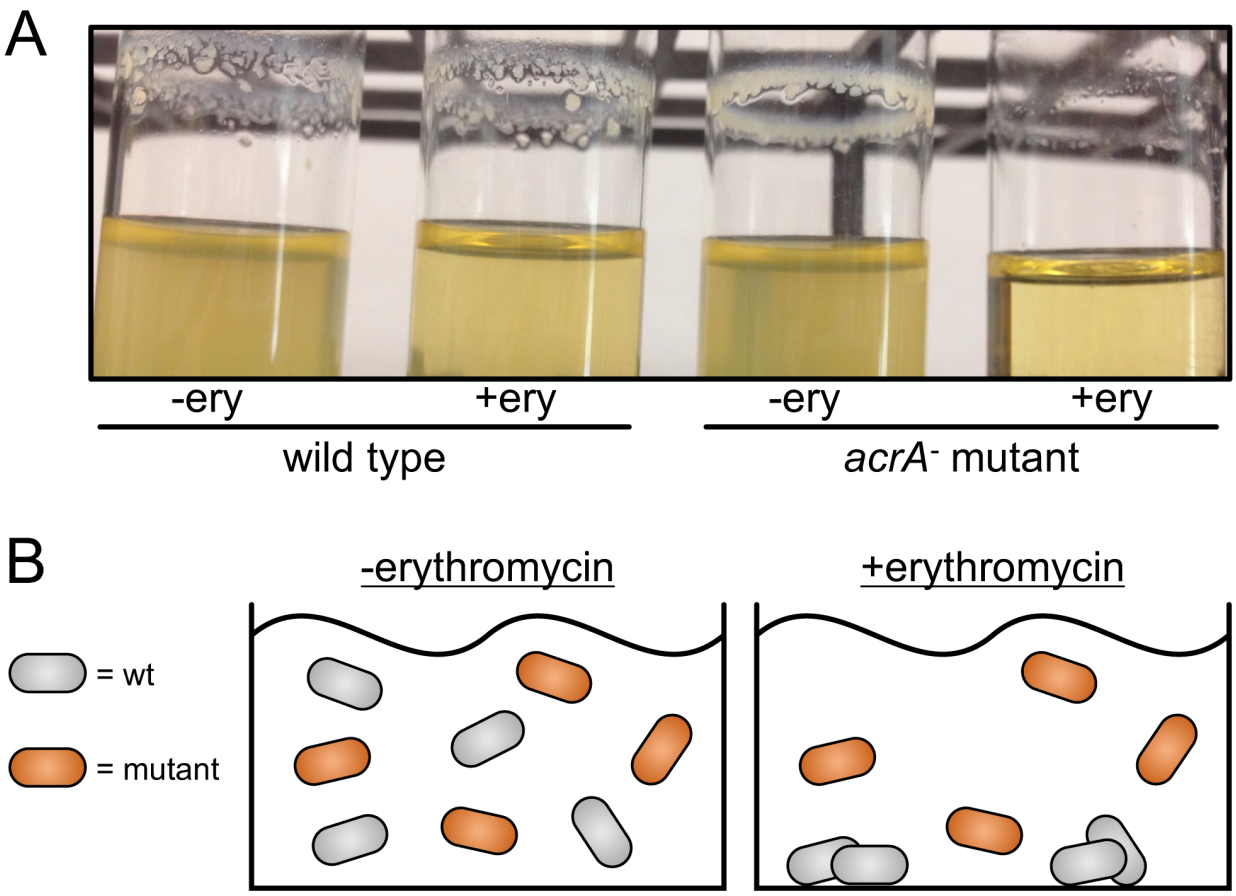

Figure S7. Factors that mediate *A. actinomycetemcomitans* erythromycin resistance. (A) The wild type and hyper-sensitive *acrA*$^-$ mutant were grown with half-MIC erythromycin (1 µg/ml, +ery) or the solvent control (same volume of pure ethanol, -ery). (B) Screen for mutants that do not attach in response to erythromycin. Attachment-defective mutants become enriched in the liquid phase specifically in the presence of antibiotic (right). wt, wild type.

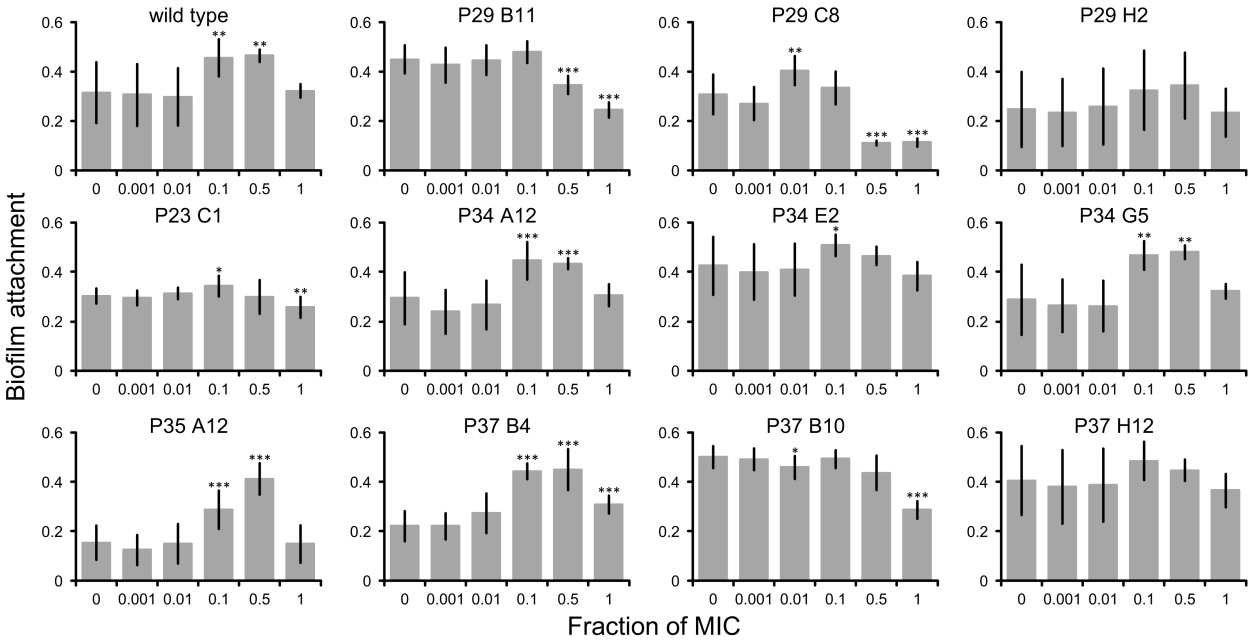

Figure S8. Factors that mediate *A. actinomycetemcomitans* erythromycin-induced attachment. Biofilm attachment was measured using crystal violet. X axis, fraction of minimum inhibitory concentration (MIC) for erythromycin. Y axis, absorbance (A_{620}) reading for crystal violet bound to biofilms. Overlying data labels indicate the plate and well in the ordered library that was the source of the mutant (see Figure 3 for gene products). Error bars represent standard deviation (n = 11-16). Significance was determined for each concentration compared to no erythromycin addition. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$ (two-tailed Student's *t* test).

Table S1. Summary of Tn-seq analysis for essential genome.

Sample	Rep	Total reads	With IR[a]	Processed[b]	Mapped[c]	Sites[d]	No slip[e]	In both[f]
VT1169	1	13,574,870	11,531,485 (85%)	10,480,596 (91%)	9,084,230 (87%)	44,329	43,564	32,344
	2	9,893,314	8,881,261 (90%)	8,649,894 (97%)	7,711,040 (89%)	38,553	37,748	
624	1	15,346,282	11,862,440 (77%)	10,189,279 (86%)	3,921,979 (38%)	39,000	38,643	26,530
	2	22,724,657	18,817,125 (83%)	16,916,684 (90%)	12,801,472 (76%)	37,325	36,273	
no antibiotic	-	7,637,188	6,919,034 (91%)	6,452,821 (93%)	5,942,808 (92%)	2,634	2,013	-
erythromycin	-	8,035,917	6,743,078 (84%)	5,740,553 (85%)	5,202,209 (91%)	2,516	1,949	-

Percentages are relative to the value in the adjacent left column.

[a]Reads containing the transposon inverted repeat (IR) sequence (CCAACCTGTTA) with 0 mismatches in its expected location (IR end position at base 38-40) within the read

[b]Reads remaining after trimming 3' low-quality bases (phred score + 33 < 20), trimming poly-C tails, and removing reads <20 bases long

[c]Reads mapping end-to-end to the *A. actinomycetemcomitans* genome with high quality (MAPQ > 39)

[d]Sites identified after accounting for the 2 bp duplication associated with mariner insertion events (by adding 2 to the genome position of reads mapping to the negative strand)

[e]Sites remaining after correcting for polymerase slippage by collapsing adjacent sites onto the site with the highest read count (local maximum)

[f]Sites present in both replicates (1 aliquot of the mutant pool before amplification, 1 after amplification)

Table S2. Smoothing to correct for distance from origin.

Sample	Replicate	Max:min*
VT1169	1	1.24
	2	1.97
624	1	1.15
	2	1.27
no antibiotic	-	1.20
erythromycin	-	1.17

*Fold difference between the maximum and minimum values of the LOESS regression for insertion abundance vs. location, indicating how much insertions close to the origin were inflated relative to those close to the terminus prior to smoothing

Table S3. Gene essentiality in *A. actinomycetemcomitans* VT1169 and 624.

Strain	Total genes	Essential genes[a]	Essential orthologs in both strains[b]	Essential orthologs in 1 strain only[b]	Essential accessory genes[b]
VT1169	1,912	413 (22%)	319 (77%)	86 (21%)	8 (2%)
624	2,164	386 (18%)	319 (83%)	54 (14%)	13 (3%)

[a]Percentages are relative to the value in the 'Total genes' column.

[b]Percentages are relative to the value in the 'Essential genes' column.

93 **Table S4. Metabolic pathways that are essential in *A. actinomycetemcomitans*.**

Carbohydrate metabolism	Glycolysis
	Pyruvate dehydrogenase
	Pentose phosphate pathway
	Phosphoribosyl pyrophosphate (PRPP) biosynthesis
	Fructose catabolism
	Acetate fermentation
Energy metabolism	ATP synthase
	Cytochrome oxidase
Lipid metabolism	Fatty acid biosynthesis
	Phospholipid biosynthesis
Nucleotide metabolism	Purine metabolism
	Pyrimidine metabolism
Amino acid metabolism	Glutamine/glutamate biosynthesis
	Glycine/serine biosynthesis
	Threonine biosynthesis[a]
	Cysteine biosynthesis[a]
	S-adenosyl methionine (SAM) metabolism
	Lysine/diaminopimelate biosynthesis
	Chorismate biosynthesis
	Glutathione biosynthesis[a]
Glycan metabolism	Lipopolysaccharide biosynthesis
	Peptidoglycan biosynthesis
Cofactor and vitamin metabolism	Riboflavin metabolism
	Coenzyme A (CoA) biosynthesis
	Lipoic acid biosynthesis[b]
	Tetrahydrofolate (THF) metabolism
	Heme biosynthesis
	Menaquinone biosynthesis
Terpenoid backbone metabolism	Isoprenoid biosynthesis
Genetic information processing	RNA polymerase
	Ribosome
	Aminoacyl-tRNA biosynthesis
	RNA degradation
	DNA replication
	DNA repair
Membrane transport	Sec-dependent protein export
	Thiamine transport
	Spermidine transport
	Arginine transport
	Cystine transport
	Peptide transport
	Zinc transport[a]

	Lipopolysaccharide transport
	Lipoprotein transport

[a]Essential in *A. actinomycetemcomitans* 624 only

[b]Essential in *A. actinomycetemcomitans* VT1169 only

Table S5. Potential *A. actinomycetemcomitans*-specific genes that lack an ortholog in DEG.

VT1169	624	Product
VT1169_1070	624_2202	polysaccharide biosynthesis protein
VT1169_0067	624_1100	sodium:glutamate symporter GltS
VT1169_0520	624_2137	thiamine transporter component ThiB
VT1169_0129	624_0989	sigmaE negative regulator RseC
VT1169_0156	624_1789	prepilin peptidase CpaA
VT1169_1349	NA*	fatty acid metabolism regulator FadR

*Did not lack an ortholog in DEG.

Table S6. Summary of Tn-seq analysis for ordered library.

MP	Row	Col	Total reads	With IR[a]	Processed[b]	Mapped[c]	Tags[e]	Top tags[f]	Sites[d]
1	A		6,148,148	4,830,936 (79%)	3,649,129 (76%)	NA	45,957	468	NA
1	B		5,373,820	4,531,279 (84%)	3,491,154 (77%)	NA	39,567	468	NA
1	C		3,887,289	3,251,179 (84%)	2,536,589 (78%)	NA	33,022	468	NA
1	D		6,358,670	5,090,909 (80%)	3,999,603 (79%)	NA	40,989	468	NA
1	E		5,608,446	4,667,131 (83%)	3,595,898 (77%)	NA	46,329	468	NA
1	F		4,969,310	4,203,787 (85%)	3,219,153 (77%)	NA	39,214	468	NA
1	G		4,566,115	3,829,211 (84%)	3,034,204 (79%)	NA	38,681	468	NA
1	H		2,492,905	1,964,043 (79%)	1,510,392 (77%)	NA	29,692	468	NA
1		1	3,545,682	2,447,125 (69%)	1,888,617 (77%)	NA	25,242	312	NA
1		2	4,499,009	3,839,363 (85%)	3,105,883 (81%)	NA	32,024	312	NA
1		3	4,021,073	3,348,562 (83%)	2,676,740 (80%)	NA	28,033	312	NA
1		4	1,234,204	869,763 (70%)	680,343 (78%)	NA	13,083	312	NA
1		5	3,401,107	2,354,929 (69%)	1,900,598 (81%)	NA	20,561	312	NA
1		6	1,644,076	1,133,790 (69%)	880,485 (78%)	NA	12,469	312	NA
1		7	3,119,598	2,490,508 (80%)	1,924,580 (77%)	NA	20,123	312	NA
1		8	4,321,024	3,649,591 (84%)	2,785,342 (76%)	NA	21,526	312	NA
2		9	2,764,821	2,206,111 (80%)	1,741,078 (79%)	NA	18,909	312	NA
1		10	1,808,258	1,358,814 (75%)	1,033,058 (76%)	NA	16,152	312	NA
1		11	5,647,526	4,519,559 (80%)	3,362,231 (74%)	NA	38,515	312	NA
1		12	7,128,590	5,956,554 (84%)	4,627,746 (78%)	NA	42,519	312	NA
2	A		3,164,558	2,542,368 (80%)	2,006,113 (79%)	NA	26,697	480	NA
2	B		2,437,736	1,860,128 (76%)	1,481,401 (80%)	NA	19,600	480	NA
2	C		1,861,927	1,363,173 (73%)	1,061,002 (78%)	NA	18,919	480	NA

2	D		2,472,193	1,911,489 (77%)	1,528,554 (80%)	NA	22,258	480	NA
2	E		5,805,898	4,543,344 (78%)	3,336,726 (73%)	NA	26,941	480	NA
2	F		3,145,652	2,518,957 (80%)	2,003,930 (80%)	NA	28,317	480	NA
2	G		1,908,165	1,319,514 (69%)	979,268 (74%)	NA	16,016	480	NA
2	H		1,774,856	1,293,266 (73%)	1,015,537 (79%)	NA	14,854	384	NA
2		1	2,942,490	2,427,646 (83%)	1,967,551 (81%)	NA	29,300	768	NA
2		2	3,592,573	2,995,991 (83%)	2,452,254 (82%)	NA	29,569	768	NA
2		3	2,726,438	2,185,599 (80%)	1,781,679 (82%)	NA	27,188	768	NA
2		4	3,685,855	3,043,665 (83%)	2,498,167 (82%)	NA	33,499	768	NA
2		5	3,241,506	2,603,686 (80%)	2,120,114 (81%)	NA	26,313	672	NA
no antibiotic[g]			7,637,188	6,919,034 (91%)	6,452,821 (93%)	5,942,808 (92%)	44,658	NA	2,634

Percentages are relative to the value in the adjacent left column.

[a]Reads containing the transposon inverted repeat (IR) sequence (CCAACCTGTTA) with 0 mismatches in its expected location within the read (IR end position at base 39)

[b]Reads remaining after trimming 3' low-quality bases (phred + 33 < 35), trimming poly-C tails, and removing reads <12 bases long

[c]Reads mapping end-to-end to the *A. actinomycetemcomitans* genome with high quality (MAPQ > 39)

[d]Sites identified after accounting for the 2 bp duplication associated with mariner insertion events (by adding 2 to the genome position of reads mapping to the negative strand)

[e]Unique 12-bp sequences next to transposon insertions

[f]Most abundant tags corresponding to the number expected if only 1 mutant was present per well

113 [g]Control for erythromycin screen (see Table S1)

Table S7. Ordered transposon mutant library.

Insertion mutants arrayed into wells	3,744 mutants
Wells successfully mapped*	3,245 wells
Mutants per well	
1 mutant per well	981 wells
2 mutants per well	990 wells
3 mutants per well	665 wells
4 mutants per well	352 wells
>4 mutants per well	257
Wells occupied per mutant	
Found in only 1 well	1,120 mutants
Found in 2 wells	57 mutants
Found in 3 wells	9 mutants
Found in 4 wells	80 mutants
Found in >4 wells	265 mutants
Unique insertions	1,531
Within coding genes (excluding 3' 10%)	990
Intergenic	541
Coding genes hit internally (of 1,912)	626
Coding genes not hit internally	1,286

*Found to contain a mutant(s) using CP-CSeq

Table S8. Summary of Tn-seq results for erythromycin screen.

Fold change vs. control	Analysis level	Total sites or genes	Inter-genic sites	Genic sites	Genes hit internally[a]	Combined[b]
Negative	Sites	14	6	8	7	9
	Genes[a]	7	-	-	-	
Positive	Sites	147	49	98	90	90
	Genes[a]	15	-	-	-	

[a]Not within the 3' 10% of the gene

[b]Total genes that changed in the combined results

Table S9. COG enrichment analysis for genes that increased in abundance in the presence of erythromycin.

COG category	Function	Fold enrichment[a]	P value[b]
P	Inorganic ion transport	1.9	0.07
V	Defense mechanisms	6.4	0.001

[a]Percent of the essential genome made up by a COG divided by the percent of the whole genome made up by same COG

[b]Two-tailed Fisher's exact test

Script S1. vlookup.awk shell script mimicking vlookup function in Excel.

Usage: vlookup.awk file1 file 2, where file1 is the binary relationship file and file2 is the test file

Source: http://unix.stackexchange.com/questions/88550/vlookup-function-in-unix

```
FNR==NR{
  a[$1]=$2
  next
}
{ if ($1 in a) {print $1, a[$1]} else {print $1, "NA"}  }
```