Multilocus Sequence Typing for Genotyping of *Vibrio parahaemolyticus*

Sara Urmersbach[1], Madura Sanjeevani Gonsal Koralage[2], Thomas Alter[1], Ute Messelhäusser[3], Duangporn Pichpol[4], Stephan Huehn[1]
[1]Institute of Food Hygiene, Free University Berlin, Germany
[2]Government Veterinary Office, Walikanda, Polonnaruwa, Sri Lanka
[3]Bavarian Health and Food Safety Authority, Oberschleißheim, Germany
[4]Faculty of Veterinary Medicine, Chiang Mai University, Thailand

Background

Vibrio (*V.*) *parahaemolyticus*, a naturally present potentially pathogenic bacteria, is isolated frequently from seawater, sediment and raw or insufficiently cooked seafood (e.g. shellfish and bivalves). Owing to the high filtrating activities of bivalves the level of *V. parahaemolyticus* was found to be 200-fold higher than in the surrounding seawater (DePaola *et al.* 2000). Consumption of or contact to raw or undercooked seafood, that contains *Vibrio* in appropriate numbers, can lead to infections (gastroenteritis, wound infections and septicaemia).

Isolates with (A) different geographic origin were analysed via MLST concerning their genetic relatedness and possible distribution patterns. One focus was the genomic analysis of (B) isolates gained from shrimp samples of three Sri Lankan regions.

Material and Methods

Bacterial-Strains

In total 78 *V. parahaemolyticus*-isolates were analysed.
Thereof, 30 strains derived from seawater, bivalves and shrimp of different geographic origin. Additionally, 46 strains were isolated from shrimp of farms located in three Sri Lankan regions (Fig. 1).
One isolate per farm and per shrimp-pond was applied to MLST analysis, respectively.
The Japanese clinical strains ATCC17802 and RIMD2210633CM29 served as reference strains.

Multilocus Sequence Typing (MLST)

In MLST analysis the nucleotide sequences of multiple reference genes are compared. Therefore, internal fragments of the genes *dna*E, *gyr*B, *rec*A, *dtd*S, *pnt*A, *pyr*C and *tna*A were amplified via PCR and sequenced (Fig. 2A; González-Escalona *et al.* 2008).
The received sequence data were analysed with Bionumerics (Applied Maths, Version 6.01) and compared to already published sequences on the PubMLST web page (http://pubmlst.org/vparahaemolyticus; Jolley *et al.* 2004). Based upon the concatenated nucleotide sequences a dendrogram was created via UPGMA-Analysis. The resulting similarity matrix served as basis for the generation of Minimum Spanning Trees.

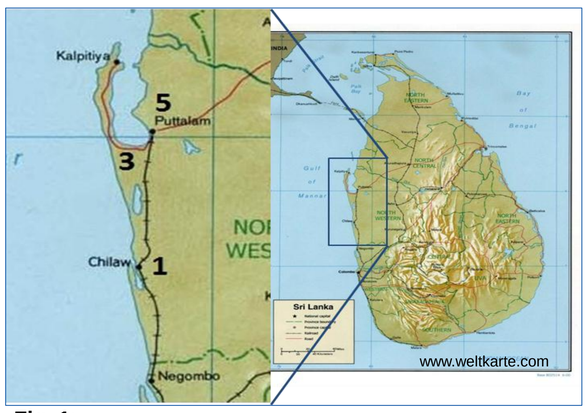

Fig. 1:
Map of Sri Lankan Sampling Areas
(1 Chillaw, 3 Madurankuliya, 5 Puttalam)

*rec*A and *gyr*B

Dividing the original fragments of *gyr*B and *rec*A (Fig. 2B) into two fragments led to complete allelic profiles. Therefor two inner primer were designed, allowing PCR amplification (Fig. 2C) and sequencing with the conditions published by González-Escalona (2008).

eBURST-Analysis

The eBURST algorithm identifies groups of related genotypes on the basis of their allelic profiles, called sequence type (ST) and predicts the descent from the presumably founder to other genotypes in this group (http://eburst.mlst.net/).

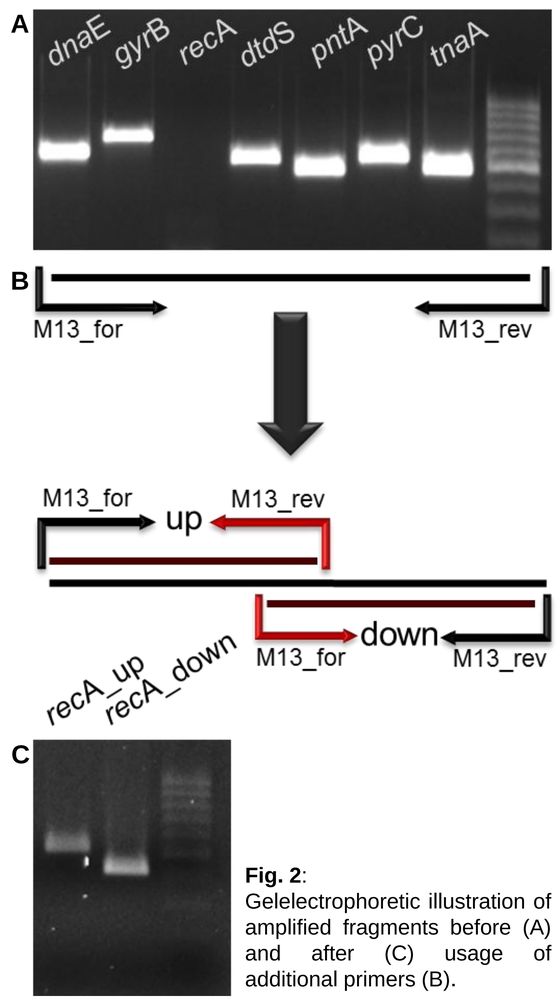

Fig. 2:
Gelelectrophoretic illustration of amplified fragments before (A) and after (C) usage of additional primers (B).

Results

Most of the analysed isolates originated from Asia (88%), especially West-Asia.
Sequencing of 78 isolates led to identification of 65 new alleles: 7 for *dna*E, 15 for *gyr*B, 9 for *rec*A, 9 for *dtd*S, 6 for *pnt*A, 10 for *pyr*C und 9 for *tna*A. A complete allelic profile and thus an assignment of a ST was revealed for 69 isolates. Most of them (n= 52) belonged to new STs, consisting of either recombination of already published or combinations with new STs.

A – Influence of geographic origin on the MLST-Sequence Types

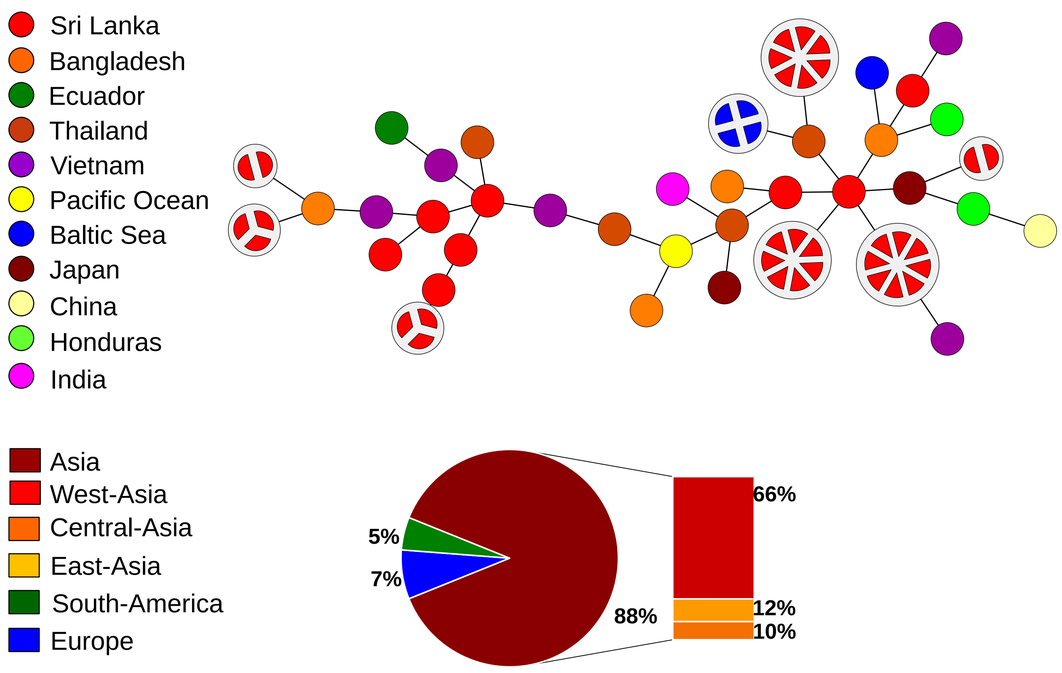

Fig. 3A:
Minimum Spanning Tree depending on sequence differences of STs. Pie chart of geographic origin. Colouring according to the region of origin.

The Minimum Spanning Tree reveals no clear clustering based on the geographic origin of the isolates (Fig. 3A). Strains from Sri Lanka formed two clusters, which are connected by strains of diverse origin (Japanese reference strains and food isolates).

Isolates from food samples could be assigned to 21 STs (8 already published, 13 new) and were distributed over the whole Minimum Spanning Tree. Isolates from South-America where located in both clusters, the two STs of European strains were located in one cluster but not nearby.

B – Distribution of MLST-Sequence Types within three regions of Sri Lanka

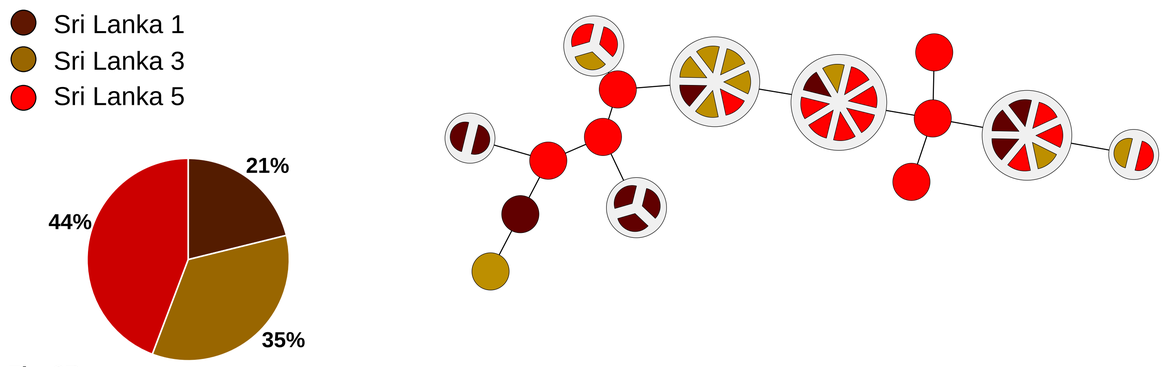

Fig. 3B:
Minimum Spanning Tree and pie chart of isolates from Sri Lanka; Colouring according to sampling area.

The majority of West-Asian isolates (n= 43) originated from shrimp-samples out of three Sri Lankan regions (Fig. 3B). They belonged to 15 STs. Ten of these derived from isolates out of one region, two appeared in two, and three in all regions. Isolates with the same ST and region originated mostly from different farms but at least from different ponds.

C – eBURST-Analysis

eBURST-Analysis of all published *V. parahaemolyticus* STs led to the detection of a new clonal complex (CC; Fig. 4). The founder, ST412 (trh$^+$tdh$^-$), was isolated from an Ecuadoran shrimp sample. All other strains of the CC were isolated in the USA. Single Locus Variants (SLV) were from environmental samples: ST313 (trh$^+$tdh$^+$), ST314 (trh$^+$tdh$^+$) and ST315 (trh$^-$ tdh$^+$).The Double Locus Variant (DLV) ST43 (trh$^+$tdh$^+$) was found in environmental and clinical samples. The Satellite ST44 (trh$^+$tdh$^+$) was only found in clinical samples.

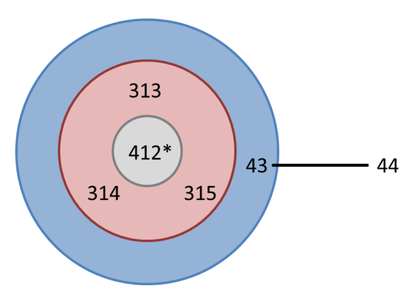

Fig. 4:
Graphical output of eBURST-analysis (http://pubmlst.org/)

Discussion

Often strains are not typable owing to missing sequences of *rec*A and *gyr*B fragments (Chao 2011). With the application of newly designed primer *gyr*B and *rec*A fragments of former non-typable isolates could be analysed via MLST.
The distribution of sequence types is independent from the geographic origin of analysed isolates. Strains from Asia as well as from Europe and South-America were genetically diverse and no dominating ST was identified.
Owing to the high fraction of Asian strains (88%) the distribution of European and South-American strains will be restudied after analysis of further non-Asian isolates.
Sri Lankan strains from shrimp farms of three nearby localized regions possessed a high genetic diversity and no local dominating ST was found. However, strains with the same ST were isolated in different regions.
Owing to the lack of more published data of Sri Lankan isolates and the naturally high genetic diversity of environmental isolates, most of the strains isolated possessed new STs (94%).
Close relatedness of environmental and clinical isolates was shown in Minimum Spanning Trees and also discovered by eBURST-analysis, thus emphasising the high pathogenic potential of environmental and food isolates of *V. parahaemolyticus*.

Acknowledgement: The project was founded by BMBF (VibrioNet-Project).

Literature:
Chao 2011, Origin of *Vibrio parahaemolyticus* O3:K6 pandemic clone
DePaola et al. 2000, Environmental investigations of *Vibrio parahaemolyticus* in Oysters after outbreakes in Washington, Texas and New York (1997 and 1998)
Feil 2004, eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from
González-Escalona et al. 2008, Determination of Molecular Phylogenetics of Vibrio parahaemolyticus Strains by Multilocus Sequence Typing
Jolley et al. 2004, mlstdbNet – distributed multi-locus sequence typing (MLST) databases; *Vibrio parahaemolyticus* MLST web page (http://pubmlst.org/ vparahaemolyticus) developed by Keith Jolley and sited at the University of Oxford

Institute of Food Hygiene • FU Berlin • Königsweg 69 • 14163 Berlin • Germany • Tel. 030 - 838 - 62538 • Fax 030 - 838 - 62552 • sara.urmersbach@fu-berlin.de • www.vetmed.fu-berlin.de