



Validation of a random *Vibrio parahaemolyticus* genomic library by selection of quinolone resistance in a heterologous host

Title	Validation of a random <i>Vibrio parahaemolyticus</i> genomic library by selection of quinolone resistance in a heterologous host
Author(s)	O'Boyle, Nicky;Boyd, Aoife
Publication Date	2022-05-10
Publisher	Microbiology Society
Repository DOI	https://doi.org/10.1099/mic.0.001189

Validation of a random *Vibrio parahaemolyticus* genomic library by selection of quinolone resistance in a heterologous host

Nicky O'Boyle^{1,2,*} and Aoife Boyd^{1,*}

Abstract

Vibrio parahaemolyticus is a shellfish-borne pathogen that is a highly prevalent causative agent of inflammatory gastroenteritis in humans. Genomic libraries have proven useful for the identification of novel gene functions in many bacterial species. In this study we prepared a library containing 40 kb fragments of randomly sheared *V. parahaemolyticus* genomic DNA and introduced this into *Escherichia coli* HB101 using a commercially available low copy cosmid system. In order to estimate coverage and suitability of the library and potentially identify novel antimicrobial resistance determinants, we screened for the acquisition of resistance to the fluoroquinolone norfloxacin – a phenotype exhibited by *V. parahaemolyticus* but not the heterologous *E. coli* host. Upon selection on solid medium containing norfloxacin, 0.52% of the library population was resistant, consistent with the selection of a single resistance locus. End-sequencing identified six distinct insert fragments. All clones displayed fourfold increased norfloxacin MIC compared with *E. coli* HB101 carrying an empty vector. The common locus contained within resistant clones included *qnr*, a previously described quinolone resistance gene. These results indicate that the library was unbiased, of sufficient coverage and that heterologous expression was possible. While we hope that this library proves useful for identifying the genetic determinants of complex phenotypes such as those related to virulence, not all norfloxacin resistance genes were detected in our screen. As such, we discuss the benefits and limitations of this approach for identifying the genetic basis of uncharacterized bacterial phenotypes.

INTRODUCTION

V. parahaemolyticus is the world's leading cause of seafood-borne bacterial gastroenteritis [1, 2]. The organism may be found in a planktonic state, embedded in sediments or in close association with shellfish including oysters, clams and shrimp [3, 4]. The most prevalent route of human infection is the consumption of contaminated seafood [5]. Extensive research has been carried out in attempts to characterize the mechanisms of virulence, which have enabled *V. parahaemolyticus* to become a successful human pathogen [2, 6]. Such studies generally require prediction of protein functionality based upon sequence conservation with proteins from other species [7]. Importantly, of the 4832 predicted protein coding regions in the *V. parahaemolyticus* RIMD2210633 genome, 1846 were annotated as hypothetical proteins in 2003 [8]. While understanding of gene function in *V. parahaemolyticus* has increased since then, the 99% identical *V. parahaemolyticus* FDAARGOS_191 genome, which was sequenced in 2017 still had 922 genes classified as hypothetical. Furthermore, transposon insertion site sequencing (TIS) has demonstrated that 91 out of 565 genes that contribute to fitness *in vitro* have no ascribed function and 32 of 230 genes specifically required for optimum *in vivo* fitness were of unknown function [9]. It is clear that the lack of functional annotations for such coding regions represents a significant bottleneck in our understanding of important *V. parahaemolyticus* biological processes. In the era of high-throughput genome sequencing it is becoming increasingly important that we derive meaningful biological information from genome sequence data.

Received 09 March 2022; Accepted 25 April 2022; Published 10 May 2022

Author affiliations: ¹Discipline of Microbiology, School of Natural Sciences and Ryan Institute, National University of Ireland Galway, Galway, H91 TK33, Ireland; ²School of Microbiology, University College Cork, National University of Ireland, Cork, T12 K8AF, Ireland.

***Correspondence:** Aoife Boyd, aoife.boyd@nuigalway.ie; Nicky O'Boyle, noboyle@ucc.ie

Keywords: *Vibrio parahaemolyticus*; genomic library; norfloxacin; resistance; *qnr*.

Abbreviations: BLAST, Basic Local Alignment Search Tool; c.f.u., colony forming units; CMR, comprehensive microbial resource; HPLC, high performance liquid chromatography; IPTG, isopropyl β -D-1-thiogalactopyranoside; MIC, minimum inhibitory concentration; PFGE, pulsed field gel electrophoresis.

One supplementary figure and one supplementary table are available with the online version of this article.

001189 © 2022 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License. This article was made open access via a Publish and Read agreement between the Microbiology Society and the corresponding author's institution.

The functional screening of genomic libraries allows for unbiased mining of microbial genomes for biologically relevant target genes. In this study we generated such a library using randomly sheared genomic DNA from *V. parahaemolyticus*. The use of the high capacity cosmid vector pWEB-TNC allowed for stable ligation of large (~40 kb) fragments, thereby increasing the possibility of co-expression of multiple genes within a single insert – a factor which would be essential for biological processes involving multi-protein complexes or pathways. Before adapting our library to screen for novel genes involved in complex biological processes, we developed a simple growth based selection assay using norfloxacin resistance – a *V. parahaemolyticus* phenotype with a previously studied genetic basis [10–12]. Selection of library clones using a concentration of norfloxacin, which was found to be inhibitory to the heterologous host, allowed for validation of library coverage, genetic diversity of insert fragments and confirmation of expression of a *V. parahaemolyticus* target gene *qnr*.

METHODS

Bacterial strains and culture conditions

V. parahaemolyticus RIMD2210633 was grown in LBN broth (LB +3% NaCl) and LBN agar (LBN broth +1.5% agar). *E. coli* strains were grown in LB broth and LB agar. Concentrations of 100 µg ml⁻¹ ampicillin or 20 µg ml⁻¹ chloramphenicol were included for vector selection. Where required, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to exponential phase cultures after 1.5 h growth and expression was induced for a further 1.5 h at 37 °C.

Genomic library preparation

A 40 kb genomic library of *V. parahaemolyticus* RIMD2210633 was prepared using the pWEB-TNC Cosmid Cloning Kit (Epicentre, USA) according to the manufacturer's instructions. First, *V. parahaemolyticus* genomic DNA was isolated using the Wizard Genomic DNA Purification Kit (Promega, USA). To optimize shearing, samples of genomic DNA were passed through a 51 µm bore syringe needle (Hamilton Messtechnik, Germany) several times. PFGE of sheared DNA showed a reduction in molecular weight towards the desired 40 kb range after 2–5 passages (Fig. S1a, available in the online version of this article). For library preparation, 50 µg genomic DNA was sheared by passaging through the 51 µm bore syringe needle four times. DNA was next end-repaired and size-selected by PFGE. Following staining with SYBR gold, 40 kb fragments were excised under blue light (Fig. S1b). DNA was purified and resuspended in TE buffer before ligating into pWEB-TNC. Ligated cosmids were packaged using Max Plax lambda packaging extract. The titre (c.f.u. µl⁻¹) of the packaged cosmids was determined by infection of 100 ml *E. coli* HB101 with 10 µl packaged cosmids and selecting on LB+ampicillin overnight. Aliquots of packaged cosmids providing approximately 50×coverage of the *V. parahaemolyticus* genome with a 99% probability of containing a given sequence were prepared. The number (*N*) of clones required for 1×coverage of a given genome was calculated by the following formula: $N = \ln(1/P) / \ln(1/f)$, where *P* is the desired probability expressed as a fraction (0.99 for 99%) and *f* is the proportion of the genome contained within a single clone (4.0×10⁴ bp/5.2×10⁶ bp). An aliquot of packaged, titered cosmids was adsorbed on to *E. coli* HB101 and the bacteria were spread on to LB agar +ampicillin. Following overnight growth, 2.5×10⁴ c.f.u. (approximately 50×coverage, with 99% probability) were harvested and pooled. Glycerol stocks were prepared and stored at –80 °C.

Selection of norfloxacin-resistant library clones

Glycerol stocks of the genomic library were serially diluted and spread plated on LB agar and LB agar +0.1 µg ml⁻¹ norfloxacin. The plates were incubated at 37°C overnight. The numbers of resistant colonies were enumerated and expressed as a percentage of the total. Eight resistant clones were selected and sequenced with pWEB.For (5'-TCGTCTTCAAGAATTCGCGGC-3') and pWEB.Rev (5'-ACGACTCACTATAGGGAGAGG-3') primers. Unclipped sequence data was aligned with the *V. parahaemolyticus* RIMD2210633 genome using NCBI BLAST. Coding regions were analysed using the CMR, PSORTb and Pfam databases [13–15].

Confirmation of norfloxacin resistance by MIC

MICs were determined for each of the norfloxacin-resistant clones by incubation in LB broth containing doubling dilutions of norfloxacin from 0.3 µg ml⁻¹ to 0.02 µg ml⁻¹ in a 96-well microtitre plate. Growth was then monitored at 37°C for 24 h using the Sunrise Absorbance Reader (Tecan, Switzerland). The MIC was determined as the minimum concentration of norfloxacin which inhibited growth. Experiments were carried out three times in triplicate and the reported MIC was obtained in at least two of three replicates.

Cloning and expression of *emr* genes

Primers VPA0097.For (5'-CACCATGAGCGGCGCATC-3') and VPA0098.Rev (5'AAACGCCTCTCAAGCTGACAC-3') were designed for amplification and cloning of *emrAB* into the expression vector pET101 as recommended by Invitrogen. Insert DNA was amplified, gel purified, TOPO cloned into pET101 and then transformed into *E. coli* Top 10. Ampicillin-resistant colonies were screened for insert presence by PCR and then confirmed by sequencing. The recombinant vector was then transformed

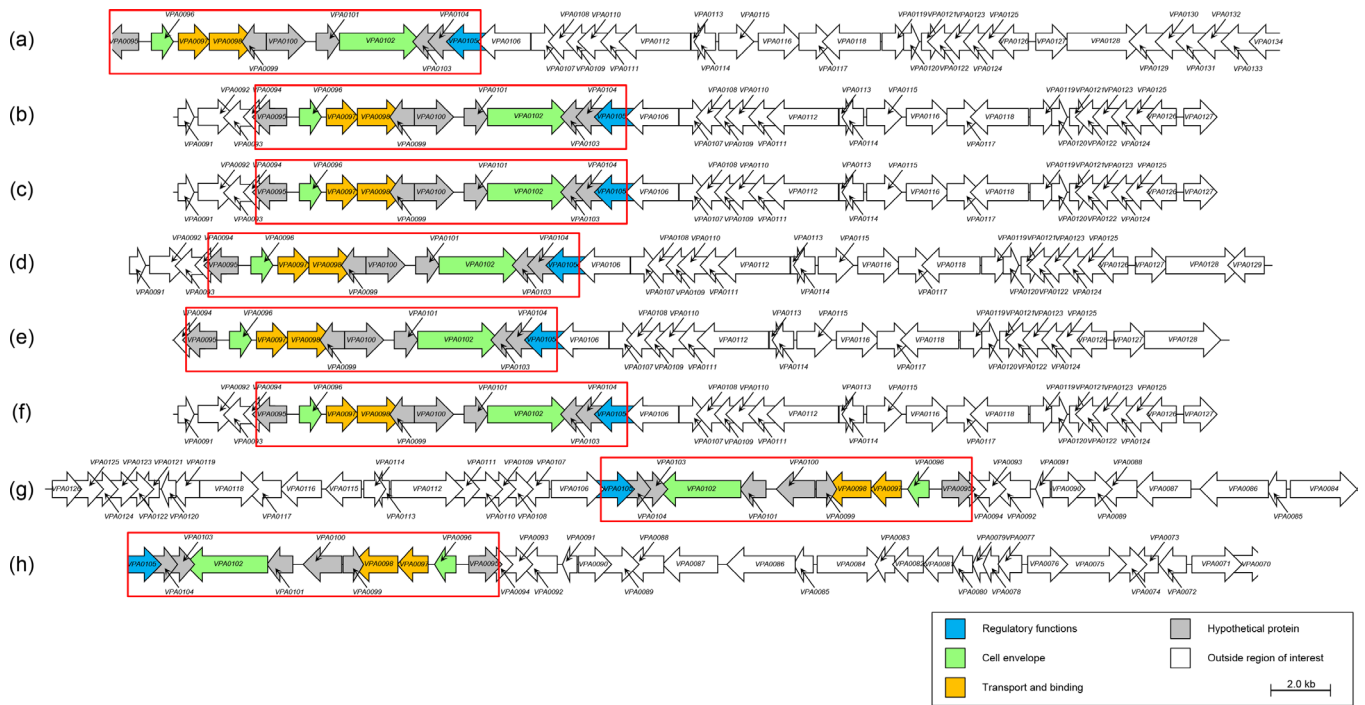


Fig. 1. Schematic representation of the coding regions from norfloxacin-resistant clones. Insert DNA from eight norfloxacin-resistant genomic library clones as identified by BLAST alignment of insert extremities with *V. parahaemolyticus* genome. (a) N1, (b) N2, (c) N3, (d) N4, (e) N5, (f) N6, (g) N7, (h) N10. The common locus contained within all clones is indicated by the red bounding box.

into *E. coli* BL21/DE3. For assessment of the role of *emrAB* in norfloxacin resistance, 0, 100, 500 or 1000 μM IPTG was added to *E. coli* BL21/DE3(pET101) and BL21/DE3(pET101 +EmrAB) throughout MIC experiments.

RESULTS AND DISCUSSION

High capacity cosmid/fosmid-based genomic libraries can be produced using DNA from pure cultures of single bacterial species [16–19] or from complex metagenomic DNA samples such as those associated with soil [20, 21] or gut [22, 23] communities. Studies employing 25–40 kb genomic DNA inserts from single species typically involve the isolation, and banking, of between 1000 and 2500 clones before subsequently screening for the gain of targeted phenotypes and sequencing insert DNA [16–19]. As we intended to employ *en masse* antibiotic selection to validate our library, we bypassed the banking of individual clones and prepared a pooled library of tenfold higher coverage than that used in other similar studies – consisting of 2.5×10^4 c.f.u. of pWEB-TNC-RIMD2210633 transduced *E. coli* HB101 as outlined in Methods.

To determine the appropriate concentration for selection the MICs of norfloxacin for *V. parahaemolyticus* and *E. coli* HB101 were assessed. *V. parahaemolyticus* displayed an MIC of $5 \mu\text{g ml}^{-1}$, while *E. coli* HB101 displayed an MIC of $0.08 \mu\text{g ml}^{-1}$. Therefore, a concentration of $0.1 \mu\text{g ml}^{-1}$ was employed for plate-based selection of norfloxacin-resistant (Nor^R) library clones. Spread plating of the library on medium without selection and medium containing norfloxacin yielded a 0.52% rate of resistance (1.6×10^5 resistant c.f.u. ml^{-1} from a total of 3.1×10^7 c.f.u. ml^{-1}). Given the 5.12 Mb genome of *V. parahaemolyticus* and a ~40 kb insert size, selection of one resistance locus per unit coverage would yield 0.77% resistance. As such, the rate of resistance was in reasonable agreement with the selection of a single resistance locus.

Eight Nor^R clones were randomly selected for sequence analysis. Given our prediction of a single locus being the basis of resistance selection, we reasoned that this small number would be sufficient to assess coverage and/or bias within the library. The extremities of the inserts were sequenced using pWEB.For and pWEB.Rev primers. Sequence data were aligned to the *V. parahaemolyticus* genome. Fig. 1 illustrates the predicted coding regions identified. A common region was observed in all eight clones (red box, Figs. 1 and 2), indicating a strong likelihood of a role for the selected locus in norfloxacin resistance. Only three clones contained identical inserts (N2, N3 and N6, Fig. 1) indicating that growth during selection did not result in bias. The isolation of six genetically distinct clones confirmed that a good level of coverage was attained (at least sixfold).

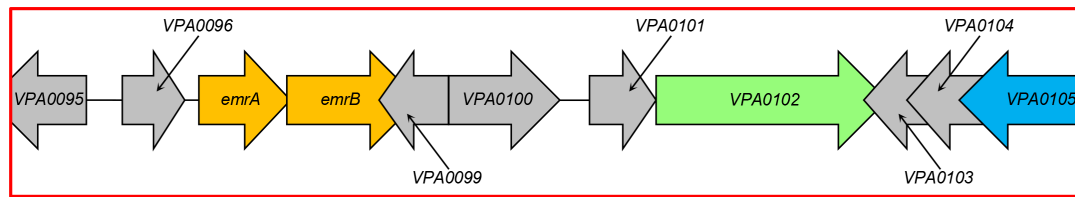


Fig. 2. Expanded view of the common locus required for norfloxacin resistance.

The six genetically distinct Nor^R clones were next assessed for MIC of norfloxacin. Each of these clones displayed an MIC of 0.3 $\mu\text{g ml}^{-1}$ norfloxacin – fourfold higher than the MIC of 0.08 $\mu\text{g ml}^{-1}$ observed for *E. coli* HB101 (pWEB) (Fig. 3). The dynamics of growth in 0.08 $\mu\text{g ml}^{-1}$ norfloxacin were similar for all resistant clones (Fig. 3a), further suggesting that resistance was the result of expression of a similar quinolone resistance mechanism by each clone.

In order to identify the specific gene/genes responsible for the gain of resistance to norfloxacin in *E. coli* HB101, bioinformatic analysis was focused on the common region (Fig. 2) present in all clones. Putative open reading frames and functional categories were analysed using the Comprehensive Microbial Resource (CMR) database [13]. pSORTb was used to predict subcellular localization [14]. Pfam domains were also analysed in attempts to infer functionality based on sequence motifs conserved with other species [24]. A summary of this analysis is illustrated in Table 1. The predicted properties of three genes (*VPA0095*, *VPA0097* and *VPA0098*) suggested that they could be associated with fluoroquinolone resistance.

VPA0097 and *VPA0098* are predicted to encode a putative Emr major facilitator superfamily multi-drug efflux pump (Fig. 2, Table 1). Emr pumps are composed of an inner membrane component (*EmrB*), which confers substrate specificity, a periplasmic component (*EmrA*) and the outer membrane pore protein (*TolC*) [25]. To assess a potential role for the Emr efflux system in norfloxacin resistance, *emrAB* were cloned downstream of an IPTG inducible promoter. BL21 (pET101) and BL21 (pET101-*emrAB*) were equally inhibited by 0.04 $\mu\text{g ml}^{-1}$ norfloxacin at 0, 100, 500 and 1000 μM IPTG (Table S1), indicating that heterologous expression of *EmrAB* does not contribute to norfloxacin resistance.

VPA0095 was annotated as encoding a hypothetical protein with a predicted DNA-binding pentapeptide repeat domain (Table 1) and as such, a direct involvement in norfloxacin resistance was not immediately apparent. Following further in-depth literature searches, a potential role in fluoroquinolone resistance became evident. *VPA0095* has been reported to encode a Qnr family quinolone resistance protein [11, 12]. The precise mechanism of action of Qnr proteins is unknown, however it has been shown that QnrA from *Klebsiella pneumoniae* inhibits formation of the gyrase-DNA-quinolone complex [26]. *K. pneumoniae* QnrA shares 58% amino acid identity with the Qnr encoded by *VPA0095* and as such the proteins may share a common mechanism of action. Indeed, expression of Qnr from *V. parahaemolyticus* CIP71.2 in *E. coli* Top 10 increased its norfloxacin MIC from 0.032 to 0.38 $\mu\text{g ml}^{-1}$ [11], a value close to that observed in our library clones. We considered employing the pET101 expression system to confirm that expression of *V. parahaemolyticus* RIMD 22210633 Qnr was responsible for the elevated MIC observed in our library clones, however as mentioned, heterologous expression of Qnr from *V. parahaemolyticus* CIP71.2 in *E. coli* has

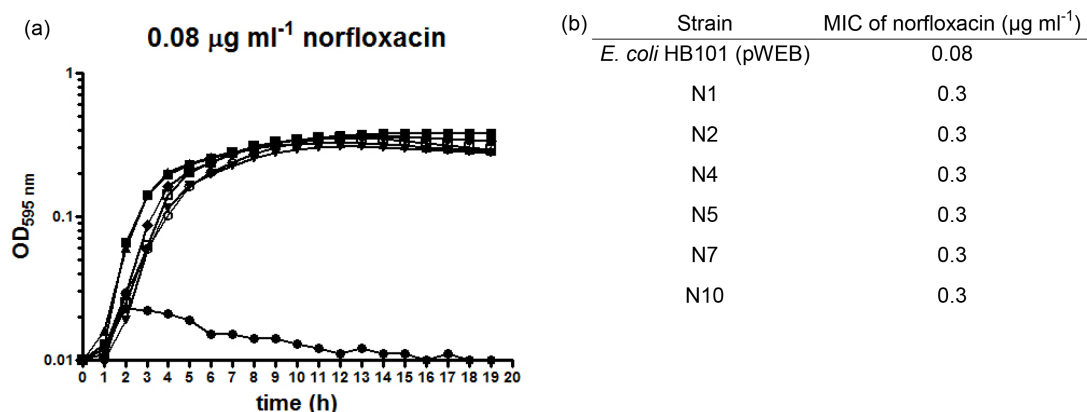


Fig. 3. Growth dynamics and MIC of norfloxacin resistant clones under norfloxacin selection. (a) Growth curve analysis of library clones N1 (■), N2 (▲), N4 (▼), N5 (◆), N7 (◻), N10 (◼) cultured in LB +0.08 $\mu\text{g ml}^{-1}$ norfloxacin compared with the empty vector control *E. coli* HB101 (pWEB) (●). (b) Summary of norfloxacin MICs for genetically distinct library clones.

Table 1. Summary of bioinformatic analysis carried out on the common region contained within the insert DNA of norfloxacin-resistant clones

Gene	Putative operon	Functional category	JCVI annotation	Pfam domains	Localization (score)
VPA0095	None	Hypothetical protein	Putative pentapeptide repeat protein	1.3 pentapeptide repeat domains: Domains which mimic the structure of DNA. Family member MfpA binds to and inhibits DNA gyrase.	C ³ . (8.96)
VPA0096	VPA0096, VPA0097, VPA0098	Cell envelope	Outer membrane protein OprG	1. Signal peptide: Signal for secretion through the CM. 2. OmpW: Family of outer membrane proteins. The receptor for colicins in <i>E. coli</i> . May also form an outer membrane channel or be involved in cell wall biogenesis.	OM ⁴ . (10.0)
VPA0097	VPA0096, VPA0097, VPA0098	Transport and binding	HlyD family secretion protein	1.HlyD: Bacterial secretion protein. Members include proteins which translocate bacterial haemolysins. Multi-domain hit for the multi-drug efflux domain EmrA. Also contains two catalytic biotinyl/lipoyl-dependent carboxylase domains	CM ² . (9.82)
VPA0098	VPA0096, VPA0097, VPA0098	Transport and binding	Putative EmrB/QacA family efflux protein	1.MFS1: Major facilitator superfamily efflux protein domain. Transport a wide variety of substrates. Multi-domain hit for EmrB. Commonly encoded in the same locus as EmrA. Multi-drug resistance. Family members confer resistance to quaternary ammonium compounds, fatty acids and tetracenomycin C.	CM ² . (10.0)
VPA0099	None	Hypothetical protein	Conserved hypothetical protein	1.ROF: Modulator of Rho-dependent transcription termination. ROF binds to Rho and inhibits Rho-dependent termination.	Unknown
VPA0100	None	Hypothetical protein	Metallo- β -lactamase protein	1.Lactamase B2: Part of the β lactamase superfamily. Predicted Zinc-dependent β lactam specific hydrolase.	C ³ . (8.96)
VPA0101	VPA0101, VPA0102	Hypothetical protein	Conserved hypothetical protein	None	C ³ . (8.96)
VPA0102	VPA0101, VPA0102	Hypothetical protein	Putative lipoprotein	1.PRK10137: α glucosidase domain. Also domain for trehalase activity. Recycles trehalose to glucose. Trehalose is a glucose dimer which can be broken down under conditions of stress.	Unknown
VPA0103	VPA0103, VPA0104, VPA0105	Hypothetical protein	Expressed protein	1.Macro: ADP ribose binding domain. In <i>E. coli</i> , family member YmdB down-regulates RNase III activity by interacting with the region of the protein required for dimerization/activation.	Unknown
VPA0104	VPA0103, VPA0104, VPA0105	Hypothetical protein	Ring cleaving di-oxygenase	1.Glyoxalase: Domain found in bleomycin resistance proteins, glyoxalase I and ring cleaving di-oxygenase, which incorporate both atoms of O ₂ into substrates using metal ions as co-factors.	C ³ . (9.97)
VPA0105	VPA0103, VPA0104, VPA0105	Regulatory functions	Putative LysR transcriptional regulator	1. HTH1: DNA binding helix turn helix domain. Members include resistance regulators AmpR and CatM. 2. LysR: Substrate binding domain. LeuO type domain. LeuO activates leucine synthesis and biofilm formation in <i>V. cholerae</i> .	C ³ . (9.97)

Gene names and functional categories were derived from primary annotations in the CMR database. Putative operons were predicted by analysing terminator/promoter positions, ribosomal binding sites and directions of transcription using region view and gene graphic features on the CMR database. Annotations were derived from JCVI annotations in the CMR database. Conserved protein domains were analysed using the Pfam, protein families database and the NCBI eLASP conserved domains feature. Signal peptides and transmembrane helices were detected using the Pfam database and the CMR TmHMM analysis feature. Subcellular localization was predicted using the pSORTb computational analysis tool. Abbreviations: ¹ EC, extracellular; ² CM, cytoplasmic membrane; ³ C, cytoplasm; ⁴ OM, outer membrane.

already been tested and confirmed to result in elevated norfloxacin resistance [11]. Given the high degree of Qnr amino acid identity (214/216 residues) between strains CIP71.2 and RIMD2210633, we believe it is reasonable to conclude that carriage and expression of VPA0095 is responsible for the norfloxacin resistance phenotype observed in our study.

Our study also reveals some important limitations of the pWEB-TNC-based library system. The *V. parahaemolyticus* NorM efflux pump has been shown to confer resistance to ciprofloxacin and norfloxacin [10]. Expression of *V. parahaemolyticus* NorM by *E. coli* KAM3 increased the MIC of norfloxacin from 0.03 to 0.24 $\mu\text{g ml}^{-1}$ [10]. The lack of detection of NorM by library selection could indicate that some *V. parahaemolyticus* proteins may not be optimally expressed or correctly localised for functionality. Alternatively, a protein which is toxic to the heterologous host may be encoded in the region of NorM, thereby removing *norM* containing clones from the population. Increased retention of inserts and functional expression may be achieved by reducing the insert size, employing inducible promoter driven expression or using stable single copy fosmid-based systems [27]. These alterations would however result in decreased likelihood of functional expression of multi-protein complexes, and may result in lower phenotypic resolution due to decreased heterologous protein concentrations, in the case of fosmid-based systems.

Overall this study describes the preparation of a cosmid-based genomic library, which offers an appropriate insert size, genome coverage, lack of bias and functional expression of heterologous proteins. A similar pWEB-based *Photobacterium asymbiotica* genomic library with 37 kb inserts was successfully employed for the detection of genes, which contribute to virulence during invertebrate infection [16]. A number of macromolecular structures with roles in pathogenesis were detected, including type-three secretion systems and type-IV pili. As an invertebrate model of infection has been developed for *V. parahaemolyticus* [28], the functional screening of a genomic library such as that described herein may allow the identification of novel virulence factors. While our data have revealed limitations in terms of the lack of universal heterologous expression, in comparison to more sensitive

and high-throughput functional genome-mining approaches such as TIS, our cosmid-based library provides a simple, inexpensive and rapid means of screening for gain of function phenotypes in a heterologous host.

Funding information

This work was supported by Science Foundation Ireland (SFI), grant no. 08/RFP/GEN1225.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

- Nair GB, Ramamurthy T, Bhattacharya SK, Dutta B, Takeda Y, *et al*. Global dissemination of *Vibrio parahaemolyticus* serotype O3:K6 and its serovariants. *Clin Microbiol Rev* 2007;20:39–48.
- O'Boyle N, Boyd A. Manipulation of intestinal epithelial cell function by the cell contact-dependent type III secretion systems of *Vibrio parahaemolyticus*. *Front Cell Infect Microbiol* 2014;3:114.
- Joseph SW, Colwell RR, Kaper JB. *Vibrio parahaemolyticus* and related halophilic Vibrios. *Crit Rev Microbiol* 1982;10:77–124.
- Chowdhury MAR, Yamanaka H, Miyoshi S, Shinoda S. Ecology and seasonal distribution of *Vibrio parahaemolyticus* in aquatic environments of a temperate region. *FEMS Microbiol Lett* 1990;74:1–9.
- Daniels NA, Shafaie A. A review of pathogenic *Vibrio* infections for clinicians. *Infect Med* 2000;17:665–685.
- Broberg CA, Calder TJ, Orth K. *Vibrio parahaemolyticus* cell biology and pathogenicity determinants. *Microbes Infect* 2011;13:992–1001.
- O'Boyle N, Houeix B, Kilcoyne M, Joshi L, Boyd A. The MSHA pilus of *Vibrio parahaemolyticus* has lectin functionality and enables TTSS-mediated pathogenicity. *Int J Med Microbiol* 2013;303:563–573.
- Makino K, Oshima K, Kurokawa K, Yokoyama K, Uda T, *et al*. Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V cholerae*. *Lancet* 2003;361:743–749.
- Hubbard TP, Chao MC, Abel S, Blondel CJ, Abel Zur Wiesch P, *et al*. Genetic analysis of *Vibrio parahaemolyticus* intestinal colonization. *Proc Natl Acad Sci U S A* 2016;113:6283–6288.
- Morita Y, Kodama K, Shiota S, Mine T, Kataoka A, *et al*. NorM, a putative multidrug efflux protein, of *Vibrio parahaemolyticus* and its homolog in *Escherichia coli*. *Antimicrob Agents Chemother* 1998;42:1778–1782.
- Poirel L, Liard A, Rodriguez-Martinez J-M, Nordmann P. Vibriaceae as a possible source of Qnr-like quinolone resistance determinants. *J Antimicrob Chemother* 2005;56:1118–1121.
- Saga T, Kaku M, Onodera Y, Yamachika S, Sato K, *et al*. *Vibrio parahaemolyticus* chromosomal qnr homologue VPA0095: demonstration by transformation with a mutated gene of its potential to reduce quinolone susceptibility in *Escherichia coli*. *Antimicrob Agents Chemother* 2005;49:2144–2145.
- Davidson T, Beck E, Ganapathy A, Montgomery R, Zafar N, *et al*. The comprehensive microbial resource. *Nucleic Acids Res* 2010;38:D340–5.
- Yu NY, Wagner JR, Laird MR, Melli G, Rey S, *et al*. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 2010;26:1608–1615.
- Punta M, Coghill PC, Eberhardt RY, Mistry J, Tate J, *et al*. The Pfam protein families database. *Nucleic Acids Res* 2012;40:D290–301.
- Waterfield NR, Sanchez-Contreras M, Eleftherianos I, Dowling A, Yang G, *et al*. Rapid Virulence Annotation (RVA): identification of virulence factors using a bacterial genome library and multiple invertebrate hosts. *Proc Natl Acad Sci U S A* 2008;105:15967–15972.
- Mehmood MA, Gai Y, Zhuang Q, Wang F, Xiao X, *et al*. *Aeromonas caviae* CB101 contains four chitinases encoded by a single gene *chi1*. *Mol Biotechnol* 2010;44:213–220.
- Battle SE, Meyer F, Rello J, Kung VL, Hauser AR. Hybrid pathogenicity island PAGI-5 contributes to the highly virulent phenotype of a *Pseudomonas aeruginosa* isolate in mammals. *J Bacteriol* 2008;190:7130–7140.
- Yi C, Chen J, Hu D, Song B. First report about the screening, characterization, and fosmid library construction of *Xanthomonas oryzae* pv. *oryzae* strain with resistance to Fubianezuofeng. *Pestic Biochem Physiol* 2020;169:104645.
- Charlop-Powers Z, Banik JJ, Owen JG, Craig JW, Brady SF. Selective enrichment of environmental DNA libraries for genes encoding nonribosomal peptides and polyketides by phosphopantetheine transferase-dependent complementation of siderophore biosynthesis. *ACS Chem Biol* 2013;8:138–143.
- Pang H, Zhang P, Duan C-J, Mo X-C, Tang J-L, *et al*. Identification of cellulase genes from the metagenomes of compost soils and functional characterization of one novel endoglucanase. *Curr Microbiol* 2009;58:404–408.
- Cheng G, Hu Y, Yin Y, Yang X, Xiang C, *et al*. Functional screening of antibiotic resistance genes from human gut microbiota reveals a novel gene fusion. *FEMS Microbiol Lett* 2012;336:11–16.
- Feng Y, Duan C-J, Pang H, Mo X-C, Wu C-F, *et al*. Cloning and identification of novel cellulase genes from uncultured microorganisms in rabbit cecum and characterization of the expressed cellulases. *Appl Microbiol Biotechnol* 2007;75:319–328.
- Finn RD, Bateman A, Clements J, Coghill P, Eberhardt RY, *et al*. Pfam: the protein families database. *Nucleic Acids Res* 2014;42:D222–30.
- Piddock LJV. Multidrug-resistance efflux pumps - not just for resistance. *Nat Rev Microbiol* 2006;4:629–636.
- Tran JH, Jacoby GA, Hooper DC. Interaction of the plasmid-encoded quinolone resistance protein Qnr with *Escherichia coli* DNA gyrase. *Antimicrob Agents Chemother* 2005;49:118–125.
- Shashikant CS, Carr JL, Bhargava J, Bentley KL, Ruddle FH. Recombinogenic targeting: a new approach to genomic analysis—a review. *Gene* 1998;223:9–20.
- Krachler AM, Ham H, Orth K. Outer membrane adhesion factor multivalent adhesion molecule 7 initiates host cell binding during infection by gram-negative pathogens. *Proc Natl Acad Sci U S A* 2011;108:11614–11619.