RECONSTRUCTION AND ANALYSIS OF THE TRANSCRIPTOME REGULATORY NETWORK OF CLOSTRIDIUM BOTULINUM TYPE A3 STR. LOCH MAREE

ROJA B., THAMANNA L. AND CHELLAPANDI P.*

Industrial Systems Biology Lab, Department of Bioinformatics, School of Life Sciences, Bharathidasan University, Tiruchirappalli 620 024, Tamil Nadu, India

(Received 7 November, 2022; Accepted 15 December, 2022)

Key words: Regulon, Transcriptome, Clostridium botulinum, Botulism, Promoter, Gene expression, Virulence

Abstract—Clostridium botulinum type A3 str. Loch Maree is a clinically important strain that produces botulinum neurotoxin type A3 and causes foodborne, infant, and wound botulism worldwide. Studying the mechanism underlying the virulence of this organism is imperative to understand its antibacterial resistance and discovering new drugs or inhibitors. The biochemical and molecular characteristics of this organism have been intensively studied, but their gene regulatory mechanisms are unclear. Hence, we reconstructed the transcriptional regulatory network from the complete genome of this organism and analyzed interactive genes from the identified hub module using a knowledge-based bottom-up approach. The biological reliability, topological properties, and robustness of the regulatory network model were validated with network parameters, followed by gene ontology terms and literature support. The reconstructed regulatory network consisted of 12 transcriptional regulators associated with 2369 coding genes. ResD, SpoOA, ComK, CcpC, DinR, DegU, CitT, CodY, GerE, GltC, GltR, IolR, and LevR were identified as transcriptional regulators from this organism homologous to Bacillus subtilis. These regulators have been shown to control beta-lactamase, methyl-accepting chemotaxis protein, DNA replication protein DnaD, sensor histidine kinase, and putative membrane proteins of this organism. This study also predicted all possible promoter sites in regulated genes and their associated molecular functions. We conclude that a global regulatory network model of this organism provides insights into its growth physiology and virulence elicitation in the human intestinal environment.

INTRODUCTION

Clostridium botulinum strains produce botulinum toxin, which is the most potent neurotoxin that causes flaccid and spastic paralysis. This bacterium can colonize the intestinal tract and cause foodborne botulism and intestinal toxemia botulism. C. botulinum type A strains have a typical system for adaptation, survival, and virulence in host responses (Carter and Peck, 2015; Mazuet et al., 2015). Among type A strains, type A3 strains are more prevalent and clinically important in human botulism cases. The levels of botulinum neurotoxin type A (bont/A) expression and toxin complex formation depend on the growth phase of the organism, which is controlled by its complex regulatory networks (Dineen et al., 2003; Kouguchi et al., 2006; Artin et al., 2008; Ihekwaba et al., 2016). The expression of neurotoxic genes depends on BotR, which is located before the neurotoxic gene cluster. BotR is an alternative RNA polymerase sigma factor that functions in the transition phase (Couesnon et al., 2006).

The BotR regulon and the transcription factor Spo0A are co-regulated in the synthesis and activation of botulinum toxin (Cooksley et al., 2010; Shin et al., 2006). Several two-component signal transduction regulators have been reported to regulate bont/A gene expression (Connan et al., 2015). CBO0366/CBO0365 is a cold-induced two-component regulator that contributes to fatty acid biosynthesis, the oxidative stress response, and iron uptake, which in turn induces bont/A gene expression (Zhang et al., 2014a, b). CBO0789 represses bont/A gene expression by blocking botR-directed transcription (Zhang et al., 2013). The formation of neurotoxins is associated with general metabolism and quorum-sensing systems. This
suggests that nutritional and environmental stress factors are required for the complete virulence of this organism. It is unclear how regulatory pathways mediate the influence of nutrition and the environment on neurotoxin production.

Regulons are a fundamental component of global regulatory networks, and promoters regulate the expression of target genes by activation or repression (Liu et al., 2016). The expression of target genes depends on regulatory inputs determined by protein and DNA interactions (Schacht et al., 2014). A collection of regulatory interactions between transcription factors and their target genes is referred to as the transcription regulatory network. Several systems biology approaches have been used to construct transcription regulatory networks and identify transcriptional units of bacterial genomes (Prathiviraj and Chellapandi, 2020; Bharathi and Chellapandi, 2022). Global regulatory networks reveal the mechanisms underlying the molecular virulence and host adaptations of human pathogens at the system level (Latorre et al., 2014; Kim et al., 2015; Han et al., 2021).

Previous studies have reconstructed the transcriptional regulatory network of C. difficile and C. botulinum type A1 for spore formation and neurotoxin production, respectively (Saujet et al., 2014; Ihekwaba et al., 2016). Genome-wide datasets and transcriptional components are important for developing a biologically reliable global regulatory network model for type A3 strains. Therefore, this study aimed to reconstruct a genome-wide transcriptional regulatory network for C. botulinum type A3 str. Loch Maree (CBL). The resulting model would decipher the biochemical and biophysical characteristics of regulated genes under identified regulators to understand the molecular virulence mechanisms in specific environmental niches.

MATERIALS AND METHODS

Dataset

Complete genome sequences of CBL (Accession: CP000962; CP000963) were retrieved from MetaCyc v25.1 (Caspi et al., 2016).

Regulon prediction

Bacteria share a common ancestor in their fundamental transcription factors and binding sites. Therefore, B. subtilis 168 was chosen as a template genome to discover homologs and uncharacterized transcriptional regulators of CBL (Reuß et al., 2017).

Experimentally defined regulons in the model bacteria present in PRODORIC 1.3.1 (Dudek and Jahn, 2022) were used to predict transcription factors in the CBL genome. A promoter analysis tool from Virtual Footprint 3.0 (Grote et al., 2009) was used to identify homologous regulators and associated target genes with promoter sites based on a position weight matrix, as defined below.

\[ m(b, l) = f(b, l) \cdot R_{\text{sequence}}(l) \]

where \( m(b, l) \) is a position weight matrix and \( f(b, l) \) is the frequency of each base \( b \) at position \( l \) in the aligned binding sites, which uses a penalty function dependent on the sample size \( n \). The position weight matrix for the gene sequence was used to calculate the similarity score of the predicted promoter sites. The position weight matrix was set to a 0.8 sensitivity/threshold and 0.9 core sensitivity/size. This tool searched for all orthologous patterns in regulons recognized for functional exploration by dividing observed orthologs from expected orthologs. A threshold score (t) was assigned to adjust the accuracy of the position weight matrix search. ProdoNet was used to visualize the regulation of individual transcription factors in PRODORIC.

Regulatory network reconstruction

The global regulatory network of the CBL was created using a traditional probabilistic weight matrix technique. An Excel file containing the dataset’s summary data on predicted regulons, TFBS, regulated genes, and proteins was created (Supplementary File 1). Cytoscape 3.4.0 was used to model the regulon-gene interaction networks (Shannon et al., 2003). All orthologous clusters containing the common genes were combined. Functional descriptions of regulon-associated genes were assigned based on clusters of orthologous groups, gene ontology terms, and literature. The network analyzer module was used to analyze the descriptive topologies and hierarchical network attributes of each network.

RESULTS AND DISCUSSION

The reconstructed global regulatory network consisted of 2383 nodes and 5463 nodes of varying sizes based on the predicted scores for regulons (Fig. 1). This network contained 12 predicted transcriptional regulators targeting 2369 genes from the CBL genome. It also describes the regulatory
status of target genes regulated by respective regulators, similar to *B. subtilis* 168 (Table 1). The network robustness was validated using topological parameters. The functions of the predicted regulators were categorized using gene ontology terms. The transcriptional regulators ResD, SpoOA, ComK, CcpC, DinR, DegU, CitT, CodY, GerE, GltC, GltR, IolR, and LevR are found in CBL and are homologs to *B. subtilis* 168. These regulators have been shown to regulate putative membrane proteins, beta-lactamase, methyl-accepting chemotaxis protein, DNA replication protein DnaD, and sensor histidine kinase.

The reconstruction and comprehensive characterization of transcriptional regulatory networks of *C. botulinum* are important for clarifying the mechanisms of intoxication and prioritizing targets for novel therapeutics and food preservatives (Saujet *et al.*, 2014; Ihekwaba *et al.*, 2016). ResD response regulators activate the transcription of diverse genes encoding enzymes involved in O₂-reduction and repairing of oxidized damaged molecules in toxigenic *Clostridia* (Henares *et al.*, 2014). It can also redirect their central metabolism onto pathways with less O₂-sensitive microenvironments in *C. botulinum* (Morvan *et al.*, 2021). Spo0A is a highly conserved transcriptional regulator that plays a key role in initiating sporulation, virulence, and host interactions in toxigenic *Clostridia* (Deakin *et al.*, 2012; Pettit *et al.*, 2014; Kirk *et al.*, 2014). ComK activates the downstream competence genes in the competence gene regulatory network for prophage excision in *C. difficile* (Serrano *et al.*, 2016). CcpA is a pleiotropic regulator involved in carbohydrate and amino acid metabolism, facilitating a link between carbon and nitrogen pathways. It may repress toxin expression in response to the phosphotransferase system sugar availability in CBE, similar to *C. difficile* (Antunes *et al.*, 2011; 2012).

In this organism, DinR functions as a repressor by binding to the promoter region of each SOS gene and the diffocin operator. It can induce the expression of diffocin, a bactericidal agent, and the SOS response in DNA damage repair, such as in *C. difficile* (Gebhart *et al.*, 2012). DegU controls the serial expression of genes involved in flagellum and biofilm formation in *C. difficile*, similar to *B. subtilis* (Wang *et al.*, 2019). CitT may regulate citrate uptake in CBE; however, its regulatory mechanism remains unclear. CodY is a globally conserved regulator that positively regulates botulinum neurotoxin, sporulation, and pathogenicity in type A strains (Zhang *et al.*, 2014). The predicted CodY regulator in CBL may be involved in glucose/pyruvate metabolism and changes from the exponential growth phase to the stationary growth phase (Sonenshein, 2005; Girinathan *et al.*, 2021). This promoter stimulates toxin gene transcription and BoNT/A synthesis in type A2 strains by binding to the promoter of the ntnh-bont operon at high levels of GTP (Chapeton *et al.*, 2020). A key survival mechanism of CBL is its ability to form heat-
resistant spores. The predicted GerE regulator controls the genes involved in early-stage sporulation (Kirk et al., 2012). GltR and GltC regulate botulinum neurotoxin synthesis and toxin complex formation in strains type A and B (Fredrick et al., 2017). IolR controls myo-inositol metabolism and inhibits biofilm formation in non-proteolytic C. botulinum (Stringer et al., 2013).

ResD, SpoOA, ComK, CcpC, IolR, and LevR control the β-lactamase gene family (CLK_0366) in CBL. The penicillin-resistant type A strain may produce β-lactamase, conferring its antibacterial susceptibility to infant botulism under these transcription regulators. This is in agreement with a previous study (Barash et al., 2018). ResD, ComK, CcpC, Hpr, IolR, and LevR regulate methyl-accepting chemotaxis proteins (CLK_0562), which undergo reversible methylation during the adaptation of Clostridial cells to environmental stimuli (Salah Ud-Din and Roujeinikova, 2017). ResD, SpoOA, CcpA, ComK, Hpr, and LevR are associated with the putative DNA replication protein, DnaD. It is a potential target for antimicrobials against drug-resistant organisms (van Eijk et al., 2017; Oliveira Paiva et al., 2020). ResD, SpoOA, CcpC, GerE, Hpr, and IolR regulators interact with the sensor histidine kinase of CBL. Sensor histidine kinase is responsible for Spo0A phosphorylation and initiation of sporulation in C. botulinum and requires interactions between the Spo0A domains and other conserved proteins (Wörner et al., 2006). Interestingly, ComK, CcpC, Hpr, TnrR, IolR, and LevR interacted with putative membrane proteins (CLK_3081) in CBL. Surface membrane proteins are potential vaccine targets for proteolytic and non-proteolytic C. botulinum (Muhammad et al., 2014; Prathiviraj et al., 2016; Prisilla and Chellapandi, 2019; Bhardwaj et al., 2019).

CONCLUSION

The pathogenesis and synthesis of virulence factors in C. botulinum should be explored in the human intestinal environment to prioritize novel therapeutic targets for human botulism. The global regulatory network of CBL serves as a computational framework for a better understanding of its molecular virulence mechanisms and host adaptability during the intoxication process in the context of host responses. CodY is a known regulator of CBL, which regulates botulinum neurotoxin, sporulation, and pathogenicity. The present study predicted new 11 transcriptional regulators and characterized their regulated gene-associated physiological functions in CBL. The predicted regulators are homologs to those found in B. subtilis. Hub gene network analysis predicted five target genes that mediate antibacterial susceptibility, botulinum neurotoxin formation, and sporulation in CBL. Hence, it provides insight into the transcriptional regulatory processes in the human gut, leading to the establishment of personalized medicine for human botulism.

ACKNOWLEDGMENTS

The authors would like to thank the Tamil Nadu State Council for Higher Education (RGP/2019-20/BDU/HECP-0042), the Government of Tamil Nadu for their financial assistance.

Conflict of interest

The authors confirm that this article has no conflicts of interest.

<table>
<thead>
<tr>
<th>PRODORIC ID</th>
<th>Regulator</th>
<th>Molecular functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>MX000009</td>
<td>ResD</td>
<td>Regulation of oxygen limited condition</td>
</tr>
<tr>
<td>MX000019</td>
<td>SpoOA</td>
<td>Regulation of entrance into sporulation and inhibition of DNA replication</td>
</tr>
<tr>
<td>MX000023</td>
<td>ComK</td>
<td>Regulation of genetic competence and DNA uptake</td>
</tr>
<tr>
<td>MX000024</td>
<td>CcpC</td>
<td>Catabolite control regulation of TCA branched enzymes in the TCA cycle</td>
</tr>
<tr>
<td>MX000025</td>
<td>DinR</td>
<td>Regulation of DNA damage repair</td>
</tr>
<tr>
<td>MX000030</td>
<td>DegU</td>
<td>Regulation of transition growth phase, flagellum and biofilm formation</td>
</tr>
<tr>
<td>MX000044</td>
<td>CitT</td>
<td>Regulation of citrate uptake</td>
</tr>
<tr>
<td>MX000045</td>
<td>CodY</td>
<td>Regulation of major metabolic and virulence pathways upon nutrient limitation</td>
</tr>
<tr>
<td>MX000047</td>
<td>GerE</td>
<td>Regulation of spore coat proteins</td>
</tr>
<tr>
<td>MX000048</td>
<td>GltC</td>
<td>Regulation of glutamate biosynthesis</td>
</tr>
<tr>
<td>MX000049</td>
<td>GltR</td>
<td>Regulation of glucose metabolism</td>
</tr>
<tr>
<td>MX000051</td>
<td>IolR</td>
<td>Regulation of myo-inositol metabolism and inhibition of biofilm formation</td>
</tr>
<tr>
<td>MX000052</td>
<td>LevR</td>
<td>Regulation of levans and fructose metabolism</td>
</tr>
</tbody>
</table>
Ethics approval and consent to participate

The need for ethical approval and individual consent was waived.

Supplementary data

REFERENCES


