FEBS Letters 586 (2012) 1617–1621

journal homepage: www.FEBSLetters.org

Genomic analysis of ICE*Vch*Ban8: An atypical genetic element in *Vibrio cholerae*

Elisa Taviani [a,1], Matteo Spagnoletti [b,1], Daniela Ceccarelli [b], Bradd J. Haley [a], Nur A. Hasan [a], Arlene Chen [a], Mauro M. Colombo [b], Anwar Huq [a], Rita R. Colwell [c,d,*]

[a] *Maryland Pathogen Research Institute, University of Maryland, College Park, MD 20742, USA*
[b] *Dipartimento Biologia e Biotecnologie Charles Darwin, Sapienza Università di Roma, Rome, Italy*
[c] *John Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA*
[d] *Center for Bioinformatics and Computational Biology, University of Maryland, College Park, MD 20742, USA*

ARTICLE INFO

Article history:
Received 8 February 2012
Revised 28 March 2012
Accepted 30 March 2012
Available online 18 April 2012

Edited by Takashi Gojobori

Keywords:
Integrative conjugative elements
Genomic islands
Lateral gene transfer
Vibrio cholerae

ABSTRACT

Genomic islands (GIs) and integrative conjugative elements (ICEs) are major players in bacterial evolution since they encode genes involved in adaptive functions of medical or environmental importance. Here we performed the genomic analysis of ICE*Vch*Ban8, an unusual ICE found in the genome of a clinical non-toxigenic *Vibrio cholerae* O37 isolate. ICE*Vch*Ban8 shares most of its genetic structure with SXT/R391 ICEs. However, this ICE codes for a different integration/excision module is located at a different insertion site, and part of its genetic cargo shows homology to other pathogenicity islands of *V. cholerae*.

1. Introduction

Recent comparative genomics studies revealed that prokaryotic genomes are dynamic entities capable of acquiring and discarding large amounts of genetic material via lateral gene transfer (LGT) and recombination [1]. The role of LGT in the evolution of *Vibrio cholerae*, an autochthonous inhabitant of riverine and marine environments, as well as a human pathogen, is extensively documented [1–4]. In this species, major virulence genes and several important adaptive functions are known to be clustered in regions of the chromosome, laterally acquired from conspecific or distantly related organisms [5,6]. The extent of LGT was confirmed by the presence of Integrative conjugative elements (ICEs) and a large set of Genomic Islands (GIs) in *V. cholerae* sequenced genomes, contributing to genomic plasticity and generating a significantly heterogeneous group of strains in this species [1,5].

GIs are not self-transmitting gene clusters, since they are devoid of transfer genes, but they can excise and form circular intermediates [7–9]. They are of significant interest since they often encode genes of clinical importance, such as virulence factors, otherwise known as pathogenicity islands (PAI) [8]. Among all the *V. cholerae* associated PAIs, *Vibrio* pathogenicity island 2 (VPI-2) is present in toxigenic O1 strains, and occasionally in non-O1, non-O139 *V. cholerae* [10,11]. VPI-2 is a 57.3 kb island integrated at a *tRNA-Ser*, encoding 52 ORFs including proteins for release, transport, and catabolism of sialic acid [11].

Integrative conjugative elements (ICEs) are a class of mobile elements integrated into the prokaryotic host chromosome. They comprise mosaic genetic element with plasmid-, phage-, and transposon-like features organized in functional modules and are capable of self-transfer from a donor to a recipient cell via conjugation [12]. These elements were first described in *V. cholerae* with the discovery of the SXT element [13] and have since been found in the majority of *V. cholerae* and related *Vibrio* spp. [14,15]. All ICEs that have been described in *V. cholerae* to date belong to the SXT/R391 family [16], which share a conserved genetic scaffold of 52 genes that incorporates unique sequences coding for resistance to antibiotics and heavy metals, new toxin/antitoxin systems, restriction/modification systems, and alternative metabolic pathways [15,16].

In this study, we report the genomic analysis of ICE*Vch*Ban8, an unusual ICE present in the genome of the non-toxigenic clinical isolate *V. cholerae* O37 MZO-3, collected in Bangladesh in 2001 [1,15]. Unlike most *V. cholerae* non-O1, non-O139 isolates, strains of serogroup O37 have been shown to have epidemic potential [17,18]. *V. cholerae* MZO-3 lacks CTXΦ [1] but shows the presence and/or major rearrangements of two of the main pathogenicity

* Corresponding author at: Institute for Advanced Computer Studies, University of Maryland, 3103 Biomolecular Science Building 296, College Park, MD 20742, USA. Fax: +1 301 314 6654.
E-mail address: rcolwell@umiacs.umd.edu (R.R. Colwell).
[1] Equal contribution.

0014-5793/$36.00
http://dx.doi.org/10.1016/j.febslet.2012.03.064

islands found in *V. cholerae* 7th pandemic isolates: a novel version of *Vibrio* seventh pandemic island 2 (VSP-2) [19] and the replacement of VPI-2 with ICE*Vch*Ban8, which is the subject of the present work.

2. Materials and methods

2.1. ICE assembly, annotation and comparative genomics

Nucleotide sequence of ICE*Vch*Ban8 was obtained from the genome sequence of *V. cholerae* O37 strain MZO-3 (AAUU00000000) [1]. Gaps between two contigs were closed by manual editing using Consed and subjected to custom primer walk and PCR amplification. Complete nucleotide sequence of ICE*Vch*Ban8 was deposited in GenBank under Accession no. JQ345361.

ICE genetic analysis was accomplished in four steps: (i) ICE*Vch*Ban8 nucleotide sequence was aligned with SXT using NUCmer [20]; (ii) ORFs were identified and annotated using the RAST annotation pipeline (http://rast.nmpdr.org); (iii) ICE sequence and genetic organization was compared with published ICEs using the Artemis Comparative Tool (ACT) (www.sanger.ac.uk/Software/ACT); and (iv) similarity in nucleotide and protein sequences for ICE*Vch*Ban8 was determined as % nucleotide or amino acid identity with other ICEs in GenBank employing BLASTN and BLAST-PSI [21].

2.2. PCR assays

Chromosomal DNA was extracted as previously described [22]. Amplification was performed in an automated thermocycler (Bio-Rad MJ-Mini), and PCR conditions were performed in 50 μl reaction mix contained 1.3 U of DreamTaq DNA polymerase (Fermentas), according to manufacturer's instructions. See Supplementary Table 1 for primer and amplicon details. To confirm accurate assembly of the two contigs and to close gaps between them, we designed different primers: pMZO_1F/pMZO3_1R for contig 1 assembly; pMZO3_1F2 /pMZO3_54R for contigs 1 and 54 junction. EnterotoxinB-F/s063R primer pair was designed to localize the 3′-end of Hotspot 4 insertion and AcriF/AcriR to confirm the presence of the acriflavine resistance gene.

Primers Ban8attP-F/Ban8attP-R were designed to investigate the capacity of ICE*Vch*Ban8 to form a circular intermediate (*attP*), and primers Ban8attB-F/Ban8attB-R were designed to detect the empty site on the bacterial chromosome (*attB*). The four primers were specifically designed to allow the combination of Ban8attP-F/Ban8attB-R and Ban8attP-R/Ban8attB-F in order to amplify the *attL* and *attR* junctions of the integrated ICE, respectively. Positive controls for excision experiments were strains *V. cholerae* O139 MO10 carrying SXT [13] and *V. cholerae* O1 7452 carrying ICE*Vch*Ind5 [23], using primers P4/P5 for *attP* amplification, as previously described [24].

Amplicons to be sequenced were directly purified from PCR by Nucleospin extract kit (Macherey–Nagel), according to manufacturer's instructions. DNA sequences were obtained using an Applied Biosystems DNA sequencer 3730.

3. Results

3.1. Assembly and genomic organization of ICEVchBan8

Comparative analysis of the genome of *V. cholerae* O37 MZO-3 revealed sequences belonging to the genetic backbone of ICE*Vch*Ban8 on two contigs: nt. 27439–105212 of contig 1 (Accession no. NZ_AAUU01000001.1), and nt. 1–25409 of contig 54 (Accession no. NZ_AAUU01000001.54). The overall genetic organization of this element had already been identified as similar to SXT/R391 ICEs, and named ICE*Vch*Ban8 [15]. Assembly of the ICE was accomplished using NUCmer with SXT^{M010} as reference and it revealed one gap at the junction between the two contigs that was resolved by PCR and confirmed by sequencing (Fig. 1). Our analysis revealed a 103.381 bp sequence (Accession no. JQ345361). The complete sequence of ICE*Vch*Ban8 was submitted to RAST pipeline for annotation and resulted in identification of 106 ORFs. Annotation was refined manually, comparing it with SXT and other ICEs of the SXT/R391 family found in *V. cholerae* [15]. ICE*Vch*Ban8 contains 47 out of 52 conserved core genes responsible for transfer and regulation of integration/excision in the SXT/R391 ICEs backbone (Fig. 1) [15]. Genes *xis* and *int* were replaced by a new integration/excision module (as described in the next section), whereas genes *s024*, *s025* and *s026* are missing. Genes *s024* to *s026* were previously annotated as hypothetical proteins, and reported as not involved in ICE conjugative transfer [25], suggesting a non-functional role in SXT biology.

Four operons of *tra* genes, encoding the conjugative apparatus (*traIDJ*, *traLEKBA*, *traCFWUN*, and *traFHG*) share at least 97% similarity at the nucleotide level with the *tra* clusters in the conserved SXT/R391 ICE backbone. As noted for all SXT-related ICEs, specific inserted genes were identified in five hotspots and four variable regions within the core backbone (Fig. 1) [15]. In ICE*Vch*Ban8, Hotspot 1 and Hotspot 2 have the same molecular arrangement as in SXT, with an additional gene in Hotspot 2 coding for a hypothetical protein, also found in ICE*Vsp*Por2 and ICE*Val*Por1 from *V. splendidus* and *V. alginolyticus* [26]. Hotspot 3, and variable regions I, II and III do not have any insertions. At the 3′ end of the element, the variable region IV is disrupted by the rearrangement of *setR*/*eex* backbone genes. Overall, ICE*Vch*Ban8 holds some major rearrangements in its genetic structure, such as the substitution of the integration/

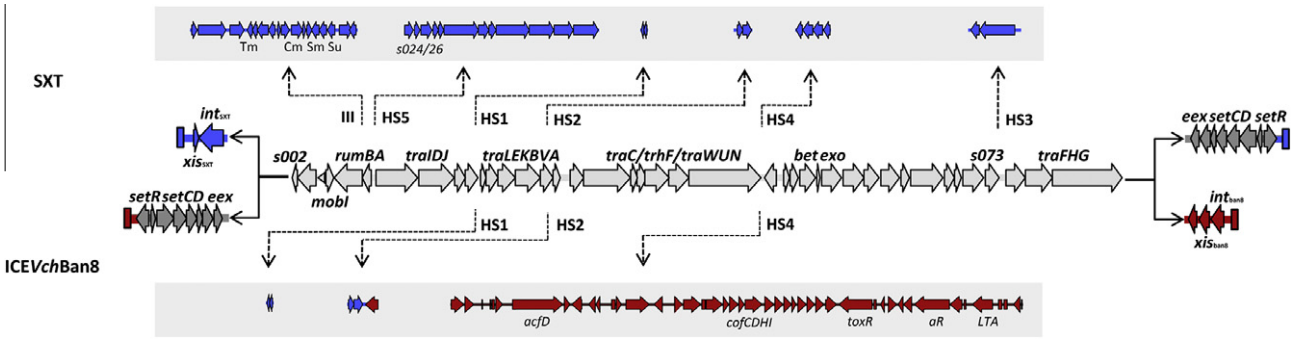

Fig. 1. Structural comparison between SXT and ICE*Vch*Ban8. A genetic map of core genes shared between the two ICEs (light gray) is shown. Specific regions for SXT and ICE*Vch*Ban8 are depicted in blue and red, respectively. Dark gray indicates inverted genes *setR*/*eex*. Hotspot insertions are indicated by dotted arrows, as shown in the above (SXT) and below (ICE*Vch*Ban8) the shared backbone.