Management of multipartite genomes: the Vibrio cholerae model.

Marie-Eve Val, Alfonso Soler-Bistué, Michael J Bland, Didier Mazel

To cite this version:
Manuscript Number:

Title: Management of multipartite genomes: the Vibrio cholerae model

Article Type: 22 Growth&Develop: prokaryotes 2014

Corresponding Author: Prof. Didier Mazel

Corresponding Author's Institution:

First Author: Marie-Eve Val

Order of Authors: Marie-Eve Val; Alfonso Soler-Bistué; Michael J Bland; Didier Mazel
Management of multipartite genomes: the *Vibrio cholerae* model

Marie-Eve Val¹,², Alfonso Soler-Bistué¹,², Michael J. Bland¹,² and Didier Mazel¹,²

¹Institut Pasteur, Unité Plasticité du Génome Bactérien, Département Génomes et Génétique, Paris, France; ²CNRS UMR3525, Paris, France

* For correspondence: mazel@pasteur.fr

Abstract:

A minority of bacterial species has been found to carry a genome divided among several chromosomes. Among these, all *Vibrio* species harbor a genome split into two chromosomes of uneven size, with distinctive replication origins whose replication firing involves common and specific factors. Most of our current knowledge on replication and segregation in multi-chromosome bacteria has come from the study of *Vibrio cholerae*, which is now the model organism for this field. It has been firmly established that replication of the two *V. cholerae* chromosomes is temporally regulated and coupled to the cell cycle, but the mediators of these processes are as yet mostly unknown. The two chromosomes are also organized along different patterns within the cell and occupy different subcellular domains. The selective advantages provided by this partitioning into two replicons are still unclear and are a key motivation for these studies.

1. Introduction

Bacterial genomes are composed of two types of replicons: chromosomes, which are by definition essential, and plasmids, which are dispensable. Most bacteria have one single circular chromosome, varying in size from barely 100 kb [1] for obligatory intracellular species to over 13 Mb [2], but bacteria with multiple chromosomes are frequent (about 10% of bacteria with sequenced genomes) and arose in several taxa. Indeed, bacteria with two or more chromosomes have been identified in diverse prokaryotic phyla including Chloroflexi, Deinococcus - Thermus, Spirochaetes, and Proteobacteria (α-, β- and γ- classes), suggesting that they have arisen independently, many times in the
course of evolution. In general, these chromosomes are circular, but in some cases such as in *Agrobacterium*, one of the two chromosomes is linear [3]. Among the proteobacterial families that have been found to carry multiple chromosomes, the Vibrionaceae family, which includes the *Vibrio, Photobacterium, Listonella*, and *Aliivibrio* genera, shows the highest consistency in terms of genome structure. Indeed, all species characterized since the original description of this specific genome organization in *V. cholerae* and *V. parahaemolyticus* [4,5], have been found to carry 2 chromosomes of uneven sizes [6]. It is now commonly accepted that this second chromosome derives from the domestication of a plasmid in the ancestor of the current Vibrionaceae families, after transfer of essential genes from the chromosome to this replicon [7].

Multiple chromosome maintenance and replication have been studied to a certain extent in bacteria from the other groups, such as *Burkholderia, Rhizobium* or *Brucella*, but most of our current knowledge on these topics has come from studies in *V. cholerae*. Like all Vibrionaceae, *V. cholerae* has one large chromosome, called the primary chromosome or chromosome 1 (chr1), due to the relatedness of its replication initiation machinery and its control with the one of other γ-proteobacteria such as *E. coli*, and one secondary chromosome, called chromosome 2 (chr2), which possesses a distinct replication initiation set-up.

2. Replication initiation of *V. cholerae* chr1 and chr2

In *V. cholerae*, the factors responsible for controlling replication initiation of the two chromosomes are distinct [8]. The minimal replication origin of *V. cholerae* chr1 (ori1) is fairly similar to the canonical *E. coli* chromosomal origin (oriC) [9]. Like oriC, ori1 contains binding sites for DnaA (DnaA boxes), the main initiator of replication that promotes the unwinding of bacterial chromosomal origins [10]. It also carries an IHF binding site, and several GATC sites for methylation by DNA adenine methyltransferase (Dam), which regulates the timing of re-initiation through sequestration of hemimethylated sites by SeqA [11]. ori1 is able to functionally replace oriC in *E. coli* [12,13] suggesting that similar processes are likely to govern chr1 replication initiation in *V. cholerae*.
Replication initiation of chr2 is very different from that of chr1 or *E. coli* oriC. This novelty has motivated its extensive study over the past decade and an understanding of chr2 replication is likely to benefit understanding of how megaplasmids have been domesticated by bacteria to become secondary chromosomes. Replication of chr2 is triggered by a specific initiator, RctB, that is conserved within all Vibrionaceae but shares no homology with any other replication initiators [9]. RctB has several functional forms. RctB binds and hydrolyzes ATP but unlike DnaA, the ATP-bound form of RctB is inactive [14]. It binds to DNA as monomer or as a dimer [15,16]. Finally, RctB is most probably remodeled by chaperones DnaJ and DnaK similar to iteron-carrying plasmid’s Rep initiators [15,17].

RctB concentration is the rate-limiting component of chr2 initiation [18]. As with chr1, the homeostatic system that sets the copy number of chr2 and corrects over-replication and under-replication is based on negative-feedback control of the availability of the initiator. For chr2, *V. cholerae* has integrated a complex regulatory mechanism to control the level and activity of RctB. Chr2 replication control appears to be a combination of mechanisms similar to those found in iteron-like plasmids, which are commonly known as initiator autoregulation, initiator titration and origin handcuffing [19]. The origin region of chr2 is divided into three functional units: (i) *rctB*, encoding the initiator of replication; (ii) *ori2*, the minimal origin of replication; and (iii) *incII*, a negative regulatory region which contains a transcribed but non-translated ORF *rctA* [9] (Figure 1). *ori2* is organized similarly to iteron-bearing plasmid origins. It contains six iterons, which are 12-mer repeated initiator binding sites. RctB binds as a monomer to these iterons and promotes the unwinding of *ori2* for initiation [14]. The contiguous incII region negatively regulates chr2 replication. It contains five regulatory iterons (11- and 12-mers) and two 39-mer motifs, one of which is found in *rctA* [20]. RctB can also bind efficiently to 39-mer motifs where it serves as a negative regulator of *ori2* initiation and regulates the timing of chr2 replication[15,20]. The inhibitory activity of the 39-mers is very powerful and is mainly conveyed through two mechanisms: initiator titration and origin handcuffing [15,20,21]. Structure-function studies of RctB have shown that the carboxy-tail of RctB is dispensable for initiation but essential for down-regulation of replication [15,22,23]. Iterons found outside *ori2* in the *incII* locus have a regulatory function. These regulatory iterons serve as titration sites for RctB; additionally their precise arrangement and orientation help to restrain the strong inhibitory activity of the
39-mer motifs [20]. The 39-mers, conversely, enhances handcuffing of the regulatory iterons with themselves [20].

Another important regulatory mechanism of chr2 replication takes place in rctA [21]. Transcription of rctA attenuates its own 39-mer inhibitory activity, presumably by interfering with RctB binding. RctB, alternatively, binds to regulatory iterons located in the rctA promoter region, repressing rctA transcription [21,24]. This mechanism of transcriptional interference participates in adjusting the level of available RctB.

RctB also auto-regulates its own expression through binding at another RctB binding motif called the 29-mer in the promoter region of rctB [25]. The 29-mer functions as a transcription operator, from which RctB exerts a negative feedback regulation on its own transcription [24]. The 29-mer also participates in the control of ori2 initiation through handcuffing with ori2 iterons [25]. The 39- and 29-mer motifs are closely related. Indeed, the 29-mer is a truncated version of the 39-mer and can be functionally replaced by a 39-mer [15,25]. Recently, a genome-wide study of RctB binding by chromatin immunoprecipitation and microarray (ChIP-chip) shows that RctB also binds to external sites, notably a span of 74 kb on chr2 containing six RctB binding sites (five iterons and one 39-mer motif) that negatively regulate ori2 replication [26]. This locus could be reminiscent of the E. coli datA titration locus [27,28], being able to titrate RctB and inhibit ori2 replication. A second site is located on chr1 and was found to enhance ori2 replication [26] using an as yet unknown mechanism.

3. Cell-cycle-dependent regulation of chromosome replication

The cell cycle is defined by events that occur only once per generation: chromosome duplication, chromosome segregation and cell division. Analogous to that of eukaryotes, the bacterial cell cycle is divided into three stages: cell birth to chromosome replication initiation (B), chromosome replication (C) and termination of replication to cell division (D). Chromosomes usually replicate at a fixed time once per cell cycle while plasmids usually initiate replication several times over the entire bacterial cell cycle [29]. In contrast, megaplasmids and secondary chromosomes, i.e. RepABC megaplasmids of α-proteobacteria and chr2 of V. cholerae, replicate once per cell cycle [30,31]. In V. cholerae, chr1 initiates at the onset of the replication period while initiation of chr2 is
delayed and occurs when 2/3rd of the replication period has already been completed. Because chr2 is 1/3rd the size of chr1, both chromosomes terminate their replication at the same time which signals the end of the C period [32]. Presently, we don’t know how and why this termination synchrony occurs. Even if the two chromosomes are independently maintained, this observation indicates that chr1 and chr2 communicate to coordinate their replication. ori1 and ori2 are very divergent but they contain some common features such as a DnaA box and a binding site for IHF, suggesting that DnaA and IHF can be used by both chromosomes for replication initiation and could play a role in their coordination [9]. The DnaA box is conserved in the ori2 of other Vibrio species, implying an evolutionary purpose [9]. However, in V. cholerae, DnaA overproduction doesn’t seem to impact chr2 replication [8].

The origin region of chr2 has an overrepresentation of Dam methylation sites (Figure 1). Like on ori1, the hemi-methylated state of ori2 is extended, thereby ori2 is also subjected to sequestration by SeqA which prevents immediate re-initiation [12]. Interestingly, unlike ori1, Dam methylation is strictly essential for chr2 initiation [12,33]. Indeed, all iterons (initiation and regulatory) contained in the origin region of chr2 have a Dam methylation site that needs to be fully methylated to bind RctB [12]. RctB binding to iterons functions, thereby, in a cell-cycle dependent manner. Alternatively, the 39-mer (including the 29-mer) does not need to be methylated to bind RctB [25]. Thus, there is interplay between methylation-dependent processes (involving RctB/iteron interaction) and methylation-independent processes (involving RctB/39-mer) that results in an equilibrium of RctB, which allows the correct timing of chr2 replication initiation. The plasmid iterons lack methylation sites and their replication is not linked to the cell cycle. Dam regulation of chr2 replication provides some clues as to how the replication is limited to a specific time of the cell cycle [25]. Chr2 has integrated a very sophisticated plasmid-derived initiation regulation system and improved it to behave in the cell like a chromosome in order to curb replication to once per cell cycle. Dam and SeqA together regulate the initiation and re-initiation of chr1 and chr2 and thus could serve as common regulators of chr1 and chr2 replication to help participate in the coordination of their replication.

4. Chr1 and chr2 distribution and segregation in V. cholerae
The two chromosomes of *V. cholerae* are longitudinally arranged in the cell [34]. While chr1 seems to be spread along the entire longitudinal axis of the cell, chr2 is restricted to the younger half of the cell. In newborn cells, chr1 extends from the old pole to the new pole and chr2 extends from midcell to the new pole [34] (Figure 2).

Most bacteria, with the notable exception of *E. coli*, carry parAB genes in their chromosome that participate in chromosome partition [35]. Each of the two *V. cholerae* chromosomes has been found to encode a specific partition system, namely ParAB1 and ParAB2, which recognize distinct sites exclusively carried on their cognate chromosome [36]. *para* codes for an ATPase and *parB* for the sequence specific DNA binding protein that binds to its specific *parS* sites. Several *parS* sites are usually found within the chromosomal *oriC* region and in its vicinity [35]. The genes were initially found in low copy number plasmids, and were shown to facilitate their segregation into daughter cells. However, their role in bacterial chromosome segregation is still controversial, as the knockout of Par systems leads to variable phenotypes depending of the species [37-40]. In *V. cholerae*, deletions of *parAB1* or of *parS1* sites do not affect chr1 partitioning [36]. Conversely, ParAB2 is essential for chr2 segregation, and its disruption leads to chr2 loss and cell death [41].

The location of *ori1* at the old pole has been found to be mainly mediated by an interaction between ParA1 and a specific pole anchor protein, HubP [42], which is in agreement with previous observations that ParA1 is essential for *ori1* polar location [43]. The displacement of *parS1* sites at a distant loci from *ori1* shifts the polar location to the region carrying the relocated *parS1* sites, confirming ParAB1’s role in mediating polar location [34]. Nevertheless, in absence of ParA1, *ori1* is kept in the old pole region, but in a less precise manner. In this case, replication was shown to drive the longitudinal organization of chr1 [34].

Deletion of *parB1* causes *parA1*-dependent over-initiation of chr1 replication that appears to be mediated through direct interaction with DnaA [44] in a manner similar to what has been described in *B. subtilis* [45]. There is an analogous regulatory cross-talk for chr2 where ParB2 contributes to chr2 replication regulation, thereby enabling linkage between replication and segregation. ParB2 binding to *parS2*-B within the *rctA* ORF nearby the 39-mer RctB-binding motif interferes with *rctA* replication inhibitory activity [23] (Figure 1). ParB2 spreads from *parS2*-B into the *rctA* 39-mer which likely interferes with RctB binding [46]. ParB2 was also reported to promote replication by
direct binding to a more distant 39-mer contained in \textit{incII} (interestingly without requiring spreading from parS2-B). On this 39-mer, ParB2 competes with RctB to restrain its activity [46]. Alternatively, binding of RctB to \textit{rctA} activates \textit{parAB2} expression [23]. These binding fluctuations underlie a regulatory network controlling both replication and partitioning of chr1 and chr2, demonstrating how chromosome replication and origin segregation are intimately intertwined. While replication and segregation of chr1 and chr2 origins occurs at different spatial and temporal points, replication and segregation of their termini are synchronous and occur at midcell [32,34] (Figure 2). There may be a control mechanism that coordinates chr1 and chr2 to synchronize their termination and late stages of segregation at midcell. One of the last steps of chromosome segregation before cell division involves the resolution of dimeric chromosomes that are frequently produced by homologous recombination between sister-chromatids following DNA damage [47]. In \textit{V. cholerae}, dimers of chr1 and chr2 are resolved by the action of the same machinery, XerC and XerD site-specific recombinases at the \textit{dif} sites (dif1 and dif2), located into the \textit{ter} regions of chr1 and chr2, respectively [48]. This recombination event is controlled by FtsK, a DNA translocase. FtsK is associated with the division apparatus and therefore links the resolution of dimeric chromosomes to the end of the cell cycle, which is cell division. A recent study of FtsK activity in \textit{V. cholerae} suggest that in addition to its role in dimer resolution, it also serves to facilitate the segregation of a specific region of sister chromosomes across the division septum [49].

5. Perspectives and outstanding questions

The question of the selective advantage of multipartite bacterial genomes is still unsolved. However, this organization is clearly stable, especially in \textit{Vibrio} species. Chromosome co-integrates have been observed in \textit{Sinorhizobium meliloti} at low rates and have been found to spontaneously revert to the multiple chromosome organization [50]. In \textit{V. cholerae}, due to its essential role in chr2 replication, Dam depletion can only be overcome by chromosome fusion [51]. Fusions have been observed to occur either by homologous recombination between identical IS copies or through recombination
between the dif sites of the two chromosomes, providing a possible escape route for chr2 replication inactivation. It has also been established that chromosome fusion through dif recombination occurred naturally at detectable levels, but were subjected to counter selection [51]. This phenomenon was further studied through the construction of a synthetic V. cholerae derivative possessing a single chromosome by fusing chr1 with chr2 in a calculated manner to conserve the “ori-ter” axial symmetry, gene synteny, strand bias and the polarities of the original replichores. In this strain, where the replication of the fused chromosome initiates at ori1 and finishes in the terminus of chr2 near dif2, while ori2 and dif1 have been removed, the effect on the generation time is minimal [33]. As this difference is clearly insufficient to explain the counter selection of natural fusions, it could be imagined that the presence of two different replication origins on a single replicon deleterious to the cell.

A second issue that is not yet understood is why bacteria with multiple chromosomes always carry different types of replication origins for the main and secondary chromosomes. A V. cholerae mutant where chr1 and chr2 are initiated at two identical DnaA-regulated ori1 origins is not impaired for growth [33]. This means that distinct replication systems are not a prerequisite for the viability of bacteria with multipartite genomes. Another pending question is the understanding of the mechanism which ensures the proper timing of chr2 replication firing to synchronize the replication termination of the two chromosomes. Finally, the physical organization of the two chromosomes, in terms of compaction and macrodomains within the cell is completely unknown. Vibrio species carry the same set of dam-associated genes involved in DNA maintenance and chromosome macrodomain (MD) organization [52], which in particular, organizes three of the E. coli MDs [53]. MatP, which organizes the E. coli Ter macrodomain, is conserved in V. cholerae, and matS binding sites are distributed in the ter regions of both chr1 and chr2 [54]. This suggests that each of the two chromosomes carries a Ter MD, but any similarity between E. coli chromosome organization and the organization of the two V. cholerae chromosomes has yet to be established. V. cholerae genomes also carry a YfbV ortholog, which in E. coli binds to two 12 bp sites called tidL and tidR and is essential for the left and right MD organization [55]. However, these two sites are not found in the V. cholerae genome, suggesting that either these MDs are absent in vibrios, or that YfbV recognizes an unrelated sequence. Thus, it would be very interesting to determine the potential MD organization and how it differs from the single
chromosome organization found in *E. coli*. In addition, the presence of specific contacts between different chromosome regions should be investigated, as well as any potential common structures involving domains of both chromosomes.

**Acknowledgements**

This work was supported by a grant from the French National Research Agency (ANR-10-BLAN-131301). Research was funded by the Institut Pasteur, the Institut National de la Santé et de la Recherche Médicale (INSERM) and the Centre National de la Recherche Scientifique (CNRS). ASB is funded by EMBO (ALTF-1473-2010) and Marie Curie Actions (BMC FP7-PEOPLE-2011-IIF). MJB was supported by a stipend from the Pasteur-Paris University (PPU) International PhD program.

**Figure Legends**

**Figure 1**: Schematic map of the origin region of *V. cholerae* chr2, which is divided into three units: *rctB* (gene encoding for the initiator of replication), *ori2* (the minimal functional origin), *inclII* (a negative regulatory region including the non-translated ORF rcta). RctB binding sites are indicated in red if Dam-methylation is required and in blue if methylation is not required for binding. The *parA2* gene flanking the origin of chr2 as well as the parS2-B site (in rctA) are indicated to illustrate the regulatory crosstalk between replication and segregation of chr2.

**Figure 2**: A, Model for the arrangement of chromosomal DNA in *V. cholerae* newborn cells. The two chromosomes are longitudinally arranged: Ori1 (the origin region of chr1) is tethered at the old pole, Ter1 (the terminus region of chr1) is at the new pole, Ori2 (the origin region of chr2) is at midcell and Ter2 (the terminus region of chr2) is closer to the new pole. B, Model for the chronological order of duplication and segregation of the Ori and Ter regions of chr1 and chr2 (based on [34]).
References


34. David A, Demarre G, Muresan L, Paly E, Barre FX, Possoz C: The two Cis-acting sites, parS1 and oriC1, contribute to the longitudinal organisation of *Vibrio cholerae* chromosome I. *PLoS Genet* 2014, 10:e1004448.


Figure 1
Jha et al, 2014 (*)
The authors show that a restricted region in the C-ter domain of RctB is sufficient for both DNA binding and protein dimerization. This 71-aa region contains two contiguous domains which bind to either iterons or 39-mers and a dimerization domain which overlaps both DNA binding domains.

Baek et al, 2014 (*)
In this work, new sites for RctB binding were found outside of the origin of chr2. On chr2, a 74 kb stretch of DNA containing five iterons and one 39-mer was found to inhibit ori2 replication. On chr1, an atypical RctB binding site not related to either iterons or 39-mers acts as an enhancer of ori2 initiation.

Val et al, 2012 (**) 
In this work, large genome rearrangements were generated in V. cholerae to study the conserved vibrio two chromosome genome organization. Mutants with a single chromosome, with equally sized chromosomes and with chromosomes controlled by identical origins were studied to address issues about genome organization and maintenance.

David et al, 2014 (**) 
In this work, the cellular location and segregation pattern of 19 V. cholerae genomic loci on both chromosomes as a function of cell length have been monitored. This work gives a precise view of chr1 and chr2 dynamics, demonstrating that both chromosomes follow a longitudinal mode of organization and that both partition and replication machineries contribute to chr1 origin polar recruitment.

Yamaichi et al, 2012 (**) 
HubP, a multi-domain transmembrane protein that is anchored at the pole has been identified as establishing polar identity. HubP is required for the proper polar localizations of ParA1 and modulates the localization of ori1. HubP is also required for the proper cellular positioning of two other polar factors, ParC (chemotactic machinery), and FlhG (flagellum).

Venkova-Canova, 2013 (*)
In this work, the mechanism of two regulatory cross talks linking V. cholerae chr2 replication and segregation has been unraveled.

Val et al, 2014 (*)
This work reports the isolation of spontaneous natural mutants of *V. cholerae* with a single fused chromosome. The selective advantage of this new genome configuration is enforced by the depletion of the Dam methylase which is essential for chr2 replication.

**Thiel et al, 2012 (**)**

By moving large chromosomal DNA segments on the genome of *E. coli*, the authors have unraveled a new site-specific insulation system that restricts to the Ter region the consequences of the MatP-mediated constraining effect. Two sequences (*tidL and tidR*) flanking both sides of the Ter region and a DNA binding protein (*YfbV*) were found to be required to isolate the Ter from the other parts of the chromosome.
Highlights

- All *Vibrio* species carry a genome divided in multiple chromosomes.
- Replication of the two *V. cholerae* chromosomes is temporally regulated and coupled to the cell cycle.
- The two chromosomes have distinctive replication origins whose replication firing involves common and specific factors.
- The two *V. cholerae* chromosomes are organized along different patterns within the cell and occupy different subcellular domains.