MICROBIOLOGY

MICROBIOLOGY SOCIETY

OPEN ACCESS

A comparative genomic and phenotypic study of *Vibrio cholerae* model strains using hybrid sequencing

Øyvind M. Lorentzen[1,*], Christina Bleis[1] and Sören Abel[1,2,*]

Abstract

Next-generation sequencing methods have become essential for studying bacterial biology and pathogenesis, often depending on high-quality, closed genomes. In this study, we utilized a hybrid sequencing approach to assemble the genome of C6706, a widely used *Vibrio cholerae* model strain. We present a manually curated annotation of the genome, enhancing user accessibility by linking each coding sequence to its counterpart in N16961, the first sequenced *V. cholerae* isolate and a commonly used reference genome. Comparative genomic analysis between *V. cholerae* C6706 and N16961 uncovered multiple genetic differences in genes associated with key biological functions. To determine whether these genetic variations result in phenotypic differences, we compared several phenotypes relevant to *V. cholerae* pathogenicity like genetic stability, acid sensitivity, biofilm formation and motility. Notably, *V. cholerae* N16961 exhibited greater motility and reduced biofilm formation compared to *V. cholerae* C6706. These phenotypic differences appear to be mediated by variations in quorum sensing and cyclic di-GMP signalling pathways between the strains. This study provides valuable insights into the regulation of biofilm formation and motility in *V. cholerae*.

Impact Statement

This study utilizes hybrid sequencing and manually curated annotation to provide a high-quality genome of *V. cholerae* C6706 that serves as a user-friendly tool to study this commonly used model strain. We uncover genotypic and phenotypic variance compared to *V. cholerae* N16961, which is widely used as a reference genome. Notably, distinct biofilm formation and motility patterns between the strains appear linked to differences in quorum sensing and c-di-GMP signalling pathways.

INTRODUCTION

Vibrio cholerae is the enteric pathogen that causes cholera, an acute diarrhoeal disease that can spread explosively and cause devastating outbreaks [1]. Annually, approximately 3 million people are infected, and 95 000 deaths are associated with the disease [2], making this pathogen a serious global health threat.

So far, the world has experienced seven cholera pandemics, spreading the pathogen globally [1, 3]. The most commonly studied strains today all originate from the seventh pandemic and belong to the O1 serotype, biotype El Tor. The El Tor biotype is the currently dominating *V. cholerae* variant, responsible for the ongoing seventh pandemic [1, 4]. Among them are *V. cholerae* N16961, C6706, E7946, A1552 and 2010EL-1786 [5–9]. Even though they belong to the seventh pandemic, they differ in origin and belong to different transmission waves within the seventh pandemic [4]. *V. cholerae* C6706 and A1552 were isolated from

an outbreak in South America in the 1990s, while *V. cholerae* 2010EL-1786 stems from the 2010 Haiti cholera outbreak [5–9]. N16961 and E7946 were isolated in Asia in the 1970s. N16961 was the first sequenced *V. cholerae* isolate and has been used as a reference genome for the seventh pandemic *V. cholerae* El Tor strains [7]. While other *V. cholerae* variants have been sequenced more recently [9–12], including a *V. cholerae* C6706 strain [9–12], the original N16961 annotation is still widely used [13–19].

V. cholerae survives in the aquatic environment year-round in endemic regions and infects humans through contaminated water and food [1, 20]. Throughout its life cycle, the pathogen alternates between the aquatic environment and the human host. To successfully survive these transitions, *V. cholerae* depends on multiple cellular processes [1, 20, 21]. Biofilm formation plays an important role in the life cycle of *V. cholerae*. Growth in biofilms facilitates survival in the aquatic environment and increases the acid tolerance and infectivity of *V. cholerae* [21–25]. Two of the key regulators of biofilm formation in *V. cholerae* are quorum sensing (QS) and the cyclic diguanylate (c-di-GMP) signalling system. QS is a key bacterial cell–cell communication system that governs collective behaviour in bacteria in response to cell population density and composition [19, 26]. Low cell density (LCD) induces biofilm formation through phosphorylation of the response regulator LuxO, which in turn simultaneously represses HapR and activates AphA, which are high-cell- and low-cell-density master QS regulators [19, 27–29]. HapR and AphA repress and stimulate the expression of biofilm genes, respectively. In contrast, high cell density leads to dephosphorylation of LuxO, which results in the repression of AphA and the expression of HapR, thereby inhibiting biofilm formation [19, 27–29]. QS states can differ between different seventh pandemic *V. cholerae* strains. For instance, the N16961 carries a frameshift mutation in HapR, which disrupts the QS signalling pathway and locks the strain into the state it assumes under LCD conditions [7, 30, 31]. One of the mechanisms through which QS represses biofilm formation in *V. cholerae* involves the regulation of c-di-GMP-metabolizing enzymes [28, 32–34]. C-di-GMP is a conserved nucleotide-based second messenger that regulates the biofilm formation and motility in bacteria [35–37]. In short, c-di-GMP upregulates the biofilm formation and inhibits motility. C-di-GMP regulates the biofilm formation and motility through multiple regulatory mechanisms, which include flagellum synthesis and function, production of pilis and adhesins and secretion of extracellular polymeric substances [35–37]. Strikingly, many Gram-negative pathogens with complex life cycles also boast complex c-di-GMP signalling systems [37]. In particular, *V. cholerae* possess more than 60 putative c-di-GMP-metabolizing proteins [35, 36].

One of the first barriers *V. cholerae* encounters in the human host is the acidic environment of the stomach [20, 38]. To overcome acidic conditions, *V. cholerae* employs a response called the 'acid tolerance response' (ATR) [39]. This is a cascade of physiological responses that enable the bacterium to counteract the damaging effects of acid stress. One extensively studied component of ATR is the Cad system (*CadABC*), which maintains the intracellular pH balance by converting H^+ and lysine to cadaverine [20, 39–41]. After passage through the stomach, the pathogen enters and colonizes the small intestine to cause the characteristic symptoms of the disease [1, 20]. To summarize, the pathogenicity of *V. cholerae* depends on multiple physiological processes including acid tolerance, motility, chemotaxis, biofilm formation and QS [1, 19–21].

In this study, we *de novo* assembled the genome of *V. cholerae* C6706 using short- and long-read next-generation sequencing (NGS) technologies. As an effort to improve upon the prior version of the C6706 genome, we also manually curated the genome annotation and linked all ORFs with the corresponding ORF in the N16961 genome. We thereby created an accessible genome with a link to the N16961 reference genome. Finally, we conducted a genetic and phenotypic comparison of *V. cholerae* C6706 and N16961 to reveal the differences between the two strains, thus contributing to the understanding of *V. cholerae* biology.

METHODS

Strains, growth media and oligonucleotides
All bacterial strains used in this work are listed in Table 1. Unless otherwise noted, cultures were grown from single colonies in Lysogeny Broth (Miller) (LB) media (10 g l⁻¹ of tryptone, 10 g l⁻¹ of sodium chloride and 5 g l⁻¹ of yeast extract) at 37 °C with shaking at 700 r.p.m. Media were supplemented with carbenicillin (50 µg ml⁻¹) when *V. cholerae* strains carrying plasmids were grown to maintain the plasmid. For the acid killing assay, the pH of the medium was adjusted to pH 4.6 with a 0.1 M solution of HCl. All oligonucleotides were synthesized by Sigma-Aldrich and purified to remove salts.

Genomic DNA extraction
One millilitre of liquid culture of *V. cholerae* in the exponential phase was pelleted by centrifugation and resuspended in 50 µl of double-distilled sterile H_2O. Afterwards, 600 µl of lysis buffer (SDS 2% and 0.1 M EDTA) was added, and the solution was incubated for 5 min at 80 °C to lyse the cells. After incubation, the solution was left to cool down to room temperature. Three microlitres of RNaseA were then added, and the solution was incubated for 60 min at 37 °C. Afterwards, the solution was cooled down to room temperature again before 200 µl protein precipitation solution (7.5 M ammonium acetate) was added. The solution was incubated on ice for 5 min followed by centrifugation to precipitate proteins. The DNA-containing supernatant was recovered, while the remaining pellet was discarded. Genomic DNA (gDNA) was then precipitated by adding 600 µl of isopropanol and gentle mixing. The precipitated gDNA was isolated from the tube with a glass pipette and washed in 600 µl of 70% EtOH. Lastly,

Table 1. Strains and plasmids used in this study

Strain no.	Strain background	Vector	Insert	References
3479	*V. cholerae* C6706 QS deficient	None	None	[6]
1810	*V. cholerae* N16961	None	None	[7]
3802	*V. cholerae* C6706 QS proficient	None	None	[30]
2832	*V. cholerae* C6706 QS deficient	pANG01	None	This study
3002	*V. cholerae* C6706 QS deficient	pOML27	VC1295-mRuby2	This study
3867	*V. cholerae* N16961	pOML27	VC1295-mRuby2	This study
3868	*V. cholerae* N16961	pANG01	None	This study
2844	*E. coli* DH5αpir	None	None	[85]
2988	*E. coli* DH5αpir	pOML27	VC1295-mRuby2	This study
2831	*E. coli* DH5αpir	pANG01	None	This study

the DNA was centrifuged one more time, and the supernatant was discarded. The remaining DNA pellet was carefully dried at room temperature for 30 min and dissolved overnight in Milli-Q filtered water.

Whole-genome data acquisition, assembly and annotation

The sequencing service was provided by the Norwegian Sequencing Centre (www.sequencing.uio.no), a Norwegian national technology platform hosted by the University of Oslo and supported by the 'Functional Genomics' and 'Infrastructure' programmes of the Research Council of Norway and the Southeastern Regional Health Authorities.

To create the paired-end fragment libraries for Illumina NGS, gDNA was sheared on a Covaris E220 instrument aiming for a 350 bp fragment size. After fragmentation, the sample was transferred to a 96 plate for a half-volume Kapa Hyper library preparation kit (Roche). Unique dual indexing (Illumina UD adaptor plate, Illumina) was used in the ligation reaction, followed by a clean-up and one round of PCR (four cycles) to boost the library amount. The final library was cleaned twice to get rid of any leftover adapters/dimers. The quality was checked on a Fragment Analyzer instrument using a standard NGS kit (AATI). The library was sequenced on an HiSeq 4000 instrument (150 bp paired end reads) and quantified using quantitative PCR (qPCR) (Kapa Library Quantification Kit, Kapa/Roche).

To conduct PacBio SMRT sequencing, gDNA was sheared to 12 kb fragments using g-tubes from Covaris. The library was prepared using Pacific Biosciences protocol for SMRTbell with PacBio Barcoded Adapters for Multiplex SMRT Sequencing. The sample was pooled together with ten other samples in roughly equimolar ratios. The final library was size selected using 0.45× AMPure PB beads. The library was sequenced on a Pacific Biosciences Sequel instrument using Sequel Polymerase v3.0, SMRT cells v3 LR and Sequencing chemistry v3.0. Loading was performed by diffusion. Sequencing was performed on one SMRT cell. Following sequencing, reads were demultiplexed using the barcoding pipeline on SMRT Link with 26 as the minimum barcode score. Finally, HGAP4 assembly was performed using the Assembly (HGAP4) pipeline on SMRT Link. Assembly was run using 4 Mb as the expected genome size.

The final genome was assembled with the hybrid assembler Unicycler with both short and long reads as input sequences [42]. The genome was automatically annotated with the Rapid Annotation and Subsystem Technology (RAST) web tool hosted on PATRIC [43, 44]. Gene annotations in strain C6706 were linked to the previously described *V. cholerae* strains by conducting a BLAST search with every identified coding sequence (CDS) in C6706 against known genomic features in a selection of *V. cholerae* strains [45]. To improve upon the automated annotation of the C6706 genome, all CDS annotated as hypothetical proteins by RAST were subjected to an additional protein BLAST search to identify the gene function through comparison of homologues with an *E*-value <0.05 in other bacterial strains.

Attachment assay

The attachment was used as a proxy for the biofilm formation and determined using a crystal violet (CV) attachment assay as described before but in 24-well plates instead of 96-well plates [46]. Briefly, 2 ml LB medium was inoculated 1 : 100 with an overnight culture of *V. cholerae* and grown statically for 24 h. After 24 h, the bacterial culture was removed, and the plate was gently submerged in filtered water thrice to remove non-adherent cells. Afterwards, plates were dried for 1 h at 55 °C. To quantify the biofilm formation, wells were stained with 2 ml of 0.1% CV for 10 min. After staining, the CV solution was removed, and the plates were again submerged in filtered water to remove non-adherent dye and left to dry at room temperature. The CV was

dissolved from the stained biomass with 70% ethanol. Finally, the dissolved dye concentration was quantified by measuring the absorbance at 595 nm in a Spark multimode plate reader (Tecan).

Motility on semi-solid agar plates

Colony size on semi-solid agar plates was used as a proxy for motility and was assessed as previously described [47]. Briefly, semi-solid agar plates were made with 60 ml of LB media containing 0.3% agarose in 140-mm-diameter Petri dishes. One microlitre of cells grown to the exponential phase in LB medium was stabbed into the surface of the swarmer plates with a sterile inoculation loop. The plates were incubated at 37 °C for 12 h. Afterwards, plates were imaged. To quantify the motility, the images were imported into Fiji, and the diameter of each individual swarm was measured in Fiji with the line and measure function [48].

Variation analysis and genome comparison

The variation analysis was performed with the variation analysis tool hosted by PATRIC [44]. Mutations were called using standard settings with BWA-mem-strict and FreeBayes as aligner and SNP caller, respectively. Mutations called in homopolymeric regions were omitted. In addition, mutations called when comparing our short read sequences of *V. cholerae* C6706 against the our assembled C6706 genome (BioProject accession PRJNA1109855) (Table S1) were omitted. Genome comparison was conducted with progressiveMauve [49].

Fluctuation assays

Wild-type *V. cholerae* C6706 and N16961 were grown overnight. A dilution series of the overnight cultures were created and plated onto an LB plate with or without 50 µg ml^{-1} rifampicin. After overnight growth at 37 °C, the rifampicin-resistant and the total number (resistant and susceptible) of colonies were counted. The mutation rate was calculated with the FALCOR web application using the Ma–Sandri–Sarkar maximum likelihood estimator (MSS-MLE) method [50]. The Student's t-test was used to determine the significance between the observed number of mutational events in C6706 and N16961 based on nine biological replicates [51].

Acid killing assay

Survival in acidic media was determined as previously described [39]. Briefly, two 100 ml LB cultures were inoculated 1 : 100 with a starter culture of *V. cholerae* C6706 and N16961, respectively. The two cultures were grown to the exponential phase and subsequently diluted to OD_{600} of 0.0125. These were used as inoculum into acidic media (LB adjusted to pH 4.6 with 0.1 M HCl). Samples were taken out at time points 0 min, 25 min, 40 min and 60 min, and c.f.u. values were enumerated by plating on LB agar. The experiment was performed in biological triplicates.

Bacterial fitness measurements

To start the growth curve experiments, overnight cultures were diluted 1 : 100 into 300 µl LB medium and pipetted into a 100-well honeycomb plates (Oy Growth Curves Ab Ltd., Finland). Growth curves were recorded by measuring OD at 600 nm (OD_{600}) every 4 min for 24 h at 37 °C with continuous shaking in a Bioscreen C instrument (Oy Growth Curves Ab Ltd., Finland). Bacterial fitness was quantified as the area under the curve (AUC) in the growth curves with the flux package in R [52]. Each biological replicate is always based on three technical replicates.

Cloning of VC1295 for the expression and fluorescent microscopy

All primers used in this study are listed in Table 2. *VC1295* was amplified by PCR with its native promotor from gDNA extracted from *V. cholerae* N16961 using primers #153 and #154. The PCR was conducted with Phusion polymerase (New England Biolabs). Afterwards, *VC1295* with its native promotor was cloned onto the pENTRY vector pMaRo1 through a Gibson assembly reaction,

Table 2. Primers used in this study

No.	Name	Sequence (5′ to 3′)	Ref.
3	P0003_pEntry_F	gatctcgggccccaaataat	[54]
4	P0004_pEntry_R	gcagctggatggcaaataat	[54] This study
153	P0153_VC1295_F	TTTTATAATGCCAACTTTGTACAAAAAAGCAGGCT*gtatcttaacagtattccttgataca*	This study
154	P0154_VC1295_R	TCTTATAATGCCAACTTTGTACAAGAAAGCTGGGT*tgcggcatctttaaagtgtgct*	This study
159	P0159_pANG_Ins_F	cagtgccaacatagtaagccag	This study
160	P0160_pANG_Ins_R	gaagcatttatcagggttattgtctc	This study

which was conducted at 50 °C for 60 min [53]. The correct sequence was confirmed by Sanger sequencing by Macrogen Europe using primers #3 and #4.

The final vector construct carrying *VC1295-mRuby2* under the control of its native promotor was assembled by a Gateway cloning reaction between the pENTRY vector containing *VC1295* and a pDESTINATION vector carrying *mRuby2* [54, 55]. The correct assembly was confirmed by PCR using primers #159 and #160. This vector was then conjugated into *V. cholerae* via the intermediate strain *Escherichia coli* SM10(λpir) [56].

V. cholerae strain carrying a plasmid expressing *VC1295*-mRuby2 was grown to the exponential phase. Upon reaching the exponential phase, 2 µl of the suspension was spotted on a 1% agarose patched and imaged with a DeltaVision Elite (GE Healthcare) inverted microscope with a DeltaVision CMOS camera and a UPlanFLN 100× PH NA 1.30 phase-contrast objective (Olympus). The images were processed with softWoRx (GE Healthcare), Fiji and Photoshop CS6 (Adobe) to find a focused slide, crop the area of interest and adjust the levels [48].

RESULTS

Building a circularized *V. cholerae* C6706 genome

We created a closed circularized genome for *V. cholerae* C6706 with a hybrid sequencing approach that combines both long- and short-read sequencing technology (Fig. 1) [42]. Long-read sequencing with PacBio sequencing yielded 504131 subreads with a mean length of 6801 bps generating a total of 1.89×10^9 bases. The assembly yielded two contigs totalling 4090295 bp, with an average read depth of 444×. Short-read sequencing using Illumina technology generated a total of 15841698 reads, with an average length of 151 bps, resulting in an average read depth of 585×. The resulting genome consists of two closed, circular chromosomes of 3019923 and 1070364 bps, respectively. Genome quality evaluation hosted by PATRIC [44] confirmed the assembled genome's completeness at 100%, with a minimal contamination of 0.6%. Additionally, the genome sequence is of high quality, with coarse and fine consistency values reaching 99.9 and 99.5%, respectively (Table 3) [57]. Automatic annotation with the RAST, hosted by PATRIC, yielded 3788 CDS [43, 44]. To increase the usability, every CDS in the C6706 genome has also been linked to the corresponding CDS in *V. cholerae* N16961 by searching for homologues by BLAST (Table S2).

V. cholerae C6706 carries multiple genetic changes compared to N16961

The *V. cholerae* El Tor variant C6706 is a commonly used strain by researchers and represents the wild-type in many studies [13, 17, 18, 58, 59]. For a long time, only the genome of *V. cholerae* N16961, another El Tor isolate, was available [7], and researchers have used and are still using it as the de facto reference genome [13–18] despite the potential differences between the El Tor variants [4, 9, 10]. To identify the genetic differences between *V. cholerae* C6706 and N16961, we conducted a genome comparison and a variant analysis. The paired-read library of Illumina short-read sequences from our hybrid sequencing of *V. cholerae* C6706 was used as input. The updated closed genome of *V. cholerae* N16961 (GenBank accession numbers: LT906614 and LT906615; PATRIC Genome ID: 243277.254) was used as a reference genome. This identified 76 differences, which included 6 deletions, 3 insertions and 67 point mutations. Of the 67 point mutations, there were 44 nonsynonymous mutations and 23 synonymous mutations (Table S3). A genome comparison of *V. cholerae* C6706 and N16961 showcased that C6706 also contained the West African–South American (WASA)-1 prophage specific to the WASA clade of *V. cholerae* isolates, which are not present in N16961 [4, 60]. Similar to what has been observed in other *V. cholerae* strains, chromosome 1 contained a large inversion (Fig. S1, available in the online version of this article) [61, 62]. A comparison of the previously sequenced *V. cholerae* C6706 showcased that the inversion varied between assemblies (Fig. S2). Finally, C6706 contained a different variant of the *Vibrio* seventh pandemic island II (VSP-2), which differs from N16961 in the region encompassing VC0511-VC0515 [4, 63, 64]. Overall, the genomes of C6706 and N16961 are very similar except for the two abovementioned mobile genetic elements and inversion in chromosome 1.

Several of the identified mutations in C6706 were located in genes previously linked to important biological functions, such as biofilm formation, motility, QS, genome stability and repair and acid tolerance (Table 4). Among the mutated genes was the previously published point mutation (Gly333Ser) in LuxO, which induces a deficiency in QS that locks *V. cholerae* into the same state; it assumes under LCD conditions [30]. We also found multiple mutations in c-di-GMP-metabolizing genes, potentially affecting the c-di-GMP turnover of *V. cholerae* C6706 relative to N16961. This included a frameshift mutation in VC1295, an HD-GYP-domain protein predicted to degrade c-di-GMP [65]. In addition, we identified mutations in genes coding for proteins important for flagellar function and chemotaxis. Lastly, we identified mutations in CadB, an enzyme crucial for acid adaptation in several enteric pathogens and RecA, which is essential for the global response to DNA damage (SOS response), DNA repair and mutagenicity in bacteria [39, 41, 66–68].

Our genetic comparison of *V. cholerae* C6706 and N16961 identified a selection of mutations likely affecting the physiology and adaptability of these strains. To understand the role of these mutations, we employed phenotypic assays to accurately quantify differences in key bacterial phenotypes between *V. cholerae* C6706 and N16961.
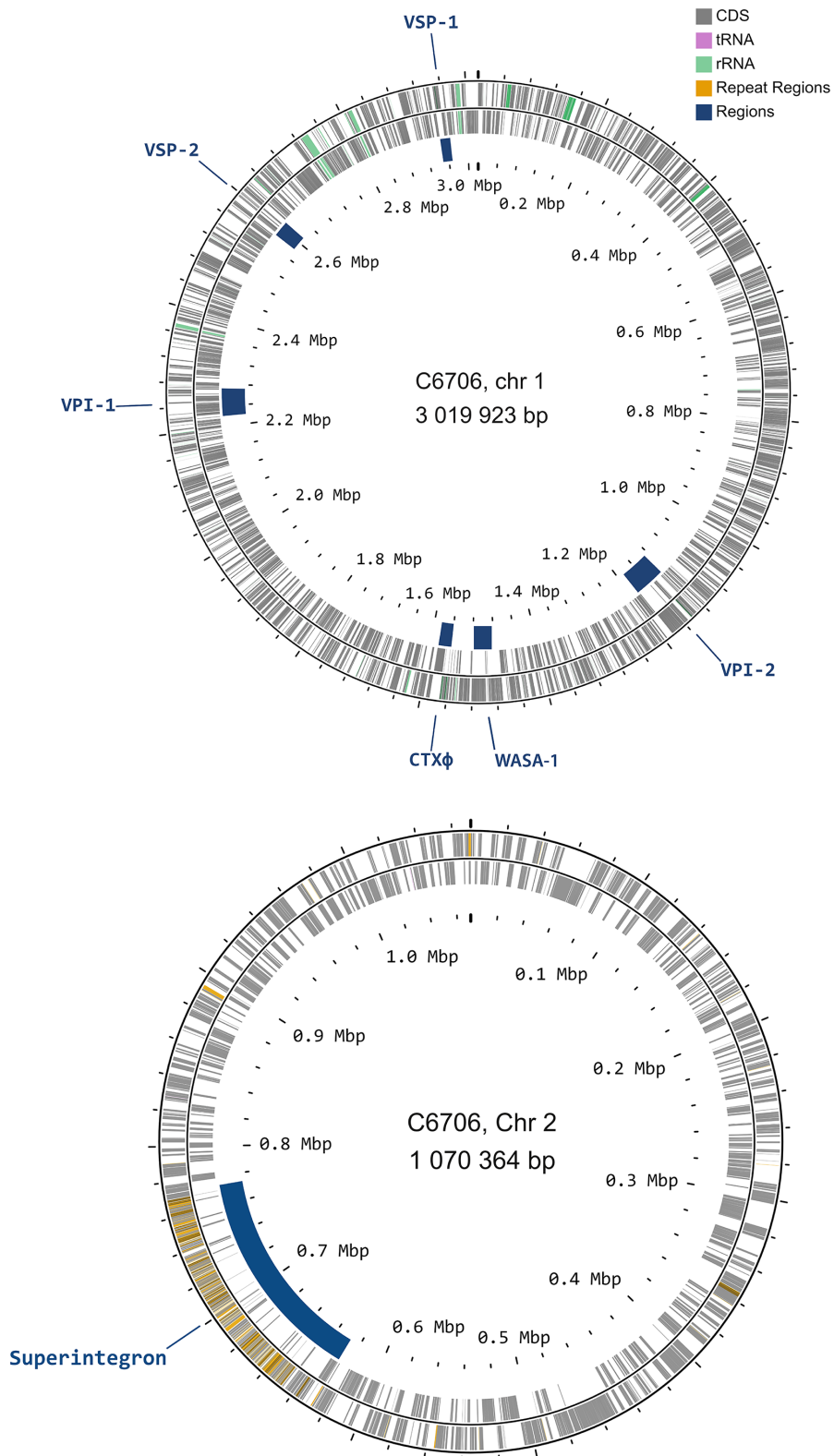
Fig. 1. Maps of chromosome 1 and 2 of *V. cholerae* C6706. Each grey bar represents a coding sequence (CDS), pink bar represents tRNAs, green bars represent rRNAs and orange bars represent repeat regions. Important mobile genetic elements ("regions") are underlined in blue. Maps were created with CGView [86, 87].

Table 3. Sequencing and genome assembly metrics and genome quality control

PacBio sequencing	
Subreads	504131
Subread mean length (bp)	6801
Total reads (bp)	1.89×10^9
Assembly	
2 contigs total (bp)	4090295
Average read depth	444×
Illumina sequencing	
Reads total	15841698
Mean length (bp)	151
Average read depth	585×
RAST	
Completeness (%)	100
Coarse consistency (%)	99.9
Fine consistency (%)	99.5
Contamination (%)	0.6

Mutation rate is increased in *V. cholerae* C6706

The *recA* gene codes for a protein essential for DNA repair and has been shown to affect the mutagenicity of bacteria [66, 67, 69]. The A914G substitution identified in C6706 is a missense mutation that results in a tyrosine-to-cysteine change at position 305 in the C-terminal domain of RecA [70]. To assess whether this mutation affects the mutation rate in *V. cholerae* C6706, we quantified the mutation rate and number of mutations per culture (*m*) with a rifampicin fluctuation assay [71]. The mutation rate and *m* were calculated using the FALCOR web application, employing the MSS-MLE method [49]. The mutation rate was higher in *V. cholerae* C6706 (7.85×10^{-9}) compared to N16961 (5.65×10^{-9}) (Fig. 2a), and we observed a significant increase in mutations per culture (*m*) in *V. cholerae* C6706 (39.3) compared to N16961 (18.2) (*P*=0.0236 in two-tailed t-test comparing the natural logarithm of *m* of C6706 and N1691) (Fig. 2b). We did not identify any additional mutations in genes predicted to be involved in genome stability and repair. This suggests that the observed difference in mutation rates could stem from functional differences in the RecA variants.

V. cholerae C6706 and N16961 have similar tolerance to acidic conditions

To understand the impact of the point mutation in *cadB* (Thr229Ile), we compared the acid tolerance of *V. cholerae* C6706 and N16961. Both strains were exposed to a pH of 4.6 in LB medium, and c.f.u. values were enumerated over time (Fig. 3). The two strains did not grow in pH 4.6. Instead, c.f.u. values declined over time, indicating cell death. There was no significant survival difference between *V. cholerae* C6706 and N16961 in acid at any time point (*P*=0.4–0.9, two-tailed t-test), indicating that the observed mutation in *cadB* does not significantly affect the acid tolerance of *V. cholerae* C6706.

Biofilm formation and motility differ between *V. cholerae* C6706 and N16961

V. cholerae C6706 harboured multiple genetic changes in genes predicted to regulate biofilm formation and motility, including several c-di-GMP-associated genes, *luxO*, and multiple proteins involved in flagellar function and chemotaxis (Table 4). In bacteria, biofilm formation and motility are often inversely regulated, where increasing the biofilm formation leads to decreased motility and vice versa [35, 36, 72, 73]. To investigate if the sum of these genetic changes affected biofilm formation and motility in *V. cholerae* C6706 and N16961, we quantified these phenotypes. As already stated, N16961 contains a non-functional *hapR* due to a frameshift mutation that locks it in an LCD QS state, independent of the actual cell density [7, 30, 31]. This defective QS is known to affect biofilm formation and motility [22, 28]. The strain of *V. cholerae* C6706 sequenced in this study is locked in the same QS state, but due to a different mutation in *luxO* [30]. To compare N16961 to the naïve, QS-proficient C6706, we also included a variant with harbouring wild-type *luxO* [22]. We conducted a variant analysis to identify the genetic differences between QS-proficient C6706 (GenBank accession numbers: CP064350, CP064351; PATRIC Genome ID: 948564.8) sequenced

Table 4. Selected genes with changes from N16961 to C6706

Gene accession no. N16961	Gene accession no. C6706	Nucleotide change	aa change	Gene name	Function	Cellular process
VC0280	VCC2568	C686T	Thr229Ile	*cadB*	Lysine/cadaverine antiporter membrane protein CadB	Acid adaptation
VC0543	VCC2308	A914G	Tyr305Cys	*recA*	RecA protein	Stress response
VC0543	VCC2271	217_218insT*	Val74fs†	*hapR*	QS regulator of virulence	QS
VC0653	VCC2204	T1664C	Val555Ala	*rocS*	Diguanylate cyclase/phosphodiesterase	c-di-GMP signalling
VC1021	VCC1859	G997A	Gly333Ser	*luxO*	LuxO global QS regulator	QS
VC1295	VCC1597/1598	1153delT‡	Phe385fs	N/A	c-di-GMP phosphodiesterase (HD-GYP domain)	c-di-GMP signalling
VC1399	VCC1499	837_838insA	Gly279_ Lys280fs	*cheR*	Chemotaxis protein methyltransferase CheR	Motility/chemotaxis
VC1653	VCC1185	C1675T	His559Tyr	*vieS*	Response regulator VieS	c-di-GMP signalling
VC1967	VCC0880	A277G	Thr93Ala	N/A	Methyl-accepting chemotaxis sensor/transducer protein	Motility/chemotaxis
VC2191	VCC0672	C863T	Pro288Leu	*flgK*	Flagellar hook-associated protein FlgK	Motility/chemotaxis
VC2208	VCC0655	G283A	Ala95Thr	*flgT*	Flagellar protein FlgT	Motility/chemotaxis
VC2338	VCC0536	C1237T	Leu413Phe	*lacZ*	Beta-galactosidase	Metabolism
VCA0557	VCCA0844	A926G	Asp309Gly	N/A	Diguanylate cyclase	c-di-GMP signalling
VCA0931	VCCA0133	G1207A	Ala403Thr	N/A	c-di-GMP phosphodiesterase (HD-GYP domain)	c-di-GMP signalling
VCA1084	VCCA0275	G339A	Met113Ile	*lapB*	Type I secretion system ATPase, LssB family LapB	Type II secretion systems

*'ins' indicates an insertion between the two given positions within a gene.
†'fs' indicates a frameshift after the first aa that is affected by the change.
‡'del' indicates a deletion of the indicated base at the indicated position.

by Weng *et al.* [11] and QS-deficient C6706 (BioProject accession PRJNA1109855) strains, which did not reveal any additional mutations other than the expected G997A in *luxO* (Table S4).

The QS-deficient variant of *V. cholerae* C6706 formed 85 and 53% more biofilm compared to QS-proficient variant of *V. cholerae* C6706 and QS-deficient N16961, respectively (Fig. 4a, one-way ANOVA, $P<0.0001$). In addition, N16961 formed 20% more
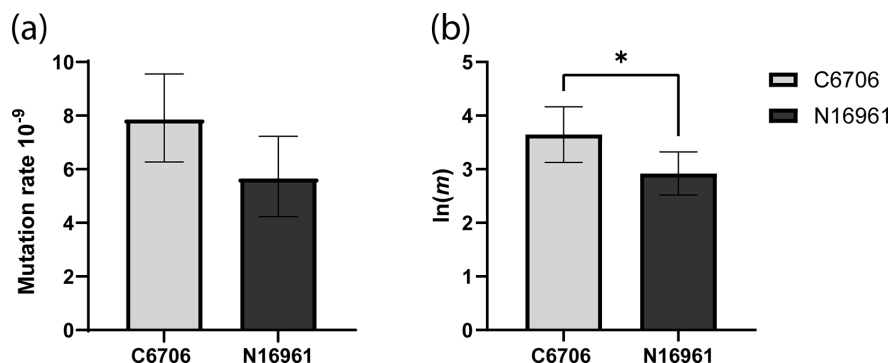

Fig. 2. Comparison of mutation rate and the natural logarithm of number of mutations per culture [ln(m)] of *V. cholerae* C6706 and N16961 measured with a rifampicin fluctuation assay [50]. (**a**) Mutation rate of *V. cholerae* C6706 and N16961. (**b**) Ln(m) of *V. cholerae* C6706 and N16961. Bars depict the means based on nine biological replicates. Error bars represent the 95% confidence intervals. Statistical significance was tested using an unpaired two-tailed T-test, *, $p<0.05$.
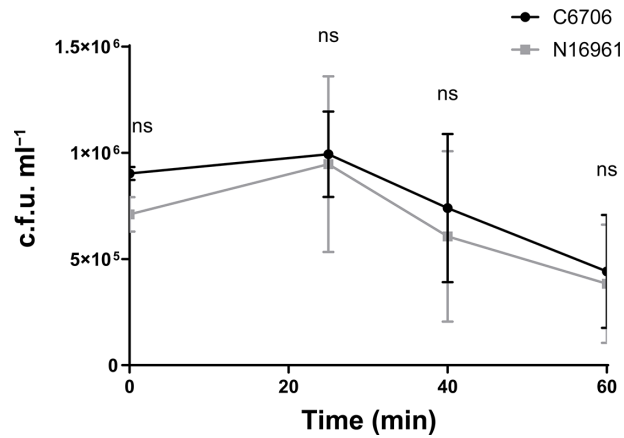
Fig. 3. Both *V. cholerae* strains, C6706 and N16961, are equally sensitive to acid. C.f.u. ml^{-1} measurements over time of *V. cholerae* C6706 and N16961 exposed to pH 4.6 in LB medium. The dots represent the mean of three biologically independent replicates and error bars indicate the standard deviation. Statistical significance was tested using a two-tailed T-test for each individual time point; non-significant (ns), $p>0.05$.

biofilm compared to the QS-proficient variant of C6706 (Fig. 4a, one-way ANOVA, $P<0.01$). The QS-deficient *V. cholerae* C6706 exhibited 14% lower motility compared to *V. cholerae* N16961 (Fig. 4b, one-way ANOVA, $P<0.0001$). In contrast, QS-deficient C6706 exhibited a 12% increase in motility compared to QS-proficient C6706 (Fig. 4b, one-way ANOVA, $P<0.0001$). To ensure that the differences in biofilm formation and motility were not due to large variations in bacterial fitness, we recorded growth curves. Comparisons of these curves revealed small differences in the AUC. Although these differences were statistically significant, the resulting minor changes in relative fitness (QS-deficient C6706:1.00, N16961:0.96, QS-proficient C6706:1.03) are unlikely to account for the substantial differences observed in biofilm formation and motility (Fig. 4c, Table S5).

QS-deficient C6706 and N16961 had inverse behaviour when comparing biofilm formation and motility, where the strain with the highest level of biofilm formation had the lowest motility. In contrast, QS-proficient *V. cholerae* C6706 did not exhibit this pattern and had both the lowest biofilm formation and the lowest motility out of the three strains. The observed differences in biofilm formation and motility between the QS-proficient and QS-deficient variants of *V. cholerae* C6706 are likely due to the differences in QS state, as QS is known to impact biofilm formation and motility [22, 28]. The same explanation could explain the differences between QS proficient and N16961. However, the QS-deficient variant of *V. cholerae* C6706 and N16961 are locked in the same LCD state. Therefore, the difference in biofilm formation and motility between these two strains is likely not mediated by differences in QS state, but caused by additional mutations (e.g. mutations in the c-di-GMP signalling system or flagellar protein).
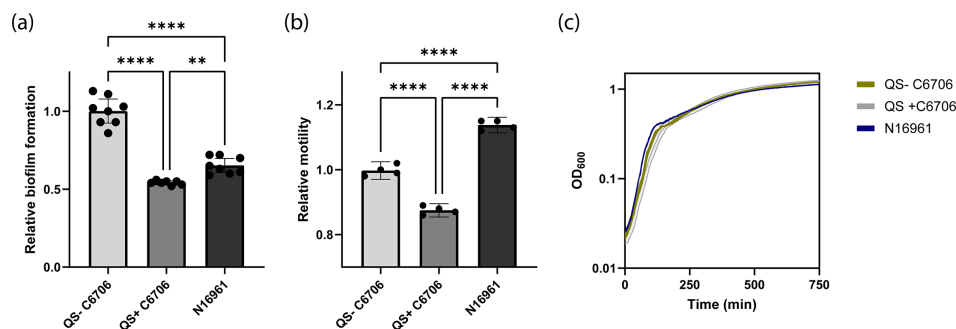

Fig. 4. Comparison of the capacity to form biofilms, motility on semi-solid agar plates, and growth of the naturally QS-deficient N16961, a laboratory acquired QS deficient C6706 (QS- C6706), and a naturally QS proficient C6706 (QS+ C6706). (**a**) Relative biofilm formation normalized to QS- C6706 after static growth at 37 °C for 24 hours in 24-well plates. The bars represent the mean of eight (QS- C6706 and N16961) or seven (QS+ C6706) biological replicates and the error bars indicate the 95% confidence interval. Statistical significance was tested using a one-way ANOVA followed by Tukey's multiple comparisons tests. **, $p<0.01$; ****, $p<0.0001$. (**b**) Relative motility normalized to QS- C6706 after growth on semi-solid agar for 12 hours. The bars represent the mean of four biological replicates and the error bars indicate the 95% confidence interval. Statistical significance was tested using a one-way ANOVA followed by Tukey's multiple comparisons tests. ****, $p<0.0001$. (**c**) Growth curves of N16961, QS-deficient *V. cholerae* C6706, and QS-proficient C6706 measured as OD$_{600}$ over time.
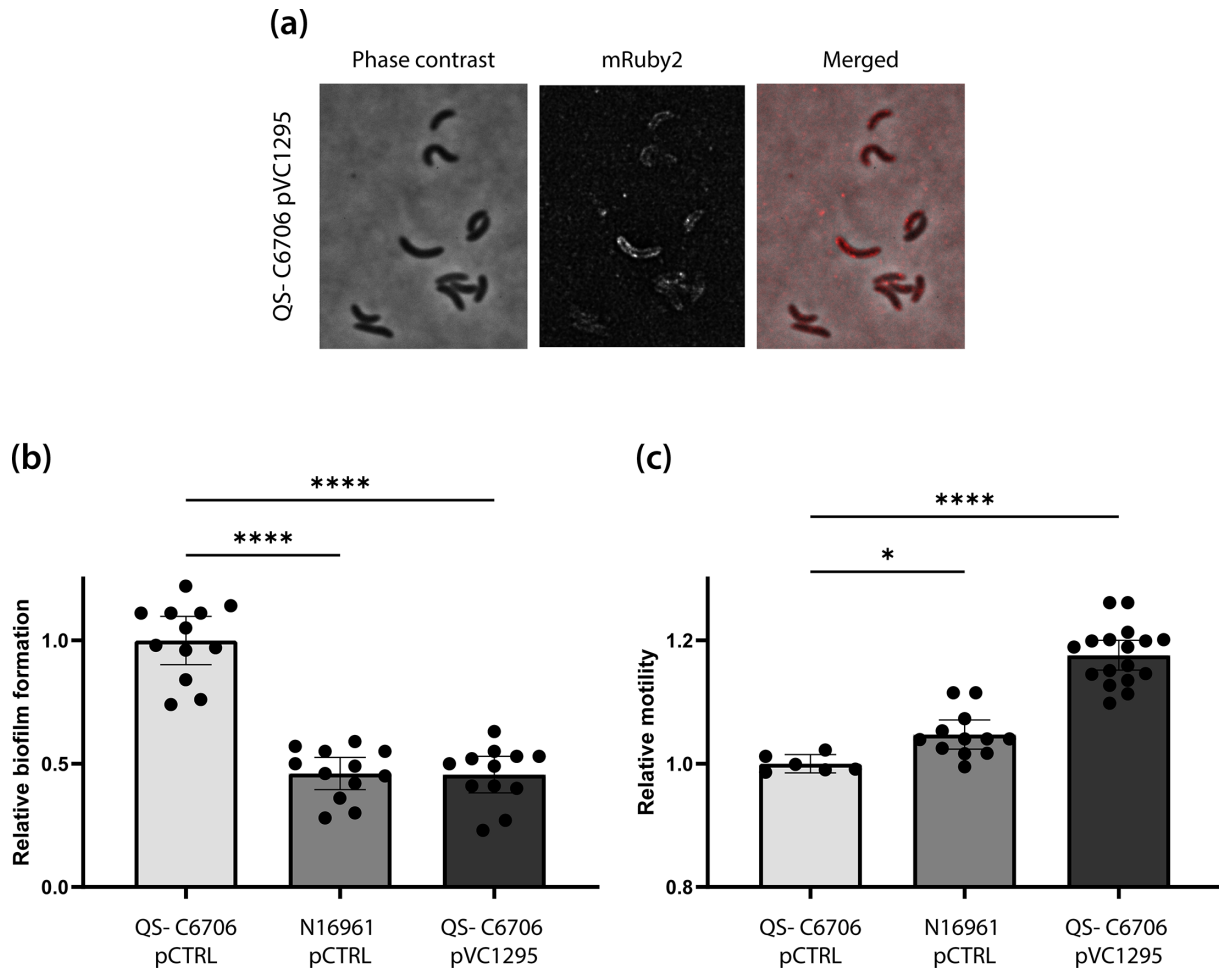
Fig. 5. VC1295 is expressed in *V. cholerae* C6706, inhibits biofilm formation and increases motility. (a) Fluorescent images of *V. cholerae* C6706 expressing VC1295 fused to a C-terminal mRuby2 on a medium-copy number plasmid under the control of its native promoters in the exponential growth phase (QS− C6706 pVC1295). A phase contrast (PC), a corresponding fluorescence (mRuby2) image and an overlay of both images (merged) are shown. (b) Biofilm formation of QS− C6706 pCTRL, N16961 pCTRL and QS− C6706 pVC1295 normalized to QS− C6706 pCTRL after static growth at 37 °C for 24 h in 24-well plates (n=12). (c) Motility of the same strains test in (b) after growth on semi-solid agar at 37 °C for 12 h ($n \geq 6$). The bars represent the mean and the error bars represent the 95% CI. Statistical significance was tested using a one-way ANOVA followed by the Dunnet multiple comparison correction. *$P<0.05$; ****$P<0.0001$.

Altered c-di-GMP signalling and QS shape biofilm and motility in *V. cholerae* strains C6706 vs. N16961

A comparison of N16961, QS-deficient *V. cholerae* C6706 and QS-proficient C6706 indicated that QS was not the sole cause of the observed differences in biofilm formation and motility. We, therefore, wanted to investigate whether variations in the c-di-GMP signalling system, particularly the frameshift mutation in the putative c-di-GMP-degrading protein VC1295, contributed to the observed phenotypic differences. First, we investigated the expression of VC1295 under the control of its native promotor. In agreement with Koestler and Waters [74] and McKee *et al.* [65], fluorescence microscopy demonstrated that VC1295 was expressed and translated in *V. cholerae* C6706 (Fig. 5a). Furthermore, the protein was correctly translocated to its expected subcellular compartment in the cell membrane.

V. cholerae N16961 contains a functional copy of VC1295 and is naturally QS deficient. Therefore, we next investigated the biofilm formation and motility in N16961 harbouring an empty control vector (N16961 pCTRL) and QS-deficient *V. cholerae* C6706 expressing either a plasmid-born VC1295 (QS− C6706 p*VC1295*) or an empty control vector (QS− C6706 pCTRL). Comparing QS-deficient *V. cholerae* C6706 harbouring pVC1295 with N16961 harbouring pCTRL, the two strains displayed identical biofilm formation (Fig. 5b), while QS− C6706 p*VC1295* demonstrated a 12% increase in motility compared to N16961 pCTRL (Fig. 5c). This indicates that the plasmid-borne VC1295 can complement the frameshifted genomic VC1295. Therefore, VC1295 seemed to be expressed and enzymatically active under the tested biofilm and motility assay conditions, and the frameshift mutation in

VC1295 in *V. cholerae* C6706 seems to contribute to the differences in the biofilm formation and motility between C6706 and N16961.

DISCUSSION

In this study, we have assembled a high-quality, carefully annotated genome of *V. cholerae* C6706 and employed it to conduct a genetic and phenotypic comparison of *V. cholerae* C6706 and N16961. The assessment of genome quality in our assembled genome aligned well with the criteria suggested by Parello *et al.* [57] for defining a high-quality genome with a low contamination score (0.6%), a fine consistency score of 99.5% and a completeness score of 100%. A genetic comparison demonstrates how these strains, isolated from different outbreaks in the seventh pandemic, carry mutations that alter their physiology. Indeed, when comparing our assembled C6706 genome with the genome of N16961, we identified multiple mutations in genes with the potential to affect *V. cholerae* biology (Table 4). One set of striking differences was the multiple mutations in enzymes metabolizing c-di-GMP, a well-known regulator of biofilm and motility in bacteria [35–37]. In addition, we observed mutations in genes involved in acid stress, stress response and genome stability, flagellum biosynthesis and chemotaxis. The comparison of our genome to the previously assembled genomes of *V. cholerae* C6706 only identified the known laboratory-acquired mutation in *luxO* [11, 30]. However, a large inversion in chromosome 1 had occurred (Fig. S1 and S2), which seems to occur intermittently in *V. cholerae* seventh pandemic strains, as this has occurred in multiple *V. cholerae* model strains [10, 61, 62]. Importantly, a previous study did not find any fitness costs associated with large inversions in chromosome 1 [61].

To better understand if these mutations really affect the biology of *V. cholerae* C6706, we quantified the selected phenotypes. Firstly, this revealed a difference in mutation rate between the two strains (Fig. 2), which could potentially be attributed to a point mutation (C305Y) in RecA. This is a ubiquitous recombinase conserved throughout the bacterial kingdom, which evolves at a slow speed [70, 75, 76]. The protein is involved in multiple important cellular processes including the SOS response, genome stability and repair and homologous recombination [66–69]. An increased mutation rate can be beneficial in some circumstances as it has been shown to accelerate evolution and increase the rate of adaptation [77–79].

The most striking phenotypic differences between the strains were differences in the biofilm formation and motility (Fig. 4a, b). QS-deficient *V. cholerae* C6706 formed increased biofilms and larger colonies on semi-solid agar plates compared to the QS-proficient C6706. These findings are in line with the current literature that QS inhibits both motility and biofilm formation [22, 28]. Prior studies have demonstrated that QS regulates the biofilm formation through control of multiple c-di-GMP-metabolizing genes [27, 28, 32–34, 80]. Typically, c-di-GMP regulates motility and biofilm formation inversely, i.e. high c-di-GMP concentrations increase biofilms and decrease motility, while low c-di-GMP concentrations do the opposite [35–37]. How QS inhibits both biofilm formation and motility at the same time in *V. cholerae* remains unclear. In certain cases, specific c-di-GMP-metabolizing enzymes asymmetrically regulate both biofilm formation and motility [35, 81]. Consequently, QS might influence c-di-GMP-degrading enzymes that primarily inhibit biofilm formation without necessarily promoting increased motility. Alternatively, additional signalling pathways could also be involved, thereby modulating biofilm formation or motility independently of c-di-GMP.

V. cholerae N16961 exhibited higher motility than both variants of C6706 (Fig. 4b). This is in agreement with a previous study that showed that *V. cholerae* N16961 had increased motility compared to South American *V. cholerae* isolates closely related to C6706 [82]. In addition to increased motility due to QS deficiency, this could come from an altered c-di-GMP metabolism with decreased levels of c-di-GMP in *V. cholerae* N16961, e.g. due to the presence of a functional copy of the putative c-di-GMP-degrading enzyme VC1295. Alternatively, the genetic differences in multiple flagellar and chemotaxis-related proteins could also contribute to the observed difference in motility.

Interestingly, QS-deficient *V. cholerae* C6706 formed more biofilm and had lower motility compared to N16961 even though they should both be locked in the same QS state (Fig. 4a, b) [30]. This indicates that additional signalling pathways other than QS are mediating the observed phenotypic differences. The inverse effect on biofilm formation and motility resembles the effect of increased levels of c-di-GMP [35–37]. While some c-di-GMP-metabolizing enzymes in *V. cholerae* are known to be regulated by QS, many of them are seemingly regulated independently of QS [32, 80, 83]. Therefore, we hypothesized that the differences between QS-deficient C6706 and N16961 could be due to the differences in the strains' QS-independent c-di-GMP signalling systems. Indeed, *V. cholerae* C6706 contains multiple mutations in putative c-di-GMP-metabolizing enzymes (Table 4). This includes a frameshift mutation in VC1295, a functional c-di-GMP-degrading enzyme [65]. Indeed, the expression of a functional copy of VC1295 in QS-deficient C6706 reversed the observed differences in biofilm formation and motility (Fig. 5b, c), indicating that the observed differences were at least partly due to the differences in c-di-GMP metabolism between the strains, although the difference could also be due to additional QS-related signalling pathways as *V. cholerae* C6706 and N16961 contain different mutations, LuxO and HapR, respectively, in the QS signalling pathway. Even though the mutations lead to the same QS state, we cannot exclude that they have additional confounding downstream effects. Altogether, the results of this work are consistent with a model where the observed differences in biofilm formation and motility are due to a combination of effects from QS and c-di-GMP signalling.

V. cholerae N16961 and QS-proficient *V. cholerae* C6706 exhibited differences in the biofilm formation and motility, and the difference in motility exceeded the biofilm formation (Fig. 4a, b). In QS-proficient *V. cholerae* C6706, this might be attributed to QS-mediated repression of biofilm formation [28, 32, 34, 80]. In addition, it harbours a frameshift mutation in the active c-di-GMP-degrading enzyme VC1295 (Table 4) [65]. This leads to putatively increased c-di-GMP levels and increased biofilm formation, which could partially balance out the QS-mediated biofilm repression (Fig. 4a). N16961 lacks QS-mediated biofilm repression but also lacks the frameshift mutation in VC1295. For motility, the QS state and c-di-GMP levels of QS-proficient *V. cholerae* C6706 would both act to repress motility (Fig. 4b) [22, 36, 84]. In contrast, the QS state and putatively lower c-di-GMP levels of *V. cholerae* N16961 would both promote motility (Fig. 4b) [22, 36, 84]. Therefore, the observed differences in biofilm formation and motility are consistent with the identified genetic changes between *V. cholerae* C6706 and N16961 (Table 4).

In conclusion, our study offers a genomic analysis of *V. cholerae* C6706, utilizing a hybrid sequencing approach. This yielded a high-quality genome of *V. cholerae* C6706, which we carefully annotated and cross referenced to N16961. We believe that this will be a valuable resource for the scientific community and represents an improvement of the previous version of the *V. cholerae* C6706 genome [11]. By characterizing genotypic and phenotypic differences between *V. cholerae* C6706 and N16961, we have uncovered the potential targets of adaptive evolution in the seventh cholera pandemic. Furthermore, the comparison of the biofilm formation and motility between *V. cholerae* C6706 and N16961 sheds further light on the complex interplay of factors regulating biofilm formation and motility in *V. cholerae* seventh pandemic strains.

Author contributions
The project was conceptualized by O.M.L. and S.A. Funding was acquired by S.A. The experiments were carried out and the data analysed by O.M.L. and C.B. under the supervision of S.A. The manuscript was drafted by O.M.L. and C.B. and edited and revised by all authors.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References
1. Kanungo S, Azman AS, Ramamurthy T, Deen J, Cholera DS. *Lancet* 2022;399:1429–1440.
2. Ali M, Nelson AR, Lopez AL, Sack DA. Updated global burden of cholera in endemic countries. *PLoS Negl Trop Dis* 2015;9:e0003832.
3. Barua D. History of cholera. In: Barua D and Greenough WB (eds). *Cholera*. Boston, MA: Springer US.pp. 1–36.
4. Mutreja A, Kim DW, Thomson NR, Connor TR, Lee JH, *et al*. Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature* 2011;477:462–465.
5. Yildiz FH, Schoolnik GK. Role of rpoS in stress survival and virulence of *Vibrio cholerae*. *J Bacteriol* 1998;180:773–784.
6. Thelin KH, Taylor RK. Toxin-coregulated pilus, but not mannose-sensitive hemagglutinin, is required for colonization by *Vibrio cholerae* O1 El Tor biotype and O139 strains. *Infect Immun* 1996;64:2853–2856.
7. Heidelberg JF, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, *et al*. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae. Nature* 2000;406:477–483.
8. Miller VL, DiRita VJ, Mekalanos JJ. Identification of toxS, a regulatory gene whose product enhances toxR-mediated activation of the cholera toxin promoter. *J Bacteriol* 1989;171:1288–1293.
9. Reimer A, Domselaar G, Stroika S, Walker M, Kent H, *et al*. Comparative genomics of *Vibrio cholerae* from Haiti, Asia, and Africa. *Emerg Infect Dis* 2011;17:2113.
10. Matthey N, Drebes Dörr NC, Blokesch M. Long-read-based genome sequences of pandemic and environmental *Vibrio cholerae* strains. *Microbiol Resour Announc* 2018;7:e01574-18.
11. Weng Y, Bina XR, Bina JE. Complete genome sequence of *Vibrio cholerae* O1 El Tor strain C6706. *Microbiol Resour Announc* 2021;10.
12. Creasy-Marrazzo A, Saber MM, Kamat M, Bailey LS, Brinkley L, *et al*. Genome-wide association studies reveal distinct genetic correlates and increased heritability of antimicrobial resistance in *Vibrio cholerae* under anaerobic conditions. *Microb Genom* 2022;8:mgen000905.
13. Rivera-Chávez F, Mekalanos JJ. Cholera toxin promotes pathogen acquisition of host-derived nutrients. *Nature* 2019;572:244–248.
14. Sit B, Srisuknimit V, Bueno E, Zingl FG, Hullahalli K, *et al*. Undecaprenyl phosphate translocases confer conditional microbial fitness. *Nature* 2023;613:721–728.
15. Dalia AB, McDonough E, Camilli A. Multiplex genome editing by natural transformation. *Proc Natl Acad Sci USA* 2014;111:8937–8942.
16. Jaskólska M, Adams DW, Blokesch M. Two defence systems eliminate plasmids from seventh pandemic *Vibrio cholerae*. *Nature* 2022;604:323–329.
17. Vidakovic L, Mikhaleva S, Jeckel H, Nisnevich V, Strenger K, *et al*. Biofilm formation on human immune cells is a multicellular predation strategy of *Vibrio cholerae*. *Cell* 2023;186:2690–2704.
18. Prentice JA, van de Weerd R, Bridges AA. Cell-lysis sensing drives biofilm formation in *Vibrio cholerae*. *Nat Commun* 2024;15:2018.
19. Bridges AA, Prentice JA, Wingreen NS, Bassler BL. Signal transduction network principles underlying bacterial collective behaviors. *Annu Rev Microbiol* 2022;76:235–257.
20. Conner JG, Teschler JK, Jones CJ, Yildiz FH. Staying alive: *Vibrio cholerae's* cycle of environmental survival, transmission, and dissemination. *Microbiol Spectr* 2016;4.
21. Silva AJ, Benitez JA. *Vibrio cholerae* biofilms and cholera pathogenesis. *PLoS Negl Trop Dis* 2016;10:e0004330.
22. Zhu J, Mekalanos JJ. Quorum sensing-dependent biofilms enhance colonization in *Vibrio cholerae. Dev Cell* 2003;5:647–656.

23. Gallego-Hernandez AL, DePas WH, Park JH, Teschler JK, Hartmann R, *et al.* Upregulation of virulence genes promotes *Vibrio cholerae* biofilm hyperinfectivity. *Proc Natl Acad Sci U S A* 2020;117:11010–11017.

24. Rita T, Bharathi P, Andrew C. Growth in a biofilm induces a hyperinfectious phenotype in *Vibrio cholerae* . *Infect Immun* 2010;78:3560–3569.

25. Faruque SM, Biswas K, Udden SMN, Ahmad QS, Sack DA, *et al.* Transmissibility of cholera: *In vivo* -formed biofilms and their relationship to infectivity and persistence in the environment . *Proc Natl Acad Sci USA* 2006;103:6350–6355.

26. Waters CM, Bassler BL. QUORUM SENSING: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol* 2005;21:319–346.

27. Bridges AA, Bassler BL. The intragenus and interspecies quorum-sensing autoinducers exert distinct control over Vibrio cholerae biofilm formation and dispersal. *PLoS Biol* 2019;17:e3000429.

28. Hammer BK, Bassler BL. Quorum sensing controls biofilm formation in *Vibrio cholerae*. *Mol Microbiol* 2003;50:101–104.

29. Rutherford ST, van Kessel JC, Shao Y, Bassler BL. AphA and LuxR/HapR reciprocally control quorum sensing in vibrios. *Genes Dev* 2011;25:397–408.

30. Stutzmann S, Blokesch M. Circulation of a quorum-sensing-impaired variant of *Vibrio cholerae* strain C6706 masks important phenotypes. *mSphere* 2016;1:e00098-16.

31. Meibom KL, Blokesch MDolganovNAWu C-Y, Schoolnik GK. Chitin induces natural competence in *Vibrio cholerae* . *Science* 2005;310:1824–1827.

32. Waters CM, Lu W, Rabinowitz JD, Bassler BL. Quorum sensing controls biofilm formation in *Vibrio cholerae* through modulation of cyclic di-GMP levels and repression of vpsT. *J Bacteriol* 2008;190:2527–2536.

33. Hammer BK, Bassler BL. Distinct sensory pathways in *Vibrio cholerae* El Tor and classical biotypes modulate cyclic dimeric GMP levels to control biofilm formation. *J Bacteriol* 2009;191:169–177.

34. Zhao X, Koestler BJ, Waters CM, Hammer BK. Post-transcriptional activation of a diguanylate cyclase by quorum sensing small RNAs promotes biofilm formation in *Vibrio cholerae*. *Mol Microbiol* 2013;89:989–1002.

35. Hengge R. Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol* 2009;7:263–273.

36. Conner JG, Zamorano-Sánchez D, Park JH, Sondermann H, Yildiz FH. The ins and outs of cyclic di-GMP signaling in *Vibrio cholerae*. *Curr Opin Microbiol* 2017;36:20–29.

37. Römling U, Galperin MY, Gomelsky M. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol Mol Biol Rev* 2013;77:1–52.

38. Hornick RB, Music SI, Wenzel R, Cash R, Libonati JP, *et al.* The broad street pump revisited: response of volunteers to ingested cholera vibrios. *Bull NY Acad Med* 1971;47:1181–1191.

39. Merrell DS, Camilli A. The cadA gene of Vibrio cholerae is induced during infection and plays a role in acid tolerance. *Mol Microbiol* 1999;34:836–849.

40. Singh A, Barnard TG. Surviving the acid barrier: responses of pathogenic *Vibrio cholerae* to simulated gastric fluid. *Appl Microbiol Biotechnol* 2015;100:815–824.

41. Meng SY, Bennett GN. Nucleotide sequence of the *Escherichia coli* cad operon: a system for neutralization of low extracellular pH. *J Bacteriol* 1992;174:2659–2669.

42. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 2017;13:e1005595–e1005595..

43. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, *et al.* The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 2008;9:75.

44. Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, *et al.* PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res* 2014;42:D581–D591.

45. Madden T. The BLAST Sequence Analysis Tool. 2002 Oct 9. The NCBI Handbook. Bethesda (MD): National Center for Biotechnology Information (US). 2002.

46. O'Toole GA. Microtiter dish biofilm formation assay. *JoVE* 2011:e2437.

47. Zhu J, Miller MB, Vance RE, Dziejman M, Bassler BL, *et al.* Quorum-sensing regulators control virulence gene expression in *Vibrio cholerae* . *Proc Natl Acad Sci USA* 2002;99:3129–3134.

48. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012;9:676–682.

49. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 2010;5:1–17.

50. Hall BM, Ma C-X, Liang P, Singh KK. Fluctuation analysis CalculatOR: a web tool for the determination of mutation rate using Luria-Delbruck fluctuation analysis. *Bioinformatics* 2009;25:1564–1565.

51. Rosche WA, Foster PL. Determining mutation rates in bacterial populations. *Methods* 2000;20:4–17.

52. Jurasinski G, Koebsch F, Guenther A, Beetz S. Flux rate calculation from dynamic closed chamber. R package version 03-01; 2022. https://cran.r-project.org/package=flux

53. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA III, *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 2009;6:343–345.

54. House BL, Mortimer MW, Kahn ML. New recombination methods for *Sinorhizobium meliloti* genetics. *Appl Environ Microbiol* 2004;70:2806–2815.

55. Hartley JL, Temple GF, Brasch MA. DNA cloning using in vitro site-specific recombination. *Genome Res* 2000;10:1788–1795.

56. Simon R, Priefer U, Pühler A. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram negative bacteria. *Nat Biotechnol* 1983;1:784–791.

57. Parrello B, Butler R, Chlenski P, Olson R, Overbeek J, *et al.* A machine learning-based service for estimating quality of genomes using PATRIC. *BMC Bioinformatics* 2019;20:486.

58. Abel S, Abel zur Wiesch P, Chang H-H, Davis BM, Lipsitch M, *et al.* Sequence tag–based analysis of microbial population dynamics. *Nat Methods* 2015;12:223–226.

59. Cameron DE, Urbach JM, Mekalanos JJ. A defined transposon mutant library and its use in identifying motility genes in *Vibrio cholerae* . *Proc Natl Acad Sci USA* 2008;105:8736–8741.

60. Morais LLC de S, Garza DR, Loureiro ECB, Vale ER, Santos DSA de S, *et al.* Population and genetic study of Vibrio cholerae from the amazon environment confirms that the WASA-1 prophage is the main marker of the epidemic strain that circulated in the region. *PLoS One* 2013;8:e81372.

61. Val M-E, Marbouty M, de Lemos Martins F, Kennedy SP, Kemble H, *et al.* A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. *Sci Adv* 2016;2:e1501914.

62. Matthey N, Stutzmann S, Stoudmann C, Guex N, Iseli C, *et al.* Neighbor predation linked to natural competence fosters the transfer of large genomic regions in *Vibrio cholerae*. *elife* 2019;8:e48212.

63. Nusrin S, Gil AI, Bhuiyan NA, Safa A, Asakura M, *et al.* Peruvian *Vibrio cholerae* O1 El Tor strains possess a distinct region in the Vibrio seventh pandemic island-II that differentiates them from the prototype seventh pandemic El Tor strains. *J Med Microbiol* 2009;58:342–354.

64. Taviani E, Grim CJ, Choi J, Chun J, Haley B, *et al.* Discovery of novel *Vibrio cholerae* VSP-II genomic islands using comparative genomic analysis. *FEMS Microbiol Lett* 2010;308:130–137.

65. McKee RW, Kariisa A, Mudrak B, Whitaker C, Tamayo R. A systematic analysis of the *in vitro* and *in vivo* functions of the HD-GYP domain proteins of *Vibrio cholerae*. *BMC Microbiol* 2014;14:272.

66. Thi TD, Lopez E, Rodriguez-Rojas A, Rodriguez-Beltran J, Couce A, *et al.* Effect of recA inactivation on mutagenesis of *Escherichia coli*

exposed to sublethal concentrations of antimicrobials. *J Antimicrob Chemother* 2011;66:531–538.

67. Cirz RT, Chin JK, Andes DR, de Crécy-Lagard V, Craig WA, *et al*. Inhibition of mutation and combating the evolution of antibiotic resistance. *PLoS Biol* 2005;3:e176.

68. Friedberg EC, Walker GC, Siede W. *DNA Repair and Mutagenesis*. ASM Press, 2005.

69. Kanie S, Horibata K, Kawano M, Isogawa A, Sakai A, *et al*. Roles of RecA protein in spontaneous mutagenesis in *Escherichia coli*. *Genes Genet Syst* 2007;82:99–108.

70. Bell JC, Kowalczykowski SC. RecA: regulation and mechanism of a molecular search engine. *Trends Biochem Sci* 2016;41:491–507.

71. Krašovec R, Richards H, Gomez G, Gifford DR, Mazoyer A, *et al*. Measuring microbial mutation rates with the fluctuation assay. *J Vis Exp* 2019.

72. Teschler JK, Nadell CD, Drescher K, Yildiz FH. Mechanisms underlying *Vibrio cholerae* biofilm formation and dispersion. *Annu Rev Microbiol* 2022;76:503–532.

73. Flemming H-C, Wingender J, Szewzyk U, Steinberg P, Rice SA, *et al*. Biofilms: an emergent form of bacterial life. *Nat Rev Microbiol* 2016;14:563–575.

74. Koestler BJ, Waters CM. Bile acids and bicarbonate inversely regulate intracellular cyclic di-GMP in *Vibrio cholerae*. *Infect Immun* 2014;82:3002–3014.

75. Rocha EPC, Cornet E, Michel B. Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genet* 2005;1:e15.

76. Karlin S, Brocchieri L. Evolutionary conservation of RecA genes in relation to protein structure and function. *J Bacteriol* 1996;178:1881–1894.

77. Ram Y, Hadany L. The evolution of stress-induced hypermutation in asexual populations. *Evolution* 2012;66:2315–2328.

78. Swings T, Van den Bergh B, Wuyts S, Oeyen E, Voordeckers K, *et al*. Adaptive tuning of mutation rates allows fast response to lethal stress in *Escherichia coli elife* 2017;6:e22939.

79. Taddei F, Radman M, Maynard-Smith J, Toupance B, Gouyon PH, *et al*. Role of mutator alleles in adaptive evolution. *Nature* 1997;387:700–702.

80. Srivastava D, Harris RC, Waters CM. Integration of cyclic di-GMP and quorum sensing in the control of vpsT and aphA in *Vibrio cholerae*. *J Bacteriol* 2011;193:6331–6341.

81. Hengge R. High-specificity local and global c-di-GMP signaling. *Trends Microbiol* 2021;29:993–1003.

82. Satchell KJF, Jones CJ, Wong J, Queen J, Agarwal S, *et al*. Phenotypic analysis reveals that the 2010 Haiti cholera epidemic is linked to a hypervirulent strain. *Infect Immun* 2016;84:2473–2481.

83. Walker LM, Haycocks JRJ, Van Kessel JC, Dalia TN, Dalia AB, *et al*. A simple mechanism for integration of quorum sensing and cAMP signalling in *Vibrio cholerae elife* 2023;12:RP86699.

84. Srivastava D, Hsieh M-L, Khataokar A, Neiditch MB, Waters CM. Cyclic di-GMP inhibits *Vibrio cholerae* motility by repressing induction of transcription and inducing extracellular polysaccharide production. *Mol Microbiol* 2013;90:1262–1276.

85. Platt R, Drescher C, Park SK, Phillips GJ. Genetic system for reversible integration of DNA constructs and lacZ gene fusions into the *Escherichia coli* chromosome. *Plasmid* 2000;43:12–23.

86. Stothard P, Wishart DS. Circular genome visualization and exploration using CGView. *Bioinformatics* 2005;21:537–539.

87. Grant JR, Stothard P. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acid Res* 2008;36:W181–W184.