J. Dairy Sci. 100:1–12
https://doi.org/10.3168/jds.2017-12801

Potential factors involved in virulence of *Cronobacter sakazakii* isolates by comparative transcriptome analysis

Yingwang Ye,*†[1] Xiyan Zhang,* Maofeng Zhang,* Na Ling,† Haiyan Zeng,† Jina Gao,* Rui Jiao,* Qingping Wu,† and Jumei Zhang†
*School of Food Science and Technology, Hefei University of Technology, Hefei, 230009, China
†State Key Laboratory of Applied Microbiology, South China (the Ministry–Province Joint Development), Provincial Key Laboratory of Microbiology Culture Collection and Application, Guangdong Institute of Microbiology, Guangzhou, 510070, China

ABSTRACT

Cronobacter species are important foodborne pathogens causing severe infections in neonates through consumption of contaminated powdered infant formula. However, the virulence-associated factors in *Cronobacter* are largely unknown. In this study, the transcriptome analysis between highly virulent *Cronobacter sakazakii* G362 and attenuated L3101 strains was used to reveal the potential factors involved in virulence. The total transcripts were grouped into 20 clusters of orthologous group categories and summarized in 3 gene ontology categories (biological process, cellular component, and molecular function). In addition, the differentially expressed genes (DEG) between these isolates were analyzed using Volcano plots and gene ontology enrichment. The predominant DEG were flagella-associated genes such as *flhD*, *motA*, *flgM*, *flgB*, and *fliC*. Furthermore, the expression abundance of outer membrane protein or lipoprotein genes (*ompW*, *slyB*, *blc*, *tolC*, and *lolA*), potential virulence-related factors (*hly* III and *hha*), and regulation factors (*sdiA*, *cheY*, *Bss*, *FliZ*) was also significantly different between G362 and L3101. Interestingly, 3 hypothetical protein genes (ESA_01022, ESA_01609, and ESA_00609) were found to be expressed only in G362. Our findings provide valuable transcriptomic information about potential virulence factor genes, which will be needed in future molecular biology studies designed to understand the pathogenic mechanism of *Cronobacter*.

Key words: *Cronobacter*, transcriptome analysis, virulence factor, flagella, outer membrane protein

INTRODUCTION

Cronobacter spp. (formerly known as *Enterobacter sakazakii*) are foodborne pathogens associated with severe infections in neonates through consumption of contaminated powdered infant formula (FAO/WHO, 2008; CDC, 2012), and hazards of *Cronobacter* species in dairy products are of great concerns. Recently, it has been reported that the genus of *Cronobacter* consisted of 7 species: *Cronobacter sakazakii*, *Cronobacter malonaticus*, *Cronobacter turicensis*, *Cronobacter muytjensii*, *Cronobacter dublinensis*, *Cronobacter condimenti*, and *Cronobacter universalis* (Iversen et al., 2008; Joseph et al., 2012), and genomic analysis showed that only strains from *C. sakazakii*, *C. malonaticus*, and *C. turicensis* were associated with severe infections (Kucerova et al., 2010).

Pagotto et al. (2003) first described the virulence differences among *Cronobacter* isolates. In addition, essential roles of outer membrane proteins (ompA and ompX) were observed in invasion and adhesion to Caco-2 and INT-407 (Mohan Nair and Venkitanarayanan, 2007; Mittal et al., 2009; Kim et al., 2010). The cell-bound zinc-containing metalloprotease encoded by the *zpx* gene was considered to be the potential virulence factors (Kothary et al., 2007). Putative virulence genes on the related RepFIB plasmids harbored by *Cronobacter* species were characterized by Franco et al. (2011). The LysR-type transcriptional regulator homolog was found to play an important role in the virulence of *C. sakazakii* ATCC29544 (Choi et al., 2012). Emami et al. (2012) reported that, compared with wild-type mice infected by *C. sakazakii*, depletion of PMN and macrophages in newborn mice caused increased recruitment of dendritic cells in the intestine. Our groups determined membrane protein expression differences for enzV, ompX, lptE, pstB, and OsmY between G362 and L3101, which were potentially involved in virulence (Ye et al., 2015). Ye et al. (2016) further reported that the potential proteins (Dps, ompA, and LuxS) might be related to virulence

Received February 25, 2017.
Accepted July 11, 2017.
[1]Corresponding author: yeyw04@mails.gucas.ac.cn

of *C. sakazakii* strains. However, global information about the molecular basis of virulence in *Cronobacter* is currently not available at the transcriptome level.

The RNA-seq technology has been proven to be a powerful tool for transcriptome analysis (Dugar et al., 2013). This direct sequencing methodology enables absolute gene expression information at the mRNA level. In addition, this method is more sensitive for detecting low-abundant transcripts and small changes in gene expression under different conditions ('t Hoen et al., 2008; Wilhelm et al., 2008). The RNA-seq technology has been recently applied to determine comparative transcriptome analysis in *Campylobacter jejuni* (Chaudhuri et al., 2011) and *Salmonella enterica* serovar Typhi (Perkins et al., 2009). The responses of *C. sakazakii* to garlic-derived organosulfur compounds were determined using high-throughput whole-transcriptome sequencing and confocal micro-raman spectroscopy (Feng et al., 2014).

In this study, Illumina mRNA sequencing (RNA-seq) was used for transcriptome analysis between highly virulent *Cronobacter sakazakii* G362 and attenuated L3101 strains. Our findings will contribute to our understanding about its pathogenicity and valuable information for revealing aspects of virulence in *C. sakazakii*.

MATERIALS AND METHODS

C. sakazakii Isolates Used in This Study

Cronobacter sakazakii G362 and L3101 strains were identified using the International Organization for Standardization method, API-20E, chromogenic media, and PCR targeting *rpoB* (Stoop et al., 2009) and sequencing of *fusA* (Baldwin et al., 2009). The virulence of *Cronobacter* strains G362 (virulent isolate) and L3101 (attenuated isolate) were determined through intraperitoneal injection of mice followed by histopathologic analysis of specific organs, which was previously reported by Ye et al. (2016), and the 2 isolates were further used for transcriptome analysis.

RNA Extraction of G362 and L3101

Cronobacter sakazakii G362 and L3101 strains were inoculated into tryptic soy broth (**TSB**, Huankai, Guangzhou, China) for 16 h at 37°C. The total RNA was extracted from 1.0 mL of TSB enrichment culture with RNA isolation kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Extraction of RNA was performed in triplicate for each isolate. Then, genomic DNA and rRNA of the RNA samples were depleted using DNase I (Invitrogen, Carlsbad, CA) and a Ribominus Transcriptome Isolation Kit (Invitrogen) according to the manufacturer's instructions.

cDNA Library Preparation for RNA-seq

Construction of cDNA libraries was performed using Truseq RNA sample preparation kit according to the manufacturer's instructions (Illumina, San Diego, CA) as previously described by Dong et al. (2013). Then, the final purified DNA was captured on an Illumina flow cell for cluster generation and cDNA libraries were sequenced on the Illumina sequencing platform HiSeq 2000.

Data Processing and Analysis

The raw data were filtered to remove "dirty" reads or low-quality reads. The clean reads obtained using RNA-seq were compiled using bowtie2–2.1.0 (Slashdot Media, La Jolla, CA) read aligner software. Then, the reads were aligned to complete genome of *C. sakazakii* ATCC BAA-894. The number of each read was normalized to reads per kilobase of exon model per million mapped reads (**RPKM**), and thus the RPKM values were considered as the final expression levels for each gene between *C. sakazakii* G362 and L3101. Thereafter, all transcripts were subject to analysis of clusters of orthologous groups (**COG**), Kyoto Encyclopedia of Genes and Genomes (**KEGG**), and gene ontology (**GO**).

Analysis of Differential Expression Genes

A statistical analysis of the frequency of each read was performed to compare gene expression differences between G362 and L3101. According to the statistical framework, the *P*-values of differential expression significance of each transcript between G362 and L3101 were calculated according to the AC statistical framework (Audic and Claverie, 1997; Xu et al., 2013). Then, *P*-values less than 0.05 were regarded as the threshold to judge the significant differences in gene expression between strains at the mRNA level.

Quantitative Analysis of Partial Differential Expression Genes Using Real-Time PCR

For real-time PCR, primers for *fliC*, *ompW*, *sdiA*, *flgM*, and *flhD* genes are listed in Table 1, and the 16S rRNA gene was used as the internal control. Two strains were inoculated into TSB for overnight at 37°C, and total RNA was extracted from 1.0 mL of culture using a bacterial RNA extraction kit (Biomiga, San Di-