Systematic and Applied Microbiology 38 (2015) 16–25

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.de/syapm

# VibrioBase: A MALDI-TOF MS database for fast identification of *Vibrio spp.* that are potentially pathogenic in humans

René Erler[a,*], Antje Wichels[a], Ernst-August Heinemeyer[b], Gerhard Hauk[c], Martin Hippelein[d], Nadja Torres Reyes[e], Gunnar Gerdts[a]

[a] *Alfred Wegener Institute for Polar and Marine Research, Biologische Anstalt Helgoland, Kurpromenade 201, 27498 Helgoland, Germany*
[b] *Governmental Institute of Public Health of Lower Saxony, Lüchtenburger Weg 24, 26603 Aurich, Germany*
[c] *Regional Office for Health and Social Affairs of Mecklenburg-Western Pomerania, Gertrudenstraße 11, 18057 Rostock, Germany*
[d] *University Medical Center Schleswig-Holstein, Central Facility: Medical Investigation Office and Hygiene, Brunswiker Straße 4, 24105 Kiel, Germany*
[e] *GFZ German Research Centre for Geosciences, Helmholtz Centre Potsdam, Telegrafenberg, 14473 Potsdam, Germany*

ARTICLE INFO

*Article history:*
Received 21 April 2014
Received in revised form 20 October 2014
Accepted 27 October 2014

*Keywords:*
*Vibrio*
MALDI-TOF MS
Database
Whole-cell fingerprinting
Marine
Pathogenic

ABSTRACT

Mesophilic marine bacteria of the family *Vibrionaceae*, specifically *V. cholerae*, *V. parahaemolyticus* and *V. vulnificus*, are considered to cause severe illness in humans. Due to climate-change-driven temperature increases, higher *Vibrio* abundances and infections are predicted for Northern Europe, which in turn necessitates environmental surveillance programs to evaluate this risk. We propose that whole-cell matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) profiling is a promising tool for the fast and reliable species classification of environmental isolates. Because the reference database does not contain sufficient *Vibrio* spectra we generated the VibrioBase database in this study. Mass spectrometric data were generated from 997 largely environmental strains and filed in this new database. MALDI-TOF MS clusters were assigned based on the species classification obtained by analysis of partial *rpoB* (RNA polymerase beta-subunit) sequences. The affiliation of strains to species-specific clusters was consistent in 97% of all cases using both approaches, and the extended VibrioBase generated more specific species identifications with higher matching scores compared to the commercially available database. Therefore, we have made the VibrioBase database freely accessible, which paves the way for detailed risk assessment studies of potentially pathogenic *Vibrio spp.* from marine environments.


## Introduction

Vibrios are a heterogeneous group within the class Gammaproteobacteria and are typically found in marine and coastal environments throughout the world [27]. Well-known features of *Vibrio* bacteria include the formation of biofilms in locations such as the surfaces of marine eukaryotes and the accumulation in filtering organisms such as blue mussels and oysters [42,78]. Some species are pathogenic in marine invertebrates, fish and mammals [5]. The main species associated with gastrointestinal illness, wound infection and septicemia in humans are *Vibrio parahaemolyticus*, *Vibrio vulnificus* and *Vibrio cholerae* [21]. The incidence of *Vibrio* infections can be increased by anthropogenic changes in the environment. One potential risk is the increasing marine traffic whereby epidemic strains can spread rapidly via ballast water [25,26]. Another serious risk addressed recently by the IPCC is global warming, with a predicted temperature rise of at least 1.8 °C by 2100 [33]. Two climate change-related aspects are to be considered: the spread and frequency of marine infectious diseases [30] as well as the positive correlation between potentially pathogenic *Vibrio spp.* and the seawater temperature [10,51,72]. This increase applies mainly to non-cholerae *Vibrio spp.* yet not considered a priority and not included in surveillance programs [7,43,55]. In particular increasing seawater temperatures in the form of heat waves [76] might lead to a higher incidence of non-cholerae-related *Vibrio spp.* infections [8]. Therefore, it is important to screen temperate waters for potentially pathogenic *Vibrio spp.*, which would benefit the analysis of human health risks.

Sensitive and specific diagnostic tools for autonomous monitoring programs would facilitate such analyses [14]. Culture-independent methods, such as quantitative real-time PCR, are important in such studies and are already used in *Vibrio* population analyses [53,57,67]. The disadvantage of these methods is that they rely on the DNA sequences of the target genes. This

* Corresponding author. Tel.: +49 4725 8193273; fax: +49 4725 8193283.
*E-mail address:* Rene.Erler@awi.de (R. Erler).

http://dx.doi.org/10.1016/j.syapm.2014.10.009
0723-2020/

dependence necessitates the design of species-specific primers for these target genes and the elimination of cross-reactions [65,76]. If these obstacles can be overcome, culture-independent methods can be used in next-generation marine monitoring programs [11].

However, culture-dependent methods are currently used in routine surveillance programs because these methods involve well-established techniques such as biochemical identification systems and the enumeration of colony forming units [68]. Biochemical methods have been used extensively for the classification of *Vibrio spp.* [4,49]; however, the accuracy of commercially available test systems is limited due to the high phenotypic diversity of *Vibrio spp.* [50]. Therefore, new molecular identification methods based on conventional culture-dependent techniques would eliminate false-positive identification results [20]. Oberbeckmann et al. used this approach and proposed polyphasic molecular species identification for *Vibrio* population analyses [52]. This study confirms earlier observations that the 16S RNA ribosomal gene sequence-based method lacks sufficient resolution for the accurate identification of closely related *Vibrio spp.* and partial nucleotide sequence data of the functional gene *rpoB* is more suitable for identification of environmental *Vibrio* strains at the species level [37,69]. This is of particular importance for monitoring programs: species that are closely related but nonetheless differ with regard to their pathogenicity can be distinguished using *rpoB* sequence analysis; making risk assessment reports more accurate. In this DNA-based method, sequence similarities lower than 85% indicate different genera, and sequence similarities between 96.5% and 98% can be used as species classification cut-offs [2,35]. However, DNA isolation and sequencing require significant effort, which precludes the use of this method in large-scale monitoring programs.

Whole-cell matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) profiling allows a high sample throughput and measurements require less effort and cost [28]. This technique has been used to distinguish species of the genus *Vibrio* [23,31] and species of other taxa despite high 16S rRNA similarity [56]. The key aspect of this method is the rapid generation and comparison of mass spectra. The peaks of these mass spectra represent abundant cellular proteins. These proteins are separated during the passage through a flight tube in which the time-of-flight is equal to the mass to charge ($m/z$) ratio of each protein. Comparable mass spectra with species-specific peaks were initially generated by Claydon et al. [18], Holland et al. [32] and Krishnamurthy et al. [39]. The quality of these spectra was improved by the optimization of sample preparation procedure and matrix composition [77].

Bruker Daltonics Inc. used this technique to develop the Biotyper™ system, which generates mass spectra from the biomass of bacterial colonies within minutes, followed by a calculation of matching scores with reference spectra filed in a database [46]. According to the manufacturer, species identification results with matching scores higher than 2.0 are probable and those over 2.3 highly probable, whereas scores between 1.7 and 2.0 only allow accurate identifications at the genus level. Bruker Daltonics Inc. provides a large Biotyper™ reference database for species classification. Each database entry contains a main spectrum generated from at least 20 summarized and processed single spectra of individual bacterial strains. This database was mainly generated for medical laboratory diagnostics because the main spectra largely correspond to infective bacterial species isolated from clinical samples. Therefore, this database lacks coverage of environmental strains, which has led to limited ecological studies based on whole-cell MALDI-TOF MS-data.

The current Biotyper™ database (Version 3.3.1.0) contains 94 *Vibrionaceae* main spectra, including seven *V. parahaemolyticus* and five *V. vulnificus* but no *V. cholerae* main spectra. Database extension with additional main spectra can increase the matching score and the discriminatory power of MALDI-TOF MS-based classification [15,17,40,66]. In this study, we focus on the applicability of MALDI-TOF MS in monitoring programs by including a total of 997 largely environmental *Vibrio* strains originating from different sampling sites in the North and Baltic Seas. The goal of this study was to expand the Biotyper™ database with these additional strains for the analysis of *Vibrio spp.* in environmental samples. Therefore, we used a two-fold approach in which we acquired data using MALDI-TOF MS and *rpoB* sequence analysis. The *rpoB* sequences were used to assign novel MALDI-TOF MS main spectra to *Vibrio* species because these sequences have been extensively used to accurately differentiate *Vibrio* bacteria. The extended MALDI-TOF MS VibrioBase was examined by comparing the MALDI-TOF MS and *rpoB* data, which includes verification of species identifications by cluster analyses. VibrioBase was also compared with the Biotyper™ database to test whether the MALDI-TOF MS species matching scores were improved. In addition, *V. alginolyticus* strains were analyzed in detail to test whether MALDI-TOF MS can distinguish intraspecific groups.

## Materials and methods

### Strains

Table S-1 in the supplementary part provides an overview of the strains used in this study, their sampling locations, origins and times. All 1036 environmental strains originated from surveillance programs, performed from 2001 to 2012 in Western and Northern Europe. The strains were isolated from water, plankton, seafood and sediment. The environmental *Vibrio spp.* strains were kindly provided by the Governmental Institute of Public Health of Lower Saxony (NLGA), Regional Office for Health and Social Affairs of Mecklenburg-Western Pomerania (LAGuS), Centre for Environment, Fisheries & Aquaculture Science (CEFAS), National Institute for Public Health and the Environment (RIVM) and Lower Saxony State Office for Consumer Protection and Food Safety (LAVES). Other strains were isolated from environmental samples obtained during a survey supervised by the University Medical Center Schleswig-Holstein (UKSH) and during a research cruise to the North and Baltic Seas. Strains from an in-house culture collection of the Alfred Wegner Institute were also included. Type strains and clinical strains were provided by the university hospital in Dresden and the hospital in Bremerhaven-Reinkenheide, or purchased from the German Collection of Microorganisms and Cell Cultures (DSMZ).

### Cultivation

All strains were grown on marine broth 2216 medium [41], containing 50% or 75% seawater (1.6% or 2.4% sodium chloride). The environmental strains were incubated at 37 °C. The type strains were incubated at the temperature recommended by the DSMZ (http://www.dsmz.de/).

### *rpoB* sequence analysis for species identification

*rpoB* fragments were amplified using the primers rpoB458F and rpoB2105R, as described previously [52], sequenced using the primers rpoB458F, rpoB1110F and rpoB2105R, as described by Tarr et al. [69] and Hazen et al. [31]. The resulting sequences with a minimum length of 1550 bp, including 13 sequences of representative type strains, were added to a phylogenetic tree constructed from 180 *rpoB* reference sequences obtained from GenBank using the ARB software package (version 5.5) provided by the technical university of Munich [44]. The phylogenetic relationships were