Improving phage genome annotation to understand phage biology: the case of *Pseudomonas aeruginosa* LES prophages

Enrique González-Tortuero[a], Revathy Krishnamurthi[b], Joanne L. Fothergill[b], Ian B. Goodhead[a], Heather E. Allison[b] and Chloe E. James[a]

[a]School of Science, Engineering and Environment, University of Salford, e.gonzaleztortuero@salford.ac.uk; [b]Institute of Infection, Veterinary and Ecological Sciences (IVES), University of Liverpool

Introduction

- *P. aeruginosa* Liverpool Epidemic Strain (LES) is a major cause of mortality and morbidity in cystic fibrosis
- LES harbours five active prophages involved in fitness and survival
- 76.5% of LES prophage genes are hypothetical proteins
- Aim: to re-annotate the LES prophages and to improve the gene function prediction

Methods

- All five LES prophages were re-annotated using the VIGA bioinformatics pipeline (González-Tortuero *et al.* 2018; Fig. 1)

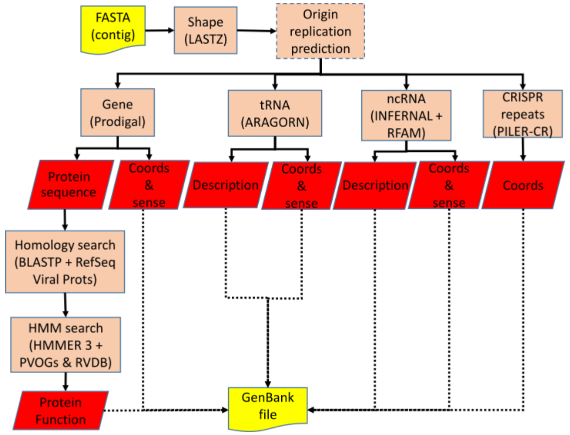

Figure 1: **Flowchart of VIGA pipeline.**

Results

1. Prophage re-annotation

- For all prophages, we predicted the gene function based on the viral life cycle (Fig. 2)
- ncRNA elements were identified in these prophages (tRNAs in prophages 2 and 5 and a putative regulatory element in prophage 4)

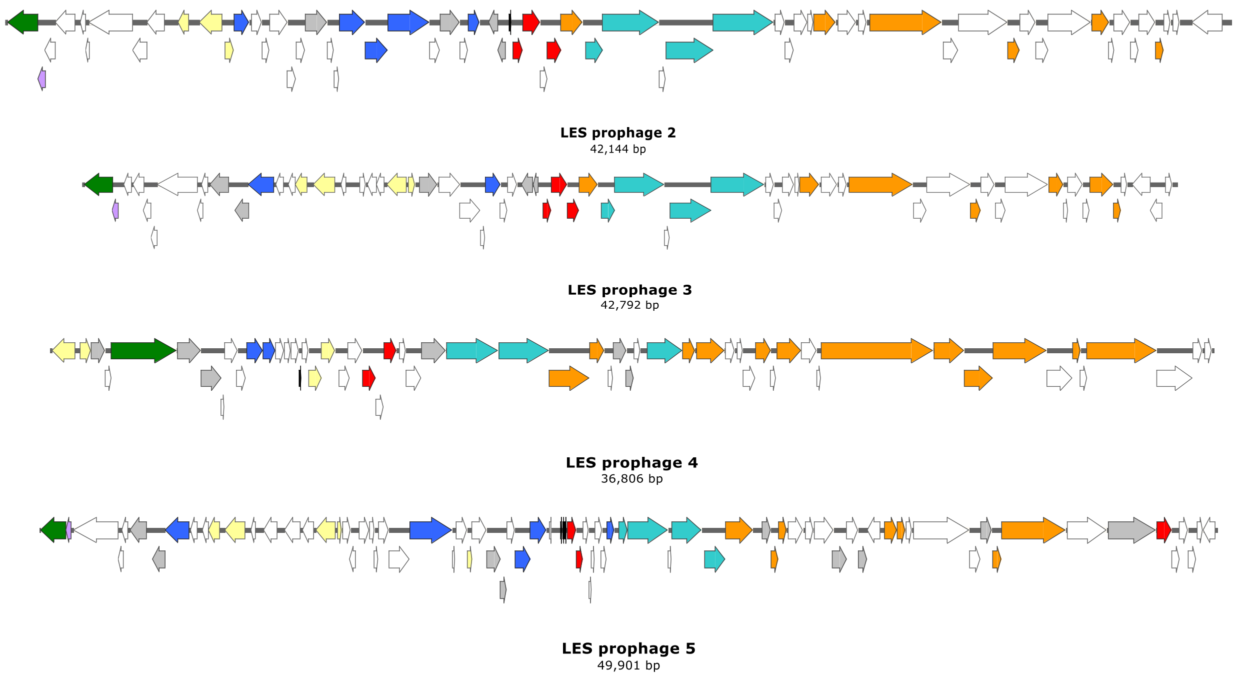

Figure 2: **Genomes of the *P. aeruginosa* LESB58 lambdoid prophages.** Genes were coloured according to the viral lifecycle (integration, recombination, DNA metabolism, regulation, lysis, packaging, structural and others). White arrows indicate genes of unknown function.

González-Tortuero E, Sutton TDS, Velayudhan V, Shkoporov AN, Draper LA, Stockdale SR, Ross RP, Hill C (2018) VIGA: a sensitive, precise and automatic de novo VIral Genome Annotator. bioRxiv 277509

2. Improvements on genome annotation

- After running VIGA, we retrieved 1.17-1.43 times more genes than in the original annotation (Tbl. 1)
- Combining VIGA and knowledge of phage biology, we were able to assign 0.98-2.34 times more gene functions than in the original annotation (Tbl. 1)

Table 1: Proportion of hypothetical genes / total number of CDSs for all prophages before and after running VIGA

Prophage	Original	Reannotation
2	27/44	36/63
3	38/52	37/64
4	35/48	26/56
5	54/65	42/76
6	9/12	10/13

Future steps

- Mapping the expression profiles of LES prophages under inducing and non-inducing conditions
- Validating these annotations based on wet-lab experiments.