Leibniz-Institut für Ostseeforschung Warnemünde

Universität Rostock Traditio et Innovatio

Predictors of *Vibrio vulnificus* occurrence: A machine learning approach

Cumulative dissertation for the award of the academic degree Doctor rerum naturalium (Dr. rer. nat.) of the Faculty of Mathematics and Natural Sciences of the University of Rostock

Submitted by

David Jeroen Riedinger

Supervisor:
Prof. Dr. Matthias Labrenz

Rostock, 2024

https://doi.org/10.18453/rosdok_id00005322

Reviewers:

Prof. Dr. Matthias Labrenz, Universität Rostock, Institut für Ostseeforschung

PD Dr. Holger Scholz, Robert Koch-Insitut

Defended: 2024

Table of content

Summary

Marginal seas and estuaries are undergoing significant transformations due to ongoing anthropogenic impacts, alterations in the hydrological cycle, increased cyanobacterial bloom frequency and rising temperatures. The genus *Vibrio*, and especially *V. vulnificus* stands to profit from these changes. This lethal human pathogen is found worldwide as a natural part of the microbial communities in these systems. Currently, infection rates are low; however, it is one of the most lethal bacterial human pathogens. An increase in temporal and spatial range in temperate brackish waters worldwide is anticipated, increasing both the infection risk and the economic burden, necessitating the development of mitigation strategies. While they might provide an avenue to limit the proliferation of *V. vulnificus*, the influence of currents, aquatic marcophytes, eutrophication and the microbial community remain unknown. The current spread, environmental conditions most suitable for *V. vulnificus,* monitoring methods, and mitigation measures are explored in this thesis. Using freely available archived 16S rRNA sequencing data, a tentative poleward spread of *V. vulnificus* occurrence over the last decade could be detected. Two random forest models, predicting relative abundance of *V. vulnificus*, either using remote sensing data or the microbial community were trained. On a global scale, besides temperature and salinity, chlorophyll *a*, and to a lesser extent, currents affect *V. vulnificus* relative abundance. Members of the Vibrionaceae and Pseudoalteromonadaceae family were shown to be the best microbial predictors. A fast and affordable method to validate model predictions was developed, utilizing a two-step agar-based approach that can distinguish *V. vulnificus* from other *Vibrio* spp. with over 80% accuracy within 48 hours. The Baltic Sea is an ideal site for studying mitigation measures as surface water temperatures are expected to rise by 2-4 °C by the end of the century. It is influenced by eutrophication, modifying the microbial community dynamics and *Zostera marina*, a seagrass that creates a unique habitat and has been shown to potentially reduce presumed *V. vulnificus* colonies at a local scale. Using both molecular and cultivation methods over the Baltic Sea salinity gradient, abundance inside and outside seagrass meadows was found to be similar, while eutrophication seemed to stimulate *V. vulnificus* abundance in the natural environment. Collectively, these results suggest a potential cascade starting with the influx of inorganic nutrients and resuling in *V. vulnificus* blooms. The inorganic nutrients stimulate cyanobacterial blooms, which are already increasing due to climate change, which in turn stimulate *V. vulnificus*, potentially originating from the sediment, through release of bioavailable organic matter during decay. Consequently, reducing eutrophication induced algal blooms might be part of a successful *V. vulnificus* mitigation strategy, not only in the Baltic Sea, but worldwide.

Introduction

Genus *Vibrio*

The widely dispersed bacterial genus *Vibrio* comprises over 130 identified species[1], three of which are strongly associated to human infections[2]. Now, many of these species have been isolated and described, but *Vibrio* first gained prominence when Italian anatomist Filippo Pacini documented a *Vibrio* as cholera bacillus amidst a cholera outbreak in Florence in 1854. sBelonging to the phylum Gammaproteobacteria, *Vibrio* bacteria are gram-negative, motile, slightly curved rods, exhibiting facultative anaerobic metabolism. They are generally mesophilic and display a preference for halophilic environments, although exceptions exist, such as the *V. cholera*[3].

Members of the genus *Vibrio* inhabit a broad spectrum of niches across freshwater, estuarine, and marine ecosystems. Within these environments, they participate in processes such as nutrient cycling[4,5], through the decomposition of complex and abundant biological polymers[6] (e.g. chitin[7], collagen[8] and complex hydrocarbons[9–11]) and nitrogen fixation[12], and the synthesis of polyunsaturated fatty acids, essential compounds for the aquatic food web organisms[13]. This diversity of roles is paired with a high diversity in lifestyles; they can grow as free-living organisms, in association with a wide variety of other organisms, such as bobtail squids[14] and sharks[15], and as biofilms. Consequently, they are able to respond rapidly to environmental changes, are ubiquitous and are likely benefactors of global human induced changes.

The three major human pathogenic *Vibrio* are *V. cholerae* serogroups O1 or O139, *V. parahaemolyticus* and *V. vulnificus*. While *V. cholera* thrives predominantly in low-salinity environments (0.5 – 3.5[16]), *V. vulnificus* is most abundant in moderate salinity (5 - 16[17,18]) and *V. parahaemolyticus* dominates at more marine salinities (optimum between 24-3[19]). Cholera, caused by ingestion of *V. cholerae* contaminated water, has an estimated 2.86 million cases and 95,000 deaths per year worldwide[20]. It represents a global threat to public health, exacerbated by inadequate access to clean water and sanitation facilities in many at risk areas and after calamities, such as earthquakes[21] and floods[22]. *V. parahaemolyticus* is more common in developed areas and is one of the leading causes of human acute gastroenteritis in relation to seafood[19,23,24]. Less commonly, it can also lead to wound infections and septicemias. While currently *V. vulnificus* infection case numbers are relatively low (e.g. < 40 cases per year, on

average, in the USA[25]), it is the most lethal of these common *Vibrio* pathogens, with lethality rates up to 50% after wound infection[26].

Vibrio vulnificus

Ingesting seafood, notably raw oysters, contaminated with *V. vulnificus,* or direct contact in brackish water through a wound[27,28], may lead to a severe and rapidly progressing systemic infection. Clinical manifestations of this condition encompass fever, chills, nausea, hypotensive septic shock, and the development of secondary lesions on the extremities of affected individuals (Figure 1)[29,30]. *V. vulnificus* can generally be isolated at water temperatures between 13 and 30°C, although they endure at higher temperatures[31,32]. They can be free living or attached to both animate and inanimate particles, frequently associating with zooplankton, oysters, fish and extending their habitat to sediments and macro-algae where they form biofilms[33–35].

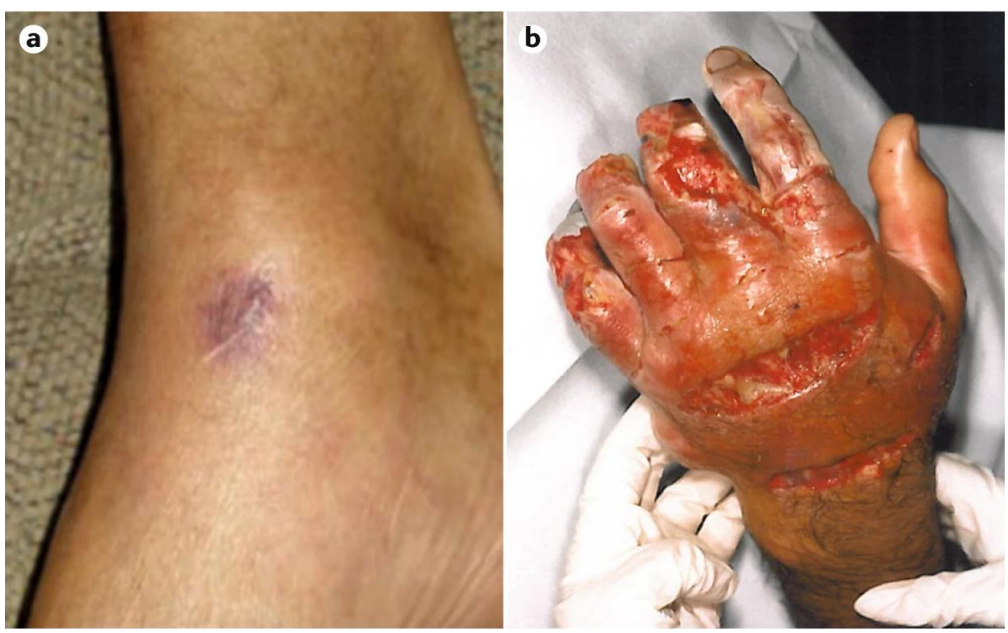


Figure 1. Even small entry sites are sufficient for fatal consequences. *(a) wound, that after infection with* V. vulnificus *led to fatal outcomes ≤48h. (b) life-altering consequences of a* V. vulnificus *infection, amputations are often necessary. Adapted from Baker-Austin et al., 2018[2].*

Attempts have been made to quantify the direct and indirect costs of *V. vulnificus* infections. It was already recognized as the most costly marine-borne pathogen in the U.S. accounting for approximately a third (233 million $) of all yearly costs (352 million $) related

to seafood-borne illnesses[36]. Additional costs occur from direct and recreational exposure of 28 million $ in the U.S. alone. These are only the costs of the cases with a known etiology and real costs might be higher. A substantial portion of these costs arises from premature death. Further costs are associated with limited harvesting or obligatory processing steps of seafood during high-risk seasons. Simulation of just the U.S. Gulf coast under restriction proposed by the FDA in 1996, indicated that the total quantity of oysters harvested might decrease by 40%[37]. Increased periods of high risk and similar restrictions on a global scale would lead to a loss in food supply and economic activity in many regions.

The total costs of Vibriosis, including cases by non- *V. vulnificus Vibrio*, is projected to reach $5.2 billion under RCP4.5 and $7.3 billion annually under RCP8.5 in the U.S. alone[38]. This estimate does not take into account the additional potential loss of economic activity from tourism. However, prompted by a fatal *V. vulnificus* infection following a swim off Warnemünde in the summer of 2018, German authorities issued unprecedented advisory, cautioning immunodeficient individuals against swimming in the Baltic Sea. This led to extensive media coverage and potential adverse effects for tourism in these coastal areas. A decrease in tourism can negatively affect coast-adjacent industries such as hotels, restaurants, seafood markets and service providers[38].

Furthermore, negative effects are noticed in aquaculture, a source of food and economic productivity in many areas. 8 Dutch eel farms had 23 severe outbreaks between 1996 and 2009, leading to loss of productivity, one infected eel farmer and an increased risk for fish processors[39]. Analogously, *V. vulnificus* is also widespread in aquaculture in Taiwan, their primary source of seafood, posing a threat to certain segments of the food supply[40]. Similar problems have occurred in Denmark, with a *V. vulnificus* outbreak in a eelfarm in 2004[41], in Spain[42], and in a Tilapia farm in India[43]. Determining the total economic costs is challenging due insufficient tracking of infections. Nevertheless, the undeniable potential risk to human health from consuming these farmed fish remains a concern.

Not all *V. vulnificus* are pathogenic

The low number of infection cases is likely due to two factors: firstly, not all strains of *V. vulnificus* are pathogenic, and secondly, not all individuals are susceptible to infection. Multiple studies have shown that infections predominantly occur in males[26,44], often above the age of 40[32], with preexisting risk factors, such as alcohol-associated liver cirrhosis or hepatitis

leading to elevated serum iron levels[45,46], or who are otherwise immuno-compromised[44]. *V. vulnificus* itself can be categorized into three biotypes, mostly based on biochemical and pathogenic characteristics. Strains within biotype 1 are responsible for the majority of human infections, while strains frequently linked to eel vibriosis, though they can also pose a threat to humans, are categorized within biotype 2[47]. Biotype 3 seems to be a genetic mosaic of the two aforementioned biotypes[48] and are endemic to Israel. The precise factors contributing to the distinction between pathogenic (C-genotypes) and non-pathogenic strains (E- genotypes) remain unclear. Several possible virulence factors have been suggested. These include extracellular hemolysin/cytolysin[49] (*vvhA* gene), an elastolytic protease (*vvpE*)[50], the ability to acquire iron from transferrin[51], the existence of a polysaccharide capsule[52], and the presence of an endotoxic lipopolysaccharide (LPS)[53].

Distribution of *V. vulnificus* infections

Despite the near global distribution of *V. vulnificus*, the abundance and frequency of necrotizing skin and soft tissue infections has a strong geographical pattern (with the majority of cases between 0 and 30°N) (Figure 2). Highlighting where most of the contacts between susceptible humans and C-genotypes occur. Case number and occurrence is notably higher in subtropical regions in the northern hemisphere than anywhere else. This might be due to a combination of high sea surface temperatures (SST), a subtropical monsoon climate pattern, more common in the northern hemisphere, and high population density. High temperatures do not only increase total *V. vulnificus* abundance, but might specifically increase the pathogenic fraction[54]. Warm currents, such as the Kuroshio Stream, in the subtropical pacific coastal region and the Gulf Stream in the subtropical Atlantic coastal region thus likely increase risk, this aligns with high numbers of infection cases (Figure 2).

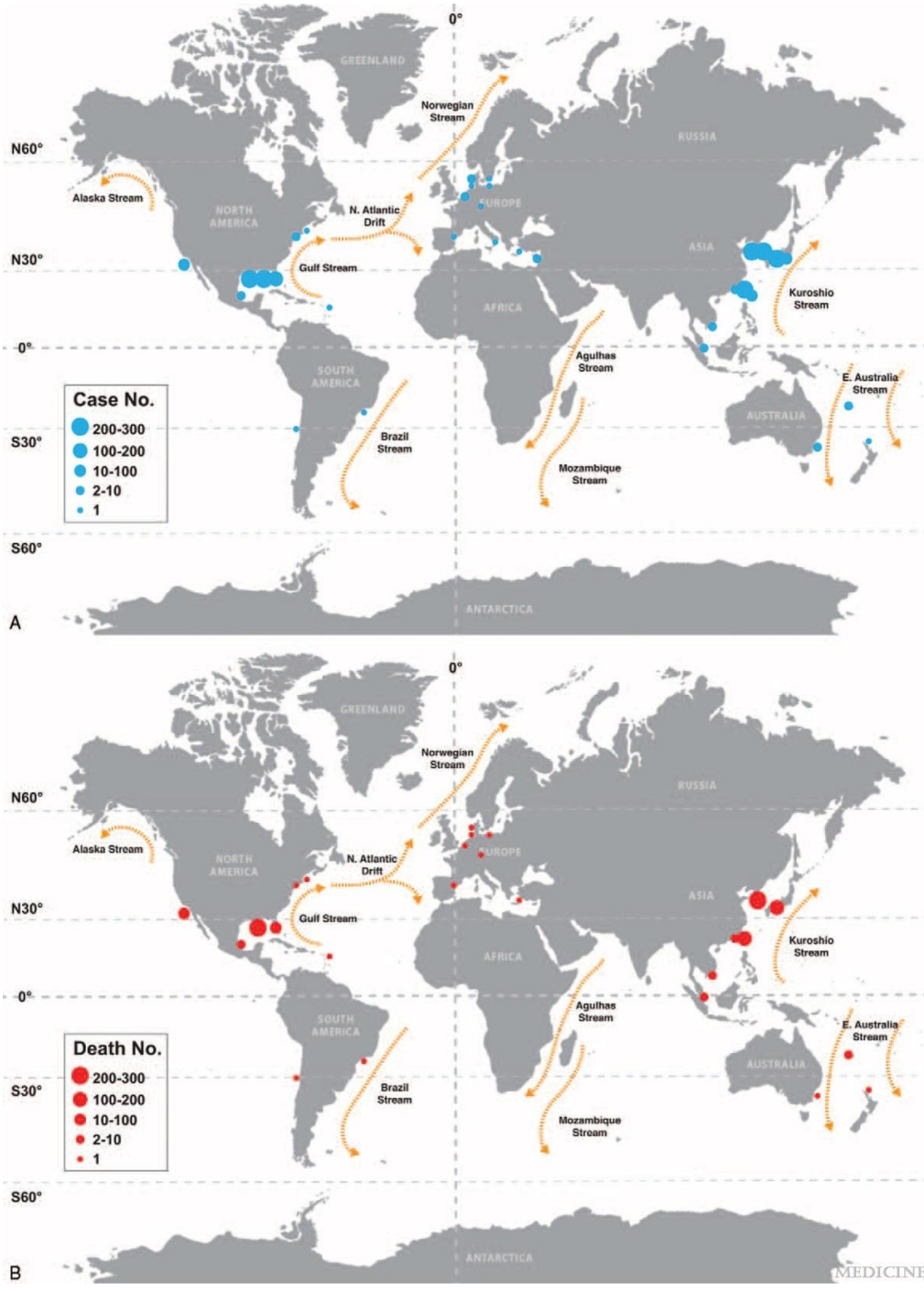


Figure 2. Spatial distribution of the reported (A) total and (B) fatal episodes of* V. vulnificus *necrotizing skin and soft tissue infections between 1966 and 2014. *Adapted from Huang et al., 2016* [54].

V. vulnificus infection risk under anthropogenic impacts

Global warming is expected to lead to higher abundances of *V. vulnificus* globally and a concomitant increased infection risk[55,56]. In fact, infection cases are already rising in for example the Eastern US[57] and the Baltic Sea region (Figure 3)[55]. Heatwaves in Northern Europe during 1994, 2003, 2006, 2010, 2014, and 2018 resulted in a rise in *Vibrio* wound infections in the region (Figure 3)[55], providing a glimpse into the future of coastal regions.

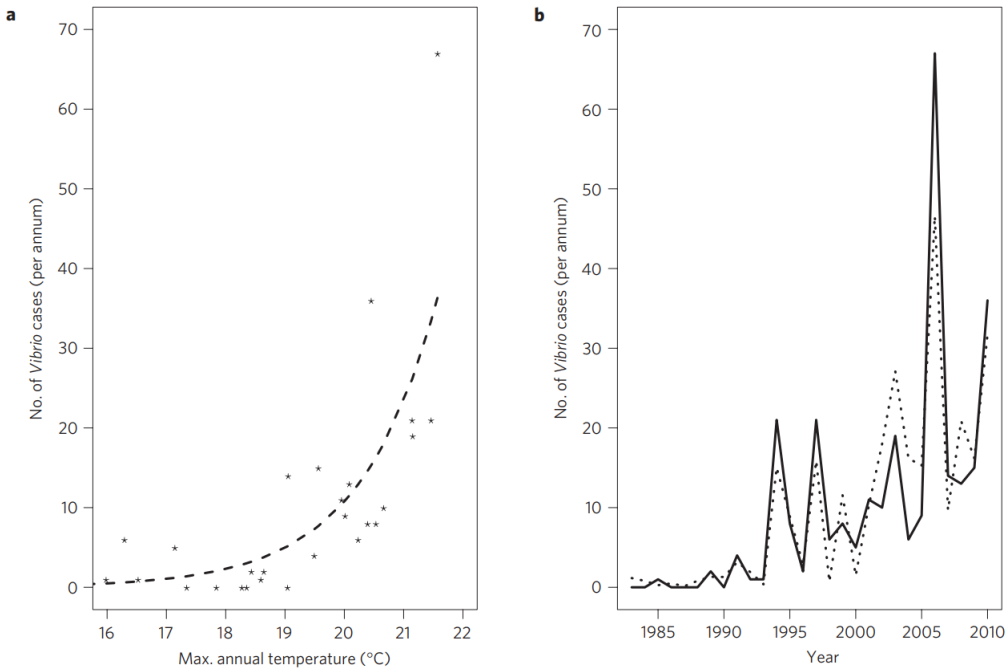


Figure 3. Relationship between* Vibrio *infection cases and maximum sea surface temperature in the Baltic Sea. *(a). Relationship between* Vibrio *infections reported and maximum annual SST. Stars are observed data, and the dashed line represents model predictions. (b) Time series of* Vibrio *cases. The solid line indicates observed cases, while the dotted line represents model predictions. Adapted from Baker-Austin et al., 2013*[55].

The number of infections was likely stimulated by additional antropogenic impacts, such as coastal freshening, eutrophication, and ecosystem disturbances. A positive relationship between eutrophication, phytoplankton blooms and *V. vulnificus* has been hypothesized before[58]. The release of labile dissolved organic from decaying phytoplankton blooms, the increase of zooplankton, specifically copepods, and the ability to avoid protozoan predation by attaching to phytoplankton might all stimulate *V. vulnificus*.

Recently, tropical seagrasses have been described to reduce the concentration of potentially harmful bacteria in the water column[59] and have been linked to a reduction of presumptive *V. vulnificus* colony counts in the Baltic Sea[60], although again the mechanism of this apparent reduction remains unclear. Non-exclusive mechanisms may involve (i) a rise in sedimentation rate facilitated by hydrodynamic attenuation of the canopy, leading to a reduction in particle load[61–63] and the associated bacteria, (ii) the reduction of total bacteria load through direct filter feeding or the feeding on host plankton[64,65], or exudation of allopathic chemicals that directly inhibit bacterial growth[66]. While the exact combination of mechanisms remains elusive, seagrass loss might have increased *V. vulnificus* infection risks. With the combined effects of warmer, fresher coastlines, with larger phytoplankton blooms due to increased eutrophication and removal of potential nature-based filters, such as seagrass, the *V. vulnificus* infection risk might rise more than currently predicted.

Especially in the Baltic Sea, one of the regions where the anthropogenic (climate) effects will be felt earliest and most intensely, as it is, in many ways, a time machine for the future of coastal oceans[67]. With current levels of warming, acidification, eutrophication and deoxygenation that are anticipated in coastal areas worldwide[67–69]. With one of the fastest aging populations in the world[70], a tourism industry already grappling with disruptions from extreme environmental events[71], and a lateral surface salinity gradient ranging from 25 to 5 from southwest to northeast[72] – an environment conducive to *V. vulnificus* – this setting becomes ideal for studying the potential impact of *V. vulnificus* in the future and to evaluate potential solutions

Seagrass and eutrophication reduction to limit *V. vulnificus* proliferation

A pilot study in the western Baltic Sea has shown a reduction of total bacterial counts and presumed *V. vulnificus* colonies in seagrass beds compared to control stations outside of seagrass beds (Figure 4a,b)[60]. If this proves effective using more specific measurements than presumed *V. vulnificus* colony counts and over a wider range of environmental conditions, seagrass might provide a local solution to the increasing *V. vulnificus* problem at popular recreational beaches. Expanding seagrass meadows can occur naturally or through targeted renaturation efforts

On a regional level, a strategy could involve mitigating eutrophication in coastal waters. DOC derived from phytoplankton has been shown to stimulate *Vibrio* colony counts in

an experimental setting (Figure 4c, d, e). Managing nutrient inflow and subsequently reducing phytoplankton blooms, along with the associated release of organic nutrients, may offer a viable approach to controlling *V. vulnificus* blooms. As most eutrophication originates from land-based sources, reducing eutrophication requires changes in our approach to agriculture and wastewater management.

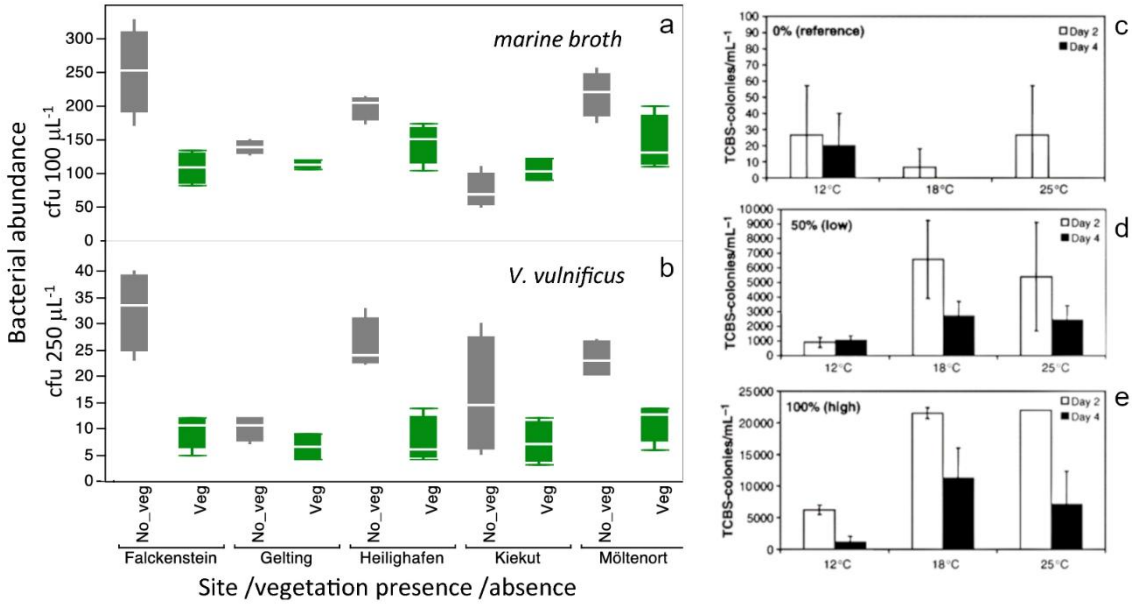


Figure 4. Potential mitigation of* V. vulnificus *proliferation through seagrass or reduction of eutrophication. *(a, b) Total colony forming units and presumed* V. vulnificus *colonies at 5 sites in the South-Western Baltic Sea. Comparison between sites with seagrass and without (N=4). Adapted from Reusch et al., 2021*[60]. *(c, d, e)* Vibrio *spp. abundances at different temperatures and DOC concentration treatments. (a) 0 mg C* L^{-1} *DOM addition. (b) 2.1 mg C* L^{-1} *DOM addition. (c) 4.2 mg C* L^{-1} *DOM addition. All DOM originates from* Nodularia spumigena*. Error bars are SD (N=3). Adapted from Eiler et al., 2007*[58].

V. vulnificus can be detected through culture-based or molecular methods

In order to evaluate current risk and the effectiveness of potential mitigation measured, methods to quantify *V. vulnificus* in the environment are essential. Isolation on selective media is the preferred method of many regulatory entities worldwide[73]. It benefits from low costs, ease of operation and providing a culture for further analysis. For potentially pathogenic *Vibrio* spp. various selective media, such as Thiosulfate-citrate-bile salts-sucrose agar (TCBS) and CHROMagar *Vibrio* (CHROMagar; Paris, France) are used[74]. They allow for the growth of a select number of *Vibrio* spp., while excluding closely related genera. In TCBS this is, for

example, achieved through thiosulfate and sodium citrate inhibiting Enterobacteria and sodium cholate and ox bile inhibiting gram-positive bacteria and retarding the growth of *Enterococci*. One downside is that they require an additional molecular step to verify the identity of the organism.

This caveat can be mitigated by using molecular methods directly. The potential virulence factor, *vvhA* is regularly used to quantify *V. vulnificus* using digital droplet polymerase chain reaction (ddPCR). This quantification does not directly determine *V. vulnificus* cells, as the *vvhA* gene may not be present in every cell or might exist in multiple copies within a single cell. Therefore, only the gene copy numbers per milliliter can be accurately calculated.

V. vulnificus occurrence can also be assessed using 16S rRNA sequencing. This approach involves comparing (zero-distance) operational taxonomic units (z)OTUs or amplicon sequence variants (ASVs), derived from environmental DNA samples that have undergone amplification and sequencing in the V3-V5 regions of the 16S rRNA gene. By comparing these (z)OTUs with known *V. vulnificus* 16S rRNA gene sequences and existing databases such as SILVA[75], they can be taxonomically identified. Given the minimal differences in 16S rRNA sequences between microbes, with just a few bases difference between *V. vulnificus* and sister branches in the V3-V5 region, classification should to be done conservatively. (z)OTUs should only be identified as *V. vulnificus* if they exhibit 100% alignment and sequence similarity to an established *V. vulnificus* 16S rRNA sequence. While this method does not provide absolute abundances of V. *vulnificus* and is affected by database size and quality, it offers a measure of relative *V. vulnificus* abundance and concurrently captures a portion of the microbial diversity present in the sample.

Analyzing microbial communities

Analyzing microbial communities through amplicon sequencing

Currently next-generation sequencing (NGS) is a widely used tool to assess the microbial diversity in environmental samples. Amplicon sequencing, a type of NGS, allows a cost-effective measurement of the relative abundance[76] of most microbes using universal primers (primers reviewed in Klindworth et al., 2013[77]). It relies on the fact that specific regions within genomic DNA contain both conserved and variable sequences. The conserved regions are common between different prokaryotic groups in the 16S rRNA and in eukaryotic

groups in the 18S rRNA, and provide binding sites for universal primers. After amplification, the variable regions can be used to differentiate between different taxa. The sequences are amplified using polymerase chain reaction (PCR), which are repeated cycles of denaturation, primer annealing and elongation, resulting in exponential amplification of the targeted region (Figure 4b), ensuring sufficient target DNA is available for sequencing. To sequence these fragments, they are attached to a flow cell and amplified (Figure 4e-g). This attachment is achieved through the addition of adapters, short synthetic DNA sequences (Figure 4d), which also provide binding sites for the primers. These adapters also contain unique sequences, identifying the origin sample of a DNA sequence allowing for parallel sequencing. The flow cell has complementary sequences to the attached adapters, allowing the DNA fragments to hybridize and attach to the cell. These fragments bind to an adjacent binding site, forming a bridge (Figure 4e). A complementary strand is synthesized and the double stranded fragment denatured. This cycle is repeated to create clusters of copies of the original DNA strand. Subsequently, primer and fluorescently labeled nucleotides with a terminator are added to the flow cell, as the nucleotide is incorporated, a fluorescent signal specific to it type is detected by a high-resolution camera (Figure 4f). Subsequently, the terminating group is removed, allowing the next nucleotide to attach to the strand and give a signal. These images generated by the camera are processed to identify the fluorescent signal in each cluster and translated back into the associated base, resulting in sequences of nucleotides of the original DNA fragment (Figure 4g).

Characteristics of microbial datasets

Amplicon sequencing datasets are however not perfect representations of the sampled community. The reasons for these differences are discussed in the order they are encountered from sampling to analysis. When sampling any environment, the *in situ* community might respond to the act of sampling itself[78], leading to alterations in the community itself. It is typically impossible to exhaustively sample a biological community, necessitating an effort to obtain a representative sample instead[79] (Figure 5a). Once a sample suitable for the intended study goal has been taken, bacteria should be deactivated as rapidly as possible, while limiting the damage to their DNA[80,81], otherwise the target community will continue to change. Additional errors can be introduced during DNA extraction, due to inter-species and even inter-strain variations in DNA extraction efficiency[82], these differences largely originate from varying efficiency in cell wall and membrane lysis and contaminant removal (Figure 5c). Contaminant removal is necessary as other biomolecules and remaining organic acids can

inhibit DNA amplification[83]. The amplification of extracted DNA through PCR introduces additional biases[84], influenced by properties of the PCR template (such as template concentration and GC content), primer selection (including primer coverage and potential mismatches), choice of polymerase, and the specifics of the PCR protocol (such as annealing temperature and the number of PCR cycles) (Figure 5b, c). Additional biases are introduced during PCR, where substitution (the replacement of one base by another), replication (multiple readings of the same base), and deletion (a base is not read) can occur. The relative abundance measured is further influenced by differences in taxa-specific 16S rRNA operon copy numbers per genome[85] (Figure 5c). Different sequencing technology have different length, quality of DNA reads, and sequencing depth, changing the inferred community composition[86]. This has substantial implications for accurately representing the rare biosphere[87].

The sequences need to be further processed in order to obtain a representative community. This can be done in with a variety of different pipelines leading to different inferred communities[88,89]. Generally reads with low quality scores are removed and primer and adapter sequences are trimmed, resulting in sequences with a largely fixed and uniform length (Figure 5h). Due to both ends of a DNA fragment being sequenced, they need to be merged using the overlapping region. Merged reads from taxonomically equivalent cells might have slightly different sequences, either due to natural variation or due to PCR/sequencing errors. A common way to deal with this is to denoise sequences into ASVs or OTUs using one of the many available algorithms, in this thesis, both DADA2[90] and UNOISE[91] are used (Figure 5i,j). These algorithms differ in their approach, with DADA2 constructing a parametric error model based on the results of an individual sequencing run. This statistical model estimated error rates for each base positions and derived “true” sequences based on this. UNOISE clusters identical sequences, and then analyzes the differences between these clusters. Sequences that are uncommon but closely resemble highly abundant ones are considered erroneous variants of the more abundant sequence. They are then grouped together to form an OTU. The amount of DNA fragments (reads) per OTU only gives the relative abundance of this OTU compared to the other OTUs in the same sample. They do not give insight into cell numbers or biomass. Taxonomic classification of these OTUs not only makes them comparable across studies, but also allows for interpretation using the existing knowledge on taxa. This is achieved by comparing the sequences to established databases, for example SILVA is used in this thesis for prokaryotes and PR2 for eukaryotes. Consequently, the resulting microbiome dataset is an

incomplete compositional representation of the true diversity in the sampled environment, with an unknown error rate[92].

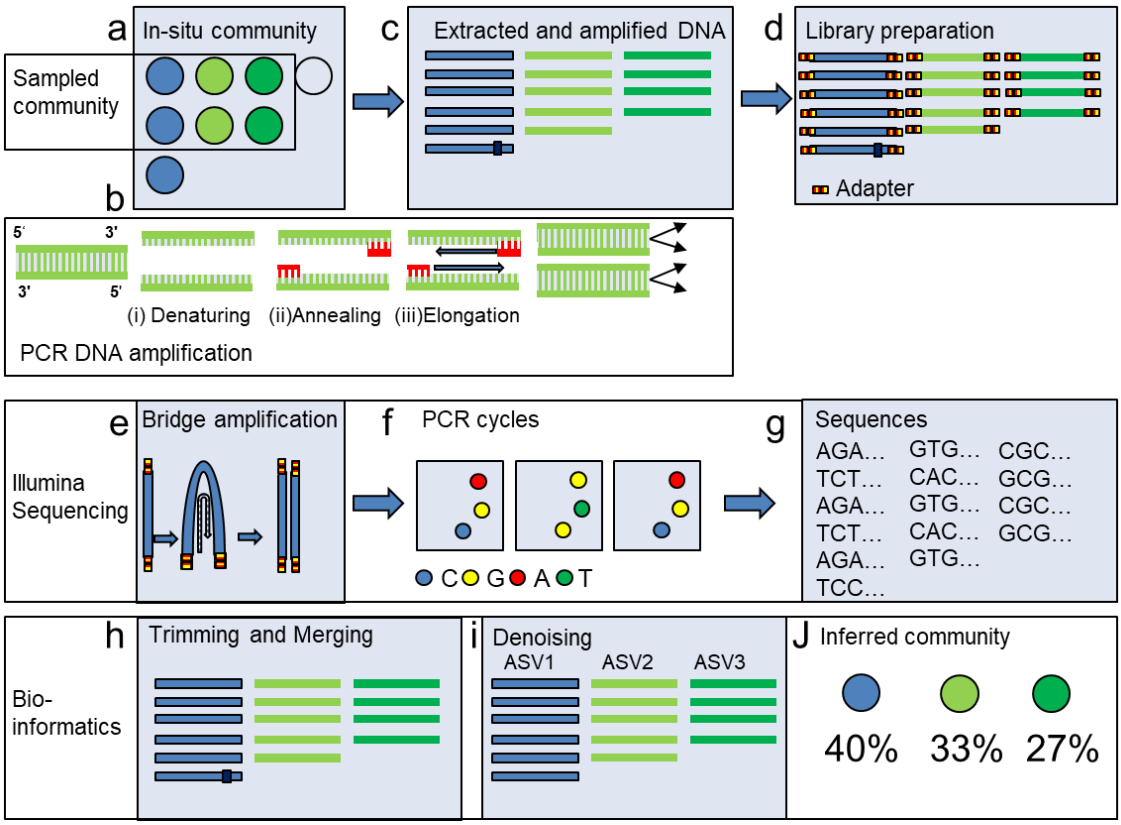


Figure 5. Workflow from sampling an in-situ community to the detected community. *(a) sampling of the* in-situ *community samples a subset of the real community. During extraction, DNA extraction efficiency differs between species and amplification (b) errors can occur leading to a deviation in the DNA compared to the sampled community (c). (d) adapters are attached to both ends of the DNA fragment. (e) Bridge amplification creates clusters of forward and reverse strands of the DNA fragments on the flow cell. (f) PCR cycles with fluorescent probes allow the forward and reverse reads to be read. (g) This leads to forward and reverse sequences of each input DNA fragment, including potential errors from the PCRs. (h) Reads are trimmed and forward and reverse reads are merged. (i) Similar sequences are combined into ASV/OTUs (denoising) limiting the effect or erroneous base additions. (j) Final inferred community deviates from in-situ community but gives a good representation of the major groups*

Despite these limitations, NGS gives us an unprecedented view of the microbial community. The usage of universal primers allows for a good overview of the microbial diversity regardless of prior knowledge of composition and the ability to simultaneously and rapidly process multiple samples for a relatively low cost have ensured that there is currently

no other methodology adopted as widely in environmental microbiology for community analysis. NGS facilitates the detection of non-cultivable organisms and allows for diversity estimates otherwise not possible by capturing a significant part of the microbial diversity in a sample with one analysis.

Microbial communities analyzed to understand *V. vulnificus*

V. vulnificus interacts with the microbes around it. When trying to understand the interaction between *V. vulnificus* and its environment, the microbial community forms crucial part of the picture. Due to the diverse range of physiological traits within bacterial communities, they can, and will, adapt and respond rapidly to environmental changes. Consequently, certain segments of the microbial community, which are most suited for a given set of conditions, will increase in abundance, while others may decrease. For instance, the presence of explosives in the environment modifies the microbial community sufficiently that it can be predicted from the community composition[93]. When the relative increase or decrease of abundance of specific groups of microbes are found to explain considerable parts of the variance in (relative) abundance of *V. vulnificus* and the environmental preferences or role of these organisms are known, they can be used to hypothesize about conditions favoring *V. vulnificus* proliferation. Going one step further, it might even be possible to conclude something about the past conditions of a water mass using the microbial community. When low concentrations of inorganic nutrient are measured, but cyanobacteria are very abundant. One can conclude that inorganic nutrient conditions were likely very high previously, but have been depleted by the cyanobacterial bloom. This, in turn, enables us to estimate the conditions that potentially trigger *V. vulnificus* blooms in a particular water mass.

Machine learning as a tool for high dimensional datasets

Microbial community datasets and additional environmental data are frequently characterized by high dimensionality, containing more features than samples. They are generally sparse, given that a significant portion of microbes is not detected in the majority of samples. The data is compositional and has various data-distributions across its components. Within both the microbial and associated environmental data, features often display collinearity or covariance, further adding to the intricacies of the dataset[94].

A possible way to understand such a dataset would be through descriptive statistics and statistical inference. Descriptive statistics aim to present a dataset more concisely, offering

insights through metrics such as mode, mean, and standard deviation. Statistical inferences on the other hand "creates a mathematical model of the data-generation process to formalize understanding or test a hypothesis about how the system behaves"[95]. While a significant advantage of statistical inference is the interpretability of the chosen model, it is based on *a priori* assumptions about the relationship between variables.

Contrastingly, a machine-learning (ML) model operates without a predefined underlying model for the system. Instead, it models the relationship between the 'x' and 'y' based on the patterns present in the data using general purpose algorithms[95]. The output is a new function that aims to optimize a specific metric. This function is derived using a part of the data and subsequently evaluated on a separate part and the model effectiveness in predicting the previously unseen portion of the data is used to assess how well it captured the underlying mechanism[96]. While statistical inference and ML might result in the same function, the essential difference is that former is model-driven while the latter is data-driven.

This approach becomes particularly helpful when we take into account the features of a microbial dataset. The relationships are often complex and include multiple features, interdependencies between various microbial species and environmental factors. A number of environmental factors might influence the abundance of a specific microbe of interest differently depending on the presence or absence of a cluster of other microorganisms. Such a relationship is difficult to test for when the underlying relationship needs to be assumed *a priori*, but might become apparent from a data-driven approach. Advantages of ML are that the associations between data are more flexible in their form, they have scaling properties compatible with highly dimensional datasets containing a variety of datatypes and are able to utilize heuristics through useful approximation to patterns recognized in the data[97,98].

An issue when deriving the function purely from the data is overfitting. Complex models relying on limited amounts of training data can capture noise that is specific to that data as a genuine effect. Consequently, the model does not generalize to other situations and it is of limited value. Testing against a hold-out set of data, or multiple holdout sets through cross validation is used in most ML approaches to limit this. For cross validation, multiple models are trained on different subsets and evaluated against different holdout subsets (reviewed in Berrar, D. 2019[99]). Although computationally intensive, each trained model will be evaluated against a different test set, reducing the impact of the specific data points included in either the train and test sets. Additional safeguards are frequently implemented, including

regularization[100], which penalizes the model's complexity, feature selection to discard features and simplify the model[101], and the integration of ensemble methods[102]. These strategies collectively increase the model's likelihood of generalization to new, unseen data.

Random forests

Random Forest[96] (RF) is a machine learning algorithm based on an ensemble of decision trees (Figure 6), which are weak predictors themselves, but together make a much better prediction. These models can serve different purposes, functioning both as classifiers or as regression tools capable of predicting the abundance of a target (y) using predictors $X = (X_a, X_z)$.

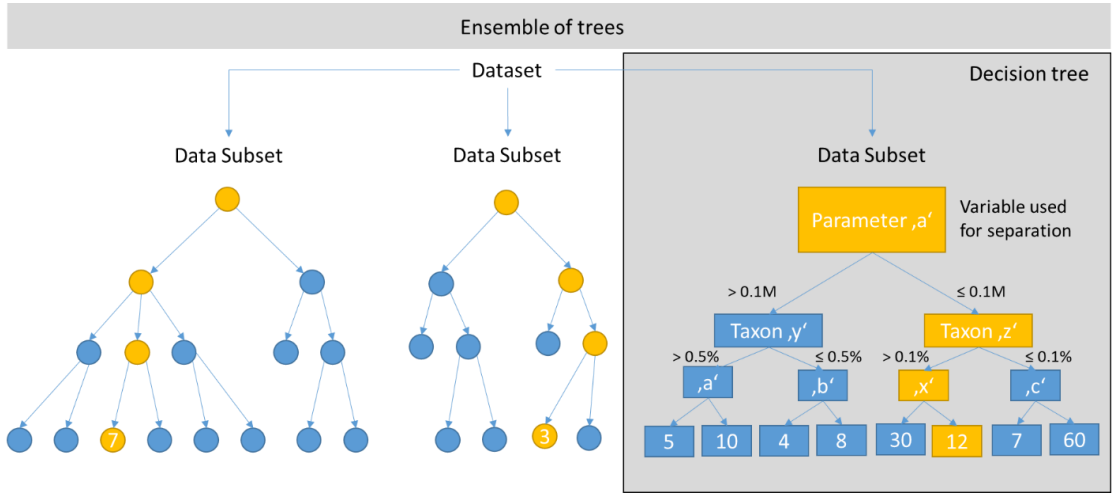


Figure 6. an example decision tree and RF ensemble. *The RF aims to quantify the abundance of the response variable using a combination of microbial community data and independent environmental data. Each tree uses a different subset of samples and features to best predict the response variable in the training data. The splits are determined during training by calculating separation capability (e.g. mean square error or mean absolute error). The variable achieving the highest degree of separation is selected first, in our example tree "Parameter 'a'". While a single tree will overfit on its subset of features, the ensemble of trees will generalize fairly well by averaging the outcomes of all trees in the case of regression. In this example, while the trees predict different numbers, the output of the model would be 7.3.*

In a tree, a first decision is made at a root, parameter 'a' in the example tree (Figure 6). Two edges originate from this and lead to further nodes (Taxon 'y' or 'z') in the example. By making a series of "yes/no" or "over/under" splits, each time maximizing the homogeneity of the target variable in the new nodes and the splitting function. A subset of data that fulfills the edge conditions for a node are used to sprout new edges from the new nodes, again maximizing the homogeneity of the target variable in the subsequent nodes. This continuous

until either (i) a predefined number of splits is made, (ii) the target variable in a node is homogeneous, (iii) splitting does not increase the value of the splitting function.

A single decision tree trained on the entire dataset is prone to overfitting, capturing noise and artifacts. To mitigate this, random forests split the observations and features randomly among multiple trees, promoting diversity and reducing overfitting. To achieve accurate predictions, the model can be optimized by tuning "hyperparameters", such as varying the number of splits made in a tree, the amount of features considered when looking for the best split and the quantity of trees. This can strongly affect the ultimate performance of the model. The optimization process involves training numerous iterations of the model on the training data and evaluating it against the training data, with each iteration featuring a different set of these parameters. The final model, with the best performance, can subsequently be tested against the test set to determine performance on an unseen dataset.

Theoretically, every prediction made by a RF can be reconstructed, practically this is not feasible. Single models often contain hundreds to thousands of trees[103], each of which would have to be viewed manually. Examining individual trees sequentially does not lead to understanding underlying mechanisms; however, we can still gather understanding of the system modelled from the variable importance

Using the found patterns to draw conclusions

Developing a model that accurately predicts a variable '*y*' is the first step. However, understanding the underlying reasons for the model's performance can improve our understanding of the biological relationships. Additionally, important variables can be used to validate a model using pre-existing domain knowledge.

We can achieve this through calculating feature importance, the contribution a specific *X* makes to the final prediction of '*y*'. Multiple measures are available to quantify the importance of a variable, such as permutation-based methods (Mean decrease in accuracy)[104,105] or the reduction of impurity at splits (e.g. Gini importance)[106]. When permuting X_a and thereby breaking its link to *Y*, increases prediction error, the variable is considered important, when the prediction error remains the same, the variable is less important. This is complicated however, when X_a correlates strongly with X_b. In that case, permuting X_a will not increase the error as much as X_b contains similar information. This effect is exasperated when X_a also correlates to (X_c,..., X_u), as few trees remain without one of the correlated variables.

This holds for positively correlated variables, while strongly negatively correlated variables lead to larger increases in prediction error[107].

Feature selection can help alleviate some of these issues. By reducing the amount of features, the chance of spurious correlations is reduced, the model complexity is lowered and the amount of correlation between predictive variables is limited. Recursive feature elimination (RFE) was used in this thesis. RFE iteratively trains a model, assesses feature importance, eliminates the least important feature(s), retrains the model, until a user defined endpoint, either in feature number or performance is attained. The final retained feature(s) of a correlated group can be better compared to features that exhibit fewer or no covariations.

Description of research aims

As a result of climate change *V. vulnificus* is expected to spread poleward, increase in abundance, and benefit from a longer season. Due to the large public health risk associated with these developments, three needs become apparent:

1. Tracking the global distribution of *V. vulnificus* and its spread & improve existing early warning tools
2. Develop tools for quick identification of *V. vulnificus* in the natural environment
3. Develop strategies to mitigate the problem

It is hypothesized that *V. vulnificus* is distributed worldwide along the coastline, with the highest abundance in the (sub)tropical regions and spreading to the poles. This is tested in Chapter I using global archived 16S rRNA data. This data is subsequently used to create an early warning model based on freely available remote sensing data. If a high-risk area is identified by this model, a method for convenient, fast and inexpensive *V. vulnificus* detection in the environment is beneficial for local governments or other stakeholder to take action (Chapter II). If an area is often identified high risk, and *V. vulnificus* is detected regularly, a need exists for mitigation strategies. Two such potential strategies are tested, (i) to reduce *V. vulnificus* abundance through seagrass, or (ii) through the reduction of eutrophication (Chapter III and IV).

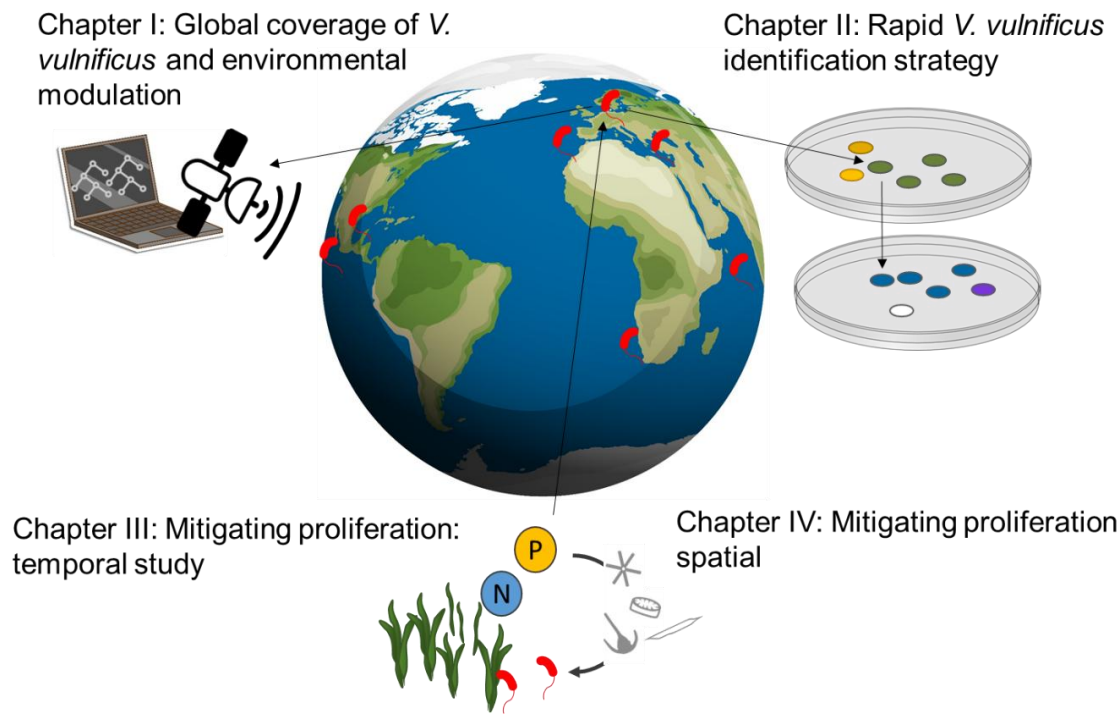


Figure 7. Graphical overview of chapters. *Ranging from prediction (Chapter I), to validation (Chapter II) and mitigation (Chapter III & IV) of* V. vulnificus *proliferation*

Conducted experiments and analysis

For chapter I, 70,640 archived samples from ENA were reanalyzed and the reprocessed sequences were compared to a custom database to identify samples that contained *V. vulnificus*. Subsequently, temperature, salinity, chlorophyll-a (chla) and surface current data was collected using remote sensing data. The goal was to quantify the prevalence of *V. vulnificus* on a global scale and to illustrate a potential poleward spread. As a secondary goal, through using the remote sensing data, a global RF model for the prediction of relative *V. vulnificus* abundance was trained. A tertiary goal was the identification of good microbial predictors of relative *V. vulnificus* abundance using a 2-step RFE-RF approach for the prediction of *V. vulnificus*. The results are described in Chapter I.

For chapter II, a set of potential *V. vulnificus* isolates from the Baltic Sea, largely covering its salinity gradient, collected in 2021 and 2022 was grown on both thiosulphate-citrate-bile salt sucrose (TCBS) agar (Merck KGaA, Darmstadt, Germany) and CHROMagar™

(CHROMagar™, Paris, France). The suspected potentially pathogenic *Vibrio* isolates were identified through multiplex real-time PCR. The results are described in Chapter II

For chapter III, the conditions promoting and regulating *V. vulnificus* over the summer season were tested at three different salinities by sampling a coastal station in Finland, Germany and Denmark between May and October 2022. Sampling was performed inside, at the border of, and outside of seagrass meadows. Absolute and relative abundances of *Vibrio* spp. and *V. vulnificus,* measured using ddPCR, were compared to physico-chemical parameters and the microbial community. Physico-chemical parameters measured were temperature, salinity, oxygen concentration, chla, ammonium, phosphate, nitrate and bacterial counts. The microbial community was analyzed using 16S rRNA amplicon sequencing. Wilcoxon Rank Sum Test was used to compare abundances inside, at the border of, and outside of seagrass meadows. Spearman rank correlation was used to quantify the relationship between *Vibrio* spp. and *V. vulnificus* in the water column and physico-chemical parameters, and a prediction model was constructed using stepwise multiple regression. The potential impact of sediment resuspension on *V. vulnificus* abundance, was estimated by identifying differentially abundant OTUs in the sediment using a negative binomial Wald test and using their relative abundance in water samples as an indicator.

For chapter IV, samples covering the salinity gradient of the Baltic Sea were collected inside, near and far from seagrass meadows at different eutrophication levels. *V. vulnificus* abundance was quantified relatively using 16S rRNA sequencing and absolute numbers were approximated using bacterial cell counts and through ddPCR. The potential mitigating effects of both seagrass and differences in eutrophication were tested. Samples inside, near and outside of seagrass meadows were compared using Wilcoxon Rank Sum Test. A eutrophication index was made using a PCA and *V. vulnificus* was shown to vary with it. A RF model was trained on the combined prokaryotic, eukaryotic and environmental dataset to predict both the approximated *V. vulnificus* cell counts and the *vvhA* gene counts mL^{-1}. Feature importance was determined using an RFE algorithm.

Chapter I: Poleward spread, environmental modulation, and predictive modeling of global *Vibrio vulnificus* abundance

The following chapter was submitted to Communications Earth & Environment as:

David Riedinger, Christiane Hassenrueck, Daniel Herlemann, Matthias Labrenz. Poleward spread, environmental modulation, and predictive modeling of global *Vibrio vulnificus* abundance

David Riedinger's contribution to the written manuscript was ~90%

Abstract

The proliferation of the potentially pathogenic *Vibrio vulnificus* due to climate change poses a growing public health threat worldwide. To evaluate its current distribution and determine predictors of relative abundance, we present a global reanalysis of archived 16S rRNA gene sequencing data. Two random forest models were trained to predict relative abundance using either satellite data or the prokaryotic community. A poleward shift of *V. vulnificus* presence over a 10-year period was revealed. The models indicated that, besides temperature and salinity, both chlorophyll-a and surface currents likely played a role in relative *V. vulnificus* abundance and that non-*vulnificus Vibrio* spp. and *Pseudoaltermonas* spp. were bacterial predictors of *V. vulnificus*. Collectively, this indicates decaying phytoplankton blooms stimulate *V. vulnificus* relative abundance. The poleward trend could accelerate due to climate change-induced changes in phytoplankton dynamics and currents; factors that should be integrated into predictive modelling to estimate future risks.

Introduction

Human-induced greenhouse gas emissions are influencing our climate[1], already resulting in a 1.2 °C increase in the global mean temperature since the pre-industrial era[2]. Despite the Paris Climate Agreement's objective to restrict the rise in global average temperature to two degrees at maximum[108], there is likely a 1.5 °C increase by the early 2030s[1].

The impacts of climate change may be particularly pronounced along the world's coastlines, critical interfaces between natural ecosystems and human populations, representing significant sources of economic activity, tourism, and leisure[4]. Besides the known threat of sea-level rise and increased storm and flood intensity, there is also a rising concern regarding human disease from pathogens[5], such as *Vibrio* spp. These gram-negative bacteria, commonly present in marine and estuarine waters[2], flourish in warm, brackish environments[7].

Of particular concern is *Vibrio vulnificus,* the most lethal human pathogenic *Vibrio*, which favors temperatures >15 °C and mesohaline conditions[8]. Climate change will likely increase both the geographic range and prolong the growth period of *V. vulnificus*, leading to an increase in infections in temperate regions at higher latitudes[7]. An infection can swiftly progress to necrosis, necessitating urgent surgical tissue removal or limb amputation in approximately 10 % of cases[9] and mortality rates as high as 18 %[10], with fatalities often occurring within 48 hours of exposure[11].

Despite the expected increase in range and growing season, the associated poleward spread of infections, and the known lethality, drivers of *V. vulnificus* likely global distribution - or the distribution itself - are not fully understood. Amplicon sequencing, a high-throughput sequencing method, allows for detailed taxonomic composition analysis of prokaryotic communities, but studies are generally limited in their temporal and spatial extent[12-14]. The steadily increasing volume of data available in public archives presents opportunities for large-scale re-analyses of thousands of datasets. This approach also enables the determination of the relative abundance of *V. vulnificus* within the overall bacterial population and assesses the extent of its spread and co-occurrence with other species at an increased spatiotemporal scale

and resolution. However, a comprehensive re-analysis of environmental amplicon data to monitor the spread of this pathogen is still missing.

We inferred the global distribution of *V. vulnificus* from 70,640 archived amplicon sequencing runs, hereafter referred to as datasets, in combination with freely available remote sensing data. This allowed the modelling of relative *V. vulnificus* abundances based on environmental factors at the global scale, the localization of global hotspots, and the observation of shifts in spatial distribution over time.

Results

We investigated the global distribution of *V. vulnificus* using publicly available sequencing datasets. Originally, 2,647,697 archived datasets were screened based on associated metadata, resulting in 70,151 datasets acquired from ENA. An additional 415 datasets were included from Riedinger et al. 2024[12] and 74 from the now public study accession PRJNA1011541 from the Baltic Sea. The analyzed datasets ranged in latitude from 78° S to 83.1° N and spanned to globe laterally. Our analysis revealed the near global presence of *V. vulnificus*, with 1,969 datasets where *V. vulnificus* presence was identified. Among these, 676 datasets were amplified in the V3-V4 region, while 1,321 datasets were amplified in the V4 region. Hotspots of relative *V. vulnificus* abundance were found along the East Asian, US-east, and southern Baltic Sea coastlines (Fig. 1a). The southernmost *V. vulnificus* presence was detected in 2019 (study accession: PRJNA903199), while the northernmost was identified in 2012 (study accession: PRJNA322089). *V. vulnificus* was found increasingly northward every year between 2013 and 2021 except in 2015 and 2020. The datasets with the highest relative abundance were concentrated in the tropics, with 45% of all detections occurring in tropical regions. The majority (80%) of all *V. vulnificus* detections were in the northern hemisphere, exhibiting a median relative abundance of 0.004%. In contrast, detections in the southern hemisphere had a higher median relative abundance of 0.02%. The highest observed relative abundance of *V. vulnificus* (3.23%) was recorded in French Polynesia in the year 2016. The median relative abundance of *V. vulnificus* varied per year, ranging from 0.0029% in 2021 to 0.011% in 2022. However, there was no discernible trend indicating an increase or decrease in relative abundance over time. The global median relative abundance was 0.0056%. The

proportion of datasets where *V. vulnificus* was detected was highest in 2021 and 2018 and lowest in 2013 (Suppl. Fig. 1).

V. vulnificus was detected between sea surface salinities (SSS) of 4.33 and 39.80. Occurrence at the lowest SSS was in 2021 in a river mouth draining into the Baltic Sea near Peenemünde, Germany, and at the highest SSS in a bay near Hurghada, Egypt, in the northern Red Sea. Among the datasets containing *V. vulnificus* where temperature data were acquired, 99.4% had temperatures above 15°C, 99.1% had temperatures above 18°C, and the median temperature was 19.8°C. *V. vulnificus* was detected between 5.4 °C in Portsmouth, UK, in March 2022 (Study accession: PRJNA587708) and 32 °C in a sewage outflow in Jeddah, Saudi Arabia, in October 2015 (Study accession: PRJNA386576).

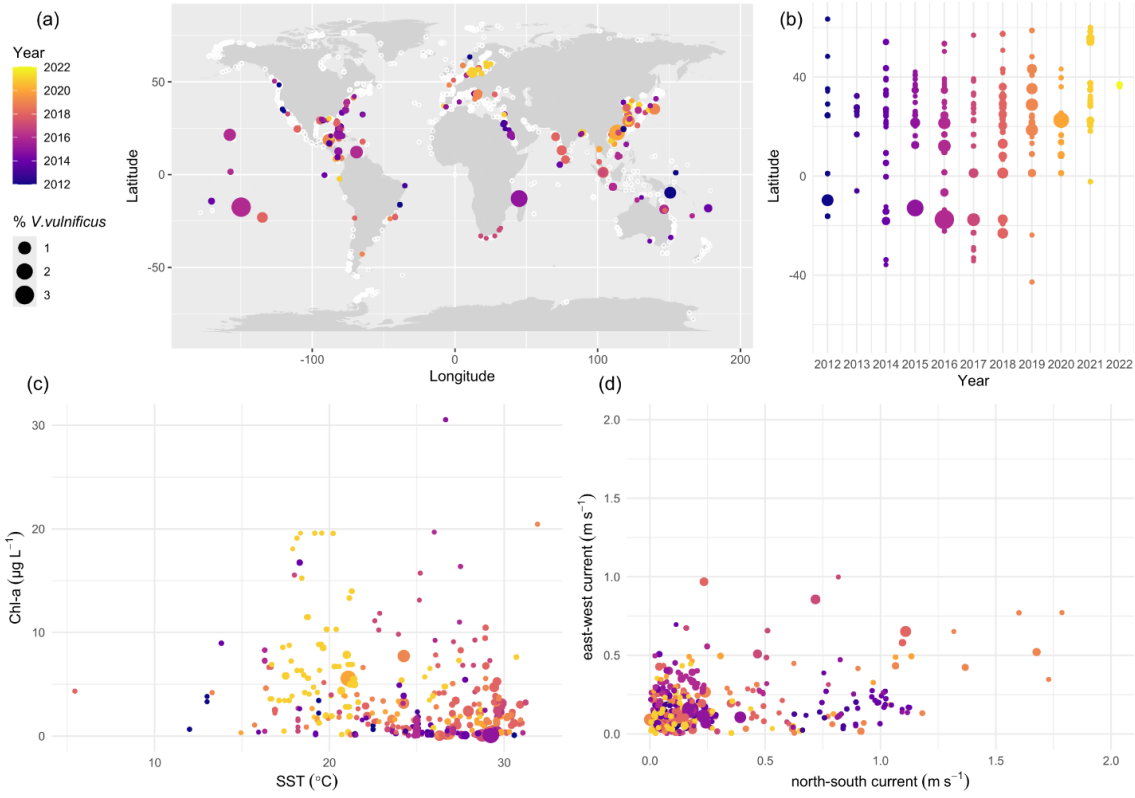


Figure 4. Spatial and temporal distribution of* V. vulnificus *globally. *(a) Worldwide spatial distribution and relative abundance over time, with white dots representing 70,640 screened datapoints. (b) Relative abundance of* V. vulnificus *over time by latitude. (c) Relative abundance of* V. vulnificus *versus SST and chl-a. (d) Relative abundance of* V. vulnificus *versus both north-south and east-west current speeds.*

High relative abundances were typically linked to elevated temperatures, whereas at lower temperatures between 20 and 25°C, similar relative abundances were only reached at elevated concentrations of chl-a (Fig. 1c). Generally, chl-a values ranged between 0.029 and 30.54 µg L^{-1} with a relative abundance of *V. vulnificus* of 0.0032% at the highest chl-a concentration. Considering current speeds (Fig. 1d), 94.3% of the datasets with *V. vulnificus* originated from locations where currents were observed to be under 1 m/s along both the north-south and east-west axes. Generally, higher relative abundances were associated with lower current speeds. No *V. vulnificus* was detected in areas with east-west currents exceeding 1 m/s or north-south currents exceeding 1.8 m/s.

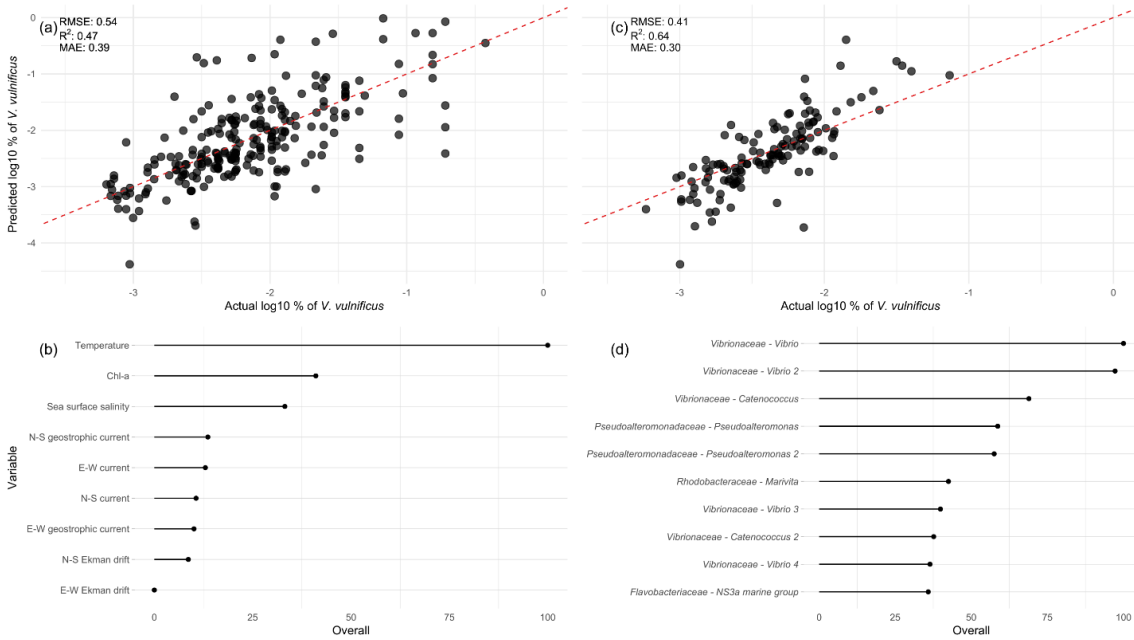
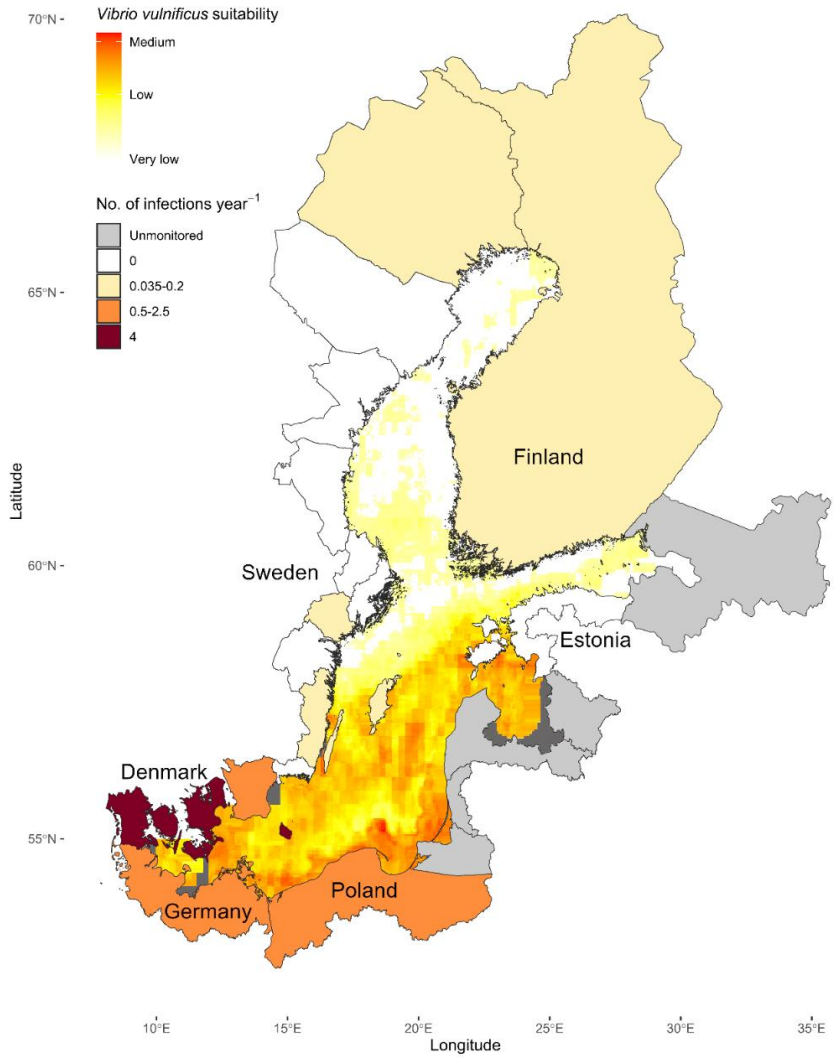


Figure 5. Performance and feature importance of the global* V. vulnificus *model based on remote sensing data (a + b) and on the prokaryotic community (c + d). *(a + c) Model prediction on the y-axis compared to the values in the test set (x-axis), the red-dashed line is the 1:1 line. Performance measures on the test dataset are given. (b + d) Feature importance scaled between 0 and 100, of the top predictor variables.*

The relative abundance of *V. vulnificus* was predicted using a random forest (RF) model based on remote sensing data for 1208 datasets, including those amplified in either the V3-V4 or V4 region, with complete predictor values, and using the prokaryotic community data from 676 datasets amplified only in the V3-V4 region. The model based on the remote

sensing data explained almost 50% of the variance, both on the test dataset and using 10-fold cross-validation (Fig. 2a, SI Table 1). SST was found to be the best predictor, followed by chl-a and SSS (Fig. 2b). E-W Ekman drift held the lowest explanatory power for the relative abundance of *V. vulnificus*. Members of the prokaryotic community predictive for *V. vulnificus* were represented by 400 sequence-based taxonomic features. In the top 10 predictors, six representatives of the Vibrionaceae family were found, four of which belonged to the genus *Vibrio* and two to the genus *Catenococcus*. Further, two Pseudoalteromonadaceae *Pseudoaltermonas*, Rhodobacteraceae *Marivita*, and Flavobacteriaceae NS3a marine group were found to be important.



Previous page: Figure 3. Averaged prediction of* V. vulnificus *suitability in the Baltic Sea over August 2021 based on remote sensing data using the RF. *Infection cases in the Baltic Sea region normalized to official monitoring year or years for which data could be acquired based on Gyraite et al., 2024. The data for gray areas was incomplete and a prediction could thus not be made.*

Using satellite data from August 2021, the global model shown in Figure 2a+b was employed for the Baltic Sea (Fig. 3), as this region provides excellent environmental conditions for *V. vulnificus*. Consequently, it relatively comprehensive data sets are available for Sweden and Germany and to a lesser extent other Baltic Sea counties (Gyraite et al. 2024), which can be used as validation for our model. The *V. vulnificus* suitability was predicted to be highest at the southwestern/southeastern Baltic Sea coast, and the lowest in the Gulf of Bothnia and the Gulf of Finland. In fact, the southern coast is also the region with the highest infection rates recorded over the past three decades. The high-risk areas were associated with moderate SSS (Suppl. Fig. 3b) and high SST (Suppl. Fig. 3c). Current speeds in the Baltic Sea were generally low in August (Suppl. Fig. 3d), and chla values, integrated over the same period, did not exhibit a clear relationship with *V. vulnificus* suitability (Suppl. Fig. 3a). Since phytoplankton blooms are characterized by high chla concentrations over only short periods, the resulting signal when averaged will be minimal. The regions classified as very low-risk have low SSS.

Discussion

In this study, we obtained 1969 global 16S amplicon datasets from archived sequencing data, in which *V. vulnificus* could be detected. This resulted in a global overview of *V. vulnificus* occurrence, spread, and changes in relative abundance. We modelled the relative abundance of *V. vulnificus* using a RF, providing high-resolution prediction of environmental suitability. Globally important prokaryotic predictors were determined using a combination of RFE and RF, indicating that decaying phytoplankton blooms might stimulate *V. vulnificus* proliferation.

The projected poleward expansion of *V. vulnificus* attributed to global warming[7] is visible between 2013 and 2021 when only non-host associated *V. vulnificus* are taken into account. The most northerly detected *V. vulnificus* in 2012 was in association with a copepod, while the northernmost non-host-associated *V. vulnificus* was detected in 2021 in the Baltic

Sea. Within the same timeframe a similar pattern, albeit less clear, emerges for the southward spread. The most southern dataset (study accession: PRJNA903199) originated from a study investigating the adaptation of microbial communities in Patagonia (Argentina) to massive DOM turnover due to decay of the invasive brown algae *Undaria pinnatifida*. The presence at such a southern point could be attributed to this influx of DOM[12]. The sampling density is, however, much lower in the southern hemisphere, limiting the reliability of the perceived southward trend. Sampling bias is unlikely to contribute significantly to the poleward shift in *V. vulnificus* detection due to the substantial number of analyzed datasets in more polar regions each year than the most northerly positive dataset (Fig. 1a & Suppl. Fig. 2). Few sampling points north of 60.12° latitude were recorded in 2021, indicating the possible existence of more polar *V. vulnificus* that were not included in the datasets (Suppl. Fig. 2). There was no notable change in the median relative abundance of *V. vulnificus* per year over the 10-year period. However, the proportion of detections reached its peak in 2021, followed by 2018 (15.3 and 6.3% respectively; Suppl. Fig. 1), coinciding in 2021, with a higher frequency of detections between latitudes 53°N and 66°N.

The global patterns of hotspots on the East Asian, US-east, and southern Baltic Sea coasts coincide with densely populated areas, indicating major risks for the coastal population there. As previous studies have proposed, high population density and the associated eutrophication itself might amplify *V. vulnificus* abundance while also escalating infection risk through heightened exposure[12]. An increase in infections at these hotspots has been noted on the US coast[15], the Baltic Sea[16], and the East Asian coast[17]. An aging population in the Baltic Sea region[18], East Asia[19], and the USA[20] will exacerbate these infection rates.

These problems will only further intensify due to the increased proliferation of *V. vulnificus* at higher temperatures, which was found to be the crucial factor in our study and in existing literature[21]. This explains why the majority of datasets with high relative abundance originated from the tropics. In this tropical region, solar radiation and temperatures are globally at their maximum and constantly high. The distribution therein, however, is likely modulated by primary productivity (approximated by chl-a), SSS, and possibly current speeds at the regional level.

The importance of chl-a in addition to SST, when predicting relative *V. vulnificus* abundance, is in accordance with recent literature[12-13]. The relationship between chl-a and *V. vulnificus* is non-linear and potentially dependent on other environmental parameters, as a direct relationship was not detected. Possibly, as hypothesized earlier, it is not the high concentrations of phytoplankton itself that lead to high relative *V. vulnificus* abundances but rather the breakdown of it and the release of dissolved organic matter[12,22]. This would also explain why lower current speeds generally favor a higher relative abundance of *V. vulnificus*. The more stagnant waters are usually more stratified and consequently, they foster larger surface blooms[23] which after depletion of the local inorganic nutrients, decompose. The organic nutrients, presumably important for the rapid growth of *V. vulnificus*[12,24], subsequently lead to an increase in the relative abundance.

This theory is supported by the predictive power of early microbial colonizers for the relative abundances of *V. vulnificus*. Other members of the genus *Vibrio* are generally early colonizers of new particles and rapidly respond to changing environmental conditions[25]. They belong to some of the fastest-growing bacteria[26] and can form massive, but short-lived, blooms when conditions are favourable[27]. A similar image arises from the high predictive capacity of *Pseudoaltermonas*, which has been previously connected to decaying blooms that also led to an increase in *Vibrio*[28]. Members of this genus have already been shown to cause cell lysis in the algal bloom, leading to its decay[29]. Such an algal bloom and plentiful organic material also stimulates members of the *Roseobacter* clade (Rhodobacteraceae *Marivita*)[30]. Flavobacteriaceae have been previously shown to be a relevant predictor *V. vulnificus* in RF models[12], and are known to be part of the rapid succession of bacterioplankton species after a spring bloom[31]. *Catenococcus* has been previously detected together with *Pseudoalteromonas* and *Vibrio,* in association with Indian white shrimp (*Penaeus indicus*)[32]. Various shrimp species are known vectors for *V. vulnificus*[33], potentially contributing to its spread or increasing its abundance in the water. Additionally, since *Catenococcus* and *Vibrio* are not easily distinguished based solely on16S rRNA gene regions and *Catenococcus* is further disputed as a genus separate from *Vibrio*[34], the earlier inferences about members of the genus *Vibrio* may also apply to Catenococcus.

We introduce a novel model that enables the global prediction of *V. vulnificus* based on near real-time (NRT), globally accessible, and freely available remote sensing data: SST, chl-a, SSS, and ocean current speed data. Our model accounts for nearly 50% of the variance in *V. vulnificus* relative abundances. Modelling is conducted only on datasets where *V. vulnificus* was detected because a dataset without detection does not equate absence of *V. vulnificus*. The lack of detection could be due to insufficient sequencing depth, an incomplete comparison database, or DNA loss from extraction biases. Therefore, absence is ambiguous, while presence is more definitive. While we are introducing a bias, restricting the model to presence-only observations, it is possible to disentangle the factors influencing the proportion of *V. vulnificus* at a much higher resolution, as this approach a priori excludes any conditions that are not conducive to *V. vulnificus*. The model's performance indicates that, despite inherent errors in satellite-based observations[109], it covers a substantial portion of the parameters that influence either *V. vulnificus* abundance or are correlated with it.

Presently, risk assessments are conducted by modeling the environmental suitability for *Vibrio* spp. in general, rather than specifically for the potentially highly pathogenic *V. vulnificus*. While some of these might themselves be good predictors of *V. vulnificus*, the higher specificity of our model is particularly valuable for groups susceptible to a *V. vulnificus* infection. The model proposed by Baker-Austin et al., 2012[7], which is primarily based on SST and validated on historical infection data, is limited to 280 infections for validation as of their publication. The recorded cases of *V. vulnificus* infections are likely an underestimation of the actual number of infections[36]. In our model, data volume is increased by using 1208 global datasets, increasing its robustness. Using ML, we address more intricate interplay among the variables than the generalized linear models currently in use[7]. Primary production and the degradation of biological matter likely influence the proliferation of *V. vulnificus* differently, at various SSTs and current speeds. SST alone did not fully explain the variance in *V. vulnificus* relative abundance, with chl-a and current speed being relevant predictors. These parameters are globally available in NRT and can thus be used for risk prediction along coastlines worldwide. Predictions might improve in areas that recently experienced a large plankton bloom[12], where our model would predict higher risks than the existing model. Areas with high SST, high current speeds, and low nutrient supply exhibit, according to the presented results, a

lower risk than would be deduced solely from SST. Following the findings of our models, it is reasonable to expect that the poleward trend of *V. vulnificus* due to climate will be exacerbated by predicted increases in (harmful) algal blooms at higher latitudes due to climate change[37]. Cyanobacterial blooms, which have been previously implicated in *V. vulnificus* proliferation[12,24], will benefit substantially from the changing climate in brackish regions[38]. This underscores the benefit of incorporating additional parameters beyond SST and SSS into *V. vulnificus* suitability modeling.

Despite the model performance, issues with using remote sensing data to predict *V. vulnificus* remain. Due to a preference for brackish waters, *V. vulnificus* are often found in (small) estuaries where SSS can rapidly change with the tides. This variability makes it challenging for remote sensing methods to accurately capture SSS with only daily measurements. This problem, albeit to a lesser extent, affects all other parameters as well. Additionally, the applied bounding box method can result in the inclusion of satellite data from outside the estuary, which often has much higher salinity levels, leading to inaccuracies. Additionally, the data mining approach could not reliably exclude datasets from such as coastal fish farms, non-surface waters, and sewage outlets which have very different local conditions compared to those detected by remote sensing. However, the extensive amount of data available from archived and satellite sources helps to compensate for the increased noise in the dataset, resulting in robust predictions.

The comparison of predicted environmental suitability for *V. vulnificus* in the Baltic Sea, averaged over August—the month with the highest infection risk—with reported infection cases demonstrates this robustness. Assuming SSS[39], SST[40], phytoplankton dynamics, and current speed[41] patterns are relatively regular, a similar risk profile for the Baltic Sea can be assumed each summer. This aligns with the recorded infections per year, with regions with higher rates also having an elevated environmental suitability for *V. vulnificus*. Monitoring efforts remain incomplete in numerous countries and are entirely absent in significant portions of the eastern Baltic Sea region. Furthermore, the precise etiology of many *Vibrio* spp. infections remains unknown, complicating accurate comparisons. Gyraite et al., 2024 predicted that there were unreported cases in Lithuania and Poland over the last three decades. These

predictions align with the modelled environmental suitability for *V. vulnificus*. Additional validation is advisable using the increasing number of infection cases with a known origin as they become available. The similarity in regional patterns does, however, indicate that the proposed model may have a role in infection prediction as well as *V. vulnificus* environmental suitability prediction.

Methods

The full workflow is schematically represented in Figure 4 outlining the analysis described in more detail in the following section.

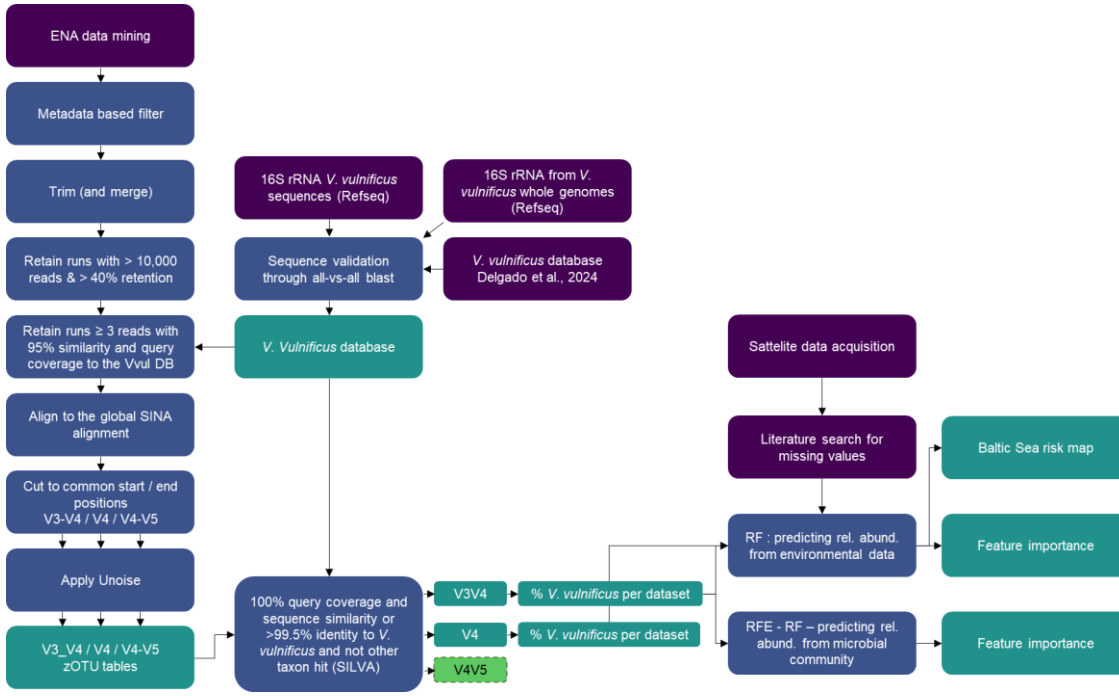


Figure 4. Overview of the workflow used in this study. Initial data acquisition (purple), processing steps (blue), final products (teal).

Obtaining raw reads and initial data selection

Metadata for raw amplicon reads (2,647,697) with specific criteria (tax_tree(408169) AND library_strategy = "AMPLICON", AND library_selection="PCR" AND library_layout = "PAIRED" AND instrument_platform = "ILLUMINA" AND library_source =

"METAGENOMIC") were obtained through an advanced search on the European Nucleotide Archive (ENA) on October 24, 2023. These underwent an initial metadata-driven selection, excluding data sets lacking latitude and longitude information. Subsequently, a positive selection based on presumed environmental origin (NCBI taxon_id) and negative keyword-based selection was performed on completed descriptive columns (full list of terms in the repository), and those with non-marine and/or non-prokaryotic sample submission checklists, host_tax_ids, or non-16S rRNA amplified datasets were excluded. Datasets further than 50km from the coast were removed using a high-resolution coastline (10m) dataset from natural earth (V. 4.1.0). Samples collected for both Riedinger et al., 2024[12] and an associated study (study accession PRJNA1011541) were additionally included. 70,640 runs, for which Fastq files were retrieved, were retained in total for initial analysis. These are referred to as datasets.

Trimming, merging, and filtering of raw reads

Despite restricting the data selection to paired-end runs, data retrieved from the SRA was available in both merged and unmerged form. Merged and unmerged reads were quality trimmed using a sliding window approach implemented in bbduk.sh of the bbmap software suite version-39.01[42], with a window size of 4 and average quality 15, and unmerged datasets were subsequently merged. Reads shorter than 248 base pairs were discarded and datasets with < 10,000 reads or more than > 60% of reads lost during quality control were removed. The remaining datasets were then compared (blast-2.13.0[43]) to a custom database comprising of 238 16S rRNA *V. vulnificus* sequences sourced from RefSeq[44] and from a custom *V. vulnificus* database[45]. Only datasets with ≥3 reads which had more than 95% similarity and query coverage to the custom database were retained and reformatted to fasta files for further analysis. This data collection represented 26357 data sets of likely 16S rRNA origin, which might contain *V. vulnificus*. Datasets were processed using Snakemake version 6.15.1[46] as a workflow manager on the compute cluster Emmy of the Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG), Germany, of the Alliance for National High-Performance Computing (NHR).

Alignment and identification of the amplified region

Using ARB (V. 7.0)[47], positions were identified where *V. vulnificus* could be differentiated from its closest relatives, such as *V. pommerensis* and *V. navarrensis*. This was achieved by comparing *V. vulnificus* type strains to known *V. pommerensis* and *V. navarrensis* sequences and type strains of sister branches (*Vibrio thalassae, Vibrio madracius, Vibrio mediterranei,* and *Vibrio shilonii*). Position 10266 within the SILVA alignment was identified as the most divergent between *V. vulnificus* and the other sequences. Position 15968 in the SILVA alignment further allows for the distinction between *V. vulnificus* and *V. pommerensis* and *V. navarrensis* only while 16421 to 16427 allow for the differentiation between *V. vulnificus* and type strains of the sister branches mentioned above. Consequently, only the V3-V5 regions are suitable for the differentiation between *V. vulnificus* and other *Vibrio* based on the 16S rRNA gene. Retained datasets were aligned to the global SILVA[48] alignment using SINA[49] (V. 1.7.2) and revealed three distinct clusters covering the identified positions (Suppl. Fig. 4), presumably generated using 3 primer pairs binding to *Escherichia coli* 16S rRNA positions 341 – 806, 515 - 806, and 515 - 926, corresponding to the hypervariable regions V3-V4, V4 and V4-V5. These were cut using seqtk[50] (V. 1.3.) to the corresponding SILVA alignment positions (6472 – 23441, 13862 – 23446, and 13862 – 27656). This resulted in 25730 datasets left for analysis, 10703 in the V3-V4 region, 5444 in the V4, and 9583 in the V4-V5.

Denoising and *V. vulnificus* detection

Per amplified region, reads were dereplicated and denoised into zOTUs using the UNOISE[51] algorithm implemented in VSEARCH (V. 2.27.0)[52]. After dereplication, singletons were removed and UNOISE was used to denoise reads into zOTUs (minsize = 10 and alpha = 3). The remaining reads with an abundance between 2 and 9 were mapped to the generated zOTUs[51]. The zOTUs were then compared to the SILVA database (V. 138.1) and the custom *V. vulnificus* 16S rRNA gene database. Only zOTUs with either 100% query coverage and similarity to the custom database or those with > 99.5% similarity and no other species hit in SILVA were classified as *V. vulnificus*. Only datasets with at least three *V. vulnificus* reads

were retained for further analysis. No such zOTUs were found in the V4-V5 datasets and therefore datasets amplified in this region were not included in the further analysis

For all samples where *V. vulnificus* was detected, the relative sequence abundance of all zOTUs identified as *V. vulnificus* were summed and taxonomy was assigned to the remaining zOTUs using DADA2[53] (V. 1.26) assign_taxonomy() with SILVA[48] (V. 138.1) as the reference database. A bootstrap cutoff of 50 was chosen, and all chloroplast, mitochondrial, and eukaryotic zOTUs were removed.

Remote sensing data

To retrieve satellite-based environmental data for datasets where *V. vulnificus* was detected, the E.U. Copernicus Marine Service was used. Chl-a, SST, SSS, Ekman currents, geostrophic currents, and general current values were obtained (Table 1).

Table 1. Data sources for remote sensing data acquisition

Data	Product	DOI
chl-a	cmems_obs-oc_glo_bgc-plankton_my_l4-gap-free-multi-4km_P1D	10.48670/moi-00281
SST	C3S-GLO-SST-L4-REP-OBS-SST	10.48670/moi-00169
SSS	cmems_obs-mob_glo_phy-sss_my_multi_P1D	10.48670/moi-00051
Ekman currents + average currents	cmems_obs_mob_glo_phy-cur_my_0.25deg_P1D-m	10.48670/mds-00327
Geostrophic currents	cmems_obs-sl_glo_phy-ssh_my_allsat-l4-duacs-0.25deg_P1D	10.48670/moi-00148

Only datasets with valid collection dates (a distinct year, month, and day) were included, due to many datasets being coastal there was no value for some of the parameters

from satellite-based measurements. A bounding box method was used, where the values were averaged over a 10x10km square surrounding the original sampling point. When data was missing after remote sensing and screening the metadata available from ENA, the original publication linked to the archived dataset was used to find environmental data if possible.

Machine learning

The relative abundance of *V. vulnificus* was predicted using RF (Caret[54] V. 6.0.94), (i) based on the satellite data, for both hits from V3-V4 and V4 *V. vulnificus* detections and (ii) based on the prokaryotic community using only the V3-V4 datasets where *V. vulnificus* was detected. The inclusion of only those datasets introduces a positive sampling bias and means the models are in essence predicting not the presence, but the relative abundances when the conditions for the presence of *V. vulnificus* were met. The data for both models was split into train/test (0.8/0.2) and validated, using the log 10 transformed data target variable, once on 10-fold CV on the training set and subsequently on the test set. The prediction based on remote sensing data was performed without any further feature selection with 9 predictors on 1208 datapoints (mtry 5, ntree 1000). To improve both prediction accuracy and decrease runtime when using the prokaryotic community, a two-step feature selection was implemented on the train dataset. First zOTUs not present in at least 30 samples were removed from the feature space. The impact on the structure of the total prokaryotic community was estimated to be minimal based on the limited effect on the difference between Bray-Curtis dissimilarity before and after the filter (Suppl. Fig. 5). A further selection was made on the zOTUs with RFE[55] using 10-fold CV. 400 optimal features were selected and used to train a RF (mtry = 20, 1000 trees). Important variables were selected using the varImp() function.

The satellite-based model was subsequently used to predict the *V. vulnificus* suitability in the Baltic Sea in August 2021. Since the training data lacks negative datasets, the model invariably predicts some level of suitability. Consequently, the category labeled "very low" (depicted as white in the plot) corresponds to the lowest 25th percentile of these predictions. This corresponds to a relative abundance of *V. vulnificus* in the prokaryotic community < 0.0014 %. The highest predicted value, defined as medium suitability is 0.0057%, which is 15% of the global average relative abundance.

Code availability

The full analysis code is available under [https://git.io-warnemuende.de/riedinge/datamining_vulnificus_DR_CH.git]. Metadata-based filtering is performed in [DataMining_dataCheck_GlobalVulnificus_round3], bioinformatics processing in [Gitbash_R_Datamining_approach / Snakemake_wf], ML, statistics and plots generation in [Global_vul_rescue_replotting.R / Final_gif_gen_script.R], and remote sensing data acquisition under [Remote_sensing_workflow].

Data availability

A full list of 16S rRNA gene amplicon sequencing data run accessions that were found to contain *V. vulnificus* can be found in the repository. Runs originate from the following study accessions: PRJEB34343, PRJNA623254, PRJNA630412, PRJNA631060, PRJNA632439, PRJDB10706, PRJDB13015, PRJNA707228, PRJDB14309, PRJDB14340, PRJDB16011, PRJEB57240, PRJEB26390, PRJEB27168, PRJEB37200, PRJEB38452, PRJEB33004, PRJEB40757, PRJEB45256, PRJEB8682, PRJEB52828, PRJNA562054, PRJEB49001, PRJNA564579, PRJNA565073, PRJNA576217, PRJNA587892, PRJNA563869, PRJNA592571, PRJNA592522, PRJNA591065, PRJNA575947, PRJNA600742, PRJNA602157, PRJNA587708, PRJNA609879, PRJNA612127, PRJNA612509, PRJNA637472, PRJNA643495, PRJNA645586, PRJNA658993, PRJNA663922, PRJNA665805, PRJNA668061, PRJNA670572, PRJNA674960, PRJNA678525, PRJNA681520, PRJNA629477, PRJNA683924, PRJNA667067, PRJNA688121, PRJNA689652, PRJNA701450, PRJNA718137, PRJNA726636, PRJNA733060, PRJNA723986, PRJNA728349, PRJNA730095, PRJNA555077, PRJNA732997, PRJNA744215, PRJNA747833, PRJNA612045, PRJNA779745, PRJNA782021, PRJNA705154, PRJNA783808, PRJNA784289, PRJNA791535, PRJNA748089, PRJNA752974, PRJNA766413, PRJNA800805, PRJNA771562, PRJNA772072, PRJNA307596, PRJNA780656, PRJNA824287, PRJNA827494, PRJNA791513, PRJNA278075, PRJNA382610, PRJNA836615, PRJNA836712, PRJNA282701, PRJNA279146, PRJNA839924, PRJNA825348, PRJNA830329, PRJNA280923, PRJNA870476, PRJNA872882, PRJNA877068, PRJNA878934, PRJNA888416, PRJNA893097, PRJNA903199, PRJNA876773, PRJNA908141, PRJNA873249, PRJNA928105, PRJNA938291, PRJNA939845, PRJNA928313, PRJNA945202, PRJNA950592, PRJNA909298, PRJNA967682, PRJNA922724, PRJNA506563, PRJNA879770, PRJNA976638, PRJNA1003239, PRJNA380120, PRJNA414763, PRJNA324410, PRJNA338729, PRJNA324417, PRJNA322089, PRJNA351850, PRJNA354655, PRJNA382809, PRJNA358038, PRJNA386576, PRJNA341929, PRJNA388809, PRJNA388812, PRJNA379850, PRJNA247822, PRJNA401168, PRJNA408048, PRJNA408271, PRJNA413618, PRJNA401637, PRJNA419831, PRJNA517146, PRJNA431510, PRJNA431050, PRJNA453627, PRJNA436216, PRJNA472049, PRJNA479721, PRJNA483963, PRJNA486858, PRJNA495084, PRJNA498285, PRJNA498402, PRJNA510085, PRJNA509355, PRJNA525234, PRJNA478695, PRJNA530378, PRJNA530696, PRJNA543129, PRJNA544301, PRJNA546573, PRJNA548204, PRJNA562056, PRJNA548158, PRJNA421139, PRJEB68222

This study has been conducted using E.U. Copernicus Marine Service Information; 10.48670/moi-00281, 10.48670/moi-00169, 10.48670/moi-00051, 10.48670/mds-00327, 10.48670/moi-00148.

Author contributions

David Riedinger and Christiane Hassenrück conceived of the idea and performed the bioinformatics analysis. David Riedinger performed the remote sensing analysis and the machine learning and wrote the manuscript. Matthias Labrenz conceived of the original project of which this was a spinoff and supplied invaluable suggestions on modelling goals. All authors read and critically commented the manuscript

Conflict of interest

The authors declare no competing interests

Acknowledgments

This work resulted from the BiodivERsA project “Pathogenic *Vibrio* bacteria in the current and future Baltic Sea waters: mitigating the problem” (BaltVib), funded by the European Union and the Federal Ministry of Education and Research, Germany (grant 16LC2022A). The authors gratefully acknowledge the computing time made available to them on the high-performance computer "Emmy" at the NHR Center NHR-NORD@Göttingen. This center is jointly supported by the Federal Ministry of Education and Research and the state governments participating in the NHR (www.nhr-verein.de/unsere-partner).

References

1. Masson-Delmotte, V. P. *et al.* Ipcc, 2021: Summary for policymakers. in: Climate change 2021: The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change. (2021).
2. Pritchard, H. D. & Turner, J. State of the global climate in 2020. (2021).
3. Kinley, R. Climate change after Paris: from turning point to transformation. *Clim. Policy* **17**, 9–15 (2017).
4. Cabana, D., Rölfer, L., Evadzi, P. & Celliers, L. Enabling climate change adaptation in coastal systems: A systematic literature review. *Earth’s Futur.* **11**, (2023).
5. Weatherdon, L. V, Magnan, A. K., Rogers, A. D., Sumaila, U. R. & Cheung, W. W. L. Observed and projected impacts of climate change on marine fisheries, aquaculture, coastal tourism, and human health: an update. *Front. Mar. Sci.* **3**, 48 (2016).
6. Baker-Austin, C. *et al. Vibrio* spp. infections. *Nat. Rev. Dis. Prim.* **4**, 1–19 (2018).

7. Baker-Austin, C. *et al.* Emerging *Vibrio* risk at high latitudes in response to ocean warming. *Nat. Clim. Chang.* **3**, 73–77 (2013).
8. Baker-Austin, C., Stockley, L., Rangdale, R. & Martinez-Urtaza, J. Environmental occurrence and clinical impact of *Vibrio vulnificus* and *Vibrio parahaemolyticus*: a European perspective. *Environ. Microbiol. Rep.* **2**, 7–18 (2010).
9. Dechet, A. M., Yu, P. A., Koram, N. & Painter, J. Nonfoodborne *Vibrio* infections: An important cause of morbidity and mortality in the United States, 1997-2006. *Clin. Infect. Dis.* **46**, 970–976 (2008).
10. Ralston, E. P., Kite-Powell, H. & Beet, A. An estimate of the cost of acute health effects from food- and water-borne marine pathogens and toxins in the USA. *J. Water Health* **9**, 680–694 (2011).
11. Baker-Austin, C. & Oliver, J. D. Rapidly developing and fatal *Vibrio vulnificus* wound infection. *IDCases* **6**, 13 (2016).
12. Riedinger, D. J. *et al.* Control of *Vibrio vulnificus* proliferation in the Baltic Sea through eutrophication and algal bloom management. *Commun. Earth Environ.* **5**, 246 (2024).
13. D., B. K. *et al.* Genomic diversity of *Vibrio* spp. and metagenomic analysis of pathogens in Florida Gulf coastal waters following Hurricane Ian. *MBio* **0**, e01476-23 (2023).
14. Möller, L., Kreikemeyer, B., Gerdts, G., Jost, G. & Labrenz, M. Fish as a winter reservoir for *Vibrio* spp. in the southern Baltic Sea coast. *J. Mar. Syst.* **221**, (2021).
15. Archer, E. J. *et al.* Climate warming and increasing *Vibrio vulnificus* infections in North America. *Sci. Rep.* **13**, 1–11 (2023).
16. Brehm, T. T. *et al.* Non-cholera *Vibrio* species — currently still rare but growing danger of infection in the North Sea and the Baltic Sea. *Internist* **62**, 876–886 (2021).
17. Huang, K.-C. *et al.* Distribution of fatal *Vibrio vulnificus* necrotizing skin and soft-tissue infections: a systematic review and meta-analysis. *Medicine (Baltimore).* **95**, (2016).
18. Biermann, U. & Stiller, S. Migration in the Baltic Sea region in the context of demographic change. *J. East-West Bus.* **19**, 105–121 (2013).
19. Bloom, D. E. & Finlay, J. E. Demographic change and economic growth in Asia. *Asian Econ. policy Rev.* **4**, 45–64 (2009).
20. Colby, S. & Ortman, J. M. The baby boom cohort in the United States: 2012 to 2060. (US Department of Commerce, Economics and Statistics Administration, US …, 2014).
21. Baker-Austin, C., Trinanes, J., Gonzalez-Escalona, N. & Martinez-Urtaza, J. Non-cholera vibrios: the microbial barometer of climate change. *Trends Microbiol.* **25**, 76–84 (2017).
22. Greenfield, D. I. *et al.* Temporal and environmental factors driving *Vibrio Vulnificus* and *V. Parahaemolyticus* populations and their associations with harmful algal blooms in South Carolina detention ponds and receiving tidal creeks. *GeoHealth* **1**, 306–317 (2017).
23. Chiswell, S. M., Calil, P. H. R. & Boyd, P. W. Spring blooms and annual cycles of phytoplankton: a unified perspective. *J. Plankton Res.* **37**, 500–508 (2015).
24. Eiler, A., Gonzalez-Rey, C., Allen, S. & Bertilsson, S. Growth response of *Vibrio cholerae* and other *Vibrio* spp. to cyanobacterial dissolved organic matter and temperature in brackish water. *FEMS Microbiol. Ecol.* **60**, 411–418 (2007).
25. Kesy, K., Labrenz, M., Scales, B. S., Kreikemeyer, B. & Oberbeckmann, S. *Vibrio* colonization is highly dynamic in early microplastic-associated biofilms as well as on field-collected microplastics. *Microorganisms* **9**, 76 (2020).
26. Baker-Austin, C., Trinanes, J. & Martinez-Urtaza, J. The new tools revolutionizing *Vibrio* science. *Environ. Microbiol.* **22**, 4096–4100 (2020).
27. Gilbert, J. A. *et al.* Defining seasonal marine microbial community dynamics. *ISME J.* **6**, 298–308 (2012).
28. Yang, C. *et al.* Bacterial community dynamics during a bloom caused by *Akashiwo sanguinea* in the Xiamen sea area, China. *Harmful Algae* **20**, 132–141 (2012).

29. Kim, J. D., Kim, J. Y., Park, J. K. & Lee, C. G. Selective control of the prorocentrum minimum harmful algal blooms by a novel algal-lytic bacterium *pseudoalteromonas haloplanktis* AFMB-008041. *Mar. Biotechnol.* **11**, 463–472 (2009).
30. Budinoff, C. R., Dunlap, J. R., Hadden, M. & Buchan, A. *Marivita roseacus* sp. nov., of the family Rhodobacteraceae, isolated from a temperate estuary and an emended description of the genus *Marivita*. *J. Gen. Appl. Microbiol.* **57**, 259–267 (2011).
31. Hahnke, R. L. *et al.* Dilution cultivation of marine heterotrophic bacteria abundant after a spring phytoplankton bloom in the North Sea. *Environ. Microbiol.* **17**, 3515–3526 (2015).
32. Patil, P. K., Vinay, T. N., Aravind, R., Avunje, S. & Vijayan, K. K. Effect of *Bacillus* spp. on the composition of gut microbiota in early life stages of Indian white shrimp, *Penaeus indicus*. *J. Appl. Aquac.* **35**, 622–632 (2023).
33. Ji, H. *et al.* Occurrence and characteristics of *Vibrio vulnificus* in retail marine shrimp in China. *Food Control* **22**, 1935–1940 (2011).
34. Parks, D. H. *et al.* GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* **50**, D785–D794 (2022).
35. Lunetta, R. *et al.* Remote sensing and geographic information system data integration- Error sources and research issues. *Photogramm. Eng. Remote Sensing* **57**, 677–687 (1991).
36. Baker-Austin, C. & Oliver, J. D. *Vibrio vulnificus*: new insights into a deadly opportunistic pathogen. *Environ. Microbiol.* **20**, 423–430 (2018).
37. Wells, M. L. *et al.* Harmful algal blooms and climate change: Learning from the past and present to forecast the future. *Harmful Algae* **49**, 68–93 (2015).
38. Wurtsbaugh, W. A., Paerl, H. W. & Dodds, W. K. Nutrients, eutrophication and harmful algal blooms along the freshwater to marine continuum. *Wiley Interdiscip. Rev. Water* **6**, 1–27 (2019).
39. Kniebusch, M., Meier, H. E. M. & Radtke, H. Changing salinity gradients in the Baltic Sea as a consequence of altered freshwater budgets. *Geophys. Res. Lett.* **46**, 9739–9747 (2019).
40. Bradtke, K., Herman, A. & Urbański, J. A. Spatial and interannual variations of seasonal sea surface temperature patterns in the Baltic Sea. *Oceanologia* **52**, 345–362 (2010).
41. Jędrasik, J. & Kowalewski, M. Mean annual and seasonal circulation patterns and long-term variability of currents in the Baltic Sea. *J. Mar. Syst.* **193**, 1–26 (2019).
42. Bushnell, B. BBMap: a fast, accurate, splice-aware aligner. (2014).
43. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
44. McEntyre, J. & Ostell, J. The NCBI handbook. *Bethesda Natl. Cent. Biotechnol. Inf.* (2002).
45. Delgado, L. F., Labrenz, M. & Andersson, A. F. *Vibrio* 16S rRNA gene sequences. *Zenodo* (2024).
46. Köster, J. & Rahmann, S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520–2522 (2012).
47. Ludwig, W. *et al.* ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**, 1363–1371 (2004).
48. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
49. Pruesse, E., Peplies, J. & Glöckner, F. O. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**, 1823–1829 (2012).
50. Heng, L. Seqtk: a fast and lightweight tool for processing FASTA or FASTQ sequences.
51. Edgar, R. C. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv* 081257 (2016).
52. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).

53. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
54. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **28**, 1–26 (2008).
55. Gregorutti, B., Michel, B. & Saint-Pierre, P. Correlation and variable importance in random forests. *Stat. Comput.* **27**, 659–678 (2017).

Chapter II: Combined TCBS and CHROMagar analyses allow for basic identification of *Vibrio vulnificus* within a 48 h incubation period in the coastal Baltic Sea

The following chapter was published in the Journal Microorganisms as:

Conor Christopher Glackin, Susann Dupke, Sharath Chandra Thota, David Riedinger, and Matthias Labrenz. (2024). Combined TCBS and CHROMagar analyses allow for basic identification of *Vibrio vulnificus* within a 48 h incubation period in the coastal Baltic Sea. Microorganisms, 12(3), 614.

David Riedinger's contribution to the written manuscript was ~40%

Abstract

With rising infection rates in recent years, *Vibrio vulnificus* poses an increasing threat to public safety in the coastal brackish Baltic Sea. It is therefore important to monitor this organism and assess the *V. vulnificus* infection risk on a more regular basis. However, as the coastline of the Baltic Sea is 8000 km long and shared by nine nations, a convenient, fast, inexpensive, yet efficient *V. vulnificus* identification method is essential. We evaluated the effectiveness of a two-step agar-based approach consisting of successive *Vibrio* isolation and cultivation on thiosulphate-citrate-bile salt sucrose (TCBS) agar and CHROMagar™ *Vibrio* for *V. vulnificus* in comparison with *V. cholerae*, *V. parahaemolyticus*, and *V. alginolyticus*. Our study contains isolates from water and sediment across a broad expanse of the Baltic Sea including 13 locations and two different summers, the time of year during which *Vibrio* infections are usually much more frequent. Confirmation of isolate species identity was carried out using molecular analyses. The two-step agar plating method performed well across different locations and timeframes in correctly identifying *V. vulnificus* by more than 80%, but the sensitivity in other *Vibrio* species varied. Thus, our approach yielded promising results as a potential tool for early *V. vulnificus* detection across a broad timeframe and transect of the Baltic Sea and potentially other brackish environments.

Introduction

Bacteria of the genus *Vibrio* are ubiquitous members of marine ecosystems and occur in coastal, estuarine, brackish, and freshwaters as well as in sediments, often in association with higher organisms [1]. The genus consists of more than 130 species and around a dozen of these have been demonstrated to be human pathogens [2,3,4,5,6]. From these, *Vibrio cholerae* serotypes O1/O139 can cause the well-known disease cholera. Beside those two, the most important potentially pathogenic non-cholera causing *Vibrio* species are the 'big four' species, consisting of *V. cholerae* [different non-O1/O139 serotypes], *V. vulnificus*, *V. parahaemolyticus*, and *V. alginolyticus* [7]. These organisms are common pathogens present in marine and estuarine waters, sediment, or plankton and can cause infections in humans which are usually associated with the consumption of raw or undercooked shellfish or by wound infections in marine or brackish water. Their characteristic disease patterns are gastroenteritis (*V. cholerae*, *V. vulnificus*, and *V. parahaemolyticus*), ear infections (*V. cholerae* and *V. alginolyticus*), wound infections (*V. vulnificus*, *V. cholerae*, *V. alginolyticus*, and *V. parahaemolyticus*), or sepsis (*V. vulnificus*) [2]. Increasing water temperatures can lead to both the increased abundance of *Vibrio* spp. and the occurrence of accordant potential pathogenic *Vibrio* species. Importantly, and this differentiates *Vibrio* spp. in general from the major foodborne pathogens *Salmonella*, *Listeria*, *Escherichia coli* O157, and *Campylobacter*, it has been shown for the USA between 2006 and 2013 that it is the only group currently increasing in occurrence [8]; thus, the behaviour or appearance of members especially of *Vibrio* spp. change in the current environmental conditions. Consequently, with the rise in global sea surface temperatures, there is growing concern about the potential impact of these changes on *Vibrio* populations and the associated public health risks [9].

V. vulnificus infections of humans are often severe [>90% of all cases] and have mortality rates as high as 50%, especially in immunodeficient individuals [2,7,10]. *V. vulnificus* blooms have been recorded on numerous occasions [11,12,13] and it is known that its optimal growth conditions involve water temperatures exceeding around 18 °C and salinity levels ranging from 5 to 25 practical salinity units (PSU) [14]. This provides *V. vulnificus* with optimal growth conditions in the Baltic Sea. The Baltic Sea is a semi-enclosed marginal sea of the Atlantic located in northern Europe, with a coastline of approximately 8000 km and covering an area of 415,266 km^2. Saline inflows through the North Sea produce a 2000 km long lateral surface salinity gradient throughout the whole Baltic Sea, ranging from high salinities [>25] in the transition zone of the Kattegat to low salinities [<5] in the Gulf of Bothnia [15]. The Baltic Sea

faces one of the highest warming rates in marine ecosystems worldwide; thus, it is considered a high-risk environment for *Vibrio* infections [16]. Indeed, infection numbers have increased significantly along the Baltic coast in recent years, particularly during heatwaves [13]. During the extremely warm summers of 1994, 2003, and 2006, for instance, a plethora of reports emerged documenting *Vibrio*-associated wound infections linked to recreational exposure in this area and included numerous fatalities [17]. Besides tourists and travellers, the Baltic Sea has almost 30 million people living within 50 km of its coastline with an increasing susceptible risk [18]. This highlights the need to delve into the development of *Vibrio* monitoring tools or early warning systems for *Vibrio* occurrences in the Baltic Sea.

Crucial for the understanding of *Vibrio* spp. distribution is an accurate identification of *Vibrio* species that most frequently cause infections in humans, but especially of the most harmful *V. vulnificus*. Pinpointing a timeframe of increased risk of *Vibrio* infection is the next step in curbing the increasing numbers of infections and also has wide-ranging ecological and economic effects [2,19]. In the context of the Baltic Sea, where diverse *Vibrio* species coexist, precise as well as simple and cost-effective identification becomes more critical. Established methods, such as isolation on agar plates, already exist and have been widely used for *Vibrio* identification [20,21,22]. However, often these methods may lack the specificity and sensitivity required to distinguish between closely related species [21] or require higher media and labour costs [23]. TCBS was one of the first selective media used for the isolation of Vibrios [24] and is widely used to isolate *Vibrio* from environmental samples, including the Baltic Sea [25,26]. In contrast to other *Vibrio* selective media, cellobiose-polymyxin B-colistin agar and its modified formulas, modified cellobiose-polymyxin B-colistin agar and cellobiose-colistin agar, TCBS is commercially available and is less time-consuming, requiring only a boiling step [20].

Thus, the objective of this study was to test a two-plate thiosulphate-citrate-bile salt sucrose [TCBS] and CHROMagar™ cultivation approach as an easy, cheap, and efficient species-specific tool to identify potentially pathogenic *Vibrio* in the Baltic Sea. Specification was performed on *V. cholerae*, *V. parahaemolyticus*, *V. vulnificus*, and V. *alginolyticus* identifications, with a broader spatial scale for *V. vulnificus*. Cultivation-based analyses were validated by *Vibrio* spp. specific multiplex-PCR or genus-identifying sequencing on various spatiotemporal levels of the Baltic Sea.

Materials and Methods

Sampling Areas

To evaluate the presence of potentially pathogenic *Vibrio* in the Baltic Sea based on the combined agar identification approach, two Baltic Sea monitoring campaigns took place, one in 2021 and one in 2022, incorporating a total of 13 sampling stations. A temporal study took place at four locations (**Figure 1**; locations 1–4, Latitude and Longitude: [54.14636° N, 11.84315° E], [54.15148° N, 11.88636° E], [54.16666° N, 11.96379° E], [54.18248° N, 12.07630° E]) on a 17 km stretch of the northern German coastline across eleven weeks (4 July 2022 to 15 September 2022) and documented *V. vulnificus*, *V. parahaemolyticus*, *V. cholera*, and *V. alginolyticus*. A spatial study focused on *V. vulnificus* abundance over an almost 1000 km range in the Baltic Sea. It took place from 26 July 2021 to 1 September 2021 and encompassed nine sites along the coastlines of Germany, Poland, Finland, and Estonia (**Figure 1**; locations 5–13, see **Supplementary File S3** for coordinates).

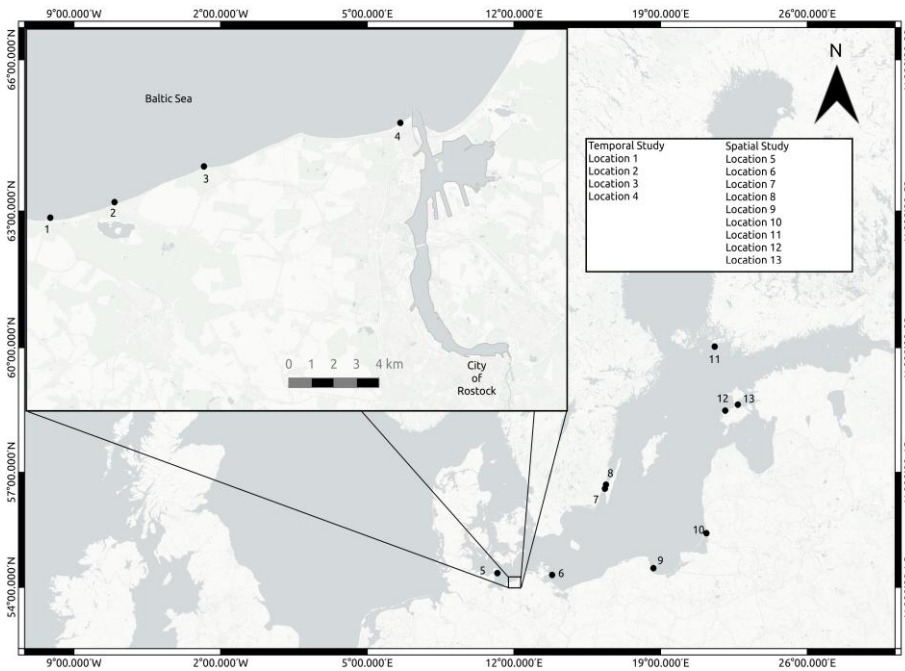


Figure 1. Map of the study area in the Baltic Sea. *Temporal sampling stations were 1-4 and spatial sampling stations were 5-13.*

Environmental Sampling

Water samples were collected in both campaigns, whereas sediment samples were exclusively gathered during the spatial approach. The workflow from the sampling location to molecular identification and comparison is documented in Figure 2. In the temporal study, surface water (less than 50 cm from surface) was collected at a water depth of around 1 m in 6 replicates. At each station, 15 mL Falcon tubes were dipped into the water with rubber gloves. In Heiligendamm (location 1), the samples were taken 150 m from the shore and ca. 3 m depth and were taken every Tuesday at the same time for the duration of the study. In Börgerende (location 2), Nienhagen (location 3), and Warnemünde (location 4), the samples were taken on the beach at around 1 m depth each Monday and Thursday, for harmonisation always three hours after sunrise, for the duration of the study. Further information on these sampling sites can be found in Supplementary File S2. The spatial study sampling sites consisted of coastal locations along the Baltic Sea (Riedinger et al. in revision). For the spatial study, six replicate water samples were collected by SCUBA divers ca. 20 cm above the sediment with 100 mL syringes and six replicates of the top 1 cm of sediment were collected with 50 mL Falcon tubes. Environmental data of the spatial station are available in Supplementary File S3 and at IOWMeta (doi.io-warnemuende.de/10.12754/data-2023-0010).

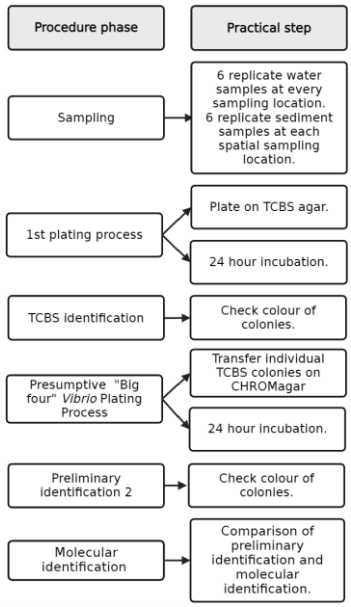


Figure 2. Workflow of the spatiotemporal studies.

All samples were transferred to a 4 °C cooler immediately and stored (maximally 8 h) until processing. Associated physical parameters salinity, temperature, and water depth were measured using a CTD48M (Sea & Sun Technology, Trappenkamp, Germany) during sampling at all stations.

Vibrio spp. Isolation and Culture-Based Identification

For *Vibrio* spp. isolation and identification, thiosulfate citrate bile sucrose (TCBS) agar (Merck KGaA, Darmstadt, Germany) as well as chromogenic agar selective for *Vibrio* spp. (CHROMagar™, Paris, France) were used. Agar plates were prepared according to the manufacturer's protocols. For the isolation of the *Vibrio* spp. from the sediment, the overlying water in the Falcon tube was removed, and after homogenisation, a subsample of 10 g (dry-weight determined accurately after lyophilization) was taken from each sample, and in new sterile 50 mL Falcon tubes, 40 mL of double sterile filtered station water was added. Through five ultrasonic pulses of 10 s at 25% capacity at 5 s intervals from the Bandelin SONOPULS HD 2200.2 (Bandelin, Berlin, Germany), sediment-attached bacteria were detached. Following vortexing and sediment settling, water aliquots of 50, 100, or 200 µL were spread on TCBS agar in six biological replicates. For direct plating, each water sample was thoroughly shaken, and 200 µL of the sample was aseptically inoculated onto a TCBS agar plate, which was subsequently evenly spread across the surface. For indirect plating, 2 mL of each water sample was filtered through a 0.22 polycarbonate filter (Isopore™, Merck Millpore Ltd., Cork, Ireland) and this filter was placed on the agar plates. After incubation for 24 h at 37 °C, colonies were quantified and preliminarily categorised according to Table 1.

Table 1. Species identification of colony forming units (CFUs) according to colour on TCBS and CHROMagar

TCBS CFU Colour	**CHROMCFU Colour**	***Vibrio* Preliminary Identification**
Green	Blue	*V. vulnificus*
Green	Mauve	*V. parahaemolyticus*
Yellow	Blue	*V. cholerae*
Yellow	White	*V. alginolyticus/*Other *Vibrio* Spp.

The preliminarily identified *Vibrio* colonies on TCBS were restreaked onto a ¼ wedge of CHROMagar™ *Vibrio* plates and incubated for 24 h at 37 °C. The final culture-based identification was documented according to the combined colours of colonies grown on TCBS agar and CHROMagar™ *Vibrio* (Table 1).

The colonies of presumptive *V. cholerae, V. parahaemolyticus, V. vulnificus*, and *V. alginolyticus* were suspended in 1 mL of Marine Broth (Roth, Germany). The mixture was left for 24 h at 37 °C. Subsequently, 200 µL of this culture was combined with 300 µL of 50% glycerol, yielding a final concentration of 30% glycerol. These isolates were shock frozen in liquid nitrogen and stored at −80 °C. For recovery, strains were re-cultured either on Columbia agar at 37 °C or in the case of no growth on Difco™ Marine Agar (BD Diagnostics, Sparks, MD, USA) at 28 °C for 24–48 h.

Molecular Identification of *Vibrio* spp.

Genomic DNA from *Vibrio* spp. colonies was extracted using the DNeasy Blood and Tissue Kit (Hilden, Germany) according to the following protocol. A previously frozen bacterial isolate was recultivated on Columbia agar or Difco™ Marine Agar and incubated for 24–48 h. For DNA extraction, an inoculation loop full of colony material was removed by tapping five to ten single colonies. The collected material was transferred into 180 µL ATL buffer and mixed homogeneously by pipetting up and down. Then, 20 µL proteinase K was added, followed by incubation of the sample for 60 min at 56 °C and shaking at 450 rpm in a thermomixer. Subsequently, 200 µL AL buffer was added, and the sample was mixed again by pipetting up and down followed by incubation for another 10 min at 70 °C and shaking at 450 rpm. Further processing was carried out according to the manufacturer's instructions. The purified DNA was eluted in two steps with 100 µL of elution buffer EB by centrifugation at 600 rpm for 1 min each, so that a total of 200 µL DNA eluate was generated. The DNA was stored at 4 °C. The primer and probe sequences for the three multiplex real-time PCR systems are shown in Table 2. The localization of the primers and probes for *V. cholerae* detection correspond to the sequence of the superoxide dismutase (*sodB*) gene of *V. cholerae* NCTC8457 [GenBank AAWD01000215], as well as the sequence of the cholera toxin (*ctxA*) gene of *V. cholerae* strain B [GenBank AY376267]. The localisation of the primers and probes for *V. parahaemolyticus* detection correspond to the sequence of the toxin regulator (*toxR*) gene of *V. parahaemolyticus* strain KP34 [GenBank DQ845170] and for *V. vulnificus* detection to the sequence of the cytolysin–hemolysin (*vvhA*) gene [GenBank AY046900]. Amplicon lengths

for each primer were as follows: 145 bp for *sodB*, 116 bp for *ctxA*, 114 bp for *toxR*, and 118 bp for *vvhA*. In addition, isolates were screened for *V. cholerae* serogroups O1 and O139 [27].

Any bacterial isolate which could not be classified based on the PCR targeting specific markers of *V. cholerae*, *V. vulnificus*, or *V. parahaemolyticus* was subjected to RNA polymerase beta subunit (*rpoB*) sequence determination. Species identification using PCR-based amplification of the *rpoB* gene and analysis of the products were performed as described earlier in Tarr et al. [28] and Schirmeister et al. [29].

Finally, 16S rDNA fragment sequencing was performed on the seven bacterial isolates that could not be identified using *rpoB* gene sequencing. The *rpoB* as well as the 16S rRNA gene fragment sequences were aligned with nucleotide sequences in the GenBank database using the Basic Local Alignment Search Tool (BLAST) search algorithm.

Table 2. List of primers used per species*. All primers were originally used in Messelhäusser et al. [27].*

Target Species	Primer / Probe	Sequence (5´- 3´)	Localisation
V. cholerae	*sodB* f	AAGACCTCAACTGGCGGTA	276-294
	sodB r	CAGCAAAAGAACCGAATGCT	420-401
	sodB TM	**Cy5**-GCAGGTTTGGAACCACACTT-**BHQ-2**	311-330
	ctx f	AGTTCATTTTGGGGTGCTTG	369-388
	ctx r	GGAAACCTGCCAATCCATAA	484-465
	ctx TM	**FAM**-CATCGTAATAGGGGCTACAGAGA-**BHQ-1**	400-422
V. parahaemolyticus	*toxR* f	CCAGAAGCGCCAGTAGTACC	149-168
	toxR r	AAACAGCAGTACGCAAATCG	262-243
	toxR TM	**FAM**-TGTGGCTTCTGCTGTGAATC-**BHQ-1**	181-200
V. vulnificus	*vvhA* f	ACCAAGTTTGGGGCCTAGAT	389-408
	vvhA r	GCTAAGTTCGCACCACACTG	506-487
	vvhA TM	**Cy5**-CCGATCGTTGTTTGACCGTA-**BHQ-2**	440-459

Statistical Analyses

To test for sampling days that could be considered outliers, the Interquartile Range method was implemented in R statistical package version 4.3.2 using ggplot2:: geom_boxplot.

Results

Environmental Parameters

For the duration of the temporal study, the temperature ranged between 15 °C and 23 °C and salinity ranged between 9 and 17 PSU in locations 1–4. In the spatial study (locations 5–13), the temperature ranged between 19 °C and 21 °C and salinity ranged between 6 and 10 PSU. A summary of environmental parameters from the temporal study and spatial study can be found in Supplementary Files S1 and S3.

Identification of Bacterial Isolates on TCBS Agar and CHROMagar™

In the temporal study, a total of 1245 colonies were cultured and isolated on TCBS agar and transferred to CHROMagar™ *Vibrio* plates. Based on the combined colour code identification, 455 of these colonies (37%) were presumed to be *V. parahaemolyticus* whilst 214 (17%), 201 (16%), and 180 (14%) were presumed to be *V. alginolyticus, V. cholerae, and V. vulnificus*, respectively (Figure 3). The remaining 195 colonies (16%) were mixed cultures or unidentified using the colour code for species identification on TCBS and CHROMagar *Vibrio* (Figure 3). The vast majority of colonies originated from water samples taken at locations 2, 3, and 4, with 399, 477, and 315, respectively. The remaining 54 colonies were isolated from location 1 water samples which showed a considerably lower *Vibrio* spp. abundance.

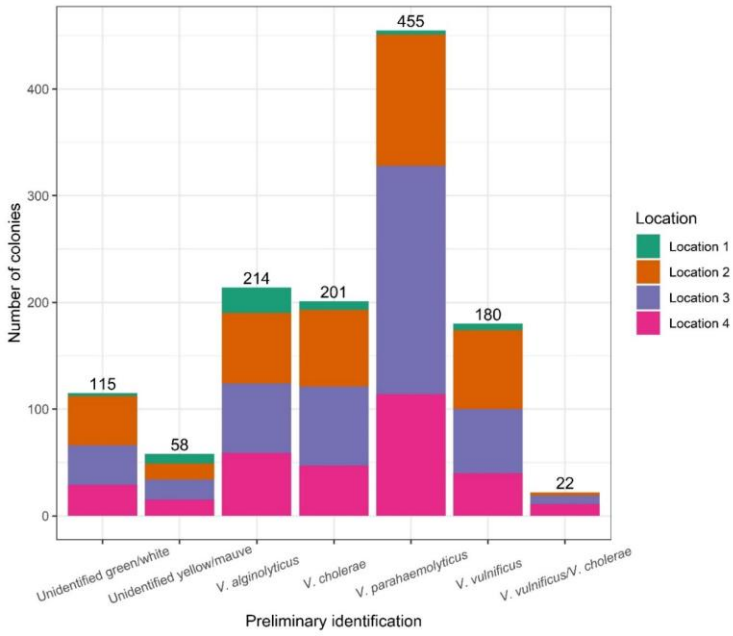


Figure 3. Total number of colonies isolated during the temporal study. *The different colours indicate the location. The preliminary identification is based on the combined TCBS/CHROMagar* Vibrio *colour code (see Table 1).*

Molecular verification of the culture-dependent *Vibrio* spp. identifications yielded very different values for the individual species. While *V. parahaemolyticus* and *V. vulnificus* were correctly identified with values above 80% by the two-plate TCBS agar/CHROMagar™ *Vibrio* approach, it was considerably lower for *V. alginolyticus* with 30%. This method also showed poor predictive ability with a value of 5% identification for *V. cholerae* (Figure 4). Of the misidentified presumptive *V. cholerae* colonies, 36.5 % belonged to *V. aestuarianus* and 51.0% to *V. diazotrophicus* based on molecular analyses. *V. vulnificus* colonies were cultured in each of the eleven weeks and were consistently isolated from samples across the three beach locations (locations 2–4).

In the spatial study from 2021, 86 colonies from eight different locations in the Baltic Sea presumed to be *V. vulnificus* were successfully cultured (Supplementary File S3). From these, 93% were accurately identified by the two-plate TCBS agar/CHROMagar™ *Vibrio* approach (Figure 4). Comparing correctly identified isolates from both the temporal study and spatial study, it became clear that *V. vulnificus* could be regularly identified correctly at a high level using the two-plate method. For *V. parahaemolyticus*, this method also showed high predictive power, correctly identifying 88% of colonies (Figure 4). In contrast, the identification level of *V. alginolyticus* and *V. cholerae* appeared to vary at an already low level (Figure 5). The number of isolates genetically identified in each location is shown in Table S1.

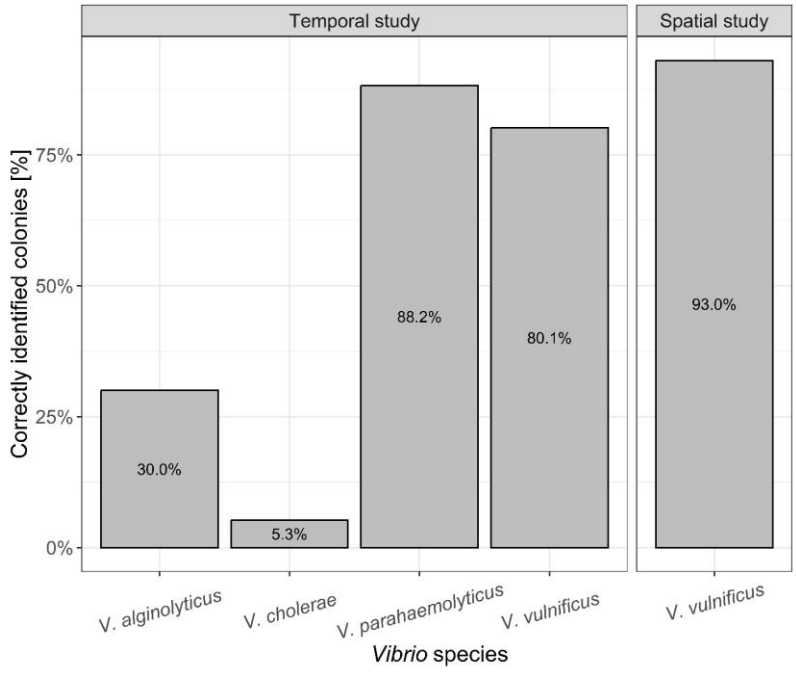


Figure 4. **Vibrio** ***species correctly assigned using the combined TCBS and CHROMagar*** **Vibrio** ***identification method and confirmed using molecular analysis.***

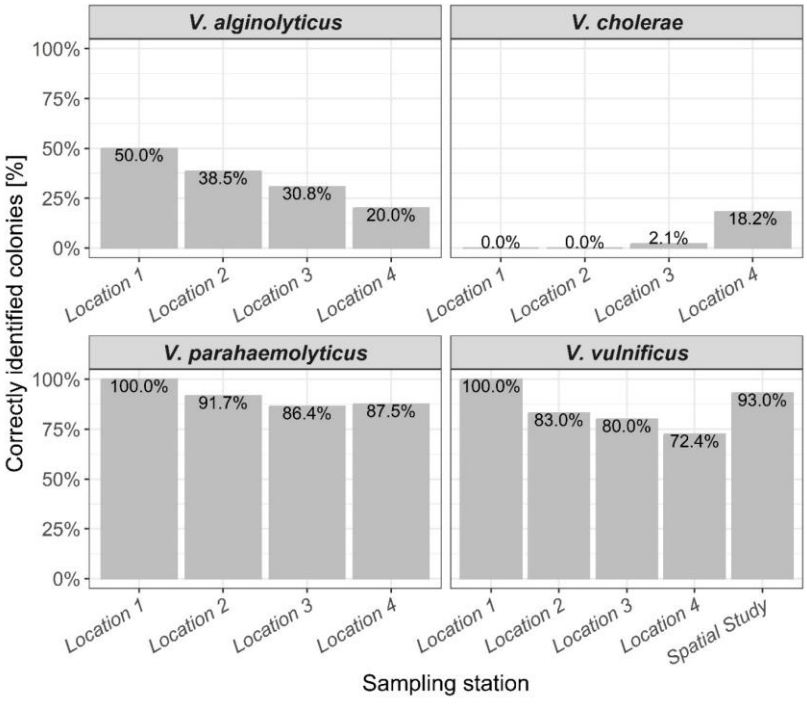


Figure 5. **Vibrio** ***species correctly assigned using the TCBS and CHROMagar*** **Vibrio** ***identification method and confirmed using molecular analysis.*** *Graph is split by species and location.*

Highest *Vibrio* spp. Abundance per Day

Of the total *Vibrio* spp. colonies isolated, it was found that 96% were extracted from water samples taken from locations 2, 3, and 4. Further investigation highlighted that one day in particular yielded a significantly higher number of colonies at these locations (Figure 6). In total, there were 22 days in which samples were taken at each of these locations and 21 July 2022 accounted for 15% (192 colonies) of the total colonies cultured. This is a sharp rise from the total of 14 colonies cultured across all locations on the 18 July. Of the 37 presumptive *V. vulnificus* samples cultured on this day, 24 were identified using molecular sequencing. A total of 18 of these 24 isolates (75%) were confirmed to be *V. vulnificus*, demonstrating similar predictive results to the overall spatiotemporal analysis within this day.

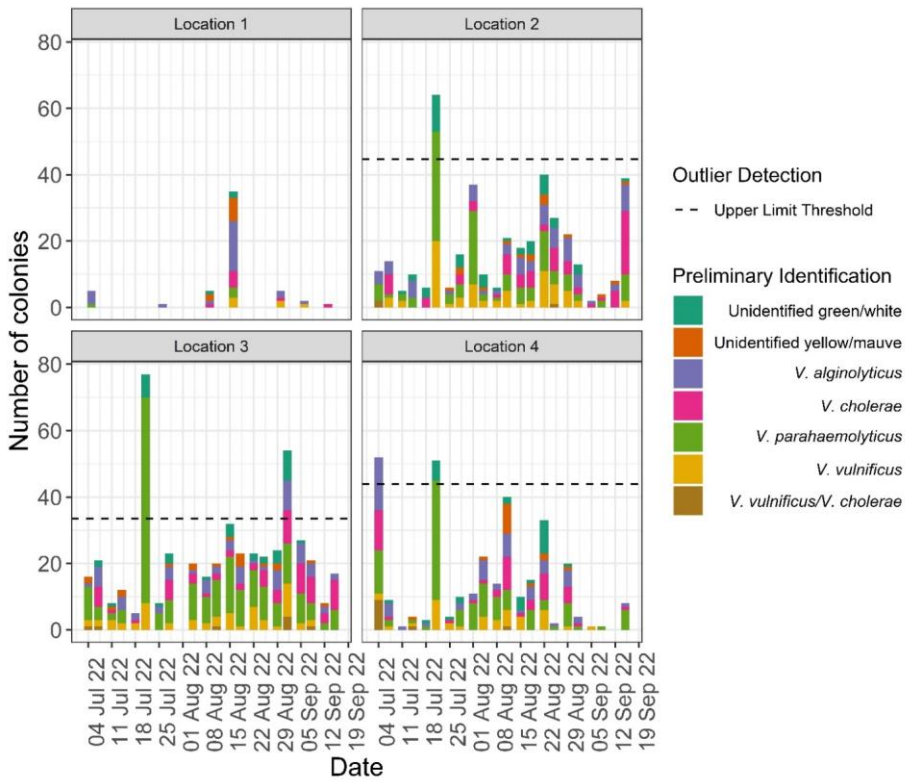


Figure 6. Total species split by location and coloured by presumed identity on TCBS and CHROM-agar* Vibrio *using the colour coding scheme (see Table 1). *Unidentified green/white and unidentified yellow/mauve indicate the colours observed on TCBS and CHROMagar, respectively.* V. vulnificus/V. cholerae *refers to colonies that could not definitively be identified as either species using the colour code. Outlier detection using upper boxplot range is shown (Figure S1).*

Discussion

Routine monitoring for *V. vulnificus* in the Baltic Sea is critical to provide a warning system for the public when the risk of infection is potentially high. This study examined using TCBS agar and CHROMagar™ *Vibrio*-based *V. vulnificus* identifications to achieve this, spanning across two separate spatiotemporal sampling projects. For evaluation, this agar-based method was also tested on *V. cholerae*, *V. parahaemolyticus*, and *V. alginolyticus* in the temporal study. The two-plate agar method proved to be accurate in identifying *V. vulnificus* across a broad range of locations and timeframes within the Baltic Sea, although it had contrasting results when used on other species of *Vibrio*.

V. vulnificus presents the largest threat for open wound infections in the Baltic Sea [30]. In total, 222 *V. vulnificus* samples were genetically identified across twelve locations and two different timeframes in 2021 and 2022, giving a comprehensive overview of this method of identification in the Baltic Sea. The overall correct identification of *V. vulnificus* was 85% and was consistent in correct identification across sampling sites (Figure 5). Thus, our study indicates that the use of TCBS agar followed by CHROMagar™ *Vibrio* to preliminarily identify *V. vulnificus* in the Baltic Sea produces applicable results.

In recent years, there have been numerous attempts to provide the accurate identification of *V. vulnificus* using a culture-based approach, which differ greatly in their success rate. TCBS agar alone has been shown to be not sufficiently selective enough for environmental *Vibrio* samples [31]. Thus, the vast majority of more recent research involves using TCBS agar, CHROMagar™ *Vibrio*, or Cellobiose polymyxin B colistin (CPC) agar with varying results in environmental samples. The highest performing agar approach for *V. vulnificus* so far was a triple plating method [23], producing a 92.8% accuracy on environmental water and oyster samples. Other comprehensive studies include a study by Froelich et al. [32] where CPC+ was used to culture presumptive *V. vulnificus* on agar. Results showed that the average yearly rate of samples confirmed to be *V. vulnificus* ranged from 0% to 45.7%. In another study by Froelich et al. [33] on oyster meat, they compared four different medium methods for presumptive *V. vulnificus*, resulting in correct identification rates between 44% and 81%, as confirmed by PCR. A study in the Mediterranean using TCBS and CPC found, with 3.7% and 7.6%, much lower accuracy levels of these two agar methods, respectively [34].

V. vulnificus aside, *V. cholerae*, *V. parahaemolyticus*, and *V. alginolyticus* are responsible for the majority of other *Vibrio* wound infections in the Baltic Sea [11,13]. In stark contrast to *V. vulnificus*, the percentage of *V. cholerae* correctly identified was 5%. This highlights a huge difference in identifying the two species responsible for the most *Vibrio* infections in the Baltic Sea using the two-plate TCBS agar/CHROMagar™ *Vibrio* approach method. The vast majority of the misidentified bacterial isolates were *V. aestuarianus* and *V. diazotrophicus*, reflecting that the *V. cholerae* colony colours using this agar method are close between the two species. In the case of *V. parahaemolyticus*, the TCBS agar/CHROMagar™ *Vibrio* yielded high correct identification results, whereas the presumptive *V. alginolyticus* colonies were misidentified the majority of the time, further highlighting the mixed results of this agar identification method on different *Vibrio* species.

Spatiotemporal Analysis

In the temporal study, 1245 different presumptive *Vibrio* spp. colonies were grown. In terms of overall *Vibrio* spp. cultivated, station 1 had considerably fewer colonies than the other three locations. The notable difference here was the water depth at the sampling areas. Although all samples were surface water, the difference in depth was around 2 m (3 m vs. 1 m) and the distance from the shoreline was 150 m for location 1 and around 10 m for locations 2–4. There are several hypotheses as to why this discrepancy occurred. Numerous environmental parameters, most notably temperature and salinity, have been associated with contributing to *Vibrio* spp. abundance, depending on species, habitat, and geographic location [35,36,37,38,39,40,41,42,43,44]. The similarity in both temperature and salinity in our project suggests that other environmental parameters or processes may have contributed to the change in abundance. Dissolved oxygen [45,46,47], chlorophyll [39,48,49,50], and plankton [39,51,52,53,54] have also been found to be important in the ecology of *Vibrio* spp. Given that location 1 was further from the shore and sediment, it is also possible that turbidity, increased nutrient loads, and increased sediment bacteria resuspension also played a role in the difference in *Vibrio* spp. abundance between location 1 and locations 2–4.

In the three beach locations, the temporal dynamics were similar in both overall colonies cultured and species correctly identified (Figure 4), with the 22 July showing the highest number of colonies cultivated across all locations. This day was an outlier in all three locations with regard to overall colonies cultured, highlighting the consistency in the results along the 17 km stretch of coast where the *Vibrio* spp. summer surveillance took place. Finding

a significantly higher number of presumptive *Vibrio* isolates in all locations demonstrates the need to further pinpoint *Vibrio* blooms, given that they are present in the marine environment throughout the summer months. This once again points to the need for a simple and fast *Vibrio* spp. identification option. The establishment of a principal monitoring or even an early warning system for *V. vulnificus* in the Baltic Sea is of paramount importance due to the potential public health risks and ecological consequences associated with this pathogenic bacterium. *V. vulnificus* is a well-recognised human pathogen and infections have been documented in increasing numbers along the Baltic coast in recent years [30]. Tourism and economic impacts associated with the danger of infections have been described recently [5,55]. Current research suggests a correlation between sea temperature and *V. vulnificus* abundance and this has major implications for the Baltic Sea given that it is one of the fastest warming seas in the world [9].

This study introduces valuable insights into the suitability of early detection methods for different *Vibrio* species in the region. The reduction of false positive identification in culture-based methods enhances predictive power but can also significantly decrease the cost of laboratory equipment and procedures that are necessary for molecular identification of *Vibrio* spp. This is especially useful in areas or situations where molecular analysis is not possible or is too time consuming, and the price of cultivating colonies in this study using the two-plate TCBS agar and CHROMagar™ *Vibrio* approach was less than USD 3 per sample. An unexpected outcome of the agar identification was the high percentage of *V. aestuarianus* and *V. diazotrophicus* misidentified as *V. cholerae* and this suggests that presumptive identification varies between species and that alternative agar methods may be better in identifying these organisms. The development of improved monitoring or an early warning system for *Vibrio* spp. in the Baltic Sea may be produced using an amalgamation of agar methods and other environmental, chemical, and biological parameters to create predictive models.

Conclusions

Our study documents a quick and straightforward method of isolating presumptive *V. vulnificus* strains using a two-plate TCBS agar and CHROMagar™ *Vibrio* approach. Similar correct identification results, confirmed by molecular analyses, across various locations and timeframes in two different years suggests that this method can be used as a general marker for further research into this topic in the Baltic Sea and probably other brackish environments.

Varying results documented with other *Vibrio* species indicate that this method is not a 'one size fits all' approach to *Vibrio* spp. identified and other agar methods may yield more consistent results.

Supplementary Materials

The following supporting information can be downloaded at **https://www.mdpi.com/article/10.3390/microorganisms12030614/s1**, Information on sampling sites, physical parameters, and CFU identification can be found in Supplementary File S1 and Supplementary File S2. Supplementary File S3 documents the station ID, longitude, latitude, date, sampling depth, temperature, and salinity of the corresponding isolates. Table S1. Colonies per *Vibrio* species from each location of isolates correctly identified on TCBS and CHROMagar and confirmed using molecular analyses. Figure S1. Data are represented as boxplots where the middle line is the median, the lower and upper hinges correspond to the first and third quartiles, the upper whisker extends from the hinge to the largest value no further than 1.5× IQR from the hinge (where IQR is the inter-quartile range), and the lower whisker extends from the hinge to the smallest value at most 1.5× IQR of the hinge. Please see function geom_boxplot in R (ggplot2).

Funding

This work was funded by the German Federal Ministry of Education and Research (BMBF), in the context of Ocean Technology Campus Rostock, grant number 03ZU1107KA (OTC Genomics). It also resulted from the BiodivERsA project 'Pathogenic *Vibrio* bacteria in the current and future Baltic Sea waters: mitigating the problem' (BaltVib), funded by the European Union and the Federal Ministry of Education and Research, Germany (grant 16LC2022A).

Data Availability Statement

The original contributions presented in this study are included in the article and the associated Supplementary material. In addition, data associated with the spatial study are available at IOWMeta (doi.io-warnemuende.de/10.12754/data-2023-0010).

Author contributions

Conceptualization, C.C.G. and M.L.; formal analysis, C.C.G., D.R. and T.S.C.; investigation, C.C.G., S.D., D.R. and T.S.C.; resources, M.L. and S.D.; data curation, C.C.G.,

D.R. and T.S.C.; writing—original draft preparation, C.C.G.; writing—review and editing, C.C.G., M.L., D.R., S.D. and T.S.C.; visualization, C.C.G.; supervision, M.L.; project administration, M.L.; funding acquisition, M.L. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The authors want to thank Heike Benterbusch, Erik Zschaubitz, Clara Nietz, Lara Renner, and the rest of the OTC Genomics sampling team who made it possible to consistently and meticulously collect and process the water samples utilised in this project. We also thank Jana Michel for excellent support of the laboratory work. We are grateful to the crew and captain of the R/V Elisabeth Mann Borgese (EMB283), the research divers of the Leibniz Institute for Baltic Sea Research, the University of Rostock and the Estonian University of Life Sciences, to Jonas Nilsson for assisting with sampling in Sweden, to Victor Fernández-Juárez for supporting *Vibrio* cultivation, and to Adam Woźniczka, Kasia Piwosz, and Jarone Pinhassi for providing lab space.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Le Roux, F.; Blokesch, M. Eco-evolutionary dynamics linked to horizontal gene transfer in Vibrios. *Annu. Rev. Microbiol*. **2018**, 72, 89–110.
2. Baker-Austin, C.; Oliver, J.D.; Alam, M.; Ali, A.; Waldor, M.K.; Qadri, F.; Martinez-Urtaza, J. *Vibrio* spp. infections. *Nat. Rev. Dis. Primers* **2018**, 4, 8.
3. Potdukhe, T.V.; Caffrey, J.M.; Rothfus, M.J.; Daniel, C.E.; Swords, M.E.; Albrecht, B.B.; Jeffrey, W.H.; Waidner, L.A. Viable putative *Vibrio vulnificus* and *parahaemolyticus* in the Pensacola and Perdido Bays: Water column, sediments, and invertebrate biofilms. *Front. Mar. Sci.* **2021**, 8, 645755.
4. Ceccarelli, D.; Amaro, C.; Romalde, J.L.; Suffredini, E.; Vezzulli, L. *Vibrio* species. In *Food Microbiology*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2019; pp. 347–388.
5. Sampaio, A.; Silva, V.; Poeta, P.; Aonofriesei, F. *Vibrio* spp.: Life strategies, ecology, and risks in a changing environment. *Diversity* **2022**, 14, 97.
6. Thompson, F.L.; Austin, B.; Swings, J. The Biology of Vibrios; ASM Press: Washington, DC, USA, 2006.

7. Baker-Austin, C.; Trinanes, J.; Gonzalez-Escalona, N.; Martinez-Urtaza, J. Non-cholera Vibrios: The microbial barometer of climate change. *Trends Microbiol*. **2017**, 25, 76–84.
8. Crim, S.M.; Iwamoto, M.; Huang, J.Y.; Griffin, P.M.; Gilliss, D.; Cronquist, A.B.; Cartter, M.; Tobin-D'Angelo, M.; Blythe, D.; Smith, K.; et al. Incidence and trends of infection with pathogens transmitted commonly through food--Foodborne Diseases Active Surveillance Network, 10 U.S. sites, 2006–2013. MMWR Morb. Mortal. Wkly. Rep. **2014**, 63, 328–332.
9. Baker-Austin, C.; Trinanes, J.A.; Taylor, N.G.H.; Hartnell, R.; Siitonen, A.; Martinez-Urtaza, J. Emerging *Vibrio* risk at high latitudes in response to ocean warming. *Nat. Clim. Chang*. **2013**, 3, 73–77.
10. Jones, M.K.; Oliver, J.D. *Vibrio vulnificus*: Disease and pathogenesis. *Infect. Immun*. **2009**, 77, 1723–1733.
11. Amato, E.; Riess, M.; Thomas-Lopez, D.; Linkevicius, M.; Pitkänen, T.; Wołkowicz, T.; Rjabinina, J.; Jernberg, C.; Hjertqvist, M.; MacDonald, E.; et al. Epidemiological and microbiological investigation of a large increase in vibriosis, northern Europe, 2018. *Euro Surveill.* **2022**, 27, 2101088.
12. Baker-Austin, C.; Trinanes, J.A.; Salmenlinna, S.; Löfdahl, M.; Siitonen, A.; Taylor, N.G.H.; Martinez-Urtaza, J. Heat wave-associated Vibriosis, Sweden and Finland, 2014. *Emerg. Infect. Dis.* **2016**, 22, 1216–1220.
13. Brehm, T.T.; Berneking, L.; Sena Martins, M.; Dupke, S.; Jacob, D.; Drechsel, O.; Bohnert, J.; Becker, K.; Kramer, A.; Christner, M.; et al. Heatwave-associated *Vibrio* infections in Germany, 2018 and 2019. *Euro Surveill*. **2021**, 26, 2002041.
14. Oliver, J.D. The biology of *Vibrio vulnificus*. Microbiol. Spectr. **2015**, 3.
15. Maar, M.; Møller, E.F.; Larsen, J.; Madsen, K.S.; Wan, Z.; She, J.; Jonasson, L.; Neumann, T. Ecosystem modelling across a salinity gradient from the North Sea to the Baltic Sea. *Ecol. Model.* **2011**, 222, 1696–1711.
16. Belkin, I.M. Rapid warming of large marine ecosystems. *Prog. Oceanogr*. **2009**, 81, 207–213.
17. Frank, C.; Littman, M.; Alpers, K.; Hallauer, J. *Vibrio vulnificus* wound infections after contact with the Baltic Sea, Germany. *Euro Surveill.* **2006**, 11, E060817.1.
18. New World Encyclopedia. Baltic Sea—New World Encyclopedia, 2023UTC. Available online: **https://www.newworldencyclopedia.org/p/index.php?title=Baltic_Sea&oldid=1122453** (acce ssed on 15 January 2024).
19. Schütt, E.M.; Hundsdörfer, M.A.J.; von Hoyningen-Huene, A.J.E.; Lange, X.; Koschmider, A.; Oppelt, N. First steps towards a near real-time modelling system of *Vibrio vulnificus* in the Baltic Sea. *Int. J. Environ. Res. Public Health* **2023**, 20, 5543.
20. Choopun, N.; Louis, V.; Huq, A.; Colwell, R.R. Simple procedure for rapid identification of *Vibrio cholerae* from the aquatic environment. *Appl. Environ. Microbiol.* **2002**, 68, 995–998.
21. Di Pinto, A.; Terio, V.; Novello, L.; Tantillo, G. Comparison between thiosulphate-citrate-bile salt sucrose (TCBS) agar and CHROMagar *Vibrio* for isolating *Vibrio parahaemolyticus*. *Food Control* **2011**, 22, 124–127.

22. Tamura, K.; Shimada, S.; Prescott, L.M. *Vibrio* agar: A new plating medium for isolation of *Vibrio cholerae*. *Jpn. J. Med. Sci. Biol.* **1971**, 24, 125–127.
23. Williams, T.C.; Froelich, B.; Oliver, J.D. A new culture-based method for the improved identification of *Vibrio vulnificus* from environmental samples, reducing the need for molecular confirmation. *J. Microbiol. Methods* **2013**, 93, 277–283.
24. Oliver, J.D. Chapter 17 Culture media for the isolation and enumeration of pathogenic *Vibrio* species in foods and environmental samples. In *Rapid Methods in Food Microbiology*; Hope, C.F.A., Adams, M.R., Eds.; Elsevier: Amsterdam, The Netherlands, 1989; pp. 249–269. ISBN 0079-6352.
25. Gyraite, G.; Katarzyte, M.; Schernewski, G. First findings of potentially human pathogenic bacteria *Vibrio* in the south-eastern Baltic Sea coastal and transitional bathing waters. *Mar. Pollut. Bull.* **2019**, 149, 110546.
26. Gomez-Gil, B.; Roque, A. Isolation, Enumeration, and Preservation of the Vibrionaceae. In The Biology of Vibrios; Thompson, F.L., Austin, B., Swings, J., Eds.; ASM Press: Washington, DC, USA, 2006; pp. 13–26. ISBN 9781683671787.
27. Messelhäusser, U.; Colditz, J.; Thärigen, D.; Kleih, W.; Höller, C.; Busch, U. Detection and differentiation of *Vibrio* spp. in seafood and fish samples with cultural and molecular methods. Int. *J. Food. Microbiol.* **2010**, 142, 360–364.
28. Tarr, C.L.; Patel, J.S.; Puhr, N.D.; Sowers, E.G.; Bopp, C.A.; Strockbine, N.A. Identification of *Vibrio* isolates by a multiplex PCR assay and rpoB sequence determination. *J. Clin. Microbiol.* **2007**, 45, 134–140.
29. Schirmeister, F.; Dieckmann, R.; Bechlars, S.; Bier, N.; Faruque, S.M.; Strauch, E. Genetic and phenotypic analysis of *Vibrio cholerae* non-O1, non-O139 isolated from German and Austrian patients. *Eur. J. Clin. Microbiol. Infect. Dis.* **2014**, 33, 767–778.
30. Fleischmann, S.; Herrig, I.; Wesp, J.; Stiedl, J.; Reifferscheid, G.; Strauch, E.; Alter, T.; Brennholt, N. Prevalence and distribution of potentially human pathogenic *Vibrio* spp. on German North and Baltic Sea coasts. *Front. Cell Infect. Microbiol.* **2022**, 12, 846819.
31. Lotz, M.J.; Tamplin, M.L.; Rodrick, G.E. Thiosulfate-citrate-bile salts-sucrose agar and its selectivity for clinical and marine *Vibrio* organisms. *Ann. Clin. Lab.* **1983**, 13, 45–48.
32. Froelich, B.A.; Weiss, M.J.; Noble, R.T. The evaluation of four recent culture-based methods for the isolation and enumeration of *Vibrio vulnificus* bacteria from oyster meat. *J. Microbiol. Methods* **2014**, 97, 1–5.
33. Froelich, B.A.; Williams, T.C.; Noble, R.T.; Oliver, J.D. Apparent loss of *Vibrio vulnificus* from North Carolina oysters coincides with a drought-induced increase in salinity. *Appl. Environ. Microbiol.* **2012**, 78, 3885–3889.

34. Arias, C.R.; Aznar, R.; Pujalte, M.J.; Garay, E. A comparison of strategies for the detection and recovery of *Vibrio vulnificus* from marine samples of the western Mediterranean coast. *Syst. Appl. Microbiol.* **1998**, 21, 128–134.

35. Colwell, R.R. Global climate and infectious disease: The cholera paradigm. *Science* **1996**, 274, 2025–2031.

36. DePaola, A.; Nordstrom, J.L.; Bowers, J.C.; Wells, J.G.; Cook, D.W. Seasonal abundance of total and pathogenic *Vibrio parahaemolyticus* in Alabama oysters. *Appl. Environ. Microbiol*. **2003**, 69, 1521–1526.

37. Grimes, D.J.; Johnson, C.N.; Dillon, K.S.; Flowers, A.R.; Noriea, N.F.; Berutti, T. What genomic sequence information has revealed about *Vibrio* ecology in the ocean—A review. *Microb. Ecol.* **2009**, 58, 447–460.

38. Huq, A.; West, P.A.; Small, E.B.; Huq, M.I.; Colwell, R.R. Influence of water temperature, salinity, and pH on survival and growth of toxigenic *Vibrio cholerae* serovar 01 associated with live copepods in laboratory microcosms. *Appl. Environ. Microbiol.* **1984**, 48, 420–424.

39. Johnson, C.N.; Flowers, A.R.; Noriea, N.F.; Zimmerman, A.M.; Bowers, J.C.; DePaola, A.; Grimes, D.J. Relationships between environmental factors and pathogenic Vibrios in the Northern Gulf of Mexico. *Appl. Environ. Microbiol.* **2010**, 76, 7076–7084.

40. Kelly, M.T. Effect of temperature and salinity on *Vibrio vulnificus* occurrence in a Gulf Coast environment. *Appl. Environ. Microbiol.* **1982**, 44, 820–824.

41. Stauder, M.; Vezzulli, L.; Pezzati, E.; Repetto, B.; Pruzzo, C. Temperature affects *Vibrio cholerae* O1 El Tor persistence in the aquatic environment via an enhanced expression of *GbpA* and *MSHA* adhesins. *Environ. Microbiol. Rep.* **2010**, 2, 140–144.

42. Lobitz, B.; Beck, L.; Huq, A.; Wood, B.; Fuchs, G.; Faruque, A.S.; Colwell, R. Climate and infectious disease: Use of remote sensing for detection of *Vibrio cholerae* by indirect measurement. *Proc. Natl. Acad. Sci.* USA **2000**, 97, 1438–1443.

43. Zimmerman, A.M.; DePaola, A.; Bowers, J.C.; Krantz, J.A.; Nordstrom, J.L.; Johnson, C.N.; Grimes, D.J. Variability of total and pathogenic *Vibrio parahaemolyticus* densities in northern Gulf of Mexico water and oysters. *Appl. Environ. Microbiol.* **2007**, 73, 7589–7596.

44. Tamplin, M.; Rodrick, G.E.; Blake, N.J.; Cuba, T. Isolation and characterization of *Vibrio vulnificus* from two Florida estuaries. *Appl. Environ. Microbiol.* **1982**, 44, 1466–1470.

45. Igbinosa, E.O.; Obi, C.L.; Okoh, A.I. Seasonal abundance and distribution of *Vibrio* species in the treated effluent of wastewater treatment facilities in suburban and urban communities of Eastern Cape Province, South Africa. *J. Microbiol.* **2011**, 49, 224–232.

46. Parveen, S.; Hettiarachchi, K.A.; Bowers, J.C.; Jones, J.L.; Tamplin, M.L.; McKay, R.; Beatty, W.; Brohawn, K.; Dasilva, L.V.; DePaola, A. Seasonal distribution of total and pathogenic *Vibrio parahaemolyticus* in Chesapeake Bay oysters and waters. *Int. J. Food Microbiol.* **2008**, 128, 354–361.

47. Ramirez, G.D.; Buck, G.W.; Smith, A.K.; Gordon, K.V.; Mott, J.B. Incidence of *Vibrio vulnificus* in estuarine waters of the south Texas Coastal Bend region. *J. Appl. Microbiol.* **2009**, 107, 2047–2053.

48. Caburlotto, G.; Haley, B.J.; Lleò, M.M.; Huq, A.; Colwell, R.R. Serodiversity and ecological distribution of *Vibrio parahaemolyticus* in the Venetian Lagoon, Northeast Italy. *Environ. Microbiol. Rep.* **2010**, 2, 151–157.

49. Deter, J.; Lozach, S.; Derrien, A.; Véron, A.; Chollet, J.; Hervio-Heath, D. Chlorophyll a might structure a community of potentially pathogenic culturable Vibrionaceae. Insights from a one-year study of water and mussels surveyed on the French Atlantic coast. *Environ. Microbiol. Rep.* **2010**, 2, 185–191.

50. Julie, D.; Solen, L.; Antoine, V.; Jaufrey, C.; Annick, D.; Dominique, H.-H. Ecology of pathogenic and non-pathogenic *Vibrio parahaemolyticus* on the French Atlantic coast. Effects of temperature, salinity, turbidity and chlorophyll a. *Environ. Microbiol.* **2010**, 12, 929–937.

51. Asplund, M.E.; Rehnstam-Holm, A.-S.; Atnur, V.; Raghunath, P.; Saravanan, V.; Härnström, K.; Collin, B.; Karunasagar, I.; Godhe, A. Water column dynamics of *Vibrio* in relation to phytoplankton community composition and environmental conditions in a tropical coastal area. *Environ. Microbiol.* **2011**, 13, 2738–2751.

52. Martinez-Urtaza, J.; Huapaya, B.; Gavilan, R.G.; Blanco-Abad, V.; Ansede-Bermejo, J.; Cadarso-Suarez, C.; Figueiras, A.; Trinanes, J. Emergence of Asiatic *Vibrio* diseases in South America in phase with El Niño. *Epidemiology* **2008**, 19, 829–837.

53. Rehnstam-Holm, A.S.; Godhe, A.; Härnström, K.; Raghunath, P.; Saravanan, V.; Collin, B.; Karunasagar, I. Association between phytoplankton and *Vibrio* spp. along the southwest coast of India: A mesocosm experiment. *Aquat. Microb. Ecol.* **2010**, 58, 127–139.

54. Turner, J.W.; Good, B.; Cole, D.; Lipp, E.K. Plankton composition and environmental factors contribute to *Vibrio* seasonality. *ISME* J. **2009**, 3, 1082–1092.

55. Novriadi, R. Vibriosis in aquaculture. *Omni Akuatika* **2016**, 12, 1–2.

Chapter III: Temperature, sediment resuspension, and salinity drive the prevalence of *Vibrio vulnificus* in the coastal Baltic Sea

The following chapter was accepted in mBio as:

Víctor Fernández-Juárez, David J. Riedinger, Joao Bosco Gusmao, Luis Fernando Delgado-Zambrano, Guillem Coll-García, Vasiliki Papazachariou, Daniel P.R. Herlemann, Christian Pansch, Anders F. Andersson, Matthias Labrenz, Lasse Riemann: Temperature, sediment resuspension, and salinity drive the prevalence of *Vibrio vulnificus* in the coastal Baltic Sea

David Riedinger's contribution to the written manuscript was ~15%

Abstract

The number of *Vibrio*-related infections of humans, e.g., by *Vibrio vulnificus*, has increased along the coasts of the Baltic Sea. Due to climate change, vibriosis risk is expected to increase. It is, therefore, pertinent to design a strategy for mitigation of the vibriosis threat in the Baltic Sea area, but a prerequisite is to identify the environmental conditions promoting occurrence of pathogenic *Vibrio* spp., like *V. vulnificus*. To address this, we sampled three coastal Baltic sites in Finland, Germany, and Denmark with salinities between 6 and 21 from May to October 2022. The absolute and relative abundances of *Vibrio* spp. and *V. vulnificus* in water were compared to environmental conditions, including the presence of the eelgrass *Zostera marina*, which has been suggested to reduce pathogenic *Vibrio* spp. abundance. In the water column, *V. vulnificus* only occurred at the German station between July and August at salinity 8.1-11.2. Temperature and phosphate (PO_4^{3-}) were identified as the most influencing factors for *Vibrio* spp. and *V. vulnificus*. The accumulation of *Vibrio* spp. in the sediment, and the co-occurrence with sediment bacteria in the water column, indicates that sediment resuspension contributed to *V. vulnificus* abundance. Interestingly, *V. vulnificus* co-occurred with cyanobacteria taxa, as well as specific bacteria associated with cyanobacteria. Although we found no reduction in *Vibrio* spp. or *V. vulnificus* associated with eelgrass beds, our study underscores the importance of extended heatwaves and sediment resuspension, which may elevate the availability of phosphate (PO_4^{3-}), for *Vibrio* spp. levels at intermediate salinities in the Baltic Sea.

IMPORTANCE

Elevated sea surface temperatures are increasing the prevalence of pathogenic *Vibrio* at higher latitudes. The recent increase in *Vibrio*-related wound infections and deaths along the Baltic coasts is therefore of serious health concern. We used culture-independent data generated from three Baltic coastal sites in Denmark, Germany, and Finland from May to October (2022), with a special focus on *V. vulnificus*, and combined it with environmental data. Our temporal model shows that temperature combined with sediment resuspension drive the prevalence of *V. vulnificus* at intermediate salinities in the coastal Baltic Sea.

Introduction

There are more than 130 recognized *Vibrio* spp. (https://www.bacterio.net/genus/vibrio), with approximately a dozen of them being associated with human illness, e.g., *V. alginolyticus*, *V. cholerae*, *V. parahaemolyticus*, and *V. vulnificus*. During recent years, the prevalence of non-cholera *Vibrio*-related wound infections and deaths has increased dramatically at higher latitudes, particularly along the brackish Baltic Sea's salinity gradient that goes from 2 to 25 (1-3). Many of these cases are caused by *V. vulnificus*, commonly known as the flesh-eating bacterium, infecting open wounds, and causing septicemia and necrosis, with a mortality rate of 25%, primarily among elderly, i.e., > 65 years of age (1,4). In the Nordic and Baltic countries, more than 1,055 cases of vibriosis, including *V. vulnificus*, were reported between 2014 and 2018, with most cases occurring during the summer (1,5,6), underscoring that vibriosis is a significant public health challenge in the region and that enhanced preventive measures are needed.

Vibrio spp. are commonly found free-living in the water column and in association with seagrasses, plankton, and aquatic animals (7). *Vibrio* spp. can also be frequently found in biofilms on sediment-trapped particles, which offer nutrient-rich environments (8-10). Seagrass meadows may play a role in promoting the sedimentation of particles and stabilizing the sediment through the reduction of turbulent flow, consequently reducing *Vibrio* spp. abundances in the water column (11, 12), but data are sparse. *Vibrio* spp. distribution exhibits strong seasonality, influenced by factors such as temperature, salinity, eutrophication, phytoplankton biomass, and turbidity (7, 13–17). The pathogen *V. vulnificus* thrives in brackish waters with salinities ranging from 5 to 25, and temperature is suggested to be the most influential factor affecting its presence and proliferation (15, 16). The increase in Baltic Sea

surface temperatures is more than three times higher than the global ocean warming average (18), and global warming will conceivably stimulate the proliferation of pathogenic *Vibrio* spp., e.g., *V. vulnificus*, in the region, with a significant risk of proliferation occurring above 20 °C (19).

Therefore, the pathogenic "*Vibrio* problem" in the Baltic Sea is predicted to be exacerbated by future climate change (20–22). From a public health perspective, there is a pertinent need for identifying the environmental conditions promoting outbreaks of pathogenic *Vibrio* spp., e.g., *V. vulnificus*, and ultimately to determine a strategy for mitigation of the vibriosis threat in the Baltic Sea area. In this study, we used culture-independent data generated from three Baltic coastal sites in Denmark, Germany, and Finland (from within to outside of eelgrass meadows, the water column, and sediment) with contrasting salinity and nutrient levels from May to October (2022) to determine drivers of *Vibrio* spp., with a special focus on the pathogen *V. vulnificus*.

Results

Temporal dynamics of environmental and biological factors

From May to October 2022, we conducted monthly sampling along the Baltic Sea salinity gradient in Denmark (13-21.3), Germany (8-11.2), and Finland (6.3-6.5) (Table 1, Fig. 1A, B). The sea surface temperature (SST) ranged from 7 to 21.5 °C (Fig. 1C). The highest Chl *a* concentration of about 5 mg m^{-3} was observed in late summer and fall in Finland and Germany (Fig. 1D). Phosphate (PO_4^{3-}) and nitrate (NO_3^-) exhibited a consistent pattern, both peaking during July and August at the German station, with concentrations reaching 0.16 µM and 0.14 µM, respectively (Fig. 1E, F). On the other hand, ammonia (NH_4^+) displayed a west-east gradient, reaching a maximum concentration of 0.3 µM (Fig. 1G). The highest bacterial abundance was from July to September in Denmark and Germany (ca. 10^6 cells ml^{-1}; Fig. 1H). All the water metadata are provided in Table S1.

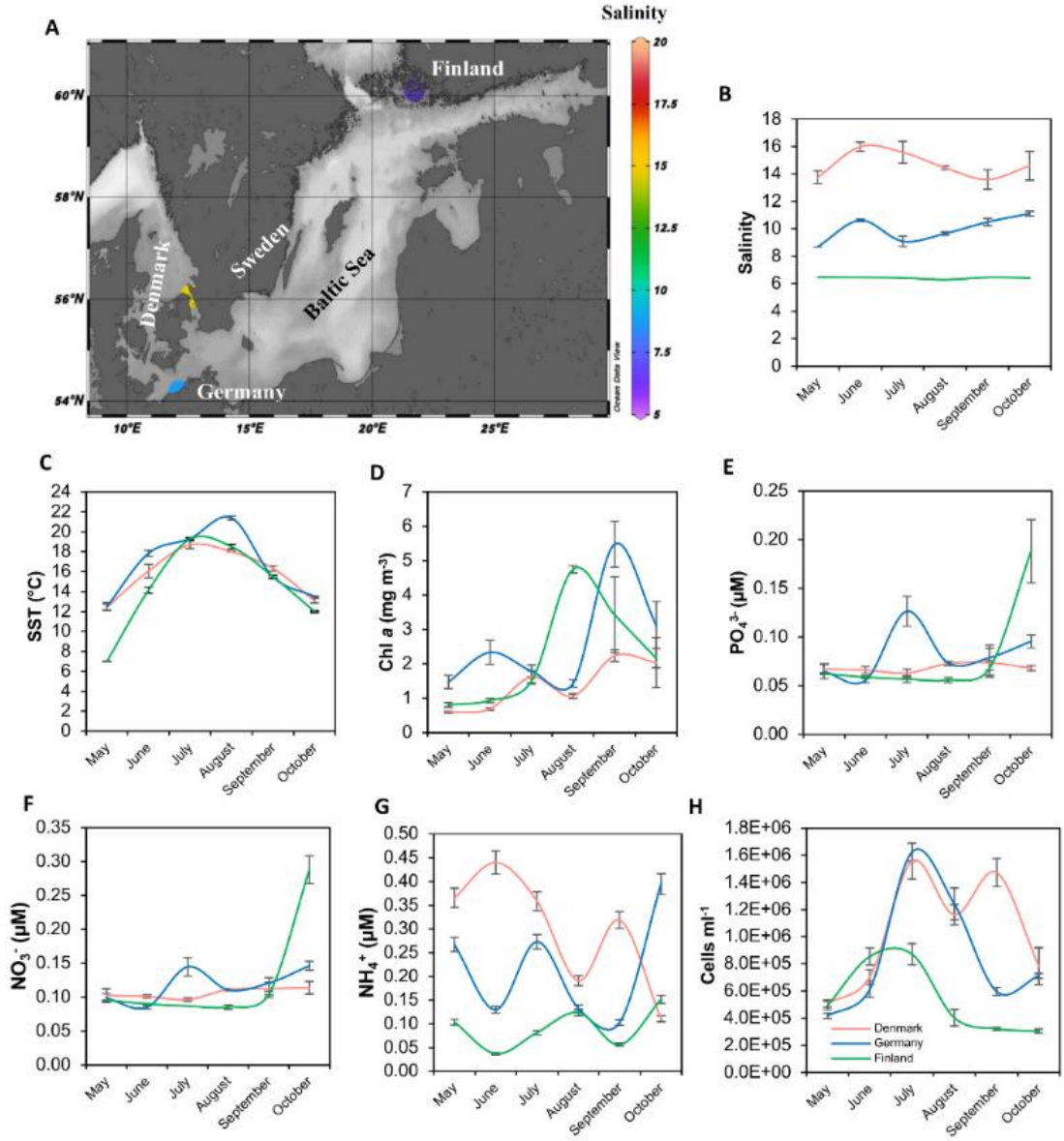
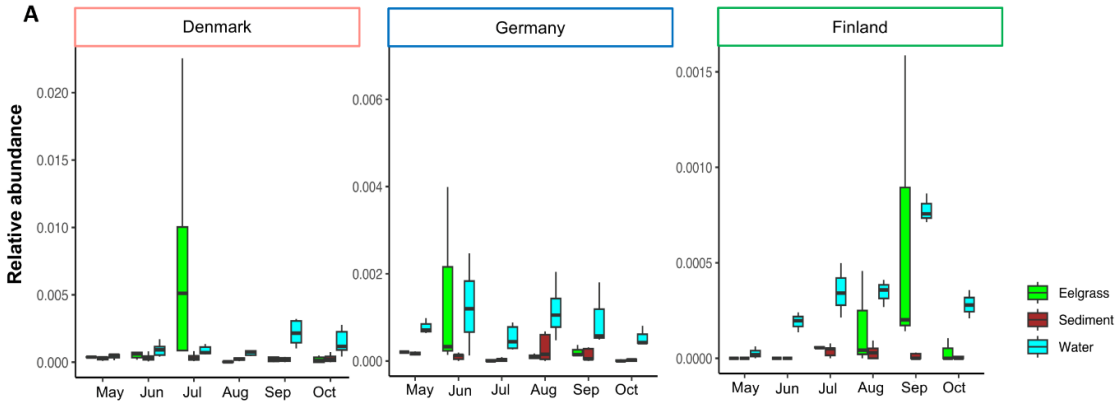


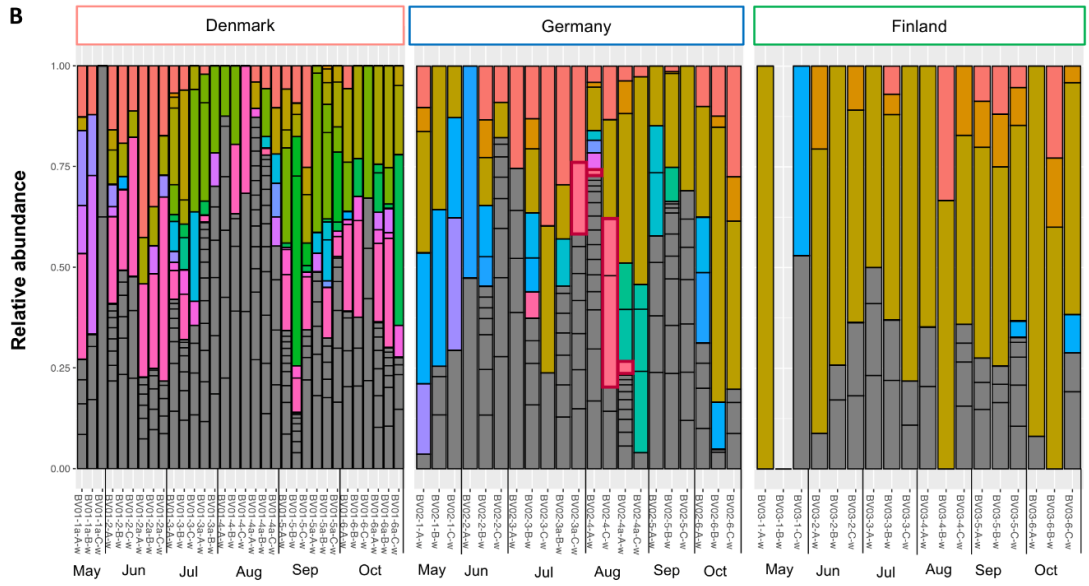
Figure 1. Sampling sites and temporal dynamics of selected parameters in the water column. *A) Map showing sampling stations in Denmark, Germany, and Finland. Seasonal dynamics of B) salinity, C) sea surface temperature (SST, °C), D) chlorophyll a (Chl a, mg m^{-3}), E) bacterial abundance (cells ml^{-1}), and F) PO_4^{3-} (µM). Values are the average of the nine replicates from the substations ("A", "B", and "C") +- the standard error between the replicates. The average data from the two monthly sampling points in Denmark and Germany (July and August) is presented to enable comparisons with the Finnish station. In (A), the color codes represent the average value in the water column across the sampling period. Detailed information about the stations and substations is provided in Table S1.*

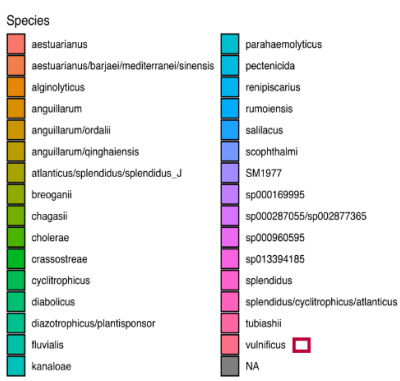
Vibrio spp. composition and diversity in the Baltic Sea

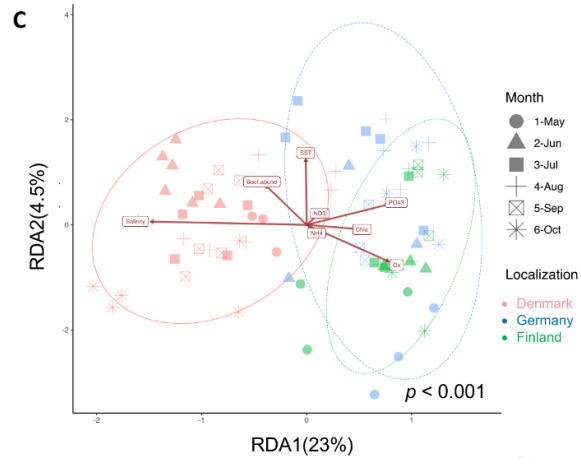
Based on Illumina analyses, we generated a total of 143,935 ASVs from 75 water, 204 sediment, and 78 eelgrass samples collected from May to October 2022. Importantly, triplicate DNA water samples from each substation were combined, resulting in one DNA water sample sequenced per substation and sampling point. The relative abundance of *Vibrio* spp. seemed to increase in Finland only during summer (Fig. 2A). There was no notable enrichment in relative abundance of *Vibrio* spp. in sediment and eelgrass, in comparison to water samples where *Vibrio* spp. accounted for an average of 0.03% of the total microbial community (Fig. 2A). The Shannon diversity of *Vibrio* spp. was higher in the water compared to the eelgrass and sediment environments (Fig. S1A). Notably, there was a significant difference in *Vibrio* spp. composition between water and sediment samples (Fig. S1B).

The 236 ASVs assigned as *Vibrio* spp. in the water column were annotated into 31 different species (Fig. 2B). The *Vibrio* community in the water differed between stations and changed over the season (Fig. 2B, C; $p < 0.001$). Indeed, the *Vibrio* spp. Shannon diversity increased across all stations during the summer season, with the Finnish station, i.e., with the lowest salinity, having the lowest diversity (Fig. S1C). Over the season, in Denmark, the *Vibrio* community was dominated by "ASV 6,774" (*V. cyclitrophicus*), "ASV 7,595" (*V. Atlanticus* or *V. splendidus*), and "ASV 7,725" (*V. chagasii*), while "ASV 1,107" (*V. anguillarum* or *V. ordalii*) and "ASV 14,013" (*V. anguillarum*) dominated in Germany and Finland (Fig. 2B). These ASVs accounted for ca. 40% of the difference in *Vibrio* communities between Denmark, Finland, and Germany (Fig. 2C; SIMPER analysis). SST, salinity, inorganic nutrients (PO_4^{3-}, NO_3^-, and NH_4^+), Chl *a*, bacterial abundance, and oxygen explained 24% of the changes in the *Vibrio* community over the season (Fig. 2C; $F = 2.45$, $p < 0.001$). The "ASV 38,490" and "ASV 30,041", annotated as *V. vulnificus*, were detected between July and August in Germany and accounted for an average of 5% and up to 50% of the *Vibrio* community (approximately 0.01% of the total microbial community). However, no ASVs affiliated with *V. vulnificus* were detected in Denmark or Finland.


A
Denmark
Germany
Finland
Relative abundance
May
Jun
Jul
Aug
Sep
Oct
Eelgrass
Sediment
Water


B
Denmark
Germany
Finland
Relative abundance
May
Jun
Jul
Aug
Sep
Oct


Species
aestuarianus
aestuarianus/barjaei/mediterranei/sinensis
alginolyticus
anguillarum
anguillarum/ordalii
anguillarum/qinghaiensis
atlanticus/splendidus/splendidus_J
breoganii
chagasii
cholerae
crassostreae
cyclitrophicus
diabolicus
diazotrophicus/plantisponsor
fluvialis
kanaloae
parahaemolyticus
pectenicida
renipiscarius
rumoiensis
salilacus
scophthalmi
SM1977
sp000169995
sp000287055/sp002877365
sp000960595
sp013394185
splendidus
splendidus/cyclitrophicus/atlanticus
tubiashii
vulnificus
NA


C
RDA2(4.5%)
RDA1(23%)
p < 0.001
Month
1-May
2-Jun
3-Jul
4-Aug
5-Sep
6-Oct
Localization
Denmark
Germany
Finland

Previous page: Figure 2. **Vibrio *community composition and dynamics from May to October (2022) in water and sediment at the Danish, German, and Finnish stations.*** *A) Relative abundance of* Vibrio *spp. on the Continued Figure 2: water column, sediment, and eelgrass microbial population through the season. Values are the median and vertical black lines indicate the position of the lower and upper quartiles of all the values collected in the three substations during the whole campaign in eelgrass, sediment, and water, and in the water column. B)* Vibrio *communities in the water column. Sample ID clustered by month is shown on the plot. Two monthly sampling points were made in Denmark and Germany (July and August), and they are represented by an "a". Within each color, each segment represents the relative abundance of amplicon sequence variants (ASVs) having the same taxonomy. NA represents* Vibrio *spp. that could not be annotated to species level.* Vibrio vulnificus *is highlighted in the plot by red squares. C) Redundancy analyses (RDA)-constrained ordination plot of the Vibrio spp. community in the water column. Salinity, sea surface temperature (SST, °C), oxygen (Ox, mg* l^{-1}*), PO_4^{3-} (µM), NH_4^+ (µM), NO_3^- (µM), and bacterial abundance (cells* ml^{-1}*) were chosen as constraint variables. Eigen-values are 17.09 for RDA_1 and 4.02 for RDA_2. The significance of the RDA model was assessed by an ANOVA-like permutation test, and ellipses enclose sample groups (Denmark, Germany, and Finland, n = 75). For (C), clr-transformed data was used to perform the analysis.*

Temporal dynamics of *Vibrio* spp. and *V. vulnificus*

Gene copies per ml seawater, i.e., 16S rRNA (*Vibrio* spp.) and *vvhA* (*V. vulnificus*), were quantified as a proxy for the prevalence of *Vibrio* spp. and *V. vulnificus*. They varied significantly over the season (Fig. 3A-C, $p < 0.05$), and correlated with the ASVs abundance – calculated from the relative abundance multiplied by total bacterial abundance (cells ml^{-1}) (23) – ($r^2 = 0.61$ and 0.63, respectively; Fig. S2A, B). In Denmark, *Vibrio* spp. peaked in July (up to 1,000 copies ml^{-1}) and September (up to 5,000 copies ml^{-1}, Fig. 3A). In Germany, the highest abundance of *Vibrio* spp. was observed from July to September, reaching ca. 2,000 copies ml^{-1} (Fig. 3B). In Finland, *Vibrio* spp. reached their peak levels in July and August, but only reached 800 copies/ml (Fig. 3C). Importantly, and corroborating the 16S rRNA gene sequencing results, *V. vulnificus* was only quantifiable in Germany, peaking in summer with up to 30 copies ml^{-1} when SST values ranged from 19.2 to 21.4 °C (Fig. 3B), while it was below the detection limit at the Danish and Finish stations.

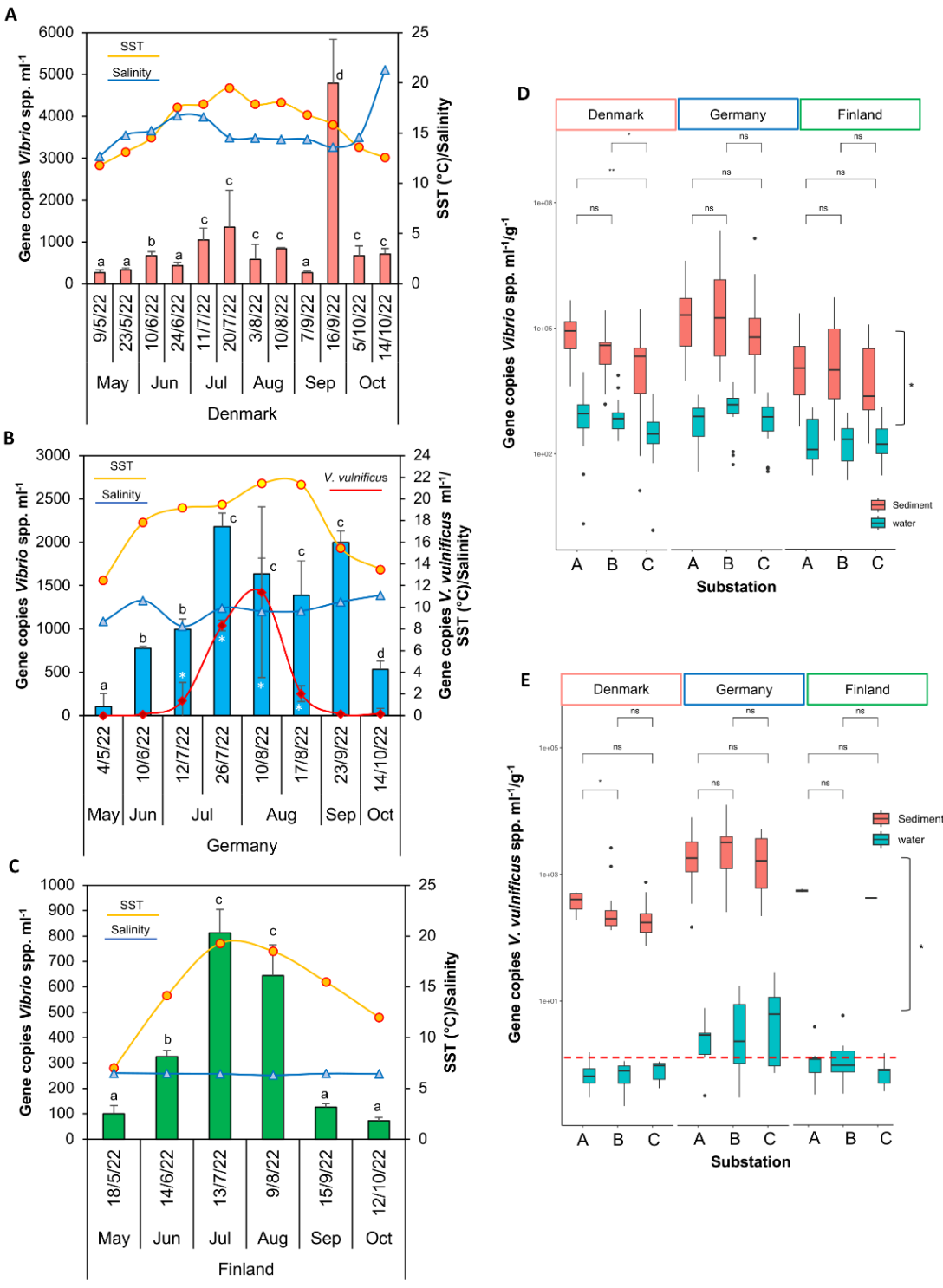
The sampling design was based on earlier findings suggesting a negative impact of *Z. marina* on *Vibrio* spp., including pathogenic species (12). However, in our study, in Denmark, substations "A" and "B", located in the middle of the eelgrass bed and 15 m away, respectively, had higher levels of *Vibrio* spp. compared to substation "C", located 100 m away from the eelgrass, both in the water column and sediment (Fig. 3D). In contrast, no discernible impact was observed on *V. vulnificus* in the water column (i.e., in the German station) (Fig. 3E).

The abundance of *Vibrio* spp. and *V. vulnificus* genes were 100- and 1000-fold higher, respectively, in the sediment than in the water (per weight [g] vs. per volume [ml]; Fig. 3D, E), and abundances in sediments correlated with abundances in water (Table 2). Sediment composition differed between stations, where the German sediment contained slightly more fine materials, i.e., clay and silt (Table S2), and the highest *Vibrio* spp. abundance (> 10^6 gene copies g^{-1} during the summer months, i.e., 10 times more than in Denmark and Finland, Figs. 3D, S3A-C). *Vibrio vulnificus* had the highest abundance in Germany, with > 10^3 gene copies g^{-1} sediment in August. It was also detectable in Denmark, with up to 10^2 gene copies g^{-1} sediment, but was absent in Finland (Fig. S3A, B). Interestingly, in the sediment *V. vulnificus* remained quantifiable below 15 °C in Germany and Denmark (Fig. S3A, B).

Environmental drivers and predictors of *Vibrio* spp. and *V. vulnificus*

We investigated the environmental factors associated with the abundance of *Vibrio* spp. and *V. vulnificus* – gene copy number– within the water column across the sampled areas. A principal component analysis (PCA) revealed distinct environmental characteristics among the stations (Fig. 4A). Correlations between the variables analyzed are provided in Table S3. The abundances of *Vibrio* spp. and *V. vulnificus* were positively correlated with SST and negatively correlated with Secchi depth (Table 2, Fig. 4A). *Vibrio* spp. correlated positively with salinity, Chl *a*, PO_4^{3-}, NO_3^-, and NH_4^+ but negatively with oxygen (Table 2, Fig. 4A). *V. vulnificus* was positively correlated with PO_4^{3-} and NO_3^- (only at the German station), and restricted to a salinity of 8.1-11.2 (Table 2, Fig. 4A).

The environmental and biological factors that correlated with the gene copy numbers of *Vibrio* spp. and *V. vulnificus* were used to identify their key predictors. Our stepwise multiple regression model showed that SST, PO_4^{3-}, *Vibrio* spp. gene copy number from sediment, and SSWWSH were key predictors of *Vibrio* spp. (Fig. 4B; $r^2 = 0.33$, $p < 0.001$). For *V. vulnificus*, SST, PO_4^{3-}, and turbidity were the main predictors (Fig. 4C; $r^2 = 0.23$, $p < 0.001$). The fact that *Vibrio* spp. gene copy number from sediment and turbidity appeared as key drivers of *Vibrio* spp. and *V. vulnificus* abundance, respectively, suggest a potential connection with sediment resuspension.


A
Gene copies Vibrio spp. ml-1
SST (°C)/Salinity
SST
Salinity
9/5/22
23/5/22
10/6/22
24/6/22
11/7/22
20/7/22
3/8/22
10/8/22
7/9/22
16/9/22
5/10/22
14/10/22
May
Jun
Jul
Aug
Sep
Oct
Denmark
B
Gene copies Vibrio spp. ml-1
Gene copies V. vulnificus ml-1/
SST (°C)/Salinity
V. vulnificus
4/5/22
10/6/22
12/7/22
26/7/22
10/8/22
17/8/22
23/9/22
14/10/22
Germany
C
18/5/22
14/6/22
13/7/22
9/8/22
15/9/22
12/10/22
Finland
D
Denmark
Germany
Finland
Gene copies Vibrio spp. ml-1/g-1
Sediment
water
Substation
E
Gene copies V. vulnificus spp. ml-1/g-1

Previous page: Figure 3. Dynamics of* Vibrio *spp. abundance in the water column at the three stations (A) Denmark, (B) Germany, and (C) Finland from May to October 2022. *Abundances were proxied by ddPCR gene Continued Figure 3: quantification, and* V. vulnificus *was only above the detection limit at the German station. Bars indicate* Vibrio *spp. abundance, the red line indicates* V. vulnificus *abundance, the orange lines indicate SST [yellow fill in (B) indicates the temperatures at which* V. vulnificus *was detected], and the blue lines indicate salinity. Values are average +- standard error (n = 9, i.e., combining the triplicates by substation). Comparison of gene copies of* Vibrio *spp. (D) and* V. vulnificus *(E) in water and sediment samples obtained from substations; within ("A"), edge of ("B", 15 m), and far from ("C", 100 m) the eelgrass. Data from all the time points at the specific substation in Denmark (n = 36), Germany (n = 24), and Finland (n = 18) were combined. The values are the median, and vertical black lines indicate the position of the lower and upper quartiles. In (E), the red line indicates gene copies are under the detection limit. In (A), (B), and (C), letters show pairwise analysis among the variables (i.e., month) at each station, and in (D) and (E), asterisks indicate pairwise significant differences between substations, i.e., "A", "B", and "C", or type of sample, i.e., water and sediment (*p < 0.05, ** p < 0.01, ns [not significant] = p > 0.05), using a posthoc test (Wilcoxon) after a Kruskal-Wallis test.*

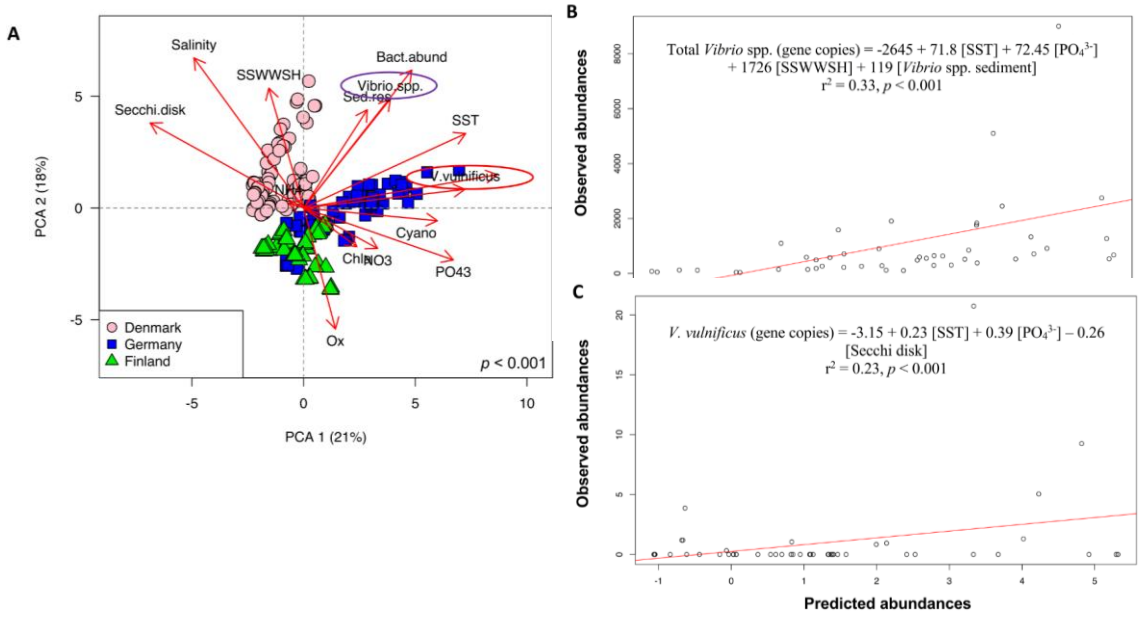


Figure 4. Environmental and biological drivers of* Vibrio *spp. and* V. vulnificus *occurrence in the water column. *A) PCA of the environmental and biological parameters measured from May to October (2022) in the Danish, Finnish, and German stations, including salinity, SST (°C), oxygen (Ox, mg l^{-1}), chlorophyll a (Chl a, mg m^{-3}), PO_4^{3-} (µM), NH_4^+ (µM), NO_3^- (µM), bacterial abundance (Bact. abund, cell ml^{-1}), Secchi depth (m), SSWWSH (m), sediment resuspension (sed.res, ASV from sediment detected in the water column), and relative abundance of potential harmful cyanobacteria, i.e.,* Cyanobium*,* Dolichospermun*, and* Nodularia *(Cyano, %). Circles highlight* Vibrio *spp. and* V. vulnificus*. Observed vs. predicted gene copies of (B)* Vibrio *spp. and (C)* V. vulnificus *from May to October (2022). Predictions were generated by stepwise multiple regression models. In (A), (B), (C) all water samples were included, n = 234*

The impact of sediment resuspension on *Vibrio* spp. and *V. vulnificus* in the water column

To identify a potential link between sediment resuspension and elevated prevalence of *Vibrio* spp. in the water column, using DESeq2 (24), we identified ASVs that were characteristic of sediments and used their abundance (relative abundance multiplied by bacterial abundance) in the water column as an indication of sediment resuspension (Fig. S4A). These ASVs, which accounted for 3-20% of sediment reads (< 0.1% in the water column) and remained stable during the season, belonged mainly to the phylum of *Actinobacteria*, *Chloroflexota, Desulfobacterota*, *Eisenbacteria*, *Gemmatimonadetes*, or *Myxococcota* (Fig. S4B, C). These phyla preferably inhabit the sediment (25–27). In Germany and Denmark, the *Vibrio* spp. gene copy number in water was positively correlated with the abundance of sediment bacteria, i.e., sediment resuspension, and negatively with Secchi depth (Fig. 5A, B). Interestingly, the presence of *V. vulnificus* in the water column in Germany coincided with elevated sediment resuspension (Fig. 5B, C). Sediment resuspension was particularly high in July and September (Fig. 5C), and correlated with higher PO_4^{3-} in the water column (Table S3). In Finland, there was no correlation between sediment resuspension and *Vibrio* spp., but fewer sediment bacteria were found in the water column compared to Denmark and Germany.

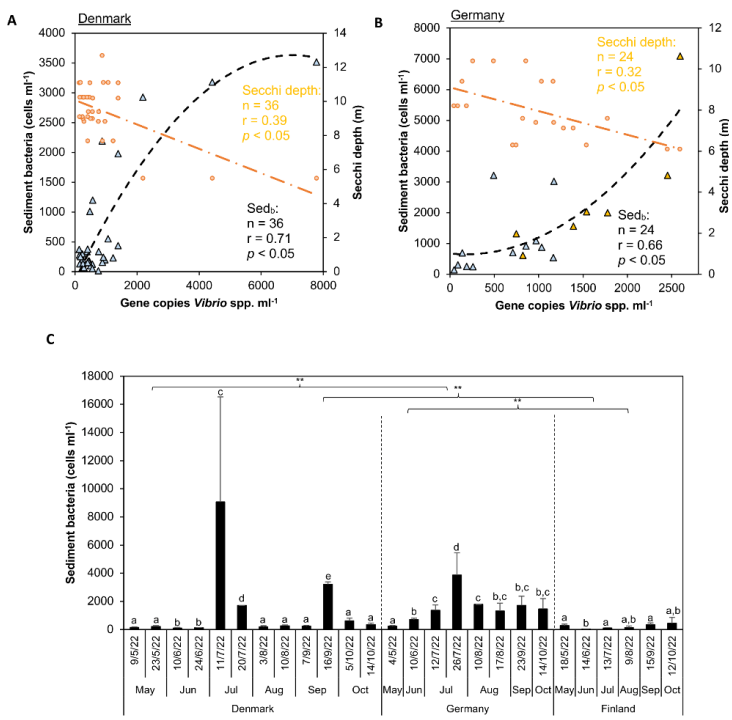


Figure 5. Regression between sediment resuspension – proxied by sediment ASVs in the water column [their relative abundances multiplied by bacterial abundance (cells ml^{-1})] – and the abundance of* Vibrio *spp. in the water column *(triangles and black curves) ($p < 0.05$). (A,B), a regression between the abundance of* Vibrio *spp. and Secchi depth (m), respectively, is inserted (dots and orange lines). (B)* V. vulnificus *is indicated by orange triangles. Note that in (A) and (B), ddPCR values from the substations (A, B, and C) were averaged to enable comparisons with the pooled water samples. C) Seasonality of sediment ASVs in the water column.*

Co-occurrence of specific microbial communities with *Vibrio vulnificus*

As in our sediment resuspension analysis, we used DESeq2 to pinpoint the ASVs that displayed significance and dominance in the presence of *V. vulnificus*. At the German station, 900 ASVs in the water column significantly changed their relative abundance when *V. vulnificus* was present (Fig. S5A; top 30 most significant taxa $p < 0.01$). These belonged to *Actinobacteriota*, *Bacteriota*, *Campylobacterota,* *Firmicutes*, *Planctomycetota*, *Proteobacteria*, SAR324, and *Verrucomicrobiota* (Fig. 6). These phyla are known to exhibit a positive correlation with cyanobacterial and algae blooms (28). Indeed, ASVs within *Cyanobacteria* constituted up to 20% of the microbial community (Fig. 6, $p < 0.05$), correlating with periods of higher *Vibrio* spp. and *V. vulnificus* gene copy numbers (Fig. 5A, 6, and Table 2). Hence, *Cyanobacteria* was one of the most abundant groups in the presence of *V. vulnificus* (Fig. 6; $p < 0.01$). The unicellular cyanobacteria *Cyanobium* spp., *Atelocyanobacterium thalassa* A (UCYN-A) and *Volcanococcus*, and the filamentous cyanobacteria *Dolichospermum* and *Nodularia* dominated the cyanobacterial community when *V. vulnificus* was present in the water (Fig. S5B). Note, that the changes in cyanobacteria community structure were related to the temporal changes (Fig. S6B).

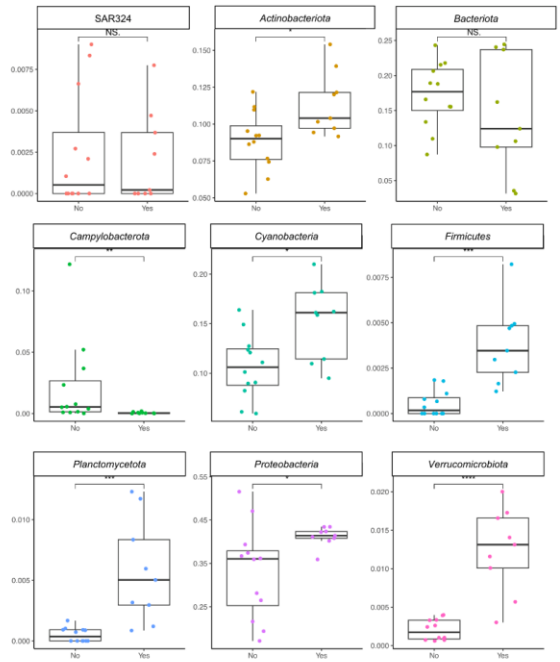


***Figure 6. Significant changes in relative abundance of bacterial ASVs in the water column in the presence or absence of* V. vulnificus.** *Approximately 900 ASVs showed notable shifts in their relative abundances when* V. vulnificus *was present at the German station. Phyla are shown. The values are the median and vertical black lines indicate the position of the lower and upper quartiles Asterisks indicate pairwise significant differences between the presence ("yes") or absence ("no") of* V. vulnificus *(*p < 0.05, ** p < 0.01, NS.] = p > 0.05), using a posthoc test (Wilcoxon) after a Kruskal-Wallis test.*

Discussion

The proliferation of *Vibrio* spp. in the Baltic Sea is a critical clinical challenge where the increased frequency of *Vibrio* infections appears to be closely linked to regional climatic trends and episodes of unusually warm weather (4, 29). Within the stations sampled, we found that *V. vulnificus* was almost exclusively confined to medium salinity (8.1 – 11.2), and episodes with elevated SST, high phosphate (PO_4^{3-}) concentration, and sediment resuspension served as indicators for *V. vulnificus* in the area. Our study did not show a reduction in *Vibrio* spp. associated with eelgrass beds, as suggested in an earlier Baltic Sea study (12), but rather indicates that vibriosis may be linked to sediment resuspension and its associated effects, e.g. elevated phosphate.

Salinity is an important driver of *Vibrio* spp. communities

Low salinity caused a decrease in *Vibrio* spp. diversity but promoted specific *Vibrio* taxa. For instance, the fish pathogen *V. anguillarum* – known from aquaculture and larviculture industries in brackish waters (30)– dominated the *Vibrio* population in Finland and Germany (salinity 6.3-11.2). In contrast at the more saline Danish station (salinity 13-21.3), the *Vibrio* spp. diversity was elevated and the mussel pathogens *V. cyclitrophicus* and *V. chagasii* appeared during the summer months. In comparison, *V. vulnificus* was only present at the German station (salinity 8.1-11.2). Therefore, as previously suggested (7, 15, 31), salinity appears an important selective driver for *Vibrio* spp., but the proliferation of particular pathogens, e.g., *V. vulnificus*, is governed by the interplay of a range of environmental and biological factors (see below).

The importance of temperature for *Vibrio vulnificus* in the Baltic Sea

The analysis of the *Vibrio* community showed a strong seasonality and sea surface temperature (SST) was a key predictor for *Vibrio* spp. and *V. vulnificus* in the water column. This is consistent with previous observations of *Vibrio* spp. peaking during summer in other regions, such as the Barnegat Bay, Chesapeake Bay, Monterrey Bay, Gulf of Mexico, North Sea, Atlantic Coast, or Sydney Harbour Estuary (15, 16, 19, 32–35*).* In contrast to *Vibrio* spp. that were found even below 10 °C, *V. vulnificus* occurred in the water column only when temperatures were above 19 °C, and peaked at 22 °C. Interestingly, this aligns with a predicted increased probability of *V. vulnificus* in the 15–20 °C range and a large risk of proliferation above 20 °C (19). Hence, future longer heatwaves in the Baltic Sea region will likely increase

the frequency of *V. vulnificus* during the summer season, as already reported in some recent studies (20–22), making *Vibrio* infections a putative increasing public health concern as this is the peak season for tourists and swimmers. However, unlike these earlier studies (20–22), *V. vulnificus* was almost exclusively found at the German station, at a medium salinity of 8.1–11.2. This observation carries significant implications as *V. vulnificus* is known to exist naturally across a broad spectrum of salinity levels, ranging from 5 to 25 (15, 36), with its ideal habitat falling within the 10 to 15 range (16). Hence, other environmental drivers than temperature and salinity regulate the prevalence of *V. vulnificus*.

Role of phosphate for *Vibrio vulnificus* in the Baltic Sea

PO_4^{3-} was also a primary predictor for forecasting *V. vulnificus*. Consequently, regions such as the German station situated in the Mecklenburg Bight at the mouth of the Warnow River, known for substantial nutrient influx due to agricultural and industrial runoff (37, 38), are likely particularly prone to *V. vulnificus* outbreaks. These results deviate from previous findings suggesting that PO_4^{3-} has a minimal explanatory effect on *Vibrio* spp. variance and exhibits no correlation with *V. vulnificus* (7), despite earlier description of PO_4^{3-} being an essential nutrient for *Vibrio* growth and pathogenesis (39, 40). However, this disparity could stem from the fact that most of these studies are conducted in Atlantic Ocean estuaries, where nutrient conditions may be different (31, 41, 42). Finally, we speculate that the association of *V. vulnificus* with high PO_4^{3-} levels is indirect and *de facto* driven by sediment resuspension (discussed below).

Interestingly, we found a correlation between potentially harmful cyanobacteria and elevated levels of PO_3^{4-} at the German station (Table S3), coinciding with a substantial increase in the gene copies of *V. vulnificus*. This matched with our recent spatial study in the Baltic Sea (43). When *V. vulnificus* was present, the unicellular picocyanobacterium *Cyanobium* spp., commonly found in association with heterocystous bloom-forming cyanobacteria (44), and the filamentous and heterocystous cyanobacteria *Dolichospermum* and *Nodularia*, prevailed. They are key taxa in the harmful cyanobacterial blooms occurring every summer in the Baltic Sea Proper (45), and may release dissolved organic matter compounds specifically supporting pathogenic *Vibrio* spp. (46), as well as providing surfaces for attachment, enhancing *V. vulnificus* survival (47). Collectively, the observed correlations between nutrient levels, cyanobacteria, and *V. vulnificus* emphasize the significance of PO_4^{3-} concentration and rising

temperatures as pivotal factors influencing the prevalence and survival of *V. vulnificus* in the Baltic Sea.

Sediment as a source of *Vibrio* spp. in the water column

Our model showed that turbidity was one of the main drivers for *V. vulnificus* in the water column. Although we found specific cyanobacteria taxa co-occurring with *V. vulnificus*, our data showed a negative correlation between Chl *a* and gene copies of *V. vulnificus*, suggesting that turbidity could be mainly attributed to sediment resuspension, rather than to phytoplankton biomass. This is supported by the positive correlation between sediment bacteria co-existing with *Vibrio* spp. and *V. vulnificus* in the water column, and the fact that some water samples contained *Vibrio* communities similar to those found in the sediment, suggesting a benthic-pelagic coupling. The detection of these sediment bacteria by sequencing in the water column can be a more useful proxy for sediment resuspension than turbidity, e.g., estimated by Secchi depth, which can be highly influenced by organic matter load or phytoplankton blooms (48). Our findings are consistent with other marine studies connecting *Vibrio* spp. and sediment resuspension, whether due to human activity, e.g., ship traffic, or natural phenomena like rainfalls and storms (31, 49, 50).

We speculate that sediment resuspension can increase *V. vulnificus* levels in the water column through two related mechanisms. First, by introducing *V. vulnificus* found in association with particulate matter originating from the sediment. Sediments may harbor pathogens (51, 52), and sediments rich in silt and organic matter – as those found in Germany – may harbor a particularly high density of bacterial pathogens (51), which may be released to the water column (31). Indeed, gene copies of *Vibrio* spp. and *V. vulnificus* were 100 and 1000 times higher in the sediment than in the water, respectively, indicating that *Vibrio* populations persist and thrive in sediment (16). Our data show that *V. vulnificus* remained quantifiable in sediments below 15 °C, indicating that sediment can be a *V. vulnificus* reservoir even when the water is less warm. At these temperatures, *Vibrio* spp., including *V. vulnificus*, can enter a viable but non-culturable (VBNC) latency state, enabling survival even at low temperatures (53). However, the proportion of bacteria capable of entering this state cannot be accurately assessed using the methods employed in this study. Second, by supplying PO_4^{3-} but also organic matter to the water column, thereby promoting the growth of *V. vulnificus*. We showed that sediment resuspension events correlated with higher PO_4^{3-} into the water column. It is known that sediments can serve as a source of PO_4^{3-}, and their release may depend on their

composition, which differed between stations (Table S3), and oxygen conditions (54). The accumulation of PO_4^{3-} in sediment can be more easily released into the water column during summer, as depicted in Fig. 1E. This phenomenon can be attributed to the heightened temperatures, which augment solubility, sediment adsorption capacity, and organic matter dynamics (55), leading to changes in nutrient availability and ecological processes indirectly impacting the prevalence of this pathogenic bacterium (56). This scenario may be occurring at the German station, where the sampling location and sediment composition could be influencing these effects. From our study, sediment resuspension appears a key driver of pathogenic *Vibrio* spp. in the water column; however, it is not possible to disentangle the relative importance of the two scenarios, i.e. distinguish effects of cell, organic matter, and PO_4^{3-} transfer associated with resuspension.

Conclusions

Our study identified SST, PO_4^{3-}, and sediment resuspension as key factors controlling the occurrence of pathogenic *Vibrio* spp., such as *V. vulnificus*, in the Baltic Sea. Also, *V. vulnificus* was only present at intermediate salinity. Climate change will cause an increased frequency of natural disasters and rain-driven floods in the Baltic Sea area with associated increased sediment resuspension and elevated nutrient concentrations (57), likely leading to raised levels of vibriosis, as was observed in association with hurricanes in the United States (58). *While it conceivable that* seagrass meadows may indirectly prevent *Vibrio* spp. proliferation via effects on sedimentation, sediment surface stabilization, and mitigation of turbulent flow (12), we did not find a reduction in *Vibrio* spp. associated with eelgrass beds. Our temporal study contradicts the assumption that seagrass reduce pathogenic *Vibrio* spp. levels in the areas studied. This is supported by our earlier spatial study, in which, as in the current study, we employed culture-independent methods, and found that *V. vulnificus* abundance did not vary significantly between vegetated and non-vegetated areas *(43)*. Therefore, it could not be shown that seagrass meadows are a direct effective nature-based solution for reducing *V. vulnificus* abundance and associated infections in the Baltic Sea, and alternative solutions should be explored.

Materials and methods

Sampling

Three coastal stations along the southwest-northeast salinity gradient of the Baltic Sea were sampled from May to October 2022 (Table 1, Fig. 1A): Julebæk, in the Øresund Strait (Denmark) – located in the Hovedstaden region, an area impacted by extensive shipping and high winds and currents; Warnemünde, at the mouth of the Warnow River (Germany) – a touristic area in the Mecklenburg-Vorpommern region; Ängsö Bay, situated in the Archipelago Sea (Southwest Finland), located to the west of Ängsö Island, in an area characterized by a low population density (Table 1). Samplings were carried out once or twice per month by snorkeling or diving (SCUBA) with local boats at three sampling points per station, i.e., substation: A - in the middle, B - 15 m, and C - 100 m outside of the *Z. marina* meadow. Based on data collected in summer 2021 (43), the average abundance and length of the seagrass beds sampled in Denmark, Germany, and Finland were: 675 shoots m^{-2} and 74.26 cm, 400 shoots m^{-2} and 75.55 cm, and 225 shoots m^{-2} and 22.46 cm, respectively. When sampling, food colorant was employed to monitor the current, and divers navigated towards the sampling site against the current, while carefully controlling buoyancy to avoid sediment resuspension. First, water samples were collected by gentle suction of 1.5 L water using three acid washed 500 ml syringes per replicate, after triplicate rinsing with sample water. The water was used for the downstream analysis, i.e., DNA, chlorophylls, nutrient, and flow cytometry. Water samples were collected either 5 cm from the eelgrass leaves (to closely monitor the maximum effect of the eelgrass) at the substation A and 20 cm above the sediment level to avoid any kind of sediment resuspension at the substations B and C. Second, eelgrass leaves (only in A) and the uppermost sediment layer were collected in triplicate 50 ml Falcon tubes. All samples were stored in a cooler and processed within 2 h. Sea surface temperature (SST), salinity, and oxygen were measured at each sampling. Secchi depth and sea surface wind wave significant height (SSWWSH) values were obtained from Copernicus EMS using the three station coordinates and the date and time at which sampling was conducted (Table S1).

Sample processing

Water samples (500 ml) for DNA and chlorophyll *a* (Chl *a*) analyses were immediately filtered onto a Durapore® membrane filter, 0.22 µm (Sigma-Aldrich, MA, USA) and GF/F filter (Whatman®, Sigma-Aldrich, MA, USA), respectively, and stored at -20 °C. Filtered water was stored frozen for analysis of phosphate (PO_4^{3-}), nitrate (NO_3^-), and

ammonium (NH_4^+). For bacterial enumeration, 2 ml of water were fixed with formaldehyde 37% (v/v), incubated for 1 h at 8 °C, shock frozen, and stored at -80 °C. Sediment and eelgrass samples were immediately frozen at -20 °C.

Chlorophyll a, inorganic nutrients, bacterial abundance, and grain size

Chl *a* was extracted with 96% (v/v) ethanol and measured on a Trilogy Laboratory Fluorometer (Turner Designs, San Jose, CA, USA) calibrated with a Chl *a* standard (DHI, Denmark) following (59). NH_4^+ was quantified fluorometrically, according to (60), while PO_4^{-3} and NO_3^- were quantified using standard colorimetric methods (61 - 62). Bacteria were enumerated on a FACSCanto II flow cytometer (BD, New Jersey, USA) according to (63) using TrueCount beads (BD) to measure flow rate. Grain size from 27 sediment samples between July and September (n = 9 per station) was measured on the Mastersizer 3000 (Malvern Panalytical, Malvern, UK).

Molecular analyses

DNA extraction.

Durapore® membrane filters were thawed and ground in 2 ml Eppendorf tubes with liquid nitrogen. The sediment or eelgrass samples were washed before extraction with 1.5 ml of phosphate buffer saline (PBS) in a 2-ml tube containing approximately 1 g of sediment or 0.3 g of eelgrass to minimize contamination from the water column. DNA was extracted using the DNeasy PowerSoil Pro Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. Samples were eluted in 10 mM Tris buffer (pH 8) and DNA content was quantified using the PicoGreen dsDNA Assay Kit (ThermoFisher, MA, USA).

16S rDNA amplicon sequencing analysis and Vibrio *spp. annotation.*

DNA from triplicate water samples from each substation were combined at every sampling point, i.e., uniformly mixed, resulting in one water sample sequenced per substation. Sediment and eelgrass samples were sequenced in triplicates. Thus, DNA from 78 water, 204 sediment, and 78 eelgrass samples was sequenced at the SciLifeLab National Genomics Infrastructure (NGI, Solna, Sweden). Libraries were prepared using the primers 341F/805R (341F, 5'-CCTACGGGNGGCWGCAG-3'; 805R, 5'-GACTACHVGGGTATCTAATCC-3') targeting the hypervariable V3-V4 region of the 16S rRNA gene (64). Library preparation was

performed according to NGI protocols (https://ngisweden.scilifelab.se/methods/illumina-16s-sequencing-2/ and https://ngisweden.scilifelab.se/2021/01/tech-note-increase-complexity-of-amplicon-libraries-using-phased-primers/), and sequenced on a NextSeq2000 (NextSeq 1000/2000 Control Software 1.5.0.42699/RTA 3.10.30) with a 301nt(Read1)-10nt(Index1)-10nt(Index2)-301nt(Read2) setup using a flowcell. After sequencing, the library was demultiplexed, and phased primers, Illumina adaptors, and 16S rRNA gene primers were removed (https://github.com/biodiversitydata-se/amplicon-multi-cutadapt). Sequence analysis was performed in R Studio v 4.1.0. (http://www.r-project.org). Three of the 78 water samples did not pass the quality control. Amplicon sequence variants (ASVs) were generated with the DADA2 pipeline v.1.21.0 (59), and annotated using a curated version of 16S rRNA genes from GTDB (v.R06-RS202-1) (60). Sequences were deposited in NCBI (PRJNA1011541).

The taxonomic annotation of the 16S ASVs was done in two steps: First, we used the DADA2 function assignTaxonomy (65), and 16S sequences from GTDB to get taxonomic annotations of the ASVs to genus level (66). Second, for the ASVs annotated as *Vibrio* in the initial annotation, species-level classification was achieved by sequence comparison to a custom database (BLAST, V. 2.13.0) (67). This database includes 16S rRNA gene sequences of complete *Vibrio* genomes from RefSeq (51 species, 317 strains - including 22 *V. vulnificus* strains) (68), 41 draft *V. vulnificus* genomes from clinical isolates from the Baltic region (4, 6), and 84 draft *V. vulnificus* genomes from environmental Baltic Sea isolates (Delgado et al., unpublished). For an ASV to obtain a species-level assignment, we required perfect (100% identity) alignment of the full ASV to 16S rRNA gene(s) of a single species in the custom database. In some cases, this resulted in ambiguous species annotation. However, when considering matches for *V. vulnificus*, the associations were unambiguous, meaning that the identified ASVs perfectly aligned with 16S sequences belonging exclusively to this species.

Exact match is recommended for species-level assignments of 16S ASVs (68).The scripts used for the bioinformatics processing is available at https://github.com/lfdelzam/ASV_dada2_chunck/ and the 16S rRNA *Vibrio* database for this paper is available at https://zenodo.org/records/1087.

Quantification of Vibrio *spp. and* V. vulnificus *with digital droplet PCR (ddPCR).*

The abundance of *Vibrio* spp. and *V. vulnificus* in water and sediments was assessed by gene quantification using ddPCR. We targeted the 16S rRNA gene with 567F/680R (567F, 5'-

GGCGTAAAGCGCATGCAGGT-3'; 680R, 5'-GAAATTCTACCCCCCTCTACAG-3, (32)) for total *Vibrio* spp., and *vvhA* gene with vvh-785f/990r (vvh-785f:5'TTCCAACTTCAAACCGAACTATGAC-3', vvh-990r: 5'-ATTCCAGTCGATGCGAATACGTTG-3,(70) for *V. vulnificus*. For the ddPCR reaction, 11 µl QX200™ ddPCR™ EvaGreen® Supermix (Bio-Rad, München, Germany), 0.01 µM forward and reverse primer (final concentration), 5-15 ng DNA template, and DEPC H_2O were added in a final volume of 22 µL. Droplets were generated using a QX100 droplet generator (Bio-Rad, München, Germany) according to the manufacturer's instructions. The emulsified samples were transferred to a 96-well plate and sealed with a pierceable heat-sealing film (Bio-Rad, München, Germany). PCR was performed using a Bio-Rad C1000 Touch™ thermal cycler (Bio-Rad, München, Germany). The PCR conditions for *Vibrio* spp. were 94 °C for 5 min, 40 cycles of 94 °C for 1 min, 61.6 °C for 1 min, and 72 °C for 1 min. For *V. vulnificus*, the program was 95 °C for 5 min, followed by 40 cycles at 95 °C for 30 seconds, 58.5 °C for 1 min, and 72 °C for 1 min. In both cases, a final step of 4 °C for 5 min, 90 °C for 5 min, and an infinite hold at 4 °C was performed. The plate was then analyzed with the QX200™ Droplet Digital PCR system (Bio-Rad) using Quantasoft 1.74.09.17 software (Bio-Rad, München, Germany). Positive controls of *V. vulnificus* isolates and blanks with DEPC water were used. Based on the above described 16S rRNA gene sequencing, we applied a correction factor for *Vibrio* spp. counts, i.e., the ratio of *Vibrio* spp. and *Photobacterium* spp. 16S rRNA gene sequences, to account for the cross-reaction of the primer set 567F/680R with *Photobacterium* spp., *sensu* (71).

Community analysis, statistics, and modelling

All analyses were carried out using R Studio v.4.1.0.

Community analysis.

Sample data matrices were managed using the phyloseq package v.1.36 (72). Community analyses were exclusively carried out on ASVs within the *Vibrio* genus unless stated otherwise. Alpha diversity (Shannon index) was calculated after rarefaction. Principal component (PCA) and redundancy analyses (RDA) were done with the microViz package v.0.10.8 to measure similarity or dissimilarity between *Vibrio* communities using central log ratio (clr) transformed data.

To evaluate changes in the *Vibrio* community, we conducted a permutational multivariate analysis of variance using distance matrices (PERMANOVA, number of permutations = 999) using the Euclidean distance matrix. We also used an ANOVA-like permutation test for constrained correspondence analysis to determine the significance of the constraints (number of permutations = 999) using the anova.cca which is part of vegan v.2.6.4 (73). To evaluate pairwise differences between groups, pairwise analyses were done with PairwiseAdonis v.0.4 (74). Additionally, we performed a SIMPER analysis (number of permutations = 999) to identify the taxa responsible for differences between groups of samples, using vegan v.2.6.4.

To identify ASVs that were differentially abundant in the sediment, a negative binomial Wald test was conducted using DESeq2 v.1.32.0 on the non-rarified and non-clr transformed data (24). The sediment samples (n = 204) were compared to the respective water samples (n = 75). Despite unequal sample sizes, the data exhibited a negative binomial distribution, ensuring a robust analysis. Significance was set at $p < 0.01$. The top 26 most significant ASVs differentially abundant in the sediment at each station were used as indicators for sediment resuspension when found in the water column. Additionally, the same analysis was performed to determine the ASVs associated with the presence (n = 9) and absence (n = 12) of *V. vulnificus* (based on the ddPCR values). DESeq2 analysis was carried out with the top 10,000 most abundant ASVs (comprising almost 100% of the entire microbial community–). The ASVs detected were visualized using clustering heatmaps created with the complexHeatmap package v.2.8.0 (75).

Statistical analyses.

Homogeneity of variances for the ddPCR and relative abundance data followed a non-normal distribution (Leven's test). Therefore, non-parametric analyses were employed using Kruskal–Wallis. Pairwise comparisons were done using Wilcoxon Rank Sum Tests. Spearman's rank analysis was employed to determine relationships between the ddPCR data of *Vibrio* spp. and *V. vulnificus* in the water column and environmental and biological data (n = 234), using Bonferroni correction to adjust Spearman Rank correlation p-values. Unless otherwise stated, the significance level for all analyses, including correlations, was set to $p < 0.05$.

Prediction of Vibrio *spp. and* V. vulnificus.

To predict *Vibrio* spp. and *V. vulnificus* occurrence, we used stepwise multiple regressions, following the methodology outlined in (33, 76). The data used for training and testing the model were split, with 80% used for training and 20% for testing. The analysis was performed using the olsrr package v0.5.3 (https://CRAN.R-project.org/package=olsrr), and the predictive model was developed based on the ddPCR data obtained from the water column (n = 234). We tested that the data met the assumptions of linearity, independence, normality of residuals, and homoscedasticity. Variables exhibiting high co-linearity were eliminated. Only variables with $p < 0.05$ were included in the model. Studentized residual analyses were conducted to thoroughly examine the variance between observed and predicted abundances. The visual examination of the difference between predicted and observed abundances was calculated.

Acknowledgements

The authors acknowledge support from the National Genomics Infrastructure (NGI) in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, for assistance with massively parallel sequencing and SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science for access to the UPPMAX computational infrastructure. We would like to thank two anonymous reviewers for their thoughtful and constructive advice that improved the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author contributions

The sampling in Denmark was conducted by VF-J, GC-C, and VP, while DR and ML carried out the sampling in Germany, and JG and CP carried out the sampling in Finland. VF-J conducted all molecular analyses with assistance from ML, DR, GC-C and DH. FD conducted the bioinformatics sequence processing. VF-J performed statistical analyses with support from DR. VF-J and LR wrote the manuscript, which was edited by all authors.

Funding

This work resulted from the BiodivERsA project “Pathogenic *Vibrio* bacteria in the current and future Baltic Sea waters: mitigating the problem” (BaltVib), funded by the European Union and the Federal Ministry of Education and Research, Germany (grant 16LC2022A), the Innovation Fund Denmark (grant 0156-00001B), the Estonian Research Council (grant T210076PKKH / P200028PKKH), the Research Council of Lithuania (grant S-BIODIVERSA-21-1), the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (grant 2020-02366), the Polish National Science Centre and the Academy of Finland (grant 344743).

References

1. Baker-Austin C, Oliver JD, Alam M, Ali A, Waldor MK, Qadri F, Martinez-Urtaza J. 2018. *Vibrio* spp. infections. *Nature Reviews Disease Primers* 4.

2. Baker-Austin C, Trinanes JA, Salmenlinna S, Löfdahl M, Siitonen A, Taylor NGH, Martinez-Urtaza J. 2016. Heat wave-associated vibriosis, Sweden and Finland, 2014. *Emerging Infectious Diseases* 22:1216–1220.

3. Baker-Austin C, Trinanes JA, Taylor NGH, Hartnell R, Siitonen A, Martinez-Urtaza J. 2013. Emerging *Vibrio* risk at high latitudes in response to ocean warming. *Nature Climate Change* 3:73–77.

4. Amato E, Riess M, Thomas-Lopez D, Linkevicius M, Pitkänen T, Wołkowicz T, Rjabinina J, Jernberg C, Hjertqvist M, Macdonald E, Karloss J, Samy A-, Dalsgaard Bjerre K, Salmenlinna S, Fuursted K, Hansen A, Naseer U. 2022. Epidemiological and microbiological investigation of a large increase in vibriosis, northern Europe, 2018. *Euro Surveill* https://doi.org/10.1101/2021.11.19.21266449.

5. Gildas Hounmanou YM, Engberg J, Bjerre KD, Holt HM, Olesen B, Voldstedlund M, Dalsgaard A, Ethelberg S. 2023. Correlation of high seawater temperature with *Vibrio* and *Shewanella* infections, Denmark, 2010-2018. *Emerging infectious diseases* 29:605–608.

6. Brehm TT, Berneking L, Sena Martins M, Dupke S, Jacob D, Drechsel O, Bohnert J, Becker K, Kramer A, Christner M, Aepfelbacher M, Schmiedel S, Rohde H. 2021. Heatwave-associated *Vibrio* infections in Germany, 2018 and 2019. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* 26.

7. Takemura AF, Chien DM, Polz MF. 2014. Associations and dynamics of vibrionaceae in the environment, from the genus to the population level. *Frontiers in Microbiology* 5.

8. Chase E, Young S, Harwood VJ. 2015. Sediment and vegetation as reservoirs of *Vibrio vulnificus* in the Tampa Bay Estuary and Gulf of Mexico. *Applied and Environmental Microbiology* 81:2489–2494.

9. Matz C, McDougald D, Maria Moreno A, Yi Yung P, Yildiz FH, Kjelleberg S. 2005. Biofilm formation and phenotypic variation enhance predation-driven persistence of *Vibrio cholerae*.

10. Andrews JH, Harris RF. 2000. The ecology and biogeography of microorganisms on plant surfaces. *Annu. Rev. Phytopathol.*

11. Lamb JB, Van De Water JAJM, Bourne DG, Altier C, Hein MY, Fiorenza EA, Abu N, Jompa J, Harvell CD. 2018. Seagrass ecosystems reduce exposure to bacterial pathogens of humans, fishes, and invertebrates.

12. Reusch TBH, Schubert PR, Marten SM, Gill D, Karez R, Busch K, Hentschel U. 2021. Lower *Vibrio* spp. abundances in Zostera marina leaf canopies suggest a novel ecosystem function for temperate seagrass beds. *Marine Biology* 168.

13. Zimmerman AM, DePaola A, Bowers JC, Krantz JA, Nordstrom JL, Johnson CN, Grimes DJ. 2007. Variability of total and pathogenic Vibrio parahaemolyticus densities in Northern Gulf of Mexico water and oysters. Applied and Environmental Microbiology 73:7589–7596.

14. Martinez-Urtaza J, Blanco-Abad V, Rodriguez-Castro A, Ansede-Bermejo J, Miranda A, Rodriguez-Alvarez MX. 2012. Ecological determinants of the occurrence and dynamics of *Vibrio parahaemolyticus* in offshore areas. *ISME Journal* 6:994–1006.

15. Randa MA, Polz MF, Lim E. 2004. Effects of temperature and salinity on *Vibrio vulnificus* population dynamics as assessed by quantitative PCR. *Applied and Environmental Microbiology* 70:5469–5476.

16. Brumfield KD, Chen AJ, Gangwar M, Usmani M, Hasan NA, Jutla AS, Huq A, Colwell RR. 2023. Environmental Factors Influencing Occurrence of *Vibrio parahaemolyticus* and *Vibrio vulnificus. Applied and Environmental Microbiology* https://doi.org/10.1128/aem.00307-23.

17. Kimes NE, Grim CJ, Johnson WR, Hasan NA, Tall BD, Kothary MH, Kiss H, Munk AC, Tapia R, Green L, Detter C, Bruce DC, Brettin TS, Colwell RR, Morris PJ. 2012. Temperature regulation of virulence factors in the pathogen *Vibrio coralliilyticus*. *ISME Journal* 6:835–846.

18. Reusch TBH, Dierking J, Andersson HC, Bonsdorff E, Carstensen J, Casini M, Czajkowski M, Hasler B, Hinsby K, Hyytiäinen K, Johannesson K, Jomaa S, Jormalainen V, Kuosa H, Kurland S, Laikre L, Mackenzie BR, Margonski P, Melzner F, Oesterwind D, Ojaveer H, Refsgaard JC, Sandström A, Schwarz G, Tonderski K, Winder M, Zandersen M. 2018. The Baltic Sea as a time machine for the future coastal ocean. *Science Advances* 1–16.

19. Böer SI, Heinemeyer EA, Luden K, Erler R, Gerdts G, Janssen F, Brennholt N. 2013. Temporal and spatial distribution patterns of potentially pathogenic *Vibrio* spp. at recreational beaches of the German North Sea. *Microbial Ecology* 65:1052–1067.

20. Eiler A, Johansson M, Bertilsson S. 2006. Environmental influences on *Vibrio* populations in northern temperate and boreal coastal waters (Baltic and Skagerrak Seas). *Applied and Environmental Microbiology* 72:6004–6011.

21. Fleischmann S, Herrig I, Wesp J, Stiedl J, Reifferscheid G, Strauch E, Alter T, Brennholt N. 2022. Prevalence and distribution of potentially human pathogenic *Vibrio* spp. on German North and Baltic Sea Coasts. *Frontiers in Cellular and Infection Microbiology* 12.

22. Gyraite G, Katarzyte M, Schernewski G. 2019. First findings of potentially human pathogenic bacteria *Vibrio* in the south-eastern Baltic Sea coastal and transitional bathing waters. *Marine Pollution Bulletin* 149.

23. Andersson AF, Riemann L, Bertilsson S. 2010. Pyrosequencing reveals contrasting seasonal dynamics of taxa within Baltic Sea bacterioplankton communities. *ISME Journal* 4:171–181.

24. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15.

25. Capo E, Cosio C, Gascón Díez E, Loizeau JL, Mendes E, Adatte T, Franzenburg S, Bravo AG. 2023. Anaerobic mercury methylators inhabit sinking particles of oxic water columns. *Water Research* 229.

26. Salah-Tantawy A, Chang CSG, Liu MY, Young SS. 2022. Exploring the diversity and structural response of sediment-associated microbiota communities to environmental pollution at the siangshan wetland in Taiwan using environmental DNA metagenomic approach. *Frontiers in Marine Science* 9.

27. Jäntti H, Jilbert T, Aalto SL, Simojoki A, Mangayil R, Peura S, Rissanen AJ. 2022. The role of organic matter and microbial community controlling nitrate reduction under elevated ferrous iron concentrations in boreal lake sediments. *Hydrobiologia* 849:2145–2160.

28. Pope PB, Patel BKC. 2008. Metagenomic analysis of a freshwater toxic cyanobacteria bloom. *FEMS Microbiology Ecology* 64:9–27.

29. Schütt EM, Hundsdörfer MAJ, von Hoyningen-Huene AJE, Lange X, Koschmider A, Oppelt N. 2023. First Steps towards a near Real-Time Modelling System of *Vibrio vulnificus* in the Baltic Sea. *International Journal of Environmental Research and Public Health* 20:5543.

30. Frans I, Michiels CW, Bossier P, Willems KA, Lievens B, Rediers H. 2011. *Vibrio anguillarum* as a fish pathogen: Virulence factors, diagnosis and prevention. Journal of Fish Diseases 34:643–661.

31. Morrison BH, Jones JL, Dzwonkowski B, Krause JW. 2024. Tracking *Vibrio* : population dynamics and ecology of *Vibrio parahaemolyticus* and *V. vulnificus* in an Alabama estuary. *Microbiology Spectrum* https://doi.org/10.1128/spectrum.03674-23.

32. Thompson JR, Randa MA, Marcelino LA, Tomita-Mitchell A, Lim E, Polz MF. 2004. Diversity and dynamics of a North Atlantic coastal *Vibrio* community. *Applied and Environmental Microbiology* 70:4103–4110.

33. Oberbeckmann S, Fuchs BM, Meiners M, Wichels A, Wiltshire KH, Gerdts G. 2012. Seasonal dynamics and modeling of a *Vibrio* Community in coastal waters of the North Sea. *Microbial Ecology* 63:543–551.

34. Vezzulli L, Brettar I, Pezzati E, Reid PC, Colwell RR, Höfle MG, Pruzzo C. 2012. Long-term effects of ocean warming on the prokaryotic community: Evidence from the vibrios. *ISME Journal* 6:21–30.

35. Siboni N, Balaraju V, Carney R, Labbate M, Seymour JR. 2016. Spatiotemporal dynamics of *Vibrio* spp. within the Sydney Harbour estuary. *Frontiers in Microbiology* 7.

36. Parvathi A, Kumar HS, Karunasagar I, Karunasagar I. 2004. Detection and enumeration of *Vibrio vulnificus* in oysters from two estuaries along the southwest coast of India, using molecular methods. *Applied and Environmental Microbiology* 70:6909–6913.

37. Murray CJ, Müller-Karulis B, Carstensen J, Conley DJ, Gustafsson BG, Andersen JH. 2019. Past, present and future eutrophication status of the Baltic Sea. *Frontiers in Marine Science* 6.

38. Pfeffer CS, Hite MF, Oliver JD. 2003. Ecology of *Vibrio vulnificus* in estuarine waters of eastern North Carolina. *Applied and Environmental Microbiology* 69:3526–3531.

39. Jahid IK, Silva AJ, Benitez JA. 2006. Polyphosphate stores enhance the ability of *Vibrio cholerae* to overcome environmental stresses in a low-phosphate environment. *Appl Environ Microbiol* 72:7043–7049.

40. Conrad JW, Harwood VJ. 2022. Sewage promotes *Vibrio vulnificus* growth and alters gene transcription in *Vibrio vulnificus CMCP6*. *Microbiol Spectr* 10:e01913-21.

41. Blackwell KD, Oliver JD. 2008. The ecology of *Vibrio vulnificus, Vibrio cholerae*, and *Vibrio parahaemolyticus* in North Carolina Estuaries. *Journal of Microbiology* 46:146–153.

42. Vezzulli L, Grande C, Reid PC, Hélaouët P, Edwards M, Höfle MG, Brettar I, Colwell RR, Pruzzo C. 2016. Climate influence on *Vibrio* and associated human diseases during the past half-century in the coastal North Atlantic. *PNAS* 113:E5062–E5071.

43. Riedinger D.J, Fernández-Juárez V, Delgado L.F, Sperlea T, Hassenrück C, Herlemann D.P.R, Pansch C, Kataržytė M, Bruck F, Ahrens A, Rakowski M, Piwosz K, Stevenson A, Reusch T.B.H, Gyraitė G, Schulz-Bull1 D, Benterbusch-Brockmöller H, Kube S, Dupke S,. Andersson A.F, Riemann L, Labrenz M. Control of *Vibrio vulnificus* proliferation in the Baltic Sea through eutrophication and algal bloom management. *Communication earth & environment.* 2024

44. Celepli N, Sundh J, Ekman M, Dupont CL, Yooseph S, Bergman B, Ininbergs K. 2017. Meta-omic analyses of Baltic Sea cyanobacteria: diversity, community structure and salt acclimation. *Env. Microbiology* 19:673–686.

45. Eigemann F, Schwartke M, Schulz-Vogt H. 2018. Niche separation of Baltic Sea cyanobacteria during bloom events by species interactions and autecological preferences. *Harmful Algae* 72:65–73.

46. Eiler A, Gonzalez-Rey C, Allen S, Bertilsson S. 2007. Growth response of *Vibrio cholerae* and other *Vibrio* spp. to cyanobacterial dissolved organic matter and temperature in brackish water. *FEMS Microbiology Ecology* 60:411–418.

47. Islam MS, Mahmuda S, Morshed MG, Bakht HBM, Khan MNH, Sack RB, Sack DA. 2004. Role of cyanobacteria in the persistence of *Vibrio cholerae* O139 in saline microcosms. *Canadian Journal of Microbiology* 50:127–131.

48. Lannergård EE, Ledesma JLJ, Fölster J, Futter MN. 2019. An evaluation of high frequency turbidity as a proxy for riverine total phosphorus concentrations. *Science of the Total Environment* 651:103–113.

49. Fries JS, Characklis GW, Noble RT. 2008. Sediment-water exchange of *Vibrio* sp. and fecal indicator bacteria: Implications for persistence and transport in the Neuse River Estuary, North Carolina, USA. *Water Research* 42:941–950.

50. Padovan A, Siboni N, Kaestli M, King WL, Seymour JR, Gibb K. 2021. Occurrence and dynamics of potentially pathogenic vibrios in the wet-dry tropics of northern Australia. *Mar. Env. Research* 169.

51. Perkins TL, Clements K, Baas JH, Jago CF, Jones DL, Malham SK, McDonald JE. 2014. Sediment composition influences spatial variation in the abundance of human pathogen indicator bacteria within an estuarine environment. *PLoS ONE* 9.

52. Abia ALK, James C, Ubomba-Jaswa E, Momba MNB. 2017. Microbial remobilisation on riverbed sediment disturbance in experimental flumes and a human-impacted river: Implication for water resource management and public health in developing sub-saharan African countries. *International Journal of Environmental Research and Public Health* 14.

53. Li L, Mendis N, Trigui H, Oliver JD, Faucher SP. 2014. The importance of the viable but non-culturable state in human bacterial pathogens. *Frontiers in Microbiology* 5:1–1.

54. Zhou C, Gao Y, Zhang H, Luo M, Ma T, Li G, Vandeputte D, Leermakers M, Baeyens W. 2024. Phosphorus mobilization in sulfidic sediments in the Baltic Sea. *Science of the Total Environment* 907.

55. Wu Y, Wen Y, Zhou J, Wu Y. 2014. Phosphorus release from lake sediments: Effects of pH, temperature and dissolved oxygen. *KSCE Journal of Civil Engineering* 18:323–329.

56. Westrich JR, Ebling AM, Landing WM, Joyner JL, Kemp KM, Griffin DW, Lipp EK. 2016. Saharan dust nutrients promote *Vibrio* bloom formation in marine surface waters. *PNAS*113:5964–5969.

57. Lehmann A, Myrberg K, Post P, Chubarenko I, Dailidiene I, Hinrichsen HH, Hüssy K, Liblik T, Meier HEM, Lips U, Bukanova T. 2022. Salinity dynamics of the Baltic Sea. *Earth Sys. Dynamics* 13:373–392.

58. Shaw KS, Jacobs JM, Crump BC. 2014. Impact of hurricane Irene on *Vibrio vulnificus* and *Vibrio parahaemolyticus* concentrations in surface water, sediment, and cultured oysters in the Chesapeake Bay, MD, USA. *Frontiers in Microbiology* 5.

59. Jespersen AM, Christoffersen K. 1987. Measurements of chlorophyll- a from phytoplankton using ethanol as extraction solvent. *Arch. Hydrobiol*. 109: 445-454.

60. Holmes RM, Aminot A, Kérouel R, Hooker BA, Peterson BJ. 1999. A simple and precise method for measuring ammonium in marine and freshwater ecosystems.

61. Hansen HP, Koroleff F. 2007. Determination of nutrients, p. . In Grasshoff, K, Kremling, K, Ehrhardt, M (eds.), *Method of seawater analysis.*

62. Helaleh MI, Fujii S, Korenaga T. 2001. Column silylation method for determining endocrine disruptors from environmental water samples by solid phase micro-extraction. *Talanta*.

63. Brussaard CPD, Payet JP, Winter C, Weinbauer MG. 2010. Quantification of aquatic viruses by flow cytometry, p. 102–109. In *Manual of Aquatic Viral Ecology*.

64. Herlemann DPR, Labrenz M, Jürgens K, Bertilsson S, Waniek JJ, Andersson AF. 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME Journal* 5:1571–1579.

65. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* 13:581–583.

66. Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil PA, Hugenholtz P. 2022. GTDB: An ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Research* 50:D785–D794.

67. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*.

68. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badretdin A, Bao Y, Blinkova O, Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey KM, Murphy MR, O'Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C, Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE, Vatsan AR, Wallin C, Webb D, Wu W, Landrum MJ, Kimchi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD, Pruitt KD. 2016. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research* 44:D733–D745.

69. Edgar RC. 2018. Updating the 97% identity threshold for 16S ribosomal RNA OTUs. Bioinformatics 34:2371–2375.

70. Panicker G, Call DR, Krug MJ, Bej AK. 2004. Detection of pathogenic *Vibrio* spp. in shellfish by using multiplex PCR and DNA microarrays. *Applied and Environmental Microbiology* 70:7436–7444.

71. Möller L, Kreikemeyer B, Gerdts G, Jost G, Labrenz M. 2021. Fish as a winter reservoir for *Vibrio* spp. in the southern Baltic Sea coast. *Journal of Marine Systems* 221.

72. Barnett D, Arts I, Penders J. 2021. microViz: an R package for microbiome data visualization and statistics. Journal of Open Source Software 6:3201.

73. Oksanen J SG Blanchet FG. Vegan: community ecology package. v.2.6.4.

74. Martinez Arbizu P. 2020. Pairwise multilevel comparison using adonis. R package version 0.4.

75. Gu Z. 2022. Complex heatmap visualization. iMeta 1.

76. Kirschner AKT, Pleininger S, Jakwerth S, Rehak S, Farnleitner AH, Huhulescu S, Indra A. 2018. Application of three different methods to determine the prevalence, the abundance and the environmental drivers of culturable *Vibrio cholerae* in fresh and brackish bathing waters. *Journal of Applied Microbiology* 125:1186–1198.

Chapter IV: Control of *Vibrio vulnificus* proliferation in the Baltic Sea through eutrophication and algal bloom management

The following chapter was published in the Communication Earth & Environment as:

David J Riedinger, Victor Fernández-Juárez, Luis F Delgado, Theodor Sperlea, Christiane Hassenrück, Daniel PR Herlemann, Christian Pansch, Marija Kataržytė, Florian Bruck, Alwin Ahrens, Marcin Rakowski, Kasia Piwosz, Angela Stevenson, Thorsten BH Reusch, Greta Gyraitė, Detlef Schulz-Bull, Heike Benterbusch-Brockmöller, Sandra Kube, Susann Dupke, Anders F Andersson, Lasse Riemann, Matthias Labrenz (2024): Control of *Vibrio vulnificus* proliferation in the Baltic Sea through eutrophication and algal bloom management *Communications Earth & Environment*, *5*(1), p.246

David Riedinger's contribution to the written manuscript was ~80%

Abstract

Due to climate change the pathogenic bacterium *Vibrio vulnificus* proliferates along brackish coastlines, posing risks to public health, tourism, and aquaculture. Here we investigated previously suggested regulation measures to reduce the prevalence of *V. vulnificus*, locally through seagrass and regionally through the reduction of eutrophication and consequential formation of algal blooms. Field samples collected in the summer of 2021 covered the salinity and eutrophication gradients of the Baltic Sea, one of the largest brackish areas worldwide. Physico-, biological- and hydrochemical parameters were measured and variables explaining *V. vulnificus* occurrence were identified by machine learning. The best *V. vulnificus* predictors were eutrophication-related features, such as particulate organic carbon and nitrogen, as well as occurrence of potential phytoplankton blooms and associated species. *V. vulnificus* abundance did not vary significantly between vegetated and non-vegetated areas. Thus, reducing nutrient inputs could be an effective method to control *V. vulnificus* populations in eutrophied brackish coasts.

Introduction

The Baltic Sea is a semi-enclosed marginal sea of the Atlantic located in northern Europe, with a coastline of approximately 8000 km and covering an area of 415,266 km2. Saline inflows through the North Sea produce a 2000 km long lateral surface salinity gradient throughout the entire Baltic Sea, ranging from high salinities (>25) in the transition zone of the Kattegat to low salinities (<5) in the Gulf of Bothnia[1]. The Baltic Sea is characterized by an estuarine-like circulation due to the positive freshwater budget. The drainage area of the Baltic Sea encompasses a population of approximately 85 million, and is consequently heavily influenced by eutrophication[2]. In addition, annual mean sea-surface temperatures are rising[2] and the ecosystem is expected to be increasingly affected by warming in the coming decades[3], and will be faced with extended heat wave durations[4].

These changes favour the growth of pathogenic bacteria of the genus *Vibrio* and an increase in *Vibrio* spp. abundances, infection rates, and fatal cases along the Baltic Sea coastline has been reported[5-8]. The infections of predominantly immunodeficient humans can be associated with the consumption of raw or undercooked shellfish, but in the Baltic Sea, they frequently manifest as skin infections resulting from direct contact with coastal brackish water. Only a low number of infections are currently associated with *Vibro vulnificu*s in the Baltic Sea9, but these are usually severe and often lethal[6,10,11].

Temperature and salinity are widely accepted as the two primary regulators of *V. vulnificus* abundance and distribution[12-15], but factors related to eutrophication, such as elevated dissolved organic carbon [DOC[16]] concentration or dinoflagellate blooms[17], have been shown to stimulate *V. vulnificus* growth in laboratory settings. Due to its preference for intermediate salinities and proliferation at water temperatures >15 °C18, *V. vulnificus* experiences optimal growth conditions during the summer in the Baltic Sea. The expected spread of *V. vulnificus* does not only pose a significant threat to public health but also to the tourist, fishing, and aquaculture industries[19,20,21]. In consequence, the question arises, whether measures can be taken to regulate *V. vulnificus* abundances.

In a natural setting, macrophytes such as seagrass (*Zostera marina*) might reduce the abundance of pelagic and potentially pathogenic *Vibrio* spp. and other pathogens within the seagrass canopy[22,23]. The underlying mechanisms for this decrease are elusive but could

include increased sedimentation rates due to hydrodynamic attenuation[24-26], filter-feeding by benthic fauna[27,28], or allopathic chemicals exuding directly from the seagrass plants[29]. Regardless of the specific causal mechanisms through which *Z. marina* beds potentially impact *V. vulnificus* abundance, they are putative nature-based solutions for reducing *V. vulnificus*[23]. Consequently, this study aims to elucidate important factors for mitigating *V. vulnificus* abundances along the Baltic coast, evaluating the potential of seagrass as a nature-based solution at the local scale and the reduction of eutrophication at the regional scale, respectively. We measured a large array of physical, biological, and hydrochemical parameters, and used three parallel methods (cultivation, amplicon sequencing, and droplet digital PCR; ddPCR) for the quantification of *V. vulnificus* within, adjacent and far from seagrass meadows. This enabled us to simultaneously assess the relationship between seagrass, environmental factors, such as temperature, salinity, and (in)organic nutrients, eukaryotic and prokaryotic microbial communities, and *V. vulnificus* over a vast salinity range, facilitating the detection of more generalizable patterns. These patterns are applicable to a wide range of environmental conditions found in estuaries and marginal seas worldwide. By combining a large number of parameters with a machine learning approach, complex and non-linear relationships are identified between the environment and *V. vulnificus* abundance.

V. vulnificus is likely stimulated by eutrophication-induced algal blooms; therefore, reducing eutrophication appears to be a promising strategy for mitigating the health risks associated this bacteria. Based on the results of this study, a regulatory effect of seagrass on *V. vulnificus* could not be determined.

Results

Environmental gradient and eutrophication index

A northeast-to-southwest salinity gradient (6 – 15.4) across the Baltic Sea was observed (for locations of sampling points, see Methods: Fig. 1), consistent with the increase of average seagrass leaf length from 22 to 109 cm. The density of the meadows (125 to 1,059 stalks per m²), depth (0.6 - 4.7 m), temperature (15.8 – 21.4 C°), dissolved oxygen (DO) (7.66 to 14.51 mg L^{-1}), pH (7.96 - 8.98) and average grain size diameter (dx50) (98 - 1,623 µm) were not structured along the salinity gradient. Abundance of heterotrophic cells, ranging between 3.44 x 10^6 and 2.68 x 10^7 mL^{-1}, varied across the Baltic Sea, with the highest and lowest abundances along the German (BV-06 to BV-11 and BV-24) and Estonian (BV-14 to BV-16

and BV-25) coasts, respectively. Among these cells, high nucleic acid (HNA) cell counts were very similar with 1.3 - 1.4 x 10^6 cells mL^{-1}, but low nucleic acid (LNA) counts varied between 1.37 x 10^6 and 1.44 x 10^7 cells mL^{-1}. Autofluorescent organisms were similarly distributed among stations, with *Synechococcus* abundances ranging from 4.09 x 10^4 to 1.12 x 10^6 cells mL^{-1}, picoeukaryotes from 1.41 x 10^3 to 5.92 x 10^4 cells mL^{-1}, and nanoeukaryotes from 2.17 x 10^2 to 4.95 x 10^4 cells mL^{-1}.

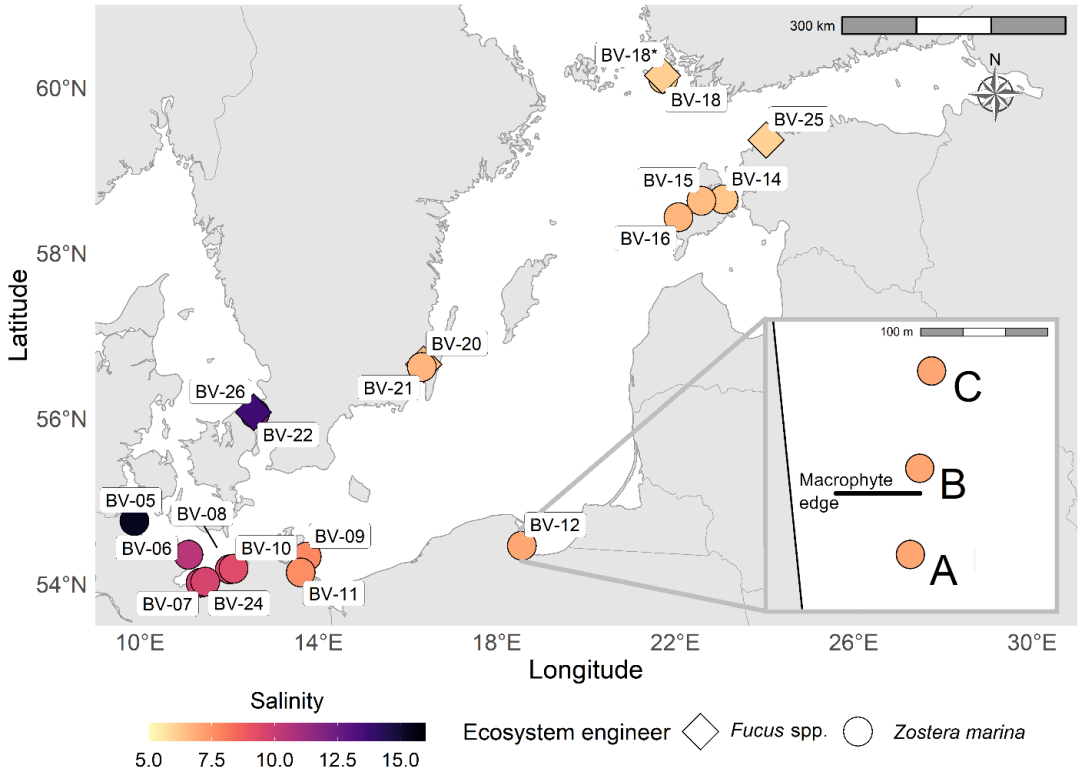


Figure 1. Map of the Baltic Sea showing sampling stations and salinity levels. *A schematic zoomed-in view of one station (BV-12) is provided to illustrate the three substations sampled at every station: A, B, and C. Station A is located within a macrophyte meadow, while station B is located 15 m, and station C is located 100 m from the meadow edge.*

Phosphate (PO_4^{3-}) and nitrate (NO_3^-) concentrations ranged from <0.1 to 0.93 and <0.2 to 1.46 µM, respectively, with the highest concentrations at the Swedish coast (BV-20, BV-21). The highest concentrations of nitrite (NO_2^-) and ammonium (NH_4^+), reaching up to 0.15 µM and 6.94 µM, were observed at the Danish coast (BV-26). Silica (SiO_2) concentrations ranged between < 2 and 48.92 µM and did not show strong geographic structuring. Particulate organic carbon (POC) varied between 11.21 and 276.83 µM, particulate organic nitrogen

(PON) between 1.64 and 41.76 µM, DOC between 180.10 and 681.30 µM, dissolved organic nitrogen (DON) between 8.30 and 30.60 µM and chlorophyll-a (chl-a) between 0.63 and 21.16 µg L^{-1}, with the highest concentrations found along the German coast. The eutrophication index was highest at the German coast (BV-11) and lowest at the Danish coast (BV-26).

No significant differences in physicochemical parameters within the water column were observed between substations A (in the macrophyte meadow), B (15 m from the macrophyte meadow), and C (100 m from the macrophyte meadow), for both *Z. marina* and *Fucus* spp. (serving as controls, see Methods) stations (Supplementary Fig. 1).

Macrophytes had no significant effect on *V. vulnificus* abundance

V. vulnificus was detected in 47% of the samples using ddPCR, in 33% through 16S rRNA gene sequencing, and both methods identified it in 20% of the cases. Across all sampled stations, local effects of *Z. marina* or *Fucus* spp. on *V. vulnificus* absolute or relative abundances were insignificant. *V. vulnificus* was equally distributed between substations A, B, and C (Wilcoxon rank sum test (Fig. 2)). This was consistent for *vvhA* gene copy numbers and 16S rRNA gene relative abundances, green colony forming units (CFUs), which are presumed, but not unequivocally identified, *V. vulnificus* colonies, and for heterotrophic bacterial cell counts (p-values > 0.05). Likewise, there was no consistent pattern between substations A, B, and C in *vvhA* gene copy number nor relative 16S rRNA gene abundances when observing the individual stations (Supplementary Figs. 2 & 3). *V. vulnificus* in sediments was also evenly distributed among substations (Supplementary Fig. 4). Based on the *vvhA* genes and 16S rRNA gene sequencing, *V. vulnificus* was not present in *Z. marina* biofilms. However, one *V. vulnificus* strain was isolated from seagrass leaves at station BV-12. In addition, other potential pathogenic *Vibrio* spp. found in the Baltic Sea were also distributed equally between substations on a Baltic Sea wide level (Supplementary Fig. 5).

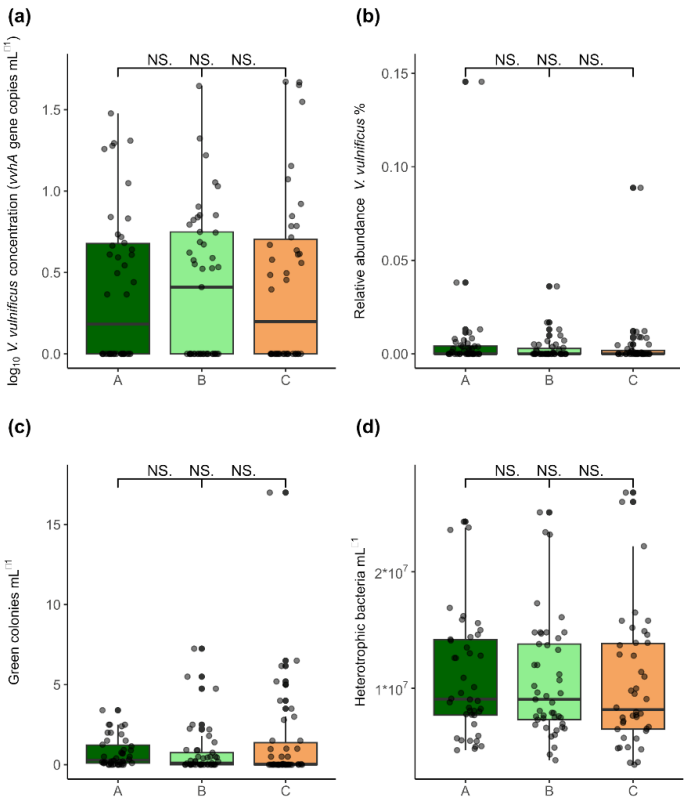


Figure 2.* V. vulnificus *abundance and microbial abundances at three substations summarized across all sampled locations. ***a*** V. vulnificus *abundances (*vvhA *gene copies* mL^{-1}*),* ***b*** *Relative* V. vulnificus *abundance based on 16S rRNA genes,* ***c*** Presumptive V. vulnificus *based on green CFUs* mL^{-1} *on TCBS agar,* ***d*** *Number of heterotrophic bacteria* mL^{-1}*. Station A is located within a macrophyte meadow, while station B is located 15 m, and station C is located 100 m from the meadow edge. Boxes represent the 25th and 75th percentiles.*

Eutrophication impacted *V. vulnificus* abundance

A positive relationship was observed between temperature and *V. vulnificus* abundance (approximated by *vvhA* gene copies mL^{-1}, Fig. 3). Within the temperature range of 19.5 – 20.5 °C, however, strongly eutrophied samples exhibited higher *vvhA* gene numbers compared to other samples in the same, as well as other, temperature ranges, coinciding with higher cell abundances of heterotrophic bacteria in general (Fig. 3). The average concentration of the *vvhA* gene in this temperature bin was higher than that at higher temperatures, and, on average, the bin exhibited higher eutrophication. The eutrophication index captured more of the variance in *V. vulnificus* abundance (36%) than temperature (12%) (Supplementary Fig. 6). The combination of both explained 43% while the more comprehensive and non-linear random forest (RF) including all parameters explained 59% of the variance.

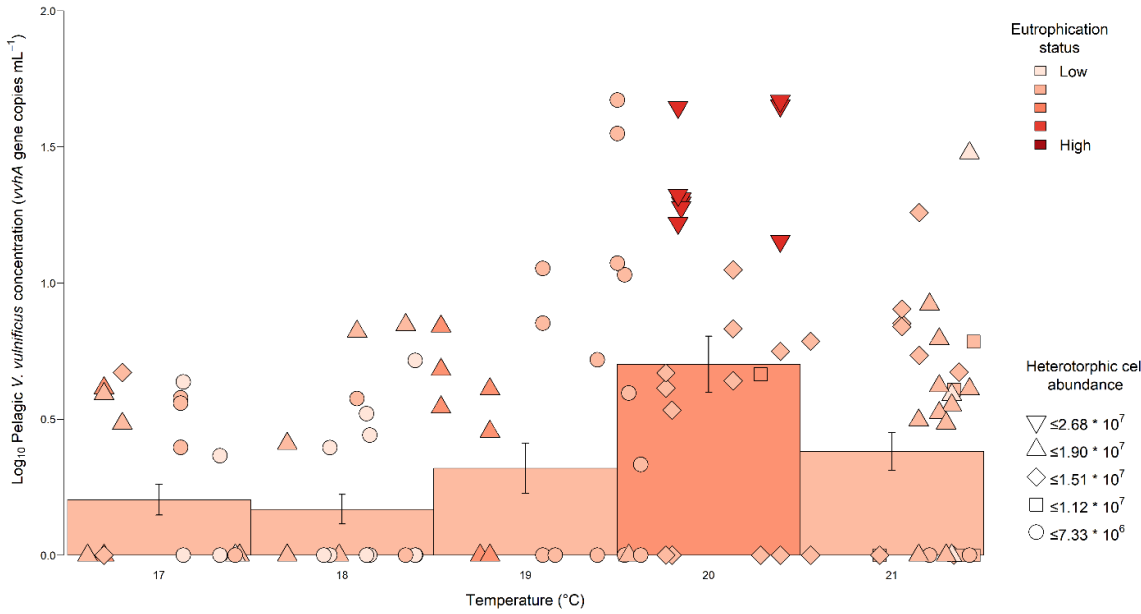


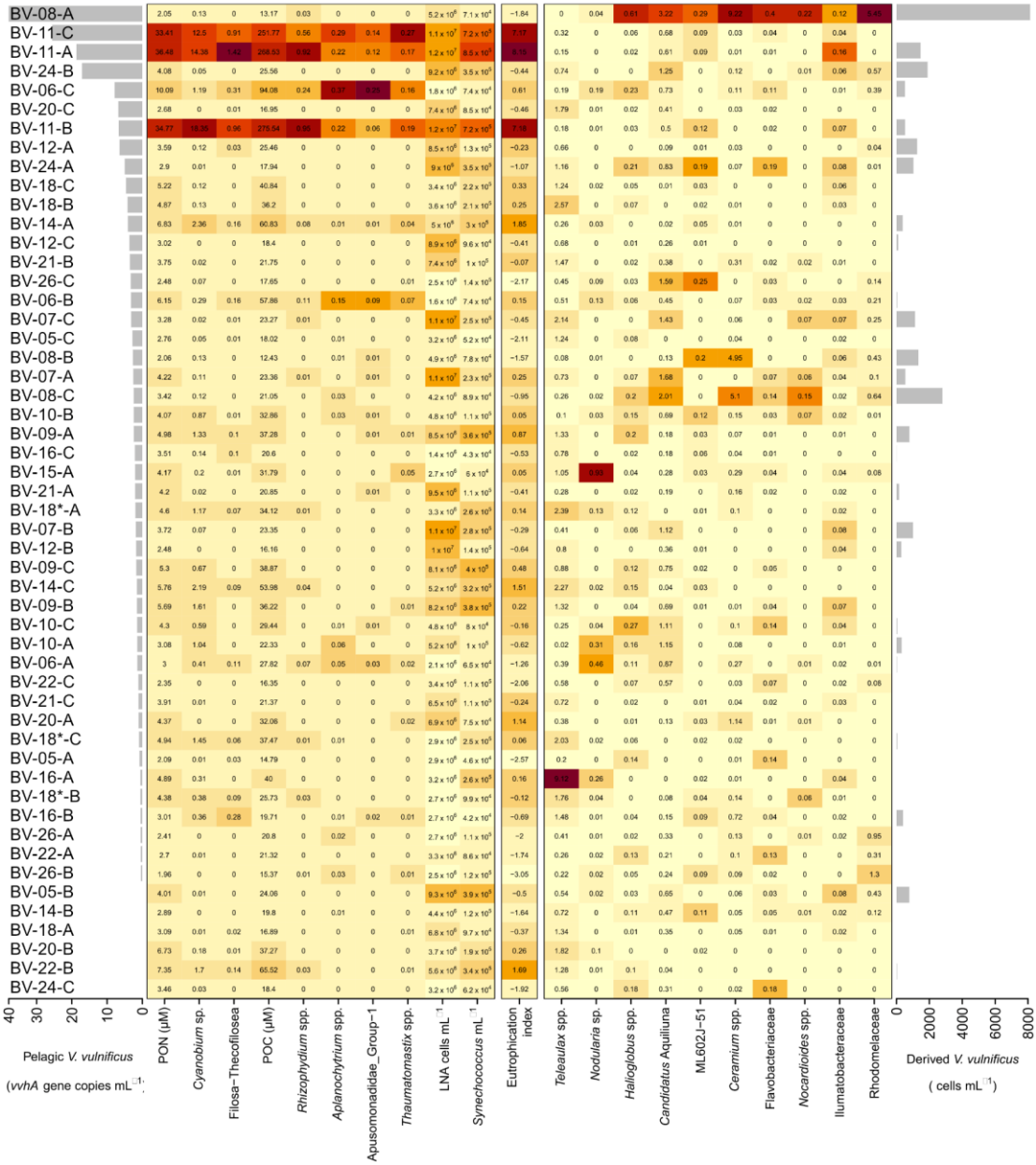
Figure 3. Relationship between temperature, eutrophication index, and pelagic* V. vulnificus *(vvhA copies mL−1): *The x-axis consists of 1-degree temperature bins for the bars and is continuous for the points, the y-axis is the log10 transformed* vvhA *gene copy numbers. Colours represent the eutrophication status per sample and are averaged per bin for the bars. The point shape indicates heterotrophic bacterial cell abundance, grouped into five bins, ranging from 3.4 × 10^6 to 2.7 × 10^7 cells mL^{-1}. Error bars represent standard deviation.*

Prevalence of *V. vulnificus* was correlated with indicators of blooming situations

Both RF models exhibited moderate performance, explaining 0.59 ± 0.32 and 0.59 ± 0.45 of the variance in the derived total *V. vulnificus* abundances and of the variation of *vvhA* gene copies mL^{-1}, respectively. Root Mean Square Error (RMSE) was 697.16 ± 597.50 and 6.33 ± 4.66 of the target values 8,130.20 and 37.93, and the mean absolute error (MAE) was 519.76 ± 307 and 4.56 ± 2.55.

Prokaryotes identified as the most important predictors for *V. vulnificus* abundance belonged, among others, to the phylum Cyanobacteria, with representatives of the genera *Cyanobium* (ASV219, ASV379) and *Nodularia* (ASV2690), as well as Actinobacteria, with *Candidatus* Aquiluna(ASV236), ML602J-51 (ASV654) and *Nocardioides* (ASV1070) (Fig. 4; Supplementary Table 1). The most important eukaryotic predictor taxa for *V. vulnificus*, which could be classified at the genus level, were the rodophyte *Ceramium* sp. (ASV71), the Cercozoa *Thaumatomastix* and *Aplanochytrium* (ASV320, ASV430), the mixotrophic organism *Teleaulax* (ASV6) and fungi *Rhizophydium* spp. (ASV312) (Fig. 4; Supplementary

Table 2). The relative abundances of these taxa, except *Teleaulax* and *Nodularia*, were notably higher at elevated *vvhA* gene concentrations and derived *V. vulnificus* cell numbers (Fig. 4). *Cyanobium* spp. encompassed up to 18.4%, and *Ceramium* up to 9.2%, of the ASV counts, respectively. The high relative abundances of these pro-and eukaryotic taxa coincided with high POC and PON, high *Synechococcus* abundance, and high LNA cell concentration (Fig. 4).



Previous page: Figure 4. Heatmap of the most important predictors for the* vvhA *gene copy number, derived* V. vulnificus *abundance, and the eutrophication index. *Predictors are ordered from high (left) to low (right) importance for the respective models. Observations are ordered according to decreasing* vvhA *gene copy numbers. Values of each predictor are scaled from 0 to 1 for color representation in the heatmap, while the original values are printed in each cell.*

Discussion

We present a comparative study of 19 coastal Baltic Sea sampling stations in summer/early fall 2021, exploring the potential of seagrass as a local solution and the reduction of eutrophication as a more regional strategy to mitigate the proliferation of potentially pathogenic *V. vulnificus*. For these stations, physical-, biological- and hydrochemical parameters were measured and variables explaining *V. vulnificus* occurrence were identified. We found that eutrophication-related parameters such as POC/PON and high *Cyanobium* sp. and *Synechococcus* sp. abundance predicted high *V. vulnificus* abundances, while the occurrence and density of *Z. marina* showed no predictive value. This implies that reducing eutrophication on a regional level could be a promising strategy for constraining further proliferation of *V. vulnificus* along the Baltic Sea coast.

The distribution and abundance of *V. vulnificus* has been extensively studied, with temperature and salinity being consistently identified as key factors[12-15]. Our study observed temperatures exceeding 15 °C and salinities ranging from 6 to 16, which are generally considered favorable conditions for *V. vulnificus*[5,30], yet abundances varied between stations. Notably, stations with a high eutrophication index tended to exhibit higher *V. vulnificus* abundances compared to those with similar or even higher temperatures but a lower eutrophication index (Fig. 3). Furthermore, the eutrophication index explained a greater proportion of the variance in the *V. vulnificus* abundance than temperature (Supplementary Fig. 6). This finding is consistent with earlier studies, which also related elevated concentrations of DOC and chl-a to high *V. vulnificus* abundances[16,30]. Chl-a was, besides being part of the eutrophication index, also closely associated with the predictors for *V. vulnificus* identified by RFE, namely POC, PON, *Cyanobium* sp. and *Synechococcus* (relative) abundance in this study (Supplementary Fig. 7). Additionally, the effectiveness of heterotrophic cell counts (LNA cells in Fig. 4) in the prediction underscored the link to eutrophication. These results align with observations from the Neuse River estuary, which faces similar problems with eutrophication as the Baltic Sea, where elevated levels of both total *Vibrio* spp. and *V. vulnificus* corresponded with high phytoplankton biomass[31] and an investigation after Hurricane Ian (October 2022),

where high phytoplankton mass, approximated by chl-a concentration, and associated zooplankton abundance, was found to potentially stimulate *V. vulnificus* proliferation[32].

Cyanobium sp., which thrives under eutrophic conditions and is known to form blooms and correlate positively with the decay phases of other blooms[33], was found to be an important *V. vulnificus* predictor. The identified ASVs for *Cyanobium* sp. comprised up to 18.7% of the total prokaryotic sequence reads at some of the highest *vvhA* gene copy numbers. Associations between pathogenic *Vibrio* spp. and (harmful) algal blooms have been observed before[34] and our results substantiate this relationship. It is plausible that eutrophication, by stimulating blooms, served as an important indirect driver of the distribution and abundance of *V. vulnificus* via the release of DOC/DON and POC/PON. The idea that the increase in DOC/DON triggered by the bloom contributed to the proliferation of *V. vulnificus* prior to sampling is supported by the predictive power of the actinobacteria *Candidatus* Aquiluna (ASV236), ML602J-51 (ASV654), and *Nocardioides* (ASV1070). Actinobacteria are commonly associated with the demise of phytoplankton blooms and the generation of cyanobacterially sourced DOC[35,36]. The stimulation of *V. vulnificus* abundance by algal blooms was also inferred from the high *V. vulnificus* abundances coinciding with high filamentous *Ceramium* sp. abundances, peaking at 9.2% of the sequence reads during both the highest *vvhA* gene concentrations and derived *V. vulnificus* cell counts at station BV-08-A (Fig. 4). *Ceramium* sp. is an opportunistic fast-growing algae, with high nutrient uptake rates[37]. Interestingly enough, the eutrophication index, POC and PON are all low at this station. This could be an indication that the blooming (in terms of relative abundance) of specific plankton, for example *Ceramium* sp. and *Cyanobium* sp. can already convey some benefit. POC / PON are still good predictors of *V. vulnificus* in that scenario, as they generally increase with phytoplankton blooms, including all the blooms that provide species-specific benefits. Regardless of the specific combination of factors that promote *V. vulnificus* proliferation, reducing eutrophication would limit the extent of algal blooms in the Baltic Sea, likely reducing infection risk. Among the top five instances of highest *vvhA* gene copy abundances and derived *V. vulnificus* cell numbers, either *Ceramium* or *Cyanobium* sp. was consistently, except at station BV-24-B, found at high relative abundance.

The high relative abundance of *Cyanobium* sp. was concurrent with other species expected to either infect or co-occur with a bloom. The parasite *Aplanochytrium* sp., which is known to bloom[38] and to correlate strongly with phototrophs[39], and the fungi *Rhizophydium* sp. (ASV 312)[40], reached their highest relative abundance (max 0.4% and 1%, respectively) at stations with high *Cyanobium* sp. and *V. vulnificus* (relative) abundance. Under conditions

conducive for extensive bacterial growth, such as during blooms, *Thaumatomastix* (ASV320) might be able to temporarily compete with other organisms[41]. *Thaumatomastix* exceeded 0.1% of the eukaryotic community only in samples where *Cyanobium* sp., *Synechococcus,* and their potential parasites were relatively abundant, further pointing to a highly productive system. This scenario coincided with elevated *V. vulnificus* abundance.

We speculate that when grazing pressures on the algal bloom reached critical levels, the bloom ceased, leading to a depletion of the organic nutrients crucial for the proliferation of *V. vulnificus*. In instances where we observed high relative abundances of the potential predator *Teleaulax* (>1%), there was a concurrent decrease in *Synechococcus* and *Cyanobium* sp. (relative) abundances and a decrease in the concentrations of PON and POC (e.g. BV-18 & BV-20). This phenomenon was consistently associated with the absence or low relative abundance of *V. vulnificus*. This may be attributed to the potential role of *Teleaulax* in depleting the *Synechoccocus* population[42,43] and thereby reducing the food sources for *V. vulnificus.*

While organic nutrients were good predictors of *V. vulnificus* abundance, inorganic nutrients were not. Additionally *V. vulnificus* was generally less correlated with inorganic nutrients (e.g. PO_4^{3-}) than heterotrophic bacteria overall (Supplementary Fig. 8). This suggests that potential bloom-forming organisms may play a pivotal role in the chain leading from inorganic nutrient input to increased *V. vulnificus* abundance by converting inorganic nutrients into organic forms.

In addition to the increase in organic constituents present during a phytoplankton bloom, the higher abundance of phytoplankton and the associated zooplankton community could potentially offer protection against grazing by bacterivorous protozoa, as previously demonstrated for *V. cholerae*[44]. *V. cholerae* population growth is normally balanced by protozoan predation, but during declining blooms, *V. cholerae* has been shown to multiply at an increased rate due to the higher availability of DOC and DON and/or change to a particle-attached lifestyle for protection from predation[45]. A similar protective effect has been shown for other species of *Vibrio*, which significantly benefitted from association to phytoplankton during periods of intense grazing[46]. We hypothesize that *V. vulnificus* might also benefit from a similar interaction.

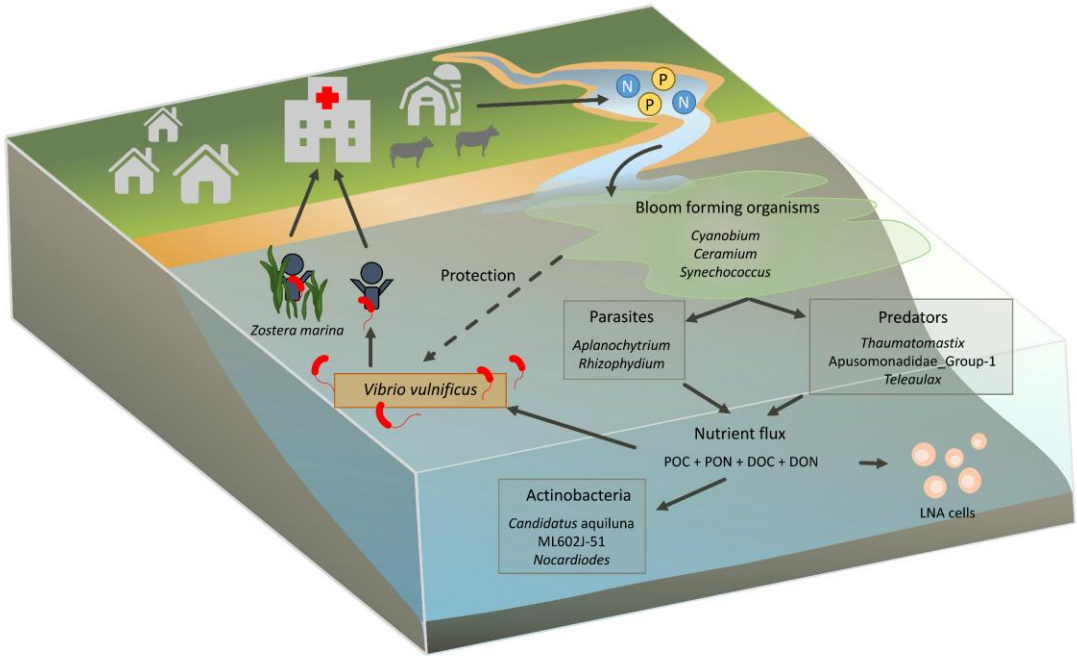


Figure 5. Overview of the suggested pathway through which eutrophication promotes* V. vulnificus *proliferation. *P and N represent the total dissolved inorganic phosphorus and nitrogen. Inorganic nutrient inflow from the land induces algal blooms, providing the organic material required for* V. vulnificus *proliferation and potentially protection from bacterivorous protozoa. The figure also displays other organisms identified as important predictors and their potential connections to blooms. Notably,* Zostera marina *could not be shown to significantly affect abundance in our study.*

The discussed microorganisms and eutrophication index exhibited greater explanatory power than temperature for *V. vulnificus* abundance. This was likely due to their sensitivity to historical and current temperatures and nutrient availability. As a result, they served as a record of the water mass that was more predictive than temperature alone. The suggested pathway through which eutrophication could stimulate *V. vulnificus* abundance is summarized in Fig. 5. The identity of the blooming organism and its predators might not be as critical as the fact that the bloom and its associated predators stimulate organic nutrient availability and might provide a degree of protection.

While eutrophication is often a regional problem, with a regional solution, such as the Baltic Sea action plan[47]. Intraregional differences exist between bays due to differences in nutrient sensitivity[48] and nutrient input from point sources[49]. Engineering solutions to eutrophication in semi-enclosed Baltic Sea bays have been tested successfully[50], potentially providing a solution for some Baltic Sea coasts. Besides allowing for a potential (multi)-

national approach to reduce the *V. vulnificus* hazard, the possible impact of eutrophication can be used to improve risk assessment and early warning by integrating the existing HELCOM Eutrophication Assessment Tool[48] into the currently available ECDC *Vibrio* viewer, which relies solely on temperature and salinity, for the Baltic Sea[51]. In other estuaries or marginal seas, the integration of chl-a data from remote sensing into early warning systems may provide an improved prediction of *V. vulnificus* abundance.

The assessment of *V. vulnificus* predictors was hampered by the observed discrepancies between the quantification methods employed. They are likely due to multiple biases associated with each methodology. These are, in no particular order: CFUs by nature of the method, underestimate the real abundance, as they are unable to quantify viable but not culturable cells, resulting in the lowest abundances measured by any of the methods (Fig. 2). Differences between *vvhA* gene copy numbers and derived cell counts can be partly attributed to multiple copies of the 16S rRNA gene[52], which range between 8 and 12 for *V. vulnificus*[53]. We did not correct for these, according to the advice of Louca et al. (2018)[54]. Additionally, not all *V. vulnificus* cells contain the *vvhA* gene, which further complicates the comparison. The discrepancy between *vvhA* and derived *V. vulnificus* abundance could be an indication that a sizeable part of the sampled *V. vulnificus* community lacks this pathogenicity gene. Further differences might arise from primer bias, PCR biases and biases in quantification of bacterial counts. When assessing the *vvhA* gene in our study, fewer instances of *V. vulnificus* are identified per station; however, *V. vulnificus* is detected at more stations. A prime example is station BV-11-C (Fig. 4), which has the second highest *vvhA* gene copy numbers, and where presumable *V. vulnificus* colonies where cultured, but no *V. vulnificus* was detected with the 16S rRNA amplicon sequencing, while at stations BV-11-A and BV-11-B *V. vulnificus* is detected by all methods. In this case, the discrepancy arose because the 16S rRNA gene sequences amplified at station BV-11-C could not unambigiously be identified as *V. vulnificus*. A similar degree of similarity was observed with a related *Vibrio* species. Additional discrepancies might arise from *V. vulnificus* variants not accounted for in the database used for comparison. This complicates the comparison between the methods, but overall, the association between algal blooms and *V. vulnificus* could still be shown across all measured parameters.

V. vulnificus was found to be absent from seagrass leaves by molecular methods and only one strain could be isolated from them, even though *V. vulnificus* is known to form biofilms[55] under a variety of conditions[56-58]. This is surprising given the release of labile DOC from seagrass leaves[59] but could be due to allelopathic chemicals exuded from the seagrass

leaves hampering the growth of *V. vulnificus,* as observed for a general bacterial community[60]. An alternative explanation is that there is a strong association between the seagrass and their associated microbiome and that *V. vulnificus* was consistently outcompeted[61,62]. *V. vulnificus* was, however, found in the benthic and pelagic environment, in similar abundances at substations inside and outside of the meadows (Fig. 2, Supplementary Figs. 1, 2), implying that the impact of seagrass meadows on *V. vulnificus* abundance, relative abundance, and culturable potential within the surrounding environment was minor in the Baltic Sea. Additionally, no effect on the relative abundance of other potential *Vibrio* pathogens could be detected within the sampled salinity range (Supplementary Fig. 5). Similar to the findings in San Diego coastal waters, our results contrast the reduction in pathogenic bacteria abundance observed by Lamb et al. in 2017[63]. Future findings may differ for other conditions, e.g., water temperatures, salinity ranges, or less eutrophied systems, which were outside our study framework. The discrepancy between our study and earlier work on the influence of seagrass on potential pathogens may also be ascribed to a difference in methodology or study area. Reusch et al. (2021) used solely chrom-agar based blue colony counts at salinities between 14.3 and 22.1[23], while we employed methods more specific to *V. vulnificus* over a different salinity gradient. The salinity of our study areas was mostly lower and may have affected the physiology of seagrass and its possible effect on *V. vulnificus*. Based on our results, it could not be shown that seagrass meadows are a direct effective nature-based solution for reducing *V. vulnificus* abundance and associated infections in the Baltic Sea. However, the current high nutrient concentrations in the Baltic Sea may have masked possible direct effects of seagrass, and probably also *Fucus* spp., on *V. vulnificus,* and indirect effects through the reduction of the nutrient load.

Our study suggests that a reduced anthropogenic nutrient input could mitigate the cascade of events that stimulate algal blooms and through that, organic nutrient availability. This in turn, would make conditions less conducive for *V. vulnificus* proliferation. This could reduce public health risks and be a critical management tool for *V. vulnificus* in the Baltic Sea, as well as in other coastal brackish water systems worldwide.

Materials and Methods

Sampling stations and SCUBA sampling strategy

Samples from fifteen *Z. marina* and, for comparison, four *Fucus* spp. stations were collected from July 25th to September 2nd, 2021, covering the German, Estonian, Finnish,

Polish, Swedish, and Danish coasts (Fig. 1). Water and sediment samples were collected inside the macrophyte meadows (substation A), as well as the leaves/fronds of the macrophyte. Analogously, water and sediment samples were collected at control stations without macrophytes, located 15 m (substation B) and 100 m (substation C) from the edge of the meadow (Fig. 1). Sampling was carried out by SCUBA or snorkel divers at water depths ranging from 0.6 to 4.7 m. Sampling was only performed up to a moderate breeze of 4 on the Beaufort scale with wave heights <0.5 m to prevent heavy mixing between substations. The prevailing current direction was visually determined at the sampling depth by releasing food colourant (Dr. Oetker 4 Back & Speisefarbe). Divers approached the sampling site against the current and maintained buoyancy to prevent sediment resuspension.

Water samples were collected following two methods. First, for DNA extraction, flow cytometry, and plating, 100 mL of water was collected in 9 replicates using rinsed syringes. The samples were collected ca. 5 cm from a macrophyte and ca. 20 cm above the sediment at substation A, and ca. 20 cm above the sediment at substations B and C. Second, sterile 5 L plastic bags, used for chl-a, inorganic nutrients, POC and PON, DOC and DON determination, were filled by swimming against the current while keeping the bag open 20 cm above the sediment. This water was split between 1 L plastic and glass bottles after triplicate rinsing. Sediment was collected in nine sterile 50 mL Falcon tubes by scraping the top 1 cm of sediment. At substation A, *Z. marina*/*Fucus* spp. samples were collected in nine replicates by pinching the connection between the root and stalk. Subsequently, the leaves/fronds were separated by hand for further processing. All samples were transferred to a 4°C cooler immediately and stored (maximally 8 h) until processing.

Macrophyte characteristics

Z. marina plant densities were determined by counting 20 x 20 cm squares in triplicate. The length of 30 leaves was measured per meadow. *Fucus* spp. densities were not determined due to the large heterogeneity within the individual beds.

Environmental parameters

Salinity, temperature, and depth were measured using a CTD48M (Sea & Sun Technology, Trappenkamp, Germany) attached to a diver. DO and pH were measured using a handheld multimeter (HQ40D Portables 2-Channel Multimeter) (HACH, Iowa, USA) with Hach Intellical LDO101 and Intellical PHC101 probes, respectively.

Water from the plastic bottles was filtered through 25Ø GF/F filters (Whatman plc, Maidstone, UK), with the filters stored for chl-a determination at -80 °C (250 – 500 mL) and the filtrate for nutrient analysis stored at -20 °C. Water from the glass bottles was filtered through pre-combusted (450 °C, 4 h) 25Ø GF/F filters for POC and PON determination and the filtrate for DOC and DON analysis, all stored at -20 °C. PO_4^{3-}, NO_3^-, NO_2^-, ammonium NH_4^+, and SiO_2 concentrations were measured using a Seal Analytical QuAAtro automated constant flow analyzer (SEAL Analytical Ltd, Nordestedt, Germany), with detection limits of 0.1, 0.2, 0.05, 0.5 and 2 µM, respectively[64]. DOC and DON were analyzed with a TOCL-CPH/TOC-VCPH TOC-Analyzer (Shimadzu, Germany)[65] and POC/PON measurements were conducted with a varioMICRO cube element analyzer (Elementar Analysensysteme, Germany). Chl-a was measured fluorometrically using a 10- AU-005-CE fluorometer (Turner, San Jose, USA) according to the HELCOM 2019[66] guidelines and corrected for phaeopigment.

Sediment grain size

Grain size was analysed after 5 d of lyophilization with a Delta 1-24 LSCplus (Martin Christ Gefriertrocknungsanlagen, Osterode am Harz, Germany). Sediment samples were sieved through a 3.5 mm sieve and visible biological elements were removed. The remaining sediment chunks were broken by grinding for a minimum of 90 s with mortar and pestle. Samples were measured using the dry cell of a Mastersizer 3000 with a range from 0.01 to 3.500 µm (Malvern Panalytical, Malvern, UK).

Microbial analyses

Vibrio spp. CFUs and *V. vulnificus* isolates were obtained from water, sediment, and macrophytes from six independent replicate samples. For this, water aliquots of 50, 100 or 200 µL were (a) plated directly onto *Vibrio* selective thiosulfate citrate bile sucrose (TCBS) agar (Merck, Darmstadt, Germany) and (b) aliquots of 2, 5, 10 or 25 mL were filtered onto 0.2 µm PC-filters (Merck-Millipore, Burlington, USA) and placed onto TCBS agar. The sediment samples were homogenized after the removal of the overlying water, a subsample of approximately 10 g (dry-weight determined accurately after lyophilization) was transferred from six sediment samples to sterile 50 mL falcon tubes, where 40 mL of double 0.2 µm-filtered station water was added. To detach bacteria, five ultrasonic pulses of 10 s at 25% capacity at 5 s intervals using the Bandelin SONOPULS HD 2200.2 (Bandelin, Berlin, Germany) were applied. After subsequent vortexing and settling of the sediment, water

aliquots of 50, 100, or 200 µL were plated on TCBS agar in six biological replicates. The same method was applied for the macrophytes. The only modification was that 5 - 20 mL of the supernatant was filtered over a PC-filter to obtain colonies. After 24 h incubation at 37 °C, CFUs of green colonies were determined for all plates.

To isolate *V. vulnificus*, green colonies from stations BV-05, BV-06, BV-09, BV-10, BV-12, BV-15, BV-20, and BV-21 were further cultivated on CHROMagar_vibrio™ (Chromagar Ltd. Paris, France) for 24 h at 37 °C and blue-colored colonies were restreaked on TCBS agar, CHROMagar, and Columbia sheep blood agar (Oxoid, Basingstoke, UK). DNA of isolates which were green on TCBS, blue on CHROMagar and confirmed to be pure cultures on blood agar, was extracted (DNeasy Blood and Tissue Kit, Qiagen, Hilden, Germany), and purity and concentration were determined using a NanoDrop Spectrophotometer (Thermo Fisher, Waltham, USA). To confirm *V. vulnificus,* the species-specific *vvhA* gene sequence was targeted using multiplex real-time PCR (5´ nuclease assay). In the same assay, an internal amplification control (KOMA) was detected[67]. Primers for the detection of the *vvhA* gene[45] are provided (Supplementary Table 3).

Enumeration of heterotrophic bacterial cells and autofluorescent phytoplankton

Four mL of water from three replicate 100 mL syringes, which were also used for the DNA extraction, was pipetted in duplicate into 5 mL sterile cryovials (VWR, Radnor, USA), and 200 µL formaldehyde (37%) was added. After homogenization and 1 h incubation at 8 °C, the samples were shock frozen and stored at -80 °C until flow cytometry analysis. Autofluorescent and heterotrophic cells were counted separately using a CytoFLEX S flow cytometer (Beckman coulter, Brea, USA) and unstained and SYBR Green I (Invitrogen, Waltham, USA) stained samples, respectively[68]. Autofluorescent cells were analysed with CytExpert software (Beckman coulter, Brea, USA) and grouped into *Synechococcus,* picoeukaryotes, and nanoeukaryotes. Stained samples were gated and clustered using FlowClust[69] after quality control by FlowAI[70] based on green and red fluorescence and side scatter[68]. The minimum Bayesian Information Criterion was reached for three clusters, showing beads, HNA, and LNA cells. The flow rate was monitored by adding beads to each sample (NFPPS-52-4K, Spherotech, Lake Forest, USA).

Molecular prokaryotic and eukaryotic community composition

Water (92 mL) from three syringes was filtered through 0.22 µm polyvinylidene fluoride membrane filters (Merck, Darmstadt, Germany), which were shock-frozen in liquid nitrogen and stored at -80°C for downstream molecular analysis. Three falcon tubes with

sediment from each substation and three falcon tubes with macrophytes from substation A were frozen at -80 °C. DNA extractions were performed using the DNeasy PowerSoil Pro Kit (Qiagen, Hilden, Germany). The sediment and macrophyte samples were thawed and homogenized. A subsample of 500 mg of sediment or 4-6 pieces of a macrophyte of 2-3 cm lengths were transferred into bead-beating tubes. These were placed on ice and sonicated twice for 7 min and bead-beaten for 30 s at 4 m/s. Subsequently, the manufacturer's instructions were followed and DNA yield was quantified using Picogreen (Thermo Fisher, Waltham, USA).

16S rRNA genes were amplified using primers, covering the prokaryotic V3–V4 hypervariable region[71], and 18S rRNA genes by targetting the eukaryotic V4 region (Supplementary Table 3)[72], with a PCR protocol derived from Latz et al. 2022[73] (Supplementary Table 4). Illumina sequencing adapters were included in the 5' ends of the 16S and 18S primers (Supplementary Table 3). Phased primers[74] were used to increase the complexity of the sequencing libraries. For 16S rRNA genes, forward and reverse primers were phased (CTAGAGT, TAGAGT, etc. for the forward and ACTACTG, CTACTG, etc. for the reverse primer), for 18S rRNA only the forward primer (ATG, TG, G, or no base) was phased. PCR thermal conditions and master mix details are provided (Supplementary Tables 4 & 5). Leftover adapters were removed using the MagSi-NGS PREP Plus Kit (MDKT00010075, magtivio BV., Nuth, the Netherlands). The purified product was indexed through a second PCR (Supplementary Table 4) following the Adapterama indexing scheme[75], pooled in equimolar ratios, and sequenced on MiSeq for 16S and 18S rRNA gene metabarcoding by SciLifeLab/NGI (Solna, Sweden). In addition to the MiSeq sequencing, the pelagic 16S rRNA gene libraries were deep-sequenced on NovaSeq 6000 (Illumina Inc, San Diego, CA, US).

Phased primer sequences were removed from the reads using a snakemake pipeline[76,77]. The pipeline (https://github.com/biodiversitydata-se/amplicon-multi-cutadapt) encompassed: the removal of read-pairs that contain Illumina adapters, exclusion of read-pairs lacking the expected primer sequences located at the 5' ends of the reads, removal of the primer sequences from the remaining reads and elimination of read-pairs that have misplaced primer sequences. DADA2[78] was used for denoising, concatenating paired-end reads, and chimera removal. The resulting amplicon sequence variants (ASVs) were taxonomically assigned with DADA2 using PR2[79] (V. 4.14.0) as a training set for 18S rRNA gene amplicons and Silva 138.1[80] for 16S rRNA gene amplicons.

Species-level classification of *Vibrio* ASVs was achieved by sequence comparison (BLAST, V. 2.13.0[81]) to a custom database. This database includes 16S rRNA gene sequences of complete *Vibrio* genomes from RefSeq[82] (51 species, 317 strains - including 22 *V. vulnificus* strains), 41 draft *V. vulnificus* genomes from clinical isolates from the Baltic Sea region[11,83], and 84 draft *V. vulnificus* genomes from environmental Baltic Sea isolates. In order for an ASV to obtain a species-level assignment to a *Vibrio* spp., we required perfect (100% identity) alignment of the full ASV to 16S rRNA genes of a single species in the custom database. None of the ASVs that perfectly matched to *V. vulnificus* 16S rRNA genes also matched perfectly to 16S genes of other species. In order to obtain derived *V. vulnificus* cells mL^{-1}, the integrated relative abundace of *V. vulnificus* ASVs was multiplied by the flow-cytometry measured heterotrophic cells per mL^{-1}.

Quantification of *V. vulnificus*

V. vulnificus was quantified using ddPCR (QX200™ Droplet Digital PCR System, Bio-Rad, München, Germany) targeting the *vvhA* gene[84] (Supplementary Table 3). PCR thermal conditions and master mix details are provided (Supplementary Tables 4 & 5). Droplets were generated using the QX100 droplet generator (Bio-Rad, Hercules, USA). Emulsified samples were transferred to a 96-well plate and sealed by a pierceable foil hot seal (BioRad, 181–4040) using PX1 PCR Plate Sealer™ (Bio-Rad, Hercules, USA) (5 s at 180 °C). The PCR was performed using a Bio-Rad C1000 Touch™ thermal cycler (Supplementary Table 4). Subsequently, the plate was analyzed with the QX100 droplet reader using the Quantasoft 1.74.09.17 software. Positive and negative controls were 50 ng of DNA of a *V. vulnificus* isolate and of a *V. harveyi* isolate, respectively. In the contamination control, template DNA was substituted with DEPC water.

Statistical analyses and machine learning

To assess the Baltic Sea wide differences in *V. vulnificus* abundance between substations A (N = 42), B (N = 45), and C (N = 42) of seagrass stations, a two-sided Wilcoxon rank sum test, assuming independence and equal variance, using the rstatix (V. 0.7.2) package in R (V. 4.3.0)[85] was performed and holm correction for multiple testing was applied. Additionally, heterotrophic bacterial cell abundances and green CFUs, inside and outside of the seagrass meadows, were compared. Sample sizes for comparison between substations are

consistent for the water column and sediment. No repeat measurements were performed, every data point is a distinct sample.

Three separate linear models for *V. vulnificus* were used to compare the explanatory power (R^2) of the traditional predictor temperature with the more integrated predictor eutrophication index and their combination, using averaged *vvhA* gene copies mL^{-1} (log10 transformed) per substation as a response.

To identify the environmental conditions associated with high *V. vulnificus* abundance, two random forests (RF) in combination with a recursive feature elimination (RFE) algorithm were employed in caret (V. 6.0.94)[86] predicting both the *vvhA* gene copies mL^{-1} and the derived *V. vulnificus* cells mL^{-1}. These models used sequencing and environmental data as predictors. Combining RF and RFE addressed correlated variables in this high-dimensional dataset and was able to detect complex and non-linear relationships between measured variables and *V. vulnificus*[87,88].

Prior to model training, the 16S and 18S rRNA relative gene proportions were pre-processed: Rare (absent in at least 25 samples) for both datasets were removed. All ASVs taxonomically classified as *Vibrio* were removed from the predictors. After pre-processing, the data set included 964 predictors, namely 314 eukaryotic, 627 prokaryotic, 14 physico-chemical, 7 biological ones, substation and macrophyte type. Samples collected from the same station were grouped, to avoid data leakage between test and train data.

Both RF models consisted of 2,000 trees with 30 randomly sampled variables as candidates at each split. The model was trained and evaluated using 10-fold cross-validation on a dataset of 52 observations, each representing the average of three biological replicates. Performance metrics used for evaluation included the mean absolute error (MAE), average absolute difference between predicted and actual values, root mean squared error (RMSE), a measure of the differences between the values predicted by the model and the actual values, and coefficient of determination (R^2) of the 10-fold cross-validation. The top 10 predictors of both models are discussed.

Eutrophication index

A eutrophication index, defined as the organic matter availability in an ecosystem[89], was derived by performing a principal component analysis (PCA), using the "FactoMineR (V 2.8)" R package[90], on the environmental parameters DOC, POC, PON, DN, chl-a, NO_3^-, and

PO_4^{3-}. All organic nutrients strongly aligned with principal component one, explaining 63%, which was chosen as the eutrophication index (Supplementary Fig. 9).

Data availability

16S and 18S rRNA gene data were archived in the European Nucleotide Archive under the accession number PRJEB68222[91] in compliance with the Minimal Information about any (X) Sequence (MIxS) standard[92] through the brokerage service GFBio[93]. Environmental data are available at IOWMeta (doi.io-warnemuende.de/10.12754/data-2023-0010)[94]. The reference *V. vulnificus* database is available under https://doi.org/10.5281/zenodo.10875108[95].

Code availability

All custom code used for the bioinformatics processing,the machine learning and analysis of the flow cytometry data in this study is available at https://github.com/lfdelzam/ASV_dada2_chunck/ and https://git.io-warnemuende.de/riedinge/Baltvib_RF_RFE

Competing interests

The authors declare no competing interests

Author contribution

M.L. initiated and supervised the project, M.L., L.R., A.F.A., D.P.R.H., C.P., T. B. H. R., M.R., G.G., D.J.R. and M.K. designed the project; M.L., D.J.R., V.F.J., D.P.R.H., K.P., S.K., A.S., G.G. and H.B.B. performed the sampling and conducted the laboratory work; L.F.D. and A.F.A. performed the bioinformatics processing; D.J.R., T.S. and C.H. analyzed the data and performed the machine learning; F.B. and A.A. carried out the ddPCR; D.S.B. analyzed the organic nutrients; S.D. performed the multiplex-PCR. The manuscript was written by D.J.R. and M.L. and reviewed by all authors.

Acknowledgments

This work resulted from the BiodivERsA project “Pathogenic *Vibrio* bacteria in the current and future Baltic Sea waters: mitigating the problem” (BaltVib), funded by the European Union and the Federal Ministry of Education and Research, Germany (grant 16LC2022A), the Innovation Fund Denmark (grant 0156-00001B), the Estonian Research Council (grant T210076PKKH / P200028PKKH), the Research Council of Lithuania (grant S-

BIODIVERSA-21-1), the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning FORMAS (grant 2020-02366), the Polish National Science Centre (grant 2020/02/Y/NZ8/00009) and the Academy of Finland (grant 344743). The authors acknowledge support from the National Genomics Infrastructure in Genomics Application Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, and SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure. We are grateful to the crew and captain of the R/V Elisabeth Mann Borgese (EMB283), the research divers of the Leibniz Institute for Baltic Sea Research, the University of Rostock and the Estonian University of Life Sciences, to Jonas Nilsson for assisting with sampling in Sweden, and to Adam Woźniczka and Jarone Pinhassi for providing lab space. Last but not least, the authors would like to thank Rita R. Colwell (University of Maryland, USA) for constructive criticism of the manuscript

References

1. Reissmann, J. H. *et al.* Vertical mixing in the Baltic Sea and consequences for eutrophication - A review. *Prog. Oceanogr.* **82**, 47–80 (2009).
2. Team, B. I. A. *et al. Second Assessment of Climate Change for the Baltic Sea Basin*. 291-292 (Springer, 2015).
3. Meier, H. E. M. *et al.* Climate change in the Baltic Sea region: A summary. *Earth Syst. Dyn.* **13**, 457–593 (2022).
4. Rutgersson, A. *et al.* Natural hazards and extreme events in the Baltic Sea region. *Earth Syst. Dyn.* **13**, 251–301 (2022).
5. Baker-Austin, C. *et al.* Emerging *Vibrio* risk at high latitudes in response to ocean warming. *Nat. Clim. Chang.* **3**, 73–77 (2013).
6. Baker-Austin, C., Trinanes, J., Gonzalez-Escalona, N. & Martinez-Urtaza, J. Non-cholera vibrios: the microbial barometer of climate change. *Trends Microbiol.* **25**, 76–84 (2017).
7. Frank, C., Littman, M., Alpers, K. & Hallauer, J. *Vibrio vulnificus* wound infections after contact with the Baltic Sea, Germany. *Wkly. releases* **11**, 3024 (2006).
8. Ruppert, J. *et al.* Two cases of severe sepsis due to *Vibrio vulnificus* wound infection acquired in the Baltic Sea. *Eur. J. Clin. Microbiol. Infect. Dis.* **23**, 912–915 (2004).
9. Brehm, T. T. *et al.* Non-cholera *Vibrio* species — currently still rare but growing danger of infection in the North Sea and the Baltic Sea. *Internist* **62**, 876–886 (2021).
10. Linkous, D. A. & Oliver, J. D. Pathogenesis of *Vibrio vulnificus*. *FEMS Microbiol. Lett.* **174**, 207–214 (1999).
11. Amato, E. *et al.* Epidemiological and microbiological investigation of a large increase in

vibriosis, northern Europe, 2018. *Eurosurveillance* **27**, 2101088 (2022).

12. Blackwell, K. D. & Oliver, J. D. The ecology of *Vibrio vulnificus*, *Vibrio cholerae*, and *Vibrio parahaemolyticus* in North Carolina Estuaries. *J. Microbiol.* **46**, 146–153 (2008).
13. Banakar, V. *et al.* Temporal and spatial variability in the distribution of *Vibrio vulnificus* in the Chesapeake Bay: a hindcast study. *Ecohealth* **8**, 456–467 (2011).
14. Johnson, C. N. *et al.* Ecology of *Vibrio parahaemolyticus* and *Vibrio vulnificus* in the coastal and estuarine waters of Louisiana, Maryland, Mississippi, and Washington (United States). *Appl. Environ. Microbiol.* **78**, 7249–7257 (2012).
15. Randa, M. A., Polz, M. F. & Lim, E. Effects of temperature and salinity on *Vibrio vulnificus* population dynamics as assessed by quantitative PCR. *Appl. Environ. Microbiol.* **70**, 5469–5476 (2004).
16. Eiler, A., Gonzalez-Rey, C., Allen, S. & Bertilsson, S. Growth response of *Vibrio cholerae* and other *Vibrio* spp. to cyanobacterial dissolved organic matter and temperature in brackish water. *FEMS Microbiol. Ecol.* **60**, 411–418 (2007).
17. Eiler, A., Johansson, M. & Bertilsson, S. Environmental influences on *Vibrio* populations in northern temperate and boreal coastal waters (Baltic and Skagerrak Seas). *Appl. Environ. Microbiol.* **72**, 6004–6011 (2006).
18. Rippey, S. R. Infectious diseases associated with molluscan shellfish consumption. *Clin. Microbiol. Rev.* **7**, 419–425 (1994).
19. Yun, N. R. & Kim, D. M. *Vibrio vulnificus* infection: A persistent threat to public health. *Korean Journal of Internal Medicine* **33** 1070–1078 (2018).
20. Baker-Austin, C., Stockley, L., Rangdale, R. & Martinez-Urtaza, J. Environmental occurrence and clinical impact of *Vibrio vulnificus* and *Vibrio parahaemolyticus*: a European perspective. *Environ. Microbiol. Rep.* **2**, 7–18 (2010).
21. Haenen, O. L. M. *et al. Vibrio vulnificus* outbreaks in Dutch eel farms since 1996: Strain diversity and impact. *Dis. Aquat. Organ.* **108**, 201–209 (2014).
22. Lamb, J. B. *et al.* Seagrass ecosystems reduce exposure to bacterial pathogens of humans, fishes, and invertebrates. *Science.* **355**, 731–733 (2017).
23. Reusch, T. B. H. *et al.* Lower *Vibrio* spp. abundances in *Zostera marina* leaf canopies suggest a novel ecosystem function for temperate seagrass beds. *Front. Mar. Sci.* **168,** 1-6 (2021).
24. Fonseca, M. S., Fisher, J. S., Zieman, J. C. & Thayer, G. W. Influence of the seagrass, *Zostera marina* L., on current flow. *Estuar. Coast. Shelf Sci.* **15**, 351–364 (1982).
25. Nowell, A. R. M. & Jumars, P. A. Flow environments of aquatic benthos. *Annu. Rev. Ecol. Syst.* **15**, 303–328 (1984).
26. Worcester, S. E. Effects of eelgrass beds on advection and turbulent mixing in low current and low shoot density environments. *Mar. Ecol. Prog. Ser.* **126**, 223–232 (1995).
27. Peterson, B. J. & Heck Jr, K. L. Positive interactions between suspension-feeding bivalves and

seagrass a facultative mutualism. *Mar. Ecol. Prog. Ser.* **213**, 143–155 (2001).

28. Gonzalez-Ortiz, V. *et al.* Interactions between seagrass complexity, hydrodynamic flow and biomixing alter food availability for associated filter-feeding organisms. *PLoS One* **9**, e104949 (2014).
29. Bodhaguru, M. *et al.* Screening, partial purification of antivibriosis metabolite sterol-glycosides from *Rhodococcus* sp. against aquaculture associated pathogens. *Microb. Pathog.* **134**, 103597 (2019).
30. Brumfield, K. D. *et al.* Environmental Factors Influencing Occurrence of *Vibrio parahaemolyticus* and *Vibrio vulnificus*. *Appl. Environ. Microbiol.* e00307-23 (2023).
31. Hsieh, J. L., Fries, J. S. & Noble, R. T. *Vibrio* and phytoplankton dynamics during the summer of 2004 in a eutrophying estuary. *Ecol. Appl.* **17**, S102–S109 (2007).
32. Brumfield, K. D. *et al.* Genomic diversity of *Vibrio* spp. and metagenomic analysis of pathogens in Florida Gulf coastal waters following Hurricane Ian. *MBio,* **14**, e01476-23 (2023).
33. Matz, C. *et al.* Biofilm formation and phenotypic variation enhance predation-driven persistence of *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 16819–16824 (2005).
34. Worden, A. Z. *et al.* Trophic regulation of *Vibrio cholerae* in coastal marine waters. *Environ. Microbiol.* **8**, 21–29 (2006).
35. Main, C. R., Salvitti, L. R., Whereat, E. B. & Coyne, K. J. Community-Level and species-specific associations between phytoplankton and particle-associated *Vibrio* species in delaware's inland bays. *Appl. Environ. Microbiol.* **81**, 5703–5713 (2015).
36. Matcher, G., Lemley, D. & Adams, J. Bacterial community dynamics during a harmful algal bloom of *Heterosigma akashiwo*. *Aquat. Microb. Ecol.* **86**, 153–167 (2021).
37. Greenfield, D. I. *et al.* Temporal and environmental factors driving *Vibrio Vulnificus* and *V. Parahaemolyticus* populations and their associations with harmful algal blooms in South Carolina detention ponds and receiving tidal creeks. *GeoHealth* **1**, 306–317 (2017).
38. Hugerth, L. W. *et al.* Metagenome-assembled genomes uncover a global brackish microbiome. *Genome Biol.* **16**, 1–18 (2015).
39. Bunse, C. *et al.* Spatio-temporal interdependence of bacteria and phytoplankton during a Baltic Sea spring bloom. *Front. Microbiol.* **7**, 517 (2016).
40. Wallentinus, I. Comparisons of nutrient uptake rates for Baltic macroalgae with different thallus morphologies. *Mar. Biol.* **80**, 215–225 (1984).
41. El-Hadary, M. H., Elsaied, H. E., Khalil, N. M. & Mikhail, S. K. Molecular taxonomical identification and phylogenetic relationships of some marine dominant algal species during red tide and harmful algal blooms along Egyptian coasts in the Alexandria region. *Environ. Sci. Pollut. Res.* **29**, 53403–53419 (2022).
42. Hamamoto, Y. & Honda, D. Nutritional intake of *aplanochytrium* (labyrinthulea, stramenopiles) from living diatoms revealed by culture experiments suggesting the new prey–predator

interactions in the grazing food web of the marine ecosystem. *PLoS One* **14**, 1–23 (2019).

43. Reñé, A. *et al.* The new chytridiomycete *Paradinomyces triforaminorum* gen. et sp. nov. co-occurs with other parasitoids during a *Kryptoperidinium foliaceum* (Dinophyceae) bloom in the Baltic Sea. *Harmful Algae* **120**, 102352 (2022).
44. Thomsen, H. A., Hällfors, G., Hällfors, S. & Ikävalko, J. New observations on the heterotrophic protists genus *Thaumatomastix* (Thaumatomastigaceae, Protista *incertae sedis*) with particular emphasis on material from the Baltic Sea. *Ann. Bot. Fenn.* **30**, 87–108 (1993).
45. Anschütz, A. A., Flynn, K. J. & Mitra, A. Acquired phototrophy and its implications for bloom dynamics of the *Teleaulax-Mesodinium-Dinophysis*-complex. *Front. Mar. Sci.* **8**, 1–18 (2022).
46. Altenburger, A. *et al.* Dimorphism in cryptophytes—The case of *Teleaulax amphioxeia*/*Plagioselmis prolonga* and its ecological implications. *Sci. Adv.* **6**, 1–9 (2020).
47. HELCOM. HELCOM Baltic Sea Action Plan (adopted by the HELCOM Ministerial meeting, Krakow, Poland 15th November 2007).
48. Ranft, S. *et al.* Eutrophication assessment of the Baltic Sea Protected Areas by available data and GIS technologies. *Mar. Pollut. Bull.* **63**, 209–214 (2011).
49. Voss, M. *et al.* History and scenarios of future development of Baltic Sea eutrophication. *Estuar. Coast. Shelf Sci.* **92**, 307–322 (2011).
50. Rydin, E., Kumblad, L., Wulff, F. & Larsson, P. Remediation of a eutrophic bay in the Baltic Sea. *Environ. Sci. Technol.* **51**, 4559–4566 (2017).
51. Semenza, J. C. *et al.* Environmental suitability of *Vibrio* infections in a warming climate: an early warning system. *Environ. Health Perspect.* **125**, 107004 (2017).
52. Klappenbach, J. A., Saxman, P. R., Cole, J. R. & Schmidt, T. M. rrndb: the ribosomal RNA operon copy number database. *Nucleic Acids Res.* **29**, 181–184 (2001).
53. Stoddard, S. F., Smith, B. J., Hein, R., Roller, B. R. K. & Schmidt, T. M. rrn DB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res.* **43**, D593–D598 (2015).
54. Louca, S., Doebeli, M. & Parfrey, L. W. Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome* **6**, 1–12 (2018).
55. Yildiz, F. H. & Visick, K. L. *Vibrio* biofilms: so much the same yet so different. *Trends in Microbiology* **17,** 109–118 (2009).
56. Joseph, L. A. & Wright, A. C. Expression of *Vibrio vulnificus* capsular polysaccharide inhibits biofilm formation. *J. Bacteriol.* **186**, 889–893 (2004).
57. Marco-Noales, E., Milán, M., Fouz, B., Sanjuán, E. & Amaro, C. Transmission to eels, portals of entry, and putative reservoirs of *Vibrio vulnificus* serovar E (Biotype 2). *Appl. Environ. Microbiol.* **67**, 4717–4725 (2001).
58. McDougald, D., Lin, W., Rice, S. & Kjelleberg, S. The role of quorum sensing and the effect of environmental conditions on biofilm formation by strains of *Vibrio vulnificus*. *Biofouling* **22**,

161–172 (2006).

59. Ugarelli, K., Chakrabarti, S., Laas, P. & Stingl, U. The seagrass holobiont and its microbiome. *Microorganisms* **5,** 1–28 (2017).
60. Guan, C. *et al.* Identification of rosmarinic acid and sulfated flavonoids as inhibitors of microfouling on the surface of eelgrass *Zostera marina*. *Biofouling* **33**, 867–880 (2017).
61. Cúcio, C., Engelen, A. H., Costa, R. & Muyzer, G. Rhizosphere microbiomes of European seagrasses are selected by the plant, but are not species specific. *Front. Microbiol.* **7**, 1–15 (2016).
62. Möller, L., Kreikemeyer, B., Luo, Z.-H., Jost, G. & Labrenz, M. Impact of coastal aquaculture operation systems in Hainan island (China) on the relative abundance and community structure of *Vibrio* in adjacent coastal systems. *Estuar. Coast. Shelf Sci.* **233**, 106542 (2020).
63. Webb, S. J., Rabsatt, T., Erazo, N. & Bowman, J. S. Impacts of *Zostera* eelgrasses on microbial community structure in San Diego coastal waters. *Elem Sci Anth*. **7**, 11 (2019).
64. Grasshoff, K., Kremling, K., Ehrhardt, M. (eds.) , Methods of Seawater Analysis - 3rd edition. *Wiley-VCH* , 159-228 (1999)
65. Lysiak-Pastuszak, E. & Krysell, M. Chemical measurements in the Baltic Sea: guidelines on quality assurance. **35,** 146 - 149 (2004).
66. HELCOM. Guidelines for monitoring of chlorophyll a. (2019).
67. Kirchner, S. *et al.* Pentaplexed quantitative real-time PCR assay for the simultaneous detection and quantification of botulinum neurotoxin-producing clostridia in food and clinical samples. *Appl. Environ. Microbiol.* **76**, 4387–4395 (2010).
68. Gasol, J. M. & Del Giorgio, P. A. Using flow cytometry for counting natural planktonic bacteria and understanding the structure of planktonic bacterial communities. *Sci. Mar.* **64**, 197–224 (2000).
69. Lo, K., Hahne, F., Brinkman, R. R. & Gottardo, R. flowClust: a Bioconductor package for automated gating of flow cytometry data. *BMC Bioinformatics* **10**, 1–8 (2009).
70. Monaco, G. *et al.* flowAI: automatic and interactive anomaly discerning tools for flow cytometry data. *Bioinformatics* **32**, 2473–2480 (2016).
71. Herlemann, D. P. R. *et al.* Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J.* **5**, 1571–1579 (2011).
72. Balzano, S., Abs, E. & Leterme, S. C. Protist diversity along a salinity gradient in a coastal lagoon. *Aquat. Microb. Ecol.* **74**, 263–277 (2015).
73. Latz, M. A. C. *et al.* Short-and long-read metabarcoding of the eukaryotic rRNA operon: evaluation of primers and comparison to shotgun metagenomics sequencing. *Mol. Ecol. Resour.* **22**, 2304–2318 (2022).
74. Fadrosh, D. W. *et al.* An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* **2**, 1–7 (2014).

75. Glenn, T. C. *et al.* Adapterama I: universal stubs and primers for 384 unique dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru & iNext). *PeerJ* **7**, e7755 (2019).
76. Köster, J. & Rahmann, S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520–2522 (2012).
77. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011).
78. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
79. Guillou, L. *et al.* The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**, D597–D604 (2012).
80. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
81. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
82. O'Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745 (2016).
83. Brehm, T. T. *et al.* Heatwave-associated *Vibrio* infections in Germany, 2018 and 2019. *Eurosurveillance* **26**, 2002041 (2021).
84. Panicker, G., Vickery, M. C. L. & Bej, A. K. Multiplex PCR detection of clinical and environmental strains of *Vibrio vulnificus* in shellfish. *Can. J. Microbiol.* **50**, 911–922 (2004).
85. Kassambara, A. rstatix: Pipe-friendly framework for basic statistical tests. (2020).
86. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **28**, 1–26 (2008).
87. Breiman, L. E. O. Random Forests. *Machine learning,* **45** 5–32 (2001).
88. Gregorutti, B., Michel, B. & Saint-Pierre, P. Correlation and variable importance in random forests. *Stat. Comput.* **27**, 659–678 (2017).
89. Nixon, S. W. Coastal marine eutrophication: A definition, social causes, and future concerns. *Ophelia* **41**, 199–219 (1995).
90. Lê, S., Josse, J. & Husson, F. FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* **25**, 1–18 (2008).
91. Labrenz M. et al 16S and 18S rRNA amplicon sequencing of coastal microbial communities across the salinity gradient of the Baltic Sea in pelagic, benthic, and biofilm environments https://www.ebi.ac.uk/ena/data/view/PRJEB68222 (2024)
92. Yilmaz, P. *et al.* Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat. Biotechnol.* **29**, 415–420 (2011).

93. Diepenbroek, M. *et al.* Towards an integrated biodiversity and ecological research data management and archiving platform: the German federation for the curation of biological data (GFBio). *Inform.* **232**, 1711-1724 (2014).

94. Labrenz M. et al. The occurrence of Vibrio spp. in the salinity gradient of shallow coastal waters of the Baltic Sea – data set including environmental and microbiological data (doi.io-warnemuende.de/10.12754/data-2023-0010) (2023)

95. Delgado, L. F., Labrenz, M., & Andersson, A. F. (2024). *Vibrio* 16S rRNA gene sequences (v1.0). Zenodo. https://doi.org/10.5281/zenodo.1087510

General Discussion

Potentially human pathogenic *V. vulnificus* are common inhabitants of estuaries and marginal seas and are likely spreading poleward[55]. Within this thesis, we used archived sequencing datasets to determine the global distribution of *V. vulnificus* in the environment and the poleward spread over the last decade. We developed a random forest model for *V. vulnificus* environmental suitability prediction based on globally freely available satellite data and one based on the prokaryotic community. We established an economical and fast two-plate method for the validation of model results in the environment and explored two different possible mitigation measures for *V. vulnificus* for coastal risk zones. With the identified tentative poleward spread of *V. vulnificus* occurrence, earlier findings based on infection rates and *V. vulnificus* suitability modelling are substantiated. The models within this thesis illustrate a likely connection between decaying phytoplankton and *V. vulnificus* abundance, the effect of which might be particularly pronounced at low current speeds. Consequently, the expected increase of *V. vulnificus* due to increasing temperatures will possibly be exasperated by a concomitant increase in phytoplankton blooms. Out of the tested mitigation measures, the reduction of eutrophication has a higher potential to limit *V. vulnificus* proliferation than seagrass meadows, by decreasing length and size of phytoplankton blooms, thereby reducing infection risks for coastal populations.

V. vulnificus is nearly cosmopolitan and can be approximated from remote sensing data

The climate change driven poleward spread of *V. vulnificus*, already detected in infection cases and within this thesis, found in the environment on a global scale will lead to increased public health risks, economic burden and a potential disruption in fisheries. Regions that currently experience relatively short risk periods, such as Delaware Bay or the south-western Baltic Sea, will see longer and more intense growing seasons for *V. vulnificus*, along with the associated problems. Increasing risks at previously largely unaffected regions, such as the north-eastern Baltic Sea, might necessitate the implementation of new regulatory measures, affecting both peoples freedom and the economic stability of coastal communities. Thus, modeling the environmental suitability for *V. vulnificus* in such regions becomes increasingly important.

With the developed RF model, which relies on satellite data, we are able to explain approximately 50 % of the variance in the relative abundance when *V. vulnificus* is present by accounting for complex relationships between environmental conditions. This is markedly different from the generalized linear models currently in use for the prediction of *Vibrio* infection risk[55]. The current models rely on salinity and temperature for the risk assessment while our models show that chla, organic carbon and nitrogen, phosphate, *Cyanobium* and currents and sediment resuspension possibly modulate the (relative) abundance of *V. vulnificus* The trained models only predict (relative) abundance and not the pathogenicity of the *V. vulnificus* population. A key assumption made is that the increase in total (relative) *V. vulnificus* abundance positively correlates with an increase in C-genotype abundance and thus infection risk. While this correlation has been clearly demonstrated along the southeastern North Carolina coast[110], they do note that C-genotypes seem to favor higher salinity than E-genotypes Additionally, total suspended matter correlated more strongly with E-genotypes than with C-genotypes, likely due to a higher propensity of these strains to attach to marine aggregates and chitin. While they conclude that the modelling abundance of total *V. vulnificus* is valuable to estimate C-genotype occurrence, the specific niche created by increased eutrophication (Found in Chapter I, III & IV) might stimulate E-genotypes more than C-genotypes. This might especially be the case in the eastern Baltic Sea, which already favors E-genotypes due to the lower salinity. Additional validation efforts in both low and high saline areas of the Baltic Sea, of the *V. vulnificus* abundance, the ratio between C- and E-genotypes, and comparisons with infection cases with a clear origin, would enhance confidence in using the newly trained model. Nonetheless, the similar geographic patterns in historical infections and predicted suitability in Chapter I indicate the model can be used for risk assessment.

Machine learning indicates decaying blooms support *V. vulnificus* growth

Increased eutrophication leads to a rise in phytoplankton and associated zooplankton populations. This might stimulate *V. vulnificus* abundance by increasing the availability of labile organic matter, providing protection from bacteriovores through attachment, and increase available chitin, a key polymer for *V. vulnificus* growth. The microorganisms found to have high explanatory power at both the global (Chapter I) and Baltic Sea scale (Chapter IV), indicate that decaying blooms might form a niche for rapid proliferation of *V. vulnificus*. It is relevant to note that in both chapters, conditions in the environment where the samples were taken were generally conducive to *V. vulnificus*. In chapter IV, this was assured by limiting the temporal and spatial extent of the sampling to the summer in the Baltic Sea, a region

characterized by brackish conditions and water temperatures around 20 °C during the sampling period. In Chapter I, only samples where *V. vulnificus* was detected were used for modeling purposes. Due to these decisions, other factors modulating (relative) *V. vulnificus* abundance, can become apparent. If a model were trained using data from many sites where *V. vulnificus* does not exist due to temperature or salinity limitations, even though all other conditions at these sites are optimal for the bacterium, the perceived effect of these favorable conditions might be diluted. Within suitable environmental conditions for *V. vulnificus*, between approximately 45 and 60 % of the variance in (relative) abundance could be explained, either from a combination of environmental and microbial (eukaryotic and prokaryotic) on a smaller Baltic Sea dataset, or solely from the prokaryotic community or environmental data from satellites on a global level.

Both the model based solely environmental data on a global scale and the combined model in Chapter IV show that chla is an important predictor of *V. vulnificus*. Interestingly, in Figure 1c, Chapter I, no obvious relationship between chla and the relative abundance of *V. vulnificus* becomes apparent. By also including the microbial community as predictors, either separately (Chapter I) or combined (Chapter IV), we were able to deduce a likely scenario that explains this. In Chapter IV, *Cyanobium* sp., a cyanobacteria that is known the bloom in the summer in the Baltic Sea, was detected as the best microbial predictor of *V. vulnificus*. If a large bloom would lead to high (relative) *V. vulnificus* concentrations directly, we would expect a clear relationship between chla and *V. vulnificus*, which is not the case. The other predictors found to be important, were species that are expected to become abundant during a decaying bloom, such as parasites (*Aplanochytrium* sp.) and fungi (*Rhizophydium* sp.). Medium chla concentrations can signify higher (relative) abundance of *V. vulnificus* when resulting from decaying blooms; however, they can also originate from regular primary production. Consequently, this connection is not evident in the plot, but the model in Chapter IV includes sufficient parameters, such as microorganisms associated with decaying blooms, to identify the nature of the medium chla concentration, increasing its importance as a predictor. That a decaying bloom might stimulate *V. vulnificus* growth is further substantiated by the high explanatory power of DOC and DN, which are expected to peak during the decay of the bloom, due to the release of organic matter from the decomposing phytoplankton cells. This hypothesis is supported by the findings in Chapter I, where the prediction of relative *V. vulnificus* abundance from the microbial community shows early colonizers of such a system

as some of the best predictors and by Chapter III, where *Cyanobium* had a high relative abundance in samples where *V. vulnificus* was present.

While the feature importance of these models allows us to hypothesize about the possible mechanisms that modulate (relative) *V. vulnificus* abundance, the environmental data that is used is noisy and covariant within itself. This noise can originate from measurements errors, introduced due to limitations of instrument precision and sampling inconsistencies, from natural variability, such as diurnal cycles, or environmental conditions that were not measured, and limitations inherent to methodologies such as amplicon sequencing. This noise can decrease the performance of our models and introduce artefacts into the model that skew feature importance.

The covariance between features, such as terrestrial inputs and salinity within a system like the Baltic Sea leads to limitations when evaluating feature importance. Nutrient concentrations are for example higher on the German and Polish coasts than in the Finish Archipelago Sea, as is salinity. Consequently, for the model, both eutrophication and salinity present similar splits in a random forest to increase the homogeneity in the next node of a tree and both parameters have similar importance. For the target variable, in our case *V. vulnificus*, one might be biologically relevant while the other is not. This can be partly deciphered by using RFE, limiting the tested range of environmental conditions to those suitable for *V. vulnificus,* and comparing model outcomes with more specific laboratory studies. The salinity levels in all samples used in both Chapter I and Chapter IV are suitable and mostly within the optimum range for *V. vulnificus*. The variable importance of POC, PN, DOC, DN, and cyanobacteria is higher than that of salinity in Chapter IV. Additionally, laboratory studies have demonstrated the impact of cyanobacterial-derived organic matter on the growth of *Vibrio* spp.[58]. Therefore, it can be concluded that there is likely a biological effect of these parameters. This is substantiated by the higher predictive power of chla in Chapter IV compared to salinity, meaning that as long as salinity is suitable for *V. vulnificus*, differences in salinity matter less to relative *V. vulnificus* abundance than chla.

The detection which (combination) of these factors influence (relative) *V. vulnificus* abundance is complicated due to the correlation between them and with a large number of microbes. Samples with high *V. vulnificus* generally have high POC, PON, DOC, DN and specific groups of microbes. *V. vulnificus* might benefit from one, a combination of (some) of them, or even something else that co-varies with all of these variables. Despite the covariance

between these features, they were all retained by RFE, indicating that the importance is possibly very high, and are all stimulated by eutrophication of coastal waters. To establish a causal relationship and discern the relative impacts of the organic nutrients, additional experiments are a necessity.

Culture-based identification of *V. vulnificus* allows for rapid testing

A warning about *V. vulnificus* at a beach can have wide-ranging consequences, for tourists, hotels, and local restaurants. It is therefore pertinent to refrain from relying solely on modeling to predict risk but to validate the model prediction *in situ*. To this end, we developed a combined TCBS and Chromagar culture-based methodology, with a focus on *V. vulnificus*, for the rapid and inexpensive identification of different potentially pathogenic *Vibrio* in both the sediment and the water column. Identification with a single agar has been shown not to be accurate for many *Vibrio*[111] and, while accurate (92.8% correct identification), established three plate methods are work and time intensive[112]. The method described in Chapter II retained some of the benefits of the one-plate method and the three-plate method, having an average accurate identification rate of 85% (93% Baltic Sea wide and 75% on the Warnemünde coast), while reducing the workload by 1/3 compared to the three-plate method. A remaining problem is the delay between sampling and results, which is a minimum of 48 h with the proposed methodology, consequently, the verification of the model prediction can, at the earliest, be made 2 days after the conditions are suitable for *V. vulnificus*. While not explicitly tested, a partial remedy might be the monitoring of the sediment. As shown in Chapter III, there is likely mobilization of *V. vulnificus* from the sediment, either through tourist or wave action in the coastal environment. Concentrations of the pathogen might increase in the sediment before swimmers are in danger. By combining model predictions, presumptive *V. vulnificus* colonies in the water column and in the sediment, a risk profile can be established and a data-driven decision on a potential warning against water recreation, fishing or seafoods can be made. This limits the economic impacts of warning too frequently and too early while allowing for the adequate protection of public health. As global coastal populations are at increasing risks of *V. vulnificus*, the low costs of this methodology are a vital part. While not as accurate as molecular screening by for example MALDI-TOF or multiplex-PCR, monitoring can be reasonably established. Combined with the model, based on freely available data, proposed in Chapter I, *V. vulnificus* infections risks can be limited at global coasts.

Reducing eutrophication limits *V. vulnificus* proliferation

Increasing *V. vulnificus* infections because of climate change necessitate mitigation strategies. This need is especially large at sites where model predictions (Chapter I) show that the *V. vulnificus* suitability is high under current conditions or will be high following IPCC[57] predictions, such as the Baltic Sea. Under the sampled conditions, representative for large areas of the Baltic Sea under summer conditions, effects of seagrass on both sediment and aquatic (relative) abundance was absent. Consequently, seagrass renaturation is unlikely to be an adequate mitigation strategy under the conditions tested. This finding might not hold under different conditions, such as higher salinity, as tested by Reusch et al., 2021[60] or areas with lower eutrophication. The lower salinities in the eastern Baltic Sea may potentially have negatively affected the physiology and length of the seagrass[113] and resulted in a different epiphytic microbial community than at higher salinities. The shorter seagrass has a reduced influence on current speeds, meaning it removes fewer particles, including those carrying *V. vulnificus*. Lower salinity levels might also lead to a reduced release of allopathic chemicals, reducing the impact of the seagrass on the microbial community. Alternatively, the low water temperatures, ranging between 15 and 18.5 °C in Reusch et al.,[60], might mean the *Vibrio* community sampled had a very low abundance of *V. vulnificus*, consequently, the measured reduction concerned other members of the genus *Vibrio.*

Previously found reduction of pathogens through seagrass might be partly the consequence of local geography. The 50% reduction in the relative abundance of potential bacterial pathogens by seagrass found by Lamb et al., 2017[59], might be due to the islands of the Spermonde Archipelago, Indonesia, acting as a point-source of pathogens due to lacking sanitation. Due the reduction of current speeds, the seagrass filtered out particles and the associated pathogens, decreasing the concentrations in the water column. *V. vulnificus* in the Baltic Sea does not have a point-source and thus similar mechanisms are less effective.

The lack of a difference between sites within and outside of seagrass meadows could, in theory, be due to a very strong effect of the seagrass meadows. Given that the control sites were only approximately 100 meters away, the reduction in potential pathogens might be so significant and widespread that the control sites were still influenced by the proximity to the seagrass meadows. This is however unlikely due to the occurrence of sampling sites within the seagrass meadows that had higher abundances of *V. vulnificus* then the sites outside. If the differences in reduction of presumptive *V. vulnificus* colonies between the western Baltic Sea

by Reusch et al.,[60] and our findings based on the whole Baltic Sea are due to differences in the ecosystem services generated by the seagrass, this might have further implication for seagrass restoration as an environmental compensation measure. The efficacy for goals of the restoration, such as carbon storage, fish nurseries, pathogen reduction might differ at different salinities and different geographies.

Regional solutions, through the reduction of eutrophication, seem more suitable as the ML models in both Chapter I and IV and the stepwise multiple regression model from Chapter III all indicate eutrophication positively influences *V. vulnificus* abundance. While eutrophication is known to casue issues such as the formation of deadzones[114], harmful algal blooms[115], biodiversity loss[114], reduced spawning success rate of commercially relevant fish[116]., these findings reveal an even stronger need for the reduction of nutrient inflows into coastal regions, not just in the Baltic Sea but also in marginal seas worldwide. Progress has already been made regarding the eutrophication reduction in the Baltic Sea, with annual waterborne nitrogen and phosphorus inputs being 21% and 41% lower respectively in 2019 compared to the average of 1995 – 2019[117]. Considerable progress has been made in reducing the nutrient influx from point sources, such as wastewater treatment plants, industry and fish farms. Further reduction of influx from diffuse sources – such as agricultural practices, forestry, dwellings and natural erosion, is difficult to achieve as tracing relative contribution of specific emitters is particularly difficult[118]. While the reduction of inorganic nutrient input represents a considerable improvement, the changes in phosphate availability are slow. This is partly due to the release of phosphate from iron–humic complexes under anoxic conditions from the Baltic Sea sediments[119], long water[120] and nutrient residence time[121]. Thus, while continued efforts to reduce Baltic Sea eutrophication are vital, they are unlikely to rapidly affect the *V. vulnificus* infection risk. Reducing eutrophication in specific coastal environments that rely strongly on tourism or shellfish production might be a more rapid solution. Remediation of eutrophic bays in the Baltic Sea has been shown to be feasible by injecting dissolved aluminum into anoxic sediment[122]. This inhibits phosphate recycling from the sediment and thus reduces phytoplankton biomass on a local scale. This might be a solution for popular recreational beaches before a generally healthier state of the Baltic Sea is achieved. These findings might also have applications on *V. vulnificus* hotspots such as Chesapeake Bay[123], where efforts to reduce eutrophication have been ongoing, with mixed succes[124]. The possible reduction of a highly pathogenic bacterium, such as *V. vulnificus*, might provide an additional impetus to rethink land use in the watershed and reduce eutrophication in coastal waters worldwide.

Observational studies cannot identify causal relationships

Temporal or spatial observations along environmental gradients can validate laboratory findings in an environmental setting or be interpreted based on a theoretical framework, identifying a causal connection based on field observation solely remains difficult. When observing the environment, we aim to affect it as little as possible, meaning we only identify association and correlations and cannot manipulate one variable to measure a possible effect on a target variable. There is a range of confounding and extraneous variables over which we have no control influencing our identified associations. We cannot conclusively state, from the performed work, that eutrophication is a "driver" of *V. vulnificus* proliferation. Our observational study does however include the complex nature of environmental communities in a natural setting. The complexity of the coastal microbial community and environment, as well as its scale, cannot be effectively replicated in the lab. If we would rely solely on laboratory studies, where conditions are controlled, we may not be working on these communities in a relevant sense. Some of the conditions that might affect *V. vulnificus* abundance in the Baltic Sea, or indeed on a global scale, are large scale currents, highly heterogeneous sediment communities, complex food web interactions, weather action, nutrient cycling, diverse macrophyte communities and associated epifauna and the spatial scale of the study site itself. *V. vulnificus* might react differently to a simulated phytoplankton bloom under laboratory conditions, when the ecological context is missing. However, due to the similar conclusions of the experimental work of Eiler et al., 2007[58] and our fieldstudy, performed on the appropriate field scale, it becomes very likely that such a connection exists.

Effectively mitigating the risk in coastal zones

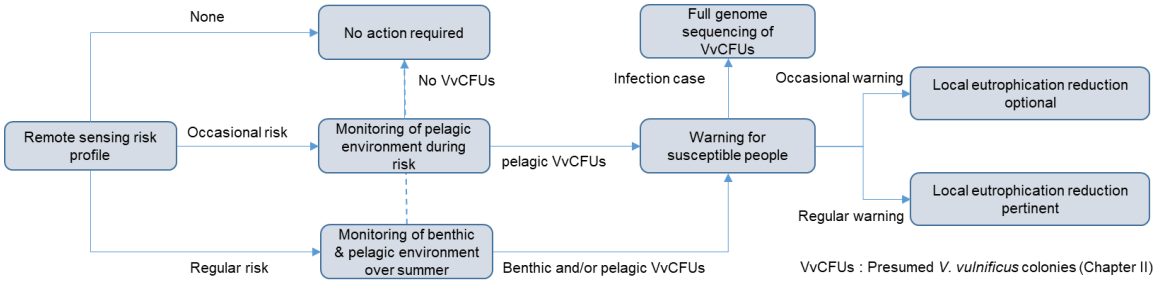


Figure 8. Possible decision tree for stakeholders. *An initial risk assessment can be made using remote sensing data, based on this assessment, proportional action can be taken.*

Figure 8 presents a decision tree designed to help stakeholders and local governments make effective decisions to mitigate public health risks while minimizing economic consequences. For occasional risk, monitoring the pelagic environment during high-risk periods, as identified by the model, may provide adequate public health protection. However, regions with extended high-risk periods will benefit from continuous monitoring of both the pelagic environment and sediment to ensure timely warnings. The minimum amount of presumed *V. vulnificus* CFUs in either the sediment or the water column to warrant a warning remains a political decision. At one extreme, issuing warnings from the first detected presumed *V. vulnificus* CFU would lead to very frequent alerts, larger economic consequences, and potentially trivializing the warnings, although it may be very helpful for the susceptible population. At the other extreme, issuing warnings only at extremely high colony counts could limit initial economic consequences but may result in inadequate protection for vulnerable individuals, more frequent infections cases and result in economic damage through media-coverage. A monitoring program will lead to a *V. vulnificus* culture collection, which can be used when an infection case is detected. The environmental cultures can be full genome sequenced thus enabling comparisons between strains from the environment and the isolated clinical strain, facilitating the identification of additional pathogenicity markers in *V. vulnificus* strains present in the environment. Additionally, the ratio between E-genotype and C-genotype colonies can be used to further fine-tune the proposed models and identify environmental conditions that lead to higher ratios of C to E-genotypes.

When warnings are infrequent, coastal communities may choose to reduce eutrophication by mitigating the impact of existing sources or implementing targeted interventions. With regular warnings, making these measures mandatory should be considered.

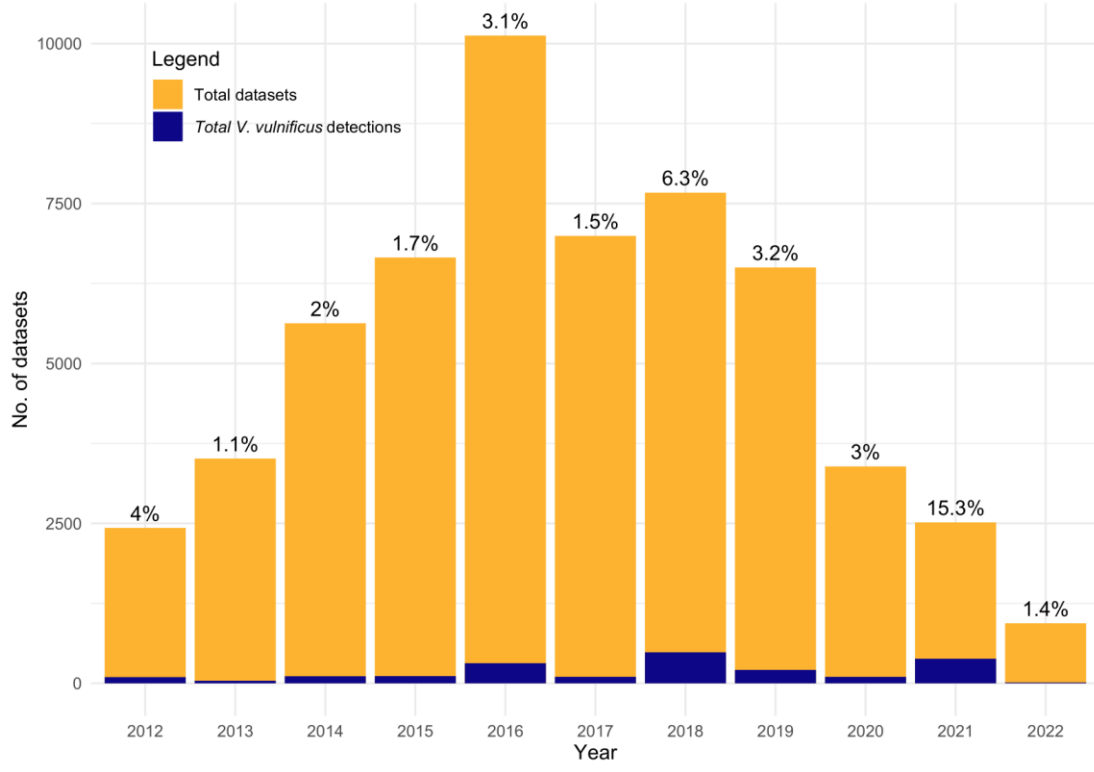
Outlook

While the reduction of eutrophication is consistently found to be a likely part of effective mitigation of *V. vulnificus* infection risk, these findings are all based on environmental observations covering many, co-varying and large environmental gradients with diverse communities. Additional studies conducted in environments where eutrophication is reduced and *V. vulnificus* is monitored are crucial. Ideally, this would involve comparative studies in relatively similar bays within the Baltic Sea, where eutrophication is reduced through engineering interventions in some areas but not in others. This approach would allow for testing the effectiveness of eutrophication reduction in mitigating *V. vulnificus* at an environmentally relevant scale.

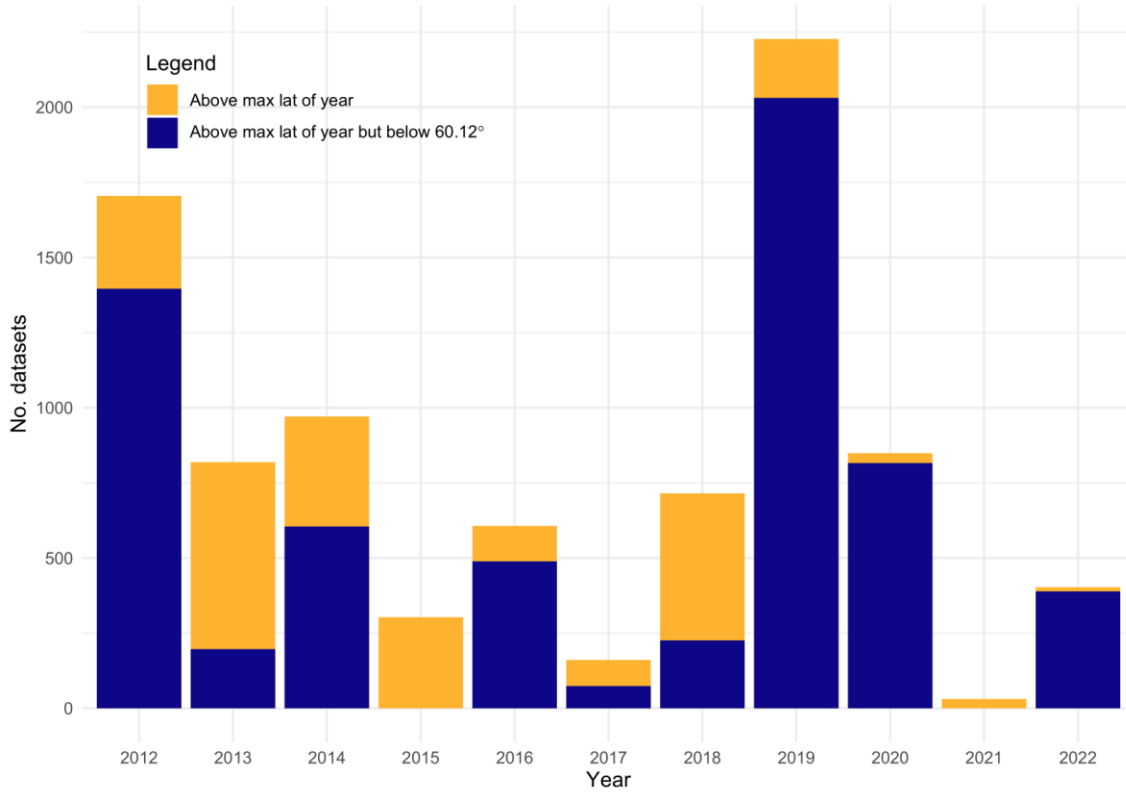
With regards to ML in environmental microbiology, the potential has been illustrated by earlier studies and again demonstrated within this thesis. A main concern is the samples sizes, which are often so small that there is a real risk of overfitting and a lower probability of detecting true effects. A solution to this is datamining, which substantially increases the samples available for model training and validation, leading to more generalizeable results, less affected by data artefacts and sampling campaign pecularities. Combining these extensive sequencing datasets with machine learning, in my view, will enhance our comprehension and forecasting of environmental reactions to changing circumstances, particularly in this instance for pathogen prediction but with broader applicability.

Appendices

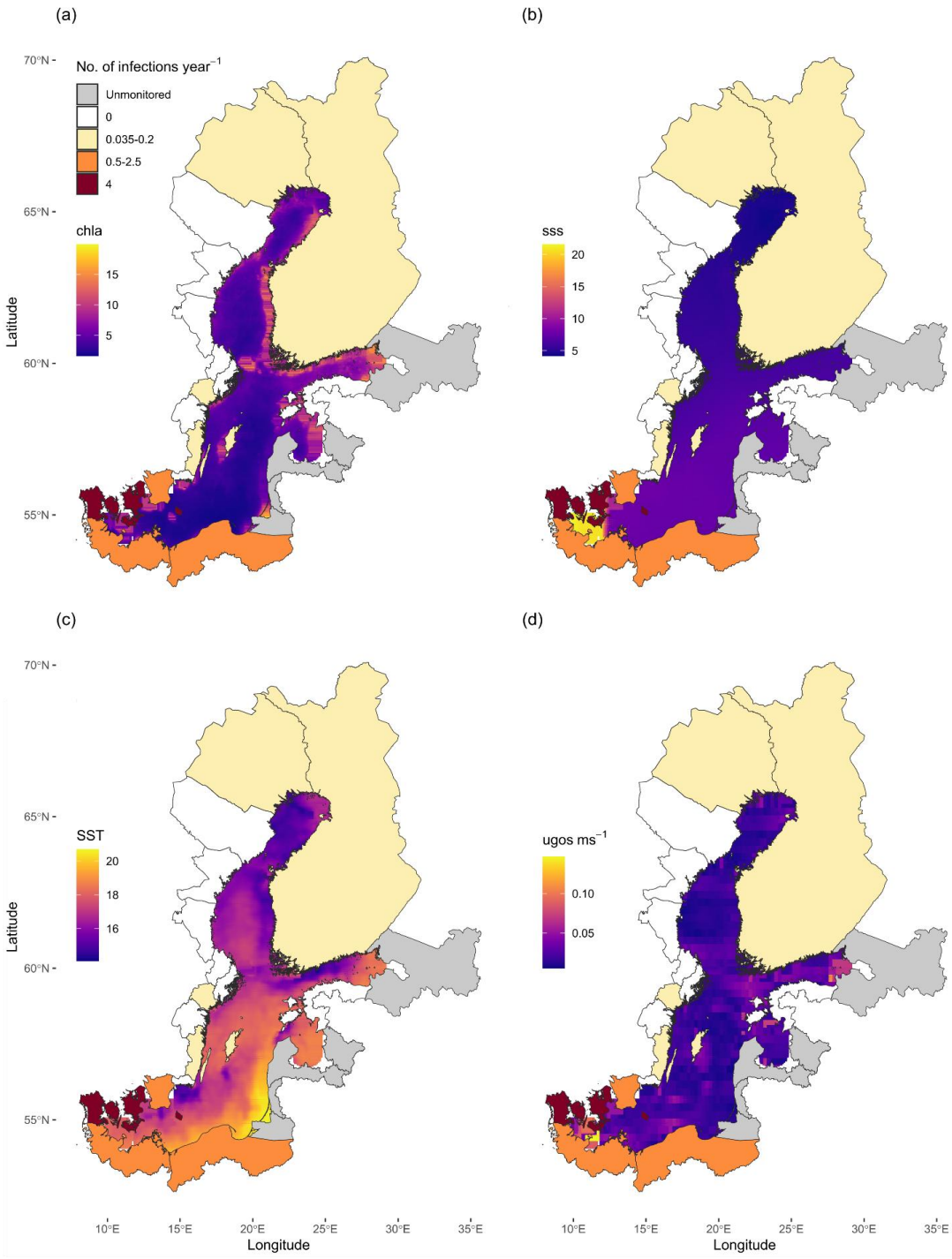
Chapter I



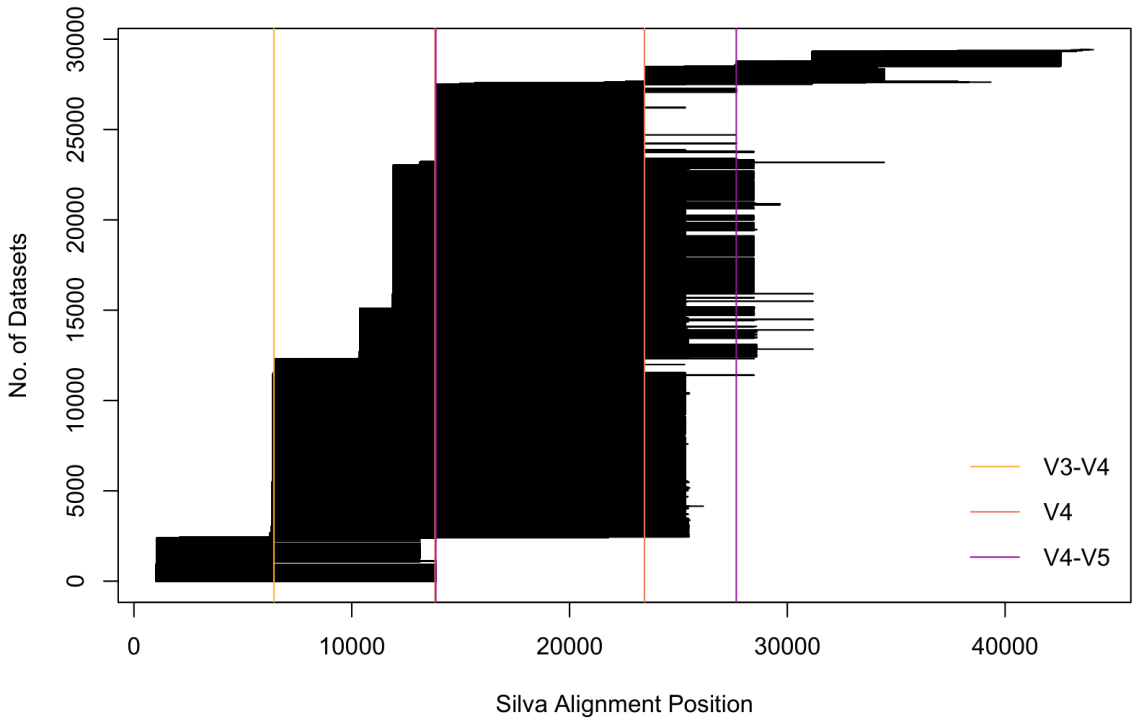
Supplementary figure 1. Datasets analyzed and datasets where V. vulnificus was detected per year. *Percentage positive samples per year is given*



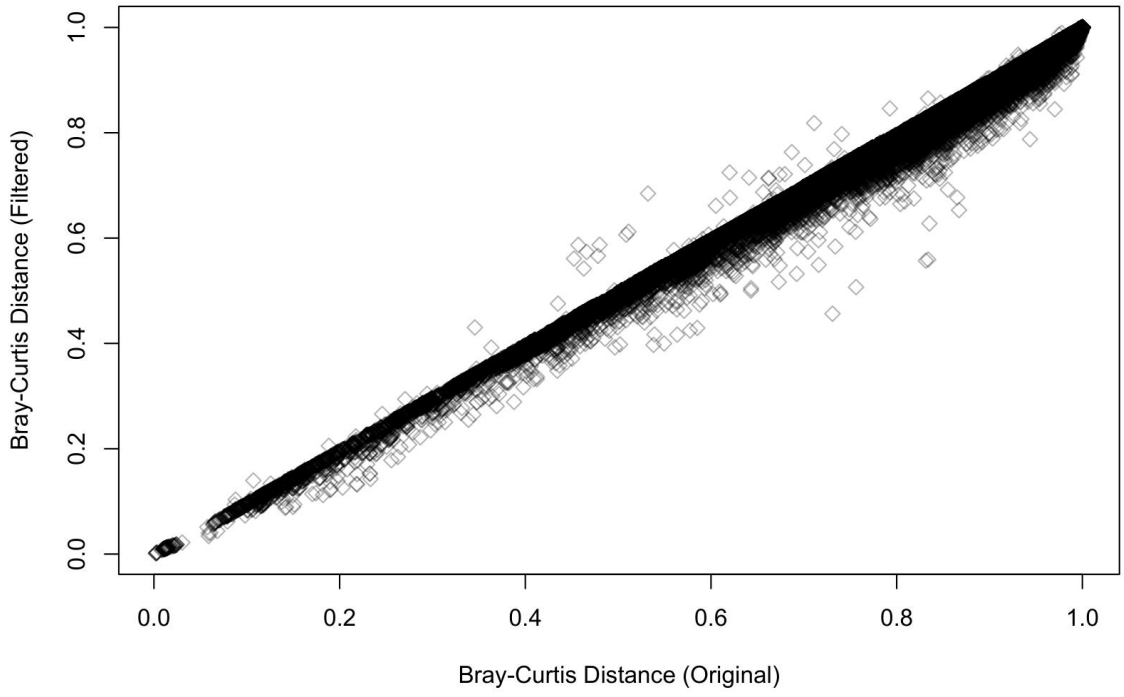
Supplementary figure 2. Datasets without* V. vulnificus *detection that are positioned further north than the most northern* V. vulnificus *detection each year, distinguishing between those samples located above and below the latitude 60.12° N (Most northern* V. vulnificus *detection within the complete data collection). *The figure shows that in all years there are datasets without* V. vulnificus *detection further north than the yearly* V. vulnificus *extent as well as the most northern* V. vulnificus *extent.*



Supplementary figure 3. Predictive features averaged over August 2021. *(a) Chl-a data (µg L^{-1}), (b) Sea surface salinity, (c) Sea surface temperature (C°), (d) Absolute east-west geostrophic sea water velocity (ugos ms^{-1}). Number of* V. vulnificus *infections normalized per year per region*



Supplementary figure 4. Common start and end positions of the datasets using known primer start and end positions. *Sequences were cut to retain maximum possible length while creating three distinct clusters*



Supplementary figure 5. Bray-Curtis dissimilarity between the original prokaryotic community and abundance filtered community

Supplementary Table 1 Performance of both RF models, evaluated by cross-validation and a hold-out test set

Model	**Mtry / Trees**	**Evaluation**	**RMSE**	**R^2**	**MAE**
Remote sensing	5	10-fold CV	0.52	0.48	0.39
	1000	Test/Train	0.54	0.47	0.39
Microbiome	20	10-fold CV	0.39	0.68	0.30
	1000	Test/Train	0.41	0.64	0.30

Chapter II

Supplementary excel files available under **https://www.mdpi.com/article/10.3390/microorganisms12030614/s1**

Temporal study: All figures and graphs contained within the Supplementary Information contain data from the Temporal study. Information on sampling sites, physical parameters, and CFU identification can be found in Supplementary File S1 and Supplementary File S2.

Spatial study: Information on sampling sites, physical parameters, and CFU identification can be found at 10.12754/data-2023-0010. Additionally Supplementary File S3 documents the station ID, longitude, latitude, date, sampling depth, temperature, salinity of the corresponding isolates.

Table S1. Colonies per* Vibrio *Species from each location of isolates correctly identified on TCBSand CHROMagar and confirmed using molecular analyses.

Campaign	*Vibrio* species	Correctly Assigned	Heiligendamm	Börgerende	Nienhagen	Warnemünde	Total
Temporal	*V. alginolyticus*	Yes	2	5	4	4	**15**
		No	2	8	9	16	**35**
	V. cholerae	Yes	0	0	1	6	**7**
		No	4	49	46	27	**126**
	V.parahaemolyticus	Yes	1	11	19	14	**45**
		No	0	1	3	2	**6**
	V. vulnificus	Yes	4	44	40	21	**109**
		No	0	9	10	8	**27**
Spatial	*V. vulnificus*	Yes	-				**80**
		No					**6**

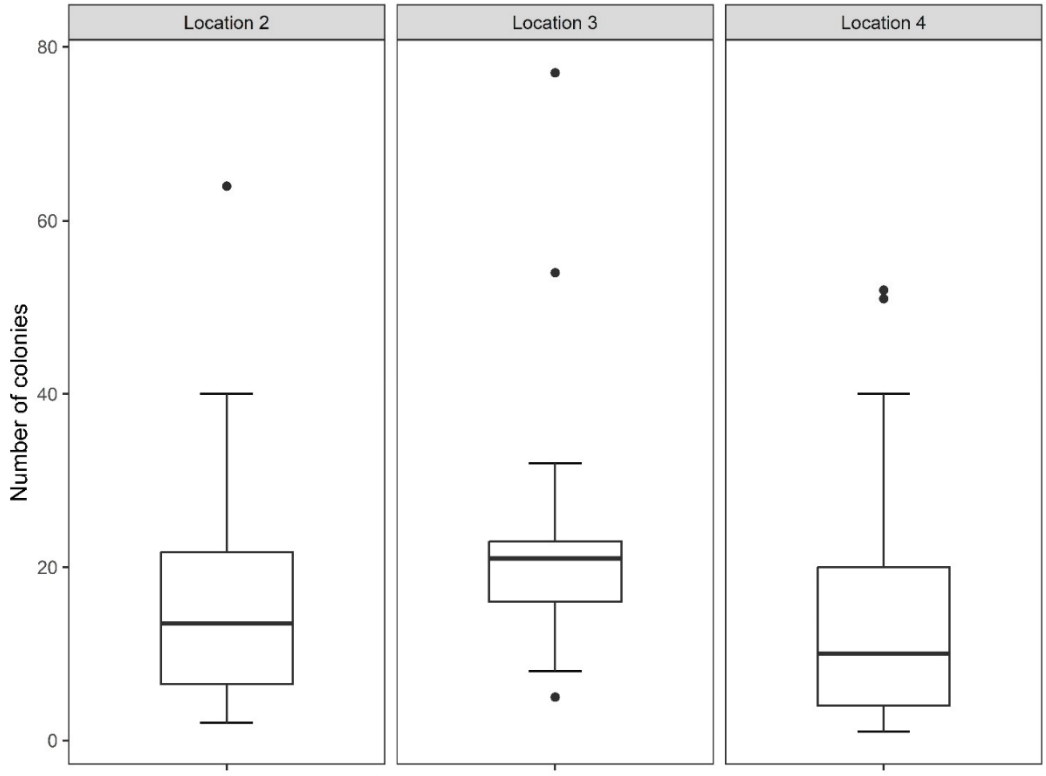


Figure S1. Data are represented as boxplots. *The middle line is the median, the lower and upper hinges correspond to the first and third quartiles, the upper whisker extends from the hinge to the largest value no further than 1.5× IQR from the hinge (where IQR is the inter-quartile range) and the lower whisker extends from the hinge to the smallest value at most 1.5× IQR of the hinge. Please see function geom_boxplot in R (ggplot2).*

Chapter III

Table S1. Metadata for water samples collected from May to October 2022. *The values presented are the averages obtained from triplicate samples. SST: sea surface temperature.*

Station	Substation	Type	Latitute	Longitude	Month	Date	Country	Salinity	SST (°C)	NH_4^+ (µM)	Heterotrophic cells mL^{-1}	NO_3^- (µM)	PO_4^{3-} (µM)
BV01-1	A	water	56.058401	12.577956	May	09/05/2022	Den	12.7	11.8	0.3	4.12E+05	0.09	0.06
BV01-1	B	water	56.058692	12.578673	May	09/05/2022	Den	12.7	11.8	0.2	3.98E+05	0.12	0.08
BV01-1	C	water	56.058401	12.580089	May	09/05/2022	Den	12.8	11.8	0.2	3.78E+05	0.12	0.08
BV01-2	A	water	56.058401	12.577956	May	23/05/2022	Den	14.6	13.1	0.8	4.76E+05	0.08	0.06
BV01-2	B	water	56.058692	12.578673	May	23/05/2022	Den	14.6	13.2	0.3	6.23E+05	0.10	0.06
BV01-2	C	water	56.058401	12.580089	May	23/05/2022	Den	15.2	13.2	0.3	5.77E+05	0.10	0.06
BV01-3	A	water	56.058401	12.577956	June	10/06/2022	Den	15.3	14.6	0.5	5.49E+05	0.11	0.07
BV01-3	B	water	56.058692	12.578673	June	10/06/2022	Den	15.3	14.6	0.1	5.47E+05	0.09	0.07
BV01-3	C	water	56.058401	12.580089	June	10/06/2022	Den	15.2	14.5	0.1	5.64E+05	0.09	0.07
BV01-4	A	water	56.058401	12.577956	June	24/06/2022	Den	16.7	17.8	0.2	1.01E+06	0.11	0.06
BV01-4	B	water	56.058692	12.578673	June	24/06/2022	Den	16.7	17.8	0.5	9.19E+05	0.11	0.07
BV01-4	C	water	56.058401	12.580089	June	24/06/2022	Den	16.7	17.2	1.2	5.28E+05	0.11	0.07
BV01-5	A	water	56.058401	12.577956	July	11/07/2022	Den	16.6	17.9	1.4	2.57E+06	0.10	0.06
BV01-5	B	water	56.058692	12.578673	July	11/07/2022	Den	16.6	17.9	0.3	2.53E+06	0.08	0.08
BV01-5	C	water	56.058401	12.580089	July	11/07/2022	Den	16.6	17.9	0.1	2.13E+06	0.08	0.08
BV01-6	A	water	56.058401	12.577956	July	20/07/2022	Den	17.5	19.8	0.1	1.20E+06	0.11	0.06
BV01-6	B	water	56.058692	12.578673	July	20/07/2022	Den	13.1	19.5	0.2	1.19E+06	0.10	0.07
BV01-6	C	water	56.058401	12.580089	July	20/07/2022	Den	13.1	19.5	0.2	1.04E+06	0.10	0.07
BV01-7	A	water	56.058401	12.577956	August	03/08/2022	Den	14.9	18.0	0.3	8.82E+05	0.10	0.07
BV01-7	B	water	56.058692	12.578673	August	03/08/2022	Den	14.9	18.0	0.1	1.34E+06	0.11	0.08
BV01-7	C	water	56.058401	12.580089	August	03/08/2022	Den	14.9	18.0	0.1	7.34E+05	0.11	0.08
BV01-8	A	water	56.058401	12.577956	August	10/08/2022	Den	14.4	18.1	0.2	1.19E+06	0.11	0.07
BV01-8	B	water	56.058692	12.578673	August	10/08/2022	Den	14.4	18.1	0.1	1.46E+06	0.12	0.07
BV01-8	C	water	56.058401	12.580089	August	10/08/2022	Den	14.4	18.1	0.2	9.43E+05	0.12	0.07
BV01-9	A	water	56.058401	12.577956	September	07/09/2022	Den	14.9	16.6	0.2	1.62E+06	0.08	0.05
BV01-9	B	water	56.058692	12.578673	September	07/09/2022	Den	14.9	16.6	0.5	1.20E+06	0.10	0.05
BV01-9	C	water	56.058401	12.580089	September	07/09/2022	Den	13.9	17.3	0.7	1.32E+06	0.10	0.05
BV01-10	A	water	56.058401	12.577956	September	16/09/2022	Den	16.6	15.6	0.2	1.80E+06	0.20	0.06
BV01-10	B	water	56.058692	12.578673	September	16/09/2022	Den	16.6	15.6	0.1	1.31E+06	0.10	0.07
BV01-10	C	water	56.058401	12.580089	September	16/09/2022	Den	18.8	16.3	0.2	1.03E+06	0.10	0.07
BV01-11	A	water	56.058401	12.577956	October	05/10/2022	Den	16.6	13.6	0.2	8.44E+05	0.10	0.06

BV01-11	B	water	56.058692	12.578673	October	05/10/2022	Den	16.6	13.6	0.1	1.24E+06	0.10	0.06
BV01-11	C	water	56.058401	12.580089	October	05/10/2022	Den	16.6	13.6	0.1	1.11E+06	0.10	0.06
BV01-12	A	water	56.058401	12.577956	October	14/10/2022	Den	21.3	12.6	0.1	4.58E+05	0.12	0.07
BV01-12	B	water	56.058692	12.578673	October	14/10/2022	Den	21.3	12.6	0.1	4.62E+05	0.12	0.09
BV01-12	C	water	56.058401	12.580089	October	14/10/2022	Den	21.3	12.6	0.0	4.63E+05	0.12	0.09
BV02-1	A	water	54.179579	12.103401	May	04/05/2022	Ger	8.7	11.8	0.1	3.89E+05	0.12	0.08
BV02-1	B	water	54.179639	12.103105	May	04/05/2022	Ger	8.7	12.5	0.5	3.33E+05	0.09	0.05
BV02-1	C	water	54.1788	12.106677	May	04/05/2022	Ger	8.7	13.2	0.1	4.47E+05	0.09	0.05
BV02-2	A	water	54.179579	12.103401	June	10/06/2022	Ger	10.4	17.4	0.0	4.20E+05	0.09	0.07
BV02-2	B	water	54.179639	12.103105	June	10/06/2022	Ger	10.7	17.7	0.2	6.70E+05	0.08	0.09
BV02-2	C	water	54.1788	12.106677	June	10/06/2022	Ger	10.7	18.4	0.2	8.03E+05	0.08	0.09
BV02-3	A	water	54.179579	12.103401	July	12/07/2022	Ger	8.3	18.9	0.1	1.45E+06	0.10	0.10
BV02-3	B	water	54.179639	12.103105	July	12/07/2022	Ger	8.2	19.4	0.2	1.60E+06	0.09	0.12
BV02-3	C	water	54.1788	12.106677	July	12/07/2022	Ger	8.3	19.2	0.2	1.43E+06	0.09	0.12
BV02-4	A	water	54.179579	12.103401	July	26/07/2022	Ger	9.9	19.5	0.4	1.22E+06	0.14	0.12
BV02-4	B	water	54.179639	12.103105	July	26/07/2022	Ger	10.0	19.5	0.4	1.83E+06	0.22	0.15
BV02-4	C	water	54.1788	12.106677	July	26/07/2022	Ger	9.9	19.4	0.4	1.94E+06	0.22	0.15
BV02-5	A	water	54.179579	12.103401	August	10/08/2022	Ger	9.5	21.5	0.1	1.27E+06	0.12	0.10
BV02-5	B	water	54.179639	12.103105	August	10/08/2022	Ger	9.7	21.3	0.2	9.70E+05	0.11	0.12
BV02-5	C	water	54.1788	12.106677	August	10/08/2022	Ger	9.8	21.5	0.1	2.32E+06	0.11	0.12
BV02-6	A	water	54.179579	12.103401	August	17/08/2022	Ger	10.1	21.6	0.2	3.29E+05	0.11	0.08
BV02-6	B	water	54.179639	12.103105	August	17/08/2022	Ger	9.8	21.9	0.2	7.96E+05	0.11	0.09
BV02-6	C	water	54.1788	12.106677	August	17/08/2022	Ger	9.2	20.5	0.1	1.27E+06	0.11	0.09
BV02-7	A	water	54.179579	12.103401	September	23/09/2022	Ger	10.2	15.2	0.1	6.17E+05	0.17	0.18
BV02-7	B	water	54.179639	12.103105	September	23/09/2022	Ger	10.2	15.4	0.0	7.20E+05	0.10	0.14
BV02-7	C	water	54.1788	12.106677	September	23/09/2022	Ger	11.0	15.8	0.2	4.90E+05	0.10	0.14
BV02-8	A	water	54.179579	12.103401	October	14/10/2022	Ger	10.8	13.4	0.4	6.83E+05	0.17	0.14
BV02-8	B	water	54.179639	12.103105	October	14/10/2022	Ger	11.4	13.5	0.5	7.51E+05	0.13	0.12
BV02-8	C	water	54.1788	12.106677	October	14/10/2022	Ger	11.2	13.6	0.0	6.99E+05	0.13	0.12
BV03-1	A	water	60.1087	21.7119	May	18/05/2022	Fin	6.5	7.0	0.2	3.96E+05	0.10	0.08
BV03-1	B	water	60.1087	21.7119	May	18/05/2022	Fin	6.5	7.0	0.1	5.16E+05	0.09	0.08
BV03-1	C	water	60.1087	21.7119	May	18/05/2022	Fin	6.5	7.0	0.1	5.49E+05	0.09	0.08
BV03-2	A	water	60.1087	21.7119	June	14/06/2022	Fin	6.4	13.5	0.1	6.59E+05	0.09	0.09
BV03-2	B	water	60.1087	21.7119	June	14/06/2022	Fin	6.4	14.4	0.1	8.87E+05	0.09	0.05
BV03-2	C	water	60.1087	21.7119	June	14/06/2022	Fin	6.4	14.5	0.0	8.46E+05	0.09	0.05
BV03-3	A	water	60.1087	21.7119	July	13/07/2022	Fin	6.5	19.1	0.1	9.06E+05	0.10	0.08
BV03-3	B	water	60.1087	21.7119	July	13/07/2022	Fin	6.5	19.2	0.1	6.05E+05	0.08	0.07
BV03-3	C	water	60.1087	21.7119	July	13/07/2022	Fin	6.5	19.5	0.1	1.19E+06	0.08	0.07

BV03-4	A	water	60.1087	21.7119	August	09/08/2022	Fin	6.3	18.5	0.0	3.73E+05	0.10	0.08
BV03-4	B	water	60.1087	21.7119	August	09/08/2022	Fin	6.3	18.3	0.4	4.95E+05	0.08	0.06
BV03-4	C	water	60.1087	21.7119	August	09/08/2022	Fin	6.3	19.1	0.0	4.41E+05	0.08	0.06
BV03-5	A	water	60.1087	21.7119	September	15/09/2022	Fin	6.4	15.7	0.1	3.31E+05	0.13	0.09
BV03-5	B	water	60.1087	21.7119	September	15/09/2022	Fin	6.4	15.3	0.0	2.94E+05	0.09	0.09
BV03-5	C	water	60.1087	21.7119	September	15/09/2022	Fin	6.4	15.7	0.0	2.83E+05	0.09	0.09
BV03-6	A	water	60.1087	21.7119	October	12/10/2022	Fin	6.4	12.2	0.2	3.24E+05	0.40	0.17
BV03-6	B	water	60.1087	21.7119	October	12/10/2022	Fin	6.4	12.0	0.1	3.52E+05	0.23	0.14
BV03-6	C	water	60.1087	21.7119	October	12/10/2022	Fin	6.4	11.9	0.2	2.90E+05	0.23	0.14

Table S2. Grain size composition. *For the analysis, 27 sediment samples obtained between July and September (n = 9 per station) were selected from the three stations. The values are the average ± standard error.*

	Grain size composition %				
	Clay	**Silt**	**Sand fine**	**Sand medium**	**Sand coarse**
Denmark	0.00 ± 0.00	4.02 ± 3.35	87.49 ± 16.95	8.49 ± 3.84	0.00 ± 0.00
Germany	0.02 ± 0.00	6.55 ± 1.62	79.64 ± 13.89	13.80 ± 3.42	0.03 ± 0.00
Finland	0.00 ± 0.00	1.16 ± 1.93	53.53 ± 12.28	45.09 ± 3.00	0.77 ± 0.00

Table S3. Spearman's rank correlation coefficients (Bonferroni corrected) between environmental parameters measured in the water column at the three stations*. Colors indicate significant correlations (green, p < 0.05, orange, p < 0.01, red, p < 0.001), and the analysis includes 234 water samples from all the stations. SSWWH: sea surface wind wave significant height.*

	1	2	3	4	5	6	7	8	9	10	11	12
SST (°C) (1)	-	-0.07	0.20	0.59	0.11	-0.02	-0.08	-0.08	-0.41	0.37	0.47	-0.18
Salinity (2)		-	-0.22	0.33	-0.32	0.30	0.17	-0.61	0.61	0.19	-0.52	0.36
Chl *a* (mg m^{-3}) (3)			-	0.06	0.33	-0.09	0.15	0.23	-0.22	0.40	0.52	-0.21
Babterial abundance (cells ml^{-1}) (4)				-	-0.07	0.19	0.02	-0.29	-0.03	0.45	0.21	0.17
PO_4^{3-} (µM) (5)					-	-0.09	0.46	0.12	-0.22	0.32	0.21	-0.04
NH_4^+ (µM) (6)						-	0.18	-0.35	0.09	0.18	-0.23	0.23
NO_3^- (µM) (7)							-	-0.11	0.10	0.25	-0.01	-0.03
Oxygen (mg l^{-1}) (8)								-	-0.08	-0.02	0.29	-0.22
Secchi depth (m) (9)									-	0.07	-0.55	0.27
Sediment bacteria (cells ml^{-1}) (10)										-	0.21	0.19
Relative abundance harmful cyanobacteria (11)											-	-0.13
SSWWH (m) (12)												-

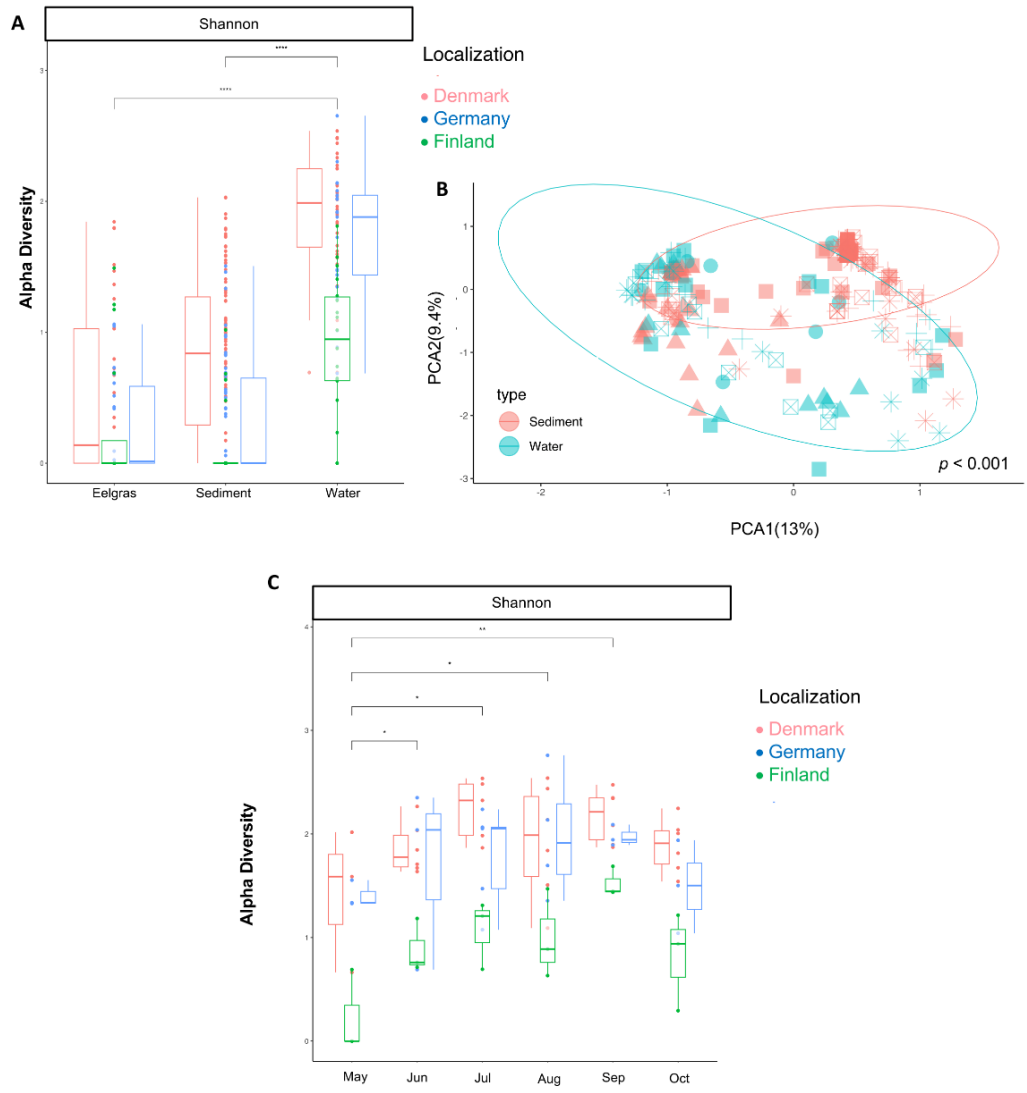


Figure S1.* Vibrio *composition in the different compartments sampled, i.e., water, sediment, and eelgrass. *A)* Vibrio *spp. alpha diversity in water, sediment, and eelgrass. B) Principal component analysis (PCA) of the* Vibrio *spp. composition in waters and sediments. Eigen-values are 10.78 for PCA1 and 7.75 for PCA2. Reported p values were calculated by a PERMANOVA test, and circles enclose sample groups (sediment and water, n = 282). C)* Vibrio *spp. alpha diversity over time in the water columns of the three stations. In (A), values are the median of all values collected at the three substations from eelgrass, sediment, and water. In (C), values are the median of the replicates from the three substations at the specific time. The vertical lines indicate the position of the lower and upper quartiles. The data from two monthly sampling points in Denmark and Germany (only July and August) were combined to enable comparisons with the Finnish station. Asterisks indicate pairwise significant differences between the type of sample or month, (*p < 0.05, ** p < 0.01, ****, p < 0.0001), using a posthoc test (Wilcoxon) after a Kruskal-Wallis test. In (A, C), colored dots represent the individual data points used to construct the box plot.*

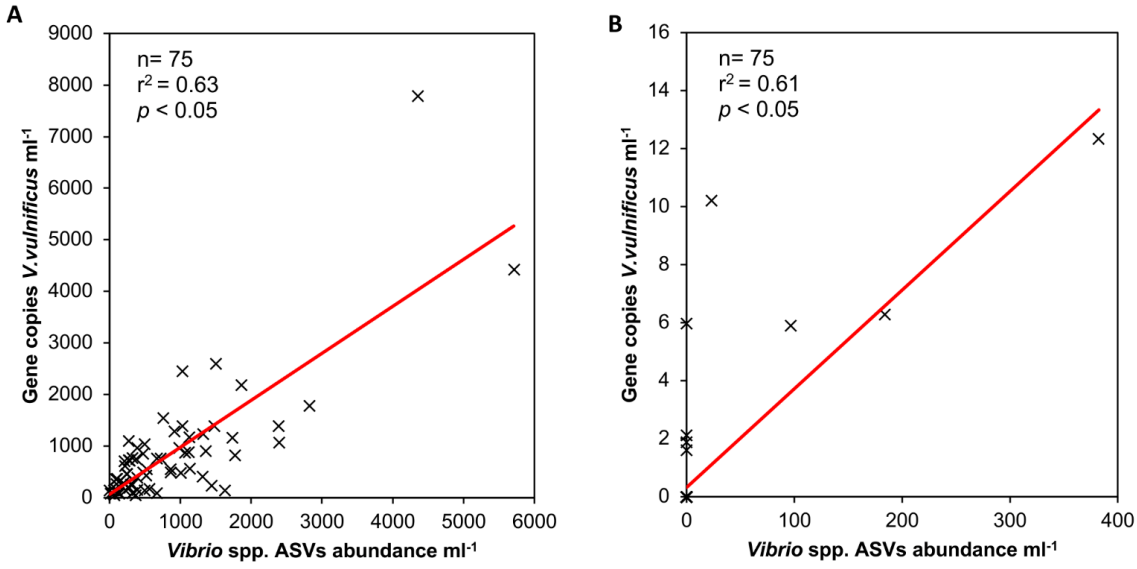


Figure S2. Regression scatter plots between* Vibrio *spp. and* V. vulnificus *ASVs abundance (obtained by multiplying the relative* Vibrio *ASVs frequencies with bacterial abundance in each sample) vs. the gene copies of* Vibrio *spp. *(A) and* V. vulnificus *(B) quantified by ddPCRs. The triplicate ddPCR values were combined by calculating their average to enable comparisons with the pooled water samples that were sequenced.*

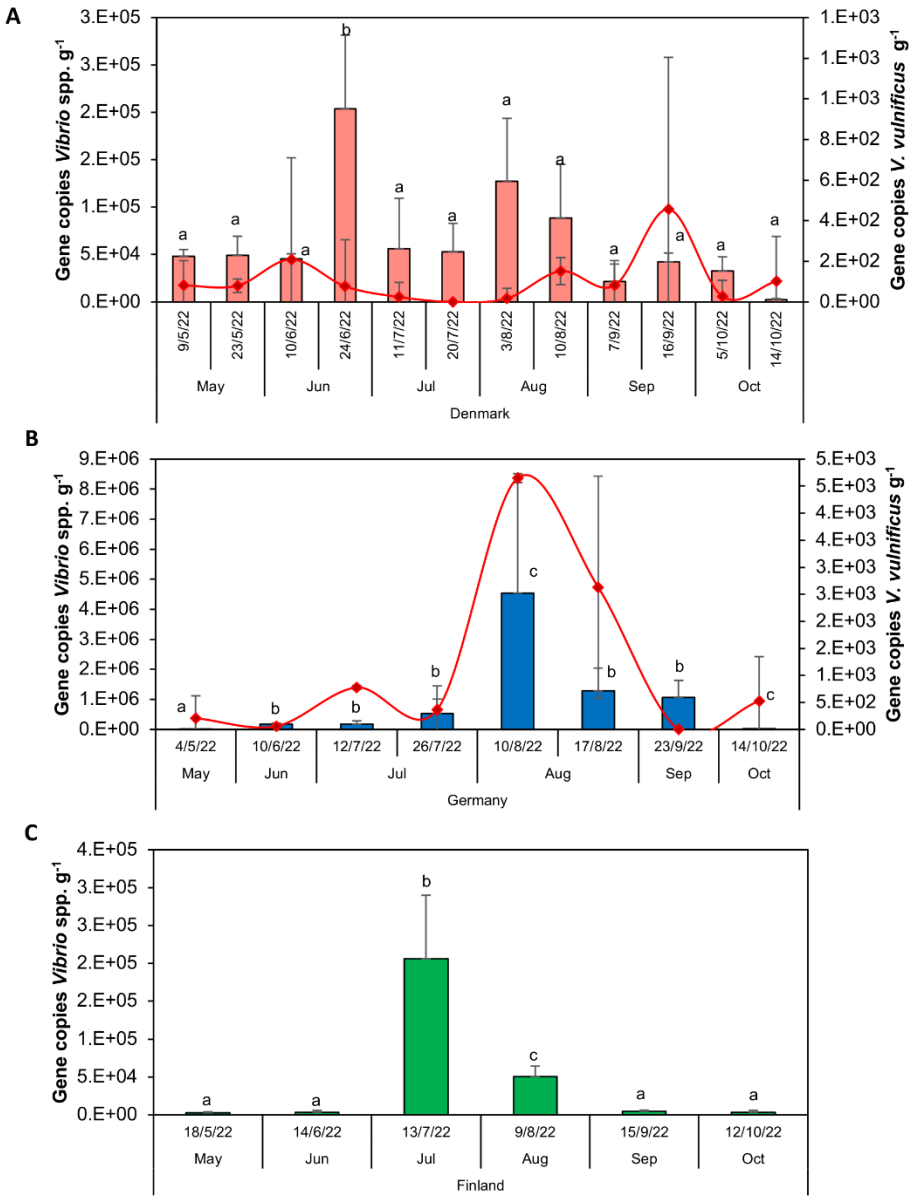


Figure S3. Temporal dynamics in the abundance of* Vibrio *spp. and* V. vulnificus *in sediment at the *A) Danish, B) German, and C) Finish stations quantified by ddPCR. The values are the average ± the standard error between the replicates (n = 9). Letters indicate pairwise analysis among the variables (i.e., month) at each station using a posthoc test (Wilcoxon) after Kruskal-Wallis for* Vibrio *spp. Bars and red lines represent the gene copies of* Vibrio *spp. and* V. vulnificus*, respectively.*

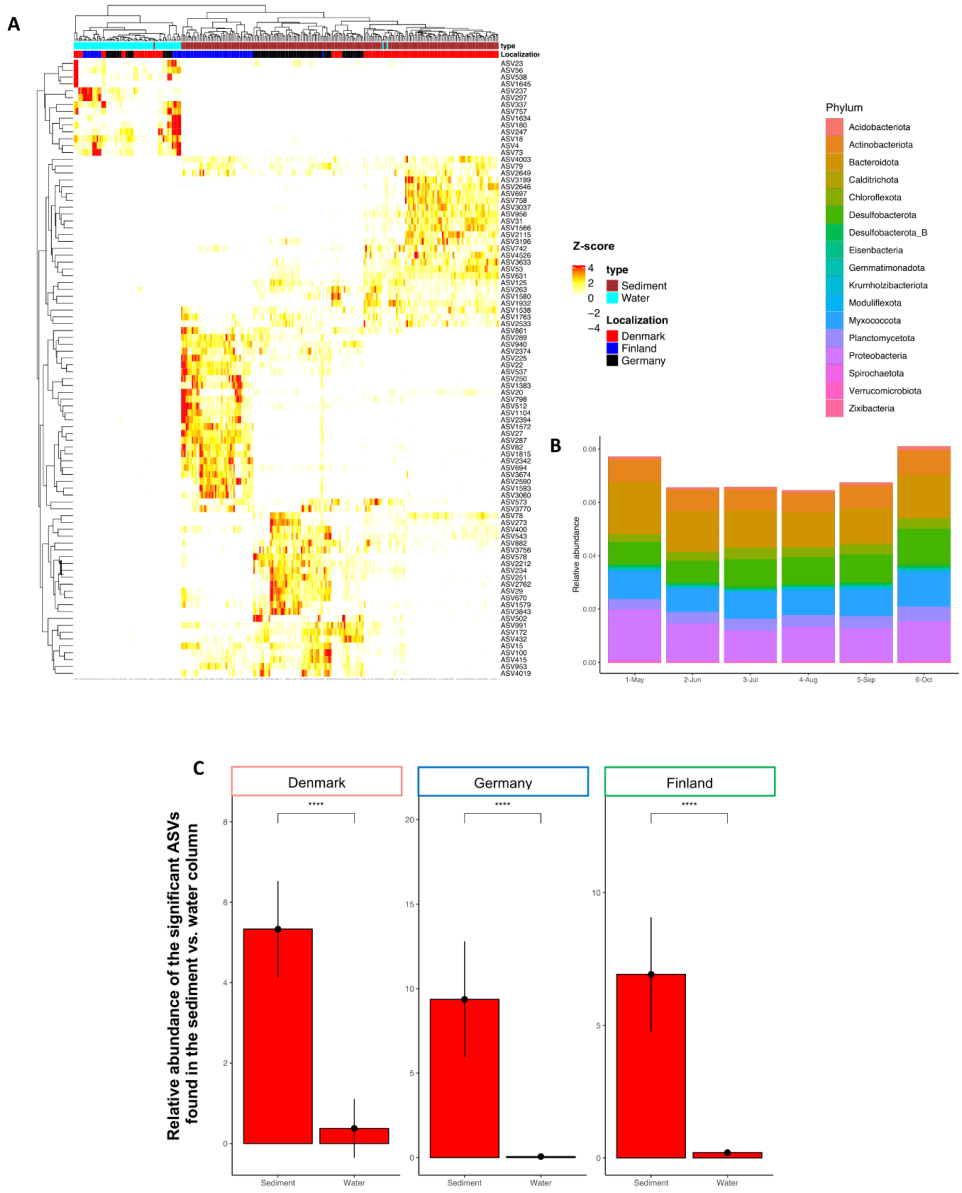


Figure S4. ASVs identified as differentially abundant in the sediment utilized as indicators for sediment resuspension in the water column (used in Fig. 5). *A) Heat map with cluster analysis of the most abundant, dominant, and significant ASVs identified by DESeq2 in the sediment (brown, n = 204) and water (blue cyan, n = 78) ($p < 0.01$). The heat map scale displays the row Z-score, i.e., the normalized taxa abundances (abundance of one ASV in one sample) - (mean abundance for that ASV))/(standard deviation of that ASV)). B) Phylum classification of the top ASVs identified as indicators for sediment resuspension. C) Comparison of the relative abundance of the ASVs detected by DESeq2 in water and sediment at each station. Values are the average ± the standard error between the replicates. Asterisks indicate pairwise significant differences between water and sediment (**** $p < 0.0001$), using a posthoc test (Wilcoxon) after a Kruskal-Wallis test.*

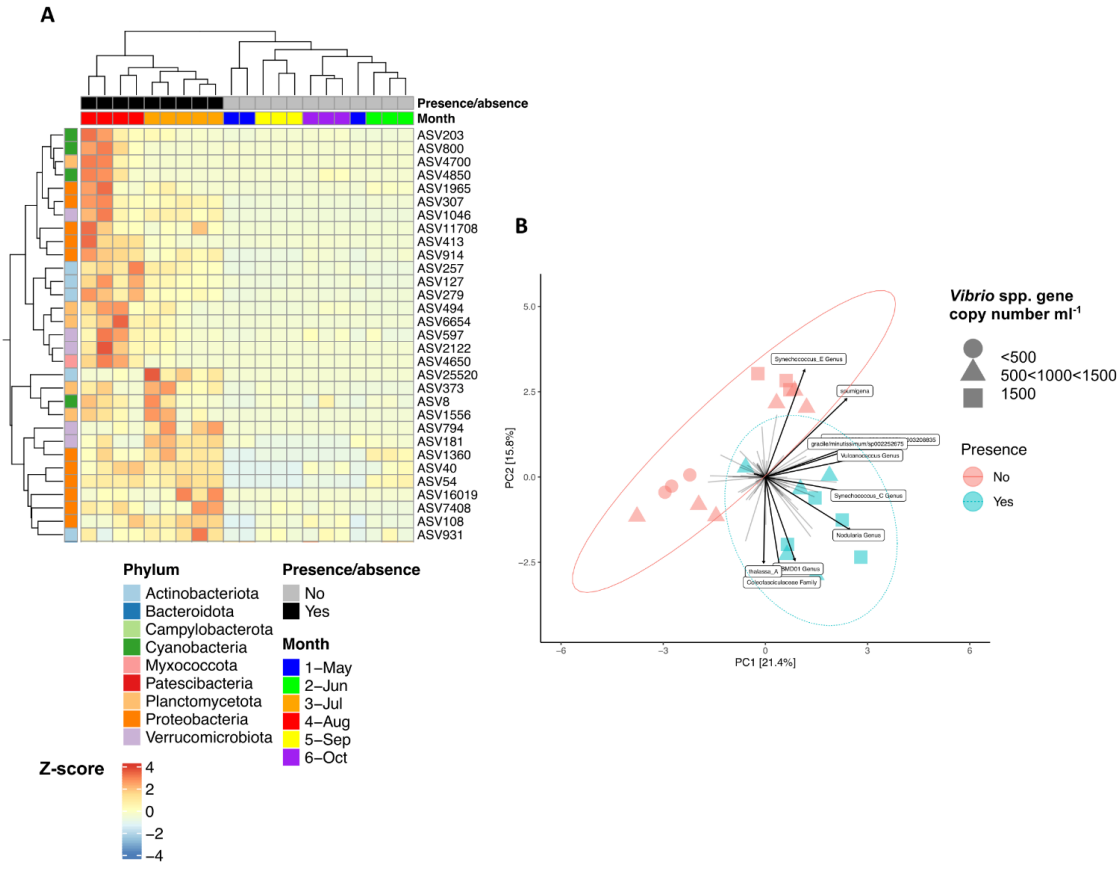
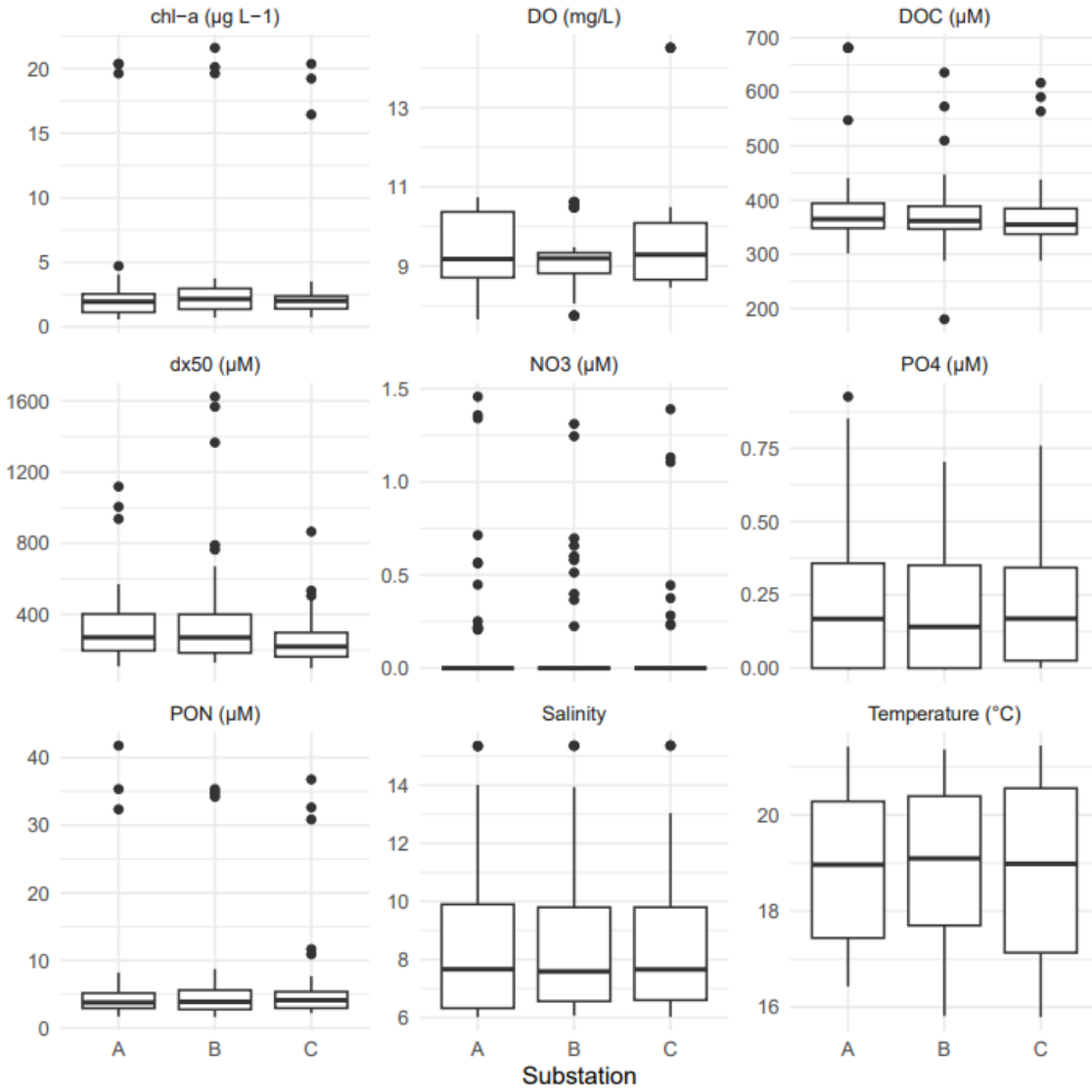
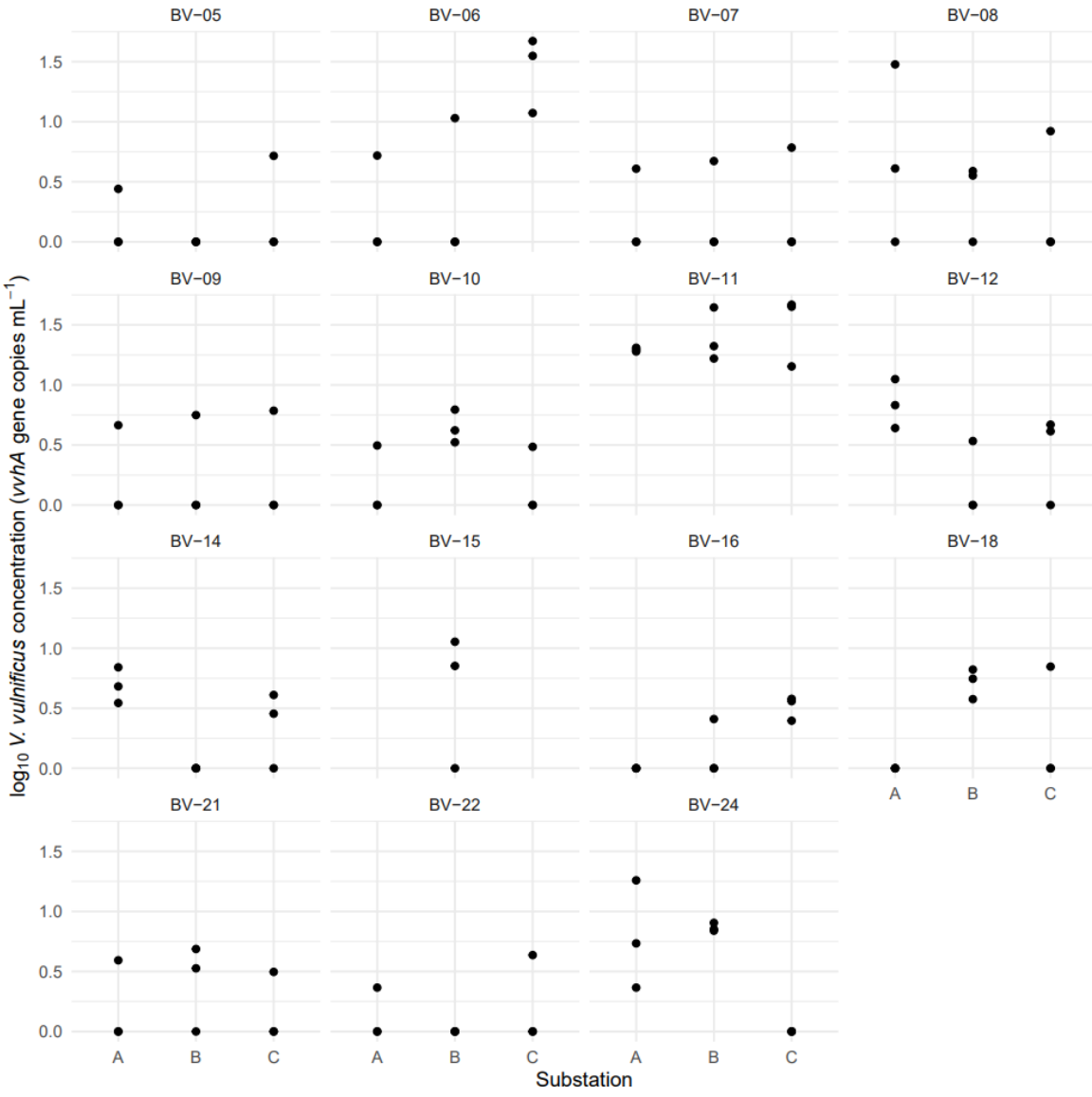


Figure S5. Dominating taxa when* V. vulnificus *was present. *A) Heat map with cluster analysis of the top 30 most significant and dominant ASVs identified by DESeq2 as differentially abundant when* V. vulnificus *was present (black) or absent (grey) ($p < 0.01$). The heatmap scale displays the row Z-score. B) PCA of the cyanobacterial community according to the presence/absence of* V. vulnificus*. Eigenvalues are 587.7 for PCA1 and 487.3 for PCA2. Arrows with cyanobacterial taxa represent the 10 most abundant taxa in the community. Reported p values were calculated by a PERMANOVA test performed with 999 permutations, and circles enclose sample groups. In A and B, "No" and "Yes" mean presence ($n = 9$) and absence ($n = 12$) of V. vulnificus. For these analyses, only Clr-transformed data from Germany was used, as V. vulnificus in the water column was only detected in this area.*

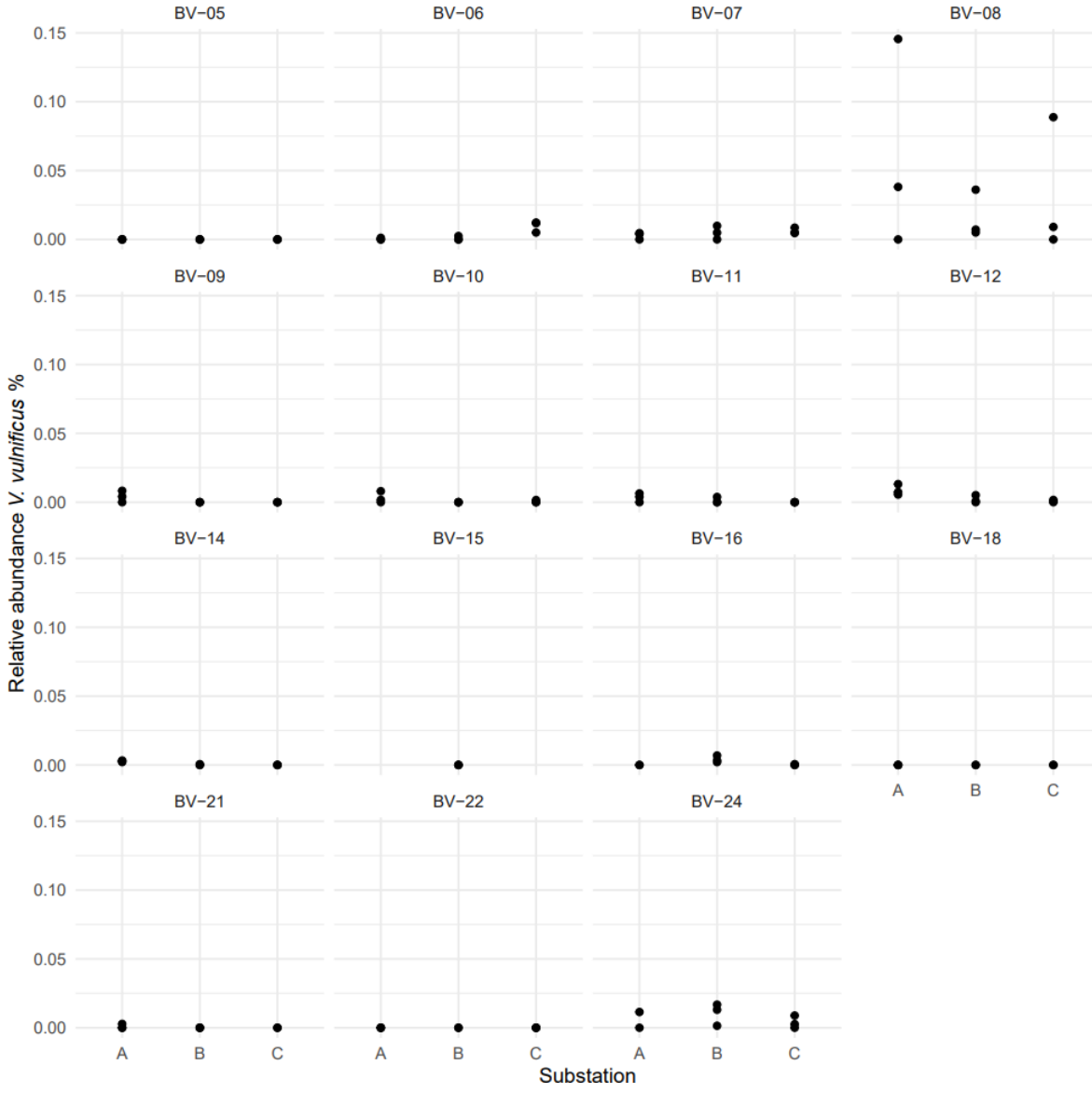
Chapter VI



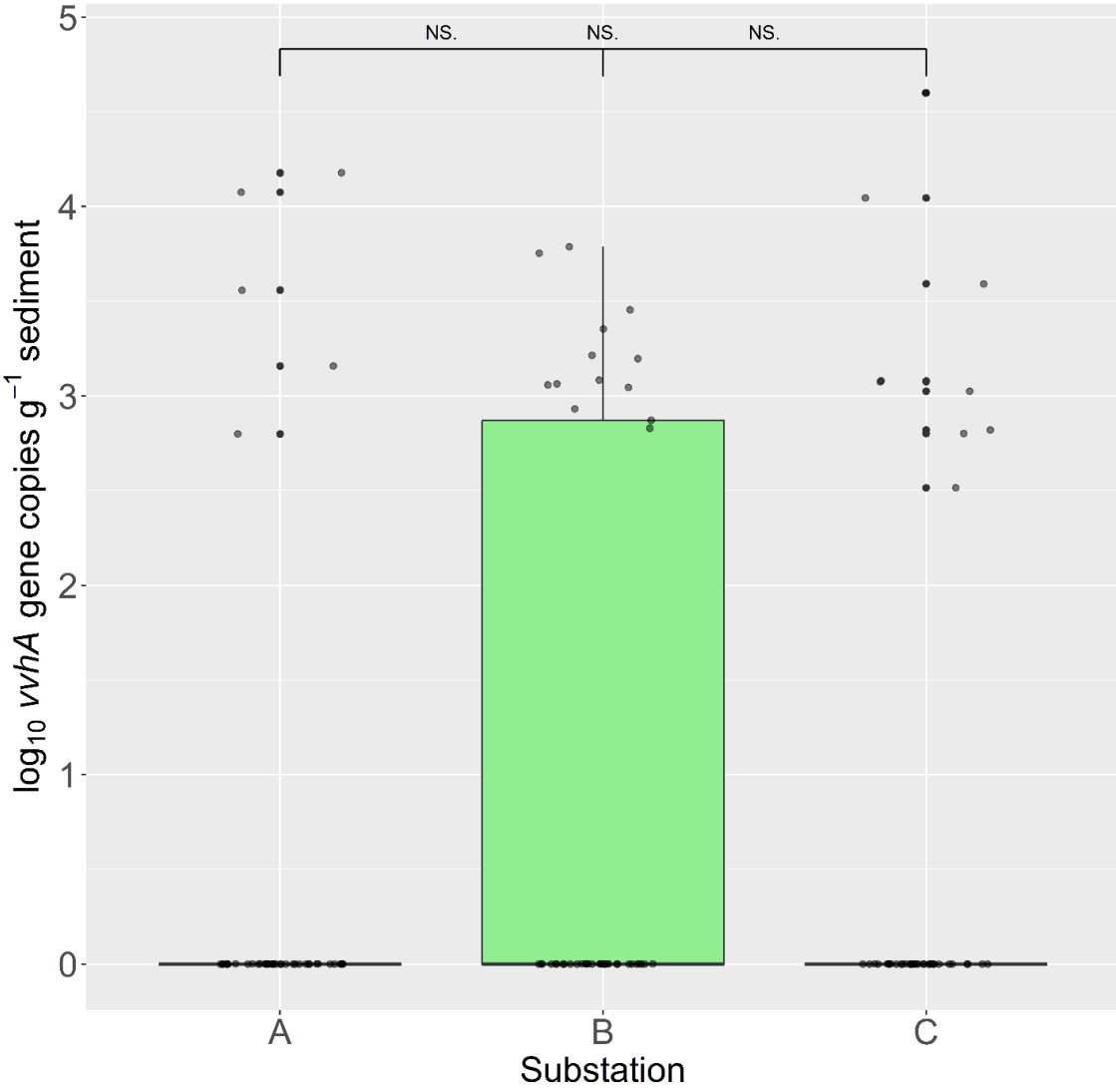
Supplementary figure 1. Distribution of selected environmental parameters per substation. *No significant differences (Wilcoxon rank sum test ($p > 0.05$) between the substations could be detected*



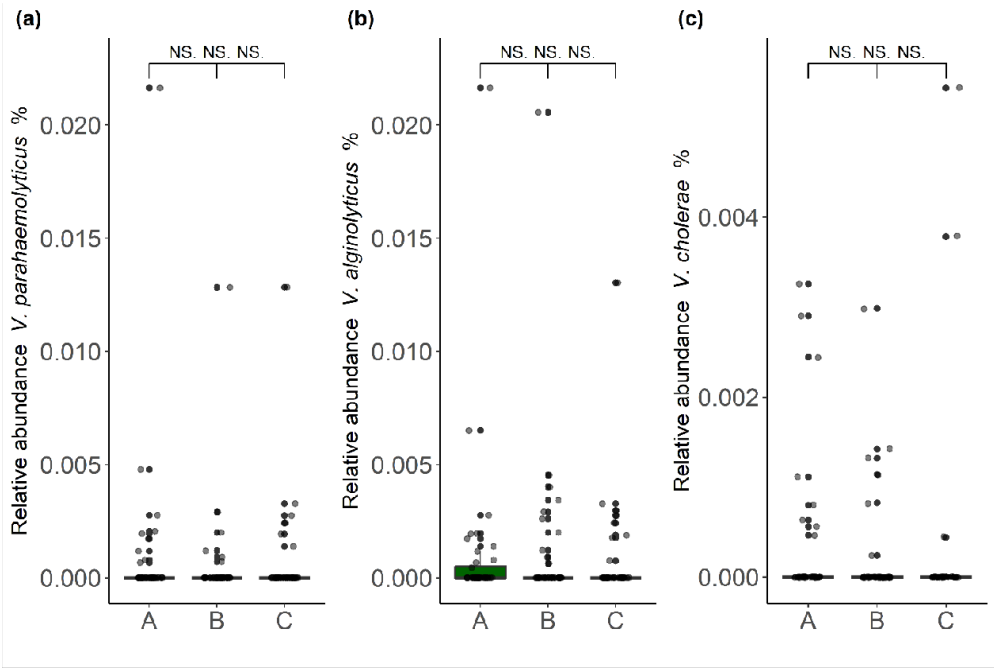
Supplementary figure 2. **vvhA** ***gene copy number per substation****.No clear trend in* vvhA *gene concentration per substation is visible. Applying Kruskal-Wallis per station shows no significant effect (p > 0.05) after Holm correction.*



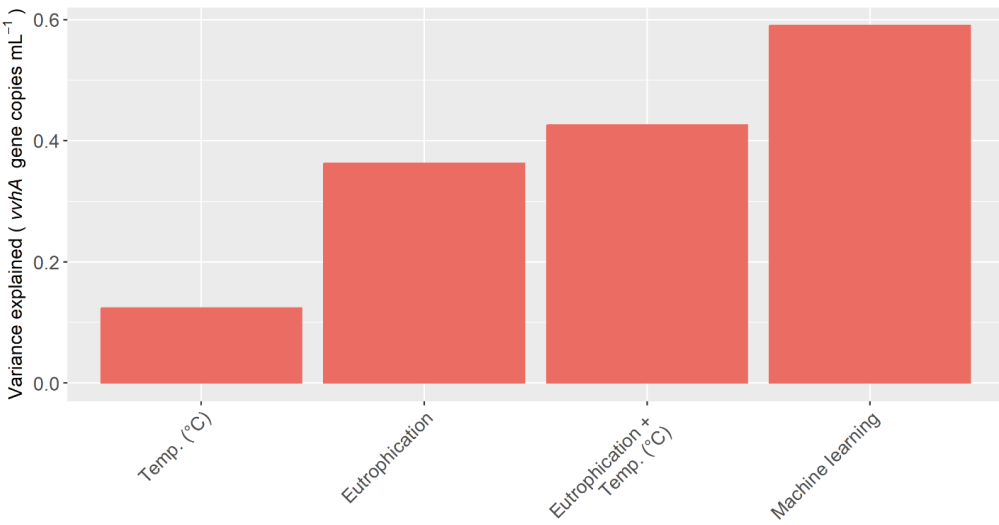
Supplementary figure 3. Relative abundance of* V. vulnificus *per substation. *No clear trend in relative abundance of* V. vulnificus *per substation is visible. Applying Kruskal-Wallis per station shows no significant effect ($p > 0.05$) after Holm correction.*



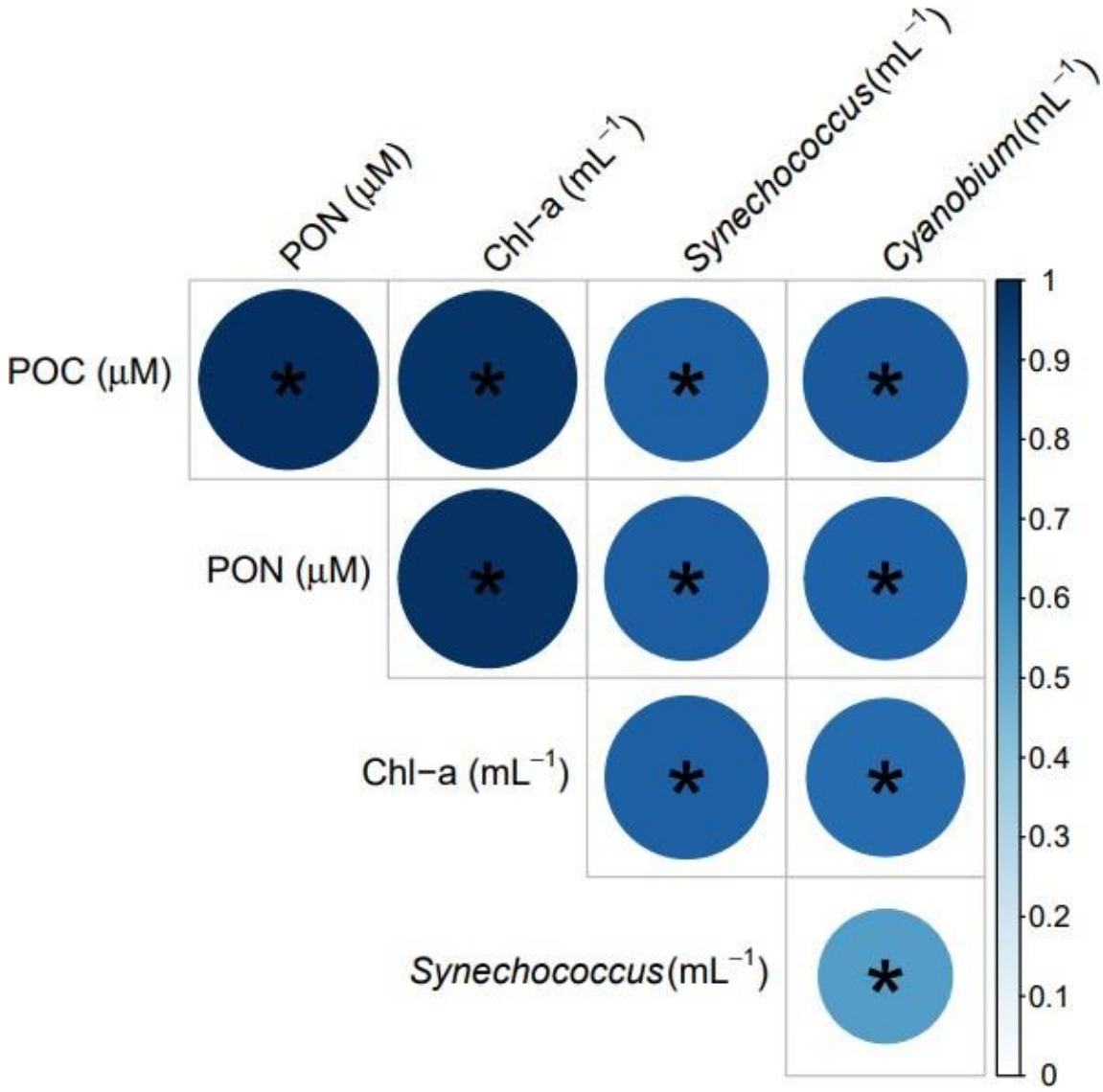
Supplementary figure 4. V. vulnificus abundance in the sediment. No statistically significant effect of substation on vvhA genecopies per gram in the sediment (Wilcoxon rank sum test ($p > 0.05$)). Boxes represent the 25th and 75th percentiles.



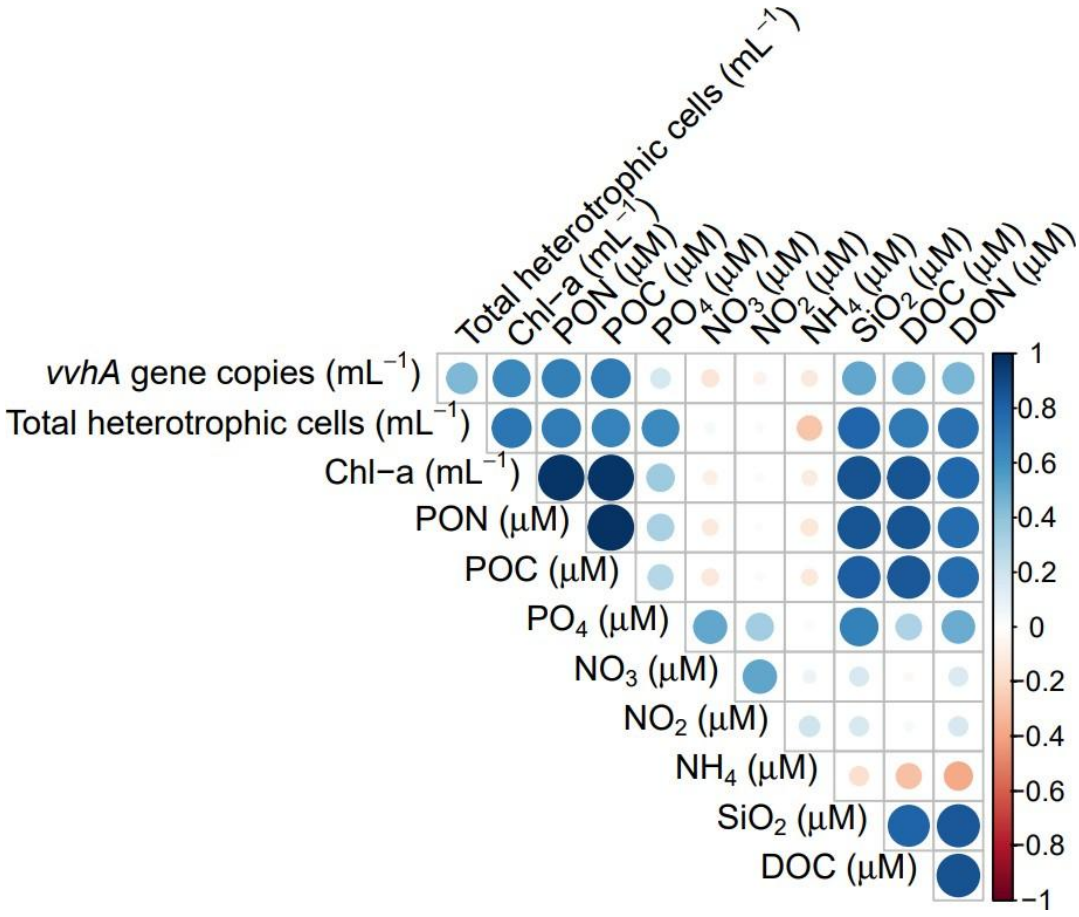
Supplementary figure 5. Relative* V. parahaemolyticus / alginolyticus / cholera *abundance per substations. *A (in the seagrass meadow), B (15m from the edge) and C (100m from the edge). Wilcoxon rank sum test (p = 0.05) with Holm correction for multiple testing. Boxes represent the 25th and 75th percentiles.*



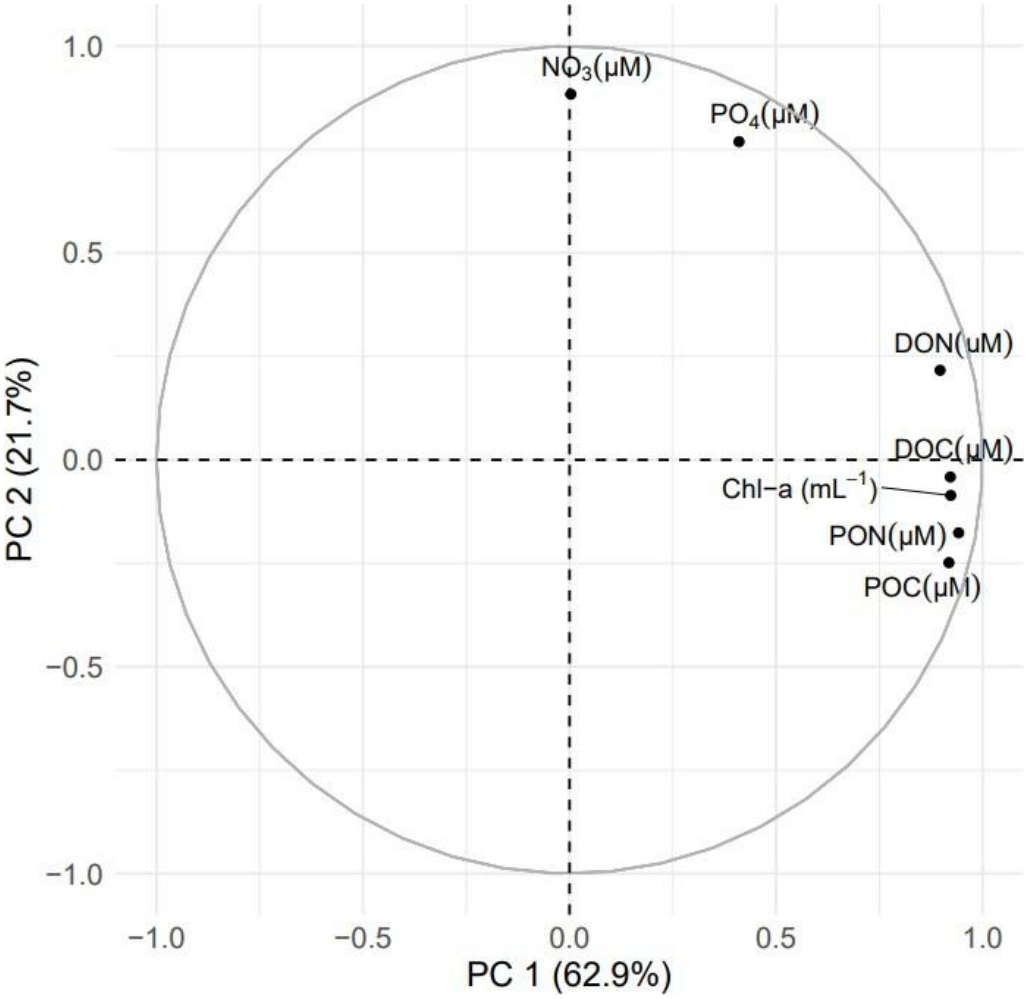
Supplementary figure 6. Variance in* vvhA *gene concentration explained (R^2) by (multiple) linear regression and machine learning. *The eutrophication index explains more of the variance in* vvhA *gene concentration than temperatures. The combination of eutrophication and temperature is marginally better. The more complex machine learning approach explains almost 60% of the variance and outperforms the other models.*



Supplementary figure 7. Correlation matrix between measured chl-a and predictors identified by RFE. *Chl-a strongly correlates with the predictors for* V. vulnificus *identified by the RFE algorithm. A star signifies a significant correlation (p = 0.05)*



Supplementary figure 8. Correlation matrix between eutrophication parameters, heterotrophic cells counts and* vvhA *gene copy numbers. *Total heterotrophic cells correlate more strongly to phosphate than* vvhA *gene copy number.*



Supplementary figure 9. PCA performed for the eutrophication index. *Dissolved and particulate organic carbon and nitrogen as well as chl-a, align with the principal axis, which explains 63% of the variance. Inorganic nutrients* NO^- *and* PO^{3-} *are more aligned to the second axis. Data was standardized.*

Supplementary Table 1. Taxonomy of predictive prokaryotes

ASV	phylum	class	order	family	genus
219	Cyanobacteria	Cyanobacteriia	Synechococcales	Cyanobiaceae	*Cyanobium* PCC-6307
379	Cyanobacteria	Cyanobacteriia	Synechococcales	Cyanobiaceae	*Cyanobium* PCC-6308
236	Actinobacteriota	Actinobacteria	Micrococcales	Microbacteriaceae	*Candidatus* Aquiluna
328	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	NA
404	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Halieaceae	*Halioglobus*
654	Actinobacteriota	Actinobacteria	Micrococcales	Microbacteriaceae	ML602J-51
1070	Actinobacteriota	Actinobacteria	Propionibacteriales	Nocardioidaceae	*Nocardioides*
1187	Actinobacteriota	Acidimicrobiia	Microtrichales	Ilumatobacteraceae	NA
2690	Cyanobacteria	Cyanobacteriia	Cyanobacteriales	Nostocaceae	*Nodularia* PCC-9350

Supplementary Table 2. Taxonomy of predictive eukaryotes

ASV	supergroup	domain	clade	order	family	genus
6	Hacrobia	Cryptophyta	Cryptophyceae	Cryptomonadales		*Teleaulax*
71	Archaeplastida	Rhodophyta	Florideophyceae	Ceramiales	Ceramiaceae	*Ceramium*
148	Archaeplastida	Rhodophyta	Florideophyceae	Ceramiales	Rhodomelaceae	
184	Rhizaria	Cercozoa	Filosa-Thecofilosea			
312	Opisthokonta	Fungi	Chytridiomycota	Chytridiomycotina	Chytridiomycetes	*Rhizophydium*
320	Rhizaria	Cercozoa	Filosa-Imbricatea	Thaumatomonadida	Thaumatomonadidae	*Thaumatomastix*
430	Stramenopiles	Sagenista	Labyrinthulomycetes	Labyrinthulida	Labyrinthulaceae	*Aplanochytrium*
608	Apusozoa	Apusomonadidae	Apusomonadidae_Group-1			

Supplementary Table 3. Primer sequences

Target	Primer	Sequence (5´–3´)	Reference
Isolate verification *vvhA* gene	vvhA f	ACCAAGTTTGGGGCCTAGAT	Messelhäusser et al. 2010 [1]
	vvhA r	GCTAAGTTCGCACCACACTG	
ddPCR	vvh-785f	TTCCAACTTCAAACCGAACTATGAC	Panicker et al. 2004 [2]
	vvh-990r	ATTCCAGTCGATGCGAATACGTTG	
16S amplification	341F	CCTACGGGNGGCWGCAG	Herlemann et al. 2011 [3]
	805R	GACTACHVGGGTATCTAATCC	
18S amplification	V4F	CCAGCASCYGCGGTAATTCC	Balzano et al. 2015 [4]
	V4RB	ACTTTCGTTCTTGATYRR	
Illumina adapters	F	ACACTCTTTCCCTACACGACGCTCTTCCGATCT	IDT DNA (IA, USA)
	R	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	

Supplementary Table 4. PCR condition

Method	Time (s)		Temperature °C
Isolate verification		120	50
		600	95
	35 cycles	15	95
		60	55
ddPCR		300	95
		30	95
	40 cycles	60	58.5
		60	72
		300	4
		300	90
16S amplification		120	98
		20	98
	28 cycles	20	54
		15	72
		120	72
18S amplification		180	95
		20	98
	30 cycles	15	52
		15	72
		120	72
Indexing PCR		120	95
		20	98
	8 cycles	30	55
		30	72
		120	72

Supplementary Table 5. PCR solutions

Target	Total Volume (µL)	Reagents	Volume(µL)
verification gene *vvhA*	25	TaqMan Environmental MasterMix 2.0	6.25
		toxR-F (10 µM)	0.75
		toxR-R (10 µM)	0.75
		toxR-TM FAM (10 µM)	0.25
		vvha-F (10 µM)	0.75
		vvha-R (10 µM)	0.75
		vvha-TM Cy5 (10 µM)	0.25
		DEPC H2O	10.25
		Template DNA	5
ddPCR	22	2x QX200 TM ddPCR™ EvaGreen® Supermix	11
		vvh-785f (10 µM)	0.22
		vvh-990r (10 µM)	0.22
		Template DNA (0.01-5 ng μL^{-1})	1
		DEPC H2O	9.56
16S rRNA	25	2x Kapa HiFi HotStart ReadyMix	12.5
		341F (10 µM)	0.75
		805R (10 µM)	0.75
		Template DNA (1 ng μL^{-1})	1
		DEPC H2O	10
18S rRNA	25	2x Kapa HiFi HotStart ReadyMix	12.5
		V4F (10 µM)	0.75
		V4RB (10 µM)	0.75
		Template DNA (5 ng μL^{-1})	1
		DEPC H2O	10

Supplementary references

1. Messelhäusser U. et al. Detection and differentiation of *Vibrio* spp. in seafood and fish samples with cultural and molecularmethods. *Int J Food Microbiol*. 2010 142:360–364. DOI https://doi.org/10.1016/j.ijfoodmicro.2010.07.020

2. Panicker, G., Vickery, M. C. L., Bej, A. K. Multiplex PCR detection of clinical and environmental strains of *Vibrio vulnificus* in shellfish. *Can. J. Microbiol*. 2004, 50, 911–922 DOI https://doi.org/10.1139/w04-085

3. Herlemann, D.P. et al. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *The ISMEjournal*. 2011, 5(10), pp.1571-1579. DOI https://doi.org/10.1038/ismej.2011.41

4.Balzano, S., Abs, E., Leterme, S.C. Protist diversity along a salinity gradient in a coastal lagoon. *Aquatic Microbial Ecology*, 74(3). 2015, pp.263-277. DOI: https://doi.org/10.3354/ame017

References

1. Gomez-Gil, B. *et al.* The family vibrionaceae. *The prokaryotes* **9**, 659–747 (2014).
2. Baker-Austin, C. *et al. Vibrio* spp. infections. *Nat. Rev. Dis. Prim.* **4**, 1–19 (2018).
3. Vital, M., Füchslin, H. P., Hammes, F. & Egli, T. Growth of *Vibrio cholerae* O1 Ogawa Eltor in freshwater. *Microbiology* **153**, 1993–2001 (2007).
4. Barbieri, E. *et al.* Occurrence, diversity, and pathogenicity of halophilic *Vibrio* spp. and non-O1 *Vibrio cholerae* from estuarine waters along the Italian Adriatic coast. *Appl. Environ. Microbiol.* **65**, 2748–2753 (1999).
5. Vezzulli, L. *et al.* Benthic ecology of *Vibrio* spp. and pathogenic *Vibrio* species in a coastal Mediterranean environment (La Spezia Gulf, Italy). *Microb. Ecol.* **58**, 808–818 (2009).
6. Charles JEUNIAUX & Marie-Françoise VOSS-FOUCART. Chitin biomass and production in the marine environment. *Biochemical Systematics and Ecology* vol. 19 347–356 (1991).
7. Pruzzo, C., Vezzulli, L. & Colwell, R. R. Global impact of *Vibrio cholerae* interactions with chitin. *Environ. Microbiol.* **10**, 1400–1410 (2008).
8. Grimes, D. J. The Vibrios: scavengers, symbionts, and pathogens from the sea. *Microb. Ecol.* **80**, 501–506 (2020).
9. ZoBell, C. E., Grant, C. W. & Haas, H. F. Marine microorganisms which oxidize petroleum hydrocarbons. *Am. Assoc. Pet. Geol. Bull.* **27**, 1175–1193 (1943).
10. Okpokwasili, G. C., Somerville, C. C., Sullivan, M., Grimes, D. J. & Colwell, R. R. Plasmid mediated degradation of hydrocarbons in estuarine bacteria. *Oil Chem. Pollut.* **3**, 117–129 (1986).
11. Grimes, D. J. *et al.* What genomic sequence information has revealed about *vibrio* ecology in the ocean-a review. *Microb. Ecol.* **58**, 447–460 (2009).
12. Urdaci, M. C., Stal, L. J. & Marchand, M. Occurrence of nitrogen fixation among *Vibrio* spp. *Arch. Microbiol.* **150**, 224–229 (1988).
13. Cottrell, M. T. & David, K. L. Contribution of major bacterial groups to bacterial biomass production (thymidine and leucine incorporation) in the Delaware estuary. *Limnol. Oceanogr.* **48**, 168–178 (2003).
14. Stabb, E. V. The *Vibrio fischeri*–Euprymna scolopes light organ symbiosis. *Biol. vibrios* 204–218 (2006).
15. Grimes, D. J. *et al. Vibrio* species associated with mortality of sharks held in captivity. *Microb. Ecol.* **10**, 271–282 (1984).
16. Singleton, F. L., Attwell, R., Jangi, S. & Colwell, R. R. Effects of temperature and salinity on *Vibrio cholerae* growth. *Appl. Environ. Microbiol.* **44**, 1047–1058 (1982).
17. Randa, M. A., Polz, M. F. & Lim, E. Effects of temperature and salinity on *Vibrio vulnificus* population dynamics as assessed by quantitative PCR. *Appl. Environ. Microbiol.* **70**, 5469–5476 (2004).
18. Kelly, M. T. Effect of temperature and salinity on *Vibrio vulnificus* occurrence in a Gulf Coast environment. *Appl. Environ. Microbiol.* **44**, 820–824 (1982).
19. Namadi, P. & Deng, Z. Optimum environmental conditions controlling prevalence of *Vibrio parahaemolyticus* in marine environment. *Mar. Environ. Res.* **183**, 105828 (2023).
20. Legros, D. Global cholera epidemiology: Opportunities to reduce the burden of cholera by 2030. *J. Infect. Dis.* **218**, S137–S140 (2018).
21. Orata, F. D., Keim, P. S. & Boucher, Y. The 2010 cholera outbreak in Haiti: how science solved a controversy. *PLoS Pathog.* **10**, e1003967 (2014).
22. Sur, D., Dutta, P., Nair, G. B. & Bhattacharya, S. K. Severe cholera outbreak following floods in a northern disrict of West Bengal. *Indian J. Med. Res.* **112**, 178 (2000).
23. Jahangir Alam, M., Tomochika, K. I., Miyoshi, S. I. & Shinoda, S. Environmental investigation

of potentially pathogenic *Vibrio parahaemolyticus* in the Seto-Inland Sea, Japan. *FEMS Microbiol. Lett.* **208**, 83–87 (2002).
24. Newton, A., Kendall, M., Vugia, D. J., Henao, O. L. & Mahon, B. E. Increasing rates of vibriosis in the United States, 1996–2010: review of surveillance data from 2 systems. *Clin. Infect. Dis.* **54**, S391–S395 (2012).
25. Baker-Austin, C., Stockley, L., Rangdale, R. & Martinez-Urtaza, J. Environmental occurrence and clinical impact of *Vibrio vulnificus* and *Vibrio parahaemolyticus*: a European perspective. *Environ. Microbiol. Rep.* **2**, 7–18 (2010).
26. Jones, M. K. & Oliver, J. D. *Vibrio vulnificus*: Disease and pathogenesis. *Infect. Immun.* **77**, 1723–1733 (2009).
27. Klontz, K. C. *et al.* Syndromes of *Vibrio vulnificus* infections: clinical and epidemiologic features in Florida cases, 1981-1987. *Ann. Intern. Med.* **109**, 318–323 (1988).
28. Dechet, A. M., Yu, P. A., Koram, N. & Painter, J. Nonfoodborne *Vibrio* infections: An important cause of morbidity and mortality in the United States, 1997-2006. *Clin. Infect. Dis.* **46**, 970–976 (2008).
29. Hlady, W. G. & Klontz, K. C. The epidemiology of *Vibrio* infections in Florida, 1981–1993. *J. Infect. Dis.* **173**, 1176–1183 (1996).
30. Chuang, Y.-C., Yuan, C.-Y., Liu, C.-Y., Lan, C.-K. & Huang, A. H.-M. *Vibrio vulnificus* infection in Taiwan: report of 28 cases and review of clinical manifestations and treatment. *Clin. Infect. Dis.* **15**, 271–276 (1992).
31. Oliver, J. D. *Vibrio vulnificus*. *Biol. vibrios* 349–366 (2006).
32. Oliver, J. D. The biology of *Vibrio vulnificus*. *Microbiol. Spectr.* **3**, 3 (2015).
33. Aznar, R., Ludwig, W., Amann, R. I. & Schleifer, K. H. Sequence determination of rRNA genes of pathogenic *Vibrio* species and whole-cell identification of *Vibrio vulnificus* with rRNA-targeted oligonucleotide probes. *Int. J. Syst. Bacteriol.* **44**, 330–337 (1994).
34. Baffone, W. *et al.* Detection of free-living and plankton-bound vibrios in coastal waters of the Adriatic Sea (Italy) and study of their pathogenicity-associated properties. *Environ. Microbiol.* **8**, 1299–1305 (2006).
35. DePaola, A., Capers, G. M. & Alexander, D. Densities of *Vibrio vulnificus* in the intestines of fish from the US Gulf Coast. *Appl. Environ. Microbiol.* **60**, 984–988 (1994).
36. Ralston, E. P., Kite-Powell, H. & Beet, A. An estimate of the cost of acute health effects from food-and water-borne marine pathogens and toxins in the USA. *J. Water Health* **9**, 680–694 (2011).
37. Zorn, D. J., Anderson, D. W., Lindrooth, R. C., Murray, B. C. & Teague, J. L. Cost of restrictions on gulf oyster harvesting for control of *Vibrio vulnificus*-Caused Disease. (1996).
38. Sheahan, M. *et al.* Examining the relationship between climate change and vibriosis in the United States: projected health and economic impacts for the 21st century. *Environ. Health Perspect.* **130**, 87007 (2022).
39. Haenen, O. L. M. *et al. Vibrio vulnificus* outbreaks in Dutch eel farms since 1996: Strain diversity and impact. *Dis. Aquat. Organ.* **108**, 201–209 (2014).
40. Tey, Y. H., Jong, K.-J., Fen, S.-Y. & Wong, H.-C. Occurrence of *Vibrio parahaemolyticus*, *Vibrio cholerae*, and *Vibrio vulnificus* in the aquacultural environments of Taiwan. *J. Food Prot.* **78**, 969–976 (2015).
41. Fouz, B., Larsen, J. L. & Amaro, C. *Vibrio vulnificus* serovar A: an emerging pathogen in European anguilliculture. *J. Fish Dis.* **29**, 285–291 (2006).
42. Biosca, E. G., Amaro, C., Esteve, C., Alcaide, E. & Garay, E. First record of *Vibrio vulnificus* biotype 2 from diseased European eel, Anguilla anguilla L. *J. Fish Dis.* **14**, 103–109 (1991).
43. Sumithra, T. G. *et al.* Pathological investigations of *Vibrio vulnificus* infection in genetically improved farmed tilapia (*Oreochromis niloticus L.*) cultured at a floating cage farm of India. *Aquaculture* **511**, 734217 (2019).

44. Baker-Austin, C. & Oliver, J. D. *Vibrio vulnificus*: new insights into a deadly opportunistic pathogen. *Environ. Microbiol.* **20**, 423–430 (2018).
45. Bross, M. H., Soch, K., Morales, R. & Mitchell, R. B. *Vibrio vulnificus* infection: diagnosis and treatment. *Am. Fam. Physician* **76**, 539–544 (2007).
46. Belkin, S. & Colwell, R. R. *Oceans and health: pathogens in the marine environment*. vol. 1 (Springer, 2006).
47. Oliver, J. D. Wound infections caused by *Vibrio vulnificus* and other marine bacteria. *Epidemiol. Infect.* **133**, 383–391 (2005).
48. Bisharat, N. *et al.* Hybrid *Vibrio vulnificus*. *Emerg. Infect. Dis.* **11**, 30 (2005).
49. Gray, L. D. & Kreger, A. S. Detection of *Vibrio vulnificus* cytolysin in *V. vulnificus*-infected mice. *Toxicon* **27**, 459–464 (1989).
50. Oliver, J. D., Wear, J. E., Thomas, M. B., Warner, M. & Linder, K. Production of extracellular enzymes and cytotoxicity by *Vibrio vulnificus*. *Diagn. Microbiol. Infect. Dis.* **5**, 99–111 (1986).
51. Simpson, L. M. & Oliver, J. D. Ability of *Vibrio vulnificus* to obtain iron from transferrin and other iron-binding proteins. *Curr. Microbiol.* **15**, 155–157 (1987).
52. Kreger, A., DeChatelet, L. & Shirley, P. Interaction of *Vibrio vulnificus* with human polymorphonuclear leukocytes: association of virulence with resistance to phagocytosis. *J. Infect. Dis.* **144**, 244–248 (1981).
53. McPherson, V. L., Watts, J. A., Simpson, L. M. & Oliver, J. D. Physiological effects of the lipopolysaccharide of *Vibrio vulnificus* on mice and rats. *Microbios* **67**, 141–149 (1991).
54. Huang, K.-C. *et al.* Distribution of fatal *Vibrio vulnificus* necrotizing skin and soft-tissue infections: a systematic review and meta-analysis. *Medicine (Baltimore).* **95**, (2016).
55. Baker-Austin, C. *et al.* Emerging *Vibrio* risk at high latitudes in response to ocean warming. *Nat. Clim. Chang.* **3**, 73–77 (2013).
56. Mora, C. *et al.* Over half of known human pathogenic diseases can be aggravated by climate change. *Nat. Clim. Chang.* **12**, 869–875 (2022).
57. Archer, E. J. *et al.* Climate warming and increasing *Vibrio vulnificus* infections in North America. *Sci. Rep.* **13**, 1–11 (2023).
58. Eiler, A., Gonzalez-Rey, C., Allen, S. & Bertilsson, S. Growth response of *Vibrio cholerae* and other *Vibrio* spp. to cyanobacterial dissolved organic matter and temperature in brackish water. *FEMS Microbiol. Ecol.* **60**, 411–418 (2007).
59. Lamb, J. B. *et al.* Seagrass ecosystems reduce exposure to bacterial pathogens of humans, fishes, and invertebrates. *Science (80-.).* **355**, 731–733 (2017).
60. Reusch, T. B. H. *et al.* Lower *Vibrio* spp. abundances in *Zostera marina* leaf canopies suggest a novel ecosystem function for temperate seagrass beds. *Front. Mar. Sci.* (2021).
61. Fonseca, M. S., Fisher, J. S., Zieman, J. C. & Thayer, G. W. Influence of the seagrass, *Zostera marina L.*, on current flow. *Estuar. Coast. Shelf Sci.* **15**, 351–364 (1982).
62. Worcester, S. E. Effects of eelgrass beds on advection and turbulent mixing in low current and low shoot density environments. *Mar. Ecol. Prog. Ser.* **126**, 223–232 (1995).
63. Nowell, A. R. M. & Jumars, P. A. Flow environments of aquatic benthos. *Annu. Rev. Ecol. Syst.* **15**, 303–328 (1984).
64. Peterson, B. J. & Heck Jr, K. L. Positive interactions between suspension-feeding bivalves and seagrass a facultative mutualism. *Mar. Ecol. Prog. Ser.* **213**, 143–155 (2001).
65. Gonzalez-Ortiz, V. *et al.* Interactions between seagrass complexity, hydrodynamic flow and biomixing alter food availability for associated filter-feeding organisms. *PLoS One* **9**, e104949 (2014).
66. Papazian, S., Parrot, D., Burýšková, B., Weinberger, F. & Tasdemir, D. Surface chemical defence of the eelgrass *Zostera marina* against microbial foulers. *Sci. Rep.* **9**, 1–12 (2019).
67. Reusch, T. B. H. *et al.* The Baltic Sea as a time machine for the future coastal ocean. *Sci. Adv.* **4**, eaar8195 (2018).

68. Markus Meier, H. E. *et al.* Climate change in the Baltic Sea region: A summary. *Earth Syst. Dyn.* **13**, 457–593 (2022).
69. Snoeijs-Leijonmalm, P., Schubert, H. & Radziejewska, T. *Biological oceanography of the Baltic Sea 683 pp.* (Springer Science & Business Media, 2017).
70. Zvidriņš, P. Demographic development in the Baltic Sea Region. *Latv. Zinātņu Akadēmijas Vēstis* **A66**, 49–61 (2012).
71. Nilsson, J. H. & Gössling, S. Tourist responses to extreme environmental events: The case of Baltic Sea algal blooms. *Tour. Plan. Dev.* **10**, 32–44 (2013).
72. Reissmann, J. H. *et al.* Vertical mixing in the Baltic Sea and consequences for eutrophication - A review. *Prog. Oceanogr.* **82**, 47–80 (2009).
73. Jones, J. L. *et al.* Comparison of molecular detection methods for *Vibrio parahaemolyticus* and *Vibrio vulnificus. Food Microbiol.* **30**, 105–111 (2012).
74. Donovan, T. J. & Van Netten, P. Culture media for the isolation and enumeration of pathogenic *Vibrio* species in foods and environmental samples. *Prog. Ind. Microbiol.* **34**, 203–217 (1995).
75. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
76. Hugerth, L. W. & Andersson, A. F. Analysing microbial community composition through amplicon sequencing: from sampling to hypothesis testing. *Front. Microbiol.* **8**, 1561 (2017).
77. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1–e1 (2013).
78. Van Gulik, W. M. Fast sampling for quantitative microbial metabolomics. *Curr. Opin. Biotechnol.* **21**, 27–34 (2010).
79. Hughes, J. B., Hellmann, J. J., Ricketts, T. H. & Bohannan, B. J. M. Counting the uncountable: statistical approaches to estimating microbial diversity. *Appl. Environ. Microbiol.* **67**, 4399–4406 (2001).
80. Choo, J. M., Leong, L. E. X. & Rogers, G. B. Sample storage conditions significantly influence faecal microbiome profiles. *Sci. Rep.* **5**, 16350 (2015).
81. Song, S. J. *et al.* Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. *MSystems* **1**, 10–1128 (2016).
82. Bonk, F., Popp, D., Harms, H. & Centler, F. PCR-based quantification of taxa-specific abundances in microbial communities: Quantifying and avoiding common pitfalls. *J. Microbiol. Methods* **153**, 139–147 (2018).
83. Weiss, S. *et al.* Tracking down the sources of experimental contamination in microbiome studies. *Genome Biol.* **15**, 1–3 (2014).
84. Schirmer, M. *et al.* Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res.* **43**, e37–e37 (2015).
85. Kembel, S. W., Wu, M., Eisen, J. A. & Green, J. L. Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. *PLoS Comput. Biol.* **8**, e1002743 (2012).
86. Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **17**, 333–351 (2016).
87. Tremblay, J. *et al.* Primer and platform effects on 16S rRNA tag sequencing. *Front. Microbiol.* **6**, 771 (2015).
88. Golob, J. L., Margolis, E., Hoffman, N. G. & Fredricks, D. N. Evaluating the accuracy of amplicon-based microbiome computational pipelines on simulated human gut microbial communities. *BMC Bioinformatics* **18**, 1–12 (2017).
89. Marizzoni, M. *et al.* Comparison of bioinformatics pipelines and operating systems for the analyses of 16S rRNA gene amplicon sequences in human fecal samples. *Front. Microbiol.* **11**, 1262 (2020).
90. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data.

Nat. Methods **13**, 581–583 (2016).
91. Edgar, R. C. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv* 081257 (2016).
92. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome datasets are compositional: and this is not optional. *Front. Microbiol.* **8**, 2224 (2017).
93. Janßen, R. *et al.* Machine Learning Predicts the Presence of 2,4,6-Trinitrotoluene in Sediments of a Baltic Sea Munitions Dumpsite Using Microbial Community Compositions. *Front. Microbiol.* **12**, 1–18 (2021).
94. Armstrong, G. *et al.* Applications and comparison of dimensionality reduction methods for microbiome data. *Front. Bioinforma.* **2**, 821861 (2022).
95. Bzdok, D. & Naomi, M. K. Statistics versus machine learning. *Nat Methods* **15**, 233 (2018).
96. Breiman, L. E. O. Random Forests. 5–32 (2001).
97. Jordan, M. I. & Mitchell, T. M. Machine learning: Trends, perspectives, and prospects. *Science (80-.).* **349**, 255–260 (2015).
98. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
99. Berrar, D. Cross-Validation. (2019).
100. Santos, C. F. G. Dos & Papa, J. P. Avoiding overfitting: A survey on regularization methods for convolutional neural networks. *ACM Comput. Surv.* **54**, 1–25 (2022).
101. Smialowski, P., Frishman, D. & Kramer, S. Pitfalls of supervised feature selection. *Bioinformatics* **26**, 440–443 (2010).
102. Hu, Q., Yu, D., Xie, Z. & Li, X. EROS: Ensemble rough subspaces. *Pattern Recognit.* **40**, 3728–3739 (2007).
103. Oshiro, T. M., Perez, P. S. & Baranauskas, J. A. How many trees in a random forest? in *Machine Learning and Data Mining in Pattern Recognition: 8th International Conference, MLDM 2012, Berlin, Germany, July 13-20, 2012. Proceedings 8* 154–168 (Springer, 2012).
104. Altmann, A., Toloşi, L., Sander, O. & Lengauer, T. Permutation importance: a corrected feature importance measure. *Bioinformatics* **26**, 1340–1347 (2010).
105. Janitza, S., Celik, E. & Boulesteix, A.-L. A computationally fast variable importance test for random forests for high-dimensional data. *Adv. Data Anal. Classif.* **12**, 885–915 (2018).
106. Nembrini, S., König, I. R. & Wright, M. N. The revival of the Gini importance? *Bioinformatics* **34**, 3711–3718 (2018).
107. Gregorutti, B., Michel, B. & Saint-Pierre, P. Correlation and variable importance in random forests. *Stat. Comput.* **27**, 659–678 (2017).
108. Kinley, R. Climate change after Paris: from turning point to transformation. *Clim. Policy* **17**, 9–15 (2017).
109. Lunetta, R. *et al.* Remote sensing and geographic information system data integration- Error sources and research issues. *Photogramm. Eng. Remote Sensing* **57**, 677–687 (1991).
110. Williams, T. C. *et al.* Different abundance and correlational patterns exist between total and presumed pathogenic *Vibrio vulnificus* and *V. parahaemolyticus* in shellfish and waters along the North Carolina coast. *FEMS Microbiol. Ecol.* **93**, 1–11 (2017).
111. Lotz, M. J., Tamplin, M. L. & Rodrick, G. E. Thiosulfate-citrate-bile salts-sucrose agar and its selectivity for clinical and marine *Vibrio* organisms. *Ann. Clin. Lab. Sci.* **13**, 45–48 (1983).
112. Williams, T. C., Froelich, B. & Oliver, J. D. A new culture-based method for the improved identification of *Vibrio vulnificus* from environmental samples, reducing the need for molecular confirmation. *J. Microbiol. Methods* **93**, 277–283 (2013).
113. Nejrup, L. B. & Pedersen, M. F. Effects of salinity and water temperature on the ecological performance of *Zostera marina*. *Aquat. Bot.* **88**, 239–246 (2008).
114. Conley, D. J. *et al.* Hypoxia-related processes in the Baltic Sea. *Environ. Sci. Technol.* **43**, 3412–3420 (2009).
115. Munkes, B., Löptien, U. & Dietze, H. Cyanobacteria blooms in the Baltic Sea: A review of

models and facts. *Biogeosciences* **18**, 2347–2378 (2021).
116. MacKenzie, B. R., Hinrichsen, H.-H., Plikshs, M., Wieland, K. & Zezera, A. S. 2000. Quantifying environmental heterogeneity: habitat size necessary for successful development of cod *Gadus morhua* eggs in the Baltic Sea. *Mar. Ecol. Prog. Ser.* **193**, 143–156 (2000).
117. Svendsen, L. & Gustafsson, B. G. Waterborne nitrogen and phosphorus inputs and water flow to the Baltic Sea 1995-2018. *HELCOM Balt. Sea Environ. Fact Sheet 2020* 1–25 (2020).
118. Svendsen, L. M. *et al.* Updated fifth baltic sea pollution load compilation. *Balt. Sea Environ. Proc.* 65–71 (2015).
119. Kuliński, K. *et al.* Biogeochemical functioning of the Baltic Sea. *Earth Syst. Dyn.* **13**, 633–685 (2022).
120. Stigebrandt, A. & Gustafsson, B. G. Response of the Baltic Sea to climate change - Theory and observations. *J. Sea Res.* **49**, 243–256 (2003).
121. Gustafsson, E., Savchuk, O. P., Gustafsson, B. G. & Müller-Karulis, B. Key processes in the coupled carbon, nitrogen, and phosphorus cycling of the Baltic Sea. *Biogeochemistry* **134**, 301–317 (2017).
122. Rydin, E., Kumblad, L., Wulff, F. & Larsson, P. Remediation of a Eutrophic Bay in the Baltic Sea. *Environ. Sci. Technol.* **51**, 4559–4566 (2017).
123. Banakar, V. *et al.* Temporal and spatial variability in the distribution of *Vibrio vulnificus* in the Chesapeake Bay: a hindcast study. *Ecohealth* **8**, 456–467 (2011).
124. Ruhl, H. A. & Rybicki, N. B. Long-term reductions in anthropogenic nutrients link to improvements in Chesapeake Bay habitat. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 16566–16570 (2010).

Scientific outreach

• Publications in peer-reviewed journals

Riedinger, D.J., Fernández-Juárez, V., Delgado, L.F. *et al*. Control of *Vibrio vulnificus* proliferation in the Baltic Sea through eutrophication and algal bloom management. *Commun Earth Environ* **5**, 246 (2024). https://doi.org/10.1038/s43247-024-01410-x

Glackin C.C, Dupke S, Chandra T.S, **Riedinger D+**, Labrenz M. Combined TCBS and CHROMagar analyses allow for basic identification of *Vibrio vulnificus* within a 48 h incubation period in the coastal Baltic Sea. *Microorganisms*. 2024; 12(3):614. https://doi.org/10.3390/microorganisms12030614

• Manuscript submitted

Riedinger, D.J., Hassenrück C., Herlemann D., Labrenz M. Poleward spread, environmental modulation, and predictive modeling of global *Vibrio vulnificus* abundance– Submitted to Communications Earth & Environment

V. Fernández-Juárez, **D. J. Riedinger**, J. B. Gusmao, L. F Delgado-Zambrano, G. Coll-García, V. Papazachariou, D.P.R. Herlemann, C. Pansch, A. F. Andersson, M. Labrenz, L. Riemann. Temperature, sediment resuspension, and salinity drive the prevalence of *Vibrio vulnificus* in the coastal Baltic Sea – Accepted in mBio

D.P. R. Herlemann, L. F. Delgado, **D. Riedinger**, V. Fernández Juárez, K. Käiro, A. F. Andersson, C. Pansch-Hattig, L. Riemann, M. Bengtsson, G. Gyraitė, M. Kataržytė, V. Kisand, S. Kube, G. Martin, K. Piwosz, M. Labrenz. Low impact of seagrass on sediment and water microbial communities under brackish conditions – Submitted to BMC microbiology: environmental microbiome

• Datasets and code

16S and 18S rRNA amplicon sequencing of coastal microbial communities across the salinity gradient of the Baltic Sea in pelagic, benthic, and biofilm environments. **ENA: PRJEB68222**

Microbial composition data collected from water and sediment samples in Denmark, Germany, and Finland over a period of time from May to October 2022**. NCBI: PRJNA1011541**

The occurrence of *Vibrio* spp. in the salinity gradient of shallow coastal waters of the Baltic Sea – data set including environmental and microbiological data - **doi.io-warnemuende.de/10.12754/data-2023-0010**

Code for the machine learning performed in Baltvib - https://git.io-warnemuende.de/riedinge/Baltvib_RF_RFE.git

Data mining project to track proliferation of *V. vulnificus* using ENA amplicon Data - https://git.io-warnemuende.de/riedinge/datamining_vulnificus_DR_CH.git

• Scientific talks

Vibrio vulnificus can be impacted by regional measures along the Baltic Sea coast – BSSC 2023 / Award for best young scientist presentation

Author contribution

The four scientific papers that are part of this thesis are either accepted or submitted to peer-reviewed journals.

Chapter I: D. Riedinger and C. Hassenrück conceived of the idea and performed the bioinformatics analysis. D. Riedinger performed the remote sensing analysis and the machine learning and wrote the manuscript. (Submitted)

Chapter II: D. Riedinger contributed to the formal analysis, sampling, data curation and reviewed and edited the first draft written by C. Glackin and all subsequent versions. (Published)

Chapter III: D. Riedinger contributed to the formal analysis, sampling, data curation and writing of the original draft and all further versions. (Accepted)

Chapter IV: D. Riedinger co-designed the project, contributed to the sampling and the laboratory work, performed the data analysis and the machine learning. D. Riedinger wrote the original draft and incorporated all comments. (Published)

....

(Place) (Date) (Signature of the author)

Rostock, 25.07.2024

....

(Place) (Date) (Signature of the supervisor)

Declaration of authorship

I hereby declare the following:

1. The opportunity for this PhD project was not communicated commercially to me. In particular, I have not engaged any organization that, for money, seeks supervisors for the drawing up of dissertations or that performs, entirely or partially on my behalf, the duties incumbent upon me regarding the examinations.

2. I hereby declare under oath that I have completed the work submitted here independently and have composed it without outside assistance. Furthermore, I have not used anything other than the resources and sources stated. In cases where I have incorporated content or text from these works, I have duly acknowledged it.

....

(Place) (Date) (Signature of the author)

Acknowledgement

First of all, I would like to sincerely thank my supervisor Matthias Labrenz for the chance to foray into microbiology from my geochemical background. There was, and still is, a whole world to be explored. I am appreciative of the continuous support and acknowledgment I have received, as well as the guidance provided. Additionally, the fieldwork in Lithuania was a positive and rewarding experience on a personal level. Christiane, I extend a heartfelt thank you for your patience and support in gradually moving me from the six slightly different copies of barely functional R scripts to GitHub, the command line, and Bash. Your ability to swallow your frustration is remarkable, and your commitment to excellence is both admirable and inspiring. Theo, thank you for the discussions, not only on machine learning but also on presentation and writing styles, scientific goals, and what it means to truly understand something. Additional thanks for proofreading my thesis. Solveig, thank you for guiding me through the same administrative maze every other month without tiring. Heike, thank you for the consistent support in the lab and mostly in the café.

Further acknowledgments, in no particular order, go out to my friends Connor, Conor, Erik, Nadya, Gizem, Anna, Kaylim, Gabriela, Paul, and Josephine, who made Rostock more colorful. To my friends in Zürich, Leiden, The Hague, and Utrecht, thank you as well.

Above all, thank you, Noémie. You’re the reason I pursued this PhD, and it honestly would not have been any fun without you.

Riedinger, David | Microbiologist / Geochemist

d.j.riedinger@gmail.com • Koraal 1, 222KL, Katwijk • (+31) 624187566

Education

Ph.D. Microbiology, IOW; Rostock University

04/2021 – 12/2024

Rostock, Germany

Thesis: *Predictors of* Vibrio vulnificus *occurence: A machine learning approach.*

M.Sc. Earth, Life and Climate, UU

09/2018-08/2020

Utrecht, the Netherlands

Geochemistry and kinetic processes

Thesis: *Barite nucleation - Measuring effects of solution stoichiometry on crystal population and critical nucleii size*

B.Sc. Earth sciences, UU, UoL

09/2015-06/2018

Utrecht, the Netherlands ; Leicester, UK

Paleoclimatology and sedimentology

Experience

Supervision Bacheror Thesis o Quantifizierung von *Vibrio vulnificus* an ausgewählten Ostseestandorten mittels droplet digital PCR	2022 Rostock

Fieldwork Baltvib o Fieldwork planning o Scientific diving o DNA/RNA/Nutrient sampling o Plating o FISH filtration	2021 Baltic Sea
Cruise EMB283 o DNA/RNA/Nutrient sampling o Station planning o Plating o FISH filtration	2021 Baltic Sea
Internship Witteveen & Bos o Modelling of CO_2 injection in thermal wells o Experimental and modelling work on passive remediation solutions for alkaline waters	2020 Den Haag
Alba Synchrotron 2020024166 - NCD-SWEET (BL 11) o SAXS – Barite nucleation o WAXS – Barite nucleation	2021 Barcelona
Cruise North Sea Monitoring Leg 1 – RV Pelagia o Boxcore - sediment fluxes o Multicore sampling	2019 North Sea

Cruise NIOZ-NICO Leg 1 – RV Pelagia	2018
○ Inorganic nutrient measurements ○ Sediment sampling ○ Organic nutrient sampling ○ Zooplankton net tows	Texel → Las Palmas

Skills

Lab

Water nutrient analysis | DNA Extraction + PCR | Plating

IT

R [Basic stats | Clustering | Machine learning (caret package) | Modelling] | Bioinformatics | Modelling | Bash | Datamining

Languages

German (native) | Dutch (native) | English (fluent) | French (conversational)

Additional

Driving license | Scientific diving license

Scientific outreach

• Publications in peer-reviewed journals

Riedinger, D.J., Fernández-Juárez, V., Delgado, L.F. *et al*. Control of *Vibrio vulnificus* proliferation in the Baltic Sea through eutrophication and algal bloom management. *Commun Earth Environ* **5**, 246 (2024). https://doi.org/10.1038/s43247-024-01410-x

Glackin C.C, Dupke S, Chandra T.S, **Riedinger D**, Labrenz M. Combined TCBS and CHROMagar analyses allow for basic identification of *Vibrio vulnificus* within a 48 h incubation period in the coastal Baltic Sea. *Microorganisms*. 2024; 12(3):614. https://doi.org/10.3390/microorganisms12030614

Riedinger, D.J., C. Hassenrück, M. Labrenz. Poleward spread, environmental modulation, and predictive modeling of global *Vibrio vulnificus* abundance– Submitted to Nature Microbiology

V. Fernández-Juárez, **D. J. Riedinger**, J. B. Gusmao, L. F Delgado-Zambrano, G. Coll-García, V. Papazachariou, D.P.R. Herlemann, C. Pansch, A. F. Andersson, M. Labrenz, L. Riemann. Temperature, sediment resuspension, and salinity drive the prevalence of *Vibrio vulnificus* in the coastal Baltic Sea – Submitted to mBio

D.P. R. Herlemann, L. F. Delgado, **D. Riedinger**, V. Fernández Juárez, K. Käiro, A. F. Andersson, C. Pansch-Hattig, L. Riemann, M. Bengtsson, G. Gyraitė, M. Kataržytė, V. Kisand, S. Kube, G. Martin, K. Piwosz, M. Labrenz. Low impact of seagrass on sediment and water microbial communities under brackish conditions – Submitted to BMC microbiology: environmental microbiome

• Datasets and code

16S and 18S rRNA amplicon sequencing of coastal microbial communities across the salinity gradient of the Baltic Sea in pelagic, benthic, and biofilm environments. **ENA: PRJEB68222**

Microbial composition data collected from water and sediment samples in Denmark, Germany, and Finland over a period of time from May to October 2022**. NCBI: PRJNA1011541**

The occurrence of *Vibrio* spp. in the salinity gradient of shallow coastal waters of the Baltic Sea – data set including environmental and microbiological data - **doi.io-warnemuende.de/10.12754/data-2023-0010**

Code for the machine learning performed in Baltvib - https://git.io-warnemuende.de/riedinge/Baltvib_RF_RFE.git

Data mining project to track proliferation of *V. vulnificus* using ENA amplicon Data - https://git.io-warnemuende.de/riedinge/datamining_vulnificus_DR_CH.git

• Scientific talks

Vibrio vulnificus can be impacted by regional measures along the Baltic Sea coast – BSSC 2023 / Price for best young scientist presentation