

Genome-Wide Identification by Transposon Insertion Sequencing of *Escherichia coli* K1 Genes Essential for *in vitro* Growth, Gastrointestinal Colonizing Capacity and Survival in Serum

Alex J. McCarthy, Richard A. Stabler, Peter W. Taylor

LEGENDS FOR SUPPLEMENTAL TABLES

Table S1. Tab 1: Genes identified by TraDIS as essential for growth of *E. coli* A192PP in Luria-Bertani (LB) broth. Systematic ID, gene identifier in annotated A192PP genome (1); strand, strand location of coding DNA sequence (CDS); gene, predicted gene annotation; size, size of CDS (bp); function, predicted function; pvalue_essential, value of essentiality determined from gamma distribution; K12, essential for growth of *E. coli* K12 MG1655 (2); EC958, essential for growth of *E. coli* ST131 urinary isolate (3); KEGG_no, KEGG orthology number; KEGG_description, KEGG predicted function; ko_no, KEGG pathway number; ko_description, KEGG pathway description; EC_no, Enzyme Commission number (EC number) for enzyme classification. Tab 2: KEGG pathways enriched for, or depleted of, *E. coli* A192PP essential genes. KEGG pathway, KEGG pathway description; whole, total number of CDS in the *E. coli* A192PP genome for each category; Whole%, percentage of CDS for each category in the *E. coli* A192PP genome; Essential, number of CDS defined as essential by TraDIS; Essential%, percentage of CDS for each category; Dif%, Essential% minus whole%; %genome, ratio Essential:Whole (D:B) X 100.

Table S2. *E. coli* K1 A192PP genes required for GI colonization. GeneID, A192PP genome systematic gene number; Norm_in, normalised read depth in input pool; Norm_MSI, normalised read depth in from TraDIS library recovered from the middle section of the small intestine (MSI) 4 h after initiation of colonization; log₂FoldChange, log₂ (Norm_out/Norm_in); * indicates number approaching negative infinity due to division of zero reads in output pool; pval, p-value; Gene, predicted gene name; Function, manually curated gene function; PROKKA function, automated functional annotation using an *E. coli* custom library.

Table S3. *E. coli* K1 A192PP genes required for survival in human serum. GeneID, A192PP genome systematic gene number; Function, manually curated gene function; PROKKA function, automated functional annotation using an *E. coli* custom library. Log₂-fold change value and a *p* value for each mutant of each gene are provided.

References

1. McCarthy AJ, Negus D, Martin P, Pechincha C, Oswald E, Stabler RA, Taylor PW. 2016. Pathoadaptive mutations of *Escherichia coli* K1 in experimental neonatal systemic infection. PLoS One 11:e0166793.
2. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2:2006.0008.
3. Phan MD, Peters KM, Sarkar S, Lukowski SW, Allsopp LP, Gomes Moriel D, Achard ME, Totsika M, Marshall VM, Upton M, Beatson SA, Schembri MA. 2013. The serum resistome of a globally disseminated multidrug resistant uropathogenic *Escherichia coli* clone. PLoS Genet 9:e1003834.

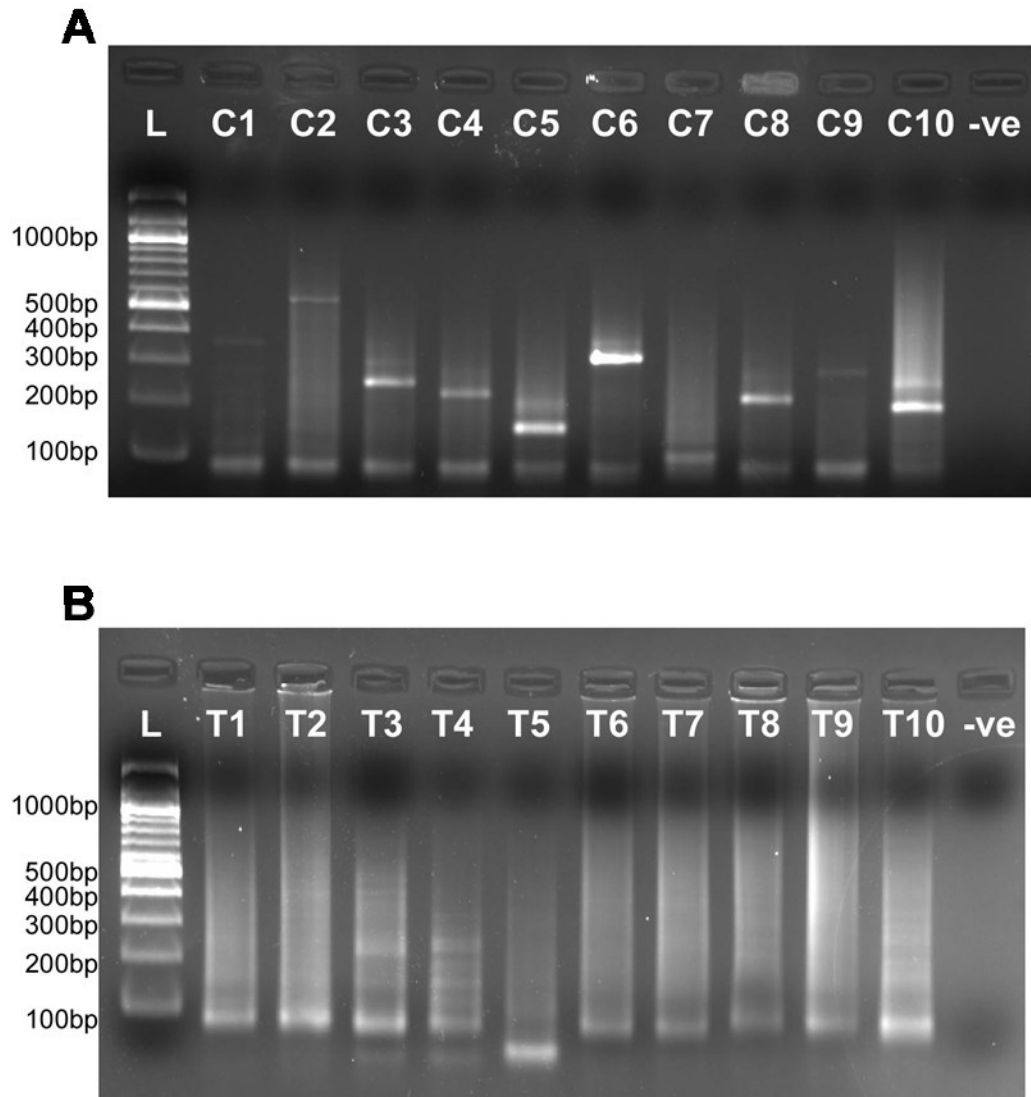
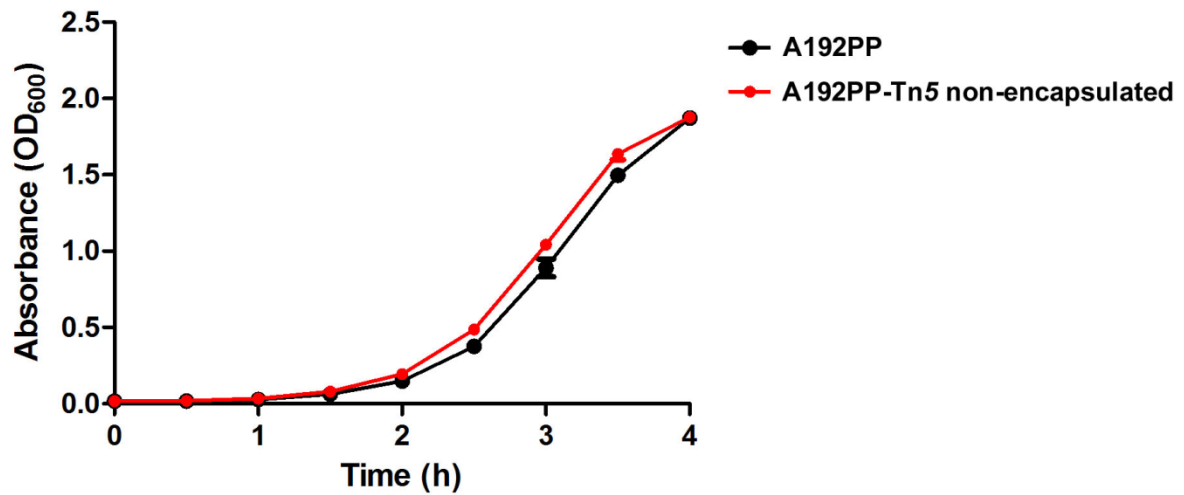
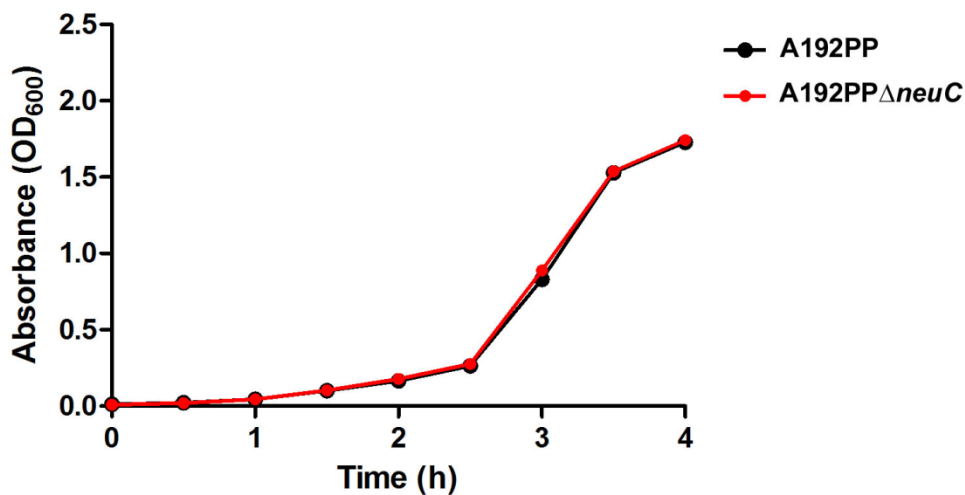


FIG S1. Linker PCR was employed to assess Tn5 insertion site diversity in: (A) ten individual adjacent colonies grown on antibiotic-supplemented Luria-Bertani agar and (B) ten individual pools of 2000-5000 colonies each.

A



B



| Strain | Mean Generation Time (min ± SD) |
|---------------------------|---------------------------------|
| A192PP | 24.45 ± 0.94 |
| A192PP-Tn5 non-capsulated | 23.64 ± 0.21 |
| A192PP | 30.67 ± 0.52 |
| A192PPΔneuC | 30.21 ± 0.941 |

FIG S2. Comparison of growth kinetics of a randomly selected non-encapsulated mutant from the Tn5 TraDIS library in MH broth (A) and a non-encapsulated single gene mutant constructed using bacteriophage λ Red recombination in LB broth (B). $n=3$ in both cases. There were no significant differences in absorbance values at any time point when the log-rank [Mantel-Cox] test was applied. Student's t test was used to evaluate generation times.

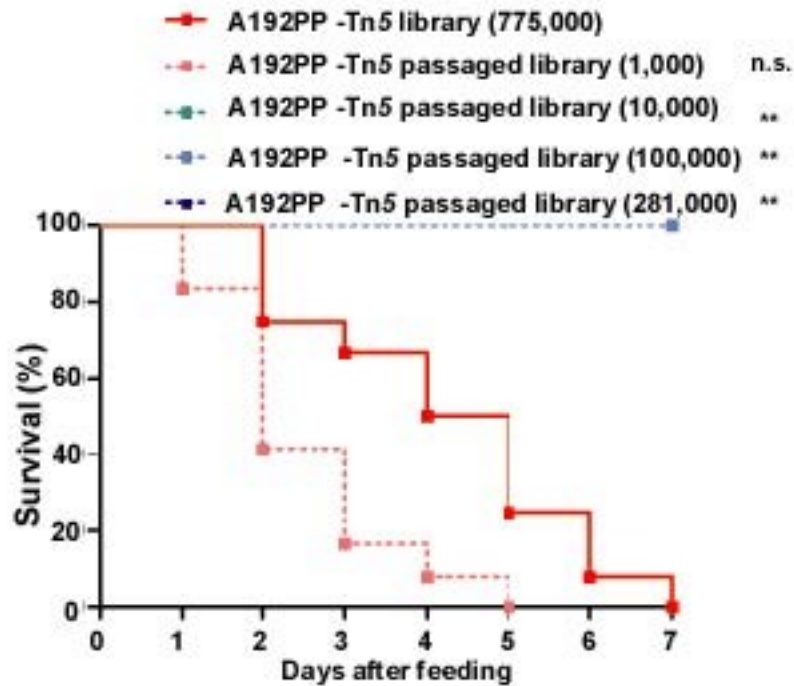


FIG S3. High-complexity cultured *E. coli* A192PP-Tn5 libraries are avirulent in neonatal rats. Survival of P2 rats colonized with *E. coli* A192PP-Tn5 libraries of differing complexities (1,000, 10,000, 100,000 or 281,000 mutants). Libraries were cultures in LB broth (8 h; 37°C) prior to initiation of colonization. Pups ($n = 12$ for each group) received $2-4 \times 10^6$ CFU by the oral route. Log-rank [Mantel-Cox] was used to compare rat survival following administration of cultured libraries with the uncultured complete library of 775,000 mutants: ns, non-significant, * $P < 0.05$, ** $P < 0.01$.

Table S4: Oligonucleotides for construction of targeted mutants

| Gene | Primer | Sequence (5' to 3') |
|-------------|-----------------|--|
| <i>lacZ</i> | <i>lacZ</i> -P1 | tggatttccttacgcgaaatacgggcagacatggcctgcccggttattatgtgtaggctggagctcttc |
| | <i>lacZ</i> -P2 | tatgttggtgaaattgtgagcgaataacaatttcacacaggatacagctcatatgaatcctccttag |
| <i>neuC</i> | <i>neuC</i> -P1 | ctagagctgaatatggaatagttcggagacttttgacaatgctaagagaagttaggctggagctcttc |
| | <i>neuC</i> -P2 | tgagaatcataacgaaagacaaaacaagcacttttttctagtcataaccatataatcctccttag |
| <i>rfaH</i> | <i>rfaH</i> -P1 | cgtaaagctttgctatccttgcgccccgattaacggataagagtcattgtgtaggctggagctcttc |
| | <i>rfaH</i> -P2 | ctggctgccaccaggatgccaatgtcaaaactgttgggattgcttcatatgaatcctccttag |
| <i>traL</i> | <i>traL</i> -P1 | gtgaaatccttcaattacaacctctgtattttccggcttcgataaagttaggctggagctcttc |
| | <i>traL</i> -P2 | cttatgataaataaaagtcgtaaaattacaattacacggacatacaaaacatataatcctccttag |
| <i>vasL</i> | <i>vasL</i> -P1 | tctgcgtcatctcaaacagcaggagcggcgctactgatggcaagtaacgcgtgtaggctggagctcttc |
| | <i>vasL</i> -P2 | aggcacatataccttattggtacataaatccccgatgttactgacttcatatgaatcctccttag |
| <i>waaW</i> | <i>waaW</i> -P1 | atagtactcatccttaattattattgtaactcagacatccatgatttttagttaggctggagctcttc |
| | <i>waaW</i> -P2 | taaaaaattaaaaggcaagcgtaaaccaacacagtcaaaacggaaccaacatataatcctccttag |
| <i>yaeQ</i> | <i>yaeQ</i> -P1 | cgtattccgttacaatggcctcctgattcgaaaggagtttcttatggcgctgtgtaggctggagctcttc |
| | <i>yaeQ</i> -P2 | actcgcatcagggatagcaacatgtcgggaatcacaatcatgaaggttcatatgaatcctccttag |
| <i>yjiG</i> | <i>yjiG</i> -P1 | gccgatgaaattcatcggcaactttgggccttttagaaatggattttgtgtaggctggagctcttc |
| | <i>yjiG</i> -P2 | acaaatcatctctgtgtgattaatgggatttcattatattcctgacatataatcctccttag |
| 0678 | 0678-P1 | tagaaagtaaaattatcggacattttatgccccacacagtcattaccgctgtaggctggagctcttc |
| | 0678-P2 | aaggcgttgatgccacacaacgcctcactgttcttttctcctatgaatcctccttag |
| 3010 | 3010-P1 | tcgcgaagaataatgatgaacttggcaaggatgattatgctgattaagtgtaggctggagctcttc |
| | 3010-P2 | tatctataacaaaaccatccggtgattttgtcatttttagccatcatatgaatcctccttag |

Table S5: Oligonucleotides for confirmation of targeted mutants

| Gene | Primer | Sequence (5' to 3') | Fragment size | |
|-------------|--------------------|-----------------------|---------------|------|
| | | | wildtype | Δ |
| <i>lacZ</i> | <i>lacZ</i> -ampF | ATGCCGGTAATAATCCACAGC | 3917 | 1600 |
| | <i>lacZ</i> -ampR | TGCCATGTCCGGTTTTCAA | | |
| <i>neuC</i> | <i>neuC</i> - ampF | GACAATGCCAGGAAAAACAAG | 1510 | 1600 |
| | <i>neuC</i> - ampR | AAACGAAATAGCGGAGATTGT | | |
| <i>rfaH</i> | <i>rfaH</i> - ampF | ACCACGGATGCCAATGTCA | 664 | 1600 |
| | <i>rfaH</i> - ampR | GTTTCATCTTTGCGATGCTGT | | |
| <i>traL</i> | <i>traL</i> - ampF | ACACGATTCTATTGGCCCTT | 873 | 1600 |
| | <i>traL</i> - ampR | GTATTTTTCCGGCTTCGCAT | | |
| <i>vasL</i> | <i>vasL</i> - ampF | TCTGCCGATCTCAGTCTGAT | 1854 | 1600 |
| | <i>vasL</i> - ampR | GGGCCACAGTCAAGAGGTAA | | |
| <i>waaW</i> | <i>waaW</i> - ampF | GGGTAATCATTGCTCATCGTG | 1308 | 1600 |
| | <i>waaW</i> - ampR | GGTAAAAGCTGTACGGCAGA | | |
| <i>yaeQ</i> | <i>yaeQ</i> - ampF | AACTCTGTTTCGCAAGGTGA | 771 | 1600 |
| | <i>yaeQ</i> - ampR | AAAACGCAGATGAATAGCCG | | |
| 0678 | 0678- ampF | TGTCAGGGAGTGAAGAGACAA | 705 | 1600 |
| | 0678- ampR | AAGTGCCTCGTTTACCGTCAT | | |
| 3010 | 3010- ampF | TTCTGTTCTAGATGCAAGGGC | 318 | 1600 |
| | 3010- ampR | ATGATGAACTTGGCAAAGGA | | |
| <i>wzzE</i> | <i>wzzE</i> -ampF | AAACGCAGACTGCGTAGAAA | 1195 | 1600 |
| | <i>wzzE</i> -ampR | GGCGCGTACCAAATACAGTCA | | |

Table S6: Oligonucleotides for construction of complemented mutants

| Gene | Primer | Sequence (5' to 3') |
|-------------|---------------------|-----------------------------------|
| <i>neuC</i> | <i>neuC</i> -sall-F | CTAGTCGTCGACGACAATGCCAGGAAAACAAG |
| | <i>neuC</i> -sphi-R | GACTAGGCATGCAAACGAAATAGCGGAGATTGT |
| <i>rfaH</i> | <i>rfaH</i> -sall-F | CTAGTCGTCGACACCACGGATGCCAATGTCA |
| | <i>rfaH</i> -sphi-R | GACTAGGCATGCGTTCATCTTTGCGATGCTGT |
| <i>traL</i> | <i>traL</i> -sphi-F | GACTAGGCATGCACACGATTCTATTGGCCCTT |
| | <i>traL</i> -sall-R | CTAGTCGTCGACGTATTTTTCCGGCTTCGCAT |
| <i>waaW</i> | <i>waaW</i> -sall-F | CTAGTCGTCGACGGTAATCATTGCTCATCGTG |
| | <i>waaW</i> -sphi-R | GACTAGGCATGCGGTAAAAGCTGTACGGCAGA |