Supplementary Methods, Results, Figures, Tables and References for Hernandez *et al.* Detecting emerging bacterial pathogens

Sample collection and processing

Eight sampling sites on the Fal estuary and English Channel coast near Falmouth (Cornwall, United Kingdom) were selected for water and sediment sampling. Four sites were estuarine (Mylor, Flushing, Penryn and Falmouth Harbour) and four truly coastal (Castle Beach, Gyllyngvase Beach, Swanpool Beach, and Maenporth Beach) (Figure S1). Samples from each site were collected on 21 June 2017 and 6 July 2017. Each sampling effort consisted of collecting water (2*50 mL) and upper 1cm layer of sediment (~25 g) from all eight sites within a time span of two hours around low-tide. Samples were collected using sterile 50mL centrifuge tubes. All samples were kept on ice during transport and processed in laboratory within an hour from collection.

Each water sample was centrifuged at 500 rpm at 4°C for 15 minutes (2 x 50 mL) to pellet sediment particles. The supernatants were then transferred to new sterile 50mL centrifuge tubes and spun down at 3500 rpm at 4°C for 30 minutes to pellet bacteria. The supernatant was discarded and the pellet in each tube resuspended in 500μL of sterile M9 buffer. Resuspensions from the two tubes per sample were then combined in a 1.5 mL Eppendorf tube to obtain a total concentrated water sample of 1mL. Bacteria were extracted from sediments by adding 10g of sediment with 10mL of sterile M9 buffer. Samples were vortexed at 3000 rpm for 2 minutes, and then centrifuged at 500 rpm at 4°C for 15 minutes to pellet sediment particles. The supernatant was then centrifuged as above and resuspended in. 1mL of M9 buffer. All samples were kept at 4°C or on ice at all times.

Flow cytometry

Concentrated water and sediment wash samples were analysed for bacterial load with the use of flow cytometry. For each sample (200 μL sample + 20 μL 2.5x SybrGold), 10 μL was analyzed for number of events using low flow rate (14 μL /min) in a BD Accuri C6 Plus flow cytometer (BD

Biosciences, San Jose, CA). The BD Accuri C6 Plus is equipped with a blue and red laser, two light scatter detectors, and four fluorescence detectors. Bacteria cells were distinguished from non-biological particles using 533/30 filter in FL1 (indicative of nucleic acid content), 670 LP in FL3 (indicative of chlorophyll content), and FSC (indicative of cell size). Data were analyzed using the BD Accuri C6 software. A non-dye sample was included to control for auto-fluorescence.

Culturable Bacteria

100μL of each water and sediment wash sample was plated onto both LB agar (Fisher BioReagents) and Coliform agar (Millipore Sigma) and incubated overnight at 37°C. CFU per plate were counted after 24 hr incubation. Colonies on coliform agar were scored as dark-blue ('*Escherichia coli*'), pink ('coliform'), or white/transparent ('other'). Units are reported in CFU/100mL for water samples and CFU/10g for sediment samples.

Galleria mellonella assay for environmental samples

Final instar waxmoth (*Galleria mellonella*) larvae (~220mg each) were purchased from Livefood UK (http://www.livefood.co.uk), stored in the dark at 4°C, and used within two weeks (all experiments utilized larvae from the same batch). A 100μL Hamilton syringe (Sigma-Aldrich Ltd) with 0.3 x 13mm needle (BD Microlance 3) was used to inject larvae with 10μL of sample into the last left proleg. New sterile needles were used for injection of each sample, and the syringe was cleaned between sample injections with 70% ethanol and sterile M9 buffer. All water and sediment wash samples (8 locations x 2 timepoints) were assayed in one experiment using 20 *Galleria* per sample. Three negative controls were used for each experiment: a no-injection control (to control for background larvae mortality), a buffer control in which larvae were injected with 10μL of M9 buffer (to control for impact of physical trauma), and a filter-sterilized sample control in which larvae were injected with filter-sterilized samples (0.22 μm syringe filters, Thermo Fisher) to control for any toxic chemicals present in the sample. Before injection, the *Galleria* larvae were separated into Petri dishes in groups of 20 and anesthetized by placing

on ice for approximately 30 minutes before injection. After injection, Petri dishes were incubated at 37°C and inspected at 24, 48 and 72 hours post-injection to record morbidity and mortality. *Galleria* larvae were considered infected if they expressed dark pigmentation (melanisation) after inoculation. Larvae were scored as dead if they did not respond to touch stimuli by blunt sterile forceps.

Galleria mellonella assay for individual clones

To isolate and identify responsible pathogens, selected samples showing high *Galleria* mortality were inoculated and incubated again as described above. Eight live larvae showing signs of infection (melanisation) were selected at random from each sample group and dissected for haemocoel collection. Before haemocoel collection, the *Galleria* larvae were anesthetized on ice for 30 minutes and the site of dissection was sterilized with 70% ethanol. Sterile micro-scissors were used to remove the last left proleg, and a drop of haemocoel was allowed to exude from the larvae before being collected with a pipette. Approximately 5-15μL of haemocoel was collected per larva and diluted into 500μL of sterile M9 buffer in an Eppendorf tube. Serial dilutions of each sample were plated onto LB- and coliform agar plates and incubated overnight at 37°C. A single clone per sample was picked from these plates and grown for 24 hours in 5mL of LB broth at 37°C at 180 rpm. Overnight cultures were diluted in broth to a turbidity equivalent to a McFarland standard of 0.5 at 625 nm as measured by spectrophotometry (Bibby Scientific Limited, Staffordshire, UK). A serial dilution was plated on LB agar and incubated for 24 hours at 37°C. LB agar plate colony counts were used to calculate CFU/mL. Individual clones were used to inoculate groups of 20 *Galleria* larvae with 10μL of $1x10^2$ CFU, $1x10^4$ CFU, and $1x10^6$ CFU. Larvae were incubated as described in the previous section and morbidity and mortality was recorded hourly after an initial 10-hour period for a period of 37 hours allowing for construction of survival plots and calculation of the 50% lethal dose value (LD_{50}).

Whole Genome Sequencing

DNA isolation, Illumina HiSeq sequencing and basic bioinformatics was performed through the MicrobeNG program in Birmingham, UK. Vials containing beads inoculated with liquid culture were washed with extraction buffer containing lysostaphin and RNase A, incubated for 25 min at 37°C. Proteinase K and RNaseA were added and incubated for 5 min at 65°C. Genomic DNA was purified using an equal volume of SPRI beads and resuspended in EB buffer. DNA was quantified in triplicates with the Quantit dsDNA HS assay in an Eppendorff AF2200 plate reader. Genomic DNA libraries were prepared using Nextera XT Library Prep Kit (Illumina, San Diego, USA) following the manufacturer's protocol with the following modifications: two nanograms of DNA instead of one were used as input, and PCR elongation time was increased to 1 min from 30 seconds. DNA quantification and library preparation were carried out on a Hamilton Microlab STAR automated liquid handling system. Pooled libraries were quantified using the Kapa Biosystems Library Quantification Kit for Illumina on a Roche light cycler 96 qPCR machine. Libraries were sequenced on the Illumina HiSeq using a 250bp paired end protocol.

Bioinformatics

Reads were adapter trimmed using Trimmomatic 0.30 with a sliding window quality cutoff of Q15 (Bolger et al 2014). De novo assembly was performed on samples using SPAdes version 3.7 (Bankevich et al 2012), and contigs were annotated using Prokka 1.11 (Seemann 2014). Kraken was used to indicate the likely species classification of our isolates (Wood and Salzberg 2014). For species with established MLST schemes, the sequence type of each isolate was identified using the Centre for Genomic Epidemiology MLST tool (Larsen et al 2012). Other isolates belonging to the same sequence types were found and accessed using databases Enterobase (Alikhan et al 2018) (*E. coli*) and (Jolley and Maiden 2010) (*P. aeruginosa*). Whole genome alignments were constructed using Progressive Mauve (Darling et al 2010). Percentage similarity was calculated using the number of SNPs found between the query isolate and the reference sequence. In the case of *V. injenesis* where a closed reference genome was not available, nucleotide identity was

confirmed using a core genome alignment to closely related draft genome. Core genomes were constructed using Roary 3.12.0 (Page et al 2015).

Differences in gene content (95% nucleotide identity threshold) between the *V. injenesis* reference strain and the *V. injenesis* clone isolated in this study were also identified using Roary 3.12.0. The presence of PCAMU-SGI1 in the P. mirabilis isolate was assessed using BLAST (Altschul et al 1990). Assemblies were screened for antimicrobial resistance related genes and virulence related genes using ABRicate (https://github.com/tseemann/abricat) using 90% and 75% nucleotide identity, and 80% length identity cut offs. The CARD database was used to identify AMR genes (McArthur et al 2013), and the VFDB virulence factor database was used to find putative virulence factors (Chen et al 2015). BRIG version 0.95 was used to represent circular draft assemblies and indicate the position of AMR and virulence genes on each contig (Alikhan et al 2011). Contigs were ordered by whole genome alignment against a reference genome using progressive mauve (Darling et al 2010). The *E. coli* isolate was aligned against K-12 substr. MG1655 (NC_000913.3), the *P. mirabilis* isolate was aligned against HI4320 (NC_010554.1), and the *P. aeruginosa* against PAO1 (NC_002516.2). The *V. injenensis* genome could not be aligned to a closed reference genome and so the contigs could not be ordered or separated by chromosome.

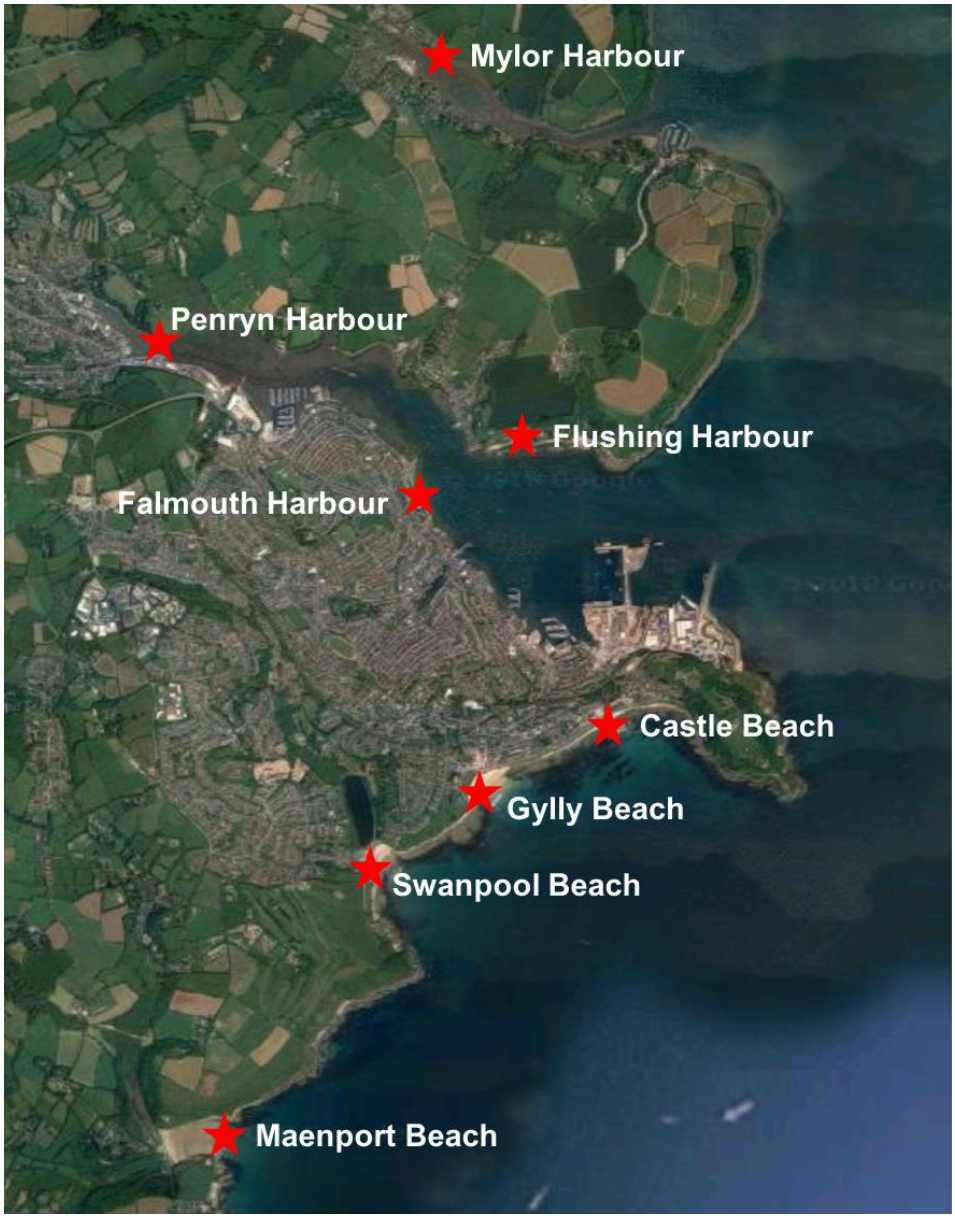
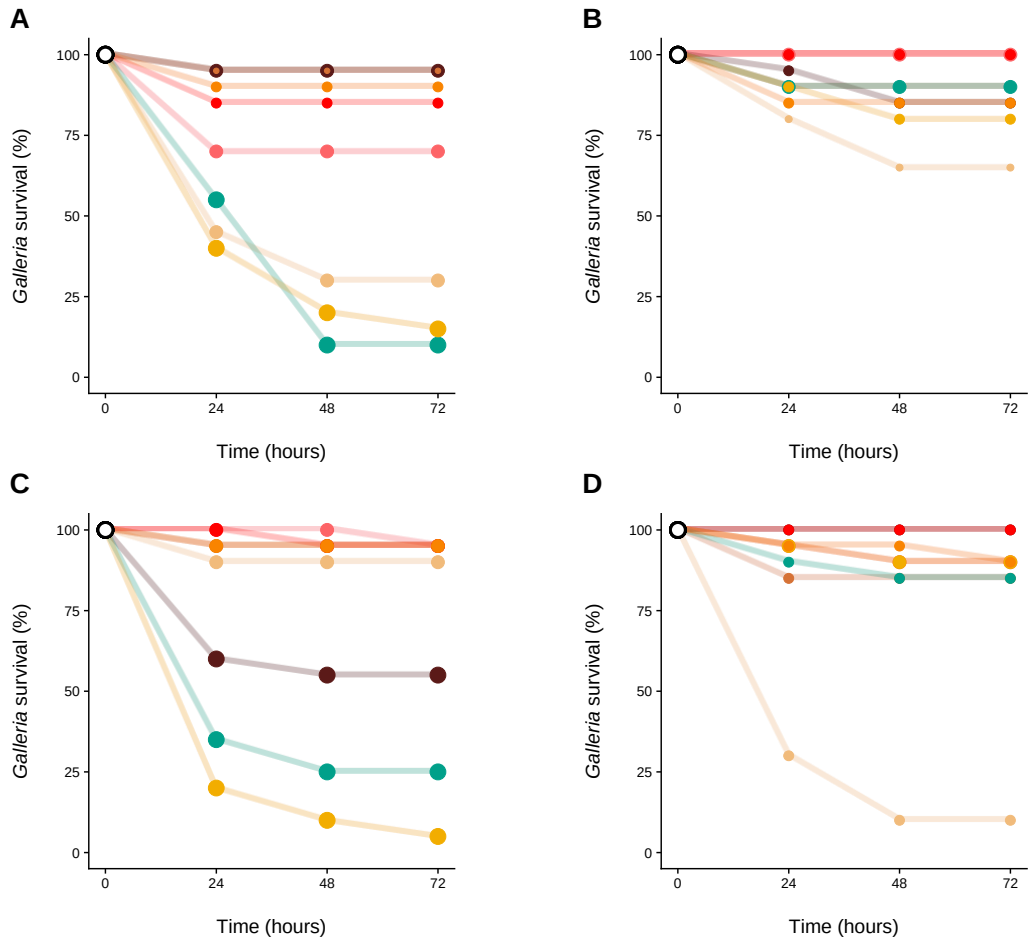


Figure S1: Sampling sites around Falmouth, Cornwall, U.K. (50°.16' N 5°.10' W).

Supplementary Results

Figure S2. *Galleria* survival (%, 20 individuals per group) 24, 48 and 72 hours post-injection with 10µl of 100-fold concentrated water (panel B, 21st June, panel D, 6st July) or sediment wash (see Supplemental Methods) (panel A, 21st June, panel C, 6st July). Different colours denote different localities.



Correlation of virulence with measures of bacterial density

Bacterial density in water and sediment wash samples was quantified in four ways: 1) total density by flow cytometry, 2) total colony counts on LB agar, 3) total colony counts on coliform agar (dark-blue, pink and white/transparent colonies that include gram-negative, non-enteric bacteria) and 4) *E. coli* and coliform colony counts (only dark-blue and pink colonies on coliform agar) (Figure S3). Generalized linear mixed effects models (glmer function from the 'glmm' R package) were used to test the additive effects of bacterial cell density ($\log_{10} + 1$) and sample type (sediment versus water) on *Galleria* mortality with a binomial error structure and random intercepts fitted for sampling date, and location nested within sampling date. Separate models were fitted for each of the four different measures of bacterial cell density. Crucially, we found that *Galleria* mortality was strongly, and positively, dependent on cell density for all different density measures (Figure S4 and Table S1). We also found that *Galleria* mortality was significantly lower (for a given bacterial count) when larvae were injected with water samples compared to sediment samples. This could point at the sediment being inhabited by different, more pathogenic bacteria. The exception to this was formed by estimates based on flow cytometry (Table S1), potentially because culture-based approaches reflect total bacterial density more reliably. There was a strong correlation between coliform cells and total counts on coliform agar with two notable exceptions: Castle Beach water and Penryn sediment (both sampled 6 July); both samples had high mortality, presumably due to these non-coliform bacteria.

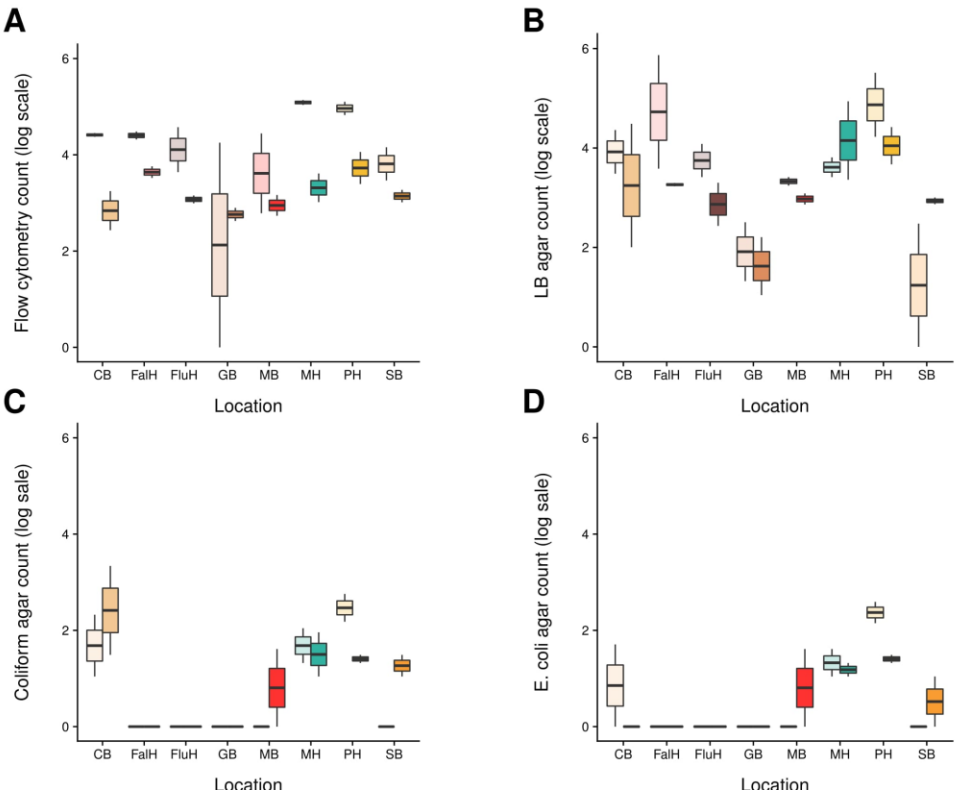


Figure S3. Boxplot summarizing variation in bacterial densities (expressed as $\log_{10}$ cells per 100 ml) across eight different locations (see text) and sample types (water versus sediment, where water samples are depicted first per location). Bacterial cell densities were enumerated using: A) flow cytometry, (B) total LB agar counts, (C) total coliform agar counts and (D) *E. coli* and coliform counts (blue and pink colonies on coliform agar).

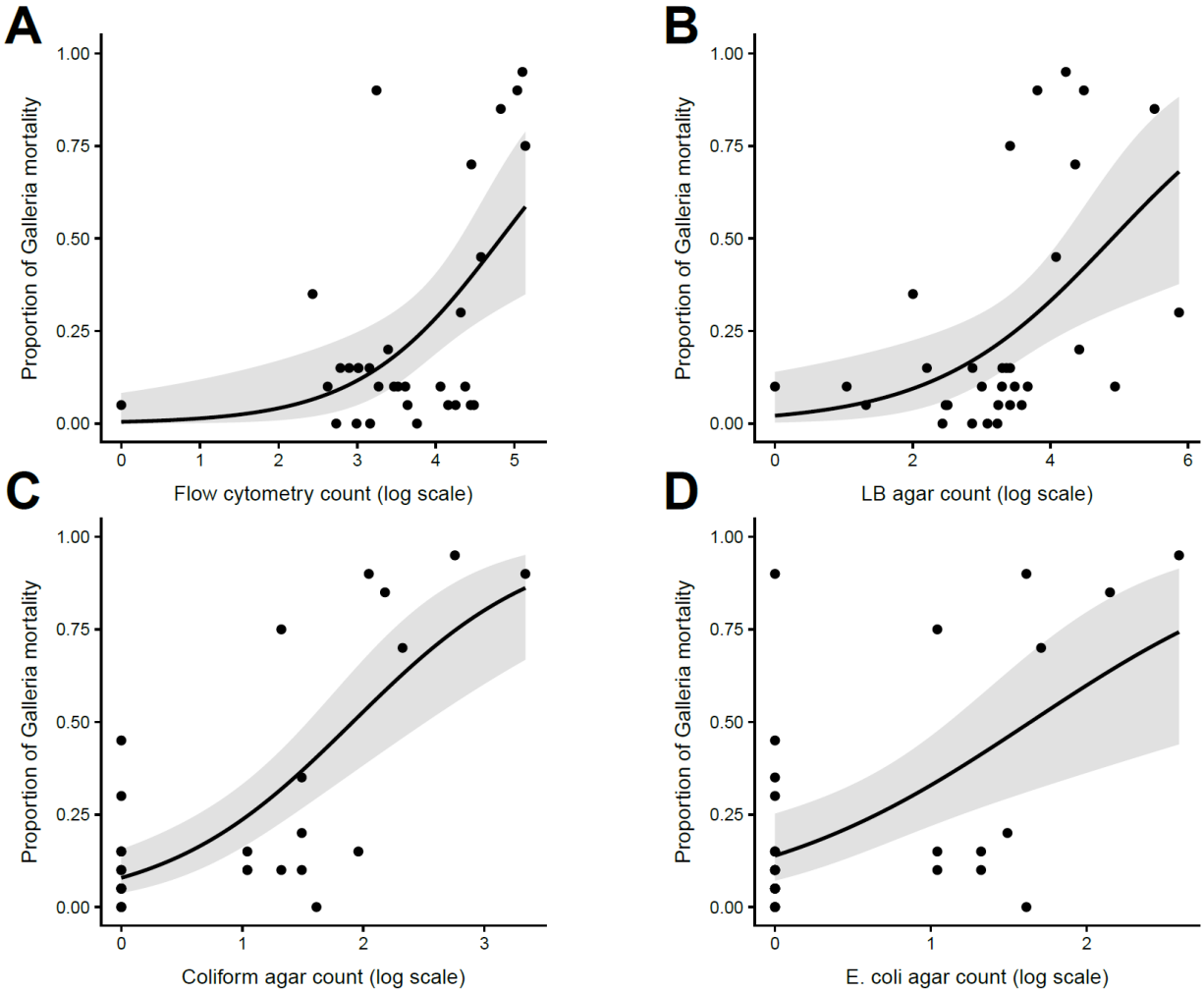


Figure S4. Plots depicting the relationship between *Galleria* mortality and bacterial cell density, estimated using: (**A**) flow cytometry, (B) total LB cell counts, (C) total coliform agar counts and (D) *E. coli* and coliform counts (blue and pink colonies on coliform agar). Lines and shaded area depict the fitted relationships ± standard error (see Table S1 for parameter estimates).

Table S1. Results of the generalized linear mixed effects models (glmer) testing the additive effects of bacterial cell density ($\log_{10}$ + 1) and sample type (sediment versus water) on *Galleria* mortality with a binomial error structure and random intercepts fitted for sampling date, and location nested within sampling data. Separate models were fitted for different estimates of bacterial density (1-4). The most parsimonious model was arrived at by sequentially deleting terms and comparing model fits using χ^2-tests of likelihood ratios.

1. Flow cytometry	Parameter estimate ± SE[§]	χ^2-test
Intercept	-4.02 ± 1.01***, z = -3.99	
Density	0.63 ± 0.23**, *z* = 2.78	$\chi^2_{1,3}$ = 25.34, *p* <0.001
Sample	-0.01 ± 0.34^{NS}, z= -0.04	$\chi^2_{1,4}$= 0.001, *p* = 0.97
Date	Standard deviation <0.001	
Date/Location	Standard deviation = 0.61	
2. LB agar counts		
Intercept	-3.00 ± 0.49***, z = -6.11	
Density	0.47 ± 0.12***, *z* = 3.85	$\chi^2_{1,4}$ = 14.48, *p* < 0.001
Sample	-0.62 ± 0.20**, z= -3.04	$\chi^2_{1,4}$ = 9.34, *p* < 0.01
Date	Standard deviation <0.001	
Date/Location	Standard deviation = 0.62	
3. Total coliform counts		
Intercept	-1.97 ± 0.17***, z = -11.73	
Density	0.78 ± 0.09***, *z* = 8.94	$\chi^2_{1,4}$ = 48.53, *p* < 0.001
Sample	-0.92 ± 0.20***, z= -4.62	$\chi^2_{1,4}$ = 43.39, *p* < 0.001
Date	Standard deviation <0.001	
Date/Location	Standard deviation < 0.001	
4. *E. coli* counts		
Intercept	-1.76 ± 0.23***, z = -7.76	
Density	0.62 ± 0.09***, *z* = 3.82	$\chi^2_{1,4}$ = 11.32, *p* < 0.001
Sample	-0.57 ± 0.20**, z= -2.82	$\chi^2_{1,4}$ = 7.99, *p* < 0.01
Date	Standard deviation <0.001	
Date/Location	Standard deviation = 0.52	

§ The 'Sample' effect in the fitted models provides a measure of the extent to which water and sediment differ in their effect on *Galleria* mortality, where negative values indicate a lower detrimental effect of water on *Galleria* mortality (i.e., smaller estimated intercept) compared to sediment (Intercept). *P* < 0.001 ***, *P* < 0.01 **.

Isolation of pathogenic clones

Individual pathogenic clones were isolated from four samples showing high mortality, genomes of which are described in the main text. The Penryn Harbour sediment (sampled 6 July) yielded a white coloured clone on coliform agar and a 'bubble-wrap' colony morphology on LB agar which was identified as *Proteus mirabilis*. Interestingly, this clone did not cause the usual melanisation in *Galleria* but instead, the larvae turned brown, indicating a different pathogenicity mechanism. Castle Beach sediment (sampled 21 June) yielded a transparent clone on LB agar which was identified as *Vibrio injenensis*. Castle Beach water (sampled 6 July) yielded a bright green clone on LB agar identified as *Pseudomonas aeruginosa* ST-667. Mylor Harbour sediment (sampled 6 July) yielded a dark-blue clone on coliform agar, which was identified as *Escherichia coli* ST-3304.

Tables with virulence- and antibiotic resistance genes found in the four genome-sequenced isolates. Hits tabulated in white are >90% nucleotide similarity (>80% coverage) and hits tabulated in grey are >75% nucleotide similarity ((>80% coverage).

Table S2 Virulence genes detected in *Proteus mirabilis* Penryn Harbour clone.

gene name	**nt identity (%)**	**coverage (%)**	**acc. nr.**	**description**
flhC	76.11	93.99	YP_001006783	flagellar biosynthesis transcription activator

Table S3 Antibiotic resistance genes detected in *Proteus mirabilis* Penryn Harbour clone.

gene name	**nt identity (%)**	**coverage (%)**	**acc. nr.**	**description**
sul1	100	100	JF969163:1054-1894	a sulfonamide resistant dihydropteroate synthase of Gram-negative bacteria linked to other resistance genes of class 1 integrons.
tet(J)	98.16	100	AF038993:1-1198	a tetracycline efflux protein expressed in Gram-negative bacteria (*Escherichia, Morganella and Proteus*).
CRP	80.41	99.84	AP009048:4153664-4154297	a global regulator that represses MdtEF multidrug efflux pump expression.
dfrA15	84.63	100	KF534911:40-514	an integron-encoded dihydrofolate reductase found in *Vibrio cholerae*
plasmid-encoded_cat_(pp-cat)	79.33	94.4	D16171:383-1043	a plasmid-encoded variant of the cat gene found in *Photobacterium damselae subsp. piscicida*

Table S4 Virulence genes detected in *Vibrio injenensis* Castle Beach clone.

*rtxC**	96.97	100	NP_231093	RTX toxin activating protein
hcp-2	83.65	99.81	NP_232418	type VI secretion system substrate
luxS	75.94	82.47	NP_230208	S-ribosylhomocysteinase
rtxB	82.17	99.49	NP_231091	RTX toxin transporter
rtxD	79.69	99.93	NP_231090	RTX toxin transporter
vipA/mglA	75	99.8	NP_232508	type VI secretion system tubule-forming protein
vipB/mglB	79.53	96.75	NP_232509	type VI secretion system tubule-forming protein
vscN2	77.37	99.92	NP_800848	type III secretion system ATPase

*= The *rtxA* gene which is the exotoxin is present at 88% similarity but in two parts (coverage 74.62% and 24.52%) suggesting it might be inactive. The Castle Beach clone carries a zonula occludens toxin (Zot) known to increase mammalian gut permeability (74% amino acid similarity to the fish pathogen *V. anguillarum* (Castillo et al 2017) which is not present in the *V. injenensis* Type Strain.

Table S5 Antibiotic resistance genes detected in *Vibrio injenesis* Castle Beach clone.

gene name	nt identity (%)	coverage (%)	acc. nr.	description
tet34	75.05	100	AB061440:306-771	causes the activation of Mg2+-dependent purine nucleotide synthesis which protects the protein synthesis pathway.
CRP	79.84	99.53	AP009048:4153664-4154297	a global regulator that represses MdtEF multidrug efflux pump expression.

Table S6 Virulence genes detected in *Pseudomonas aeruginosa* Castle Beach clone.

gene name	nt identity (%)	coverage (%)	acc. nr.	description
alg44	99.66	100	NP_252232	alginate biosynthesis protein *Alg8*
alg8	99.8	100	NP_252231	alginate-c5-mannuronan-epimerase *AlgG*
algA	98.82	100	NP_252241	phosphomannose isomerase / guanosine 5'-diphospho-D-mannose pyrophosphorylase
algB	99.19	100	NP_254170	two-component response regulator
algC	99.14	100	NP_254009	phosphomannomutase
algD	99.24	100	NP_252230	GDP-mannose 6-dehydrogenase
algE	99.66	100	NP_252234	alginate biosynthetic protein AlgK precursor
algF	99.69	100	NP_252240	alginate o-acetyltransferase
algG	99.63	100	NP_252235	outer membrane protein AlgE
algI	99.49	100	NP_252238	alginate o-acetyltransferase
algJ	99.15	100	NP_252239	alginate o-acetyltransferase
algK	99.72	100	NP_252233	alginate biosynthesis protein Alg44
algL	99.28	100	NP_252237	poly(beta-d-mannuronate) lyase precursor
algP/algR3	97.17	98.87	NP_253940	alginate regulatory protein
algQ	99.17	100	NP_253942	alginate regulatory protein
algR	99.6	100	NP_253948	alginate biosynthesis regulatory protein AlgR
algU	99.14	100	NP_249453	alginate biosynthesis protein AlgZ/FimS
algW	98.72	100	NP_253136	AlgW protein [Alginate regulation]
algX	99.16	100	NP_252236	alginate biosynthesis protein
algZ	99.35	100	NP_253949	sigma factor AlgU [Alginate (VF0091)] [Pseudomonas aeruginosa PAO1]
aprA	99.24	100	NP_249940	alkaline metalloproteinase precursor

chpA	98.66	99.92	NP_249104	still frameshift probable component of chemotactic signal transduction system
chpB	98.64	100	NP_249105	probable methylesterase
chpC	98.82	100	NP_249106	probable chemotaxis protein
chpD	99.12	100	NP_249107	probable transcriptional regulator
chpE	98.86	100	NP_249108	probable chemotaxis protein
clpV1	98.93	100	NP_248780	type VI secretion system AAA+ family ATPase
dotU1	99.33	100	NP_248768	type VI secretion system protein
exoT	99.42	100	NP_248734	type III secretion system effector
exoU	99.42	100	AAC16023	type III secretion system effector
exsA	98.69	100	NP_250404	type III secretion system regulatory protein
exsB	98.07	100	NP_250403	type III secretion system piolitin
exsC	97.49	100	NP_250401	type III secretion system regulatory protein
exsD	99.16	100	NP_250405	type III secretion system regulatory protein
exsE	96.75	100	NP_250402	type III secretion system regulatory protein
fha1	96.19	96.99	NP_248771	type VI secretion system forkhead-associated protein
fimT	97.06	100	NP_253239	type 4 fimbrial biogenesis protein
fimU	99.61	100	NP_253240	type 4 fimbrial biogenesis protein
fimV	98.27	100	NP_251805	putative Type IV pili related protein
fleI/flag	99.73	100	NP_249784	flagellar protein
fleN	99.29	100	NP_250145	flagellar synthesis regulator FleN
fleP	99.33	100	NP_249787	flagellar protein FliT [Deoxyhexose linking sugar 209 Da capping structure (AI138)]
fleQ	99.19	100	NP_249788	transcriptional regulator
fleR	98.94	100	NP_249790	two-component response regulator
fleS	99.09	100	NP_249789	two-component sensor
flgA	99.28	100	NP_252040	flagellar basal body P-ring biosynthesis protein
flgB	99.75	100	NP_249768	flagellar basal body rod protein
flgC	100	100	NP_249769	flagellar basal-body rod protein
flgD	99.16	100	NP_249770	flagellar basal-body rod modification protein
flgE	98.56	100	NP_249771	flagellar hook protein
flgF	99.07	100	NP_249772	flagellar basal-body rod protein
flgG	99.24	100	NP_249773	flagellar basal-body rod protein
flgH	98.85	100	NP_249774	flagellar L-ring protein precursor
flgI	99.46	100	NP_249775	flagellar P-ring protein precursor
flgJ	99.25	100	NP_249776	flagellar rod assembly protein/muramidase
flgK	99.12	100	NP_249777	flagellar hook-associated protein 1

flgL	99.32	100	NP_249778	flagellar hook-associated protein 3
flgM	98.46	100	NP_252041	negative regulator of flagellin synthesis
flgN	98.73	100	NP_252042	flagella synthesis protein
flhA	99.01	100	NP_250143	flagellar biosynthesis protein
flhB	98.94	100	NP_250140	flagellar biosynthetic protein
flhF	98.92	100	NP_250144	flagellar biosynthesis protein
fliA	99.46	100	NP_250146	flagellar biosynthesis sigma factor
fliC	99.52	100	NP_249783	B-type flagellin
fliD	98.95	100	NP_249785	flagellar capping protein
fliE	99.7	100	NP_249791	flagellar hook-basal body complex protein
fliF	99.56	100	NP_249792	flagellar M-ring protein
fliG	99.7	100	NP_249793	flagellar motor switch protein G
fliH	98.39	100	NP_249794	flagellar assembly protein H
fliI	99.26	100	NP_249795	flagellum-specific ATP synthase
fliJ	99.78	100	NP_249796	flagellar protein
fliK	97.2	100	NP_250132	flagellar hook-length control protein
fliL	99.62	100	NP_250133	flagellar basal body protein
fliM	99.9	100	NP_250134	flagellar motor switch protein
fliN	99.58	100	NP_250135	flagellar motor switch protein
fliO	99.56	100	NP_250136	flagellar protein FliO
fliP	99.09	100	NP_250137	flagellar biosynthetic protein
fliQ	99.63	100	NP_250138	flagellar biosynthetic protein
fliR	98.58	99.36	NP_250139	flagellar biosynthetic protein
fliS	99.47	100	NP_249786	flagellar protein
fptA	98.84	100	NP_252911	Fe(III)-pyochelin receptor precursor
hcp1	100	100	NP_248775	type VI secretion system substrate
hsiA1	98.65	100	NP_248772	type VI secretion system hcp secretion island protein
hsiB1/vipA	100	100	NP_248773	type VI secretion system tubule-forming protein
hsiC1/vipB	99.6	100	NP_248774	type VI secretion system tubule-forming protein
hsiE1	99.05	100	NP_248776	type VI secretion system hcp secretion island protein interacting with HsiB1 to form a novel subcomplex of the T6SS
hsiF1	99.22	100	NP_248777	type VI secretion system hcp secretion island protein, a gp25-like protein but not exhibit lysozyme activity
hsiG1	99.03	100	NP_248778	type VI secretion system hcp secretion island protein
hsiH1	98.57	100	NP_248779	type VI secretion system hcp secretion island protein
hsiJ1	99.33	100	NP_248769	type VI secretion system hcp secretion island protein
icmF1/tssM1	98.91	100	NP_248767	type VI secretion system protein

lasA	97.85	100	NP_250562	LasA protease precursor
lasB	98.4	100	NP_252413	elastase
lasI	99.17	100	NP_250123	autoinducer synthesis protein
lip1	99.57	100	NP_248770	lipoprotein
mbtH-like	100	100	NP_251102	MbtH-like protein from the pyoverdine cluster
motA	97.89	100	NP_253641	flagellar motor protein [Deoxyhexose linking sugar 209 Da capping structure (AI138)]
motB	99.14	100	NP_253640	flagellar motor protein [Deoxyhexose linking sugar 209 Da capping structure (AI138)]
motC	98.65	100	NP_250151	flagellar motor protein [Deoxyhexose linking sugar 209 Da capping structure (AI138)]
motD	98.43	100	NP_250152	flagellar motor protein [Deoxyhexose linking sugar 209 Da capping structure (AI138)]
motY	99.59	100	NP_252216	probable outer membrane protein precursor [Deoxyhexose linking sugar 209 Da capping structure (AI138)
mucA	99.14	100	NP_249454	alkaline metalloproteinase precursor [Alginate (VF0091)
mucB	98.95	100	NP_249455	anti-sigma factor MucA inhibitor of alg gene expression
mucC	99.78	100	NP_249456	negative regulator for alginate biosynthesis MucB [Alginate (VF0091)]
mucD	98.88	100	NP_249457	serine protease MucD precursor [Alginate regulation (CVF523)]
mucE	98.15	100	NP_252722	small envelope protein
mucP	99.04	100	NP_252339	metalloprotease protease
pchA	98.95	100	NP_252921	salicylate biosynthesis isochorismate synthase
pchB	99.67	100	NP_252920	salicylate biosynthesis protein
pchC	99.34	100	NP_252919	pyochelin biosynthetic protein
pchD	98.78	100	NP_252918	pyochelin biosynthesis protein
pchE	98.54	100	NP_252916	dihydroaeruginoic acid synthetase
pchF	98.45	100	NP_252915	pyochelin synthetase
pchG	99.14	100	NP_252914	pyochelin biosynthetic protein
pchH	98.66	100	NP_252913	ABC transporter ATP-binding protein [
pchI	98.9	100	NP_252912	ABC transporter ATP-binding protein [Pyochelin
pchR	98.99	100	NP_252917	transcriptional regulator
pcr1	98.92	100	NP_250390	type III secretion system protein
pcr2	98.92	100	NP_250391	type III secretion system protein
pcr3	98.36	100	NP_250392	type III secretion system protein
pcr4	98.79	100	NP_250393	type III secretion system protein
pcrD	98.96	100	NP_250394	type III secretion system protein
pcrG	99.66	100	NP_250396	type III secretion system cytoplasmic regulator

pcrH	99.21	100	NP_250398	type III secretion system regulatory protein
pcrR	99.31	100	NP_250395	type III secretion system regulatory protein
pcrV	98.98	100	NP_250397	type III secretion system hydrophilic translocator needle tip protein
phzA1	96.16	90.59	NP_252899	phenazine biosynthesis protein
phzA1	99.39	100	NP_252899	phenazine biosynthesis protein
phzB1	98.36	100	NP_252900	phenazine biosynthesis protein
phzC1	99.18	100	NP_252901	phenazine biosynthesis protein
phzD1	98.88	100	NP_252902	phenazine biosynthesis protein
phzE1	99.1	100	NP_252903	phenazine biosynthesis protein
phzF1	99.16	100	NP_252904	phenazine biosynthesis protein
phzG1	99.53	100	NP_252906	phenazine biosynthesis protein
phzH	98.91	100	NP_248741	phenazine-modifying enzyme
phzM	99.4	100	NP_252898	phenazine-specific methyltransferase
phzS	98.35	100	NP_252907	flavin dependent hydroxylase
pilB	99	100	NP_253216	type 4 fimbrial biogenesis protein
pilE	99.06	100	NP_253246	type 4 fimbrial biogenesis protein
pilF	98.95	100	NP_252494	type 4 fimbrial biogenesis protein
pilG	100	100	NP_249099	twitching motility protein
pilH	99.18	100	NP_249100	twitching motility protein
pilI	100	100	NP_249101	twitching motility protein
pilJ	99.51	100	NP_249102	twitching motility protein
pilK	99.2	100	NP_249103	methyltransferase PilK
pilM	99.81	100	NP_253731	type IV pilus inner membrane platform protein
pilN	99.5	100	NP_253730	type IV pilus inner membrane platform protein
pilO	99.52	100	NP_253729	type IV pilus inner membrane platform protein
pilP	99.24	100	NP_253728	type IV pilus biogenesis protein
pilQ	99.07	100	NP_253727	type 4 fimbrial biogenesis protein
pilR	99.18	100	NP_253237	two-component response regulator PilR
pilS	99.18	100	NP_253236	two-component sensor
pilT	99.61	100	NP_249086	twitching motility protein
pilU	99.74	100	NP_249087	twitching motility protein
pilV	99.1	100	NP_253241	type IV pilus biogenesis protein
pilW	98.91	100	NP_253242	type IV fimbrial biogenesis protein
pilX	99.32	100	NP_253243	type 4 fimbrial biogenesis protein
pilY1	93.35	99.43	NP_253244	type 4 fimbrial biogenesis protein
pilY2	96.26	100	NP_253245	type 4 fimbrial biogenesis protein
plcH	96.08	100	NP_249535	hemolytic phospholipase C precursor

popB	99.15	100	NP_250399	type III secretion system hydrophobic translocator pore protein
popD	97.41	100	NP_250400	type III secretion system hydrophobic translocator pore protein
popN	99.42	100	NP_250389	type III secretion system outer membrane protein
ppkA	99.23	100	NP_248764	serine/threonine protein kinase
pppA	99.45	100	NP_248765	Pseudomonas protein phosphatase
pscB	99.05	100	NP_250406	type III secretion system protein
pscC	99.5	100	NP_250407	type III secretion system secretin
pscD	99	100	NP_250408	type III secretion system basal body protein
pscE	97.06	100	NP_250409	type III secretion system cochaperone
pscF	99.22	100	NP_250410	type III secretion system needle filament protein
pscG	99.71	100	NP_250411	type III secretion system chaperone PscG
pscH	100	100	NP_250412	type III secretion system protein PscH
pscI	99.11	100	NP_250413	type III secretion system inner rod protein PscI
pscJ	99.33	100	NP_250414	type III secretion system inner MS ring protein
pscK	98.09	99.04	NP_250415	type III secretion system protein
pscL	98.61	100	NP_250416	type III secretion systemt protein
pscN	99.24	100	NP_250388	type III secretion system ATPase
pscO	99.37	100	NP_250387	type III secretion system protein
pscP	94.64	96.49	NP_250386	type III secretion system protein
pscQ	98.82	100	NP_250385	type III secretion system protein
pscR	99.85	100	NP_250384	type III secretion system protein
pscS	99.25	100	NP_250383	type III secretion system protein
pscT	99.24	100	NP_250382	type III secretion system protein
pscU	98.76	100	NP_250381	type III secretion system protein
ptxR	98.08	99.79	NP_250948	transcriptional regulator
pvcA	98.68	100	NP_250944	paerucumarin biosynthesis protein
pvcB	96.98	100	NP_250945	paerucumarin biosynthesis protein
pvcC	99.27	100	NP_250946	paerucumarin biosynthesis protein
pvcD	97.99	99.85	NP_250947	paerucumarin biosynthesis protein
pvdG	98.69	100	NP_251115	pyoverdine biosynthesis protein
pvdH	99.15	100	NP_251103	diaminobutyrate-2-oxoglutarate aminotransferase
pvdL	98.8	100	NP_251114	peptide synthase [pyoverdine (IA001)]
pvdM	94.44	99.78	NP_251083	dipeptidase precursor [pyoverdine (IA001)] [Pseudomonas aeruginosa PAO1]
pvdN	99.45	100	NP_251084	pyoverdine biosynthesis protein
pvdO	97.78	100	NP_251085	pyoverdine biosynthesis protein
pvdQ	99.13	100	NP_251075	3-oxo-C12-homoserine lactone acylase

pvdS	97.34	100	NP_251116	extracytoplasmic-function sigma-70 factor
rhlA	99.1	100	NP_252169	rhamnosyltransferase chain A
rhlB	99.61	100	NP_252168	rhamnosyltransferase chain B
rhlC	97.03	100	NP_249821	rhamnosyltransferase 2
rhlI	100	100	NP_252166	autoinducer synthesis protein RhlL [Quorum sensing]
tagF/pppB	98.68	100	NP_248766	Pseudomonas protein phosphatase
tagQ	98.25	100	NP_248760	type VI secretion associated protein
tagR	98.95	100	NP_248761	type IV secretion associated protein
tagS	99.25	100	NP_248762	type IV secretion associated protein
tagT	98.89	100	NP_248763	type six secretion associated protein
toxA	99.01	100	NP_249839	exotoxin A precursor
tse1	99.57	100	NP_250535	type VI secretion system effector Tse1 peptidoglycanhydrolase
tse2	98.53	100	NP_251392	type VI secretion system effector Tse2 [HSI-1 (Hcp-secretion island 1)
tse3	98.94	100	NP_252174	type VI secretion system effector Tse3 glycoside hydrolase
vgrG1a	98.45	100	NP_248781	type VI secretion system substrate
waaA	99.53	100	NP_253675	lipopolysaccharide core biosynthesis protein WaaP
waaC	98.88	100	NP_253698	3-deoxy-D-manno-octulosonic-acid (KDO) transferase
waaF	99.61	100	NP_253699	heptosyltransferase I
waaG	99.2	100	NP_253697	B-band O-antigen polymerase
waaP	99.5	100	NP_253696	UDP-glucose:(heptosyl) LPS alpha 13-glucosyltransferase
xcpA/pilD	99.65	99.31	NP_253218	type 4 prepilin peptidase PilD [Type IV pili
xcpP	98.16	100	NP_251794	secretion protein
xcpQ	97.62	100	NP_251795	general secretion pathway protein D [xcp secretion system (VF0084)]
xcpR	98.01	100	NP_251793	general secretion pathway protein E [xcp secretion system (VF0084)]
xcpS	99.02	100	NP_251792	general secretion pathway protein F [xcp secretion system (VF0084)]
xcpT	99.11	100	NP_251791	general secretion pathway protein G [xcp secretion system (VF0084)]
xcpU	98.65	100	NP_251790	general secretion pathway protein H [xcp secretion system (VF0084)]
xcpV	98.97	99.49	NP_251789	general secretion pathway protein I [xcp secretion system (VF0084)]
xcpW	98.88	100	NP_251788	general secretion pathway protein J [xcp secretion system (VF0084)]
xcpX	96.51	100	NP_251787	general secretion pathway protein K [xcp secretion system (VF0084)]
xcpY	98.87	100	NP_251786	general secretion pathway protein L [xcp secretion system (VF0084)]
xcpZ	99.24	100	NP_251785	general secretion pathway protein M [xcp secretion system (VF0084)]
phzB1	89.89	88.96	NP_252900	phenazine biosynthesis protein
pilC	76.17	98.31	NP_253217	still frameshift type 4 fimbrial biogenesis protein
pvdA	83.56	100	NP_251076	L-ornithine N5-oxygenase PvdA
pvdF	82.49	100	NP_251086	pyoverdine synthetase F
pvdP	85.04	99.76	NP_251082	tyrosinase required for pyoverdine maturation

Table S7 Antibiotic resistance genes detected in *Pseudomonas aeruginosa* Castle Beach clone.

APH(3')-IIb	98.64	100	X90856:388-1195	a chromosomal-encoded aminoglycoside phosphotransferase
ArmR	96.3	100	NC_002516.2:4165719-4165881	a 53-amino-acid antirepressor allosterically inhibits MexR dimer-DNA binding by occupying a hydrophobic binding cavity within the center of the MexR dimer. ArmR up-regulation and MexR-ArmR complex formation have previously been shown to upregulate MexAB-OprM.
arnA	98.99	100	NC_002516.2:3982021-3984010	modifies lipid A with 4-amino-4-deoxy-L-arabinose (Ara4N) which allows gram-negative bacteria to resist the antimicrobial activity of cationic antimicrobial peptides and antibiotics such as polymyxin.
basS	98.75	100	JQ340365:1-1435	Histidine protein kinase sensor Lipid A modification gene; part of a two-component system involved in polymyxin resistance that senses high extracellular Fe(2+)
bcr-1	98.92	100	CP012901.1:5979157-5980366	Transmembrane protein which expels bicyclomycin from the cell leading to bicyclomycin resistance. Identified in *P. aeruginosa* strains responsible for outbreaks in Brazil often appearing with blaSPM-1 another bicyclomycin resistance gene
catB7	97.66	100	NC_002516:779463-780102	chromosome-encoded variant of the cat gene
fosA	99.51	100	NC002516:1221691-1222099	enzyme that confers resistance to fosfomycin in *Serratia marcescens* by breaking the epoxide ring of the molecule. It depends on the cofactors Manganese (II) and Potassium and uses Glutathione (GSH) as the nucleophilic molecule. In *P. aeruginosa* FosA catalyzes the conjugation of glutathione to carbon-1 of fosfomycin rendering it ineffective as an antibacterial drug.
MexA	99.31	100	NC_002516:472024-473176	membrane fusion protein of the MexAB-OprM multidrug efflux complex
MexB	99.01	100	L11616:1570-4711	inner membrane multidrug exporter of the efflux complex MexAB-OprM

MexC	99.23	100	U57969:295-1459	membrane fusion protein of the MexCD-OprJ multidrug efflux complex
MexD	97.73	99.94	U57969:1486-4618	multidrug inner membrane transporter of the MexCD-OprJ complex
MexE	99.12	100	NC_002516:2808743-2809988	membrane fusion protein of the MexEF-OprN multidrug efflux complex
MexF	99.28	100	NC_002516:2810009-2813198	multidrug inner membrane transporter of the MexEF-OprN complex. mexF corresponds to 2 loci in *P. aeruginosa* PAO1 (mexF/mexB) and 4 loci in *P. aeruginosa* LESB58 (mexD/mexB).
mexG	99.78	100	NC_002516:4705956-4706403	membrane protein required for MexGHI-OpmD efflux activity
mexH	99.1	100	NC_002516:4706410-4707523	membrane fusion protein of the efflux complex MexGHI-OpmD
mexI	99.39	100	NC_002516:4707535-4710625	inner membrane transporter of the efflux complex MexGHI-OpmD
mexJ	98.91	100	NC_002516.2:4119270-4120374	membrane fusion protein of the MexJK multidrug efflux protein
mexK	98.7	100	AE004091.2:4116188-4119266	inner membrane resistance-nodulation-cell division (RND) transporter in the MexJK multidrug efflux protein
mexL	99.53	100	AE004091.2:4120469-4121108	a specific repressor of mexJK transcription and autoregulates its own expression
mexM	99.31	100	AB219523.1:22-1180	membrane fusion protein of the MexMN-OprM multidrug efflux complex
mexN	99.29	100	AB219523.1:1176-4287	inner membrane transporter of the MexMN-OprM multidrug efflux complex
mexP	98.88	100	AB219524.1:23-1181	membrane fusion protein of the MexPQ-OpmE multidrug efflux complex
mexQ	98.89	100	AB219524.1:1177-4339	inner membrane transporter of the multidrug efflux pump MexPQ-OpmE
mexV	98.67	100	AE004091.2:4903466-4904597	membrane fusion protein of the MexVW-OprM multidrug efflux complex
mexW	99.48	100	NC_002516.2:4904647-4907704	RND-type membrane protein of the efflux complex MexVW-OprM
mexX	97.71	100	AB015853:146-1316	membrane fusion protein of the MexXY-OprM multidrug efflux complex
mexY	98.31	99.81	AB015853:1331-4472	RND-type membrane protein of the efflux complex MexXY-OprM
MuxA	98.98	100	NC_002516.2:2854011-2855292	membrane fusion protein component of the efflux pump system MuxABC-OpmB in *P. aeruginosa*

MuxB	99.33	100	NC_002516.2:2850883-2854015	one of the two necessary RND components in the *P. aeruginosa* efflux pump system MuxABC-OpmB.
MuxC	98.94	100	NC_002516.2:2847776-2850887	one of the two necessary RND components of the MuxABC-OpmB efflux pumps system in *P. aeruginosa*
OpmB	99.13	100	NC_002516.2:2846283-2847780	outer membrane efflux protein in *P. aeruginosa* that shows functional cooperation with MuxABC to form the efflux pump system MuxABC-OpmB.
opmD	98.57	99.93	NC_002516:4710621-4712085	outer membrane channel protein of the efflux complex MexGHI-OpmD.
opmE	99.25	100	AB219524.1:4335-5811	outer membrane factor protein that is part of the multidrug efflux pump MexPQ-OpmE.
OpmH	99.45	100	NC_002516.2:5584101-5585550	outer membrane efflux protein required for triclosan-specific efflux pump function.
OprJ	98.61	100	U57969:4623-6063	outer membrane channel component of the MexCD-OprJ multidrug efflux complex.
OprM	99.18	100	NC_002516:476333-477791	outer membrane factor protein found in *P. aeruginosa* and *Burkholderia vietnamiensis*. It is part of the MexAB-OprM MexVW-OprM MexXY-OprM and the AmrAB-OprM complex.
OprN	98.45	100	NC_002516:2813194-2814613	outer membrane channel component of the MexEF-OprN multidrug efflux complex.
OXA-50	98.23	100	AY306130:1-790	beta-lactamase found in *P. aeruginosa*. It confers decreased susceptibility to ampicillin and ticarcillin and interestingly to moxalactam and meropenem in *P. aeruginosa* but not in *E. coli.* Also confers resistance to piperacillin-tazobactam and cephalotin.
PDC-5	99.41	100	FJ666068:1-1195	extended-spectrum beta-lactamase found in *P. aeruginosa*
PmpM	99.16	100	NC_002516.2:1472547-1473981	multidrug efflux pump belonging to the MATE family of *P. aeruginosa*. PmpM is an H+ drug antiporter and is the first reported case of an H+ coupled efflux pump in the MATE family. PmpM confers resistance to fluoroquinolones fradiomycin benzalkonium chloride chlorhexidine gluconate ethidium bromide tetraphenylphosphonium chloride (TPPCl) and rhodamine 6G
Pseudomonas_aeruginosa_CpxR	99.56	100	LT673656.1:1884345-1885023	directly involved in activation of expression of RND efflux pump MexAB-OprM in *P. aeruginosa*. CpxR is required to enhance mexAB-oprM expression and drug resistance in the absence of repressor MexR.

Pseudomonas_aeruginosa_emrE	99.1	100	NC_002516.2:1-334	a small multidrug transporter that functions as a homodimer and that couples the efflux of small polyaromatic cations from the cell with the import of protons down an electrochemical gradient. Confers resistance to tetraphenylphosphonium methyl viologen gentamicin kanamycin and neomycin.
Pseudomonas_aeruginosa_soxR	98.94	100	NC_002516.2:2503425-2503896	SoxR is a redox-sensitive transcriptional activator that induces expression of a small regulon that includes the RND efflux pump-encoding operon mexGHI-opmD. SoxR was shown to be activated by pyocyanin.
TriA	99.22	100	NC_002516.2:177307-178459	membrane protein that is fused to TriB and both are required for the triclosan efflux pump function of TriABC-OpmH in *P. aeruginosa*
TriB	99.25	100	NC_002516.2:178455-179526	membrane protein that is fused to TriA and both are required for the triclosan efflux pump function of TriABC-OpmH in *P. aeruginosa*
TriC	99.18	100	NC_002516.2:179522-182570	resistance nodulation cell division (RND) transporter that is a part of TriABC-OpmH a triclosan-specific efflux protein.
AxyY	78.44	97.74	AFRQ01000061.1:23987-27125	periplasmic adaptor protein of the AxyXY-OprZ efflux pump system in *Achromobacter spp.*

Table S8 Virulence genes detected in *Escherichia coli* Mylor Harbour.

gene name	**nt identity (%)**	**coverage (%)**	**acc. nr.**	**description**
aslA	98.25	100	AAG10151	putative arylsulfatase
chuA	99.74	100	NP_756170	Outer membrane heme/hemoglobin receptor
chuS	99.51	100	NP_756169	heme oxygenase
chuT	99.5	100	NP_756175	periplasmic heme-binding protein
chuU	99.09	100	NP_756179	heme permease protein
chuV	98.38	100	NP_756180	ATP-binding hydrophilic protein
chuW	99.33	100	NP_756176	Putative oxygen independent coproporphyrinogen III oxidase
chuX	100	100	NP_756177	putative heme-binding protein
chuY	100	100	NP_756178	ChuY
entA	98.26	100	NP_752614	23-dihydro-23-dihydroxybenzoate dehydrogenase [Enterobactin (VF0228)]
entB	99.77	100	NP_752613	isochorismatase [Enterobactin (VF0228)]
entC	98.57	100	NP_752611	isochorismate synthase 1 [Enterobactin (VF0228)]

entD	99.87	100	NP_752599	phosphopantetheinyl transferase component of enterobactin synthase multienzyme complex [Enterobactin (VF0228)]
entE	98.39	100	NP_752612	23-dihydroxybenzoate-AMP ligase component of enterobactin synthase multienzyme complex [Enterobactin (VF0228)]
entF	98.7	99.28	NP_752604	enterobactin synthase multienzyme complex component ATP-dependent [Enterobactin (VF0228)]
entS	98.32	100	NP_752609	enterobactin exporter iron-regulated [enterobactin (IA019)]
fdeC	98.28	100	YP_002390132	adhesin
fepA	99.06	100	NP_752600	ferrienterobactin outer membrane transporter [Enterobactin (VF0228)]
fepB	99.48	100	NP_752610	ferrienterobactin ABC transporter periplasmic binding protein [Enterobactin (VF0228)]
fepC	98.9	100	NP_752606	ferrienterobactin ABC transporter ATPase [Enterobactin (VF0228)]
fepD	98.92	100	NP_752608	ferrienterobactin ABC transporter permease [Enterobactin (VF0228)]
fepG	98.49	100	NP_752607	iron-enterobactin ABC transporter permease [Enterobactin (VF0228)]
fes	99.58	100	NP_752602	enterobactin/ferric enterobactin esterase [Enterobactin (IA019)]
fimA	92.08	100	NP_757241	Type-1 fimbrial protein A chain precursor
fimB	99.83	100	NP_757239	Type 1 fimbriae Regulatory protein
fimC	99.04	100	NP_757243	Chaperone protein fimC precursor
fimD	99.01	100	NP_757244	Outer membrane usher protein fimD precursor
fimE	100	100	NP_757240	Type 1 fimbriae Regulatory protein
fimF	99.81	100	NP_757245	FimF protein precursor
fimG	98.81	100	NP_757247	FimG protein precursor
fimH	98.9	100	NP_757248	FimH protein precursor
fimI	99.44	100	NP_757242	Fimbrin-like protein fimI precursor
fyuA	99.8	100	NP_405467	pesticin/yersiniabactin receptor protein
gspM	92.62	92.91	YP_404609	general secretion pathway protein M
ibeA	98.47	99.93	AAF98391	invasion protein
irp1	99.69	100	NP_405471	yersiniabactin biosynthetic protein
irp2	99.44	100	NP_405472	yersiniabactin biosynthetic protein
ompA	99.81	100	AAF37887	outer membrane protein A
set1A	94.38	99.44	YP_006098866	toxin subunit
set1B	98.39	100	YP_006098865	toxin subunit
vat	99.9	100	NP_752330	Haemoglobin protease
yagV/ecpE	98.14	99.74	NP_286006	*E. coli* common pilus chaperone
yagW/ecpD	99.09	100	NP_286007	polymerized tip adhesin of ECP fibers
yagX/ecpC	97.78	100	NP_286008	*E. coli* common pilus usher
yagY/ecpB	97.76	100	NP_286009	*E. coli* common pilus chaperone

yagZ/ecpA	98.3	100	NP_286010	*E. coli* common pilus structural subunit
ybtA	99.58	100	NP_405473	transcriptional regulator
ybtE	99.94	100	NP_405468	yersiniabactin siderophore biosynthetic protein
ybtP	99.61	100	NP_405474	lipoprotein inner membrane ABC-transporter
ybtQ	99.72	100	NP_405475	inner membrane ABC-transporter
ybtS	99.69	100	NP_405477	salicylate synthase Irp9
ybtT	99.75	100	NP_405469	yersiniabactin biosynthetic protein
ybtU	99.73	100	NP_405470	yersiniabactin biosynthetic protein
ybtX	99.69	100	NP_405476	putative signal transducer
ykgK/ecpR	99.32	100	NP_286011	regulator protein
cheW	75.11	90.36	YP_001006779	purine-binding chemotaxis protein
cheY	77.06	99.49	YP_001006774	chemotaxis regulatory protein
csgB	83.59	99.78	NP_460114	minor curlin subunit precursor curli nucleator protein
csgD	81.41	100	NP_460113	DNA-binding transcriptional regulator
csgE	79.29	98.48	NP_460112	curli production assembly/transport protein
csgF	81	99.04	NP_460111	curli production assembly/transport protein
csgG	83.45	100	NP_460110	curli production assembly/transport protein
flgG	75.81	98.93	YP_001006759	flagellar basal-body rod protein
flgH	78.99	80.7	YP_001006758	flagellar L-ring protein precursor
flhA	75.01	99.04	YP_001006770	flagellar biosynthesis protein
flhC	75.05	95.19	YP_001006783	flagellar biosynthesis transcription activator
fliG	78.54	99.5	YP_001006742	flagellar motor switch protein G
fliI	75.74	92.79	YP_001006744	flagellum-specific ATP synthase
fliM	76.33	99.5	YP_001006748	flagellar motor switch protein
fliP	77.43	98.4	YP_001006751	flagellar biosynthetic protein

Table S9 Antibiotic resistance genes detected in *Escherichia coli* Mylor Harbour.

gene name	**nt identity (%)**	**coverage (%)**	**acc. nr.**	**description**
acrB	98.83	100	NC_000913.3:481254-484404	Protein subunit of AcrA-AcrB-TolC multidrug efflux complex. *AcrB* functions as a herterotrimer which forms the inner membrane component and is primarily responsible for substrate recognition and energy transduction by acting as a drug/proton antiporter.
acrD	98.56	100	NC_007779:2586251-2589365	aminoglycoside efflux pump expressed in *E. coli*. Its expression can be induced by indole and is regulated by baeRS and *cpxAR*.
acrE	98.79	100	U00096:3413864-3415022	membrane fusion protein similar to AcrA.

acrF	96.59	100	U00096:3415033-3418138	inner membrane transporter similar to AcrB.
acrS	98.34	100	U00096:3412803-3413466	repressor of the AcrAB efflux complex and is associated with the expression of AcrEF. AcrS is believed to regulate a switch between AcrAB and AcrEF efflux.
bacA	98.17	99.76	U00096:3203310-3204132	gene that recycles undecaprenyl pyrophosphate during cell wall biosynthesis which confers resistance to bacitracin
baeR	97.23	99.72	NC_007779:2166413-2167136	response regulator that promotes the expression of MdtABC and AcrD efflux complexes.
baeS	90.46	100	AP009048:2165013-2166417	sensor kinase in the BaeSR regulatory system. While it phosphorylates BaeR to increase its activity BaeS is not necessary for overexpressed BaeR to confer resistance.
cpxA	98.47	99.56	NC_002695:4903562-4904936	a membrane-localized sensor kinase that is activated by envelope stress. It starts a kinase cascade that activates CpxR which promotes efflux complex expression.
CRP	99.05	100	AP009048:4153664-4154297	global regulator that represses MdtEF multidrug efflux pump expression.
emrA	98.13	100	AP009048:2810083-2811256	membrane fusion protein providing an efflux pathway with EmrB and TolC between the inner and outer membranes of *E. coli*.
emrB	97.99	100	U00096:2812616-2814155	translocase in the emrB -TolC efflux protein in *E. coli*. It recognizes substrates including carbonyl cyanide m-chlorophenylhydrazone (CCCP) nalidixic acid and thioloactomycin.
emrD	97.82	100	GG749185.1:116813-118004	multidrug transporter from the Major Facilitator Superfamily (MFS) primarily found in *E. coli*. EmrD couples efflux of amphipathic compounds with proton import across the plasma membrane.
emrK	97.64	100	D78168:537-1593	membrane fusion protein that is a homolog of EmrA. Together with the inner membrane transporter EmrY and the outer membrane channel TolC it mediates multidrug efflux.
emrR	99.06	100	NC_000913.3:2810770-2811301	negative regulator for the EmrAB-TolC multidrug efflux pump in *E. coli*. Mutations lead to EmrAB-TolC overexpression.
emrY	97.66	100	D78168:1592-3131	multidrug transport that moves substrates across the inner membrane of the Gram-negative *E. coli*. It is a homolog of emrB.
Escherichia_coli_acrA	99.08	100	NC_000913.3:484426-485620	subunit of the AcrAB-TolC multidrug efflux system that in *E. coli*.
Escherichia_coli_ampC	98.06	100	NC_000913.3:4377811-4378945	class C ampC beta-lactamase (cephalosporinase) enzyme described in *E. coli* shown clinically to confer resistance to penicillin-like and cephalosporin-class antibiotics.
Escherichia_coli_mdfA	96.35	100	JQ394987:1-1234	Multidrug efflux pump in *E. coli*. This multidrug efflux system was originally identified as the Cmr/CmlA chloramphenicol exporter.

evgA	99.03	100	NC_002695:3211892-3212507	when phosphorylated is a positive regulator for efflux protein complexes emrKY and mdtEF. While usually phosphorylated in a EvgS dependent manner it can be phosphorylated in the absence of EvgS when overexpressed.
evgS	96.19	100	U00096:2484374-2487968	sensor protein that phosphorylates the regulatory protein EvgA. evgS corresponds to 1 locus in *P. aeruginosa* PAO1 and 1 locus in *P. aeruginosa* LESB58.
gadW	99.86	100	CP015085.1:2551712-2552441	AraC-family regulator that promotes mdtEF expression to confer multidrug resistance. GadW inhibits GadX-dependent activation. GadW clearly represses gadX and in situations where GadX is missing activates gadA and gadBC.
gadX	93.7	100	NC_007779:3974605-3975430	AraC-family regulator that promotes mdtEF expression to confer multidrug resistance.
H-NS	99.28	100	NC_002695:1737554-1737968	histone-like protein involved in global gene regulation in Gram-negative bacteria. It is a repressor of the membrane fusion protein genes acrE mdtE and emrK as well as nearby genes of many RND-type multidrug exporters.
kdpE	95.84	99.26	NC_000913.3:721056-721734	transcriptional activator that is part of the two-component system KdpD/KdpE that is studied for its regulatory role in potassium transport and has been identified as an adaptive regulator involved in the virulence and intracellular survival of pathogenic bacteria. kdpE regulates a range of virulence loci through direct promoter binding.
marA	98.18	100	NC_007779:1621288-1621672	In the presence of antibiotic stress E. coli overexpresses the global activator protein MarA which besides inducing MDR efflux pump AcrAB also down- regulates synthesis of the porin OmpF.
mdtA	95.84	100	U00096:2154016-2155264	membrane fusion protein of the multidrug efflux complex mdtABC.
mdtB	95.81	100	U00096:2155263-2158386	transporter that forms a heteromultimer complex with MdtC to form a multidrug transporter. MdtBC is part of the MdtABC-TolC efflux complex.
mdtC	93.86	100	U00096:2158386-2161464	transporter that forms a heteromultimer complex with MdtB to form a multidrug transporter. MdtBC is part of the MdtABC-TolC efflux complex. In the absence of MdtB MdtC can form a homomultimer complex that results in a functioning efflux complex with a narrower drug specificity. mdtC corresponds to 3 loci in Pseudomonas aeruginosa PAO1 (gene name: muxC/muxB) and 3 loci in Pseudomonas aeruginosa LESB58.
mdtE	98.53	100	AP009048:3980026-3981184	membrane fusion protein of the MdtEF multidrug efflux complex. It shares 70% sequence similarity with AcrA.
mdtF	97.24	100	U00096:3660414-3663528	multidrug inner membrane transporter for the MdtEF-TolC efflux complex.

mdtG	98.37	100	NC_007779:1115841-1117068	also named YceE, appears to be a member of the major facilitator superfamily of transporters and it has been reported when overexpressed to increase fosfomycin and deoxycholate resistances. mdtG is a member of the marA-soxS-rob regulon.
mdtH	98.1	100	U00096:1124118-1125327	Multidrug resistance protein
mdtN	95.83	100	AP009048:4306557-4307589	Multidrug resistance efflux pump. Could be involved in resistance to puromycin acriflavine and tetraphenylarsonium chloride.
mdtO	97.17	100	AP009048:4304506-4306558	Multidrug resistance efflux pump. Could be involved in resistance to puromycin acriflavine and tetraphenylarsonium chloride
mdtP	97.48	100	AP009048:4303043-4304510	Multidrug resistance efflux pump. Could be involved in resistance to puromycin acriflavine and tetraphenylarsonium chloride
msbA	97.77	100	NC_000913.3:966621-968370	multidrug resistance transporter homolog from *E. coli* and belongs to a superfamily of transporters that contain an adenosine triphosphate (ATP) binding cassette (ABC) which is also called a nucleotide-binding domain (NBD). MsbA is a member of the MDR-ABC transporter group by sequence homology. MsbA transports lipid A a major component of the bacterial outer cell membrane and is the only bacterial ABC transporter that is essential for cell viability.
patA	97.03	100	NC_000913.3:3219494-3220874	PatA is an ABC transporter of *Streptococcus pneumoniae* that interacts with PatB to confer fluoroquinolone resistance.
PmrC	91.12	100	AP009048:4338625-4340269	mediates the modification of Lipid A by the addition of 4-amino-4-deoxy-L-arabinose (L-Ara4N) and phosphoethanolamine resulting in a less negative cell membrane and decreased binding of polymyxin B.
pmrE	94.09	100	U00096:2098447-2099614	required for the synthesis and transfer of 4-amino-4-deoxy-L-arabinose (Ara4N) to Lipid A which allows gram-negative bacteria to resist the antimicrobial activity of cationic antimicrobial peptides and antibiotics such as polymyxin
pmrF	97.63	100	U00096:2367071-2368040	required for the synthesis and transfer of 4-amino-4-deoxy-L-arabinose (Ara4N) to Lipid A which allows gram-negative bacteria to resist the antimicrobial activity of cationic antimicrobial peptides and antibiotics such as polymyxin. pmrF corresponds to 1 locus in Pseudomonas aeruginosa PAO1 and 1 locus in Pseudomonas aeruginosa LESB58.
tolC	98.39	100	FJ768952:1-1489	TolC is a protein subunit of many multidrug efflux complexes in Gram negative bacteria. It is an outer membrane efflux protein and is constitutively open. Regulation of efflux activity is often at its periplasmic entrance by other components of the efflux complex.

yojI	97.57	99.94	NC_000913.3:1-1645	mediates resistance to the peptide antibiotic microcin J25 when it is expressed from a multicopy vector. YojI is capable of pumping out microcin molecules. The outer membrane protein TolC in addition to YojI is required for export of microcin J25 out of the cell. Microcin J25 is thus the first known substrate for YojI.
mdtK	75.27	98.74	CP014358.1:2161326-2162751	multidrug and toxic compound extrusions (MATE) transporter conferring resistance to norfloxacin doxorubicin and acriflavine.
vgaC	83.55	100	KU302801:102200-102431	efflux protein expressed in staphylococci that confers resistance to streptogramin A antibiotics and related compounds. It is associated with plasmid DNA.

Supplemental References

Alikhan N-F, Petty NK, Zakour NLB, Beatson SA (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* **12:** 402.

Alikhan N-F, Zhou Z, Sergeant MJ, Achtman M (2018). A genomic overview of the population structure of Salmonella. *Plos Genet* **14:** e1007261.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol* **215:** 403-410.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS *et al* (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19:** 455-477.

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30:** 2114-2120.

Castillo D, Alvise PD, Xu R, Zhang F, Middelboe M, Gram L (2017). Comparative Genome Analyses of *Vibrio anguillarum* Strains Reveal a Link with Pathogenicity Traits. *Msystems* **2:** e00001-00017.

Chen L, Zheng D, Liu B, Yang J, Jin Q (2015). VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res* **44:** D694-D697.

Darling AE, Mau B, Perna NT (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *Plos One* **5:** e11147.

Jolley KA, Maiden MC (2010). BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* **11:** 595.

Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL *et al* (2012). Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol* **50:** 1355-1361.

McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ *et al* (2013). The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* **57:** 3348-3357.

Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT *et al* (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31:** 3691-3693.

Seemann T (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics***:** btu153.

Wood DE, Salzberg SL (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* **15:** R46.