**SUPPLEMENTARY INFORMATION**

**Evolution of a Plasmid Regulatory Circuit Ameliorates Plasmid Fitness Cost**

Clinton A. Elg, Erin Mack, Michael Rolfsmeier, Tom C. McLean, Olivia Kosterlitz, Elizabeth Soderling, Solana Narum, Paul A. Rowley, Christopher M. Thomas, Eva M. Top

**INDEX**

**Table**

[Table S1. Evolved pBP136Km genotypes. 1](#_Toc157176813)

[Figure S1. Upf31 SDS-Page protein gel 2](#_Toc157176814)

[Figure S2. Maximum growth rates and persistence of archetype IncP plasmid R751 in *E. coli* K-12 (R751) (pCW-LIC-*upf31*) and *E. coli* K-12 (R751) (pCW-LIC-*sacB*). 3](#_Toc157176815)

[Figure S3. Persistence of IncP plasmids pAKD1 and pALTS29 in *E. coli* K-12 4](#_Toc157176816)

[Figure S4. Restriction fragment length profiles of *dam*+ and *dam*- *E. coli* strains suggesting *upf31* does not encode a functional Dam methylase 5](#_Toc157176817)

[Figure S5. Putative *upf31* promoter in pBP136Km 6](#_Toc157176818)

[Data S1. Plasmids containing *upf31* 6](#_Toc157176819)

[Data S2. PHYRE predictions of *upf31* function 6](#_Toc157176820)

[Data S3. RNA-Seq results of K-12(pBP136Km) and K-12(pBP136KmΔ*upf31)* 6](#_Toc157176821)

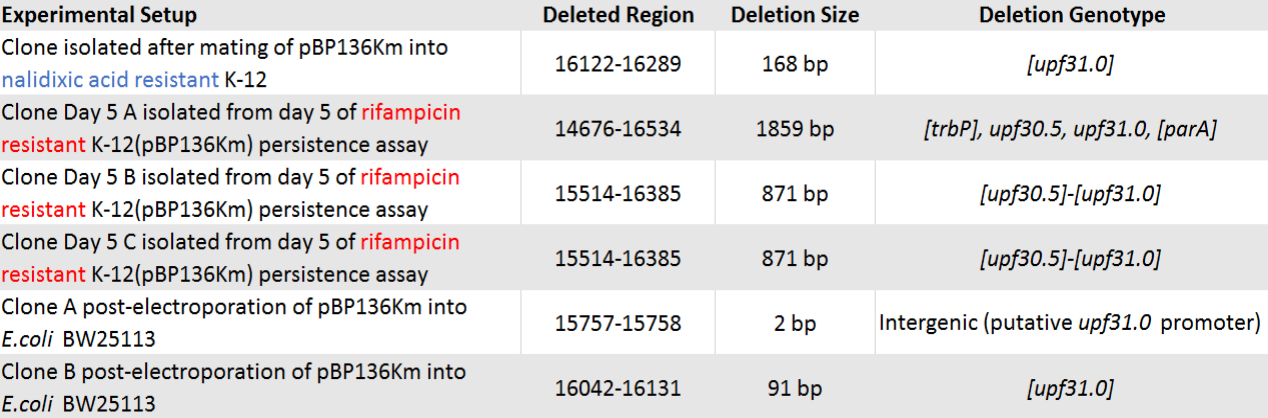
[Data S4. DAVID predictions of differentially expressed cellular pathways 6](#_Toc157176822)

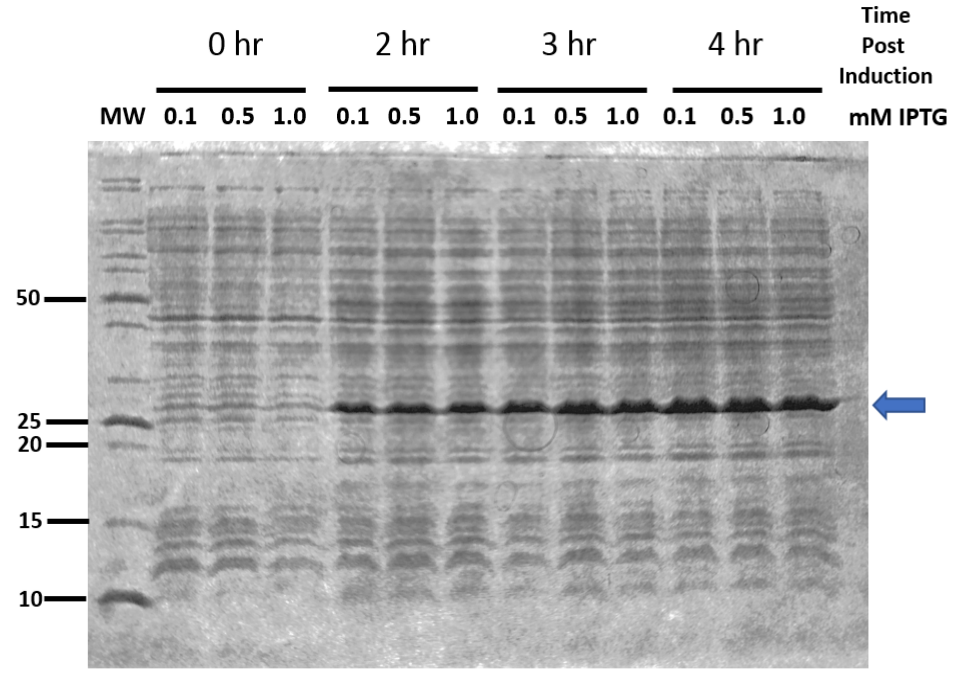
[Data S5. Interactions between AlpaFold2 predicted Upf31 and pBP136-encoded proteins quantified by self-assessment ranking score 6](#_Toc157176823)

# 

# **Table S1. Evolved pBP136Km genotypes.**

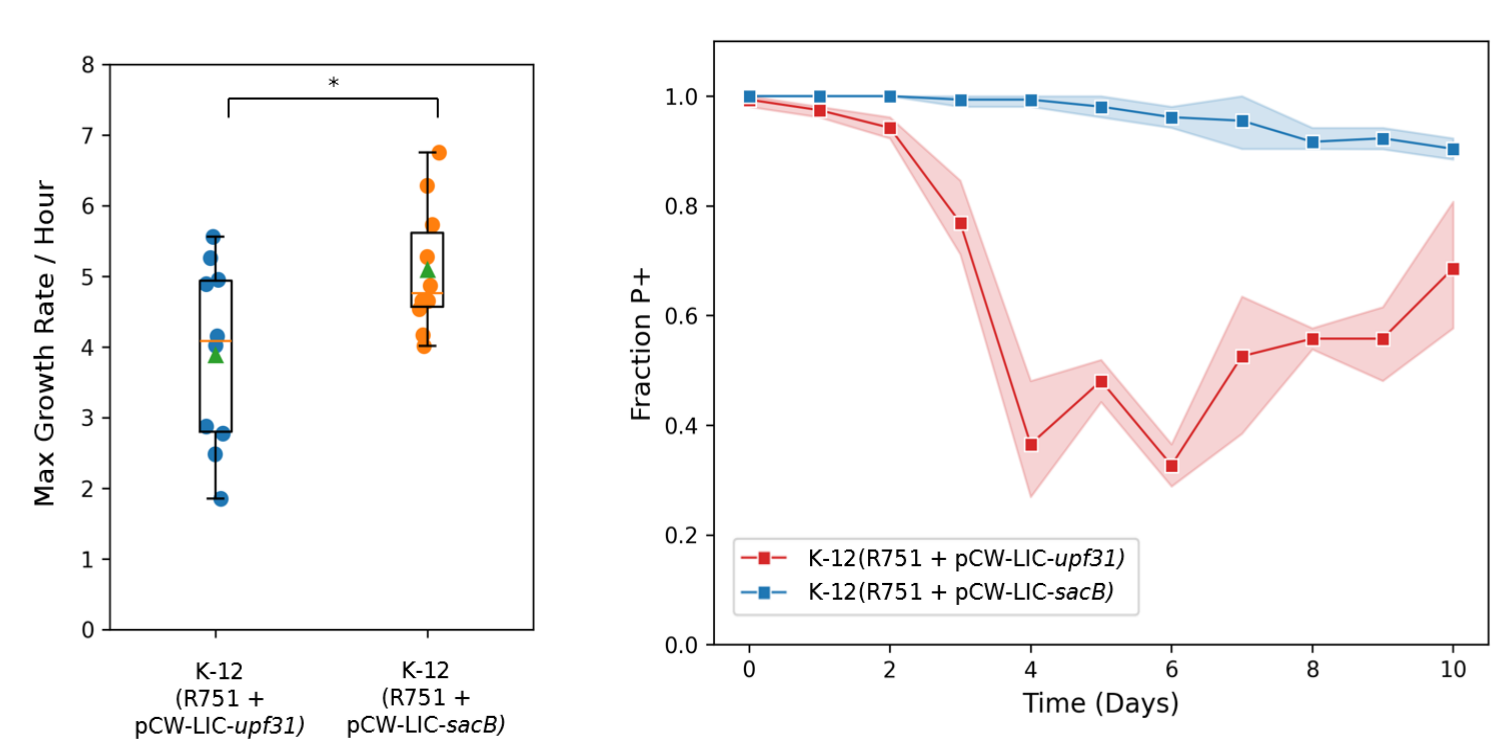
Base pair locations of deleted regions are derived from NCBI genome for pBP136Km (accession number NZ\_OR146256.1).





# **Figure S1. Upf31 SDS-Page protein gel**

Coomassie stained 12% SDS-PAGE demonstrating expression of upf31. MW is the protein mass standards (pertinent masses indicated left of the gel). The blue arrow indicates the position of the upf31 protein (expected mass 25.4 kDa).



# **Figure S2. Maximum growth rates and persistence of archetype IncP plasmid R751 in *E. coli* K-12 (R751) (pCW-LIC-*upf31*) and *E. coli* K-12 (R751) (pCW-LIC-*sacB*).**

Left: Two sample independent T-test p-value = 0.02676, *n*=10. Box is interquartile range, green triangle is mean, orange line is median, *n.s.*=p > 0.05, \*=p ≤ 0.05.

Right: The increase in persistence after day four in the K-12(R751 + pCW-LIC-*upf31)* populations is likely due to clonal interference between evolved pCW-LIC-*upf31*. A clone isolated from day 0 showed no mutations in pCW-LIC-*upf31* or R751. Two clones (CFU A and B) from day 10 were randomly selected and sequenced via Illumina short reads. The sequencing data from both clones showed plasmid heterogenicity of pCW-LIC-*upf31* and no changes in R751*.*We cannot ascertain what proportion of cells contained which pCW-LIC-*upf31*variants, but the variants we identified are as follows: (i) CFU A contained a pCW-LIC-*upf31* variant with complete deletion of the two *tac* promoters upstream of *upf31;* (ii) CFU B contained:

* A pCW-LIC-*upf31* variant with a deletion of the last 367 base pairs of *upf31* and part of the F1 origin of replication (a secondary *oriV* on PCW-LIC).
* A pCW-LIC-*upf31* variant with a 293 base pair deletion entirely within the *upf31* coding sequence.
* A pCW-LIC-*upf31* variant with complete loss of the *lacL* gene and the entire promoter region just upstream of *upf31*.

# **Figure S3. Persistence of IncP plasmids pAKD1 and pALTS29 in *E. coli* K-12**

A screenshot of a test

Description automatically generated

# **Figure S4. Restriction fragment length profiles of *dam*+ and *dam*- *E. coli* strains suggesting *upf31* does not encode a functional Dam methylase**

To test if *upf31* is a functional homologue to *dam*, we performed a comparative restriction enzyme assay. The Dam protein methylates the adenine base in ‘GATC’ sequences (*i.e. ‘*GAmTC’). We utilized two restriction enzymes, *DpnI* and *MobI*, each of which cuts GATC sequences but with different methylation specificity. *DpnI* only cuts GAmTC, while *MobI* cuts unmethylated GATC base pairs.

*Escherichia coli* K-12 natively encodes *dam* and has methylated GAmTC base pairs throughout its genome. Conversely, *E. coli* JM110 contains a *dam* knockout and contains unmethylated GATC. Thus K-12 has a *dam+* and JM110 has a *dam*- methylation profile.

The agarose gels show genomic DNA extracted from the respective strains, which appear as solid bright lines if uncut (left hand side). On the right hand side, when using enzyme *DpnI,* as expected, the DNA of K-12with its *dam+* profile is clearly cut, represented by a smear, whereas the DNA of *E. coli* JM110 is not due to its *dam-* profile; JM110 with pBP136Km and several evolved plasmid variants likewise remain uncut with *DpnI*, suggesting that *upf31* does not restore *dam+* methylation patterns.

A similar experiment was repeated, but with *MobI* instead, which cuts any unmethylated GATC sequences (far right side of the gel). Note the DNA of JM110 with its *dam-* phenotype is cut, along with that of JM110 with pBP136Km and several evolved plasmid variants. This confirms that pBP136Km with *upf31* does not restore *dam+* methylation patterns to JM110. Conversely, as the methylated GAmTC sequences in K-12 are impervious to *MobI* cuts the genomic DNA remained uncut.



# **Figure S5. Putative *upf31* promoter in pBP136Km**

# **Data S1. Plasmids containing *upf31***

See ‘supplemental\_data\_S1.xlsx’ file

# **Data S2. PHYRE predictions of *upf31* function**

See ‘supplemental\_data\_S2.pdf’ file

# **Data S3. RNA-Seq results of K-12(pBP136Km) and K-12(pBP136KmΔ*upf31)***

See ‘supplemental\_data\_S3.xlsx’ file

# **Data S4. DAVID predictions of differentially expressed cellular pathways**

See ‘supplemental\_data\_S4.xlsx’ file

# **Data S5. Interactions between AlpaFold2 predicted Upf31 and pBP136-encoded proteins quantified by self-assessment ranking score**

See ‘supplemental\_data\_S5.xlsx’ file