COMPARATIVE ANALYSIS OF *CRONOBACTER* SP. WITH RESPECT TO GENOMIC DIVERSITY AND PHYSIOLOGY

EVA KUCEROVA

A thesis submitted in partial fulfilment of the requirements Nottingham Trent University for the degree of Doctor of Philosophy. This research programme was carried out in collaboration with Nottingham University, UK and Vaccine Research Institute of San Diego, California, US

November 2010

i

ABSTRACT

In recent years, some outbreaks of bacterial infection in neonatal intensive care units have been traced to powdered infant formula contaminated with *Cronobacter* species, causing life-threatening diseases such as necrotizing enterocolitis and meningitis.

This study for the first time presents growth data and whole genome comparisons for five different species of *Cronobacter* after the taxonomic revision by Iversen *et al.* (2008). Growth data for 28 bacterial strains across 21 bacterial species in casein- and whey-dominant infant formula at temperatures 21, 27, 37, 41 and 44°C were determined, covering category A and B organisms as defined by WHO 2006. The data revealed potential of some *Cronobacter* species to grow at 44°C, posing a significant risk of infection by the bacterium when following the current formula preparation guidelines. The results were presented to the Food Standards Agency with the aim to update the current risk assessment model and improve infant formula preparation guidelines.

The first sequenced *Cronobacter* genome (*C. sakazakii* BAA-894) was used to construct a 384,030 probe oligonucleotide tiling DNA microarray covering its 4 Mb chromosome and plasmids pESA2 (31 kb) and pESA3 (131 kb). Comparative genomic hybridization (CGH) was undertaken on five *C. sakazakii* strains, and representatives of four other *Cronobacter* species. CGH highlighted 15 clusters of genes that were divergent or absent in more than half of the tested strains. Six of these were of probable prophage origin; other regions included type VI secretion systems, the O-antigen gene locus, a tellurite resistance cluster, a fimbriae cluster, and a copper resistance operon *cop*. The CGH analysis highlighted the role of horizontal gene transfer, as a significant part of the variable gene pool was due to acquisition of mobile DNA. A number of genes unique to *Cronobacter* species associated with neonatal infections (*C. sakazakii*, *C. malonaticus* and *C. turicensis*) were identified. These included a copper and silver resistance operon *cus* linked to invasion of the blood-brain barrier by neonatal meningitic strains of *Escherichia coli*. By comparing the neonatal intensive care unit outbreak strains with less virulent strains and by analysis of the variable regions, we identified a list of putative virulence factors that may improve our limited understanding of *Cronobacter*' s pathogenesis.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Prof. Steve Forsythe for his support and guidance throughout the whole course of my PhD, inspiring comments on my work and supernaturally quick corrections of my thesis. It has been a great journey for me and I thoroughly enjoyed (almost) every part of it. I would also like to thank to my external supervisor Prof. Mike McClelland for his invaluable help with CGH data analysis and with the publication of the manuscript.

Another thank you belongs to Dr. Gina Manning for being a great inspiration for me, both professionally and personally.

I would also like to express my gratitude to all my colleagues and staff in the Microbiology department for creating such a lovely working atmosphere.

A very special thank you belongs to my fellow PhD students Abiyad Baig, Gemma Croxall, Naomi Dunning, Muriel Funck, Jenny Ince, Susan Joseph and Sandra Reuter for being my true friends and giving me inspiration and support. I have enjoyed working with you so much.

An even more special thank you belongs to Dr. Tim Baker, who supported me in every possible way during the course of my PhD. Having you by my side made the writing-up of this thesis so much easier. I apologise for the lack of cooked meals and meaningful conversations that you patiently endured for several months. It is indeed reversible.

Finally, I would like to dedicate this thesis to a very special lady, my mum Dr. Jitka Kucerova. I will never be able to thank you enough for all the support you have been giving me throughout my entire life. Thank you for all the opportunities that you have given to me, it is you who made this possible.

CONTENTS

GENERAL INDEX

INDEX OF TABLES

INDEX OF FIGURES

and green indicates presence of a gene. For gene status determination see 3.4.1.9.
Csak1 = *C. sakazakii* ATCC 29544$^{\text{T}}$; Csak2 = *C. sakazakii* ATCC 12868, Cmal = *C. malonaticus* ; Cdub = *C. dublinensis*; Cmuy = *C. muytjensii*; Ctur = *C. turicensis*.

LIST OF ABBREVIATIONS

ACLAME   A CLAssification of Mobile genetic Elements

ATP   Adenosine Triphosphate

BLAST   Basic Local Alignment Search Tool

BMEC   Brain Microvascular Endothelial Cells

CAC   Codex Alimentarius Commission

CCC   Closed Circular Complex

CDC   Centers for Disease Control

CDS   Coding Sequence

CFU   Colony Forming Unit

CGH   Comparative Genomic Hybridisation

CNS   Central Nervous System

COG   Clusters of Orthologous Groups

CU   Chaperone-Usher

DNA   Deoxy-ribonucleic Acid

dNTP   Deoxynucleotid Triphosphate

EDTA   Ethylenediaminetetraacetic Acid

EPP   Estimated Probability of Presence

EPS   Exo-Polysaccharide

FAO   Food and Agriculture Organisation

GACK   Genomotyping Analysis by Charlie Kim

GEO   Gene Expression Omnibus

GYT   Glycerol Yeast Tryptone

IAHP	IcmF Associated Homologous Proteins
iNOS	Inducible Nitric Oxide Synthase
JEMRA	Joint FAO/WHO Expert Meetings on Microbiological Risk Assessment
KOH	Potassium Hydroxide
LB	Luria-Bertani medium
LOWESS	Locally Weighted Scatterplot Smoothing
LPS	Lipo-Polysaccharide
MGE	Mobile Genetic Elements
MLVA	Multiple-Locus Variable-number tandem repeat Analysis
MLST	Multi-Locus Sequence Typing
NaOH	Sodium Hydroxide
NEC	Necrotizing Enterocolitis
NICU	neonatal intensive care unit
NO	nitric oxide
NTU	Nottingham Trent University
OD	Optical density
ORF	Open Reading Frame
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PFGE	Pulsed-Field Gel Electrophoresis
PIF	Powdered Infant Formula
RABIT	Rapid Automated Bacterial Impedance Technique
RAPD	Random Amplification of Polymorphic DNA
rDNA	Ribosomal Deoxy-ribonucleic Acid
SDS	Sodium Dodecyl Sulfate

SNP	Single Nucleotide Polymorphism
TAE	Tris-acetate-EDTA (buffer)
T3SS	Type II Secretion System
T4SS	Type IV Secretion System
T6SS	Type VI Secretion S ystem
TSA	Trypticase Soy Agar
TSB	Trypticase Soy Broth
TTD	Time To Detection
VRBGA	Violet Red Bile Glucose Agar
WHO	World Health Organisation

1. INTRODUCTION

Cronobacter spp. (formerly *Enterobacter sakazakii*) are Gram-negative, peritrichous rods belonging to the *Enterobacteriaceae* family, producing characteristic yellow colonies on TSA agar (Figure 1). The organism has been associated with infections in all age groups, and with outbreaks of necrotizing enterocolitis, meningitis and septicaemia in pre-term, low birth neonates with exceptionally high mortality rates ranging from 40 to 80 percent. In recent years, some outbreaks of bacterial infection in neonatal intensive care units have been traced to rehydrated powdered infant formula (PIF) contaminated with *Cronobacter* (Biering *et al*. 1989, Caubilla-Barron *et al*. 2007). In 2002, the International Commission for Microbiological Specifications for Foods ranked the organism as a "severe hazard for restricted populations, life threatening or substantial chronic sequelae of long duration" and the organism joined the group of the most dangerous food pathogens such as *Listeria monocytogenes* or *Clostridium botulinum*. The FAO/WHO (2008) have estimated the annual incidence of *Cronobacter* infection over the period 1992-2007 to 17.6 per million population for neonates and 2.06 per million population for infants aged 1-11 months. *Cronobacter* infections have been documented in infants and young children up to 3 years of age, with at least 120 cases and 27 deaths (FAO/WHO 2008). The number of *Cronobacter*-related infections is likely to be significantly under reported, however, and the incidence is believed to be more common (FAO/WHO 2008). Better understanding of the virulence of this opportunistic pathogen is warranted to ensure the safety of products designed for the most vulnerable part of our population, neonates.

1.1. Literature review

This section will cover the brief history of *Cronobacter,* its environmental and clinical sources and its prevalence in the environment, and describe the challenges of the classification of this fascinating organism into the microbial taxonomic scheme. It will then cover the pathogenesis of the diseases caused by the bacterium and a summary of research into the virulence factors involved in the onset of the diseases. Finally, it will focus on the factors that contribute to *Cronobacter*'s remarkable persistence in powdered products, as well as the control measures designed to ensure the safety of powdered breast milk substitutes.
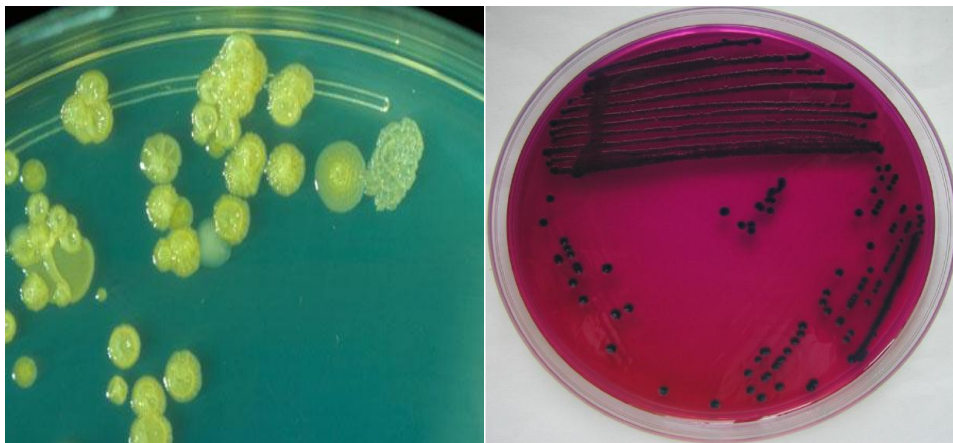
Figure 1. Culture morphology of *Cronobacter*. Left: *Cronobacter sakazakii* colonies on Trypticase Soy Agar; source: http://www.enterobactersakazakiiblog.com/articles/e-sakazakii-information/. Right: *Cronobacter* colonies on Violet Red Bile Glucose Agar; source: own culture.

1.1.1 History

In 1929, Pangalos reported a yellow-pigmented bacterium that caused a septicemia in an infant (Pangalos 1929). However, the first account mentioning a "yellow pigmented *Enterobacter cloacae*" as the causative agent in meningitis was published much later, in 1961, when Urmenyi and Franklin described two cases of terminal neonatal meningitis that occurred in England in 1957 (Urmenyi & Franklin 1961). Since then cases have been reported worldwide. In 2001, a male infant was admitted to a neonatal intensive care unit (NICU) because of low birthweight and respiratory problems. After 11 days the baby developed symptoms of meningitis and despite being given intravenous antibiotics it died 9 days later. *Cronobacter* was cultivated from the cerebrospinal fluid and further 10 cases of *Cronobacter* infection were reported on the neonatal unit (Himelright *et al*. 2002). All 10 cases were associated with the use of powdered formula. After the cultivation of *Cronobacter* from an unopened can of the same product, the manufacturer voluntarily recalled the batch in March 2002 and further recalls followed (Himelright *et al*. 2002).

Since then, contamination of powdered infant formula by *Cronobacter* has come under considerable attention of international regulatory bodies aiming to ensure the safety of this product, as well as the attention of research groups worldwide. The FAO/WHO 2004 expert committee recommended that research should be promoted to gain a better understanding of the taxonomy, ecology, virulence and other characteristics of *Cronobacter,* and since then remarkable progress has been made in the knowledge of the organism. As powdered infant formula is not a sterile product, this potential vector of infection has been in the spotlight of regulatory authorities, as well as media. It should be noted, however, that such neonatal infections are rare, and not all have been associated with reconstituted formula ingestion. Despite the fact that the focus of regulatory bodies is on the safety of PIF and other breast milk substitutes, the majority of cases of *Cronobacter* infection occur in fact in adults (FAO/WHO 2008).

1.1.2 Sources of the organism

Cronobacter is a ubiquitous organism that is present in a wide range of environments, including water, soil, and a variety of processed foods and fresh produce. The main reservoir of the organism still remains unclear, however. Iversen & Forsythe (2003) and more recently, Schmid *et al*. (2009) suggested a possible plant origin of the organism. *Cronobacter* is most commonly isolated from dried herbs and spices and is extremely resistant to dry stress, which are both expected characteristics of plant organisms.

The bacterium has been isolated from a wide range of clinical samples including cerebrospinal fluid, blood, bone marrow, sputum, urine, faeces and hospital environment. Many of the reported cases have been linked with the use of powdered infant formula (FAO/WHO 2004, Forsythe 2005, FAO/WHO 2006,) and hence the organism is considered an infant formula pathogen. Whether the contamination of the formula product is of intrinsic or extrinsic character is often unclear, however. The bacterium has also been isolated from factories producing milk powder and household vacuum cleaning bags (Kandhai *et al*. 2004) and from household utensils used for reconstitution of PIF (Muytjens & Kollee 1990, Bar-Oz *et al*. 2001, Block *et al*. 2002). Due to the ubiquitous presence of *Cronobacter* and its extreme resistance to dry conditions (Caubilla-Barron *et al*. 2007) contamination of food products including powdered infant formula is difficult to avoid.

1.1.2.1 Environmental sources and food contamination

Since the organism does not belong to the normal animal and human gut flora, it was hypothesized that it contaminates food primarily through soil, water, and vegetables and that flies and rodents may be a secondary vector of contamination (Iversen & Forsythe 2003). Hamilton and colleagues were the first to isolate *Cronobacter* from the midgut of the larvae *Stomoxys calcitrans* feeding on blood of livestock and other

domesticated animals, as well as humans (Hamilton *et al*. 2003). Interestingly, *S. calcitrans* is found in all countries reporting *Cronobacter* infections. Molloy and colleagues examined 518 samples from dairy farms, meat abattoirs and retail food stores for the presence of *Cronobacter* (Molloy *et al*. 2009). They isolated the species from 33 (6.4%) samples, but did not isolate the organism from cattle or pig faecal samples, or from dairy farm environment. The authors concluded that animal carriage is not contributing to the transmission of the pathogen and that raw milk is not the main source of infant formula contamination. Molloy *et al*. (2010) investigated whether *Cronobacter* can survive passage through a model of the bovine rumen and abomasums, and in bovine faeces in the farm environment. They concluded that *Cronobacter* spp. are unlikely to be shed in bovine faeces, but that the organism survives well in the farm environment.

Iversen and Forsythe (2003) hypothesized that the *Cronobacter* species might be of plant origin. Some physiological traits of the organism such as production of a polysaccharide capsule (Schmid *et al*. 2009, Hurrell *et al*. 2009a), production of yellow pigment (Lehner *et al*. 2006) and its extreme desiccation resistance (Caubilla-Barron *et al*. 2007, Osaili & Forsythe 2009) may enable the bacteria to stick to plant leaves, offer protection against oxygen radicals and help to survive dry periods. In order to provide evidence for the plant origin of *Cronobacter*, Schmid *et al*. (2009) investigated biochemical traits associated with plant microorganisms in nine strains representing five *Cronobacter* species. All strains were able to solubilise mineral phosphate, produce indole acetic acid and produce siderophores. In addition, the strains were able to endophytically colonise tomato and maize roots. The authors concluded that plants may be the natural habitat of *Cronobacter* spp. and that the rhizosphere might act as a reservoir of the bacterium (Schmid *et al*. 2009).

The unique ability of *Cronobacter* to survive dry stress enables it to survive on various types of dry materials, which contributes to the reservoir of the organism in the environment. Furthermore, it has been documented that some starches and proteins derived from plants that are added to powdered infant formula without

additional heat treatment can act as a source of contamination of these products (FAO/WHO 2004, FAO/WHO 2006).

1.1.2.2 Breast milk substitutes

According to the definition by Codex Alimentarius (1981), breastmilk substitutes are products intended for infants who do not have access to breastmilk. In most cases, infant formulas are used as breastmilk substitutes for normal healthy infants under 6 months, and are formulated in accordance with the appropriate Codex standards. Infant formula is a breast-milk substitute designed to satisfy the nutritional requirements of infants during the first months of life up to the introduction of complementary feeding. Human milk fortifiers are powdered supplements which can be added to expressed human milk if nutritional needs cannot be satisfied by human milk alone. This can include thickening agents as well as mineral supplements.

Cronobacter was first associated with contaminated breast milk substitutes by Muytjens *et al*. in 1983, when it was isolated from prepared milk formula and formula reconstitution equipment during a study of eight cases of meningitis and sepsis (Muytjens *et al*. 1983). In 1988 it was reported that 52.2% (n=141) powdered infant formula samples from 35 countries were contaminated with *Enterobacteriaceae*, with 14% containing *Cronobacter* (Muytjens *et al*. 1988). While the source of contamination in *Cronobacter*-related outbreaks has not always been confirmed, breast milk substitutes have been epidemiologically or microbiologically established as the source of infection in many cases (Muytjens *et al*. 1983, Biering *et al*. 1989, Simmons *et al*. 1989, Clark *et al*. 1990, Muytjens & Kollee 1990, van Acker *et al*. 2001, Weir 2002, Iversen & Forsythe 2003). A strong link between the presence of *Cronobacter* in formula milk and an outbreak of *Cronobacter* infection has been established by Centers for Disease Control (CDC) following the outbreak in a neonatal intensive care unit (NICU) in Tennessee in 2001, when a neonate died from *Cronobacter*-induced meningitis and further 10 cases of *Cronobacter* colonisation were found on the same unit (Centers for Disease Control (CDC) 2002). Later investigation revealed that the formula fed to the infant in Tennessee was a formula

6

not intended for neonates. Iversen and Forsythe (2004a) reported that 2.4% (n=82) powdered infant formulae, 10.2% (n=49) dried infant foods, 4.2% (n=72) milk powders, and 32.8% (n=122) herbs and spices contained *Cronobacter*. In 2009, an international survey for *Cronobacter* and related organisms in powdered infant formula, follow up formula and infant foods was conducted by 8 laboratories in 7 countries in response to the FAO/WHO 2008. In total, 290 products were analysed using a standardised procedure and *Cronobacter* was isolated from 3% (n=91) follow up formulae and 12% (n=199) infant foods and drinks (Chap *et al*. 2009). In another study, *Cronobacter* was isolated from just 1.4% (n=69) infant formulae and milk powders (Jaradat *et al*. 2009). The level of reported contamination is typically very low, under 1 CFU/g.

Miled-Bennour *et al*. (2010) analyzed 150 isolates previously identified as *Cronobacter* using the ISO/TS 22964 standard method, out of which 70% were isolated from the PIF production environment and 20% from PIF. Biochemical testing for indole production, malonate utilization and acid production from dulcitol and methyl-α-D-glucoside indicated that 82.5% of the strains were *C. sakazakii*, 8% *C. malonaticus*, 5% *C. muytjensii*, 3% *C. dublinensis*, and 1.5% *C. turicensis*. However, of the 47 strains isolated from PIF products or samples from PIF production plants, 91.5% were *C. sakazakii*, and 8.5% were *C. malonaticus*, which shows the prevalence of *C. sakazakii* and *C. malonaticus* in PIF factories and PIF products. Similar results were reported in other studies (Mullane *et al*. 2008a, Healy *et al*. 2009).

Powdered infant formula is not a sterile product and hence can contain viable bacterial cells. The contamination of formula products can be intrinsic, occurring during the manufacturing process, or extrinsic from equipment used for its reconstitution. The possible sources of intrinsic PIF contamination include a wide range of dry blended raw material and other environmental sources associated with the factory environment (Drudy *et al*. 2006). In summary, due to prevalence of *Cronobacter* in the milk-processing environment and its ability to persist in powdered products such as PIF, stringent microbiological criteria need to be adapted to ensure safety of these products.

1.1.2.3 Clinical sources

Cronobacter has been isolated from a wide range of clinical sources including cerebrospinal fluid, blood, bone marrow, sputum, urine, appendix, intestinal and respiratory tracts, eye, ear, wounds, ear and breast abscess, and faeces (Gurtler *et al.* 2005); and from various formula preparation utensils (Clark *et al.* 1990, Bar-Oz *et al.* 2001, Block *et al.* 2002).

Most reports on *Cronobacter* outbreaks have implicated infant formula or the hospital environment as a source of contamination, and therefore the main focus of regulatory bodies has been the safety of breast milk substitutes and equipment used for its preparation. Recent studies, however, have highlighted other sources of infection. During one outbreak at a neonatal intensive care unit (NICU) in Greece, *Cronobacter* was isolated from throat swabs of neonates colonised by the bacterium, but it was not present in their blood (Arseni *et al.* 1987). This suggests that neonatal throat flora can act as a source of contamination of the hospital environment and possibly other neonates. One of our recent studies done in collaboration with Aston University, Nottingham City Hospital and Queens Medical Centre has identified enteral naso-gastric feeding tubes used at neonatal care intensive units as a possible source of microbial infection. Moreover*, C. sakazakii* was isolated from tubes of neonate that was not fed FIP (Hurrell *et al.* 2009b). This study has also highlighted the issue of biofilm formation on the surface of enteral tubes, which can significantly increase the risk of contamination of the infant feed. Another study examining biofilm formation of *Cronobacter*, other *Enterobacteriaceae* and *Acinetobacter* on different types of enteral tubes was undertaken by our group at Nottingham Trent University (Hurrell *et al.* 2009b).

The ubiquitous nature of *Cronobacter* and its ability to form biofilms on surfaces commonly used in formula preparation such as glass, some forms of plastic and stainless steel (Iversen & Forsythe 2003, Forsythe 2005) mean an increased risk of contamination of prepared feeds by the bacterium, especially if hygienic practices are not followed.

1.1.3 Taxonomy

The taxonomical classification of the organism now known as *Cronobacter* has proven to be somewhat of a challenge. First the bacterium was known as yellow-pigmented *Enterobacter cloacae,* then as *Enterobacter sakazakii* and finally, *Cronobacter* spp. A variety of other confusing names used for this organism reflect the taxonomic uncertainty connected with *Cronobacter:* 'Urmenyi and Franklin bacillus', 'yellow coliform', 'yellow *Enterobacter'* and 'pigmented *cloacae* A organism'.

The bacterium was first believed to be a yellow-pigmented variety of *Enterobacter cloacae*, one of the many causative agents of infant meningitis. In 1974, Brenner *et al.* showed that yellow-pigmented strains of *E. cloacae* had less than 50% sequence identity with non-pigmented strains and it was suggested that yellow pigmented *E. cloacae* should comprise a new species (Brenner 1974). In 1980, a comparison of '*Enterobacter sakazakii*' to the type strains of the genera *Enterobacter* and *Citrobacter by* DNA-DNA hybridization (Farmer 1980) revealed 41% similarity to *Citrobacter freundii* and 54% similarity to *E. cloacae,* respectively. Since the organism was also phenotypically closer to *Enterobacter cloacae*, it was assigned to the *Enterobacter genus* (Farmer 1980). Farmer and colleagues designated it a new species and named it after the Japanese bacteriologist Riichi Sakazaki (Farmer 1980). "*E. sakazakii*" was later distinguished from *E. cloacae* based on differences in pigment production, biotype assignment and antimicrobial resistance (Farmer 1980, Izard *et al*. 1983). More recent analysis of both partial 16S rDNA and hsp60 gene sequencing showed that *'E. sakazakii'* isolates formed at least four distinct clusters (Iversen *et al*. 2004c), and the authors concluded that the clusters may in fact represent different closely related species. The clusters included the 15 different biogroups previously defined by Farmer (1980) based on biotyping, and a 16th biogroup was added in subsequent work (Iversen *et al*. 2006). However, full taxonomic revision required considerable further analysis. In 2007, a re-classification of *"Enterobacter sakazakii"* into a new genus *Cronobacter* based on DNA-DNA hybridization and phenotyping was proposed (Iversen *et al*. 2007) and revised in 2008

9

(Iversen *et al*. 2008). The new genus harbours five distinct species: *Cronobacter sakazakii*, *C. malonaticus, C. turicensis*, *C. muytjensii*, *C. dublinensis* and a possible sixth *Cronobacter* group, currently named 'genomospecies I' that only includes two strains. The genus would derive its name from a Greek mythological god Kronos, who was said to swallow his children at birth. With a few exceptions, the former biotypes and genomogroups correspond with the new species (Baldwin *et al*. 2009). Distinguishing between the two species *C. sakazakii* and *C. malonaticus* is still problematic, however.

All *Cronobacter* species have been associated with clinical infections in infants or adults and are considered pathogenic (FAO/WHO 2008). However, it is plausible that some species are more virulent to neonates, as determined by *in vitro* virulence studies (Townsend *et al*. 2007a, Townsend *et al*. 2008a). To date, isolates from infected neonates have been limited to *C. sakazakii*, *C. malonaticus* and *C. turicensis*, although not all *"E. sakazakii"* isolates have been classified according to the new nomenclature.

1.1.4 Molecular typing methods

Although it is possible to differentiate all six *Cronobacter* species by biochemical profiling (Iversen *et al*. 2008), molecular methods are being increasingly used as a more rapid and reliable tool to study bacterial genomic diversity and to track sources of infection. *Cronobacter* species have been successfully characterized by random amplification of polymorphic DNA (RAPD), ribotyping, pulsed-field gel electrophoresis (PFGE), multiple-locus variable-number tandem-repeat analysis (MLVA) and multiple-locus sequencing technique (MLST). Among these, PFGE is considered the gold standard for genetic typing (Nazarowec-White, Farber 1999, Mullane *et al*. 2008a) and is recommended to be used in surveillance and investigations of sources of outbreaks. Terragno *et al*. (2009) analysed 23 *Cronobacter* isolates by PFGE and showed that both enzymes used, XbaI and SpeI, generated discrete DNA band patterns that discriminated all the isolates analyzed. 16S

rDNA sequences have been traditionally used to determine phylogenetic relationships between organisms including *Enterobacteriaceae*. 16S rDNA sequencing is of limited use for more closely related organisms, however, because of minor differences in the rDNA sequence. 16S rDNA gene sequencing can distinguish between 5 *Cronobacter* spp., but difficulties can arise whilst trying to separate *C. sakazakii* and *C. malonaticus* species. As an example, automated ribotyping results defined groups of *Cronobacter* strains which corresponded to biochemical testing of the strains (Miled-Bennour *et al.* 2010). However, at least three strains clustered outside the expected group defined by biochemical testing. The authors concluded that due to the limited discrimination of the 16S rDNA sequencing method at the intra-species level, other typing methods should be considered as the standard method for *Cronobacter* strains identification. Sequencing of protein coding genes is a useful, more discriminatory alternative to ribotyping especially if sequences from multiple protein-coding genes are used. In collaboration with University of Warwick, our group developed a full MLST scheme based on sequences of 7 loci (Baldwin *et al.* 2009). The *Cronobacter* MLST analysis is based on 7 housekeeping genes; ATP synthase beta chain (*atp*D), elongation factor G (*fus*A), glutaminyl-tRNA synthetase (*gln*S), glutamate synthase large subunit (*glt*B), DNA gyrase subunit B (*gyr*B), translation initiation factor IF-2 (*inf*B) and phosphoenolpyruvate synthase A (*pps*A). This MLST scheme is accessible online (http://pubmlst.org/cronobacter/) and has been designed to cover all *Cronobacter* species. The MLST analysis has revealed a remarkably strong clonal nature in *Cronobacter*. For example, it showed that some sequence types were primarily associated with infant formula (ST3) or were primarily composed of clinical isolates (ST8). This indicates that there may be a source of *Cronobacter* which is not PIF associated and therefore focusing on PIF analysis after a *Cronobacter* outbreak may lead to oversight of the true source of the infection.

1.1.5 Pathogenesis and virulence factors

The majority of reported infections caused by *Cronobacter* occur in adults. These comprise a wide range of symptoms, from conjunctivitis, biliary sepsis, urosepsis and appendicitis to wound infection and pneumonia; infections in neonates include infant meningitis and necrotizing enterocolitis (Gurtler *et al*. 2005). At least 120 cases of *Cronobacter*-related infections and 27 deaths were reported in infants and young children up to 3 years of age (FAO/WHO 2008). Infections by *Cronobacter* can have a very severe outcome, particularly for neonates and the immuno-compromised. Although very rare, infections by *Cronobacter* spp. are a cause for public health concern due to their extreme severity and resulting mortality and morbidity. Although *Cronobacter* can cause infection in all age groups (FAO/WHO 2008), neonates and infants less than 2 months old, premature infants and those with underlying medical conditions are at the highest risk (FAO/WHO 2006, Bowen & Braden 2006). A newborn infant has a sterile gastrointestinal tract that is quickly colonised by bacteria. Furthermore, being less acidic than that of adults, the stomach of newborns, especially of those born prematurely, can enable the survival of *Cronobacter* cells. Most documented outbreaks of *Cronobacter* infection occurred in hospitals at neonatal intensive care units. However, Adamson & Rogers (1981) reported separate cases of previously healthy infants admitted to hospital with meningitis caused by *Cronobacter*, showing that infection can also be acquired in the home environment. Finally, some *Cronobacter* infections occurred in adults, namely wound infections or bacteraemia. Adult patients at increased risk include those previously treated with antibiotics, immuno-compromised and elderly patients, those with medical implants or with acute, chronic, or serious illnesses (Pitout *et al*. 1997, Sanders & Sanders 1997).

1.1.5.1 Infant meningitis

Infant meningitis can be caused by a variety of bacterial pathogens, including *Cronobacter* and its close relatives *Enterobacter cloacae* and *Citrobacter koseri*.

Cronobacter-related meningitis is characterized by a mortality rate of 40-80% (Iversen & Forsythe 2003) and generally a very poor clinical outcome (Figure 2). The bacterium causes cystic changes, abscesses, fluid collection, brain infarctions, hydrocephalus, necrosis of brain tissue and liquefaction of white cerebral matter (Iversen & Forsythe 2003). Patients surviving *Cronobacter*-related meningitis often suffer from severe neurological sequelae, such as hydrocephalus, quadriplegia and retarded neural development (Lai 2001). The infection usually arises between the fourth and fifth day after birth and it can be fatal within hours to days following the first clinical signs (Muytjens *et al*. 1983). Compared with patients suffering from *Cronobacter*-induced enterocolitis, infants in whom meningitis developed tend to have normal gestational age and birth weight (Bowen & Braden 2006).

Figure 2. Magnetic resonance scan of neonatal brain. The darker area is cerebral matter that has been liquefied as a result of *Cronobacter* infection. Source: (Farber & Forsythe 2008)

1.1.5.2 Necrotizing enterocolitis

Necrotizing enterocolitis is caused by several different pathogens. It is the most common gastrointestinal disease in newborns, affecting approximately 2-5% of premature neonates and leading to death in 10-55% cases (Peter *et al*. 1999). It is characterized by ischaemia, bacterial colonisation of the intestinal tract, and increased

levels of proteins in the gastrointestinal lumen. *Cronobacter* has been implicated as a causative agent of necrotizing enterocolitis (NEC), but its role in the pathogenesis of the disease is somewhat unclear. There are some reports of *Cronobacter* isolations from babies who developed NEC (van Acker *et al*. 2001, Caubilla-Barron *et al*. 2007). NEC is ten times more common in babies fed with formula milk compared with those fed with breast milk (Lucas & Cole 1990). This suggests that there is an association between *Cronobacter* occurrence and NEC, although until recently, the organism has not been conclusively proven to cause the disease and the pathogenesis of NEC is still poorly understood.

Hunter *et al*. (2008) demonstrated that *C. muytjensii* induces nitric oxide (NO) production in infected intestinal epithelial cells and that the NO production is responsible for the apoptosis of the intestinal cells. The increased NO production was due to the production of an isoform of NO synthase, iNOS (inducible NO synthase), which is only expressed during inflammation. The iNOS expression was suppressed when *Lactobacillus bulgaricus*, a probiotic bacterium commonly used in yoghurt cultures, was added prior to infection with the *Cronobacter*. It was also demonstrated that *L. bulgaricus* pretreatment protects against the loss of microvillus structure and other structural changes in the intestinal epithelial cells caused by *Cronobacter*. This was confirmed by an *in vivo* study, when pretreatment of rat pups with *L. bulgaricus* decreased the degree of intestinal cell injury caused by *Cronobacter* (Hunter *et al*. 2009). The authors also showed that *Cronobacter* induces apoptosis by the expression of tumor necrosis factors, apoptotic genes and iNOS genes, and that the expression of these genes was suppressed when the cell lines were pretreated with *L. bulgaricus*. It is of note, however, that these two studies have been in fact done with a single strain of *C. muytjensii* which has not been associated with neonatal infection, and has not been extended to cover other *Cronobacter* species.

1.1.5.3 Virulence studies

Still very little is known about virulence factors and pathogenicity of *Cronobacter*. It is known that *Cronobacter* strains and species vary in their virulence, as determined

by epidemiological studies and in-house mammalian tissue culture (Caubilla-Barron *et al.* 2007, Townsend *et al.* 2007a, Townsend *et al.* 2008a), but the mechanisms of virulence are unknown. The bacterium is believed to invade the tissues using pathogenic secretory factors (elastases, glycopeptides, endotoxins, collagenases and proteases), which increases the permeability of the blood-brain barrier and allows the bacterium to gain access to the nutrient-rich cerebral matter (Iversen & Forsythe 2003). Some reports suggest a similarity between the tropism of *Cronobacter* and *Citrobacter koseri* for invasion and infection of the central nervous system (Willis & Robinson 1988, Burdette & Santos 2000). Kline (1988) also noted that brain abscesses caused by *Cronobacter* and *Citrobacter koseri* were morphologically similar and may be due to similar virulence mechanisms. Pagotto *et al.* (2003) were the first to describe putative virulence factors for *Cronobacter* when they found that enterotoxin analogue compounds were produced by 4/18 strains in a suckling mouse assay. The authors found that the lethal dose was 10^8 CFU per mouse for all strains when injected intraperitoneally, but only two *Cronobacter* strains caused death when given orally. However, FAO/WHO (2004) have proposed a much smaller tentative oral dose for neonates ranging from 10^3 to 10^8 CFU. Although the production of enterotoxin by some *Cronobacter* strains described by Pagotto *et al.* (2003) is widely acknowledged, the genes encoding the putative toxin have yet to be identified. The type strain ATCC 29544[T] showed no enterotoxin production in the study, which also suggests that there are considerable differences in virulence among *Cronobacter* strains, and some strains may be non-pathogenic. Kothary *et al.* (2007) characterized a zinc metalloprotease *zpx* unique to 135 *Cronobacter* strains tested, which could allow the bacterium to penetrate the blood-brain barrier and cause meningitis. Townsend *et al.* (2007b) used the neonatal rat model to study translocation of *Cronobacter* and other bacteria from the gut to deeper tissues following endotoxin (LPS) administration. The authors suggested that LPS may enhance the translocation of bacteria from the blood to the CNS within 24 h, but the mechanism which LPS utilizes to enhance bacterial translocation across the gut and blood brain barrier was not clear. It was hypothesized that the ingested LPS may disrupt the intestinal wall barrier and promote bacterial invasion in necrotizing enterocolitis and sepsis, and

subsequently weaken the blood brain barrier leading to meningitis (Townsend *et al*. 2007b).

Townsend *et al*. (2007a) showed that *Cronobacter* can attach to intestinal Caco2 cells and survive in macrophages. The authors provided evidence that *Cronobacter* can enter the CSF circulation and suggested that the organisms cause a massive influx of inflammatory cells into the ventricles and meninges, breaking down adhesion junctions and thus gaining access to the brain parenchyma. This study showed for the first time that *Cronobacter* is able to invade capillary endothelial cells and persist in macrophages up to 96 h. It also confirmed that the ability to cross blood-brain barrier and cause CNS infection varied significantly between the strains. Kim and Loessner (2008b) suggested that the invasion of *Cronobacter* to Caco-2 cells may be receptor-mediated, as the bacterial invasion showed characteristics of saturation kinetics. The authors also concluded that bacterial *de novo* protein synthesis was required for invasion. In the same study, pretreatment of Caco-2 cells with an actin polymerization inhibitor resulted in decreased invasiveness of *Listeria monocytogenes* and *Salmonella* Typhimurium, but enhanced the invasiveness of *Cronobacter.* The authors hypothesized that this enhancement was due to the disruption of tight junction, a membrane-associated structure that acts as a barrier against the molecular exchange between epithelial cells. This was confirmed when the disruption of the tight junction by EGTA significantly increased the invasive properties of *Cronobacter*.

Townsend *et al*. (2008a) studied seven *Cronobacter* strains associated with the largest reported NICU outbreak with the most reported deaths to date. All strains were able to attach to and invade the intestinal cell line Caco2 more than *E. coli* K12 and *Salmonella* Enteritidis, while two strains, both associated with fatal cases of meningitis and NEC, showed the highest invasion rates. These two strains also showed acquired extended spectrum β-lactamase activities. Two strains were also able to replicate within macrophages, while all other strains survived inside macrophages. Hunter *et al*. (2009) examined the effects of *C. muytjensii* on intestinal epithelial cells IEC-6 an *in vivo* on formula fed rat pups subjected to hypoxia. The rat pups subjected to *Cronobacter,* formula feeding and hypoxia had 70% mortality compared to 40% mortality of the pups subjected to formula feeding and hypoxia alone. Electron

microscopy revealed that *Cronobacter* induced morphological changes in the gut associated with infection such as villus disruption and formation of gaps in the epithelium. The bacterium induced apoptosis both *in vivo* and *in vitro*, in the intestinal cells and in enterocytes of rat pups. As the organism failed to invade the intestinal cells *in vitro*, the authors hypothesized that the bacterium causes intestinal tissue damage by binding to the enterocytes and inducing apoptosis. A limitation of this study, however, is that it used a single *C. muytjensii* strain, which was not associated with neonatal infection, and that the study was not extended to cover other *Cronobacter* species.

Despite the viruence studies examining the organism´s effect on tissue culture cell lines or *in vivo* models of NEC, there is still very little information on the genetic basis of *Cronobacter*'s virulence. Several studies highlighted the role of the outer membrane protein OmpA in the invasion of both intestinal epithelial and brain microthelial cells (Singamsetty *et al*. 2008, Mittal *et al*. 2009b, Wang & Kim 2002, Kim *et al*. 2010). Singamsetty *et al*. (2008) showed that *Cronobacter* strains with deletion of *ompA* gene had significantly lower invasion in HBMEC than the wild type strain. The authors also concluded that invasion of intestinal epithelial cells by the bacterium requires microfilaments, whereas invasion of HBMEC requires microtubules. Mittal *et al*. (2009a) showed that *C. muytjensii* carrying *ompA* gene could survive inside myeloid dendritic cells, whereas *C. muytjensii* with deleted *ompA* was killed within 2 hours upon infection, although *ompA*- cells did not enter the dendritic cells. Complementation of *ompA*- cells with a plasmid carrying *ompA* restored the ability of *C. muytjensii* to persist in dendritic cells. The authors demonstrated that *ompA* was required for the survival inside, but not for the invasion of, dendritic cells. As before, this study was limited to a single strain of *C. muytjensii*, however. In a following *in vivo* study using the same *C. muytjensii* strain, Mittal *et al*. (2009b) confirmed that the expression of OmpA is crucial for the onset of meningitis using the rat pup model, when 100% of rat pups infected by OmpA+ *Cronobacter* suffered meningitis, and none of the rats infected with OmpA- developed the disease.

1.1.6 Antibiotic susceptibility

When an infection by *Cronobacter* occurs, it is essential to provide rapid antibiotic treatment. Although the bacterium tends to be more sensitive to most antibiotics compared to other *Enteriobacteriaceae,* its increasing resistance to some antibiotics causes concern. *Cronobacter*-related infections have been traditionally treated with ampicillin combined with gentamycin or chloramphenicol (Lai 2001). Unfortunately, the organism has developed resistance to ampicillin (Muytjens *et al*. 1983, Lai 2001) and gentamicin use is limited as it fails to reach sufficient concentrations in the cerebro-spinal fluid (Iversen & Forsythe 2003). In 1980, all strains tested by Farmer *et al*. were susceptible to ampicillin (Farmer 1980). Lai (2001) described five cases *of Cronobacter* infection in which one or more of the isolates were resistant to ampicillin and most cephalosporins of 1st and 2nd generation; and in 2008, Kim *et al*. (2008a) reported frequent resistance of *Cronobacter* food isolates to ampicillin and cephalotin. For this reason, the shift to carbapenems or 3rd generation cephalosporins with an aminoglycoside or trimethoprim with sulfamethoxazole was proposed. This treatment improved the outcome of *Cronobacter* meningitis, but it also caused the increase in resistance to these antimicrobials (Lai 2001). It was reported that all *Cronobacter* strains associated with an outbreak at a NICU in France were sensitive to ciprofloxacin, amikacin, gentamicin, imipenem, piperacillin, and trimethoprim and resistant to doxycycline. Two strains (n=31) were resistant to cefpodoxime, ceftazidime, and chloramphenicol (Caubilla-Barron *et al*. 2007).

First reports concerning the ability of *Cronobacter* to produce β-lactamases gave conflicting results. However, the presence of β-lactamases in *Cronobacter* was reported in a study by Pitout *et al*. (1997) when all tested strains were positive for Bush group 1 ß-lactamase (cephalosporinase). Lai (2001) reported increasing β-lactamase production among *Cronobacter* strains. Similarly, Block *et al*. (2002) reported that all *Cronobacter* isolates tested were β-lactamase positive. However, Stock and Wiedemann (2002), did not find any evidence β-lactam production in the 35 *Cronobacter* strains tested. *Cronobacter* spp, are phenotypically a very diverse group of bacteria that currently includes five distinct species. The discrepancy in the

results might be due to the different selection of the tested strains, as well as due to the limited number of strains used, as well as differences in the experimental protocol. Also, some *Enterobacter* strains express β-lactamases at very low levels, which might have not been detectable by the methods used. Two strains isolated from fatal cases in a *Cronobacter* outbreak in France were shown to possess ESBL activity, which may have been acquired by horizontal transfer from other *Enterobacteriaceae*, since it was absent in the other strains of the same pulsetype included in the study (Caubilla-Barron *et al*. 2007).

1.1.7 Contamination of powdered infant formula

Due to the ubiquitous nature of *Cronobacter* and its remarkable resistance to desiccation and heat, contamination of desiccated products including powdered milk formula is a persistent problem. Effective control measures, as well as improved reconstitution practices must be in place to ensure safety of the products. By definition, powdered infant milk formula (PIF) is a relatively hostile environment for the survival of microorganisms. There are two main reasons why the survival and growth of microorganisms is suppressed in this environment: heat treatment of PIF and the low water activity of the product. Powdered infant formula is treated by pasteurization, which is designed to reduce the living microbial flora in the product. However, *Cronobacter* and other *Enterobacteriaceae* are being frequently isolated from such products, which can be due to the imperfection in the heat propagation during the pasteurization process, or due to post-processing contamination. For these reasons, the thermotolerance of *Cronobacter*, its desiccation resistance, as well as its ability to form biofilms and adhere to surfaces are of great concern.

1.1.7.1 Thermal resistance

Cronobacter was originally considered a very thermotolerant organism, and it was even believed that it could be able to survive the pasteurization process. In one of the earliest studies, Nazarowec-White and Farber (1999) reported that *Cronobacter* was the most thermotolerant organism among *Enterobacteriaceae*. Later studies by Breeuwer *et al.* (2003), Barron & Forsythe (2007) and Osaili & Forsythe (2009) demonstrated that *Cronobacter* was not particularly thermotolerant, but that the organism has the ability to adapt to heat stress following exposure to desiccation stress.

There appears to be considerable variability in thermotolerance of *Cronobacter*. Buchanan *et al.* compared thermal resistance of *Cronobacter* strains with other *Enterobacteriaceae* and the D-value measured for "*E. sakazakii* 607" was 10 times greater than that for "*E. sakazakii* 51329" (*C. muytjensii*) (Edelson-Mammel & Buchanan 2004). A subsequent comparison of the thermal resistance of 12 *Cronobacter* strains confirmed that the organism's thermal resistance was highly variable between different strains. The isolates exhibited an almost 20-fold divergence in thermal resistance, with the strain *"E. sakazakii* ATCC 51329" (*C. muytjensii)* being the least resistant, and a clinical isolate the most resistant. The new taxonomic scheme provides an explanation for such divergence in thermal resistance; what was previously considered as large intra-species variation became variation between different species of the *Cronobacter* genus.

Most earlier investigations used *Cronobacter* cultures of non-stressed cells grown in nutritionally rich reconstituted PIF under optimal growth conditions. As it is known that exposure to stress factors can increase the thermal resistance of bacteria (Pagan *et al.* 1997) and that the contamination of PIF often involves dry materials, it is beneficial to assess the heat resistance of cells that have been exposed to desiccation stress. Chang *et al.* (2010) investigated the influence of growth phase, growth temperature, heat shock, pH and water activity of treatment medium on the heat resistance of *Cronobacter*. The authors found that heat shock significantly enhanced the bacterium's resistance to all the lethal stresses assessed. For example, survival of the heat-shocked *Cronobacter* cells was approximately 752 times that of the

nonshocked cells, and after 36 h of exposure to high osmotic environment, the survival of the heat-shocked cells was approximately 119 times that of the control cells. The maximum increase in thermotolerance was reported after 15 min at 47 °C, with a 2.5-fold increase in D_{51} value compared to non heat-shocked cells. In accordance with other studies, it was found that the heat resistance was highly strain-dependent. The acidification of the culture medium resulted in decreased heat tolerance, whereas the decrease in water activity of the media resulted in increased heat tolerance. As *Cronobacter* cells showed a greater heat tolerance in real food products than in laboratory media, the authors concluded that the heat treatment designed based on laboratory media would offer insufficient reduction of the microbial load in foods and that these treatments should be validated in alimentary products.

It is also important to note that heat resistance of bacteria can be greatly influenced by the composition of the heating medium. For instance, the higher fat content in PIF can cause the increased thermal resistance of *Cronobacter* in reconstituted PIF compared to saline, due to the lower heat conductivity of fat (Kim & Park 2007b).

1.1.7.2 Desiccation tolerance

Many bacterial species are unable to survive in PIF due to the product's low water activity. *Cronobacter* has an unusual ability to stay alive when exposed to dry conditions, as it was shown to persist in the powder for more than 2.5 years (Barron & Forsythe 2007). In general, bacteria protect themselves from increasing osmolarity by the rapid intracellular accumulation of ions, mainly K+, followed by the accumulation of compatible solutes such as proline, glycine, betaine, and trehalose (Kempf & Bremer 1998). Breeuwer *et al*. (2003) demonstrated that in the stationary phase, *Cronobacter* strains are relatively resistant to osmotic and dry stress compared with other strains of the *Enterobacteriaceae* group such as *Salmonella* and *E. coli*. In dried stationary cells the trehalose concentration increased more than five-fold and such accumulation was not observed in dried stationary *E. coli*. Furthemore, addition of trehalose to the growth medium of exponential phase *Cronobacter* improved its

21

survival after drying (Breeuwer *et al*. 2003). Barron and Forsythe (2007) prepared desiccated cultures of 10 *Cronobacter* strains and 17 strains of other *Enterobacteriaceae* to study their persistence in powdered infant formula for a period of 2.5 years. It was found that the organisms could be divided into four groups according to their persistence in dessicated state, with *C. sakazakii* strains surviving the longest. Five of the 10 *Cronobacter* strains were still recoverable after two years. Two of the five persistent strains were capsulated, and the only *Cronobacter* strains that could be recovered after 2.5 years were capsulated. The authors concluded that capsule production can have a role in persistence in dry environments over long periods of time. This was further supported by the fact that the capsulated strains of *K. oxytoca, E. vulneris,* and *Pantoea* spp. persisted in the powdered formula over two years, whereas none of the *Cit. koseri* or *E. cloacae* non-capsulated strains survived. Gurtler & Beuchat (2007b) monitored the survival of *Cronobacter* spp. in infant formula under different storage conditions for up to 12 months. In both products the *Cronobacter* spp. survived better at low water activities and a low storage temperature (4°C). Interestingly, the organism persisted less in infant formulas with higher water activities (0.43–0.50) compared with lower water activities (0.25–0.30). The persistence of the bacterium in infant formulas decreased with increasing storage temperature (4, 21 and 30 °C) and the survival rate was not affected by the composition of the formula (i.e. milk-based or soy-based). In an extensive study by several collaborating laboratories, Beuchat *et al*. (2009) examined various factors affecting growth and survival of *Cronobacter* in infant formula and other foods. The organism survived better in dried formula and cereal at low a_w (0.25–0.30) than at high a_w (0.69–0.82) and at 4 °C compared to 30 °C. The composition of infant formulas did not significantly affect the rate of growth. Riedel and Lehner (2007) examined the changes in protein synthesis following desiccation and osmotic stress in "*C. sakazakii* z236" and found an array of genes that were up-regulated during the response to desiccation, including cold-shock protein CspC, DNA repair proteins, superoxide dismutase and outer membrane proteins OmpA and OmpC, possibly linked to transport of compatible solutes similar to trehalose.

Interestingly, powdered infant formula and other milk based products may contain substances that have a protective effect on bacteria exposed to desiccation during the manufacturing of such products. For example, neither nisin or lactoferrin, natural substances known for their antimicrobial effect in food products, had detectable bactericidal activity against *Cronobacter* spp. in reconstituted PIF, whereas the same *Cronobacter* strains were sensitive to both nisin and lactoferrin in peptone water (Al-Nabulsi *et al*. 2009). The authors concluded that this might be due to the presence of divalent cations such as Ca^{2+} and Mg^{2+}, which may protect the bacteria by causing changes in the spatial structure of lactoferrin and by stabilising the bacterial membrane, or by high concentrations of iron in PIF, which can reverse the bacteriostatic effect of lactoferrin. Dancer *et al*. (2009) examined the resistance of *Cronobacter* to different stress factors including heat, drying and decreased water activity. In accordance with previous studies, the authors demonstrated that *Cronobacter* was much more resistant to environmental stresses compared to other members of *Enterobacteriaceae*. The most heat-resistant strains were also the most resistant to drying. The authors concluded that the potential cross-resistance of *Cronobacter* strains can increase the likelihood of its persistence in powdered infant formula.

The increased resistance to osmotic stress of *Cronobacter* strains compared to other *Enterobacteriaceae* points towards inherent bacterial mechanisms that help the organisms to resist dry stress. Given the possible plant origin of *Cronobacter* spp. it is plausible that these mechanisms originally evolved as survival strategies specific to plant-associated organisms.

1.1.7.3 Biofilm formation and capsule production

Biofilms have been defined as sessile communities of bacterial cells attached to a surface or to each other, usually embedded in polymeric substances produced by the bacteria (Marshall 1992). Biofilm formation is known to enhance the resistance of cells to environmental stresses and to provide protection against sanitizers (Marshall

1992). *Cronobacter* organisms often produce a heteropolysaccharide capsule composed of glucuronic acid, D-glucose, D-galactose, D-fucose and D-mannose; strains from NICU outbreaks have been reported to produce so much capsular material that on milk agar plates the colonies will drip onto the lid of inverted Petri dishes (Caubilla-Barron *et al.* 2007). Interestingly, this material has been patented for use as a thickening agent in foods (Harris & Oriel 1989, Scheepe-Leberkuhne & Wagner 1986). This capsular material could facilitate the organism's attachment to plant surfaces and, combined with a tolerance to desiccation, enable the organism to colonize plant material and survive harsh environmental conditions. Production of heteropolysaccharide capsule is also known to promote survival of the bacterium in foodstuffs for long periods of time (Iversen & Forsythe 2003, Lehner *et al.* 2005). Iversen and Forsythe (2003) suggested that the capsule produced by *Cronobacter* might also increase the organism's ability to attach to surfaces and to form biofilms. In a later study by Hurrell *et al.* (2009a) it was shown that biofilm formation did not correlate with capsule production, however.

Kim *et al.* (2007a) examined survival of *Cronobacter* dried on the surface of stainless steel and exposed to 43% relative humidity with varying temperature and medium, as well as survival of *Cronobacter* cells in biofilm. The authors reported that some of the cells in biofilms survived under all test relative humidities for up to 42 days. Their findings indicate that infant formula provides protection against desiccation for cells attached to a steel surface, as well as cells in biofilms. Kim *et al.* (2006) reported that *Cronobacter* did not form biofilms on stainless steel or tubes at 25 °C in sterile tryptic soy broth or lettuce juice. However, the pathogen did form biofilms on surfaces of both materials immersed in infant formula. This indicates that the ability of *Cronobacter* to form biofilms is affected by the composition of the media, and that it is enhanced by substances present in infant formula. The infant formula composition can also increase *Cronobacter*'s resistance to disinfectants, as shown by Kim *et al.* (2007a), who examined the effect of thirteen disinfectants commonly used in infant formula preparation areas. Populations of *Cronobacter* cells suspended in water (ca 7 log CFU/ml) decreased to undetectable levels (< 0.3 log CFU/ml) within 1−5 min of treatment with disinfectants, whereas numbers of cells in reconstituted PIF diminished

by only 0.02–3.69 log CFU/ml after treatment for 10 min. Furthermore, cells attached to stainless steel were less resistant to disinfectants. *Cronobacter* is able to adhere to silicon, latex and polycarbonate and to a lesser extent to stainless steel (Iversen & Forsythe 2003). Furthermore, *Cronobacter* has been reported to attach and form biofilm on glass and polyvinyl chloride (Lehner *et al*. 2005). All of these materials are commonly used for infant-feeding and food preparation equipment, and, if contaminated, may increase the risk of infection.

1.1.8 Safety of breast milk substitutes

Although the level of PIF contamination does not usually exceed 1 CFU per 100g, in 2007 WHO concluded that it was not possible to assess the contribution of the hygiene abuse and extended hold of reconstituted formula at non-refrigerated temperatures to infections by *Cronobacter* and that it must be assumed that even low contamination below 3 cfu/100g can lead to disease (WHO 2007). During the manufacture of PIF, *Cronobacter* can gain access from the environment, from the ingredients added or as a post processing contamination. In one of the methods used for producing powdered milk substitutes, individual ingredients are heat-treated before spray-drying and blending of the product take place. This method is prone to bacterial contamination as the drying and the blending are often the principle contamination sites for dried products (Iversen & Forsythe 2003). Furthermore, due to the survival of the cells at elevated temperatures and the capacity to grow up to 47°C, *Cronobacter* has a competitive advantage compared with other members of *Enterobacteriaceae*. The bacterium then becomes more dominant in the warm and dry environment of the PIF factory and its presence further increases the risk of post-processing contamination.

 In 2004, the European Food Safety Authority (EFSA) reviewed the known facts about *Cronobacter* and categorized it together with *Salmonella* as an organism with a

25

clear evidence of causality and concluded that there was a need to reduce its prevalence in PIF (EFSA 2004). In 2007, it was suggested that monitoring *Enterobacteriaceae* in the processing environment and in the product would be useful in reducing the prevalence of *Cronobacter* (EFSA 2007). However, *Enterobacteriaceae* direct count does not always detect *Cronobacter* presence in the PIF samples, possibly due to the low cell number and injury to the organism during dessication (Forsythe 2005). Therefore, specific test methods for the recovery of dessication-stressed *Cronobacter* cells have been developed. Since 2008, the Codex Alimentarius Commission (Codex Alimentarius Commission (CAC) 2008) microbiological criteria are applied to PIF products intended for under 6 month of age. The Codex Alimentarius Commission requirement is to test thirty 10g quantities for presence/absence of *Cronobacter*. The current recommended ISO standard method (ISO 21528-1:2004) for detection of *Cronobacter* includes an enrichment step in EE ("*Enterobacteriaceae* Enrichment") broth, but according to an earlier study (Iversen & Forsythe 2007), growth of some strains of *Cronobacter* is inhibited in this broth, which can lead to false negative results. Later Joosten *et al.* (2008) observed inhibition of growth in EE for 9 (9.3%) strains, including seven *C. sakazakii* and one *C. malonaticus* strain. The authors evaluated an alternative method which omits the EE enrichment step and proposed the revision of the standard ISO method. Previous FAO/WHO expert meetings on the microbiological risk assessment of powdered infant formula have recommended the reconstitution of PIF with water no less than 70°C, and use of the reconstituted product within 2 hours (FAO/WHO 2004, FAO/WHO 2006). The use of high reconstitution temperature should reduce the viability of *Cronobacter* spp. and *Salmonella* serovars; short storage period would limit the growth of any surviving microorganism, and therefore this procedure would reduce the risk of neonatal infection through the consumption of contaminated formula.

Recently, another product was highlighted as a potential source of *Cronobacter* infection – follow-up formula (FUF). To meet the increased nutritional needs of older infants, FUF often contains a variety of dry blended ingredients which have been linked to *Cronobacter* contamination. Although FUF is designed for infants above 6

26

months, international surveys reveal that it is given to infants less than 6 months or even less that 1 month old (FAO/WHO 2008). Given the fact that there is no hygiene control requirement for testing FUF for *Cronobacter*, FUF can be potentially hazardous if consumed by susceptible neonates. The FAO/WHO meeting held in 2008 reviewed this matter in order to establish microbiological critera for FUF, but concluded that no microbiological criteria were required (FAO/WHO 2008).

1.1.9 Control methods

Several research groups have proposed the use of different substances to limit growth of *Cronobacter* in PIF. Nair *et al*. (2004) showed that monocaprylin can be effective at inactivation of *Cronobacter* in PIF stored at refrigeration temperature and Gurtler & Beuchat (2007c) demonstrated the inhibitory effect of the lactoperoxidase system on *Cronobacter* stored at 21, 30 and 37°C. Lee & Jin (2008) investigated the effect of natural antimicrobial compounds on inactivation of *Cronobacter* and concluded that the combination of diacetyl with nisin had the greatest synergistic effect. Wakabayashi *et al*. (2008) tested the susceptibility of *Cronobacter* to several metal-bound forms of bovine lactoferrin and lactoferrin-derived peptide lactoferricin B. The authors reported that four *Cronobacter* strains were inhibited by hydrolysed lactoferrin, Cu-lactoferrin, apo-lactoferrin and lactoferricin B, but not to Fe-lactoferrin. The authors concluded that bovine lactoferrin-related compounds may be used for the inhibition of *Cronobacter* in foods. However, as the organism was not inhibited by Fe-lactoferin, this strategy would be limited to foods that do not contain iron and thus unsuitable for use in PIF.

In an interesting study, Zhao *et al*. (2008) isolated five bacteriophages of *Cronobacter* from sewage, one of which could lyse 24 (89%) of *Cronobacter* strains tested. The authors concluded that these bacteriophages could be potentially used in typing or treatment of *Cronobacter* infection. This idea has not been further explored, however. Amalaradjou *et al*. (2009) reported that trans-cinnamaldehyde (0.5%), which is one of the major components of bark extract of cinnamon, reduced *Cronobacter* count to

undetectable levels by 4 h of incubation at 37 or 23 ^{o}C and 10 h of incubation at 8 or 4°C, respectively. The authors proposed that trans-cinnamaldehyde could potentially be used to inactivate *Cronobacter* in reconstituted PIF. However, this compound has general antimicrobial activity and could therefore upset the neonatal gut flora leading to serious infections. Al-Holy *et al*. (2010) reported a more than 5-log reduction in *Cronobacter* count after 2 h and a complete elimination by the end of the storage period when lactic acid was applied together with copper (II). The authors hypothesized that lactic acid in the dissociated form may chelate copper (II) ions, thus allowing them to penetrate through the cytoplasmic membrane and exert toxicity against *Cronobacter.* Lactic acid is a weak organic acid which has been extensively used as antimicrobial agent in foods. Its undissociated form can penetrate the cytoplasmic membrane, which reduces the intracellular pH and disrupts the transmembrane proton motive force and it can also permeabilize the outer membrane of Gram-negative bacteria for other antimicrobial agents (Alakomi *et al*. 2000).

1.2. General aims

Cronobacter has been known as a separate species since 1980, and has been the focus of an intensive scientific research only since 2003. Between 2003 and 2007, our knowledge of the organism then known as *Enterobacter sakazakii*, significantly advanced, especially in the fields of physiology and detection methods. Reliable identification methods were developed and control measures were put in place to increase the safety of product designed for neonates. Pioneering works examining the organism's remarkable persistence in the environment, as well as various aspects of the bacterium virulence both *in vitro* and *in vivo,* were published. The knowledge about the genetic make-up of this fascinating organism was still very limited, however, as there was no genome of '*E. sakazakii'* sequenced at that time, and molecular biology research was primarily aimed at developing improved detection methods and phylogenetic classification of '*E. sakazakii'* strains. The re-classification of *"E.*

sakazakii" into five new *Cronobacter* species in 2008 raised queries on the previous identification methods for the organisms and the published works in general, as the experimental data were obtained for uncertain *Cronobacter* species. Previous studies on '*E. sakazakii*' will be difficult to interpret unless the strains are re-examined and re-classified according to the current taxonomic structure. There is a strong need to obtain robust data reflecting the taxonomic diversity of *Cronobacter* that would provide insight into the species behaviour in products intended for neonatal consumption, as well as its mechanisms of pathogenicity, which are largely unknown to date.

In response to the FAO-WHO (2006) recommendations, the Food Standards Agency UK (FSA) funded our group at NTU to investigate the survival and growth of *Cronobacter* and other *Enterobacteriaceae* in infant formula during reconstitution and storage, in order to update the current risk assessment for *Cronobacter* in PIF (http://www.who.int/foodsafety/micro/jemra/assessment/esakazakii/en/index.html). This required the determination of death rates at temperatures >50°C and growth rates. I undertook the measurements of the growth rates and lag times, whilst Dr. J. Caubilla-Barron measured the death rates. My research focussed on *Cronobacter* growth in infant formula and breast milk with respect to iron availability (external supervisor Dr. J. Hobman). However, during my second year of research, I was offered the opportunity to construct and analyse a *Cronobacter* microarray based on the *C. sakazakii* BAA-894 genome sequenced by Prof. M. McClelland. Consequently, my thesis does not necessarily reflect a progression in research of one topic, but instead gives an initial study in the context of control of *Cronobacter* infection in neonates through potentially contaminated PIF, and subsequently, a detailed exploration of the organism's genome and comparative analysis with other *Cronobacter* strains and species. This forms a highly novel and unique investigation of *Cronobacter* which covers the topics related to infant formula safety, as well as exploration of the detailed genomic architecture of this remarkable pathogen. In conclusion, the diversity of the *Cronobacter* group, highlighted by the recently adopted taxonomy scheme, will be addressed with respect to growth behaviour in powdered infant products and via microarray whole genome analysis, as well as *in*

silico genome comparisons both within *Cronobacter*, and within the *Enterobacteriaceae* family.

I have been appropriately recognised for my contribution to additional studies in the research team: i) *in vitro* biofilm formation in enteral feeding tubes (Hurrell *et al*. 2009a); ii) showing that enteral feeding tubes from neonatal intensive care units may act as loci for bacterial colonisation (Hurrell *et al*. 2009b). These have not been included in this thesis as they were not the main focus of my research.

1.3. General materials and methods

1.3.1 Media and reagents

Ampicillin	Sigma-Aldrich, Dorset, UK
Arabinose(L-)	Sigma-Aldrich, Dorset, UK
Chloramphenicol	Sigma-Aldrich, Dorset, UK
Glycerol	Fisher Scientific, Loughborough, UK
Glucose	Sigma-Aldrich, Dorset, UK
Isopropyl Alcohol	Sigma-Aldrich, Dorset, UK
LB Agar	Fisher Scientific, Loughborough, UK
LB broth	Fisher Scientific, Loughborough, UK
NaOH	Sigma-Aldrich, Dorset, UK
Na_2EDTA	Sigma-Aldrich, Dorset, UK
Nutrient Broth	Fisher Scientific, Loughborough, UK

Phenol/Chloroform/IsoamylAlcohol	Sigma-Aldrich, Dorset, UK
SDS	Sigma-Aldrich, Dorset, UK
Sucrose	Sigma-Aldrich, Dorset, UK
Tryptone	Oxoid, Hamphshire, UK
Tris base	Sigma-Aldrich, Dorset, UK
TSA	Oxoid, Hamphshire, UK
TSB	Oxoid, Hamphshire, UK
Yeast extract	Oxoid, Hamphshire, UK
VRBGA	Oxoid, Hamphshire, UK

1.3.2 Molecular biology kits and reagents, biochemical kits

DNA ladder (1 kb)	Promega, Southampton, UK
EcoRV	Sigma-Aldrich, Dorset, UK
Gel loading dye	Promega, Southampton, UK
Genomic DNA Buffer Set	Qiagen, Crawley, UK
GoTaq Hot Start Polymerase	Promega, Southampton, UK
ID32 E	Bio Merieux, France
One Shot Top10 chem. comp. *E. coli*	Invitrogen, Paisley, UK
Phusion Polymerase	Finnzymes, Hitchin, UK
Sigma GenElute	Sigma-Aldrich, Dorset, UK
SYBR Safe	Invitrogen, Paisley, UK
Qiagen Genomic tip – 100/G	Qiagen, Crawley, UK
Qiagen Spin Miniprep	Qiagen, Crawley, UK
Qiaquick PCR purification kit	Qiagen, Crawley, UK

TAE Buffer	Geneflow, Fradley, UK
T4 DNA ligase	Promega, Southampton, UK

1.3.3 Preparation of media

Glycerol Yeast Tryptone (GYT) Medium

To make 100 ml of GYT, 10 ml of glycerol was mixed with 0.125 grams of yeast extract and 0.25 grams of tryptone. The media was autoclaved (121°C, 15 min) and stored at room temperature until required.

Luria-Bertani Agar (LB agar)

Twenty grams of LB agar base were resuspended in 500 ml of distilled water and autoclaved (121°C, 15 min). The agar was allowed to cool to 55°C and approximately 20 ml was poured into sterile Petri dishes. The plates were then stored at room temperature until required, for a maximum of 1 week.

Luria-Bertani Broth (LB broth)

Twelve point five grams of Luria-Bertani base were resuspended in 500 ml of distilled water. The mixture was autoclaved (121°C, 15 min) and stored at room temperature until required.

Trypticase Soy Agar (TSA)

Twenty grams of trypticase soy agar base were resuspended in 500 ml of distilled water and autoclaved (121°C, 15 min). The agar was allowed to cool to 55°C and

approximately 20 ml was poured into sterile Petri dishes. The plates were then stored at room temperature until required, for a maximum of 1 week.

Trypticase Soy Broth (TSB)

Fifteen point zero grams of TSB base were resuspended in 500 ml of distilled water. The mixture was autoclaved (121°C, 15 min) and stored at room temperature until required.

Violet Red Bile Glucose Agar (VRBGA)

Nineteen point three grams of VRBGA base were resuspended in 500 ml of distilled water. The mixture was autoclaved (121°C, 15 min) and stored at room temperature until required.

1.3.4 Preparation of media with antibiotics

Addition of Ampicillin to LB Agar and LB Broth

A 10 mg/ml stock solution of ampicillin in sterile distilled water was filter sterilized using a 0.2 μm filter. The sterile antibiotic solution was added to LB agar or LB broth to a final concentration of 200 μg/ml. The plates were stored at room temperature for a maximum of 1 week.

Addition of Chloramphenicol to LB Agar and LB Broth

A 2 mg/ml stock solution of chloramphenicol in sterile distilled water was filter sterilized using a 0.2 μm filter. The sterile antibiotic solution was added to LB agar or LB broth to a final concentration of 35 μg/ml. The plates were stored at room temperature for a maximum of 1 week.

Addition of Kanamycin to LB Agar and LB Broth

A 2.5 mg/ml stock solution of kanamycin in sterile distilled water was filter sterilized using a 0.2 µm filter. The sterile antibiotic solution was added to LB agar or LB broth to a final concentration of 50µg/ml. The plates were stored at room temperature for a maximum of 1 week.

2. FOOD STANDARDS AGENCY PROJECT: BACTERIOCIDAL PREPARATION OF POWDERED INFANT FORMULAIntroduction

2.1.1 Aims

The ability of *Cronobacter* to attach to surfaces, to form biofilms and its extreme resistance to dry stress contribute to the risk of powdered infant formula contamination by the bacterium. Moreover, the composition of PIF seems to have a strong protective effect on the survival of *Cronobacter*. Finally, high levels of heat-stable lipopolysaccharide (endotoxin) present in infant formula may enhance the translocation of *Cronobacter* across the gut and blood–brain barrier and increase the risk of bacteraemia in neonates (Townsend *et al*. 2007b). This highlights the importance of safety testing of PIF, as well providing adequate guidance for safe reconstitution practices.

Breast milk substitutes including powdered infant formula (PIF) and milk fortifiers are not manufactured as sterile products and have to conform to the microbiological specifications recently revised by Codex Committee for Food Hygiene (Codex Alimentarius Commission (CAC) 2008), which demand presence/absence testing of thirty 10 g quantities of the product. Limiting the intrinsic contamination of PIF products by statutory microbiological criteria is one way to reduce health risks associated with PIF ingestion. However, at the very low levels of contamination, *Cronobacter* may still be present in the product after the microbiological testing and bacterial multiplication following reconstitution can lead to infection. In 2007, the WHO published the 'Safe preparation, storage and handling of powdered infant formula Guidelines' (WHO 2007) and recommended reconstituting of PIF with water at temperatures greater than 70°C, which should provide strong bacteriocidal effect. This recommendation was supported by several studies on thermal resistance of *Cronobacter* where it was shown that using water at $70^{o}C$ for reconstitution of PIF

had caused reductions in *Cronobacter* populations greater than 4 log (Edelson-Mammel & Buchanan 2004, Shaker *et al*. 2008). If this procedure is widely implemented, it could considerably reduce *Cronobacter* in PIF. From the combined results of our survey published in the FSA report (Caubilla-Barron *et al*. 2009) undertaken by other members of our team, it is apparent that there are various PIF reconstitution practices in place and there is a need for a robust risk model that would provide scientific basis for the new formula reconstitution guidelines.

The current risk assessment for *Cronobacter* in PIF (FAO/WHO 2006) (http://www.who.int/foodsafety/micro/jemra/assessment/esakazakii/en/index.html) is based on limited studies with a single *Cronobacter sakazakii* strain and needs to be updated. As it is reasonable to expect variation in growth behaviour both between *Cronobacter* species and between strains of the same species, there is a need to obtain growth data for diverse *Cronobacter* species as defined in the taxonomic reclassification by Iversen *et al*. (2008). It is also useful to obtain growth data for other potentially pathogenic bacteria that may occur in PIF products. This chapter examines the growth behaviour of *Cronobacter* spp., *Salmonella*, other *Enterobacteriaceae* and *Acinetobacter* sp. at a temperature range $21 - 44^{o}$C. Along with the results from a study on heat inactivation of *Cronobacter* and other *Enterobacteriaceae* undertaken by Dr. Juncal Caubilla-Baron as a part of the same study, the results were used to create a new risk assessment for *Cronobacter* in PIF and provided to the Food Standards Agency UK.

The organisms used in our study were defined by (FAO/WHO 2004, FAO/WHO 2006) as 'category A' organisms with clear evidence of causality between infant infection and infant formula contamination (*Salmonella* serovars and *Cronobacter* spp) or 'category B' organisms for which the causality was plausible, but not yet demonstrated (*Enterobacter cloacae*, *Citrobacter koseri*, *Citrobacter freundii*, *Klebsiella oxytoca*, *Klebsiella pneumoniae*, *Pantoea agglomerans*, *Escherichia vulneris*, *Escherichia coli*, *Hafnia alvei*, *Serratia marcescens*, and *Acinetobacter baumannii*). Moreover, all the organisms chosen for the study are frequently being isolated from breast milk substitutes (unpublished NTU data). Most strains used in this study were isolates from PIF or milk products and some clinical or environmental

36

strains have been used where necessary. *Enterobacter hormaechei* and *Enterobacter cancerogenus* were included in the study although there is no evidence that they have caused infections via contaminated PIF. These organisms are frequently being mis-classified as *E. cloacae* and so would not be reported as such. However, *E. hormaechei* strains have been isolated from enteral feeding products and powdered infant formula (Townsend *et al*. 2008b) and have been shown to be of clinical significance due to the report of several outbreaks of sepsis in neonatal intensive care units in Brazil and the United States (Wenger *et al*. 1990, Campos *et al*. 2007). Townsend *et al*. (2008b) also reported a neonatal outbreak of *E. hormaechei* that was originally misclassified as *Cronobacter* spp.

The temperatures 21, 27, 37, 41 and 44°C were chosen because they reflect a wide spectrum of PIF reconstitution, cooling and storage practices, as shown in Figure 2. The upper temperature range is useful for modelling growth of bacteria that have survived reconstitution at raised temperatures, and also for formulae for which lower rehydration temperatures have been recommended on the packaging. The lower temperatures are relevant for modelling bacterial growth in prepared feeds that are stored at room temperature prior to the feeding. The infant formulae tested were two most-commonly used types of breast milk substitutes, casein- and whey-dominant infant formulae.
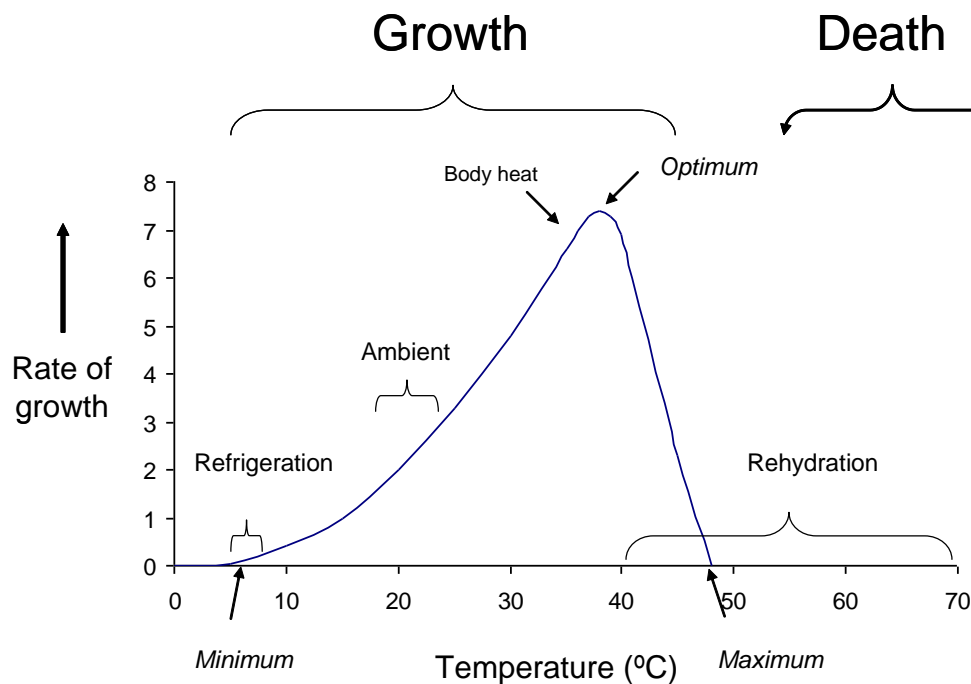
Figure 2. Example growth profile of *Cronobacter* in infant formula over a range of temperatures reflecting refrigeration, reconstitution and storage practices.

2.1.2 Growth rates of *Cronobacter* and other *Enterobacteriaceae* in powdered infant formula

The majority of existing studies examining growth of *Cronobacter* in PIF were published before the taxonomic reclassification by Iversen *et al*. (2008) and were therefore determined for uncertain *Cronobacter* species. Most of the early studies examined growth of *Cronobacter* in media other than PIF, and later studies focussed predominantly on growth behaviour of *Cronobacter* cells subjected to various forms of stress (Edelson-Mammel & Buchanan 2004, Kim & Park 2007b, Lin & Beuchat 2007, Gurtler & Beuchat 2007b, Beuchat *et al*. 2009, Chang *et al*. 2010). There is a general lack of data describing growth of other potentially pathogenic

Enterobacteriaceae in PIF, and to our knowledge, no comparison of the growth rates of different members of *Enterobacteriaceae* in reconstituted PIF has been published.

Iversen *et al*. (2004a) reported a mean doubling time for *Cronobacter* in PIF of 13.7 h at 6 ºC, 1.7 h at 21ºC and 19-21 min at 37ºC. In another similar study, the mean doubling times were determined as 13.7 h at 6°C, 1.3 h at 21°C and 22.0 min at 37°C (NTU, unpublished data). In a study by Gurtler & Beuchat (2007a) different types of reconstituted PIF inoculated with *Cronobacter* spp. were stored at 4, 12, 21, and 30°C and populations were monitored for up to 72 h. *Cronobacter* spp. did not grow in formulae stored at 4°C and growth was not particularly influenced by the composition of PIF. The most extensive study on growth behaviour of non-stressed *Cronobacter* cells in PIF was published by Lenati *et al*. (2008). The authors examined the growth of 9 strains of *Cronobacter* in breast milk, human breast milk with fortifiers and powdered infant formula at 10, 23 and 37 °C. The *Cronobacter* strains studied included clinical, environmental and food isolates, but as this study was published prior to the taxonomic reclassification by Iversen *et al*. (2008), they were not assigned to separate *Cronobacter* species. Lenati and colleagues reported the following generation times of *Cronobacter* in PIF: 29.5 h for clinical isolates and 18.3 h for environmental isolates at 10°C, 42.6 min for isolates from all sources at 23°C; and 17.4 min for all isolates at 37°C. For all isolate sources, the average generation times were shorter for cells grown in PIF than in breast milk and breast milk with fortifiers. The reported lag times were: 126.6 h for clinical and 59.0 h for environmental isolates grown at 10°C, from 3.9 to 4.7 h at 23°C and from 2.2 to 3.0 h at 37°C. At all temperatures, growth of *Cronobacter* in PIF produced higher population densities than growth in breast milk or breast milk with fortifiers. Due to the growth of *Cronobacter* in PIF at 10°C the authors suggested that powdered infant formula should not be stored at room temperature for longer than 4 h and that proper reconstitution practices, as well as temperature control are needed to control *Cronobacter*. Kandhai *et al*. (2006) used a clinical, environmental, and food isolate, as well as the ATCC 29544[T] type strain to investigate the lag time, specific growth rate, and maximum population density of *Cronobacter* strains grown in PIF. The authors reported a specific growth rate of 2.3 h and lag times lag ranging from 83.3 h at 10 °C

to 1.73 h at 37 °C. The growth phase of the cells had no effect on the specific growth rates or lag times (Kandhai *et al*. 2006).

2.2. Materials & Methods

2.2.1 Bacterial strains and their maintenance

All strains used in this study were stored at -80°C in Nutrient Broth with 10% (v/v) glycerol, subcultured on Trypticase Soy Agar and checked for purity. The list of strains is provided in Table 1.

Table 1. Bacterial strains used in the FSA study

Organism	Strain	Source	Additional comment
Acinetobacter genomosp. 13	415	Cheese and chive dip, UK	
Citrobacter freundii	153	Infant formula milk, Korea	
Citrobacter koseri	598	Infant formula milk	
Cronobacter dublinensis	582	NCTC 9844	
Cronobacter malonaticus	1212	LMG 23826	Species type strain
Cronobacter malonaticus	685	CDC	
Cronobacter muytjensii	3	ATCC 51329	Species type strain
Cronobacter sakazakii	1	ATCC 29544	Species type strain
Cronobacter sakazakii	2	ATCC 12868	
Cronobacter sakazakii	114	Unknown	
Cronobacter sakazakii	658	ATCC BAA-894	Genome sequenced strain
Cronobacter sakazakii	767	Fatal meningitic strain	Isolated from trachaea
Cronobacter turicensis	57	Milk powder, UK	
Enterobacter cloacae	49	ATCC 13047	Isolated from spinal fluid
Enterobacter cancerogenus	806	Neonatal enteral feeding tube	
Enterobacter hormaechei	790	Neonatal enteral feeding tube	
Escherichia coli	605	Unknown	
Escherichia coli	780	Neonatal enteral feeding tube	
Escherichia coli	796	Neonatal enteral feeding tube	
Escherichia vulneris	52	Milk powder, UK	
Hafnia alvei	236	Cow udder	
Klebsiella oxytoca	379	ATCC 43155	
Klebsiella pneumoniae	273	Spice	
Pantoea ananatis	44	Infant milk formula, Korea	
Salmonella Enteritidis	358	NCTC 3046	
Salmonella Senftenberg	1188	NCTC 9959	
Salmonella Anatum	1197	NCTC 5779	
Serratia marcescens	464	Unknown	

2.2.2 Introduction to impediometry

Impediometry is based on the observation that microbial metabolism changes the conductivity of culture media. Impedance can be defined as the resistance to a flow of an alternating current as it passes through a conducting medium, and is a function of conductance, capacitance and the applied frequency (Eden & Eden 1984). In microbiological impediometry, the impedance change is measured using a pair of electrodes submerged in the test sample with inoculated bacteria. Upon bacterial growth, an increase in conductance and capacitance and a decrease in impedance are observed (Silley & Forsythe 1996). These changes are caused by the conversion of neutral or weakly charged molecules of substrates to charged molecules of metabolism end-products (for example glucose to lactic acid), and the increase of particle mobility by cleavage of larger molecules to smaller molecules (for example conversion of lactate ion to a smaller more mobile bicarbonate ion) (Silley & Forsythe 1996). At some stage during microbial growth, the amount of ions produced will reach the point when an increase in impedance can be detected. This time is called the "time to detection" or TTD and occurs in the early stages of the exponential growth phase. TTD is a function of the size of the initial population, the growth kinetics of the test organism and properties of the test medium (Eden & Eden 1984). The generation time t_d of a particular organism in a defined medium is obtained by plotting the scattergram of logarithm (base2) of the bacterial concentration of the samples against the corresponding TTDs; applying a linear regression on the scattergram and using the formula $t_d = \log 2 \, / \, B$, where B is the slope of the linear regression curve. For more details on the calculation see Section 2.2.3.

The measurement can be performed in two ways, using either direct or indirect technique. In direct technique, the bacteria are in direct contact with the electrodes and the changes in the conductance of the growth medium directly reflect the changes in the test medium. In indirect technique, the growth of the microorganism is detected

by the release of carbon dioxide, which is absorbed by KOH in the test tube causing a decrease in conductivity. The indirect technique is especially useful if bacteria do not metabolise the nutrients in the medium by fermentation and therefore do not produce enough ions that could be detected by the direct impedance method.

2.2.3 Calculation of doubling time (t_d) – adapted from (Eden & Eden 1984)

This section provides a justification for the use of the fomula $t_d = \log 2 / B$ for calculation of the doubling time t_d.

If c_0 is the initial bacterial concentration, TTD is time to detection, and t_L is the duration of lag phase, it can be shown that the relationship between $\log c_0$ and $(TTD - t_L)$ is linear within a range of bacterial concentrations (as discussed below).

If c_{B0} is the bacterial concentration at the beginning of the exponential phase, after n generations the bacterial concentration will be:

$$c_B = c_{B0} \cdot 2^n \tag{1}$$

Since the number of generations is $n = t / t_d$, and

$2^n = e^{n \ln 2}$, we obtain:

$$c_B = c_{B0} e^{t \ln 2 / t_d} \tag{2}$$

As bacterial division is not occurring during the lag phase, the instantaneous bacterial concentration c_B during the exponential phase becomes:

$$c_B(t) = c_{B0}e^{(t-t_L)\ln 2/t_d} \qquad (3)$$

The equation (3) can be translated into the total concentration of ions generated by the microbial metabolism in the growth medium. If each microorganism generates K_B ions per second, the total concentration of ions at any time is:

$$c_S(t) = K_B \int_0^t c_B(t)\,dt \qquad (4)$$

Where c_S is the concentration of ions generated by microbial metabolism and K_B is a 'bacterial activity' coefficient in min^{-1}.

Although during the lag phase bacteria do not multiply, they are metabolizing and increasing the ionic concentration. From the equation (4), the ionic concentration c_S at the end of the exponential phase will be:

$$c_S(t) = K_B \int_0^{t_L} c_{B0}\,dt = K_B\,c_{B0}\,t \qquad (5)$$

During the exponential phase the total ionic concentration is the sum of the ionic concentration at the end of the lag phase (5) and the concentration of ions generated during the bacterial multiplication (3). Using equations (3) and (5) we obtain:

$$c_S(t) = K_B \int_0^{t_L} c_{B0}\,dt + K_B \int_{t_L}^{t} c_{B0}e^{(t-t_L)\ln 2/t_d}\,dt = K_B\,c_{B0}(t_L - \frac{t_d}{\ln 2}) + \frac{K_B c_{B0} t_d}{\ln 2}\,e^{(t-t_L)\ln 2/t_d} \quad (6)$$

If $t \gg t_L$, the first term in equation (5) becomes negligible. At some point in time, the concentration of ions generated by the bacteria c_S reaches the magnitude of the initial

ionic concentration c_{Si} and measurable increase in conductivity can be detected. This time is denoted as TTD.

$$c_{Si} = c_s(TTD) = \frac{K_B c_{B0} t_d}{\ln 2} \, e^{(TTD - t_{L)}) \ln 2 / t_d} \tag{7}$$

The expected relationship can be derived from equation (7):

$$\ln c_{B0} = \ln \frac{c_{Si} \ln 2}{K_B t_d} - \frac{\ln 2}{t_d} \left(TTD - t_L \right) \tag{8}$$

For convenence, this equation can be converted using decadic logarithm into:

$$\log c_{B0} = \log \frac{c_{Si} \ln 2}{K_B t_d} - \frac{\log 2}{t_d} \left(TTD - t_L \right) \tag{9}$$

Now we can see that the relationship between decadic logarithm of initial bacterial concentration and time to detection TTD (9) describes a straight line with a negative slope B and intercept A, where

$$B = \frac{\log 2}{t_d} \qquad \text{and} \qquad A = \log \frac{c_{Si} \ln 2}{K_B t_d} \tag{10}$$

Both constants depend on the particular properties of the organism and medium used. The simple equation $t_d = \log 2 / B$ derived from (10) is used for the determination of doubling time t_d from the slope B of the regression curve of the scattergram of the decadic logarithm of initial bacterial concentration and time to detection TTD as shown in Figure 3.
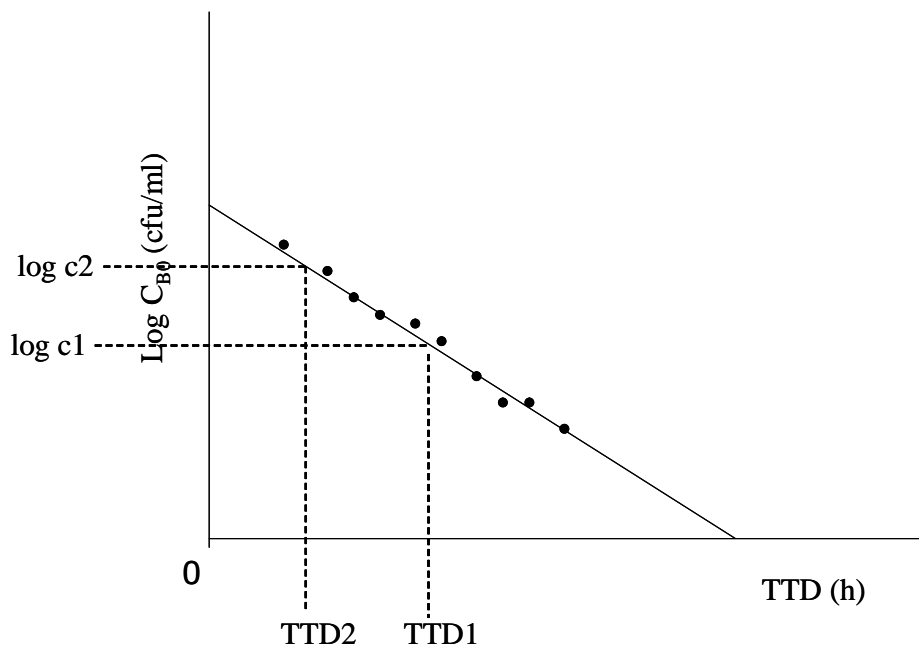
Figure 3.Relationship between the decadic logarithm of initial bacterial concentration c_0 and time to detection TTD.

2.2.4 Direct impedance method

Growth rates of all strains at 27, 37, 41 and 44°C were obtained using the RABIT system (Rapid Automated Bacterial Impedance Technique, Don Whitley Scientific Ltd., Figure 4). Bacterial strains were inoculated into 9 ml of TSA and incubated overnight at 37°C. RABIT tubes were filled with 2 ml of infant formula, which was reconstituted aseptically according to the manufacturer's instructions with sterile distilled water, and the tubes were inoculated with 100 µl of serial dilutions of the overnight cultures. Four serial dilutions in duplicate (10^{-2} – 10^{-5}) were recorded for each growth rate. The tubes were then incubated in the RABIT system at 27, 37, 41 or 44°C using the direct impedance setting whilst the impedance changes were recorded

by the system. The time to detection (TTD) and impedance curves were recorded for each sample. Impedance threshold of 8 µS was used, which means that a TTD was called if a change of impedance equal or greater than 8 µS was reached in three subsequent readings. The readings were taken every 6 minutes. The samples that produced inferior curves or curves indicating contamination (as described in 2.2.6) were disregarded in the calculation. The generation times for each strain in a given medium and temperature were calculated as described in 2.2.3.

Figure 4. RABIT system (Rapid Automated Bacterial Impedance Technique)

Note that in preliminary experiments, seven different serial dilutions ($10^{-2} - 10^{-8}$) were used for the calculation of the generation times, but after establishing that a good R^2 coefficient of the linear regression could be obtained using 3-4 first serial dilutions, we limited the samples analysed to four serial dilutions (10^{-2}-10^{-5}), which also limited the sampling error associated with using low bacterial concentrations. The 10^{-1} serial dilution was not used because for very high bacterial concentrations, the relationship

47

between bacterial concentration and TTD is not linear, but hyperbolic, and cannot be used for determination of doubling time (Eden & Eden 1984).

2.2.5 Indirect impedance method

Bacterial strains were inoculated into 9 ml of TSA and incubated overnight at 37°C. 1% molten agar (w/v) was mixed with 1% (w/v) KOH and 0.75 ml of the mixture were pipetted into the bottom of the plastic RABIT impedance tubes, not allowing the mixture to set on the sides of the tube. Two millilitres of PIF aseptically reconstituted according to the manufacturer's instructions were pipetted into sterile small glass tubes and 100 μl of a serial dilution of the overnight culture was added to the tube. Four serial dilutions in duplicate (10^{-2} – 10^{-5}) were used for each experiment. The glass tubes were carefully lowered into the RABIT impedance tubes and sealed with rubber bungs. The tubes were placed in the RABIT system and incubated at 27, 37, 41 or 44°C using the indirect impedance setting (impedance threshold -12 μS) whilst the impedance changes were recorded by the system. The results were processed in the same way as for the direct impedance method, although in this case, the impedance change was negative.

2.2.6 Quality control of impedance curves and TTDs generated by the automated RABIT system

Figure 5 illustrates a set of impedance curves for different dilutions of the same sample and TTDs used for the calculation of the doubling times. In some cases, it is possible that a TTD value is miscalled by the RABIT system. If a TTD value was miscalled due to the contamination of the test sample, the aberrant impedance curve could be easily identified by plotting the impedance curves of different dilutions of the same sample, as shown in Figure 5; the aberrant impedance curve would have a

different shape compared to the other curves of the same set. If the TTD value was miscalled due to an electrical spike or gas production, or when the threshold impedance change was reached before or after the actual exponential growth phase, it was possible to identify the TTD value from the table of impedance values available in the RABIT software package.
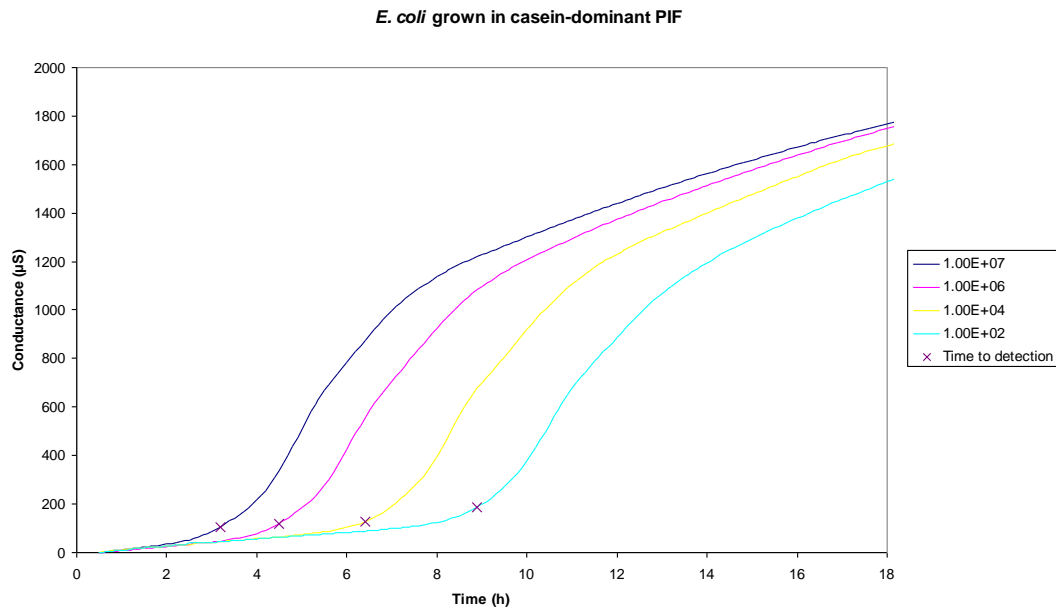
Figure 5. Impedance curves of different serial dilutions of *E.coli* 605 at 37°C

2.2.7 Determination of generation times and lag times by viable count

The use of the RABIT impedance system is limited to the temperatures above ambient. For that reason, we were obliged to choose a different method to examine bacterial growth at 21°C. Two most common methods used to monitor bacterial growth are total viable count and optical density (OD) measurement over a time course. As OD measurement cannot be used with turbid solutions such as reconstituted PIF, our experiments were based on the total viable count approach. Unlike the impedance

49

method, the viable count method allowed us to also determine lag time, a critical parameter for growth modelling.

A single colony from a fresh TSA plate was inoculated into 9 ml of TSB and incubated at 37$^{\text{o}}$C overnight. Powdered infant formula products were aseptically reconstituted according to the manufacturer's instructions in sterile distilled water. The tubes with 10 ml of the reconstituted formulae were inoculated with 100 μl of a 10^{-3} dilution of an overnight TSA culture, which corresponded to a starting inoculum of approximately 10^{5} cfu/ml. The reconstituted formulae were incubated at 21$^{\text{o}}$C over a 20 – 22 h period; samples for the viable count were taken at 1-1.5 h intervals for the first 5 hours of the experiment and then at 2 h intervals. Viable count was performed on TSA according to the Miles & Misra (1938) method after a 16 h incubation period at 37$^{\text{o}}$C. The data obtained were loaded into MicroFit 1.0 software (http://www.ifr.ac.uk/MicroFit/), which generated generation times and lag times for all datasets (Figure 6). The Microfit software tool uses the model proposed by Baranyi & Roberts (1994), which was designated as the model that provides the best fit for the majority of datasets and gives reasonably precise lag time estimates in a comparison of different models by Baty & Delignette-Muller (2004).
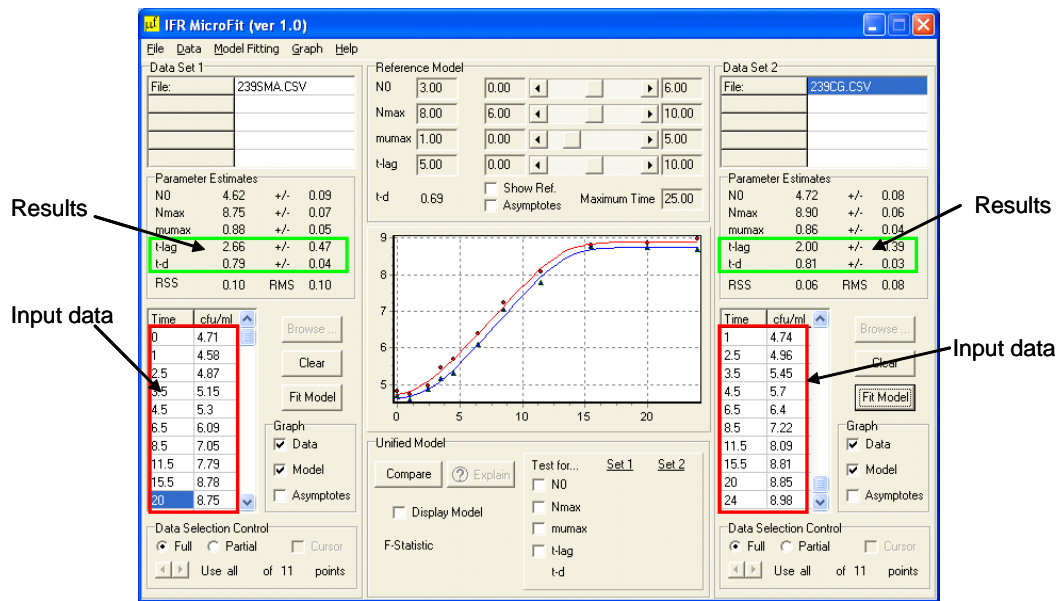
Figure 6. Example of data analysis using Microfit software

2.3. Results

2.3.1 Laboratory determination of generation times

We determined the generation times of 28 bacterial strains across 21 bacterial species in casein- and whey-dominant infant formulae at temperatures 21, 27, 37, 41 and 44°C. This temperature range was chosen because it reflects a wide spectrum of PIF reconstitution, cooling and storage practices (Figure 2). The following sections present the results for Category A organisms (*Cronobacter* and *Salmonella*) and Category B organisms defined by (FAO/WHO 2006).

51

2.3.2 Growth of Category A organisms: *Cronobacter* spp. and *Salmonella* spp.

Generation times were determined for 11 *Cronobacter* strains across 5 *Cronobacter* species (*C. sakazakii, C. malonaticus, C. muytjensii, C. turicensis* and *C. dublinensis*) at 21, 27, 37, 41 and 44°C. The results are presented in Table 2.

Table 2. Generation times of *Cronobacter* strains grown at 21-44°C in casein- and whey-dominant formulae.

Organism	Strain	PIF	Doubling time (min)				
			21°C	27°C	37°C	41°C	44°C
C. sakazakii	1	W	46	41	25	22	20
		C	47	39	22	23	30
C. sakazakii	2	W	56	40	24	24	24
		C	87	37	20	20	21
C. sakazakii	114	W	54	58	33	31	39
		C	55	60	34	39	46
C. sakazakii	566	W	57	NG	21	20	23
		C	60	46	23	19	24
C. sakazakii	658	W	86	35	20	22	28
		C	94	42	22	21	22
C. sakazakii	767	W	36	38	19	20	21
		C	54	41	21	20	23
C. malonaticus	1212	W	59	53	24	23	35
		C	76	48	28	31	31
C. malonaticus	685	W	40	55	25	25	33
		C	58	48	32	35	NG
C. muytjensii	3	W	73	36	22	24	23
		C	73	34	20	20	21
C. turicensis	57	W	48	35	20	18	NG
		C	54	39	23	21	NG
C. dublinensis	582	W	65	33	22	26	NG
		C	68	41	25	28	NG

Footnotes: C = casein-dominant PIF, W = whey-dominant PIF.   NG = growth not detected. Blue font: the shortest doubling time of the strain.

The generation times varied both between the *Cronobacter* species and between strains of each species (Table 2). There were considerable differences in upper temperature range for growth within the *Cronobacter* genera. All *C. sakazakii* strains, *C. muytjensii* and *C. malonaticus* type strain (1212) were able to grow at 44$^\text{o}$C, whereas *C. turicensis* and *C. dublinensis* were not detectable at this temperature. *C. malonaticus* 685 showed different growth behaviour in whey- and casein-dominant PIF; the organism was able to grow in whey-dominant PIF at 44$^\text{o}$C, but its growth in casein-dominant PIF was not detected (Table 2). The generation times of *Cronobacter* strains were also dependent upon formula type. A statistically significant difference in generation times between the two types of formula was observed at 21, 37 and 41$^\text{o}$C (p=0.037, p=0.018 and 0.016, Wilcoxon signed-rank test). At these temperatures, the bacteria grew faster in whey-dominant PIF, whereas no significant difference between the two formulae was noted for temperatures 27 and 44$^\text{o}$C. Interestingly, growth of *C. sakazakii* 566 was not detected in whey-dominant formula.

The most striking observation was the relatively short generation times of some *Cronobacter* strains at elevated temperatures (41 and 44$^\text{o}$C). All *Cronobacter* strains studied except *C. dublinensis* were able to grow the fastest at 41$^\text{o}$C at least in one formula type and 6/11 strains were able to grow at 44$^\text{o}$C at a rate comparable to growth at 37$^\text{o}$C.
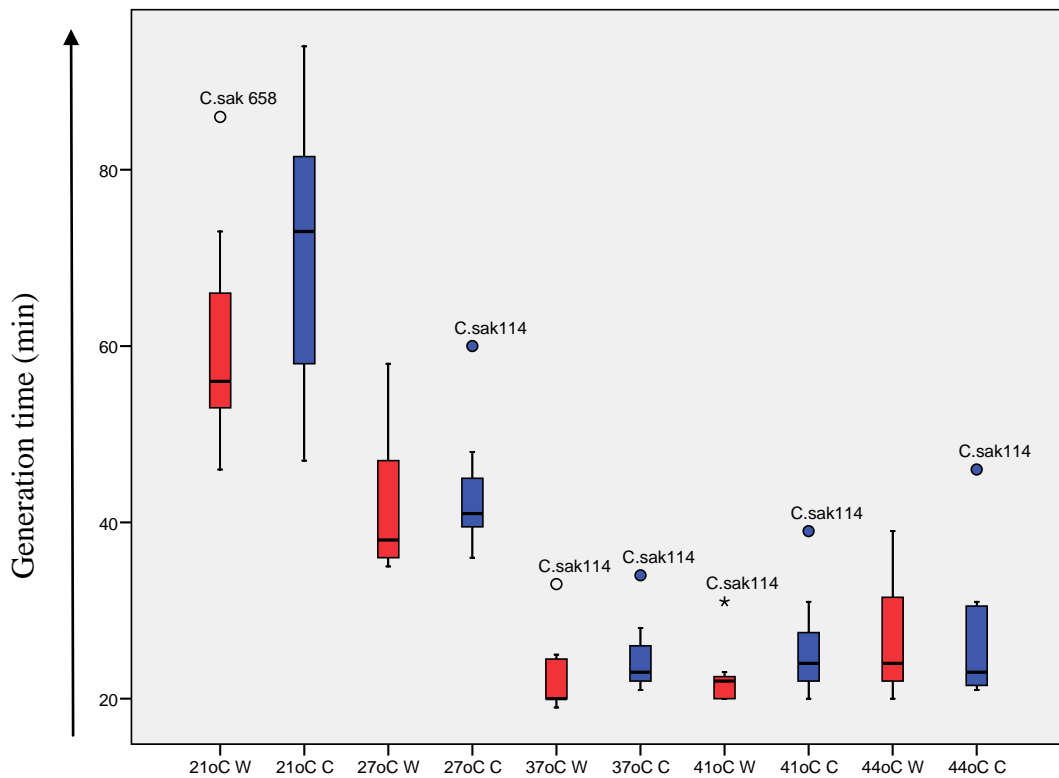
Figure 7. Distribution of generation times of *Cronobacter* (n=11) in whey-dominant and casein – dominant PIF at 21 – 41°C. Blue (C) = casein-dominant PIF, red (W) = whey-dominant PIF.

As we can see from the distribution of generation times of *Cronobacter* strains in Figure 7, the greatest variation in growth was observed at 21°C in both casein- and whey-dominant formula, there was less variation in growth rates at 27 and 44°C and, as expected, the least variation at 37 and 41°C. Except growth at 21°C, there was less variation in generation times in casein-dominant PIF then in whey-dominant PIF. *C. sakazakii* strain 114 clearly showed different growth behaviour compared to other *Cronobacter* strains, as its growth was significantly slower. Figure 7 also shows that the generation times of *Cronobacter* strains at 37, 41 and 44°C were comparable and close to the optimal rate of growth.

Figure 8 shows the impedance curves of *C. sakazakii* strain 2 in whey- and casein-dominant PIF. Surprisingly, the impedance curves of *C. sakazakii* pictured are different even though the two formulae have almost identical stated nutritional content.

Growth of *C. sakazakii* in two diferent types of PIF at 37°C
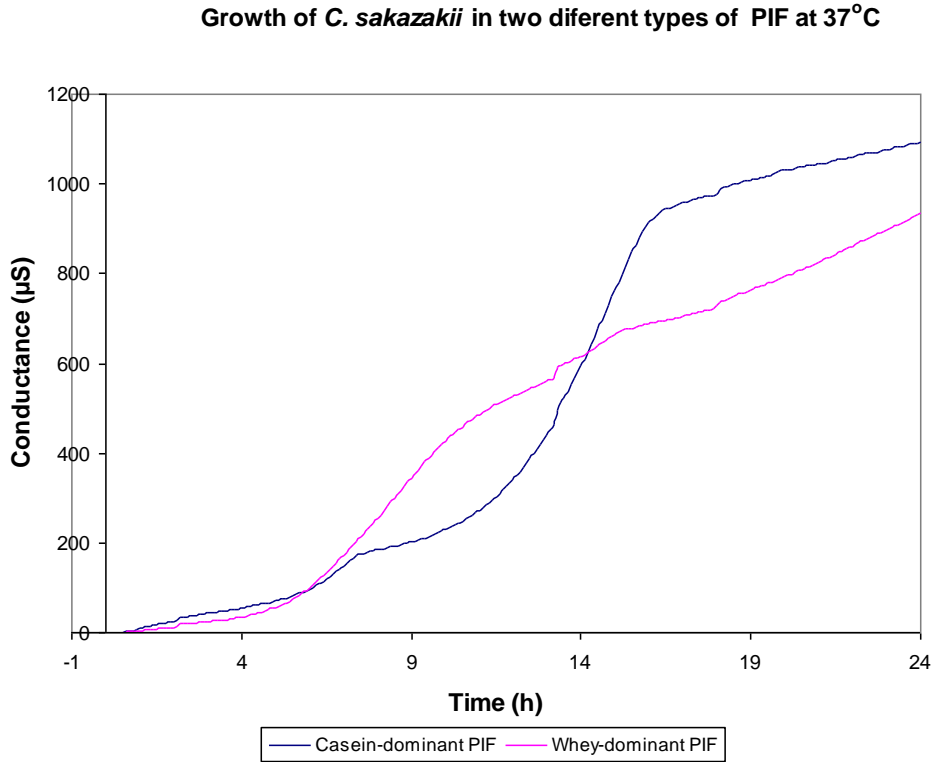
Figure 8. Impedance curves of *C. sakazakii* (strain 2) grown in casein- and whey-dominant formulae.

55

All *Salmonella* strains grew at 21, 27, 37 and 41°C, but none of them grew at 44°C, except *S.* Senftenberg, which had a generation time of 60 min in whey-dominant PIF (Table 3). Similarly to *Cronobacter* strains, growth of *Salmonella* strains was faster in the whey-dominant PIF compared to casein-dominant PIF, although it was not possible to assess the significance of this difference statistically because of limited numbers of data (Table 3).

Table 3. Generation times of *Salmonella* strains grown at 21-44°C in casein- and whey-dominant formulae.

Organism	Strain	PIF	Doubling time (min)				
			21°C	27°C	37°C	41°C	44°C
Salmonella Anatum	1197	W	43	33	19	36	ND
		C	41	40	23	39	ND
Salmonella Enteritidis	358	W	52	39	25	25	ND
		C	46	43	30	30	ND
Salmonella Senftenberg	239	W	49	35	20	27	60
		C	47	29	24	32	ND

Footnotes: C = casein-dominant PIF, W = whey-dominant PIF.   NG = growth not detected. Blue font: the shortest doubling time of the strain.

2.3.3 Growth of Category B organisms: other *Enterobacteriaceae*

Generation times were determined for 14 *Enterobacteriaceae* strains across 7 *Enterobacteriaceae* species at 21, 27, 37, 41 and 44°C (Table 4).

Table 4. Generation times of *Enterobacteriaceae* strains grown at 21-44°C in casein- and whey-dominant formulae.

Organism	Strain	PIF	Doubling time (min)				
			21°C	27°C	37°C	41°C	44°C
Citrobacter freundii	153	W	52	36	25	39	NG
		C	46	39	26	48	NG
Citrobacter koseri	598	W	46	40	23	29	NG
		C	52	41	23	23	NG
Enterobacter cancerogenus	806	W	41	38	23	29	NG
		C	43	37	24	24	NG
Enterobacter cloacae	49	W	58	39	23	26	52
		C	55	40	22	26	58
Enterobacter hormaechei	790	W	36	33	20	20	NG
		C	30	32	22	22	NG
Escherichia coli	605	W	85	40	21	20	24
		C	77	37	21	19	21
Escherichia coli	790	W	43	37	18	17	18
		C	41	37	21	20	22
Escherichia coli	796	W	43	20	20	18	20
		C	37	18	15	16	26
Escherichia vulneris	52	W	101	39	29	NG	NG
		C	110	36	28	NG	NG
Hafnia alvei	236	W	54	37	38	NG	NG
		C	55	41	NG	NG	NG
Klebsiella oxytoca	379	W	38	37	23	23	NG
		C	36	40	23	23	NG
Klebsiella pneumoniae	273	W	48	40	21	21	33
		C	52	41	22	24	42
Pantoea ananatis	44	W	109	28	33	NG	NG
		C	96	43	33	NG	NG
Serratia marcescens	464	W	41	37	27	NG	NG
		C	43	34	23	NG	NG

Footnotes: C = casein-dominant PIF, W = whey-dominant PIF.   NG = growth not detected. Blue font: the shortest doubling time of the strain.

The optimum growth temperature varied between the *Enterobacteriaceae.* For *C. freundi, C. koseri, E. cancerogenus, E. cloacae, E. vulneris, Hafnia, Pantoea and Serratia* it was close to 37°C, whereas for *E. hormaechei, E. coli* 605, 790 & 796, *K. oxytoca* and *K. pneumoniae* the optimum temperature was closer to 41°C (Table 4).

At lower temperatures (21°C, 27°C and 37°C), impedance response was detected for all strains except *Hafnia* in casein-dominant PIF*,* whereas *Hafnia, Serratia, Pantoea* and *E. vulneris*, did not grow at higher temperatures (41°C and 44°C, Table 4). Since lactose is the major sugar in both formulae and all these organisms are lactose non-fermenters, we hypothesized that these organisms might have not produced enough ionic fermentation end products and their growth could have been missed by the direct impedance method. The growth of these four species was re-determined using the indirect impedance method, whereby growth is detected by the liberation of carbon dioxide (Silley, Forsythe 1996). The indirect method confirmed the results obtained via the direct method and showed that these strains did not grow in PIF at elevated temperatures. Same values were obtained for *H. alvei* (generation time of 38 min in whey-dominant PIF and no detectable response in casein-dominant PIF) using both direct and indirect impedance methods. At 44°C growth in both types of PIF was inhibited in most of the *Enterobacteriaceae*, the only organisms which were able to grow at this temperature were *E. cloacae*, all *E. coli* strains (605, 790 & 796) and *K. pneumoniae* (Table 4).
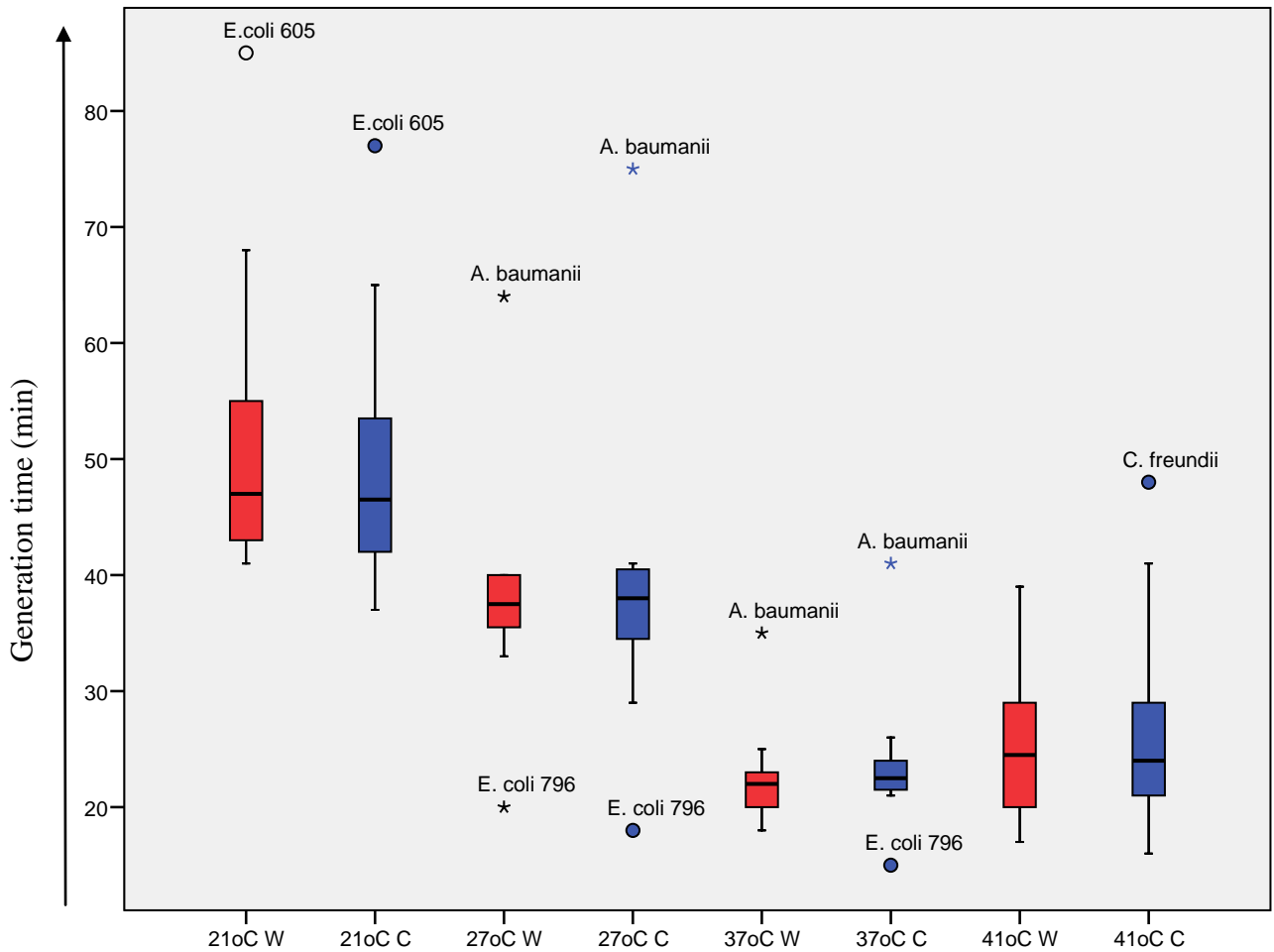
Figure 9. Distribution of generation times of *Enterobacteriaceae* (n=14) in casein- and whey-dominant PIF at 21 – 41°C. Blue (C) = casein-dominant PIF, red (W) = whey-dominant PIF.

There was a greater variation in generation times between the *Enterobacteriaceae* strains at 21 and 44°C, whereas the range of their generation times at 27 and 37°C were smaller (Figure 9). Some strains showed different growth behaviour compared to the majority of *Enterobacteriaceae*; these are shown as outliers in Figure 9. For example, *E. coli* 605 grew significantly slower in both types of formula at 21°C, *A. baumanii* grew significantly slower at 27 and 37°C in both formulae and *C. freundii* grew significantly slower at 44°C in casein-dominant PIF (Figure 9). It was also notable that *E. coli* 796 grew significantly faster in both formulae at 27°C and at 37°C

in casein-dominant PIF. Unlike for *Cronobacter* strains, no statistically significant difference between the casen- and whey-dominant formula was observed in *Entrobacteriaceae* strains (Wilcoxon signed-rank test).

Figure 10 shows that different bacterial species can have characteristic impedance curves. Besides the speed of the microbial metabolism and growth, the shape of the curves reflects the dynamics of the microbial metabolism including switching to different energy sources during fermentative metabolism.

E. coli, C. koseri and E. cloacae grown in casein-dominant PIF
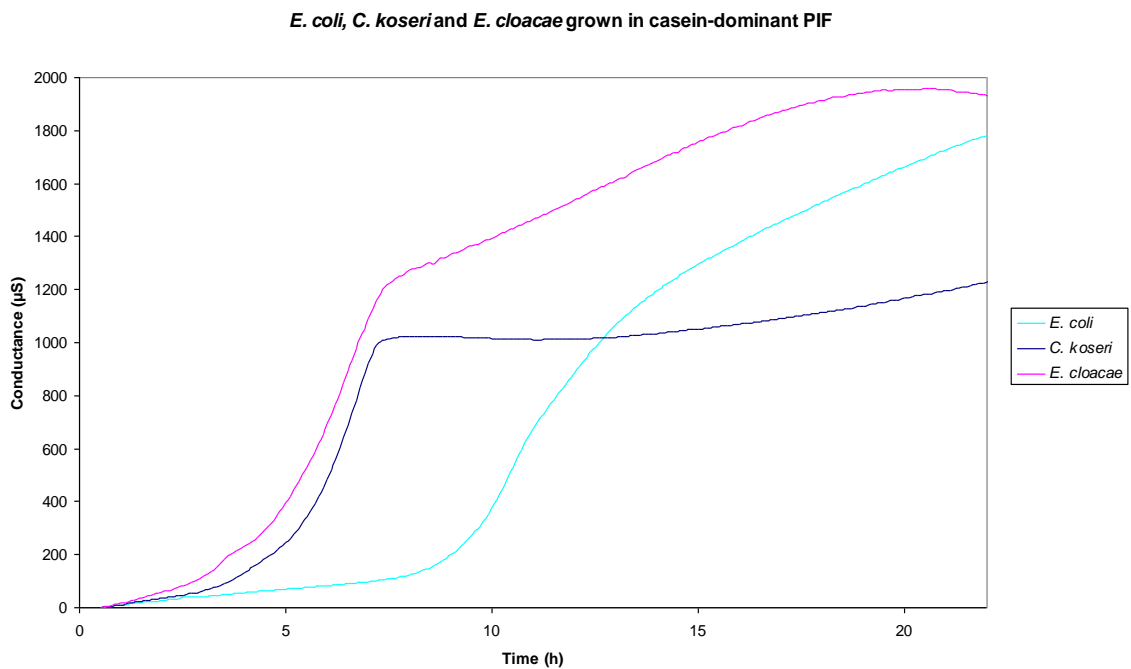
Figure 10. Impedance curves of three *Enterobacteriaceae* species in casein-dominant formula

2.3.4 Bacterial lag phase and lag time

In microbial kinetics, the lag phase is usually observed before bacterial multiplication as a delayed response to a change in environment, *i.e.* change of the culturing conditions, and can occur in both growth and inactivation processes (Swinnen *et al*. 2004). It is defined as a transition period during which the specific growth rate increases to the maximum value characteristic for the culture environment. As the lag phase duration increases with decreasing temperatures (Buchanan & Klawitter 1991) its determination is especially important for growth modelling at sub-optimal temperatures. Lag phase duration is influenced by a variety of factors such as changes of environmental conditions, the entity and the phenotype of the bacterium, the growth stage or physiological history of the cells and the inoculum size (Swinnen *et al*. 2004).

2.3.5 Laboratory determination of lag time

We determined the lag times for category A (*Cronobacter* and *Salmonella*), and category B (other *Enterobacteriaceae* and *Acinetobacter*) organisms in casein- and whey-dominant PIF at 21°C, as the lag times at 21°C were expected to be the longest and would have the greatest impact on predictive growth modelling. For details on the method see 2.2.7. The lag times of both category A and B organism are presented in Table 5.

Table 5. Lag times (h) of category A (*Cronobacter* and *Salmonella*) and category B (*Enterobacteriaceae* and *Acinetobacter*) organisms at 21°C in whey- and casein-dominant PIF.

Organism	Strain	Whey-based PIF	Casein-based PIF
Cronobacter sakazakii	1	1.6	2.7
Cronobacter sakazakii	2	1.8	0.0*
Cronobacter sakazakii	114	2.7	4.6
Cronobacter sakazakii	658	0.6*	1.5*
Cronobacter sakazakii	767	2.7	1.9
Cronobacter malonaticus	1212	2.2	1.5
Cronobacter malonaticus	685	3.2	3.2
Cronobacter sakazakii	658	1.5*	0.6*
Cronobacter muytjensii	3	1.5	2.4
Cronobacter turicensis	57	1.9	1.8
Cronobacter dublinensis	582	1.8	3.4
Salmonella Anatum	1197	1.5	2.1
Salmonella Enteritidis	358	1.8	1.5
Salmonella Senftenberg	239	2.0	2.7
Acinetobacter baumanii	415	4.0	9.1
Citrobacter freundii	153	1.3	3.6
Citrobacter koseri	598	73	36
Escherichia coli	605	1.0	1.6
Escherichia coli	780	2.7	3.6
Escherichia coli	796	2.4	3.3
Enterobacter cloacae	49	1.8	3.0
Enterobacter cancerogenus	806	2.0	3.3
Enterobacter hormaechei	790	1.3	2.8
Escherichia vulneris	52	1.3	3.3
Hafnia alvei	236	0.9	2.0
Klebsiella oxytoca	379	1.5	2.1
Klebsiella pneumoniae	273	0.7	2.0
Pantoea ananatis	44	3.2	5.4
Serratia marcescens	464	2.3	3.4

W - Whey-dominant PIF   C - Casein-dominant PIF   * - lag times not accurately predicted by Microfit

It seems apparent that organisms grown in casein-dominant PIF tend to have longer lag times compared to whey-dominant PIF. However, this difference was only statistically significant for category B organisms (*Enterobacteriaceae* and *Acinetobacter*) at the level of p=0.01 (Wilcoxon signed-rank test). There was no statistically significant difference in lag times between the casein and whey-dominant formula in category A organisms (*Cronobacter* and *Salmonella*). The mean lag time at 21°C for category A organisms in casein-dominant PIF was 2.5 h and 2.1 h in whey-dominant PIF; the mean lag time for category B organisms in casein-dominant PIF was 3.0 h and in whey-dominant PIF 1.7 h. *Acinetobacter* had remarkably long lag time in both formulae, 9.1 h in casein-dominant and 4.0 h in whey-dominant PIF.

A.   B.
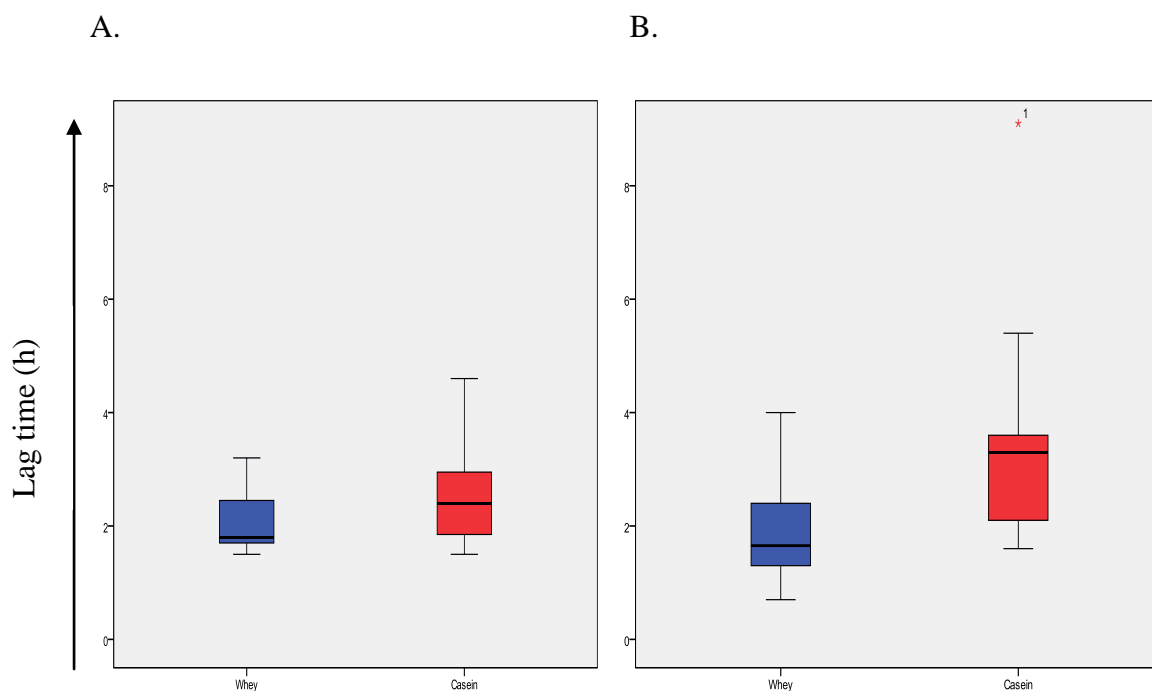
Figure 11. Distribution of lag times of (A) *Cronobacter* (n=14) and (B) selected *Enterobacteriaceae* and *Acinetobacter* (n=15) in casein- and whey-dominant formulae at 21°C. Outlier 1 = *Acinetobacter baumanii.*

2.3.6 Predictive modelling of bacterial growth & risk assessment

The growth data obtained for the Category A and B organisms obtained in our study were used to model predicted temperature profiles of *Cronobacter* and other *Enterobacteriaceae* in casein- and whey-dominant PIF (Figure 12). In combination with the heat inactivation data obtained by Dr. Juncal Caubilla-Baron, the data were used to update the JEMRA Risk Model for *Cronobacter* spp. The modelling was done by Decisionalysis Risk Consultants, Inc.

Bacterial growth and death rates in whey- and casein-dominant infant formulae for *Cronobacter* spp., *Salmonella* serovars, other *Enterobacteriaceae*, and *Acinetobacter* spp. were used with the JEMRA *Cronobacter* risk model for 7 scenarios with reconstitution temperatures between 10 and 70°C. The scenarios are listed in Appendix 1. For all scenarios, reconstitution at 70°C predicted a 1-log decrease in bacterial numbers for *Salm.* Senftenberg and *Salm.* Enteritidis, and $>10^4$ log decrease for *C. sakazakii* depending on the formula base. The model also predicted overall decreases in bacterial numbers for all scenarios when the reconstitution temperature was 60°C, although this was less than 0.5 log for some strains. In contrast, at 50°C some strains had predicted increases depending upon the scenario and the formula base.

In summary, using various scenarios, reconstitution of PIF with water at a temperature $>60°C$ (and in particular $>70°C$) was predicted to reduce the overall number of *Cronobacter* spp., *Salmonella*, other *Enterobacteriaceae* and *Acinetobacter* sp. The extent of bacterial kill was dependent upon the bacterial strain, type of formula, and scenario. The full report is available at: http://www.foodbase.org.uk/results.php?f_report_id=395.
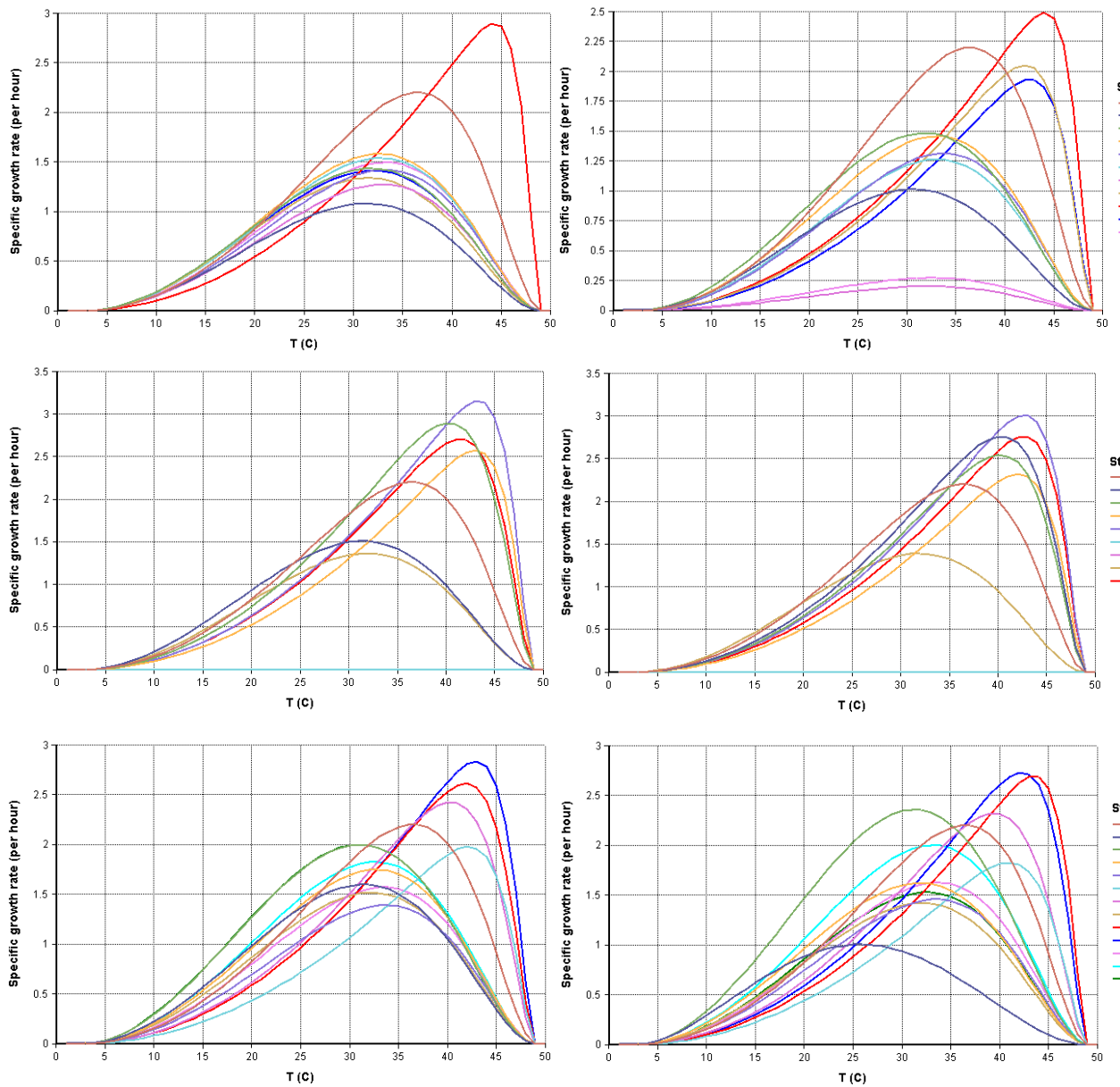
Figure 12. Predicted temperature profiles for the category A and B organisms (split into 3 sets) included in the data set. Whey- and casein-dominant infant formulae are on the left and the right, respectively. In each case the baseline model is included for comparison.

2.4. Discussion

Cronobacter species are opportunistic pathogens that can cause life-threatening disease including necrotizing enterocolitis, septicaemia and meningitis, with premature neonates and infants at a particular risk. Neonates are a high risk population due to their immature immune system which can promote opportunistic infections, lower acidity of the stomach, and absence of an established gut microflora (FAO/WHO 2004). Unlike infant formula, human breast milk contains compounds with protective properties that limit bacterial growth. Processing technologies are unable to completely eliminate potentially pathogenic bacteria in breast milk substitutes without affecting their nutritional properties and it is known that *Cronobacter* occurs as one of the contaminating organisms in PIF (Muytjens *et al*. 1988, Iversen & Forsythe 2004a, Chap *et al*. 2009, Jaradat *et al*. 2009). Moreover, the temperature abuse of reconstituted PIF was reported as a common risk factor in *Cronobacter* outbreaks in France (Coignard *et al*. 2006, Caubilla-Barron *et al*. 2007) associated with deaths of infants. For these reasons, stringent control measures including temperature control need to be put in place, which necessitates a robust risk model that accounts for the considerable phylogenetic diversity of the *Cronobacter* genus confirmed by the recent taxonomic reclassification of the organisms previously known as '*E. sakazakii*'.

The initial 2006 JEMRA Risk Model for *Cronobacter* spp. used data supplied by NTU and Prof. Marcel Zwietering (Wageningen University) from FAO/WHO 2004 and FAO/WHO 2006. The risk model was constructed prior to the taxonomic revision of *Cronobacter* (Iversen *et al*. 2008) and did not consider the *Cronobacter* species diversity. Moreover, it did not consider *Salmonella* serovars or other *Enterobacteriaceae* potentially present in breast milk substitutes. It is pertinent to note that recently there have been outbreaks of *Salmonella* infections in neonates linked to infant formula in Spain and France (Jourdan *et al*. 2008, Soler *et al*. 2008) and other *Enterobacteriaceae* examined in this study are frequent contaminants of infant

formula, and potential pathogens. It was also unknown whether their growth behaviour significantly varied between *Cronobacter* spp. or was dependent on formula type.

This study for the first time presents growth data for five *Cronobacter* species: *C. sakazakii, C. malonaticus, C. turicensis, C. muytjensii* and *C. dublinensis* at a range of temperatures reflecting wide spectrum of PIF reconstitution, cooling and storage practices in two different types of formula, casein-and whey-dominant. In addition, it presents data for *Salmonella, Enterobacter cloacae, Citrobacter koseri, Citrobacter freundii, Klebsiella oxytoca, Klebsiella pneumoniae, Pantoea agglomerans, Escherichia vulneris, Escherichia coli, Hafnia alvei, Serratia* spp., *Acinetobacter* sp, *Enterobacter hormachei* and *Enterobacter cancerogenus.* Altogether, data were obtained for all Category A and Category B organisms as defined by (FAO/WHO 2006).

We determined the generation times of 28 bacterial strains across 21 bacterial species in casein- and whey-dominant infant formula at temperatures 21, 27, 37, 41 and 44°C, including 11 *Cronobacter* strains representing different *Cronobacter* species (*C. sakazakii, C. malonaticus, C. muytjensii, C. turicensis* and *C. dublinensis*) and three strains of *Salmonella.* Please note that the generation times at 21°C were obtained by viable count, whereas all other values were obtained using impediometry, and therefore these two sets of values were not determined with the same level of precision.

There was considerable variation in doubling times among *C. sakazakii* strains and between different *Cronobacter* species at 21°C. At 37, 41 and 44°C, however, the range of generation times between all *Cronobacter* strains was quite narrow and the growths were comparable to the optimal rate of growth. This shows that *Cronobacter* is an organism well adapted for growth at a wide range of temperatures including 44°C, which inhibits growth of most other *Enterobacteriaceae.*

There were considerable differences in upper temperature range for growth within the *Cronobacter* genera. All *C. sakazakii* strains*, C. muytjensii* and *C. malonaticus* type strain (1212) were able to grow at 44°C, whereas *C. turicensis* and *C. dublinensis*

67

were not detectable at this temperature. *C. sakazakii* strain 114 had different growth characteristics compared to other *Cronobacter* strains; its growth was significantly slower at all temperatures except 21°C. Unfortunately, strain 114 was one of the three strains in our collection with unknown origin, so we cannot make any hypotheses about the possible reasons. In an extensive study on growth rates of *Cronobacter* in PIF and breast milk, Lenati (2008) reported that the strain's origin had a significant effect on growth rate. The average generation times of clinical isolates at 23 and 37°C were longer in breast milk than in fortified BM, compared to the environmental and food isiolates. The origin of strains had no effect on generation times in PIF, however.

Generation times were also determined for 14 other *Enterobacteriaceae* strains across 7 *Enterobacteriaceae* species. The optimum growth temperature *C. freundi, C. koseri, E. cancerogenus, E. cloacae, E. vulneris, Hafnia, Pantoea and Serratia* was 37°C, whereas for *E. hormaechei, E. coli* 605, 790 & 796, *K. oxytoca* and *K. pneumoniae* the optimum temperature was 41°C. Interestingly, *Hafnia, Serratia, Pantoea* and *E. vulneris* did not grow at higher temperatures (41°C and 44°C), which was confirmed by both direct and indirect impedance method. Whether the inhibition of growth is connected with the organisms lack of ability to ferment lactose or the organisms' decreased thermotolerance is unknown.

The variation in generation times between the *Enterobacteriaceae* strains at 21 and 44°C was greater compared to the variation at 27 and 37°C. Considerable differences in growth behaviour compared to the other *Enterobacteriaceae* were observed in *E. coli* 605, which grew slower in both types of formula at 21°C, *A. baumanii*, which grew significantly slower at 27 and 37°C in both formulae and in *C. freundii*, which grew significantly slower at 44°C in casein-dominant PIF. It was notable that *E. coli* 796 grew significantly faster in both formulae at 27°C and at 37°C in casein-dominant PIF.

One of the most surprising results of this study is the upper growth temperature of *Cronobacter* strains. All *Cronobacter* strains except *C. dublinensis* and *C. turicensis* were able to grow the fastest at 41°C at least in one formula type and most *Cronobacter* strains were able to double every 22 min at 44°C, which is the rate of growth comparable to growth at 37°C. Considering that the reconstitution temperature

in some feeding practices is < 50°C, this has a very important implication for the safety of PIF products. Reconstitution temperature close to 50°C would not mean a significant reduction of the microbial load in the product, it could even support microbial growth during cooling and storage of reconstituted PIF.   The results from our study suggest that most *Cronobacter* species and some other *Enterobacteriaceae* are able to grow at a surprisingly fast rate even at elevated temperature of 44°C. The results highlight the need for new guidelines for reconstitution and storage of infant formula.

None of the *Salmonella* strains tested grew at 44°C, except *S.* Senftenberg, which had a generation time of 60 min in whey-dominant PIF and growth at 41°C was slower compared to *Cronobacter* strains. Similarly to *Salmonella*, growth at 44°C was inhibited in most of the Category B organisms; the only species which were able to grow at this temperature were *E. cloacae*, all *E. coli* strains (605, 790 & 796) and *K. pneumoniae*.   This suggests that *Salmonella* and other *Enterobacteriaceae* except *E. coli* are less heat tolerant and that the current reconstitution practices would reduce the contamination in reconstituted PIF and therefore the risk of infection by these organisms. However, the risk posed by *Cronobacter* and *E. coli* potentially present in PIF would not be adequately controlled using the current reconstitution practices, as reconstitution at temperatures close to 50°C would not stop bacterial multiplication in the products. In light of these findings, new guidelines for PIF reconstitution should be considered.

Our results showed that generation times of *Cronobacter* strains were also depending upon the type of PIF used. The bacteria grew significantly faster in whey-dominant PIF at 21, 37 and 41°C, whereas there was no significant difference between the two formulae at 27 and 44°C. Similarly to *Cronobacter* strains, growth of *Salmonella* strains was faster in the whey-dominant PIF at 21, 37 and 41°C compared to casein-dominant PIF. Both *C. malonaticus* 685 and *Salmonella* Senftenberg were able to grow in whey-dominant PIF at 44°C, but they did not grow in casein-dominant PIF. In general, we observed less variation in generation times in casein-dominant PIF than in whey-dominant PIF. This suggests that whey protein may have a certain thermo-protective effect.

Enterobacteriaceae had significantly shorter lag times in whey-dominant PIF at 21°C. However, there was no significant difference between the lag times of *Cronobacter* between the two types of formula.

In summary, notable differences in generation times and lag times have been observed between casein- and whey-dominant PIF, although these should have virtually identical nutritional content according to their product labelling. The impedance curves obtained for growth of *Cronobacter* and other *Enterobacteriaceae* confirm that the bacteria do indeed have different growth behaviour in the two formulae. In both types of formulae, the protein is derived from cow's milk protein. In whey-dominant milks the protein is adjusted to a similar casein to whey ratio to human breast milk (40:60) and the formula may also have a lower mineral content, in particular sodium and potassium. In casein-dominant formulae the ratio of casein to whey is the same as in cows' milk (80:20). The different growth behaviour or *Cronobacter* and other *Enterobacteriaceae* might be attributed to different source of the key nutrients, presence of inhibitory substances in the particular formulations or a discrepancy between the product label and the product content.

The risks associated with the growth of *Cronobacter* in breast milk substitutes were further highlighted by the recent study examining growth of pathogenic bacteria inside the enteral feeding tubes published by our team at NTU (Hurrell *et al*. 2009b). The studies identified the enteral feeding tubes, which are used in intensive neonatal care units (NICU) to provide nutrition to hospitalised infants, as loci for bacterial multiplication that can significantly increase the risk of infection by *Cronobacter* or other pathogenic bacteria. As it was demonstrated in another study (Hurrell *et al*. 2009a), biofilm formation in the interior of the feeding tubes further contributes to the contamination of infant feed. It therefore seems essential to limit bacterial contamination in the original reconstituted formula before it reaches the enteral feeding tube and temperature control of the prepared feeds becomes a crucial hazard control point. However, it was shown in Hurrell *et al*. (2009b) that significant numbers of *Enterobacteriaceae* were recovered from the biofilms inside enteral feeding tubes in infants that were fed infant formula, but also breast milk and sterile ready-to-feed formula. This suggests that microbiological control should not be only

focussed on the intrinsic contamination of PIF, but also on general practices of preparation of feeds for enteral feeding.

The growth data obtained in this study were, along with the thermal inactivation data and a reconstitution practices survey collected by Dr. Juncal Caubilla-Baron, provided to the Food Standards Agency in order to update the JEMRA Risk Model for *Cronobacter* spp., which will help the regulatory bodies and formula manufacturers evaluate the risks associated with contamination of PIF by *Cronobacter* and other potentially harmful bacteria. The results of this study and the derived risk model may also prompt the UK regulatory bodies to enforce the change of the formula preparation guidelines provided with their product in order to protect the most vulnerable part of our population, infants.

3. CRONOBACTER SPP. COMPARISON BY COMPARATIVE GENOMIC HYBRIDIZATION

3.1. Introduction

3.1.1 DNA microarray technology

Microarray technologies are a popular tool to study interactions between complex biopolymers such as DNA, proteins or lipids. The term 'microarray' simply refers to the use of multiple arrays of spatially organized molecules miniaturized to a small format. DNA microarray technology is an automated, high-throughput technique developed on the basis of Southern blotting that facilitates simultaneous parallel analysis of potentially thousands of genes from a given biological species. The technology is based on hybridization of chemically or enzymatically labelled sample DNA to reporter elements bound to a surface of a microarray slide. The probes are attached or directly synthetised on a microarray slide surface; and they can consist of genomic regions such as open reading frames (ORF), short oligonucleotide probes or 'tiled' whole genomes. In a tiling array, probes (usually synthetic oligonucleotides) covering the whole genome of an organism are represented on the microarray, often with overlapping sequences. The material and spatial organization of microarray slides vary between different microarray formats, depending on the question that needs to be examined. After the hybridization reaction, the unbound probes are washed away using different stringency buffers and the remaining fluorescent signal resulting from complementary binding between oligonucleotide probes on the array probes and sample DNA is measured using a confocal microarray scanner. A reference sample is included in the experimental setup to compare the test sample results to a control condition or to a reference organism. In bacterial comparative genomic hybridisation (CGH), the reference sample is typically the same sequenced strain that was used to design the array probes, and its DNA should thus hybridize to

72

all probes on the microarray. The presence or absence of a DNA sequence in a test strain is determined by comparing the fluorescent intensities of the test strain element against the reference strain element.

The most widely used application of DNA microarrays is expression analysis, which compares the expression patterns of samples subject to different culture conditions, carrying mutated genes or originating from different tissues, whereas the Comparative Genomic Hybridization (CGH) method has been developed to examine the variation between the genomic content of closely related organisms. The CGH methodology has been used extensively for DNA copy-number analysis in humans to identify chromosomal aberrations that can lead to cancer and other diseases. In microbiology, CGH serves as a tool to examine presence and absence of particular genes or genomic regions, as well as to determine phylogenetic relationship between tested strains. CGH analysis can reveal the presence of various evolution events such as acquisition of genes by horizontal gene transfer and gene deletions on the whole-genome scale. Apart from expression analysis and CGH, microarrays can be also used for resequencing, microbial identification, single nucleotide polymorphism (SNP) analysis and drug discovery (Huyghe *et al.* 2009).

There are two basic types of elements used for the microarray design: PCR products and oligonucleotide probes. The advantages of PCR product-based microarrays are their relative simplicity, relatively low cost, and reproducibility (Dorrell *et al.* 2005). On the other hand, oligonucleotide-based DNA microarrays are not dependent on PCR to amplify the gene targets, they offer a reduced cross-hybridisation and an increase in the differentiation of homologous regions; mutant alleles or even single nucleotide polymorphisms can be detected using such arrays (Dorrell *et al.* 2005).

3.1.2 Virulence factors of *Cronobacter*

There are a limited number of studies examining the molecular basis of the virulence factors of *Cronobacter*. To date, the metalloprotease Zpx (Kothary *et al.* 2007), the

73

outer membrane proteins OmpA (Mohan Nair & Venkitanarayanan 2006, Singamsetty *et al*. 2008, Mittal *et al*. 2009b, Kim *et al*. 2010) and ompX (Kim *et al*. 2010), the O-antigen cluster (Mullane *et al*. 2008b, MacLean *et al*. 2009a, MacLean *et al*. 2009b, MacLean *et al*. 2009c, Czerwicka *et al*. 2010), biofilm formation (Iversen *et al*. 2004, Lehner *et al*. 2005, Kim *et al*. 2006, Kim *et al*. 2007a, Hurrell *et al*. 2009a), and capsule (Hurrell *et al*. 2009a) have been studied. Since *Cronobacter* is a Gram-negative member of the *Enterobacteriaceae,* it is reasonable to expect other virulence determinants to be present, such as type IV and type VI secretion systems. The following section gives a short review on the virulence factors that were examined in this project.

3.1.2.1 Lipopolysaccharide and O-antigen

Exopolysaccharides (EPS) are components of the bacterial cell envelope; they include the serotype-specific lipopolysaccharide (LPS), capsular K-antigen and other organism-specific compounds. Lipopolysaccharide (LPS), also known as bacterial endotoxin, is a major component of the outer membrane of Gram-negative bacteria. The LPS consists of three parts: lipid A, a conserved core oligosaccharide region, and finally O-antigen, which is a repetitive glycan polymer specific for each serogroup (Stenutz *et al*. 2006), consisting of oligosaccharide units that are repeated between 10 and 30 times and have between 3 to 6 sugars (Trent *et al*. 2006). As a major structure that is exposed on the bacterial surface, the O-antigen is highly immunogenic, causing strong antibody responses (Comstock & Kasper 2006). The lipid A domain, also known as endotoxin, anchors the molecule in the outer membrane and is the bioactive component recognized by the immune system during human infection. The LPS structure is represented in Figure 13.
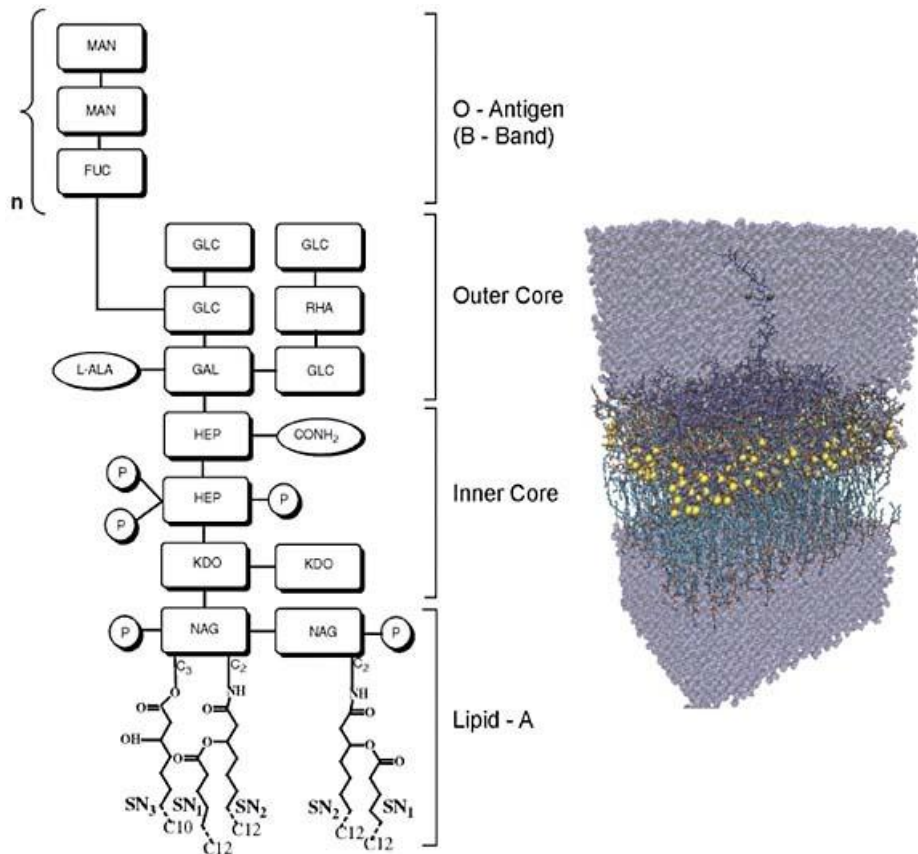
Figure 13. Schematic representation of a LPS unit. Source: (Soares *et al*. 2008). (NAG: N-acetyl-D-glucosamine; P: phosphatidyl group; KDO: 3-Deoxy-D-manno-octulosonic Acid;HEP: heptose; GAL: D-galactose; GLC: D-glucose; L-ALA: L-alanine; RHA: D-rhamnose; FUC: D-fucose; MAN: D-manose). Acyl lipid chains SN1,SN2 and SN3 are labeled (left). Atomistic model of the A-B+ LPS membrane of *Pseudomonas aeruginosa* (right). Membrane atoms are represented in"sticks", Ca++ ions in filled yellow and water in transparent blue model.

Due to their ability to mediate bacterial virulence and host response, LPS and O-antigen are considered important virulence factors.   There is a strong link between invasiveness of *Cronobacter* and LPS presence in PIF. Townsend *et al*. (2007b) showed that PIF is frequently contaminated by high levels of LPS and suggested that LPS may enhance the translocation of *Cronobacter* from the blood to the CNS during *Cronobacter*-induced sepsis, but the mechanism of such enhancement was unclear. The variability of the O-antigen gene clusters in bacterial pathogens has been successfully used as the molecular basis for serotyping (Feng *et al*. 2005, Fitzgerald *et al*. 2006). As it is implicated in avoidance of host immune system and survival in

different environments, the O-antigen locus is subject to strong selective pressure (Reeves 1993). It is known that the variations in O-antigen biosynthesis loci can arise not only from substantial genetic differences, but also from a variation in a single gene and that even slight genetic variation in this locus can have a great impact on immune recognition (Comstock & Kasper 2006).

The O-antigen gene locus usually contains three categories of genes: i) genes encoding enzymes involved in the synthesis of the O-subunit sugars, ii) genes encoding glycotransferases necessary for the assembly of sugars in the O subunit, and iii) genes encoding the transporter (*wzx*) and polymerase (*wzy)* proteins involved in the processing and assembly af the O-antigen from O subunits (Mullane *et al*. 2008b) . The O-antigen locus size depends on the sugar composition, as well as on the antigen structure (Schnaitman & Klena 1993). Mullane *et al*. (2008b) characterized the molecular structure of the O-antigen locus of two *Cronobacter* serotypes O:1 and O:2 and identified nine additional potential serotypes. The authors found that the genetic organization of the O-antigen locus in *Cronobacter* was similar to the corresponding locus in related *Enterobacteriaceae.* Each O-antigen cluster contained 11 unique ORFs.

It has been shown that bacteria can switch between smooth and rough forms of LPS; these two different forms are a result of mutations in one or more genes involved in the synthesis and polymerization of the O-antigen chains (Coimbra *et al*. 2000). In *E. coli,* the O-antigen heterogeneity was a result of mutation in the gene encoding the Wzz protein (Franco *et al*. 1996). Switching between smooth and rough forms of LPS has also been documented in *Cronobacter* (Farmer 1980). It was speculated that mutations in homologs of WxxB O-antigen chain determining protein may cause the differences in O-antigen chain lengths and therefore the existence of S and R forms.

The O-antigen is one of the few structural features of *Cronobacter* that has been investigated and it is known that it varies across the *Cronobacter* spp. In *C. sakazakii* and *C. malonaticus*, the O-antigen is composed of various branched polymers, whereas they are unbranched in *C. muytjensii*. In *C. sakazakii* BAA-894 (MacLean *et al*. 2009a) (Figure 14A) it is a branched polymer of pentasaccharide units composed of 2-acetamido-2-deoxy-D-galactose, 3-(N-acetyl-L-alanylamido)-3-deoxy-D-qui

novose, D-glucuronic acid, and D-glucose. *C. sakazakii* strain 767 is also a branched polymer but of a repeating heptasaccharide composed of 2-acetamido-2-deoxy-D-glucose, and D-galacturonic acid, L-rhamnose, and D-glucose (Czerwicka *et al.* 2010) (Figure 14B). *C. malonaticus* LPS (MacLean *et al.* 2009c) (Figure 14D) is also a branched pentasaccharide unit of 2-amino-2-deoxy-D-glucose, 2-amino-2-deoxy-D-galactose, 3-deoxy-D-manno-oct-2-ulosonic acid, D-galactose and D-glucose residues. *C. sakazakii* ATCC 29004 possesses a branched hexasaccharide O-unit with a randomly mono-O-acetylated terminal rhamnose residue in the side chain. This structure is the most similar to the O-antigen of *C. sakazakii* 767, which differs in the presence of an additional lateral alpha-d-Glcp residue on GlcNAc and the pattern of O-acetylation (Arbatsky *et al.* 2010). Whereas, *C. muytjensii* LPS (MacLean *et al.* 2009b) (Figure 14C) is a linear unbranched pentasaccharide polymer of 2-acetamido-2-deoxy-D-galactose, 2-acetamido-2-deoxy-D-glucose, 2-acetamido-3-deoxy-D-quinovose, L-rhamnose and D-glucuronic acid.

A.

→2)-β-D-Qui*p*3NR-(1→6)-α-D-Glc*p*-(1→3)-β-D-Glc*p*A-(1→3)-α-D-Gal*p*NAc-(1→

$$
\begin{array}{c}
2 \\
\uparrow \\
1 \\
\text{α-D-Glc}p
\end{array}
$$

R = *N*-acetyl-L-alanylamido

B.

$$
\begin{array}{c}
\text{α-L-Rha}p \\
1 \\
\downarrow \\
4
\end{array}
$$

$$
\begin{array}{cc}
\text{α-D-Glc}p & \text{α-D-Gal}p\text{A} \\
1 & 1 \\
\downarrow & \downarrow \\
4 & 2
\end{array}
$$

→ 3)-β-D-Glc*p*NAc-(1→3)-α-L-Rha*p*-(1→4)-α-D-Glc*p*-(1→2)-α-L-Rha*p*-(1→

$$
\begin{array}{c}
4 \\
\uparrow \\
\text{O-Ac (50 \%)}
\end{array}
$$

C.

α-D-Gal*p*A-(1←4)-α-L-Rha*p*-(2, 3 or 4 — Ac (totally ~70%)

$$
\begin{array}{c}
1 \\
\downarrow \\
2
\end{array}
$$

. →3)-α-L-Rha*p*-(1→4)-α-D-Glc*p*-(1→2)-α-L-Rha*p*-(1→3)-β-D-Glc*p*NAc-(1→

D.

→6)-β-D-Glc*p*-(1→6)-β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→4)-β-D-Kdo-(2→

$$
\begin{array}{c}
2 \\
\uparrow \\
1 \\
\text{β-D-Glc}p\text{NAc}
\end{array}
$$

Figure 14. Known O-antigen structures of *Cronobacter*. A. *C. sakazakii* BAA-894 (MacLean *et al*. 2009a); B. *C. sakazakii* 767 (Czerwicka *et al*. 2010); C. *C. muytjensii* 3270 (MacLean *et al*. 2009b); D. *C. malonaticus* (MacLean *et al*. 2009c).

3.1.2.2 Type VI secretion systems

Bacterial secretion systems deliver proteins and toxins across the bacterial membrane into the environment, or onto an eukaryotic target. There have been six secretion systems characterized to date (type I – VI), with type VI being the most newly identified. Type VI secretion systems (T6SS) have been found in almost 100 bacterial species, but have been characterized in only a few Gram-negative bacteria, including *Salmonella enterica, Escherichia coli, Rhizobium leguminosarum, Vibrio cholerae, Pseudomonas aeruginosa, Francisella tularensis,* and *Burkholderia* spp., as referenced in Filloux *et al*. (2008). T6SS are involved in adherence, cytotoxicity, host-cell invasion, growth inside macrophages and survival within the host.

Type VI secretion systems (T6SS) were originally described as systems of Gram-negative bacteria that contain two components of another secretion system, type IV secretion system (T4SS), IcmF and DotU (Filloux *et al*. 2008). Although these proteins were encoded within a conserved gene cluster that did not share homology with other T4SS components, another cluster of conserved genes was being systematically found in the vicinity of these genes, initially known as IAHP (IcmF associated homologous proteins). In the plant symbiont *Rhizobium leguminosarum*, a homologue of dotU ImpK was identified as a component of a system encoding a secretion machine which could block colonization process in pea plants (Bladergroen *et al*. 2003). In *Salmonella enterica*, a complete genomic cluster *sci* (*Salmonella enterica* centisome 7 genomic island) was described as a cluster encoding many T6SS genes. The deletion of the whole *sci* cluster caused a decreased ability of *S. enterica* to enter eukaryotic cells. However, the cluster also included genes involved in fimbrial assembly and invasion (Folkesson *et al*. 2002).

T6SS involve different genomic organisation of various combinations of genes, and the core genes of T6SS are therefore very hard to define (Cascales 2008). A lack of systematic nomenclature further complicates the study of T6SS, as homologous genes may have several names and unrelated genes can be called the same. No consensus in the model for the assembly of the T6SS has been defined yet (Cascales 2008) and T6SS gene clusters of different bacterial species vary in their organization (Filloux *et al*. 2008). The model for the T6SS system proposed by Cascales (2008) based on the reviewed literature is shown in Figure 15. According to an extensive review by Filloux *et al*. (2008), T6SS gene clusters usually encode from 12 to 25 proteins, of which only DotU and IcmF-like proteins have homologues in T4SS and have transmembrane domains. The T6SS clusters do not necessarily encode these two proteins, which may be located outside the T6SS locus and most of the other T6SS genes encode unknown proteins. The locus frequently contains genes for ATPases, which provide energy for the transport process, with ClpB homologues (named ClpV) are found most often. It has been speculated, however, that the function of ClpV might be different from ClpB proteins which solubilise aggregated proteins, such as providing energy-dependent transport of polypeptides (Yeo & Waksman 2004) or folding/unfolding components of T6SS prior to their assembly or T6SS effectors before their secretion (Filloux *et al*. 2008). The T6SS gene clusters also frequently encode a putative outer membrane lipoprotein (VasD in *V. cholerae*) that may contribute to pore formation in the membrane (Filloux *et al*. 2008). The T6SS gene clusters in some bacteria also encode serine/threonine kinases and phosphatases (Mukhopadhyay *et al*. 1999, Mougous *et al*. 2007). It has been shown that in *P. aeruginosa*, both T6SS assembly and effector secretion require phosphorylation by Ser/Thr kinase PpkA, and are prevented from dephosphorylation by Ser/Thr phosphatase PppA (Mougous *et al*. 2007). In some cases, T6SS components contribute to virulence, but independently of other T6SS components. For example, expression of T6SS genes in *Franciscella tularensis* is required for growth and survival in macrophages, but expression of these genes does not require other T6SS genes (Barker *et al*. 2009).

A single organism can have multiple copies of complete or partially complete T6SS loci, often including different types of T6SS, but it is uncertain whether these loci cooperate with each other (Filloux *et al*. 2008) and very little is known about the protein interactions of the various components of the system. The phylogenetic tree constructed by Boyer and colleagues reveals that these do not result from duplication events, and have been conserved throughout evolution, which suggests a mechanism that favours this trend (Boyer *et al*. 2009). The authors proposed that this might be explained by specialization of the systems for different purposes and different cellular targets. Many of the bacterial species with multiple T6SS clusters that are not paralogous, have multiple hosts or diverse environments (Jani & Cotter 2010).
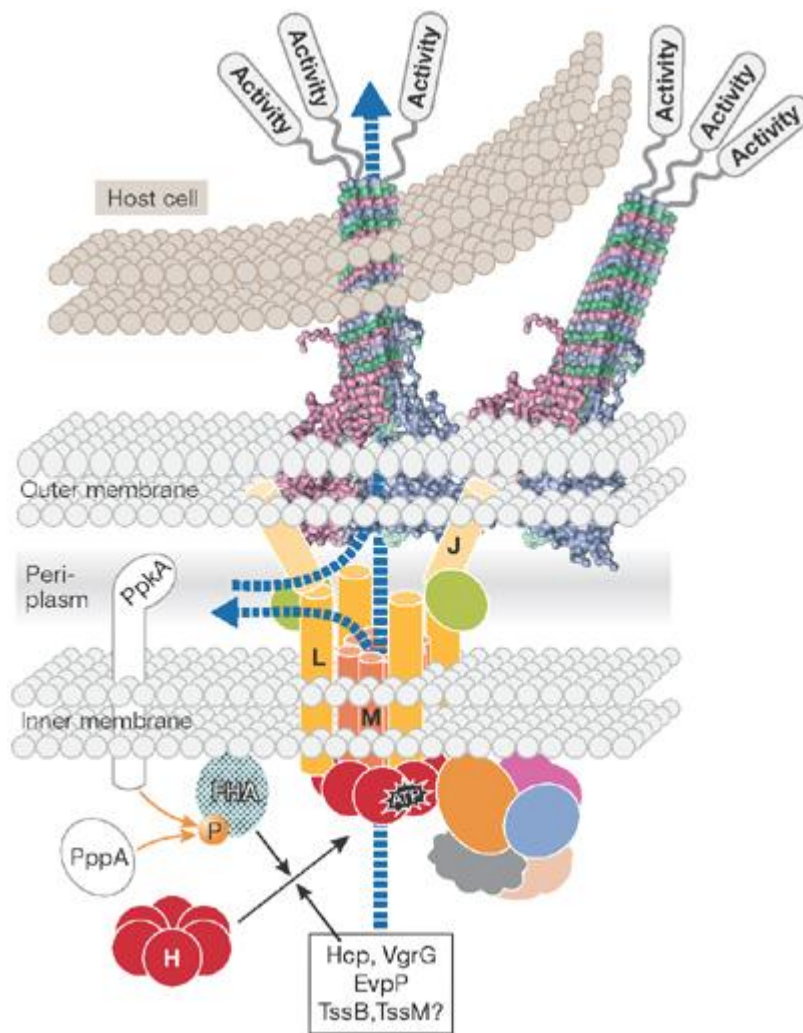
Figure 15. A model for type VI secretion system assembly and function. Source: Cascales (2008). An inner membrane channel formed by the IcmF-like and IcmH-like proteins interacts at the cytoplasmic side of the IM with a complex composed of the probable cytosolic type VI secrection (T6S) subunits and the ClpV AAA$^+$ ATPase. Recruitment of the ClpV multimer is induced by the regulation of forkhead-associated (FHA) phosphorylation through the activities of PpkA and PppA, and by the presence of the Hcp protein. A multimer of the putative lipoprotein in association with periplasmic subunits is shown at the outer membrane. Putative routes for substrate translocation are depicted through the cell envelope and the host-cell membrane (blue arrow) including a 'one-step' mechanism through a unique channel, and a 'two-step' mechanism, in which both steps are catalysed by T6S subunits with transient accumulation in the periplasm (P). This hypothetical model shows a trimeric VgrG inserted into the OM through the amino-terminal domain and puncturing the host cell through the needle-like structure formed by the central domains, releasing the activity domain into the host cytosol (for eukaryotic-like activities) or in the medium (for binding or adhesion activities). J, L, M and H (TssJ, TssL, TssM and TssH respectively) represent the T6S core components, following the nomenclature of (Shalom *et al*. 2007). Source: Cascales (2008).

82

Interestingly, there is a remarkable resemblance of some of the T6SS components (Hcp, VgrG, gp-like protein COG3501) to proteins of bacteriophage origin (Leiman *et al*. 2009). The VgrG protein domains share similarities with the proteins gp5 and gp7 of the bacteriophage T4 spike. The bacteriophage uses dimers or trimers of these proteins to puncture the bacterial envelope and inject its DNA to the cell (Rossmann *et al*. 2004). Pukatzki *et al*. (2007) found that the VgrG proteins form homotrimeric or heterotrimeric complexes, and this type of assembly suggests that the putative phage spike-like complexes may be used as a puncturing device to penetrate the bacterial cell envelope or host cell membrane (Filloux *et al*. 2008). Pukatzki and colleagues also hypothesized that the T6SS Hcp proteins, which form a hexameric rings with a central channel (Mougous *et al*. 2007) might be forming a tube for the T6SS effectors surrounding the VgrG complex.

Although many of the structural T6SS proteins have been described, there is a lack of identified effector proteins (Jani & Cotter 2010). Hcp and VgrG-like proteins were identified as potential secreted effector proteins, as they were found in culture supernatants of most bacteria with T6SS, but it was shown that their secretion is mutually dependant (Dudley *et al*. 2006, Pukatzki *et al*. 2007), suggesting that they may also act as structural components. The experimental study of T6SS is challenging, as T6SS genes are usually not induced in laboratory conditions. For example in *R. leguminosarum*, the T6SS secretion is temperature dependant, with optimum secretion at 24°C (Bladergroen *et al*. 2003) and in *S. enterica*, expression of genes from the T6SS cluster (including the SciS encoding the ImF-like protein) is induced inside macrophages (Parsons & Heffron 2005).

T6SS are considered important factors for virulence and survival within host cells (Cossart & Sansonetti 2004), although recent findings suggest that these systems are more versatile and that they can promote commensal or mutualistic relationships between bacteria (Jani & Cotter 2010). For example, it has been shown that T6SS can limit bacterial replication of virulence by changing the state of virulence to commensal or mutualistic relationship (Bladergroen *et al*. 2003, Parsons & Heffron 2005, Chow & Mazmanian 2010). The two-component response regulator system SsrAB negatively controls the expression of a T6SS gene *sciS* and positively controls

the expression of genes on *Salmonella* pathogenicity island 2 (SPI-2) including T3SS genes that mediate macrophage cell death (Garmendia *et al*. 2003). It seems that in *Salmonella*, T6SS induction limits the bacterial load in host cells and intracellular replication, which leads to bacterial persistence rather than macrophage death and damage to the host (Parsons & Heffron 2005).

3.1.2.3 Fimbriae

Interestingly, fimbriae (or pili) were first discovered as receptors for bacteriophages (Anderson 1949). In 1950, proteinaceous non-flagellar appendages were discovered on the surface on *E. coli* by electron microscopy (Houwink & van Iterson 1950). It was shown that these structures mediate adherence of *E. coli* to eukaryotic cells (Duguid *et al*. 1955). These multi-subunit, filamentous surface structures are associated with adhesion to bacteria, host cells or environmental surfaces (Fronzes *et al*. 2008). Moreover, pili can have an important role in the pathogenesis of bacterial disease, as they are also involved in host-pathogen interactions, host colonization, biofilm formation and cell-to-cell signalling (Kline *et al*. 2009, Proft & Baker 2009,). The pathways involved in pilus assembly are established pathogenicity determinants, as well as potential targets for control of infectious diseases. Type I pili in UPEC contain the FimH adhesin that is necessary for UPEC pathogenesis. FimH also binds to the mannosylated receptors on the bladder epithelium, and mutation in the residues of FimH essential for mannose binding results in decreased virulence in a mouse urinary tract infection model.

There are several types of pilus assembly systems. In Gram-negative bacteria, we distinguish the chaperone/usher (CU)-assembled pili, type IV pili, and curli, among which the chaperone-usher pathway is the best described and contains the most representatives (Nuccio & Baumler 2007, Kline *et al*. 2010). In uropathogenic *E. coli* (UPEC), the CU pili are represented by type 1 pilus and Pap (P) pilus. The components of these pilus sytems are encoded in a gene cluster that contains the genes for the pilin major and minor subunits, the adhesin and the corresponding chaperone and usher proteins (Kline *et al*. 2010). Bacterial genomes typically encode numerous

pilus operons, but these are often poorly expressed under standard laboratory conditions (Humphries *et al*. 2003). Figure 16 shows a model of type 1 pilus chaperone–usher pathway assembly.
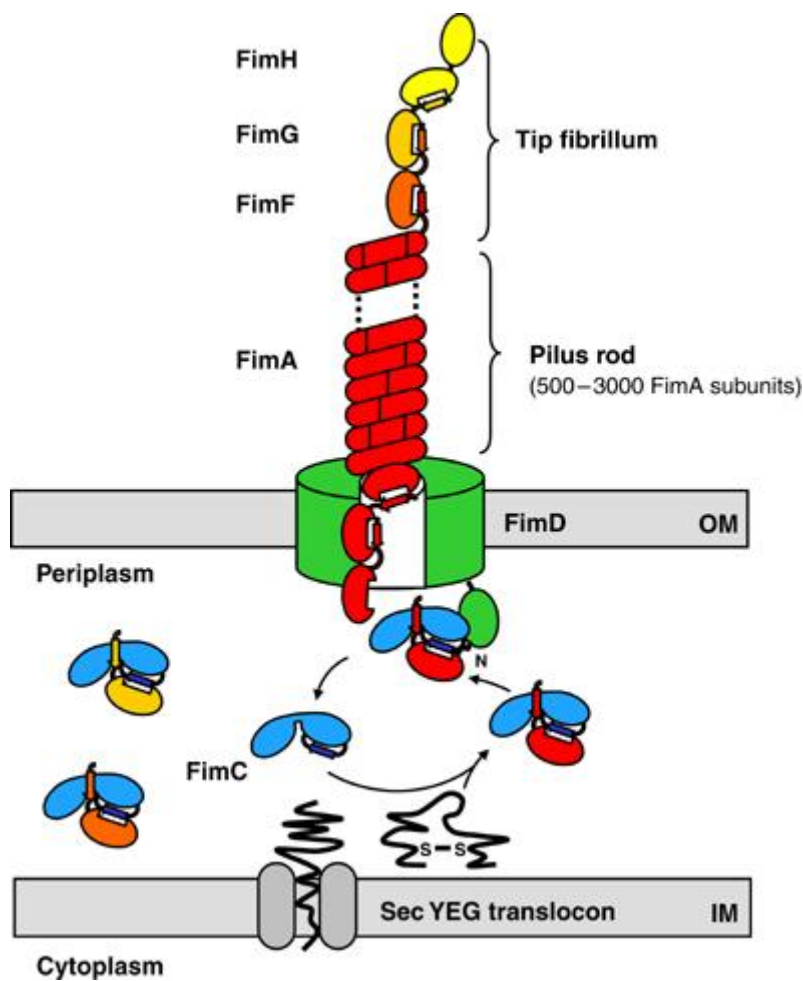
Figure 16. Schematic model of type 1 pilus assembly by the chaperone–usher pathway. Source: (Nishiyama *et al*. 2005). The periplasmic chaperone FimC forms stoichiometric complexes with the newly translocated pilus subunits (FimA, FimG, FimF, FimH). In these complexes, FimC donates its G_1 donor strand to the individual subunits, thereby completing the immunoglobulin-like fold of the subunits. FimC–subunit complexes diffuse to the assembly platform (usher) FimD, which specifically recognizes FimC–subunit complexes via its periplasmic, N-terminal segment of residues 1–139. Subsequently, FimC is released to the periplasm, and the subunit is delivered to the translocation pore of FimD, where it is supposed to interact with the previously incorporated subunit via donor strand exchange. The pilus rod, composed of FimA subunits, assembles into its helical quaternary structure on the cell surface. IM, inner membrane; OM, outer membrane. Source: (Nishiyama *et al*. 2005).

The pilin subunits are translocated across the cytoplasmic membrane to form complexes with the periplasmatic chaperone proteins that ensure their proper folding

(Jones 1993). Each chaperone-subunit complex is then targeted towards the outer membrane usher, which functions as an assembly site and a pore in the outer membrane that allows export of the growing pilus fiber to the cell surface (Thanassi *et al*. 1998), with the most distal protein being secreted the first (Nishiyama *et al*. 2005). The chaperone-pilin complex then dissociates and the subunit is added to the base of the growing pilus (Kline *et al*. 2010). The adhesin mediates attachment of the pilus to a specific receptor.

3.1.2.4 Iron uptake

Iron acts as an essential mineral for many cellular functions including electron transport, ATP production via oxidative phosphorylation, DNA metabolism, protection against oxidative stress and regulation of gene expression, but it is toxic and poorly soluble in its free ferric form (Crosa *et al*. 2004). A single bacterial cell requires from 10^5 to 10^6 ferric ions per generation to maintain the optimal 10^{-6} M internal concentration (Braun & Killmann 1999). In the presence of oxygen, the level of free iron is strictly decreased by formation of insoluble ferric hydroxides, and in mammalian hosts, its availability is further limited by host iron-binding proteins (Crosa *et al*. 2004). In human serum, the concentration of iron is regulated by transferrin to a concentration of $\approx 10^{-24}$ M (Aisen *et al*. 1978, Kretchmar *et al*. 1988). Pathogenic bacteria must be able to compete for this very limited supply of iron with their host organism in order to survive and propagate in the host. Furthermore, the decreased iron availability in host organisms can serve as a stimulus triggering expression of virulence-related genes. It has been shown that increased iron availability correlates with increased virulence of *Escherichia*, *Klebsiella*, *Listeria*, *Neisseria*, *Pasteurella*, *Shigella*, *Salmonella*, *Vibrio*, and *Yersinia*, as reviewed in (Raymond *et al*. 2003). The efficiency of iron assimilation via expression of iron uptake mechanisms is hence an important aspect of bacterial virulence.

Bacteria have evolved several mechanisms to cope with iron scarcity and scavenge iron from the host environment: use of proteases that cleave the iron-binding proteins

87

to obtain free iron, reducing Fe^{3+} to Fe^{2+} followed by the release of iron from a protein complex, and siderophore synthesis (Henderson & Payne 1994). Siderophore production is probably the best studied mechanism of iron acquisition in pathogenic bacteria. Siderophores are high-affinity iron-binding compounds produced by various bacterial species as a response to iron depletion. Their expression is regulated by Fur, a homodimer which binds to a consensus sequence Fur box. Firstly, the ferric siderophore complex binds to the receptor protein on the microbial cell surface, then the complex is translocated across the outer and inner membrane, and finally, iron is released for metabolism inside the cell (Crosa *et al*. 2004). Iron uptake via the catechol siderophore enterobactin (or entrochelin) is the best described among the siderophore-mediated iron uptake systems in prokaryotes. The biosynthesis of aerobactin requires the expression of the operon *entABCDEF* and the transport system for enterobactin is encoded by genes *fepABCD* and *fepG*. In addition, some *Enterobacteriaceae* are able to synthesize a hydroxamate siderophore aerobactin, which has also been linked to increased virulence in members of *Enterobacteriaceae* (Lafont *et al*. 1987, Martinez *et al*. 1994). Aerobactin system was first described on the large *E. coli* plasmid pColIV-K30 by (Warner *et al*. 1981) and the products of the biosynthetic pathway were identified by (de Lorenzo *et al*. 1986). Aerobactin synthesis requires expression of four genes *iucABCD*. Aerobactin is then secreted into the extracellular environment, and iron-aerobactin complex binds to a specific outer membrane TonB-dependent receptor IutA. Interestingly, the aerobactin synthesis genes *iucABCD* and the outer membrane receptor gene *iutA* are located in the same operon, whereas the genes required for trans-membrane transport which encode the periplasmic binding protein FhuB and the inner membrane permease FhuCD form a different genetic cluster that is not related to aerobactin synthesis (Crosa *et al*. 2004).

3.1.2.5 Prophages in bacterial chromosomes

Bacterial genomes typically contain multiple prophages which can constitute a substantial proportion of the total bacterial DNA (Canchaya *et al*. 2003, Casjens

2003). In contrast, sequenced *Archaea* and most intracellular eubacterial pathogens do not contain prophages or their remnants (Canchaya *et al*. 2004). It has been documented that *E.coli* O157:H7 strain Sakai contained 18 prophages making 16% of its total DNA (Canchaya *et al*. 2003). More importantly, prophages can constitute a large part of strain-specific DNA in both pathogenic and non-pathogenic bacteria (Porwollik *et al*. 2002). In a comparison of *Staphylococcus aureus* strains, prophages were the major contributors to variability (Kuroda *et al*. 2001). In *E. coli*, half of the 1-Mb DNA that is variable between *E.coli* 0157:H7 and K-12 consists of prophage DNA (Ohnishi *et al*. 2001).

Phage DNA acts as a vector for horizontal gene transfer between bacteria. Interestingly, prophages often carry additional genes (termed cargo genes, morons, or lysogenic conversion genes) that are not required for the phage life cycle, but may change the phenotype or fitness of the phage (Brussow *et al*. 2004). These cargo genes include a variety of proven or putative virulence factors (Brussow *et al*. 2004, Wagner & Waldor 2002) and even regulatory factors that alter expression of virulence genes that are not encoded by the phage (Spanier & Cleary 1980). Phage genes can encode virulence properties such as bacterial adhesion, colonization, invasion, serum resistance and resistance to phagocytes, exotoxin production, sensitivity to antibiotics and transmissibility among humans.

Table 6 provides a summary of virulence properties that may be altered by prophages. Some cases have been documented when phage tail genes had dual functions and served as adhesion proteins for the attachment of bacteria to its host, such as *Streptococcus mitis* pblAB genes (Bensing *et al*. 2001ab). Prophages can also mediate progression of many bacterial infectious diseases and influence expression of virulence factors (Breitbart & Rohwer 2005). In addition to this, prophages can contribute to the rearrangement of bacterial chromosomes by providing targets for homologous recombination between bacterial genomes that harbour related prophages (Brussow & Hendrix 2002, Nakagawa *et al*. 2003). In an extensive review on phages and evolution of bacterial pathogens by Brussow *et al*. (2004), the authors summarized that apart from introduction of new genes via lysogenic conversion or

transduction, phages can also modulate bacterial fitness by providing anchor points for genome rearrangements, by gene disruption, via protection of the bacterium from lytic infection, and by lysis of competing strains via prophage induction.

Table 6. Bacterial virulence properties altered by bacteriophages. Reproduced from (Wagner & Waldor 2002).

Organism	Mechanism	Source
	Colonization/adhesion	
Escherichia coli	The λ-encoded *lom* gene promotes adhesion to buccal epithelial cells.	(Reeve & Shaw 1979, Barondess & Beckwith 1990)
Streptococcus mitis	The SM1-encoded PblA and PblB surface proteins promote adhesion to platelets.	(Bensing *et al*. 2001b)
Vibrio cholerae	The toxin-coregulated pilus may be phage encoded.	(Karaolis *et al*. 1999)
	Invasion	
Salmonella enterica	Phage SopEφ transduces a type III secretion system effector that promotes entry into epithelial cells.	(Mirold *et al*. 1999)
	Phage Gifsy-1 encodes *gipA*, a gene that enhances survival in the Peyer's patch.	(Stanley 2000)
Streptococcus pyogenes	Hyaluronidase is phage encoded.	(Hynes & Ferretti 1989)
Staphylococcus aureus	Fibrinolysin is phage encoded.	(Sako *et al*. 1983)
	Resistance to serum/phagocytes	
Escherichia coli	The λ-encoded *bor* gene confers a survival advantage in animal serum.	(Barondess & Beckwith 1990)
Pseudomonas aeruginosa	Phages encode enzymes that alter the O antigen.	(Holloway & Cooper 1962)
Salmonella enterica	Phage Gifsy-2 encodes SodC, a superoxide dismutase.	(Figueroa-Bossi & Bossi 1999)
	Phages encode enzymes that alter the O antigen.	(Robbins & Uchida 1962, Wright 1971)
Shigella dysenteriae	Phages encode enzymes that alter the O antigen.	(Mavris *et al*. 1997, Guan *et al*.

		1999)
Staphylococcus aureus	Phages encode CHIPS, a phagocytotoxin.	Van Wamel *et al.*[a]
	Phage φPVL encodes the Panton-Valentine leukocidin.	(Kaneko *et al.* 1997)
Streptococcus pyogenes	Lysogeny up-regulates the antiphagocytic M protein.	(Spanier & Cleary 1980)

Exotoxin production		
Bordetella avium	Pertussis toxin is phage encoded in *B. avium.*	van Horne *et al.*[b]
Clostridium botulinum	Botulinum toxin is phage encoded.	(Wagner & Waldor 2002)
Corynebacterium diphtheriae	Diphtheria toxin is phage encoded.	(Holmes & Barksdale, 1969, Uchida *et al.* 1971)
Escherichia coli	The Shiga toxins are phage encoded.	(Newland *et al.* 1985, Willshaw *et al.* 1985, Huang *et al.* 1986)
Pseudomonas aeruginosa	Pseudomonas cytotoxins are phage encoded.	(Hardt *et al.* 1998)
Shigella dysenteriae	The Shiga toxin genes are associated with phage sequences, probably a defective prophage.	(McDonough, Butterton 1999)
Staphylococcus aureus	Staphylococcal enterotoxins are phage encoded.	(Casman 1965, Betley & Mekalanos 1985, Coleman *et al.* 1989)
	Staphylococcal exfoliative toxins are phage encoded.	(Yamaguchi *et al.* 2000)
	Toxic shock syndrome toxin is encoded by SapI, a mobile pathogenicity island transduced at high frequency by phage 80α.	(Lindsay *et al.* 1998)
Streptococcus pyogenes	Streptococcal pyrogenic (erythrogenic, scarlatinal) exotoxins are phage encoded.	(Johnson & Schlievert 1984, Weeks & Ferretti 1984, Goshorn & Schlievert 1989)
Vibrio cholerae	Cholera toxin is phage encoded.	(Waldor & Mekalanos 1996)

Susceptibility to antibiotics		
Staphylococcus aureus	Generalized transduction contributes to horizontal transmission of Gram-positive antibiotic-resistance genes.	(Fujii *et al.* 1988, Fischetti *et al.* 2000)

Transmission		
V. cholerae	Phage-encoded cholera toxin likely promotes transmission by stimulating copious amounts of watery diarrhea.	(Waldor & Mekalanos 1996)

[a] Van Wamel *et al*., Abstr. 101st Gen. Meet. Am. Soc. Microbiol. 2001.

[b] S. J. van Horne, D. Bjornsen, P. Carpentier, and L. M. Temple, Abstr. 101st Gen. Meet. Am. Soc. Microbiol. 2001, abstr. B-109, p. 64-65, 2001.

3.2. Targeted gene disruption using λ-red technique

Creating gene disruptions via targeted gene deletion can be achieved using a variety of techniques. Generally, linear or plasmid DNA carrying an antibiotic resistance marker flanked by regions of homology to the target gene, or just outside of it, can be transformed into bacterial cells. Thereafter a recombination event between the homologous regions is induced. The cells are then screened for acquired antibiotic resistance and the gene disruption is verified in the cells with the correct phenotype by PCR and sequencing of the target region.

There are a range of experimental techniques exploiting this strategy. The first protocols that used linear DNA for gene replacement (Jasin & Schimmel 1984, Winans *et al*. 1985) were limited to strains deficient in the RecBCD nuclease system, which would otherwise degrade the linear double stranded phage DNA (Murphy 1998). Other methods used the integration of a plasmid into the bacterial chromosome and subsequent resolution of the integrated complex, which regenerated either the wild type locus or the desired gene replacement. One of the limitations of such protocols was that plasmids that do not replicate under the conditions used for selection of targeted mutants had to be used, such as temperature-sensitive plasmids or phagemid-based vectors, so that plasmids carrying the marker were ultimately lost due to cell division. According to Murphy (1998), other disadvantages of this strategy included the low frequency of the resolution of the cointegrate, the necessity to clone the gene before the replacement and unintended gene replacements generated by the replacement allele.

The first gene replacement protocol that used bacteriophage λ recombination genes known as "Red" or "λ-Red" was developed by Murphy (1998). The λ-Red system uses the activity of λ bacteriophage genes *exo, bet* and *gam.* Assisted by the host RecA protein, the (5′-3′) exonuclease Exo binds to dsDNA and generates single stranded DNA overhangs, and the Bet protein induces recombination between these overhangs and homologous regions of the chromosome. Gam inhibits the host RecBCD exonuclease that would cause digestion of phage DNA. λ-Red recombination is stimulated by the presence of dsDNA ends, generated by a restriction endonuclease (Stahl *et al.* 1985). Datsenko and Wanner developed an improved λ-Red gene replacement method that is now widely used for generating targeted deletions in *E. coli* and in other bacteria (Datsenko, Wanner 2000). In this method, the desired chromosomal sequence is replaced by a PCR product with ~ 26-nt homology at each end to target sequences flanking the gene to be replaced. Gene replacement occurs via λ-Red recombination, mediated by the pKD46 plasmid expressing the λ-Red system under the control of an arabinose promoter. pKD46 has a temperature-sensitive replicon and it can be subsequently eliminated by growth at 37°C. However, external supervisor Dr. Jon Hobman (University of Nottingham) working in collaboration with the University of Birmingham reported limited success with the Datsenko and Wanner method in pathogenic *E. coli* strains; O157:H7 Sakai strain, CFT073 UPEC, O42 EAEC and ETEC strains (Lee *et al*. 2009). For that reason, the team developed an improved protocol for creating gene disruptions and for coupling of genes to epitope tags (Lee *et al*. 2009). The protocol was based on the "Gene gorging" method developed by Herring *et al*. (2003) which uses two plasmids. The recombineering plasmid pACBSR includes the λ-Red and I-SceI endonuclease genes which are controlled by an arabinose-inducible *araBAD* promoter. The donor plasmid (pDOC) carries an antibiotic resistance cassette, or epitope tag flanked by regions of homology to the target gene and the recognition site for the I-Sce-I endonuclease. In the presence of arabinose, the donor plasmid is cleaved by I-SceI and a linear dsDNA target for the lambda-Red system is generated. Hobman and colleagues designed a set of donor plasmids which include a kanamycin resistance cassette and the *sacB* gene, pDOC-K (Figure 17). *SacB* encodes a levansucrase, which converts sucrose to levans and has been shown to be lethal in *E. coli* and other Gram-

negative bacteria (Gay & Kado 1985). Plasmids expressing kanamycin resitance and *sacB* can be therefore used for the counterselection of true recombinants on media containing kanamycin and sucrose. An I-SceI recognition site was added to the recombineering plasmid to make pACBSCE. This means that this plasmid is cleaved after induction of expression of I-SceI and the λ-Red genes, and hence the exposure of cells to the λ-Red proteins, which might lead to undesirable chromosomal rearrangements, is limited (Lee *et al*. 2009). This method was successfully used to generate deletion mutants in enterohaemorrhagic O157:H7 (EHEC), uropathogenic CFT073 (UPEC), enteroaggregative O42 (EAEC) and enterotoxigenic H10407 (ETEC) *E. coli* strains as well as in K-12 laboratory strains. The gene doctoring method is outlined in Figure 18 reproduced from (Lee *et al*. 2009).
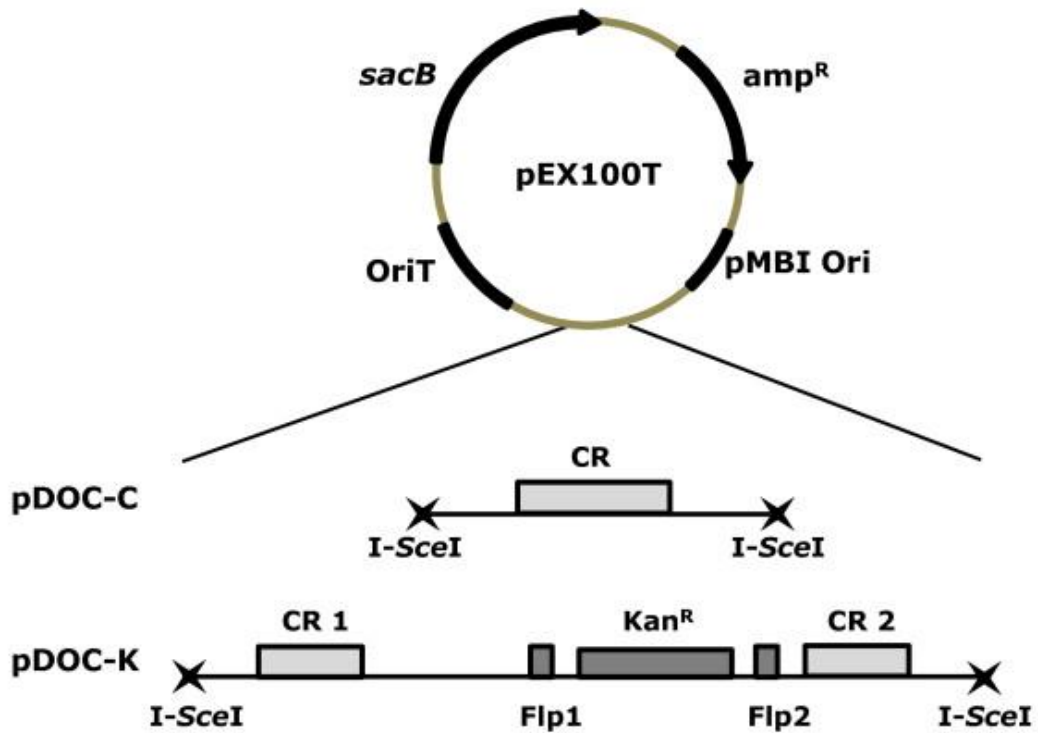
Figure 17. The pDOC donor plasmids. Taken from Lee *et al*. (2009). Circular representation of the pEX100T plasmid showing the location of the origins of replication, the *sacB* gene and the ampicillin resistance gene. Below is a linear representation of the pDOC plasmid inserts, showing the I-SceI restriction sites, cloning regions (CR, CR1 and CR2) and the Flp recognition sites flanking the kanamycin resistance cassette (Kan^R).
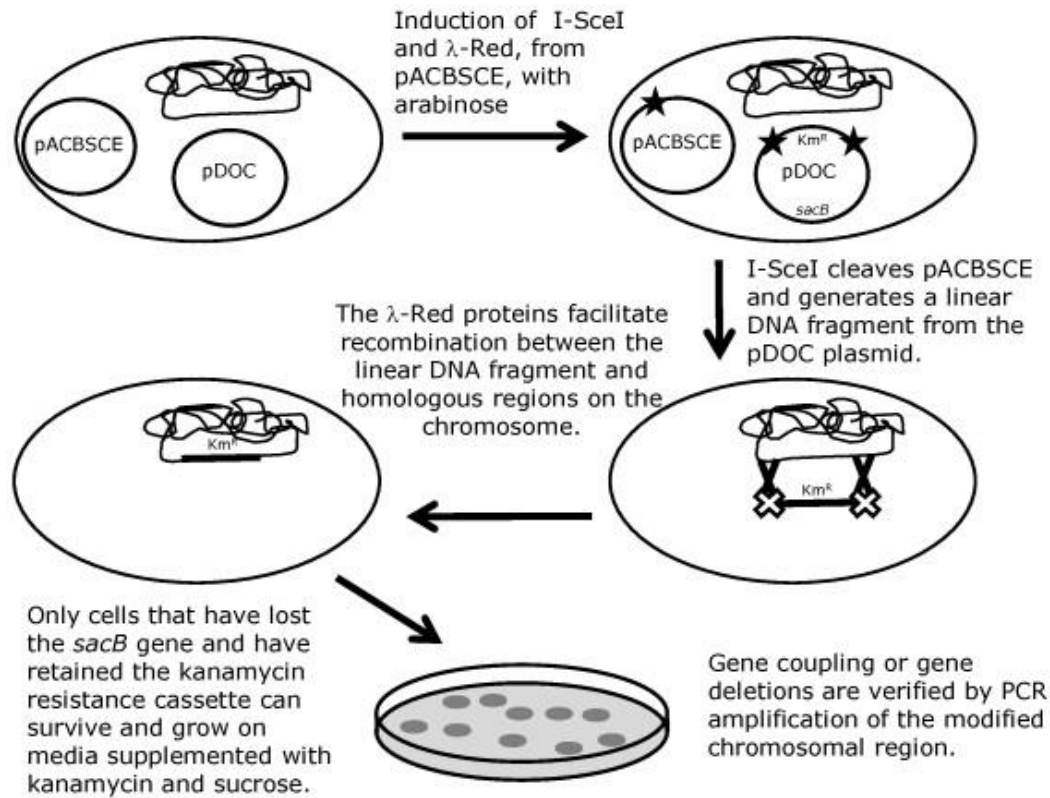
Figure 18. Gene doctoring protocol. Taken from (Lee *et al*. 2009). The pDOC donor plasmid and the recombineering plasmid pACBSCE are co-transformed into the recipient strain. Arabinose induction promotes expression of the λ-Red gene products and I-SceI. I-SceI generates a linear DNA fragment from the donor plasmid that is a substrate for recombination with the chromosome mediated by the λ-Red system. In addition, I-SceI also cleaves the pACBSCE recombineering plasmid, resulting in loss of plasmid and loss of λ-Red expression. Recombination occurs between homologous regions on the linear DNA substrate and the chromosome, transferring the kanamycin cassette, onto the chromosome. Recombinants are selected by the ability to survive and grow on LB supplemented with kanamycin and sucrose. Only true recombinants, which have lost the *sacB* gene due to donor plasmid loss and have retained the kanamycin cassette due to recombination, are able to survive and grow on this medium. Confirmation that the correct chromosomal modification has been generated, is achieved by amplifying the target region by PCR using primers that anneal adjacent to this region and electrophoresing the products on a 0.7% agarose gel. Size differences in PCR product between the wild-type and modified strain confirm whether a chromosomal modification has been achieved.

3.3. Aims

The taxonomic reclassification of *Enterobacter sakazakii* into six separate species accepted in 2008 (Iversen *et al*. 2008) confirmed the immense phenotypical diversity of the *Cronobacter* genus that was observed in many earlier studies. To date, only strains from *C. sakazakii*, *C. malonaticus* and *C. turicensis* have been associated with neonatal infections, which suggests that there are considerable differences between the virulence of *Cronobacter* species. Also, most of the studies published before 2008 have been in fact done on "*Enterobacter sakazakii*" strains, without further assigning the *Cronobacter* species according to the new taxonomy. It is probable that most of the strains studied before were in fact *C. sakazakii*, as it is the most commonly isolated *Cronobacter* species. Although earlier studies have considerabely widened our knowledge about *Cronobacter* pathogenicity, the mechanisms by which *Cronobacter* causes the disease are still largely unclear, and few virulence factors have been identified to date. Due to the severity of infant infection, a better understanding of the genomic variation between *Cronobacter* spp. is needed, and will be of interest to manufacturers of powdered infant formula, regulatory bodies, as well as those studying the evolution and diversity of pathogenicity.

The sequencing of the first strain of *Cronobacter, C. sakazakii* BAA-894 and depositing of its sequence to GenBank by my external supervisor McClelland at Washington University in September 2007 marked a new era in the research of this opportunistic pathogen. For the first time, whole genome comparative approaches such as use of DNA microarrays and whole-genome alignments were possible. The aim of my study is to use the first complete genome sequence of *C. sakazakii* strain BAA-894 isolated from powdered formula associated with a NICU outbreak (Himelright *et al*. 2002) for comparative genomic hybridization (CGH) analysis of physiological and virulence related traits across the *Cronobacter* genus, via the first *Cronobacter* whole-genome microarray. The microarray will be designed and produced by NimbleGen and a comparative genome hybridization assay will be

performed with different *Cronobacter* strains with known history and virulence potential from our internal collection representing *C. sakazakii* strains, *C. malonaticus, C. turicensis, C. dublinensis* and *C. muytjensii,* against the sequenced genome strain *C. sakazakii* BAA-89. Strains will be chosen according to their association with virulence (ie. clinical source and previous tissue culture assays) and 16S rDNA cluster grouping. The data obtained will be analysed to describe the regions of genes whose distribution is variable between the *Cronobacter* strains, describe the gene clusters with importance for *Cronobacter*'s pathogenesis, evaluate the distribution of the organism's known virulence determinants, as well as to discover new virulence factors that have not been previously identified by combining the CGH data with the strains' virulence data.

3.4. Materials and methods

3.4.1 *Cronobacter* DNA microarray and comparative genomic hybridisation

3.4.1.1 Bacterial strains

Table 7. *Cronobacter* strains used in the microarray study.

Strain	Comment	Source	16S rDNA[1]	References
C. sakazakii BAA-894	Genome sequenced strain	Powdered formula	-	(Himelright *et al.* 2002, Baldwin *et al.* 2009)
C. sakazakii ATCC 29544[T]	Species type strain	Child's throat	0.38	(Iversen *et al.* 2008, Baldwin *et al.* 2009)
C. sakazakii ATCC 12868		Unknown	0.57	(Baldwin *et al.* 2009)
C. sakazakii strain 20		Clinical	0.57	(Baldwin *et al.* 2009)
C. sakazakii strain 701	Fatal NEC III case	Peritoneal fluid	0.47	(Caubilla-Barron *et al.* 2007, Baldwin *et al.* 2009)
C. sakazakii strain 767	Fatal meningitis case	Trachea	0.47	(Caubilla-Barron *et al.* 2007, Baldwin *et al.* 2009)
C. sakazakii strain 696	NEC II case	Stools	0.57	(Caubilla-Barron *et al.* 2007, Baldwin *et al.* 2009)
C. malonaticus LMG 23826[T]	Species type strain	Breast abscess isolate	0.57	(Iversen *et al.* 2008)
C. turicensis LMG 23827[T]	Species type strain. Meningitis case	Blood isolate	1.8	(Iversen *et al.* 2008)
C. muytjensiii ATCC 51329[T]	Species type strain	Unknown	2.94	(Iversen *et al.* 2008)
C. dublinensis LMG 23823[T]	Species type strain	Milk powder plant	3.69	(Iversen *et al.* 2008)

[1] % 16S rDNA sequence difference from *C. sakazakii* BAA-894

3.4.1.2 Genomic DNA extraction

Total genomic DNA of 10 selected *Cronobacter* strains (Table 7) and the reference sequenced strain *C. sakazakii* BAA-894 (NC_009778) was isolated and purified.

Three different isolation protocols (SIGMA GenElute, SIGMA GenElute with additional purification and QIAGEN Genomic-tip 100/G) were tried. Finally, a modified QIAGEN Genomic-tip 100/G protocol was used to obtain the required DNA purity (A260/A280 ≥ 1.8 and A260/A230 ≥ 1.9). The QIAGEN protocol was modified to include an extended cell lysis (1h), extended deproteinization and an additional wash of the precipitated DNA (see protocol below).

- A single colony of *Cronobacter* from a fresh TSA plate was inoculated into 9 ml of LB broth and incubated overnight at 37°C.

- 3 ml of overnight culture were pelleted by centrifugation at 5000 g for 10 min and the supernatant was discarded.

- The bacteria were resuspended in 3.5 ml of buffer B1 by thorough vortexing at top speed.

- 80 µl of lysozyme stock solution (100 mg/ml), and 100 µl of Proteinase K stock solution were added to the bacterial suspension and incubated at 37°C for 1 h.

- 1.2 ml of Buffer B2 were added to the lysate and mixed by inverting the tube several times. The lysate was incubated at 50°C for 1 h.

- The Genomic-tip was placed over a tube using a tip holder and equilibrated with 4 ml of Buffer QBT.

- The bacterial lysate was vortexed for 10 s at maximum speed, applied to the equilibrated Genomic-tip and allowed to empty by gravity flow.

- The Genomic-tip was washed with 3 x 7.5 ml of Buffer QC, allowing the Buffer QC to move through the Genomic-tip by gravity flow.

- The genomic DNA was eluted with 5 ml of Buffer QF prewarmed to 50°C, allowing the Genomic tip to drain by gravity flow.

- The DNA was precipitated by adding 3.5 ml of isopropanol to the eluate. The sample was mixed by inverting the tube 10 times and centrifuged immediately

at >5000 x g for 15 min. The supernatant was carefully discarded and the pellets air dried for 10 min.

- The DNA was immediately transferred to a microcentrifuge tube containing 0.2 ml of DNase free water and dissolved at 55°C for 1 – 2.5 h. The samples were checked every 30 min to ensure that the shortest possible resuspension time was used to avoid degradation of the DNA.

The DNA samples were checked for fragmentation on 0.8% agarose gels electrophoresis (3.4.2.4) and checked for protein and RNA content by spectrophotometry at 280 and 230 nm, respectively, using NanoDrop (Fisher Scientific). The DNA concentration was determined by spectrophotometry at 260 nm.

3.4.1.3 Plasmid DNA extraction

The plasmid DNA was extracted according to the method described in Kado & Liu (1981). The composition of the solutions used and the extraction procedure is described below.

E buffer (for 1l)

- 4.85 g of Tris base (121.14 $g.mol^{-1}$) were resuspended in 1 l of distilled water to obtain a 40 mM solution.
- 0.74 g of Na_2EDTA (372.24 $g.mol^{-1}$) were added to the Tris base solution to obtain 2mM Na_2EDTA
- The pH was adjusted to 7.9 with glacial acetic acid

Lysis buffer (for 100 ml)

- 0.61 g of Tris base (121.14 g.mol^{-1}) were dissolved in 40 ml of distilled water and 30 ml of 10% SDS (w/v) were added to the solution.

- 8 g of NaOH (40 g.mol^{-1}) were dissolved in 100 ml of distilled water to obtain 2M NaOH solution using a 100 ml volumetric flask for accuracy. The solution could be stored for a maximum of 3 weeks.

- To bring the SDS/Tris solution to pH 12.6, 3.2 ml of the 2M NaOH solution were added and the mixture was brought to 100 ml using a volumetric flask. This solution could only be used on the day of preparation.

Plasmid DNA extraction protocol

- One colony of a *Cronobacter* strain was inoculated into 5 ml of LB broth and incubated overnight at 37°C in a shaking incubator.

- 3 ml of the overnight culture were centrifuged for 3 min at 7,000 g in a conventional centrifuge and the resulting pellet was resuspended in 50 μl of E buffer.

- 100 μl of the lysis buffer were added and the tube was gently mixed to observe lysis.

- The sample was heated at 55 ^{o}C for 60 min in a heatblock.

- 150 μl of phenol/chloroform were added to the sample and gently mixed.

- The sample was centrifuged for 30 min at 12,000 g to separate phases

- 45 μl of the top phase containing plasmid DNA were analysed by gel electrophoresis (3.4.2.4), which was run in 0.7% E buffer gel at 90 V for 1.5 h.

For samples where no bands of plasmid DNA were visible on the gel, the plasmid DNA extraction was repeated with the QIAfilter Plasmid Midi Kit, using the protocol available at http://www.qiagen.com.

3.4.1.4 *C. sakazakii* ATCC BAA-894 microarray design

A 384,030 probe oligonucleotide tiling DNA microarray which comprised the complete genomic sequence of *C. sakazakii* ATCC BAA-894 was designed for us by Fred Long and Michael McClelland (based at the Vaccine Research Institute of San Diego, San Diego, California, US). Probes were designed at an average of less than 12 base spacing on alternating strands, leading to an average of over 100 50-mer oligonucleotide probes per annotated gene. Probes were designed using the protocol below (provided by Prof. McClelland).

Every possible 50-base probe from both strands of the *C. sakazakii* genome and two plasmids pESA2 and pESA3 was tested for the ability to be manufactured by NimbleGen. Those that required too many NimbleGen cycles were shortened. Resulting candidate probes that were less than 35 bases long were omitted. The remaining 9,061,350 potential probes had an average melting temperature (Tm) of 74°C. Probes that had Tm above the average were shortened to a minimum of 35 bases. The probes were selected from the pool of 9,061,350 potential probes by selecting the best probe at 11.375-base increments, alternating between strands each time. The resulting 386,802 candidate probe sequences were analyzed to remove any probes that covered the same region as another probe due to duplications in the genome and the resulting 384,030 unique sequences were chosen for the array. Mappings between the unique probe sequences and their genome/plasmid/gene positions were stored in a separate file.

3.4.1.5 Comparative Genomic Hybridization (CGH)

Comparative hybridisation (CGH) was performed by Roche NimbleGen Inc. (http://www.nimblegen.com/) according to the method described at http://www.nimblegen.com/products/lit/cgh_userguide_v6p0.pdf. Briefly, total genomic DNA of the reference strain *C. sakazakii* BAA-894 and the test strains isolated as described in 3.4.1.2 was sheared into shorter fragments. Equal quantities of DNA fragments of each strain were labelled with a fluorescent dye; the reference DNA was labelled with Cy5 and the test sample DNA was labelled with Cy3. After the hybridization reaction using probes designed as described in 3.4.1.4 the two colour arrays were scanned using the MS 200 Microarray Scanner and the MS 200 Data Collection Software; the fluorescent intensities from both channels were calculated using the NimbleScan software.

3.4.1.6 Data normalization

Data normalization was performed by Roche NimbleGen Inc. (http://www.nimblegen.com/) according to the method described at http://www.nimblegen.com/products/lit/cgh_userguide_v6p0.pdf.

DNA from the sequenced strain, *C. sakazakii* BAA-894, was used as the internal array control to minimize any background effects. For within-array normalization, a LOWESS method (Cleveland 1979) was used as spatial correction and QSPline (Workman *et al*. 2002) was used to correct for dye bias. The raw data was deposited in GenBank GEO (accession number GSE19308).

3.4.1.7 Examples of data

Hybridization and design data were delivered in form of i) Design file; ii) Positioning file; and iii) Intensity files.

i) Design file (.ndf) is a tab-delimited text file that describes the placement of the

probes on the microarray slide in a format similar to a probe file. The unique identifier of the probes is PROBE_ID, except quality control probes, which have non-unique PROBE_ID (Figure 19).

PROBE_DESIGN_ID	CONTAINER	SEQ_ID	PROBE_SEQUENCE	MISMATCH	MATCH_INDEX	FEATURE_ID	ROW_NUM	COL_NUM	PROBE_ID	POSITION	DESIGN_ID	X	Y
7736_0350_070	BLOCK1	ESA	caccacgatgat	0	64232058	64232058	706	350	ESA_072472	446866	7736	350	706
7736_0352_070	BLOCK1	ESA	tcgcgctcgctc	0	64493314	64493314	706	352	ESA_333728	980521	7736	352	706
7736_0354_070	BLOCK1	ESA	tttacggcgcgg	0	64538519	64538519	706	354	ESA_378933	1920543	7736	354	706
7736_0356_070	BLOCK1	ESA	cgccttgttttttta	0	64302241	64302241	706	356	ESA_142655	462205	7736	356	706
7736_0358_070	BLOCK1	ESA	gcgacctcaaa	0	64389814	64389814	706	358	ESA_230228	1192348	7736	358	706
7736_0360_070	BLOCK1	ESA	gggctcgatatg	0	64451459	64451459	706	360	ESA_291873	929927	7736	360	706
7736_0362_070	BLOCK1	ESA	cgagccgctgc	0	64286812	64286812	706	362	ESA_127226	3047243	7736	362	706
7736_0364_070	BLOCK1	ESA	gcgccggcgtc	0	64395775	64395775	706	364	ESA_236189	1751212	7736	364	706
7736_0368_070	BLOCK1	ESA	gaaaggcgcgc	0	64349354	64349354	706	368	ESA_189768	1516249	7736	368	706
7736_0370_070	BLOCK1	ESA	cgacacgccaa	0	64283237	64283237	706	370	ESA_123651	790783	7736	370	706

Figure 19. Example of a Design (ndf.) file

ii) The positions (.pos) file is a tab-delimited text file that maps each probe to its position in the reference genome. Positions are keyed on PROBE_ID (Figure 20).

PROBE_ID	SEQ_ID	CHROMOSOME	POSITION	COUNT	LENGTH
ESA_072472	ESA	ESA	446866	1	50
ESA_333728	ESA	ESA	980521	1	42
ESA_378933	ESA	ESA	1920543	1	35
ESA_142655	ESA	ESA	462205	1	50
ESA_230228	ESA	ESA	1192348	1	50
ESA_291873	ESA	ESA	929927	1	39
ESA_127226	ESA	ESA	3047243	1	47
ESA_236189	ESA	ESA	1751212	1	42
ESA_189768	ESA	ESA	1516249	1	35
ESA_123651	ESA	ESA	790783	1	50
ESA_140122	ESA	ESA	3462434	1	50

Figure 20. Example of a Positioning file

iii) A pair report (.pair) provides sequence, probe, and signal intensity information for one data channel. Two pair reports were supplied for each hybridization experiment,

105

one for the test (Cy3) channel and one for the reference (Cy5) channel. Column 'PM' shows the normalised fluorescent intensity values from each hybridization experiment (Figure 21).

IMAGE_ID	GENE_EXPR_OPTION	SEQ_ID	PROBE_ID	POSITION	X	Y	MATCH_INDEX	PM	MM
18789102_532	BLOCK1	AB265687	extra_001667	123	448	508	64543775	205	0
18789102_532	BLOCK1	AB265687	extra_001952	104	33	1007	64544060	4950	0
18789102_532	BLOCK1	AF056287	extra_000166	187	741	973	64542274	1622	0
18789102_532	BLOCK1	AF056287	extra_000712	218	526	548	64542820	177	0
18789102_532	BLOCK1	AF056287	extra_001382	202	165	11	64543490	192	0
18789102_532	BLOCK1	AJ515699	extra_000198	1256	195	969	64542306	168	0
18789102_532	BLOCK1	AJ515699	extra_000206	1365	327	805	64542314	179	0
18789102_532	BLOCK1	AJ515699	extra_000288	1527	294	580	64542396	188	0
18789102_532	BLOCK1	AJ515699	extra_000366	702	718	890	64542474	204	0
18789102_532	BLOCK1	AJ515699	extra_000539	2963	485	145	64542647	211	0
18789102_532	BLOCK1	AJ515699	extra_000556	1500	660	556	64542664	197	0

Figure 21. Example of Intensity (.pair) file

After a series of problems with data delivery from NimbleGen, normalized data were analyzed using the NimbleGen software NimbleScan and SignalMap. These did not contain suitable tools for bacterial CGH data analysis, however. The delivered data were also missing a correct annotation file crucial for data analysis. Prof. McClelland (external advisor and author of *Cronobacter* sequencing project) supplied the missing annotation file.

3.4.1.8 Calculation of median $\log_2$ intensities and data visualization by WebArrayDB

For the following steps of CGH data analysis, the data were transformed into $\log_2$ (test channel intensity / reference channel intensity). The probes corresponding to the intergenic regions of the reference genome were filtered out and the intensity files and

106

the positioning files from NimbleGen were transformed to the file formats suitable for WebArrayDB (http://www.webarraydb.org/webarray/index.html) using SPSS software. Median $\log_2$ intensity ratios for each *C. sakazakii* BAA-894 gene based on the $\log_2$ ratios of separate probes were calculated by the new CGH tool available on WebArrayDB.

3.4.1.9 Dynamic cut-off determination by GACK and gene category assignment

The GACK algorithm developed by Charles Kim (Kim *et al*. 2002) sets a floating set of thresholds (cut-offs) for each hybridization set based on computational analysis of the $\log_2$ ratio data distribution. Briefly, the GACK algorithm finds the location and height of the main peak of the $\log_2$ ratio distribution and determines the coordinates of the two "side points", which are points at half of the peaks maximum height (Figure 22). Then GACK calculates the normal probability density function, which is an estimate of the data distribution if all the genes analysed were present, and which is a function of the height of the main peak and the coordinates of the two side points. Finally, the 'Estimated Probability of Presence' (EPP) function is calculated, which is simply the normal probability density function divided by the real $\log_2$ ratio distribution. Note that only the left portion of the EPP curve is considered for the cut-off determination (Figure 22).

107
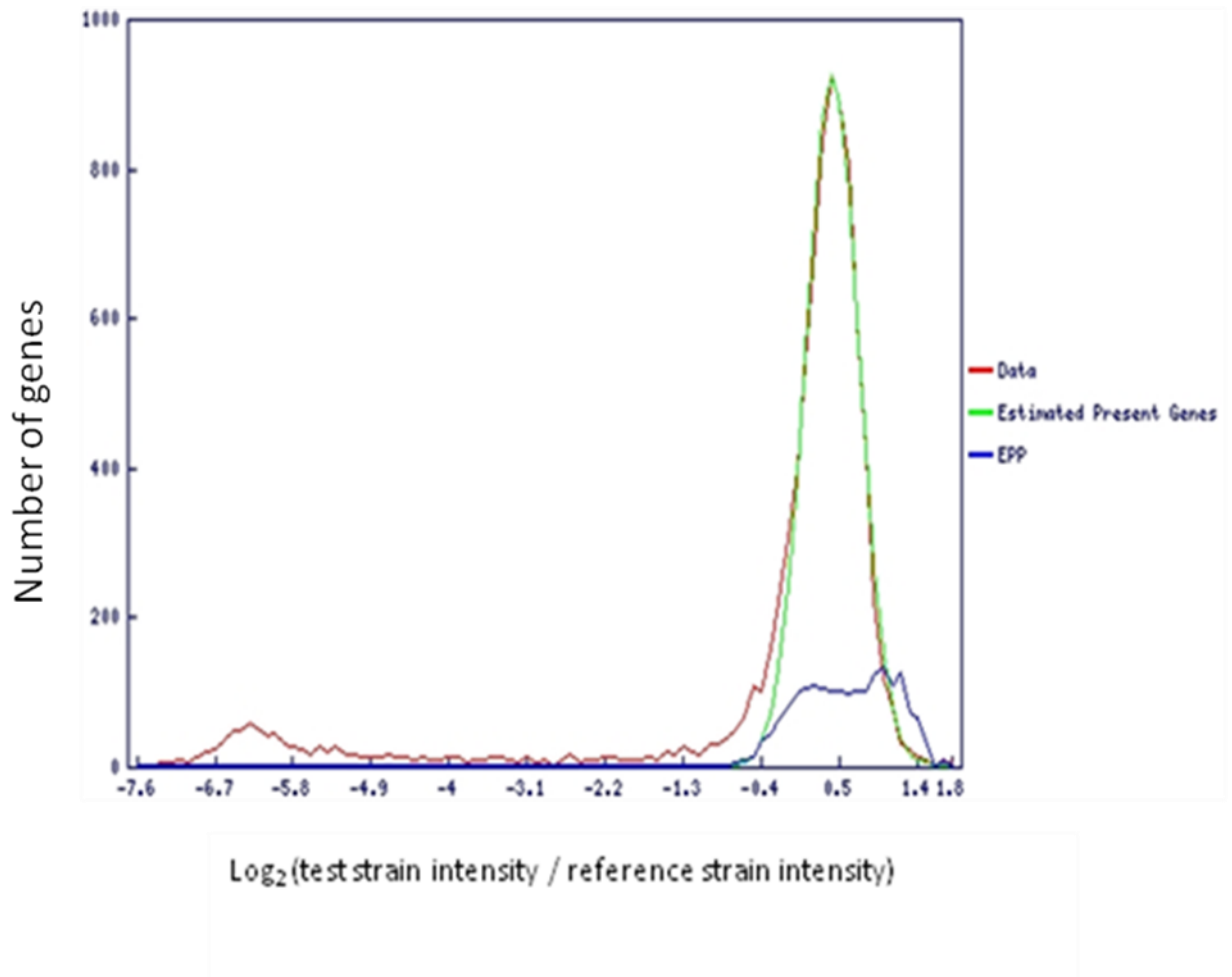
Figure 22. GACK analysis. Hybridization data for *Cronobacter sakazakii* type strain ATCC 29544. Red: histogram of $\log_2$ ratio intensity medians for 4,381 *Cronobacter sakazakii* genes; green: estimated present genes; blue: Estimated Probability of Presence (EPP) as determined by GACK.

The median $\log_2$ ratio data calculated by WebArrayDB CGH plotter were transformed to the data format supported by GACK (Figure 23).

UNIQID	NAME	GWEIGHT	EXP1
ESA_02139	ESA_02139	1	.63
ESA_02140	ESA_02140	1	.49
ESA_02141	ESA_02141	1	-.22
ESA_02142	ESA_02142	1	-4.62
ESA_02143	ESA_02143	1	-5.24
ESA_02144	ESA_02144	1	-5.01
ESA_02145	ESA_02145	1	.27
ESA_02146	ESA_02146	1	.38
ESA_02147	ESA_02147	1	.07

Figure 23. Input data for GACK analysis

Each *Cronobacter* gene was classified according to the most stringent settings of the trinary output of GACK as present (1), intermediate (0) or absent (-1). Genes classified as 'present' had a $\log_2$ intensity ratio (test/reference) greater than the cut-off value corresponding to the 100% estimated probability of presence (EPP) as calculated by GACK. The 'absent' genes had a $\log_2$ ratio equal or inferior to the cut-off value corresponding to 0% EPP. This category includes genes that are absent from the test strain, but can also include genes with a degree of sequence divergence. The 'intermediate' category includes genes whose status could not be assigned with certainty and can include genes with some degree of sequence divergence. The EPP function and cut-offs were determined separately for hybridization data for each strain.

3.4.1.10   Whole genome clustering

Clustering analysis was performed using Cluster (Eisen *et al*. 1998). Hierarchical clustering was performed using the average linkage method on the trinary matrix based on the CGH analysis (1 for present, 0 for uncertain and −1 for absent/divergent gene). The tree was visualised using the Treeview software (Eisen *et al*. 1998).

3.4.1.11　　　Prophage Identification

ACLAME was used for prophage identification (Lima-Mendez *et al*. 2008). This was accessed at http://aclame.ulb.ac.be/Tools/Prophinder/. The Prophinder algorithm uses BLASTP to identify coding sequences (CDS) that are similar to phage proteins stored in the ACLAME database (http://aclame.ulb.ac.be/). Prophinder then identifies regions with a significantly high density of phage-like proteins and compares the number of the matching CDS in these regions to the number of CDS that could be expected by chance only. Using the ACLAME annotations and biological information that is integrated in the database, Prophinder selects the phage-like regions best matching potential prophages.

3.4.1.12　　　Identification of genomic islands

Genomic islands in *C. sakazakii* BAA-894 were identified by IslandViewer (Langille & Brinkman 2009). This software tool combines several methods for gene island prediction: IslandPick (Langille *et al*. 2008), SIGI-HMM (Waack *et al*. 2006) and IslandPath (Hsiao *et al*. 2003).

3.4.2 Construction of *Cronobacter* mutants via "Gene Doctoring" recombineering method

3.4.2.1 Plasmids used in this method

Plasmids pDOC, pDOC-K, pACBSR and pACBSRSceI (Table 8) were kindly provided by Dr. Jon Hobman (University of Nottingham, UK).

Table 8. Plasmids used in the gene doctoring method.

Name	Size (kb)	Genes carried	Resistance marker
pDOC	5.8	*sacB*	*AmpR*[1]
pDOC-K	5.8	*sacB*	*AmpR, KanR*[2]
pACBSR	7.3	λ- Red system: *exo, bet, gam*	*CmR*[3]
pACBSRSceIV	7.3	λ- Red system: exo, *bet, gam,* Sce-I site	*CmR*[3]

[1] ampicillin marker for selection on media containing 200 µg/ml ampicillin

[2] kanamycin marker for selection on media containing 50 µg/ml kanamycin

[3] chloramphenicol marker for selection on media containing 35 µg/ml chloramphenicol

3.4.2.2 Plasmid extraction

Extraction of plasmid DNA was performed using the QIAprep Spin Miniprep Kit (Qiagen). Note that the standard manufacturer's protocol recommends incubation of cultures for 16 h at 37°C and using 5 ml of the culture. However, the yield of plasmid DNA was greatly increased when incubation for just 5 h at 37°C and 1.5 ml of culture was used. All solutions listed below were provided with the kit.

- A single *Cronobacter* colony from a fresh TSA plate was inoculated into 1.5 ml of LB with the appropriate selective antibiotic and incubated for 5 h at 37°C in a shaking incubator (2000 rpm).

- The bacterial cells were harvested by centrifugation at > 8000 rpm (6800 x g) in table-top microcentrifuge for 3 min at room temperature and the supernatant was discarded.

- The pelleted bacterial cells were resuspended in 250 µl Buffer P1 with added RNAse and LyseBlue reagent, and transferred to a microcentrifuge tube. The bacteria were resuspended completely by vortexing until no cell clumps remained.

- 250 μl of Buffer P2 were added and the suspension was mixed thoroughly by inverting the tube until a homogeneously coloured suspension was achieved. The lysis reaction was not allowed to proceed for more than 5 min.

- 350 μl of Buffer N3 were added and the solution was mixed thoroughly by inverting the tube. The suspension was mixed until all trace of blue had gone and the suspension was colourless.

- The suspension was centrifuged for 10 min at 13,000 rpm (~17,900 x g) in a microcentrifuge.

- The supernatant was applied to a QIAprep spin column by decanting and centrifuged for 30–60 s at 13,000 rpm (~17,900 x g). The flow-through was discarded.

- The QIAprep spin column was washed by adding 0.5 ml Buffer PB and centrifuged for 30–60 s at 13,000 rpm (~17,900 x g). The flow-through was discarded.

- The QIAprep spin column was washed by adding 0.75 ml Buffer PE and centrifuged for 30–60 s at 13,000 rpm (~17,900 x g). The flow-through was discarded and the column was centrifuged for an additional 1 min to remove residual wash buffer.

- The QIAprep column was placed in a clean 1.5 ml microcentrifuge tube. To elute DNA, 50 μl of Buffer EB were added to the center of each QIAprep spin column, let stand for 1 min, and centrifuged for 1 min.

3.4.2.3 PCR amplification of kanamycin cassette

To amplify the kanamycin cassette from pDOC-K, primers containing regions annealing to the extremities of the kanamycin cassette and to the DNA flanking the gene of interest were designed (Table 9). The forward primers were designed to contain 32 bp of the DNA immediately upstream of the start of the gene of interest, followed by an 18-bp sequence, which anneals to the 5' end of the kanamycin casette

on pDOC-K. The reverse primers were designed to contain 32 bp of the reverse complement of the DNA immediately downstream to the 3'end of the gene and a 17-bp sequence that anneals to the 3'end of the kanamycin cassette.

Table 9. List of primers used for the PCR amplification of the 1.4-kb kanamycin cassette from pDOC-K.

Name	Sequence[1]
cusA_F	TGTAAATATGCATTCAGGGCACTGAGGAGACGACG**AATTGGCTGGAGCTGCTTC**
cusA_R	GAAGGACACCCACAGGGGGTGCCCTTTCAGGTT**CTCCTTAGTTCCTATTCC**
cusB_F	ATGCCAGGTGTTTTGATTTTTTAGCGGAAAATTGT**AATTGGCTGGAGCTGCTTC**
cusB_R	GATGCCGGACTGTTGCTGTGACTCATGCAGAA**CTCCTTAGTTCCTATTCC**
cusC_F	AACCATTTACCAGGTCTGCCTGGACGAGAAGCGTT**AATTGGCTGGAGCTGCTTC**
cusC_R	GAGAGTTATCTGAAGTGAATCAAGATAACGCTG**CTCCTTAGTTCCTATTCC**

[1] The sequences annealing to the kanamycin casette are marked in **bold.**

High-fidelity Phusion® (Finnzymes) polymerase was used to generate blunt-ended PCR products suitable for cloning into pDOC. Each PCR reaction contained the following reagents:

Phusion polymerase buffer 5x	5 µl
Magnesium chloride (50mM)	5.5 µl
dNTP mix (10mM each dNTP)	1.0 µl
Forward primer (10µM)	1.25 µl
Reverse primer (10 µM)	1.25 µl
DMSO (100%)	0.8 µl
Phusion polymerase (0.02 U/ µl)	0.25 µl
Nuclease free water	To a final volume of 25 µl

The tube was gently mixed by tapping and 2 µl of the pDOC-K template (15-30 ng) were added. The manufacturer of Phusion® HotStart DNA polymerase recommends using the 2-step protocol with primers that are longer than 20 nt and have Tm values

above 69°C. All the primers used for the amplification of kanamycin cassette satisfied this condition and a 2-step protocol with the following settings was run.

Initial denaturation	98°C	30 s
Denaturation	98°C	10 s
Annealing + extension	72°C	26 s
30 cycles		
Final extension	72°C	10 min

To confirm the results of the PCR, a 5 µl aliquot of the reaction mixture was analysed by agarose gel electrophoresis (3.4.2.4). Subsequently, the PCR products were purified using the Qiaquick PCR purification kit (3.4.2.5) and their concentration was measured by spectrophotometry at 260 nm using NanoDrop.

3.4.2.4 Agarose gel electrophoresis

To prepare 1% agarose gel, 0.5 g of agarose were mixed with 50 ml of TAE buffer and 7 µl of SYBRSafe. The agarose solution was microwaved until dissolved and poured into the electrophoresis casting tray to settle. Aliquots (5 µl) of PCR samples were mixed with 1 µl of gel loading dye and loaded onto the gel, alongside 5 µl 1-kb DNA ladder. The gel was run in TAE buffer at 100 V for 40 min (5 V/cm).

3.4.2.5 Purification of PCR products

The PCR products were purified using the QIAquick PCR Purification Kit (Qiagen). The protocol is detailed below.

- 200 µl of Buffer PB with added indicator was added to 40 µl of the PCR sample and mixed.

- A QIAquick spin column was placed in a provided 2 ml collection tube. To bind DNA, the sample was applied to the QIAquick column and centrifuged for 60 s at 17,900 x g (13,000 rpm) in a microcentrifuge. The flow-through was discarded and the QIAquick column was placed back into the same tube.

- 0.75 ml of Buffer PE was added to the QIAquick column and centrifuged for 60 s. The flow-through was discarded and the QIAquick column placed back in the same tube.

- The column was centrifuged for an additional 1 min at 17,900 x g (13,000 rpm).

- The QIAquick column was placed in a clean 1.5 ml microcentrifuge tube and 30 μl of Buffer EB were added to the center of the QIAquick membrane and centrifuged the column for 1 min at 17,900 x g (13,000 rpm). To increase the concentration of the purified PCR product, this step was repeated.

3.4.2.6 Restriction of pDOC by *Eco*RV

To open the pDOC vector by *Eco*RV, the following reagents were mixed in a 1.5 ml tube:

*Eco*RV (10,000 units/ml)	1 μl
10X NE buffer	5 μl
pDOC template	up to 100ng
nuclease-free water	add to final volume of 50 μl

*Eco*RV was the last reagent added and all the components were mixed prior to adding the enzyme. The reaction mixture was incubated at 37°C for 1 hour. As the presence of the enzyme could interfere with further processing of the sample, the enzyme was heat inactivated by incubating the reaction mixture at 80°C for 20 min.

3.4.2.7 Cloning of the kanamycin cassette into pDOC

For cloning of the kanamycin cassette flanked by regions homologous to the DNA adjacent to the genes to be deleted into pDOC, the ratio of vector : insert was determined according to the following equation:

3 x ng (vector) x kb (insert) / kb (vector) = ng insert

For a 1.4 kb insert and a 4 kb vector, the ratio of insert : vector was 1 : 1 (ng).

To increase the efficiency of the cloning reaction, a higher ration insert : vector was used (1.5 : 1.0).

The following reagents were added to a 0.5 ml PCR tube and mixed:

Vector DNA	50 ng
Insert DNA	75 ng
Ligase 10 x buffer	1.5 µl
Nuclease-free water	to the final volume of 15 µl
T4 DNA ligase (10 U/ µl)	1 µl

The sample was then incubated at 10°C overnight.

3.4.2.8 PCR to check the success of cloning of the kanamycin cassette into pDOC

To confirm that the cloning of the kanamycin cassette into the pDOC plasmid was successful, PCR using the primers annealing to the DNA flanking the multicloning site in pDOC (Table 10) were used.

Table 10. Primers used to check the success of cloning of the kanamycin cassette into pDOC.

Name	Sequence	%GC[1]	Tm[1]
pDEXF	TATGCTTCCGGCTCG	60	50
pDEXR	GGATGTGCTGCAAGG	60	50

[1]...%GC and Tm were calculated using the biomath Tm calculator available at: http://www.promega.com/biomath/calc11.htm

Each PCR reaction contained the following reagents:

Phusion polymerase buffer 10x	5 µl
Magnesium chloride (50mM)	5.5 µl
dNTP mix (10mM each dNTP)	0.5 µl
Forward primer (10µM)	1.25 µl
Reverse primer (10 µM)	1.25 µl
DMSO	0.8 µl
Phusion® polymerase (0.02 U/ µl)	0.25 µl
Nuclease-free water	To the volume of 50 µl

The contents of the PCR tube was gently mixed by tapping, and 2 µl of the reaction mixture from the ligation step (3.4.2.7) were added. The manufacturer of Phusion® HotStart DNA polymerase recommends using the 3-step protocol when Tm values are below 72°C for primers shorter than 20 nt (pDEXF and pDEXR). Extension of 15 s/kb product for plasmid DNA is recommended. A 3-step PCR protocol with the temperature cycling conditions detailed below was run. The empty vector would give product of ca 200 bp, whereas the vector with the insert would give a product of 1.6 kbp.

Initial denaturation	98°C	30 s
Denaturation	98°C	10 s
Annealing	56°C	30 s
Extension	72°C	26 s
30 cycles		
Final extension	72°C	10 min

To visualise the results of the PCR, a 5 µl aliquot of the reaction mixture was analysed by agarose gel electrophoresis (3.4.2.4).

3.4.2.9 Transformation of OneShot® Top10 chemically competent *E. coli* cells

The protocol for the transformation of chemically competent One Shot® Top10 *E. coli* cells (Invitrogen) with the plasmid DNA is outlined below. The SOC medium was provided with the kit.

- A water bath was equilibrated to 42°C, SOC and LB media were warmed to room temperature.

- LB plates containing 50 µg/ml kanamycin were warmed at 37°C for 30 min

- The Top10 cells were thawed on ice.

- 1-5 µl of the plasmid DNA were added to Top 10 vial, mixed gently and incubated on ice for 30 min.

- The cells were heat shocked at 42°C without shaking for 30 s.

- 250 µl of SOC medium were added to each vial and the cells were incubated at 37°C for 1 h in a shaking incubator (225 rpm).

- 20-200 µl from each transformation were spread on a selective plate and incubated overnight at 37°C.

Note: Positive (pUC19) and negative (Top10 cells without added plasmid) controls were included in the assay.

3.4.2.10 Preparation of electrocompetent *Cronobacter* cells

- A single *Cronobacter* colony from a fresh LB plate was inoculated into a flask containing 50 ml of LB and the culture was incubated overnight at $37^{o}C$ with vigorous shaking (250 rpm).

- Two flasks with 500 ml of prewarmed LB medium were inoculated with 25 ml of the overnight bacterial culture and the flasks were incubated at $37^{o}C$ with aeration (200 rpm).

- The OD_{600} of the cultures was measured every 20 min.

- When the OD_{600} of the cultures reached 0.4, the flasks were rapidly transferred to an ice-cold water bath for 30 min.

- The cultures were transferred to ice-cold centrifuge bottles and the cells were harvested by centrifugation at 3000 rpm for 15 minutes at $4^{o}C$. The supernatant was decanted and the pellet resuspended in 500 ml of ice-cold water.

- The cells were centrifuged at 3000 rpm for 20 minutes at $4^{o}C$. The supernatant was carefully decanted and the pellet resuspended in 250 ml ice-cold 10 % glycerol.

- The cells were centrifuged at 3500 rpm for 20 min at $4^{o}C$. The supernatant was carefully decanted and the pellet resuspended in 10 ml ice-cold 10 % glycerol.

- The cells were centrifuged at 3500 rpm for 20 min for 20 min at $4^{o}C$. The supernatant was carefully decanted and the pellet resuspended in 1 ml of ice-cold GYT medium.

- 80 µl aliquots of the cell suspension were stored at $-80^{o}C$ until required

3.4.2.11 Electroporation of plasmid DNA into electrocompetent *Cronobacter* cells

- Electrocompetent *Cronobacter* cells were thawed on ice and 80 µl aliquots were pipetted into ice-cold sterile microcentrifuge tubes placed on ice.

- 1.5 µl of the plasmid DNA preparation (8-16 ng/µl) were added to each tube with electrocompetent cells and incubated on ice for 30-60 s. Using larger volumes than 1.5 µl was avoided, as it resulted in arcing during the electrical impulse.

- The DNA/cell mixture was pipetted into an ice-cold electroporation cuvette, the outside of the cuvette was wiped dry and the cuvette was placed in the electroporation device (MXcell, BIO-RAD). The electroporation apparatus was set to deliver an electrical pulse of 25 µF capacitance, 2.5 kV and 200 ohm resistance.

- The cuvette was removed as quickly as possible from the apparatus and 1 ml of LB medium (at room temperature) was added.

- The contents of the cuvette were transferred to a microcentrifuge tube and the cultures were incubated for 1 h at 37°C in a shaking incubator (200 rpm).

- Different volumes (50 and 150 µl) of the cultures were plated on LB plates supplemented with the appropriate antibiotic (35 µg/ml chloramphenicol for pACBSR or pACBSRSceI and 200µg/ml ampicillin with 35 µg/ml chloramphenicol for the pDOC construct).

- The plates were incubated overnight at 37°C.

3.4.2.12 Arabinose induction and selection of mutants

- *Cronobacter* colonies were patched on plates containing 10 % sucrose or 200 µg/ml ampicillin.

- A single fresh sucrose sensitive and ampicillin resistant colony was inoculated into 1 ml of LB supplemented with 200 µg/ml ampicillin, 35 µg/ml chloramphenicol, 50 µg/ml kanamycin and 0.5 % glucose (w/v). The cultures were incubated at 37°C for 2 hours in a shaking incubator (200 rpm). Usually four parallel tubes were processed for each experiment.

- Cells were harvested by centrifugation and re-suspended in 1 ml LB containing 0.5 % arabinose (w/v). Note that antibiotics were omitted from the growth medium at this stage. The culture was incubated at 37°C for 4-5 hours in a shaking incubator (200 rpm).

- Dilutions of the culture were plated on to LB agar plates containing 50 µg/ml kanamycin and 10 % (w/v) sucrose and incubated overnight at 30°C.

- Kanamycin resistant, sucrose insensitive recombinants were checked for donor plasmid loss by patching onto LB agar plates containing 200 µg/ml ampicillin and pACBSR or pACBSRSceI plasmid loss by patching onto LB agar plates containing 35 µg/ml chloramphenicol.

3.4.2.13   Verification of gene deletion of *cusA, cusB* and *cusC* by PCR

The samples of bacterial cultures for PCR screening were prepared as follows. A single *Cronobacter* colony was resuspended in 10 µl of PBS and the suspension was heated in the PCR thermocycler at 99°C for 10 min. Two microliters of the heated solution were used in each PCR reaction

To verify that the recombination event occurred at the desired loci, PCR primers annealing to the sites adjacent to the regions used for generating the kanamycin cassette PCR insert (3.4.2.3) were designed (Table 11). Size differences in PCR product between the wild-type and modified strain confirm whether a chromosomal modification has been achieved. The predicted PCR products for *cusA* were 3.2 kb for wild-type strain and 1.4 kb for deletion mutant. In case of primers for the *cusB* and *cusC* mutants, the PCR products generated from the wild-type strain would have a very similar size to the PCR product amplified from the deletion mutant; the reverse primers cusB_CTRL_R   and cusC_CTRL_R were moved into the *cusC* gene. The wild type PCR products would in these cases be 1.2 kb, whereas no products would be generated from deletion mutants.

Table 11. List of primers used in the PCR verification of gene deletion.

Name	Sequence	%GC	Tm
cusA_CTRL_F	AGGGCACTGAGGAGACGACGA	62	62
cusA_CTRL_R	CCCACAGGGGGTGCCCTTTC	70	62
cusB_CTRL_F	CCAGGTGTTTTGATTTTTTAGCGG	42	56
cusB_CTRL_R	ATGCCGGACTGTTGCTGTGACT	55	61
cusC_CTRL_F	GGTCTGCCTGGACGAGAAGCG	67	62
cusC_CTRL_R	CTGAAGTGAATCAAGATAACGC	41	54

Each PCR reaction contained the following reagents:

GoTaq® Flexi buffer 5x	3.0 µl
Magnesium chloride (50mM)	1.5 µl
dNTP mix (10mM each dNTP)	0.3 µl
Forward primer (10 µM)	0.7 µl
Reverse primer (10 µM)	0.7 µl
GoTaq® Hot Start Polymerase (5U/ µl)	0.1 µl
Nuclease free water	To a final volume of 15 µl

The tube was gently mixed by tapping and 2 µl of the template were added. The tubes were placed in a PCR thermocycler and a 3-step protocol with the following temperature cycling parameters was run. The 2 min denaturation step at 95°C is necessary to activate the GoTaq® Hot Start polymerase.

Initial denaturation	95°C	2 min
Denaturation	95°C	30 s
Annealing	62°C[1]	30 s
Extension	72°C[2]	1 min/kb of product
30 cycles		
Final extension	72°C	5 min

[1] gradient PCR was performed to optimize the annealing temperature for the PCR protocol. Expected products (3.2 k for *cusA* and 1.2 kb for *cusB* and *cusC*) were amplified in the whole range of annealing temperatures (57-66°C) with no visible difference in PCR product quantity

[2] 3.2 min for *cusA*, 1.2 min for *cusB* and *cusC*

To visualise the results of the PCR, a 10 µl aliquot of the reaction mixture was analysed by agarose gel electrophoresis (3.4.2.4).

3.5. Results

3.5.1 *Cronobacter* DNA microarray and comparative genomic hybridisation

3.5.1.1 Development of new CGH data analysis tool

WebArrayDB available at (http://webarraydb.org/webarray) is a database system and online cross-platform analysis suite for analysis of microarray data, developed by Prof. McClelland and his team, which allows storage of data in the repository and their online analysis. Prior to this study, the site only included tools for microarray expression data analysis, but Prof. McClelland agreed to design a CGH tool that could be used with data from NimbleGen arrays. The tool was developed by Xiao-Qin Xia and tested and optimised by myself. The new utility includes a CGH plotter which can plot the $\log_2$ ratio data against genomic location in an window of adjustable size with the possibility to colour code ranges of $\log_2$ ratio values (Figure 26) and a calculator of $\log_2$ ratio medians for each gene from $\log_2$ ratios of separate probes. The tool is now accessible on WebArrayDB upon request.
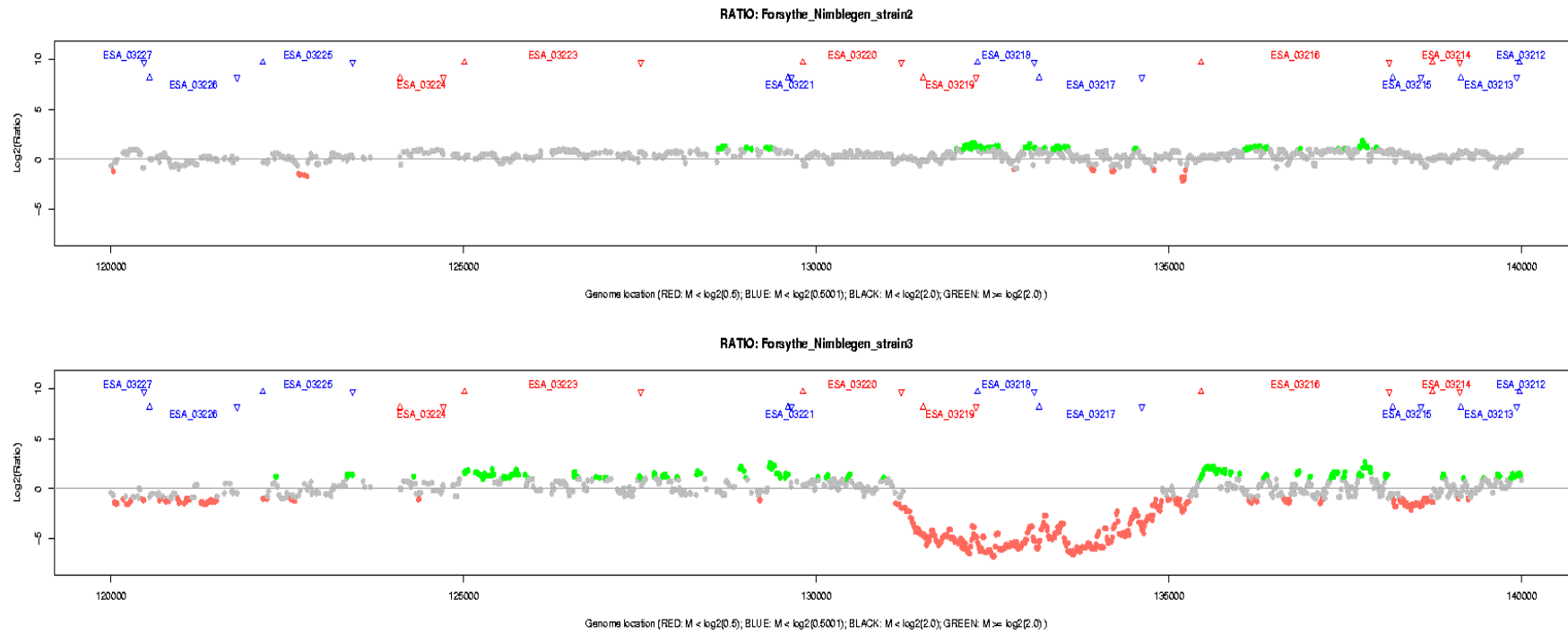
Figure 24. Example of a CGH plot in WebArrayDB. Two-channel microarray hybridization data for *Cronobacter sakazakii* ATCC 12868 (top chart) and *Cronobacter muytjensii* ATCC 51329[T] (bottom chart) hybridized against the reference strain *Cronobacter sakazakii* BAA-894 is plotted against the genome location of oligonucleotide probes. Green: gain of signal, $Log_2(Ratio)>1$; red: loss of signal, $Log_2(Ratio)<-1$. $Log_2(Ratio) = log_2(test channel intensity/reference channel intensity)$.

125

3.5.1.2 Hybridization data distributions and gene category assignment

Figure 25 shows the distribution of median $\log_2$ ratio intensity data for all strains examined by CGH. The microarray data distribution confirmed the taxonomic revision of *Cronobacter* spp. (formerly *Enterobacter sakazakii)* which separated the organisms into five different species (and one genomospecies). The $\log_2$ ratio histograms of *C. sakazakii* strains hybridised against the reference strain *C. sakazakii* BAA-894 (in red) have narrow main peaks, whereas the wider peaks of the $\log_2$ ratio distributions of the more distantly related species *C. muytjensii, C. dublinensis, C. turicensis* and *C. malonaticus* against *C. sakazakii* BAA-984 (in blue) reflect a greater sequence divergence of the test strains relative to the reference strain. The tailings on the left side of the distributions correspond to the genes that had a low test channel intensity value and therefore would be classified or absent/divergent. Figure 25 shows that the proportion of these genes is greater in the other members of the *Cronobacter* genera (in blue) compared to the *C. sakazakii* strains (in red). *C muyjensii* appears to be the most divergent compared to the reference *C. sakazakii* BAA-894.
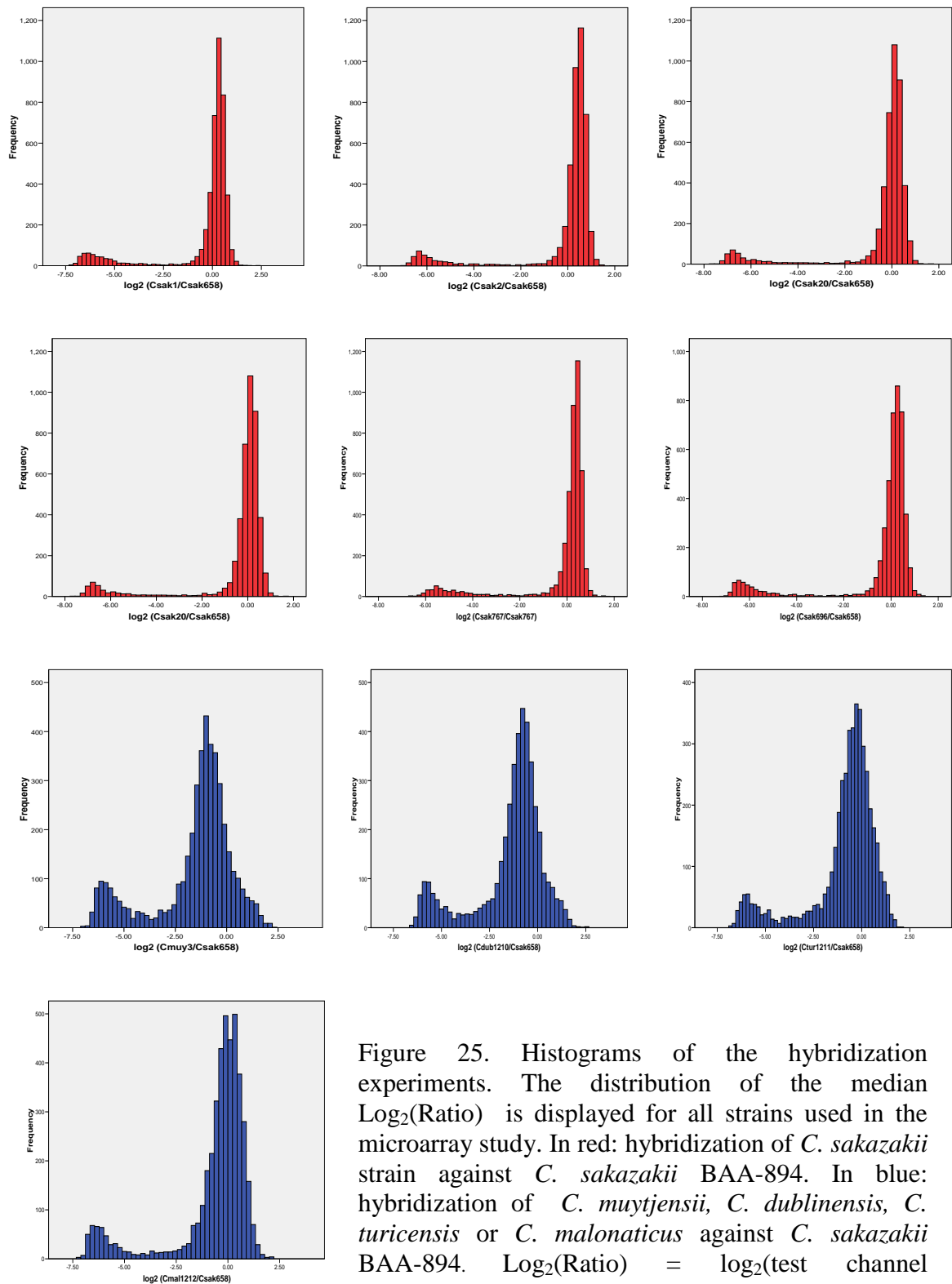
Figure 25. Histograms of the hybridization experiments. The distribution of the median Log_2(Ratio) is displayed for all strains used in the microarray study. In red: hybridization of *C. sakazakii* strain against *C. sakazakii* BAA-894. In blue: hybridization of *C. muytjensii, C. dublinensis, C. turicensis* or *C. malonaticus* against *C. sakazakii* BAA-894. Log_2(Ratio) = log_2(test channel intensity/reference channel intensity).

The cut-off values used to classify each of the 4,382 genes represented on the microarray as i) present, ii) intermediate or iii) absent, were determined as described in 3.4.1.9 and are listed in Table 12.

Table 12. Cut-off values of hybridiation ratios (test intensity / reference intensity) used for the assignment of absent, intermediate or present gene status.

Strain	Absent / intermediate	Intermediate / present
C. sakazakii ATCC 29544[T]	0.51	0.93
C. sakazakii ATCC 12868	0.54	1.03
C. sakazakii strain 20	0.45	0.85
C. sakazakii strain 701	0.62	1.01
C. sakazakii strain 767	0.66	1.05
C. sakazakii strain 696	0.46	0.89
C. malonaticus LMG 23826[T]	0.16	0.57
C. turicensis LMG 23827[T]	0.11	0.43
C. muytjensiii ATCC 51329[T]	0.10	0.31
C. dublinensis LMG 23823[T]	0.11	0.32

3.5.1.3 Core genome

Of the 4,382 unique annotated gene sequences represented on the microarray, 54.9% (2,404) were classified as present in all *C. sakazakii* strains and 43.3% (1,899) were present in all ten *Cronobacter* strains representing five *Cronobacter* species. The vast majority of these shared genes are predicted to encode cellular essential functions such as energy metabolism, biosynthesis, DNA, RNA and protein synthesis, cell division and membrane transport. It should be noted that this core gene set was defined according to the most stringent trinary classification of genes by GACK and does not include the genes classified as intermediate. When intermediate genes are included, the core gene set for *C. sakazakii* species constitutes 80.9% (3,547) genes and core gene for *Cronobacter* genus includes 75.3% (3301) genes. The proportion of genes absent from test strains compared with *C. sakazakii* BAA-894 ranged from 10.3% (453) in *C. sakazakii* strain 20 to 17.1% (751) in *C. muytjensii*. The trinary

classification of genes in 10 *Cronobacter* strains analysed by CGH is summarised in Table 13.

Table 13. Number of absent, intermediate and present genes in *Cronobacter* strains when compared to the sequenced *C. sakazakii* BAA-894.

Strain	Absent genes (%)[1]	Intermediate genes (%)[1]	Present genes[1]
C. sakazakii ATCC 29544[T]	582 (13.3)	449 (10.2)	3351 (76.5)
C. sakazakii ATCC 12868	461 (10.5)	314 (7.2)	3607 (82.3)
C. sakazakii strain 20	453 (10.3)	580 (13.2)	3349 (76.4)
C. sakazakii strain 701	521 (11.9)	569 (13.0)	3292 (75.1)
C. sakazakii strain 767	546 (12.5)	649 (14.8)	3187 (72.7)
C. sakazakii strain 696	497 (11.3)	593 (13.5)	3292 (75.1)
C. malonaticus LMG 23826[T]	522 (11.9)	608 (13.9)	3252 (74.2)
C. turicensis LMG 23827[T]	495 (11.3)	591 (13.5)	3296 (75.2)
C. muytjensii ATCC 51329[T]	751 (17.1)	522 (11.9)	3109 (70.9)
C. dublinensis LMG 23823[T]	682 (15.6)	549 (12.5)	3151 (71.9)
Absent in at least one *C. sakazakii*	835 (19.1)		
Absent in at least one *Cronobacter*	1081 (24.7)		

[1]... Percentage out of 4,382 *C. sakazakii* BAA-894 genes represented on the microarray. For definition of "absent", "intermediate" and "present" gene refer to 3.4.1.9.

3.5.1.4 *C. sakazakii* plasmids

The sequenced strain *C. sakazakii* BAA-894 contains two plasmids; pESA2 (31 kb) and pESA3 (131 kb). Thirty-eight genes were annotated on pESA2 and 127 genes on pESA3. The copy number of the plasmids was estimated from the median hybridization signals of oligonucleotides representing the plasmid compared to the sequenced genome. The ratio was (1:1.1:8.6) for the chromosome versus pESA2 versus pESA3. Thus, pESA2 exists as low copy, and pESA3 appears to be a moderate copy number plasmid. The genes on pESA2 were absent in all other strains tested except *C. turicensis*, which had 19 (61.3%) genes present, and *C. sakazakii* 696, which had 4 (12.9%) genes present. Note that it is possible that some or all of the

detected present genes are located on the chromosome in other strains. The results for genes on pESA3 are summarized in Table 14.

Table 14. pESA3 genes present in *Cronobacter* strains when compared to the sequenced strain of *C. sakazakii* BAA-894.

Strain	Present genes (%)[1]	Plasmid profile[2]
C. sakazakii ATCC 12868	103 (88.8)	+
C. sakazakii strain 696	81 (69.8)	+
C. sakazakii strains 767	77 (66.4)	+
C. sakazakii strain 20	73 (62.9)	+
C. sakazakii strain 701	61 (52.6)	+
C. malonaticus LMG 23826[T]	57 (49.1)	+([3])
C. sakazakii ATCC 29544[T]	51 (44.0)	+([3])
C. turicensis LMG 23827[T]	35 (30.2)	-
C. dublinensis LMG 23823[T]	19 (16.4)	-
C. muytjensii ATCC 51329[T]	0 (0.0)	-

[1] Of the 127 annotated genes on pESA3, 116 were analyzed by comparative genomic hybridization. The other genes did not have sufficient number of probes that passed the filters for manufacture of the array.

[2] Plasmid of a size similar to pESA3 (131 kb) was detected by gel electrophoresis of plasmid isolations.

[3] The detected plasmid was smaller in size than pESA3 (110 kb).

Plasmid profiling was performed on the *Cronobacter* strains analyzed by comparative hybridization. A plasmid of a size similar to pESA2 (31 kb) was detected in *C. sakazakii* 696 and in *C. turicensis*, which is in accordance with our CGH results. A large plasmid similar in size to pESA3 (131 kb) was visible in *C. sakazakii* strains ATCC 29544[T], ATCC 2868, 20, 696, 701, 767 and *C. malonaticus* (Table 14).

Note that genes on a multicopy or medium copy plasmid may require a different degree of divergence to be identified as absent or divergent by comparative hybridizations and it is therefore difficult to establish the accuracy of the CGH results

for plasmid genes. However, our CGH data correlated well with the experimental data, plasmids of similar size as pESA3 was detected in strains with 53-89 % present pESA3 genes as determined by CGH, smaller plasmids (110 kb) were detected in strains with 51 and 57 % present pESA3 genes and no plasmid was detected in strains with 0-35 % pESA genes.

3.5.1.5 Genome cluster analysis

 Whole-genome clustering based on the gene category assignment by GACK (3.4.1.9) revealed that *Cronobacter* strains formed two distinct phylogenetic clusters (Figure 26). All *C. sakazakii* strains formed one cluster, whereas *C. malonaticus, C. turicensis, C. dublinensis* and *C. malonaticus* formed a second, separate cluster. Within *C. sakazakii*, strains 701 and 767 were the most closely related and clustered together with strain 20. The remaining *Cronobacter* species formed sub-clusters: *C. malonaticus* clustered with *C. turicensis* and *C. dublinensis* grouped with *C. muytjensii*. The tree remained identical when adjacent genes were collapsed into a single phylogenetic character if they had the same pattern of presence and absence. The regions GR1- GR15 annotated Figure 26 will be described in the next section (3.5.1.6). Note that the clustering was based on the presence of the genes of the reference *C. sakazakii* BAA-894 in other *Cronobacter* strains, and therefore genes that were present in other *Cronobacter* strains and missing in the reference strain were not involved in the analysis.
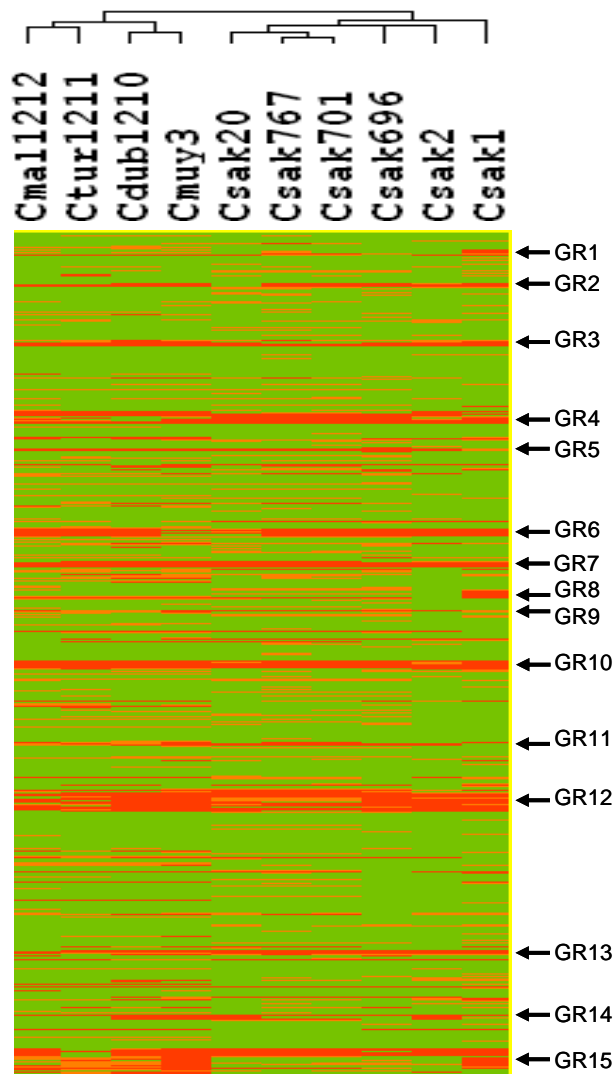
Figure 26. Whole genome clustering analysis showing the relatedness of 10 *Cronobacter* isolates with respect to the genes of the reference strain *C. sakazakii* BAA-894. Genomic regions GR1–GR15 are marked. Clustering analysis was performed using Gene Cluster (EisenSoftware). Hierarchical clustering was performed using the average linkage method on the trinary matrix based on the CGH analysis: 1 for presence, 0 for intermediate status and −1 for absence/divergence of a gene. Green indicates present genes, orange uncertain genes and red absent genes. For description of strains refer to Table 7.

3.5.1.6 Genomic regions absent in some strains of *Cronobacter*

Genes that were absent in more than half of *Cronobacter* strains relative to the sequenced strain *C. sakazakii* BAA-894 were selected for further analysis. These genes form 15 clusters of contiguous genes (based on the annotation of the reference genome). These are shown on Figure 27, where the number of strains in which a particular gene was classified as absent is plotted against the gene locus. The clusters were designated as regions GR1–GR15.
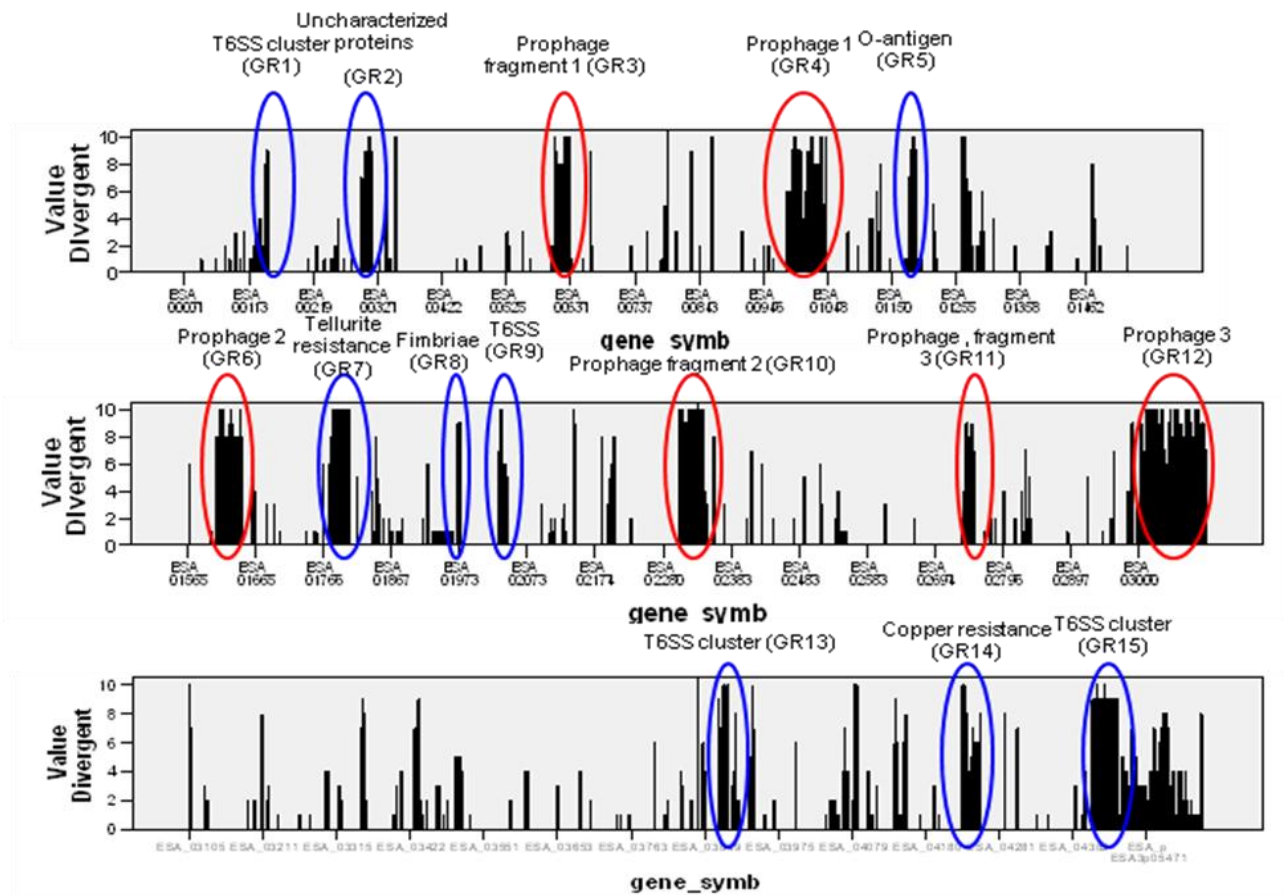
Figure 27. Distribution of the 15 variable regions across the *C. sakazakii* BAA-894 genome. Each column represents a gene classified as absent by CGH analysis in at least one strain. The height of the columns indicate the number of strains (out of 10) in which the gene was found to be absent. The major variable regions (blue) and prophages (red) are indicated in order of their appearance in the genome of *C. sakazakii* BAA-894.

3.5.1.6.1 Prophages

Of the 15 clusters, three putative prophage genomes and one prophage fragment were identified by Prophinder (Lima-Mendez *et al*. 2008) and three additional regions are probable prophage fragments based on the presence of phage protein homologues identified by BLASTP. The genome organisation of the putative prophages is represented on Figure 28.

The average GC content of the sequenced *C. sakazakii* BAA-894 genome is 56%, the GC content of prophages 1, 2 and 3 was 53, 49 and 51%, respectively. In the three prophage gene clusters (prophages 1, 2 and 3), genes encoding close homologues of known phage genes involved in integration, lysis and termination as well as head and tail structure were identified, based on the annotation in IMG-JGI (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi). Where necessary, the genes were re-annotated based on amino acid identity searches in IMG-JGI. The complete list of re-annotated putative prophage genes is available in Appendix 2. In addition, Figure 29 shows the status of all putative prophage genes in the 10 *Cronobacter* strains.

Prophage 1 (**GR4**; ESA_00990–ESA_01052). In the 46-kb putative prophage 1 (Figure 28A), 30/63 (48%) hypothetical proteins were similar to known phage proteins. Homologues of phage genes involved in integration, lysis, head morphogenesis, tail assembly and phage regulation were identified. Prophage 1 also contains a gene (ESA_00997) encoding a protein homologous to the Eae-like adhesion protein associated with the attaching and effacing phenotype. However, the aligment on the Eae protein is only partial (50%). The *eae* gene is carried by some other bacteriophages of enteric pathogens: *Salmonella* phage epsilon34, *E. coli* O157:H7 bacteriophage PhiV10, enterobacterial phage P22 and enterobacterial phage epsilon15.

Prophage 2 (**GR6**; ESA_01608–ESA_01644). Putative prophage 2 (Figure 28B) contains 37 genes, out of which 25 (68%) were homologous to known phage proteins. Prophage 2 contains several lambdoid phage genes encoding the following proteins: repressor CII (ESA_01613), replication proteins O and P (ESA_01614 and ESA_01615), the antitermination protein Q (ESA_01622), small and large subunits of the phage terminase (ESA_01632 and ESA_01633) as well as head and tail morphogenesis proteins. Head morphogenesis genes (ESA_01635–ESA_01637) were similar to head proteins of bacteriophage HK97 from the family of lambda phages. Two gene clusters in the prophage have very low average GC content; 33% (ESA_01616–ESA_01620) and 44% (ESA_01627–ESA_01631). Both these clusters

contain hypothetical proteins that showed no similarity with known proteins or functions.

Prophage 3 (**GR12**; ESA_03025–ESA_03102) was the largest (47-kb) putative prophage identified (Figure 28C). Thirty-four genes (39%) genes had homology to known phage proteins or functions. The rest of the annotated genes are conserved proteins of unknown functions or hypothetical proteins. Similarly to prophage 2, several regulatory genes characteristic for lambdoid phages were identified: repressor proteins CI, CII and CIII, early gene regulator protein, replication proteins O and P as well as N independence proteins NinBFGZ. A cluster of three O-antigen conversion genes (ESA_03026–ESA_03028) was found in putative prophage 3 between phage integrase and tail morphogenesis genes (Figure 28). The putative colicin uptake protein TolA (ESA_03048) may be involved in the internalization of the bacteriophage, as the Tol pathway can be also used for the translocation of phages into the bacterial cell (Lazdunski 1988).

Figure 3
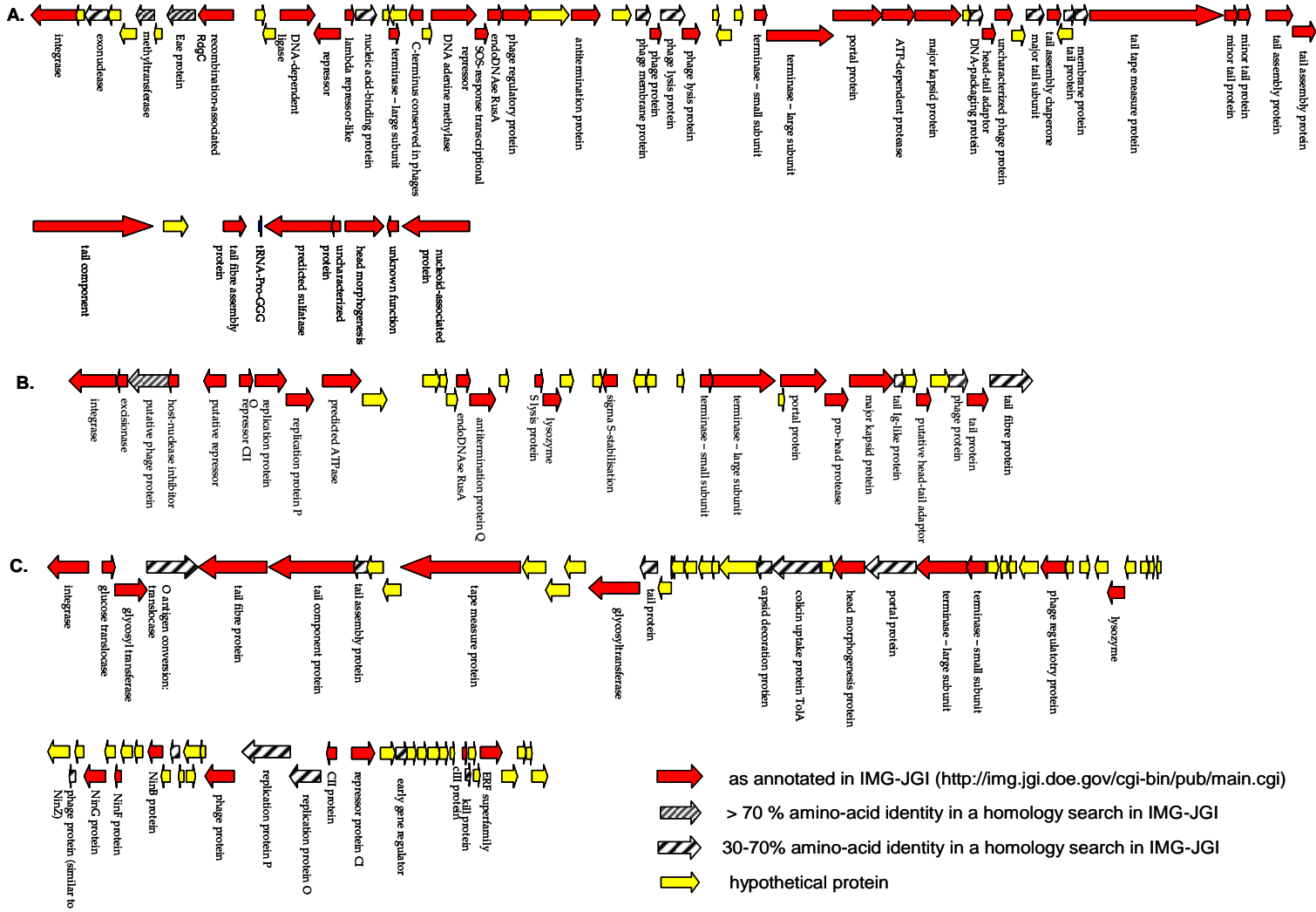
Figure 28. Gene maps of putative prophages. A. Gene map of putative prophage 1. B. Gene map of putative prophage 2. C. Gene map of putative prophage 3. Annotation of the putative prophage genes is available in Appendix 2.

Prophage fragment 1 (**GR3**; ESA_00604–ESA_00630). Eight of 19 genes in this region encode proteins associated with phages: plasmid and phage DNA primase (ESA_00620), a protein from Ash phage family (ESA_00624), the phage transcriptional regulator AlpA (ESA_00625), a putative phage capsid protein (ESA_00626), the phage transcriptional activator Ogr/Delta (ESA_00627) and phage integrase (ESA_00630). ESA_00618 was homologous to ea59 of lambda bacteriophage and ESA_00622 was homologous to a P4 phage protein. This cluster is most probably a phage remnant and may not encode a functional phage due to the absence of homologues of known structural tail proteins. The phage cluster was absent in all *Cronobacter* strains except a short region, ESA_00609–ESA_00617, that was present in *C. sakazakii* strains 701 and 767.

Prophage fragment 2 (**GR10;** ESA_02304–ESA_02339). This cluster is likely to represent another prophage remnant. The region mostly contains hypothetical proteins with unknown functions. However, eight genes showed some degree of homology to proteins of phage origin. The cluster is flanked by a gene homologous to phage methyltransferase (ESA_02304) and a gene containing a site-specific recombinase domain that is found in putative integrases/recombinases of mobile genetic elements of diverse bacteria and phages (ESA_02339). It also includes genes homologous to phage lysozyme (ESA_02309), a phage tail component (ESA_02311), a putative phage tape measure protein (ESA_02313), another unspecified phage protein (ESA_02316), a major capsid protein (ESA_02319) and a phage portal protein (ESA_02320). GR10 might be another remnant of a prophage that has previously integrated into *C. sakazakii* BAA-894 genome. The presence of putative integrase flanking the cluster suggests introduction of this cluster into the genome by horizontal gene transfer. This cluster was absent in all *Cronobacter* strains except the genes encoding phage lysozyme (ESA_02309) and a hypothetical protein (ESA_02310), present in *C. sakazakii* strains ATCC 12868 and 20, probably as a part of different prophages.

Prophage fragment 3 (**GR11**; ESA_02740–ESA_02755). The gene cluster of the 8-kb putative prophage fragment 3 comprises 16 hypothetical proteins, 10 of which

(63%) may be associated with phage functions. As it lacks genes for head and tail morphogenesis as well as phage regulatory genes, it is likely to be a non-functional phage remnant.

Figure 29. Absence/presence status of prophage genes in *Cronobacter* strains. Red indicates absence/divergence of a particular gene, orange indicates uncertain status and green indicates presence of a gene. For gene status determination see 3.4.1.9. Csak1 = *C. sakazakii* ATCC 29544[T]; Csak2 = *C. sakazakii* ATCC 12868, Cmal = *C. malonaticus* ; Cdub = *C. dublinensis*; Cmuy = *C. muytjensii*; Ctur = *C. turicensis*.

3.5.1.6.2 Non-phage regions that differ among the strains

There were nine genomic regions not associated with phage genes that were absent in more than half of the *Cronobacter* strains relative to the sequenced strain *C. sakazakii* BAA-894; these regions are described below. The complete list of genes belonging to the variable non-phage regions and their presence in tested strains is available in Appendix 2. Where necessary, the genes were re-annotated based on homology searches in IMG-JGI (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi).

GR1 (ESA_00140–ESA_00145) is a small cluster of T6SS genes which will be described in more detail in Section 3.5.1.13. ESA_00145, a secretion lipoprotein from a VC _A0113 family was present in *C. sakazakii* strains 701 and 767 associated with fatal outbreaks. The rest of GR1 was absent in all tested strains except *C. sakazakii* 696.

GR2 (ESA_00292–ESA_00310) mostly contains uncharacterized conserved proteins and was absent from all *Cronobacter* strains.

GR5 (ESA_01179–ESA_01189) contains a cluster of proteins involved in cell wall biogenesis and nucleotide sugar metabolism. It corresponds to the *C. sakazakii* O-antigen gene locus used to distinguish the two *Cronobacter* serotypes O:1 and O:2 (Mullane *et al*. 2008b). DNA microarray analysis revealed that GR5 is highly divergent; its genes were not sufficiently similar to be detected by microarray hybridization in any other *Cronobacter* strains. For more details on this region see 3.5.1.14.

GR7 (ESA_01775–ESA_01804). Most genes in GR7 were predicted to be involved in tellurite resistance. It contains homologues of tellurite resistance proteins TerA, TerC, TerD, TerY and TerZ. This region was absent from all *Cronobacter* strains tested by CGH.

GR8 (ESA_01970–ESA_01976) contains seven genes encoding pilus assembly proteins. This cluster of genes was absent in all *Cronobacter* strains except *C. sakazakii* strain 20. For more details on fimbriae see 3.5.1.11.

GR9 (ESA_02032–ESA_02041) contains genes encoding hypothetical proteins and four proteins involved in T6SS (ESA_02037–ESA_02040). These four genes were present in *C. sakazakii* strains 1 and 20 and absent in all other strains. For more details on type VI secretion systems see 3.5.1.13.

GR13 (ESA_03887–ESA_03912) is a cluster of 16 hypothetical proteins, mostly without homology to known proteins or functions. At least two genes encode homologues of T6SS components (ESA_03899 and ESA_03905). This region partly overlaps with the T6SS cluster described in 3.5.1.13. This cluster was absent in all other *Cronobacter* strains tested by CGH.

GR14 (ESA_04248–ESA_04255) encodes the putative *cop* operon involved in copper resistance. ESA_04248 encodes a protein identical to the copper resistance protein PcoE of *E. cloacae* (100% sequence identity, but only 60% homolog coverage), ESA_04249 encodes a putative copper-resistance protein CopA, ESA_04250 encodes a homolog of *E. coli* copper resistance protein CopB, ESA_04251 encodes a homolog of Cu resistance protein CopC and ESA_04252 encodes a putative copper export protein CopD. Along with CopA, CopB and CopC, CopD mediates copper resistance by sequestration of copper in the periplasm (Cha, Cooksey 1991). ESA_04253 encodes a heavy metal response regulator PcoR (CusR) and ESA_04254 encodes a putative heavy metal sensor kinase CusS. This cluster was found in *C. sakazakii* strains ATCC 29544[T] and 696, as well as *C. turicensis* and *C. malonaticus*.

GR15 (ESA_pESA3p05493–ESA_pESA3p05505) contains genes located on *C. sakazakii* plasmid pESA3 and includes components of T4SS and T6SS pathways. GR15 was absent from all strains except *C. sakazakii* strains ATCC 29544[T] and 696.

3.5.1.6.3 The differences in genome content of *C. sakazakii* and other *Cronobacter* species

A list of genes that were classified as present or intermediate in all *C. sakazakii* strains but absent in one of the other *Cronobacter* species was compiled. There were 56 genes absent in *C. malonaticus,* 91 genes absent in *C. turicensis*, 152 genes absent in *C. dublinensis* and 174 genes absent in *C. muytjensii.* Among these genes, clusters of contiguous genes were identified and denoted as "specific regions" SR1-SR13. The presence of these regions in the different *Cronobacter* species is summarized in Table 14 and the regions are described below.

SR1 (ESA_00258 and ESA_00257) encodes a putative toxin/antitoxin pair RelE/RelB. RelE acts like a translational repressor, probably through binding ribosomes (Galvani, *et al*. 2001, Pedersen *et al*. 2002). RelB stably binds RelE, presumably deactivating it. The toxin/antitoxin pairs encoded on plasmids may act like addiction modules that help to maintain a plasmid in a bacterial population. When encoded on a chromosome, however, the toxin/antitoxin system probably represents selfish DNA (Pedersen *et al*. 2002). In the reference *C. sakazakii* BAA-894, these genes were encoded on a chromosome. The microarray data do not provide the information about the location of these genes in other tested strains, however. This region was missing in all other *Cronobacter* species except *C. sakazakii.*

SR2 (ESA_01116 – ESA_01119) is a cluster of genes encoding a complete ABC-type multidrug efflux system. ESA_01116 encodes a multidrug efflux pump. ESA_01117 encodes outer membrane efflux protein from a family including TolC. ESA_01118 is the permease component of the ABC-type system and ESA_01119 encodes the ATPase component of the efflux system. This region is missing in all other *Cronobacter* species except *C. sakazakii.*

SR3 (ESA_01448 - ESA_01450) encodes three proteins from the family of fatty acid desaturases. Members of this family are involved in cholesterol biosynthesis and

biosynthesis of a plant cuticular wax, but may be implicated in other biosynthetic pathways (Arthington *et al*. 1991, Aarts *et al*. 1995). The exact function of this cluster of genes is therefore unknown. This cluster was missing in *C. malonaticus*, but present in all other *Cronobacter* species.

SR4 (ESA_02125 - ESA_02129) encodes a diverse group of proteins where no common assignment to a pathway or mechanism could be found. ESA_02125 and ESA_02128 encode acetyltransferases. ESA_02125 was present or intermediate in all species except *C. turicensis*, whereas ESA_02128 was only present in *C. sakazakii* strains and intermediate in *C. malonaticus*. ESA_02126 encodes a transcriptional regulator from the LysR family, this gene was present or intermediate in all species except *C. turicensis*. ESA_02127 is a gene encoding a putative esterase / lipase, which was present in all species except *C. muytjensii* and *C. dublinensis*. ESA_02129 encodes serine protease inhibitor ecotin. The gene for ecotin was intermediate in *C. sakazakii* strains and in *C. malonaticus*, but absent in *C. muytjensii, C. dublinensis* and *C. turicensis*.

SR5 (ESA_02538 – ESA_02542) is a cluster of fimbrial genes. See section 3.5.1.11 for details.

SR6 (ESA_02544 – ESA_02547). The genes from this cluster are involved in the metabolism of β-glucosides. ESA_02544 is a transcriptional antiterminator from the BglG family, which is involved in positive control of the utilization of different sugars by transcription antitermination (Bardowski *et al*. 1994). ESA_02545 encodes a kinase that converts β-glucosides to 6-phospho-β-glucosides and ESA_02546 encodes a 6-phospho- β-glucosidase (EC 3.2.1.86). According to the annotation in IMG-JGI (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi), ESA_02546 is specific to arbutin-6 phosphate and salicilin-6-phosphate, transforming them into β-D-glucose-6-phosphate, whereas ESA_02545 has a wider specificity to β-glucosides. It remains to be determined whether the pathway described is specific to arbutin/salicin or may convert a wider range of β-glucosides. ESA_02547 also encodes 6-phospho-β-

glucosidase (EC 3.2.1.86) which may have the same or similar function as ESA_02546. This cluster of genes was present in all strains except *C. turicensis.*

SR7 (ESA_02549 – ESA_02553) may encode a complete ABC multidrug transport system. ESA_02549 encodes a transcriptional regulator from the TetR family of proteins, which may function as a repressor that controls the level of susceptibility to hydrophobic antibiotics and detergents (Pan & Spratt 1994, Hillen & Berens 1994). ESA_02550 encodes a multidrug resistance efflux pump, ESA_02551 encodes the ATPase component and genes ESA_02552 and ESA_02553 encode the permease components. This cluster of genes was present in all *Cronobacter* strains except *C. muytjensii*.

SR8 (ESA_02616 – ESA_02618) contains genes related to mannose metabolism. ESA_02616 is an alpha-mannosidase, ESA_02617 is a gene taking part in mannosyl-D-glycerate uptake via the phosphotransferase system and ESA_02618 is mannosyl-D-glycerate transport/metabolism system repressor. This cluster was present in all species except *C. turicensis*.

SR9 (ESA_02795 – ESA_02799) is a fimbrial cluster. For more details see 3.5.1.11.

SR10 (ESA_03301 – ESA_03305). This region encodes proteins involved in the metabolism of fructose and mannose via the PTS and a putative porin KdgM. ESA_03301 encodes a isomerizing glucosamine--fructose-6-phosphate aminotransferase. The SIS domain of ESA_03301 is found in phosphosugar isomerases, phosphosugar binding proteins, and also in proteins that regulate the expression of genes involved in synthesis of phosphosugars (Teplyakov *et al*. 1998). ESA_03302, located on the opposite strand to the rest of the genes in this cluster, encodes an oligogalacturonate-specific porin protein (KdgM). ESA_03303 encodes a fructose-specific II component of the PTS system FruA, which converts fructose to fructose-1-phosphate. ESA_03304 encodes alpha-mannosidase involved in mannose degradation. ESA_03305 encodes a phosphomannose isomerase, which converts D-

mannose to β-D-glucose-6-phosphate. This cluster of genes was only present in *C. sakazakii* strains and absent in other *Cronobacter* species.

SR11 (ESA_03609 – ESA_03613) is a cluster of genes also involved in mannose metabolism. ESA_03609 encodes a putative β-galactosidase. Genes ESA_03610 and ESA_03612 encode genes involved in the *N*-acetylneuraminate and *N*-acetylmannosamine degradation pathway. ESA_03610 encodes a N-acetylmannosamine kinase and ESA_03611 encodes a N-acetylneuraminate lyase. Gene ESA_03612 encodes a transcriptional regulator from the GntR family. This cluster of genes was only present in *C. sakazakii* strains and absent in all other *Cronobacter* species.

SR 12 (ESA_04067 – ESA_04073) is a cluster of fimbrial genes. For more details see 3.5.1.11.

SR13 (ESA_04101 – ESA_04106) encodes genes that may be involved in the O-PS biogenesis. ESA_04102 encodes a glycosyltransferase involved in cell wall biogenesis, ESA_04103 encodes a putative O-antigen ligase and related enzymes. Genes ESA_04104 and ESA_04105 encode glycosyltransferases and ESA_04105 encodes a putative lipopolysaccharide heptosyltransferase III. This cluster of genes was present in all *Cronobacter* strains, but absent in *C. malonaticus.* ESA_04103 was only present in *C. sakazakii* strains. Note that this is not the O-antigen locus described by Mullane *et al*. (2008b).

Table 14. Genomic regions that differed between *Cronobacter* species.

ID	Gene loci	Description	Csak (n=7)	Cmal	Ctur	Cmuy
SR1	ESA_00257 - ESA_00258	Toxin/antitoxin RelE/RelB	YES	NO	NO	NO
SR2	ESA_01116 - ESA_01119	ABC-type multidrug efflux	YES	NO	NO	NO
SR3	ESA_01448 - ESA_01450	Fatty acid desaturases	YES	NO	YES	YES
SR4	ESA_02125 - ESA_02129	Ecotin (ESA_02129)	YES	YES	NO	NO
SR5	ESA_02538 - ESA_02542	Fimbriae	YES	NO	NO	NO
SR6	ESA_02544 - ESA_02547	β -glucosides metabolism	YES	YES	NO	YES
SR7	ESA_02549 - ESA_02553	Multidrug efflux system	YES	YES	YES	NO
SR8	ESA_02616 - ESA_02618	Mannosyl-D-glycerate uptake	YES	YES	NO	YES
SR9	ESA_02795 - ESA_02799	Fimbriae	YES	NO	NO	NO
SR10	ESA_03301 - ESA_03320	Mannose metabolism	YES	NO	NO	NO
SR11	ESA_03609 - ESA_03613	Mannose metabolism	YES	NO	NO	NO
SR12	ESA_04067 - ESA_04073	Fimbriae	YES	NO	NO	NO
SR13	ESA_04101 - ESA_04106	Cell wall biogenesis	YES	NO	YES[1]	YES[1]

[1] See main text for details about absence/presence of particular genes from this cluster

Csak = *C. sakazakii*; Cmal = *C. malonaticus* ; Cdub = *C. dublinensis*; Cmuy = *C. muytjensii*; Ctur = *C. turicensis*

3.5.1.7 Comparison of *C. sakazakii* BAA-894 to other Enterobacterial genera

The *C. sakazakii* BAA-894 genome sequence was compared to the genomes of *Citrobacter koseri* BAA-895, *Klebsiella oxytoca* VJSK009, *E. coli* K12 MG1655 and *Salmonella enterica* Typhimurium strain LT2, representing some of the most closely related genera to *Cronobacter*. This was done by Fred Long (Vaccine Research Institute of San Diego, USA). Using a threshold of identity of >85% in a 100 base window, 334 genes were present in all *Cronobacter* strains but absent or diverged in the four members of other genera. These genes included a cluster of T6SS genes (ESA_03943 - ESA_03948) which might be involved in virulence, and a putative palatinose operon (ESA_02709 - ESA_02715). Alpha-glucosidase activity, which has

been linked to palatinose metabolism, is considered as one of the major biochemical traits that distinguish *Cronobacter* from other related *Enterobacteriaceae*. The gene cluster ESA_02709 - ESA_02715 is identical to the *C. sakazakii* putative palatinose operon described in (Lehner *et al*. 2006). Genes ESA_02714 and ESA_02713 encode palatinose-binding proteins PalE and PalF. ESA_02712 encodes a palatinose ABC transporter membrane protein PalG, ESA_02711 encodes a palatinose ABC transporter ATP-binding protein palK, genes ESA_02710 and ESA_02709 encode an alpha-glycosidase PalQ and PalZ.

3.5.1.8 Genomic islands

Genomic islands in *C. sakazakii* BAA-894 were identified by IslandViewer (Langille, Brinkman 2009). Most of the mobile genetic elements identified in the genome of corresponded to the variable genomic regions discovered using our microarray data (Section 3.5.1.6). The regions that do not correspond to the variable regions encode hypothetical proteins or conserved uncharacterized proteins. The putative genomic islands are listed in Table 15, along with the corresponding variable genomic regions and their description. Note that the loci of the genomic islands may include additional genes or include less genes compared to the variable genomic regions.

Table 15. Putative genomic islands in *C. sakazakii* BAA-894

Genomic island location	Genomic region[1]	Specific region[2]	Description
ESA_00293 - ESA_00305	GR2		Uncharacterised proteins
ESA_00604 - ESA_00612	GR3		Putative prophage fragment 1
ESA_00622 - ESA_00634	GR3		Putative prophage fragment 1
ESA_00988 - ESA_00996	GR4		Putative prophage 1
ESA_01116 - ESA_01120		SR2	Toxin/antitoxin pair RelE/RelB
ESA_01179 - ESA_01188	GR5		O-antigen gene locus
ESA_01261 - ESA_01276			Hypothetical proteins
ESA_01610 - ESA_01623	GR6		Putative prophage 2
ESA_01625 - ESA_01645	GR6		Putative prophage 2
ESA_01177 - ESA_01804	GR7		Tellurite resistance cluster
ESA_02028 - ESA_02034			Hypothetical proteins
ESA_02303 - ESA_02341	GR10		Putative prophage fragment 2
ESA_02741 - ESA_02754	GR11		Putative prophage fragment 3
ESA_03009 - ESA_03106	GR12		Putative prophage 3
ESA_03428 - ESA_03434			Hypothetical proteins
ESA_04081 - ESA_04088			Hypothetical proteins
ESA_04243 - ESA_04250	GR14		Copper resistance cluster
ESA_04395 - ESA_04404			Hypothetical proteins

[1] Variable genomic regions defined in 3.5.1.6.1 and 3.5.1.6.2 .

[2] Species specific genomic regions defined in 3.5.1.6.3.

3.5.1.9 Invasion of brain microvascular endothelial cells

Because *Cronobacter* may cause fatal cases of neonatal meningitis, the status of genes identified in other organisms as associated with invasion of brain microvascular endothelial cells (BMEC) (*ibeA, ibeB, yijP* and *ompA*) in the sequenced strain was of particular interest (Prasadarao *et al*. 1996, Huang *et al*. 1999, Wang *et al*. 1999, Huang *et al*. 2001). The gene encoding OmpA was present in all *Cronobacter* strains. This protein is associated with the invasive ability of neonatal meningitic *E. coli* (Wang & Kim 2002) and *Cronobacter* (Nair *et al*. 2009, Kim *et al*. 2010). While genes *ibeA* and *yijP* produced no match in the reference strain *C. sakazakii* BAA-894, a match to *ibeB* was found. BLASTP analysis showed that this IbeB protein was homologous to the CusC protein from the copper and silver resistance cation efflux system *cusCFBA* which allows bacteria to invade BMEC (Franke *et al*. 2003). Microarray data revealed that the complete cation efflux system c*usA* (ESA_04242),

cusB (ESA_04241), *cusC* (ESA_04239), *cusF* (ESA_04240) and its regulatory genes *cusR* (ESA_04238) and *cusS* (ESA_04237) were present in strains associated with neonatal infections (*C. sakazakii* ATCC 29544^T, 701, 767, 696, *C. turicensis* and *C. malonaticus*), and absent in the other tested strains.

3.5.1.10 Comparison of *C. sakazakii* neonatal intensive care unit (NICU) outbreak strains with *C. sakazakii* type strain ATCC 29544^T

The genes that were shared by the three strains associated with *C. sakazakii* outbreaks in NICUs (BAA-894, 701 & 767) were compared to the *C. sakazakii* species type strain ATCC 29544^T; the later showed decreased virulence properties compared to strains 701 and 767 in earlier tissue culture studies (Townsend *et al*. 2008a). One hundred and forty-four genes present in the three NICU strains were absent in the type strain, 66 (46%) of them were in clusters of consecutive genes based on the annotation of BAA-894. In most of these clusters, genes encoding proteins associated with resistance to antibiotics and different forms of stress were identified, including multidrug efflux systems, genes involved in resistance to oxidative stress, and those with a putative function in resistance to metals. The complete list of genes present in NICU outbreak strains *C. sakazakii* BAA-894, 707 and 767 and absent in the *C. sakazakii* type strain ATCC 29544^T is provided in Appendix 3, genes of interest are collated below into four categories. The presence of these genes in *Cronobacter* strains tested by CGH is summarised in Table 19 in Section 3.5.1.15.

Category I. Genes encoding proteins associated with resistance to different forms of antibiotics: (i) a transcriptional regulator (ESA_01938) from the TetR family of protein repressors that control the level of susceptibility to hydrophobic antibiotics and detergents; (ii) a homologue of the CpmG protein involved in carbapenem resistance (ESA_pESA3p05435); (iii) a protein conferring resistance to antimicrobial

peptides Mig-14 (ESA_pESA3p05439); and (iv) a transcriptional regulator (ESA_pESA3p05448) involved in tetracycline resistance.

Category II. Genes encoding components of multidrug efflux systems: (i) a cationic drug transporter (ESA_01940) from the family of proteins that confer resistance to a wide range of toxic compounds; (ii) a complete bacterial ABC-transport system involved in active transport across the cytoplasmic membrane (ESA_01944–ESA_01946); and (iii) a variety of multidrug efflux components located on the plasmid pESA3 in BAA-894.

Category III. Genes involved in resistance to oxidative stress and genes with a putative function in resistance to metals: (i), a redox-sensitive transcriptional activator SoxR (ESA_00115); (ii) a glutathione S-transferase (ESA_00116); (iii) ADP-ribose pyrophosphatase involved in oxidative stress protection (ESA_pESA3p05446); (iv) an arsenate reductase (ESA_pESA3p05485); and (v) a predicted transcriptional regulator involved in mercurium resistance ESA_pESA3p05463.

Category IV. Other genes of interest include: (i) putative adhesins which are recognized as virulence factors in enteric bacteria (Campellone & Leong 2003) (ESA_00983–ESA_00986); (ii), the universal stress protein UspA (ESA_01955) which can enhance the rate of cell survival during prolonged exposure to stress conditions (Nystrom & Neidhardt 1994); (iii) a gene encoding a Type VI secretion lysozyme-related protein (ESA_02735); (iv) a gene encoding a predicted virulence SciE-type protein (ESA_02736) which affects the ability of bacteria to enter eukaryotic cells (Folkesson *et al*. 2002); and (v) genes involved in pilus assembly (ESA_03515 and ESA_03516).

3.5.1.11   Fimbriae

As the route of *Cronobacter* infection is probably through attachment and invasion of the intestinal cells, genes encoding surface appendages such as fimbriae are of interest. Four putative fimbriae clusters were identified in the genome of *C. sakazakii* BAA-

894, some of which were mentioned previously in (Healy, Huynh *et al.* 2009). These are Cluster 1 (ESA_01970 - ESA_01976), cluster 2 (ESA_02538 – ESA_02542), cluster 3 (ESA_02795 – ESA_02799) and cluster 4 (ESA_04067 – ESA_04073) (Table 16). Analysis of the CGH data showed that three of the four putative fimbrial clusters (Clusters 2, 3 and 4) were *C. sakazakii* specific, i.e. were classified as present or intermediate in all five strains of *C. sakazakii*, but were absent in *C. muytjensii, C. dublinensis, C. turicensis,* and *C. malonaticus.* Cluster 1 was present only in the reference strain and in *C. sakazakii* strain 2, which suggests that it is strain specific (Table 16). According to the BLASTP analysis, the genetic content of all fimbriae clusters was most similar to the type I chaperone/usher-assembled pilus system as defined in Kline *et al.* (2010). Genes encoding pilin FimA, chaperone FimC and usher FimD have been found in all four putative fimbriae clusters. Some degree of homology to the remaining components necessary for type-I pilus assembly (the minor tip fibrillum FimG and fimbrial adhesin FimH) was found among the remaining genes in the fimbrial clusters; these clusters may encode complete and functional pili. However, there was not enough evidence for a robust annotation of genes that may encode FimA, FimG and FimH, as these shared at least 30% amino acid identity with each other and with the corresponding proteins in the model organism for the type-I pili, *E. coli*. The presence of the putative fimbriae clusters in *Cronobacter* species according to the CGH data is summarised in Table 16.

Table 16. *C. sakazakii* BAA-894 fimbrial clusters and their presence in other *Cronobacter* strains

Cluster	Locus tag	Gene Product	Csak1	Csak2	Csak20	Csak701	Csak767	Csak696	Cmal	Ctur	Cmuy	Cdub
1	ESA_01970	Pilin FimA	-1	1	-1	-1	-1	-1	-1	-1	-1	-1
	ESA_01971	Pilin FimA	-1	1	-1	-1	-1	-1	-1	-1	-1	-1
	ESA_01972	Pilin FimA	-1	1	-1	-1	-1	-1	-1	-1	-1	-1
	ESA_01973	Pilin FimA	-1	1	-1	-1	-1	-1	-1	-1	-1	-1
	ESA_01974	Usher FimD	-1	1	-1	-1	-1	-1	-1	-1	-1	-1
	ESA_01975	Chaperone FimC	-1	0	-1	-1	-1	-1	-1	-1	-1	-1
	ESA_01976	Pilin FimA	-1	1	-1	-1	-1	-1	-1	-1	-1	-1
2	ESA_02538	Pilin FimA	1	1	1	1	1	1	-1	-1	-1	-1
	ESA_02539	Chaperone FimC	0	1	1	0	0	1	-1	-1	-1	-1
	ESA_02540	Usher FimD	1	1	1	1	1	1	-1	-1	-1	-1
	ESA_02541	Pilin FimA (FimH)	1	1	1	1	1	1	-1	-1	-1	-1
	ESA_02542	Putative minor component FimG	1	1	1	1	1	1	-1	-1	-1	-1
3	ESA_02795	Fimbrial protein	1	1	1	1	1	1	-1	-1	-1	-1
	ESA_02796	Pilin FimA	1	1	1	1	1	1	-1	-1	-1	-1
	ESA_02797	Usher FimD	1	1	1	1	1	1	-1	-1	-1	-1
	ESA_02798	Chaperone FimC	1	1	1	1	1	1	-1	-1	-1	-1
	ESA_02799	Putative fimbrial protein	1	1	0	0	0	0	0	-1	-1	-1
4	ESA_04067	Putative fimbrial protein	0	1	1	1	0	1	-1	-1	-1	-1
	ESA_04068	Fimbrial protein	-1	1	0	-1	-1	0	-1	-1	-1	-1
	ESA_04069	Fimbrial protein	0	1	1	1	1	1	-1	-1	-1	-1
	ESA_04070	Fimbrial protein	0	1	1	1	1	1	-1	-1	-1	-1
	ESA_04071	Usher FimD	1	1	1	1	1	1	-1	-1	-1	-1
	ESA_04072	Chaperone FimC	1	1	1	1	1	1	0	0	-1	-1
	ESA_04073	Fimbrial protein	0	1	1	1	1	1	-1	-1	-1	-1

1 for presence , 0 for intermediate status and −1 for absence/divergence of a gene. Csak1 = *C. sakazakii* ATCC 29544[T]; Csak2 = *C. sakazakii* ATCC 12868, Cmal = *C. malonaticus* ; Cdub = *C. dublinensis*; Cmuy = *C. muytjensii*; Ctur = *C. turicensis.*

3.5.1.12　　Iron uptake mechanisms

A list of known iron assimilation mechanisms was compiled from the published literature and their presence in *Cronobacter* strains was evaluated based on the available CGH data (Table 17).

All *Cronobacter* strains examined by CGH possess complete operons for enterobactin synthesis (*entABCDEFS*) and enterobactin receptor and transport (*fepABCDEG*), except *C. dublinensis*, in which *fepE* is absent. All *Cronobacter* species except *C. muytjensii* also possess a complete operon for aerobactin synthesis *iucABCD* and its receptor *iutA*. The operon for salmochelin synthesis was not present in the reference strain *C. sakazakii* BAA-894 used for the CGH and hence it was not possible to assess its presence in other *Cronobacter* strains (Table 17). The presence of the iron uptake genes in the newly sequenced *C. turicensis* was verified by BLASTP and corresponded to the results obtained by CGH.

Table 17. Iron uptake systems and their presence in *Cronobacter* strains

Gene	Homolog in *C. sakazakii* BAA-894 (BLAST)	Locus Tag	Csak1	Csak2	Csak20	Csak701	Csak767	Csak696	Cmuy	Cdub	Ctur	Cmal
Enterobactin synthesis - non ribosomal peptide synthesis pathway												
entA	2,3-dihydroxybenzoate-2,3-dehydrogenase	ESA_00799	1	1	1	1	1	1	0	1	1	1
entB	hypothetical protein ESA_00798	ESA_00798	1	1	1	1	1	1	1	1	1	1
entC	hypothetical protein ESA_00796	ESA_00797	0	1	0	0	0	0	1	1	1	1
entD	hypothetical protein ESA_02731	ESA_02731	1	1	1	1	1	0	0	0	1	0
entE	enterobactin synthase subunit E	ESA_02729	0	1	0	0	1	1	0	1	1	0
entF	enterobactin synthase subunit F	ESA_02727	1	1	0	0	1	0	0	1	1	1
entS	enterobactin exporter EntS	ESA_00794	1	1	1	1	1	1	1	1	1	1
Enterobactin receptor & transporters												
fepA	outer membrane receptor FepA	ESA_02730	1	1	1	1	1	1	1	1	1	1
fepB	iron-enterobactin transporter	ESA_00796	1	1	1	1	1	0	1	1	1	1
fepC	hypothetical protein ESA_00791	ESA_00791	1	1	1	1	1	0	1	1	1	1
fepD	iron-enterobactin transporter	ESA_00793	1	1	0	0	1	0	1	1	1	1
fepE	ferric enterobactin transport protein FepE	ESA_00459	1	1	1	1	1	1	0	-1	0	0
fepG	iron-enterobactin transporter permease	ESA_00792	1	1	0	1	1	0	1	1	1	1
Salmochelin synthesis												
iroB	salmochelin siderophore system	No match	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
iroC	salmochelin siderophore system	No match	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
iroD	enterobactin/ferric enterobactin esterase	No match	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
iroE	IroE protein	No match	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
iroN	outer membrane receptor FepA	ESA_01552	1	1	1	1	1	1	1	1	0	0
Aerobactin synthesis & receptor												
iucA	hypothetical protein ESA_pESA3p05547	ESA_pESA3p05547	1	1	0	0	1	1	-1	1	1	1
iucB	hypothetical protein ESA_pESA3p05548	ESA_pESA3p05548	1	1	0	0	0	0	-1	1	1	1
iucC	hypothetical protein ESA_pESA3p05549	ESA_pESA3p05549	0	0	0	0	0	0	-1	0	1	1
iucD	hypothetical protein ESA_pESA3p05550	ESA_pESA3p05550	1	1	0	0	0	1	-1	1	1	1
iutA	hypothetical protein ESA_pESA3p05551	ESA_pESA3p05551	1	1	1	1	1	1	-1	0	0	1
Other iron uptake genes												
feS	enterobactin/ferric enterobactin esterase	No match	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
fhuA	ferrichrome outer membrane transporter	ESA_03190	1	1	1	1	1	-1	-1	1	0	1
fhuB	iron-hydroxamate transporter permease	ESA_03187	1	1	1	1	1	0	1	1	1	1
fhuD	iron-hydroxamate transporter	ESA_03188	1	1	1	1	1	1	1	1	1	1
fpvA	ferrichrome outer membrane transporter	No match	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
fur	ferric uptake regulator	No match	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ibpA	heat shock protein IbpA	ESA_03960	1	1	1	1	0	1	1	1	1	1
ibpB	heat shock chaperone IbpB	ESA_03959	1	1	1	1	1	1	1	1	1	1

1 for presence , 0 for intermediate status and −1 for absence/divergence of a gene. No match indicaes absence in the reference strain *C. sakazakii* BAA-894. NA = not appropriate. Csak1 = *C. sakazakii* ATCC 29544[T]; Csak2 = *C. sakazakii* ATCC 12868, Cmal = *C. malonaticus* ; Cdub = *C. dublinensis*; Cmuy = *C. muytjensii*; Ctur = *C. turicensis*.

3.5.1.13 Type VI secretion systems

Five putative T6SS were identified in the genome of *C. sakazakii* BAA-894. The presence of the putative T6SS genes in other *Cronobacter* strains is summarised in Table 18.

Cluster 1 (ESA_00140 – ESA_00145, **GR1** as defined in 3.5.1.6.2) encodes most of the proteins that are conserved across different T6SS clusters (encoding a DotU homolog ESA_00141, Vgr homolog ESA_00142 and a putative lipoprotein from the VC_A0113 family ESA_00145). However, most T6SS clusters typically encode from 12 to 25 proteins (Filloux *et al*. 2008) and also encode a ClpV ATPase, which was not found in this cluster. In some instances, the genes encoding Vgr and DotU proteins are located outside the main T6SS locus, and their products might cooperate with proteins encoded in other loci. Intrestingly, the lipoprotein ESA_00145 was detected in the most virulent strains *C. sakazakii* 701, 767 and 696, whilst the other genes from Cluster 1 were absent from all other *Cronobacter* strains.

Cluster 2 (ESA_02035 – ESA_02041, **GR9** as defined in 3.5.1.6.2) includes genes encoding a homolog of a Vgr-type protein (ESA_02035), lipoprotein from the VC_A0113 family ESA_02038 and other genes homologous to T6SS proteins.

Cluster 3 (ESA_02735 – ESA_0240) contains genes encoding SciE-type protein (ESA_02736), Vgr-type protein (ESA_02739) and a protein homologous to phage gp7 protein (ESA_02740), all of which are frequently found in T6SS clusters. However, this cluster is adjacent to a prophage fragment described in 3.5.1.6.1 and due to the sequence similarities between T6SS and prophages it is difficult to conclude whether this cluster is a part of a T6SS.

Cluster 4 (ESA_03899 – ESA_03948, GR15 as defined in 3.5.1.6.2) is the longest and the most complete cluster of T6SS genes. Its 48 genes include all the components of T6SS typically conserved among different T6SS systems, such as genes encoding

Vgr-type proteins (ESA_03905 and ESA_03917), IcmF-type protein (ESA_03945), DotU-type protein (ESA_03946), ClpV ATPase (ESA_03921), SciE-type protein (ESA_03925), Ser/Thr protein phosphatase (ESA_03927) and Ser/Thr protein kinase (ESA_03920). This cluster may encode a complete and functional T6SS.

Cluster 5 (ESA_pESA3p05491 - ESA_pESA3p05506) located on the plasmid pESA3 in *C. sakazakii* BAA-894 encodes another putative T6SS cluster encoding some of the conserved T6SS proteins. ESA_pESA3p05494 encodes DotU-like protein, ESA_pESA3p05495 encoded a protein with a C-terminal extension with similarity to OmpA, a protein strongly associated with virulence properties of *Cronobacter* (Kim *et al*. 2010, Mittal *et al*. 2009b, Nair *et al*. 2009). ESA_pESA3p05497 encodes a ClpV ATPase and ESA_pESA3p05500 encodes a Vgr-like protein.

The Clusters 1 – 5 described here are putative T6SS clusters. It remains to be determined whether they encode functional T6SS or functional components of these.

.

Table 18. Type VI secretion system genes and their presence in *Cronobacter* species.

Locus Tag	Gene Product in *C. sakazakii* BAA-894	Csak1	Csak2	Csak20	Csak701	Csak767	Csak696	Cmuy	Cdub	Ctur	Cmal
ESA_00141	DotU-like protein	-1	-1	-1	-1	-1	1	-1	-1	-1	-1
ESA_00142	type VI secretion system Vgr family protein	-1	-1	-1	-1	-1	1	-1	-1	-1	-1
ESA_00143	type VI secretion protein, VC_A0110 family	-1	-1	-1	-1	-1	1	-1	-1	-1	-1
ESA_00144	type VI secretion protein, VC_A0111 family	-1	-1	-1	0	0	1	-1	-1	-1	-1
ESA_00145	type VI secretion lipoprotein, VC_A0113 family	-1	-1	-1	1	1	1	-1	-1	-1	-1
ESA_02035	Vgr-like protein	-1	-1	0	-1	-1	-1	-1	0	0	0
ESA_02036	hypothetical protein	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
ESA_02037	type VI secretion system Vgr family protein	0	-1	-1	-1	-1	-1	-1	0	0	-1
ESA_02038	type VI secretion lipoprotein, VC_A0113 family	1	1	-1	-1	-1	-1	-1	0	0	-1
ESA_02039	type VI secretion protein, VC_A0111 family	1	1	-1	-1	-1	-1	-1	0	0	-1
ESA_02040	type VI secretion protein, VC_A0110 family	1	1	-1	-1	-1	-1	-1	0	0	-1
ESA_02041	IcmF-related protein	1	1	-1	-1	-1	-1	-1	0	0	-1
ESA_02735	type VI secretion system lysozyme-related protein	-1	1	1	1	1	1	-1	0	0	0
ESA_02736	type VI secretion system protein, SciE, ImpE	-1	1	1	1	1	1	-1	0	0	-1
ESA_02737	hypothetical protein	-1	1	1	1	1	1	-1	-1	0	-1
ESA_02738	hypothetical protein, Rhs family protein	-1	1	1	1	1	1	-1	0	0	-1
ESA_02739	type VI secretion system Vgr family protein	-1	1	1	-1	-1	1	-1	0	0	-1
ESA_02740	hypothetical protein, phage gp7 homolog	-1	-1	0	0	-1	-1	-1	-1	1	-1
ESA_03899	Rhs family protein	-1	-1	1	0	0	-1	1	1	1	1
ESA_03900	hypothetical protein	-1	-1	-1	-1	-1	-1	0	0	0	-1
ESA_03902	hypothetical protein	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
ESA_03903	Rhs family protein	0	1	1	1	1	1	1	1	1	1
ESA_03904	hypothetical protein	1	1	1	1	1	1	1	1	1	0
ESA_03905	type VI secretion system Vgr family protein	1	1	1	1	1	1	1	1	1	1
ESA_03906	hypothetical protein	1	1	1	1	1	1	0	0	1	1
ESA_03907	hypothetical protein	1	1	1	1	1	1	1	1	1	1
ESA_03908	hypothetical protein	0	0	0	0	-1	0	-1	-1	1	1
ESA_03909	hypothetical protein	0	0	0	-1	-1	0	-1	-1	0	1
ESA_03910	hypothetical protein	0	1	1	0	0	-1	-1	-1	1	-1
ESA_03911	Rhs family protein	-1	1	-1	-1	-1	-1	1	1	1	-1
ESA_03912	hypothetical protein	-1	0	-1	-1	-1	-1	-1	1	-1	-1

Locus Tag	Gene Product in *C. sakazakii* BAA-894	Csak1	Csak2	Csak20	Csak701	Csak767	Csak696	Cmuy	Cdub	Ctur	Cmal
ESA_03913	Rhs family protein	0	1	0	0	0	-1	1	1	1	-1
ESA_03914	Ankyrin repeat	1	1	1	1	1	-1	1	0	0	-1
ESA_03915	hypothetical protein	1	1	1	1	1	-1	1	1	1	-1
ESA_03916	Uncharacterized protein conserved in bacteria	1	1	1	1	1	-1	1	1	1	-1
ESA_03917	type VI secretion system Vgr family protein	1	1	1	1	1	1	1	1	1	1
ESA_03918	Ankyrin repeat	1	1	1	1	1	1	0	0	0	0
ESA_03919	hypothetical protein	1	1	1	1	1	1	1	1	1	0
ESA_03920	Serine/threonine protein kinase	1	1	1	1	1	1	1	1	1	1
ESA_03921	type VI secretion ATPase, ClpV1 family	1	1	1	1	1	1	1	1	1	1
ESA_03922	type VI secretion protein, VC_A0111 family	1	1	1	1	1	0	1	1	1	1
ESA_03923	type VI secretion protein, VC_A0110 family	1	1	0	1	1	0	1	1	1	1
ESA_03924	type VI secretion system lysozyme-related protein	1	1	0	0	0	1	1	1	1	1
ESA_03925	SciE-like type VI secretion system protein ImpE	0	1	1	1	1	1	1	1	1	1
ESA_03926	hypothetical protein	1	1	1	1	1	1	1	1	1	1
ESA_03927	Serine/threonine protein phosphatase	1	1	1	1	1	1	1	1	1	1
ESA_03928	type VI secretion system FHA domain protein	1	1	1	0	1	1	1	1	1	1
ESA_03929	hypothetical protein	-1	0	1	1	1	-1	-1	-1	-1	-1
ESA_03930	hypothetical protein	-1	0	1	1	1	-1	-1	-1	-1	-1
ESA_03932	hypothetical protein	-1	0	1	0	1	-1	-1	-1	0	-1
ESA_03933	hypothetical protein	0	-1	1	1	0	0	0	-1	1	1
ESA_03934	type VI secretion system effector, Hcp1 family	0	1	1	1	1	1	1	1	1	1
ESA_03935	hypothetical protein	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
ESA_03936	hypothetical protein	-1	-1	-1	-1	-1	-1	0	0	-1	-1
ESA_03937	hypothetical protein	-1	-1	-1	-1	-1	-1	0	0	0	-1
ESA_03939	hypothetical protein	0	-1	-1	-1	-1	1	1	1	1	-1
ESA_03940	hypothetical protein	1	1	1	1	1	1	1	1	1	1
ESA_03941	type VI secretion protein, EvpB/VC_A0108 family	0	1	0	0	0	1	1	1	1	1
ESA_03943	type VI secretion-associated protein, ImpA family	1	1	1	1	1	1	1	1	1	1
ESA_03944	type VI secretion-associated protein, BMA_A0400 family	1	1	1	1	1	1	1	1	1	1
ESA_03945	type VI secretion protein IcmF	1	1	1	1	1	1	1	1	1	1
ESA_03946	type VI secretion system protein, DotU family	1	1	1	1	1	1	1	1	1	1
ESA_03947	type VI secretion protein, VC_A0114 family	1	1	1	1	1	1	1	1	1	1

Locus Tag	Gene Product in *C. sakazakii* BAA-894	Csak1	Csak2	Csak20	Csak701	Csak767	Csak696	Cmuy	Cdub	Ctur	Cmal
ESA_03948	type VI secretion lipoprotein, VC_A0113 family	1	1	1	1	1	1	1	1	1	1
pESA3p05491	type VI secretion protein, VC_A0107 family	0	1	1	1	1	1	-1	0	0	0
pESA3p05492	type VI secretion protein, EvpB/VC_A0108 family	1	1	1	1	1	1	-1	1	1	1
pESA3p05493	type VI secretion protein, VC_A0114 family	-1	1	-1	-1	-1	1	-1	0	-1	-1
pESA3p05494	type VI secretion system protein, DotU family	-1	1	-1	-1	-1	1	-1	0	-1	-1
pESA3p05495	Outer membrane protein, peptidoglycan-associated	-1	1	-1	-1	-1	1	-1	-1	-1	-1
pESA3p05496	type VI secretion system effector, Hcp1 family	1	1	-1	-1	-1	1	-1	1	1	-1
pESA3p05497	type VI secretion ATPase, ClpV1 family	-1	1	-1	-1	-1	1	-1	0	0	-1
pESA3p05498	hypothetical protein	-1	1	-1	-1	-1	0	-1	-1	-1	-1
pESA3p05499	hypothetical protein	-1	1	-1	-1	-1	1	-1	-1	-1	-1
pESA3p05500	type VI secretion system Vgr family protein	-1	1	-1	-1	-1	1	-1	-1	-1	-1
pESA3p05503	hypothetical protein, putative lipoprotein	1	1	-1	-1	-1	1	-1	-1	-1	-1
pESA3p05504	Uncharacterized conserved protein	1	1	-1	-1	-1	1	-1	-1	-1	-1
pESA3p05505	hypothetical protein	1	1	-1	-1	-1	1	-1	-1	-1	-1
pESA3p05506	type VI secretion lipoprotein, VC_A0113 family	1	1	1	1	1	1	-1	0	0	-1

1 for presence , 0 for intermediate status and −1 for absence/divergence of a gene. Csak1 = *C. sakazakii* ATCC 29544[T]; Csak2 = *C. sakazakii* ATCC 12868, Cmal = *C. malonaticus* ; Cdub = *C. dublinensis*; Cmuy = *C. muytjensii*; Ctur = *C. turicensis*

3.5.1.14 Lipopolysaccharide (LPS) and O-antigen locus

The chemical structure of the lipo-polysaccharide in three *Cronobacter* species has been determined and shown to be compositionally and structurally different (MacLean *et al*. 2009abc, Czerwicka *et al*. 2010). It is therefore predictable that the biosynthetic pathways will vary across the genus and this has been confirmed by CGH. The gene cluster corresponding to the O-antigen cluster described by Mullane *et al*. (2008b) (ESA_01177–ESA_01189) was examined. The genetic architecture of the O-antigen cluster in the sequenced *C. sakazakii* BAA-894 corresponds to the serotype O:1 as defined by Mullane and colleagues. According to our CGH data, two of the genes in this region, *gal*F (ESA_01177) and *rfb*B (ESA_01178) are conserved among all *Cronobacter* strains tested except *C. sakazakii* 696, whereas the rest of the genes from the O-antigen locus are highly divergent; its genes were not sufficiently similar to be detected by microarray hybridization in any other *Cronobacter* strains. This correlates with the findings of Mullane *et al*. (2008b) who showed that both serotypes O:1 and O:2 had the two genes *gal*F and *rfb*B in common, whereas the rest of the gene cluster content differed between the two serotypes.

3.5.1.15 Other virulence determinants

In order to find other *Cronobacter* genes related to virulence, the UniProtKB/Swiss-Prot (http://www.ebi.ac.uk/uniprot/database) was searched for virulence factors in the sequenced genomes of *C. sakazakii* BAA-894 and *C. turicensis* z3032 using the keyword 'virulence' and 'adhesion'. The resulting hits were searched in IMG-JGI (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi) to obtain their annotation or to identify their homologs in *C. sakazakii* BAA-894. The presence/absence of the putative virulence factors in other *Cronobacter* species and other *C. sakazakii* strains was assessed based on the CGH microarray data. The results are summarized in Table 19. Among the virulence-related genes, proteins involved in the T6SS, fimbriae assembly and belonging to the LPS and O-antigen cluster were found. These systems were examined in more detail earlier. The remaining virulence factors are described below.

Figure 30 shows the gene map of the sequenced *C. sakazakii* BAA-894 with annotated selection of variable regions, virulence determinants and MLST loci.

- ESA_02934 encodes an outer membrane autotransporter barrel domain involved in the type V autotransporter pathway implicated in virulence of Gram-negative bacteria (Henderson & Nataro, 2001).

- ESA_03365 encodes a membrane protein that affects hemolysin expression under anaerobic conditions (Cox *et al*. 2000). Hemolysin expression is a recognised virulence factor in some pathogens, as bacteria producing hemolysins can inactivate immune system by lysing white blood cells.

- ESA_02516 encodes a haemagglutinin family N-terminal domain found in a variety of haemagglutinins and haemolysins.

- ESA_02576 encodes a pectin methylesterase which catalyses the de-esterification of a major component of a plant cell wall, pectin, into pectate and methanol. It has been suggested that pathogenic bacteria may use similar mechanisms to invade eukaryotic cells to plant tissue invasion.

- ESA_03608 encodes an outer membrane autotransporter barrel domain involved in type V secretion pathway, homologous to bacterial adhesin AidA, which mediates bacterial attachment to a broad spectrum of human and other mammalian cells in some diarrheagenic *E. coli* strains and other Gram-negative bacteria. The expression of AidA was also found to dramatically enhance biofilm formation by *E. coli* on abiotic surfaces (Sherlock *et al*. 2004).

- ESA_01908 encodes a putative virulence factor SrfB of unknown function conserved in bacteria. In *Salmonella enterica*, SrfB is activated by the two-component regulatory system SsrAB encoded on the *Salmonella* pathogenicity island SPI-2 (Worley *et al*. 2000).

- ESA_02278 encodes an integral membrane protein MviN. In *Salmonella enterica,* deletion of *mviN* reduces virulence properties of the bacterium in a mouse model of typhoid disease and MviN is therefore considered a virulence factor. However, MviN is also an essential protein in many non-pathogenic bacteria (Carsiotis *et al*. 1989).

- ESA_02216 encodes a response regulator phoP from the two-component signal transduction system phoP-phoQ. The adjacent gene ESA_02217 encodes the cognate signal transduction histidine kinase phoQ.

- ESA_00878 encodes a surface lipoprotein VacJ. In *Shigella flexneri,* VacJ is required for the intercellular spreading of the pathogen (Suzuki *et al*. 1994).

- ESA_03350 encodes a protein involved in polysaccharide export. The homolog of this protein was found in the *E. coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid which may be implicated in virulence (Stevenson *et al*. 1996).

- ESA_00102 encodes a putative hemolysin activation/secretion protein HlyA that is secreted across both the cytoplasmic and outer membranes of Gram-negative bacteria.

Table 19. Putative virulence factors in *C. sakazakii* BAA-894 and their presence in other *Cronobacter* strains

Locus tag	Gene Product	Csak1	Csak2	Csak20	Csak701	Csak767	Csak696	Cmuy	Cdub	Ctur	Cmal
ESA_00102	Putative hemolysin activation/secretion protein HlyA	1	1	1	1	1	1	0	0	0	0
ESA_00115	Redox-sensitive transcriptional activator SoxR	-1	1	1	1	1	1	1	1	1	1
ESA_00116	Glutathione S-transferase	-1	1	1	1	1	1	0	1	1	0
ESA_00614–17	Restriction endonucleases	-1	-1	-1	1	1	-1	-1	-1	-1	-1
ESA_00878	VacJ, intercellular spreading of *Shigella flexneri*	1	1	1	1	1	1	1	1	1	1
ESA_00983–86	Putative adhesins	-1	-1	1	1	1	1	-1	-1	-1	-1
ESA_00997	*eae*-like adhesion protein, attaching and effacing	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
ESA_01116 –19	ABC-type multidrug efflux system	0	0	1	0	1	1	-1	-1	-1	-1
ESA_01908	Putative virulence factor SrfB	1	1	1	1	1	1	1	1	1	1
ESA_01938	TetR family, hydrophobic antibiotics and detergents	-1	1	1	1	1	1	0	0	1	1
ESA_01940	Cationic drug transporter	-1	1	1	1	0	1	0	0	0	1
ESA_01944–46	ABC-transport system	-1	1	1	1	1	1	1	1	1	1
ESA_01955	Universal stress protein UspA	-1	1	1	1	1	1	-1	-1	-1	-1
ESA_02129	Ecotin	0	0	0	0	0	1	1	1	1	1
ESA_02216	Regulator phoP	1	0	1	0	0	1	1	1	1	0
ESA_02217	Signal transduction histidine kinase phoQ	1	1	1	1	1	1	1	1	1	1
ESA_02278	Integral membrane protein MviN	1	1	1	1	1	1	-1	-1	0	-1
ESA_02516	Haemagglutinin family N-terminal domain	1	1	1	1	1	1	-1	1	1	1
ESA_02549 – 53	ABC multidrug transport system	1	1	1	0	0	1	1	1	1	1
ESA_02576	Pectin methylesterase	0	1	0	1	1	1	-1	0	0	0
ESA_02934	Outer membrane autotransporter barrel domain	1	1	1	1	1	1	0	1	0	0
ESA_03350	Protein involved in polysaccharide export	1	1	1	1	1	1	0	1	0	1
ESA_03365	Membrane protein affecting hemolysin expression	1	1	1	1	1	1	1	1	1	1
ESA_03608	Outer membrane autotransporter barrel domain	1	1	1	1	1	1	0	0	0	0
ESA_04238 - 42	Copper uptake CusABCFR	1	-1	-1	1	1	1	-1	-1	1	1
ESA_04248–55	Copper resistance	1	-1	-1	-1	-1	1	-1	-1	1	1
ESA_pESA3p05435	CpmG protein involved in carbapenem resistance	-1	1	1	1	1	0	-1	0	0	1

Locus tag	Gene Product	Csak1	Csak2	Csak20	Csak701	Csak767	Csak696	Cmuy	Cdub	Ctur	Cmal
ESA_pESA3p05439	Mig-14 (resistance to antimicrobials)	-1	1	1	1	1	0	-1	-1	0	-1
ESA_pESA3p05446	ADP-ribose pyrophosphatase , oxidative stress protection	-1	1	1	1	1	0	-1	-1	0	0
ESA_pESA3p05448	Transcriptional regulator, tetracycline resistance	-1	1	1	1	1	1	-1	0	1	1
ESA_pESA3p05463	Predicted transcriptional regulator involved in mercurium resistance	-1	1	1	1	1	1	-1	0	1	0
ESA_pESA3p05485	Arsenate reductase	-1	1	1	1	1	1	-1	-1	1	0

1 for presence , 0 for intermediate status and −1 for absence/divergence of a gene. Csak1 = *C. sakazakii* ATCC 29544[T]; Csak2 = *C. sakazakii* ATCC 12868, Cmal = *C. malonaticus* ; Cdub = *C. dublinensis*; Cmuy = *C. muytjensii*; Ctur = *C.turicensis*.
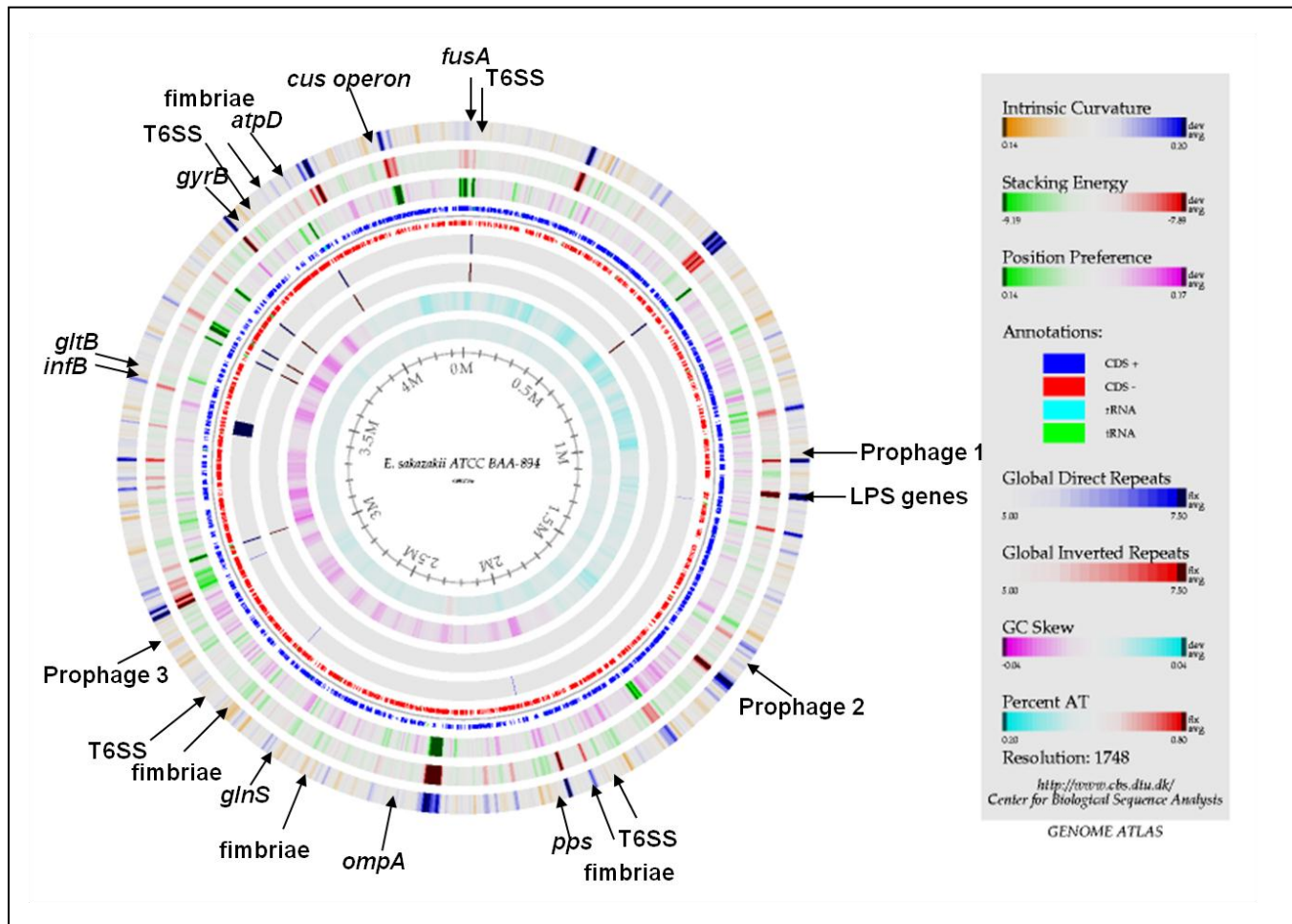
Figure 30. Gene atlas of *C. sakazakii* BAA-894 with selected annotated regions (prophages, T6SS, fimbrial clusters, LPS locus, MLST loci, other virulence determinants). The gene atlas was obtained at (http://rast.nmpdr.org/rast.cgi).

3.5.2 Construction of *Cronobacter* mutants via "Gene Doctoring" recombineering method

3.5.2.1 Choice of strains and genes

In a study examining the genes involved in bacterial meningitis, *ibeB* was identified as a gene required for *E. coli* K1 invasion of the blood-brain barrier (Huang *et al*. 1999). It was shown that the *ibeB* deletion mutant had <1% of the invasive ability of the wild-type *in vitro* and also showed significantly decreased invasivity in the *in vivo* model of *E. coli* meningitis compared to the wild-type strain. The cloning of *ibeB* back into the deletion mutant completely restored the invasive ability of the strain. BLASTP analysis showed that the protein IbeB was homologous (72% sequence identity, 99% query coverage) to the CusC protein from the copper and silver resistance cation efflux system in *C. sakazakii* BAA-894. Microarray data showed that the complete copper efflux system cusCFBA (ESA_04239 – ESA_04242) and its regulatory genes *cusR* (ESA_04238) and *cusS* (ESA_04237) were present in strains associated with neonatal infections (*C. sakazakii* ATCC 29544^{T}, 701, 767, 696, *C. turicensis* and *C. malonaticus*), and absent in all other tested strains, which had no association with clinical cases of *Cronobacter*-induced illness. These findings make the genes belonging to the copper resistance operon strong candidates for virulence determinants contributing to the invasiveness of *Cronobacter*, but this hypothesis needs to be verified.

To investigate whether *cusC* and other genes from the copper resistance operon have an effect on human brain microvascular endothelial cells invasion by *Cronobacter*, targeted gene deletion via the lambda-Red method described in 3.4.2 was attempted. Genes *cusA*, *cusB* and *cusC* were chosen for the first round of the gene deletions. *CusA* (3.2 kb) encodes an integral membrane proton-driven transporter, which forms a two-channel pump for the export of metal ions and concurrent import of protons (Munson *et al*. 2000), *cusB* (1.3 kb) encodes a membrane fusion protein (Saier *et al*. 1994) and *cusC* (1.4 kb) encodes an outer membrane factor (Dinh *et al*. 1994).

3.5.2.2 Construction of *cusA, cusB* and *cusC* deletion mutants via the "gene doctoring" method

The sequenced strain *C. sakazakii* BAA-894 and one of the most virulent strains of our collection associated with a fatal meningitis case, *C. sakazakii* 767 (Caubilla-Barron *et al*. 2007), were chosen as target organisms for the gene deletions. These strains were tested for antibiotic sensitivity by plating on LB plates containing 35 µg/ml chloramphenicol, 50 µg/ml kanamycin or 200 µg/ml ampicillin. Both strains were susceptible to all three antibiotics at this concentration. The same *Cronobacter* strains were tested for growth on LB plates containing 5 and 10 % (w/v) sucrose. All *Cronobacter* strains were able to grow at these concentrations of sucrose. The kanamycin cassette was amplified by PCR from pDOC-K (3.4.2.3) using primers containing regions homologous to the extremities of the kanamycin cassette and the DNA flanking the genes *cusA, cusB* and *cusC*. The PCR yielded the expected 1.4 kb products (Figure 31), which were then purified as described in 3.4.2.5.
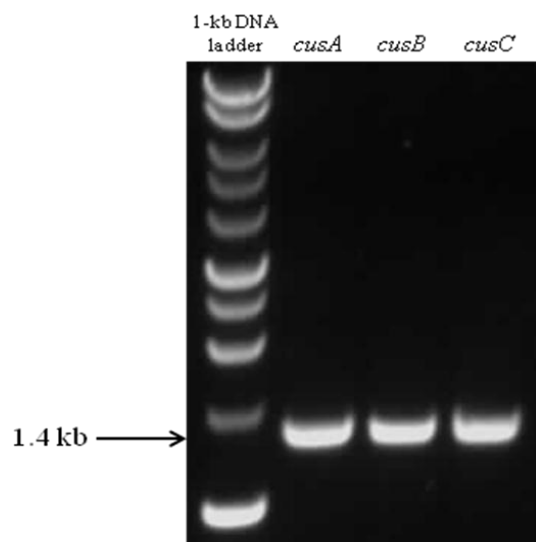
Figure 31. PCR amplification of the 1.4-kb kanamycin cassette from pDOC-K using primers specified in 3.4.2.3.

The linearization of pDOC was verified by gel electrophoresis (3.4.2.4). The linearized plasmid travelled more slowly in the gel than the supercoiled closed circular complex (CCC) uncut plasmid and appeared to have a greater size, as shown on Figure 32. Note that a supercoiled marker was not used and the 1-kb DNA ladder was used to illustrate the difference in the conformation of the plasmid.
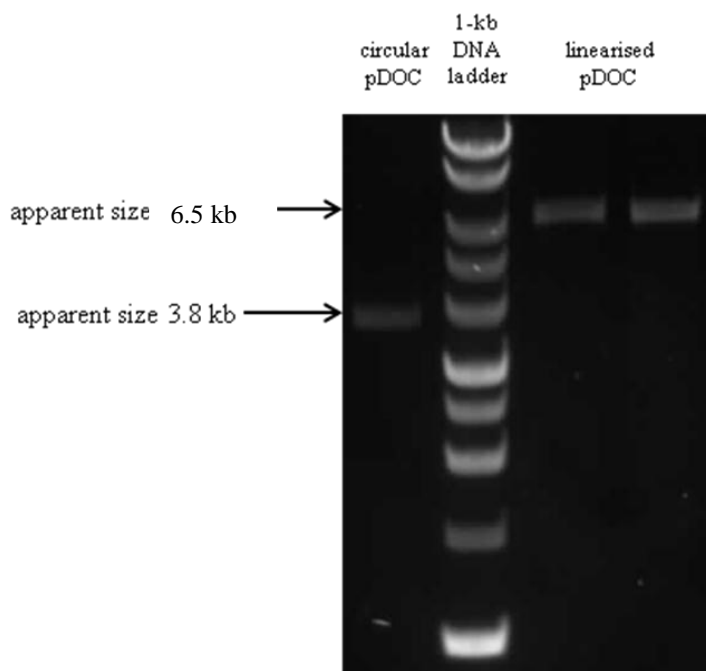
Figure 32. Digestion of pDOC with *Eco*RV.

The kanamycin cassette flanked by regions homologous to the DNA adjacent to the genes of interest was cloned into pDOC linearized by *Eco*RV as described in 3.4.2.7 and the success of the cloning was verified by PCR using the primers annealing to the DNA flanking the multicloning site in pDOC as detailed in 3.4.2.8. The empty vector would give product of ca 200 bp, whereas the vector with the insert would give a product of 1.6 kb (Figure 33). The first cloning attempts were unsuccessful, with ampicillin and kanamycin resistant transformants yielding no PCR products when screened using primers pDEXF and pDEXR. It was concluded that one of the possible problems might be degradation of ATP in the ligase buffer. When the ligase buffer

was replaced by buffer with added ATP, faint bands were visible for all cloned products. The manufacturer of T4 DNA ligase (Promega) recommends ligation at 15-20°C for 4-18 hours for blunt end products. However, incubation at 10°C significantly improved the yield of the ligation reaction. Using this modified protocol, higher quantities of the cloned products were obtained (Figure 33).
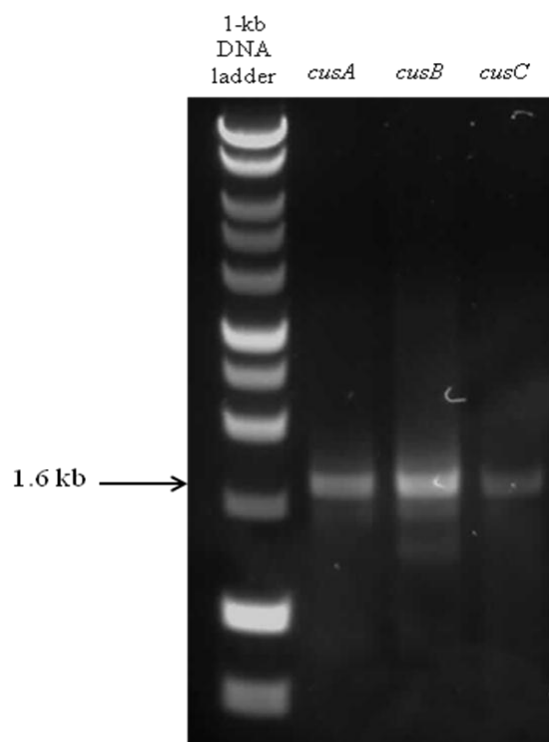
Figure 33. PCR verification of cloning of kanamycin cassette into pDOC.

To amplify the required quantity of the plasmid constructs, chemically competent One Shot Top 10 *E. coli* cells were transformed with the pDOC plasmid carrying the kanamycin cassette flanked by the short DNA sequences homologous to regions upstream and downstream of the genes to be knocked out, as described in 3.4.2.9. Plasmid DNA was isolated from the transformed cell cultures (3.4.2.2). Electrocompetent *Cronobacter* cells prepared according to (3.4.2.10) were transformed with the recombineering plasmid, pACBSRSceIV, and subsequently, the *Cronobacter* cells carrying pACBSRSceIV were made electrocompetent using the

171

same protocol and transformed with the pDOC donor construct created in 3.4.2.7, carrying regions of homology flanking genes *cusA, cusB* and *cusC*.

Cronobacter colonies carrying the recombineering plasmid pACBSRSceIV and the donor plasmid were checked for sucrose sensitivity and ampicillin resistance; all colonies were sensitive to sucrose and ampicillin resistant. Those selected isolates were grown in the presence of arabinose (3.4.2.12) to induce homologous recombination.

After arabinose induction, the mutants were plated on kanamycin and ampicillin LB agar plates. True mutants should acquire kanamycin resistance by homologous recombination with the donor plasmid at the desired loci, and become ampicillin sensitive due to loss of the donor plasmid pDOC. However, virtually all screened mutants retained ampicillin resistance, indicating that no donor plasmid loss occurred. In order to screen *Cronobacter* colonies for recombination events at the desired loci, PCR primers annealing to the sites adjacent to the regions used for generating the kanamycin cassette PCR construct were used (3.4.2.13). A PCR reaction with such primers will amplify a product corresponding to the wild-type gene, whereas PCR on a mutant with a deletion in a particular gene will amplify a product corresponding to the replacement of the gene with the kanamycin cassette. However, the control PCR yielded only the products expected in wild-type strains (3.2 kb for *cusA*, 1.2 kb for *cusB* and *cusC*).

Similar problems were encountered by the team of external supervisor Dr. Hobman using this strategy. Dr. Hobman recommended using the pACBSR plasmid used by (Herring *et al*. 2003) which does not carry the I-Sce recognition site and is therefore not self-cleaving, instead of the pACBSRSceIV plasmid. The protocol was repeated using the plasmid pACBSR. Electrocompetent *Cronobacter* cells were transformed with pACBSR and subsequently with the pDOC donor construct created in 3.4.2.7. Prior to arabinose induction, colonies were checked for sucrose sensitivity and ampicillin resistance; all colonies were sensitive to sucrose and ampicillin resistant. After arabinose induction, the transformants were plated on to kanamycin and ampicillin LB agar plates. The vast majority of the transformants were again both ampicillin and kanamycin resistant, indicating that the donor plasmid had not been

lost, and kanamycin resistance probably originated from the kanamycin cassette on the donor plasmid. Some of the transformant colonies were screened by the PCR as described in 3.4.2.13. Typically, 50 colonies for each gene deletion were screened in each round of experiment, but the results were negative. The arabinose induction was repeated, but with the same negative results.

4. DISCUSSION

4.1. Comparative genomic hybridisation method

Whole-genome sequencing information is usually only available for a relatively limited number of bacteria and to date (Oct 2010), 1344 bacterial genomes have been sequenced. Comparative genomic hybridization (CGH) is a useful alternative to sequencing that can be used to study sequence variation at the strain and species level. With respect to *Cronobacter*, only *Cronobacter sakazakii* BAA-894 had been sequenced when this project was designed; another sequence of *C. turicensis* z3032 became available after the completion of this project, in September 2009. In fact, the *Cronobacter* CGH microarray was designed only shortly after the taxonomic reclassification of "*Enterobacter sakazakii*" in 2008. Not only does our study present the first whole-genome analysis of all five known *Cronobacter* species; it is also one of the first steps towards studying virulence associated traits of this opportunistic pathogen in the context of whole genomes of diverse *Cronobacter* species.

The analysis of microarray CGH data is not a straightforward matter and is further complicated by the total lack of a standardised method for analysis of such data. The literature is dominated by methods designed for microarray expression analysis, and so are the software tools available. The most widely used application of CGH is copy-number variation analysis in eukaryotic cells, which is conceptually different from bacterial CGH analysis, and most CGH software tools are therefore inappropriate for microbial genome comparisons. It is not uncommon that research groups develop their own statistical tools for CGH data analysis to overcome this problem. The three main stumbling blocks of CGH analysis are: i) choice of software that would

conveniently manage large data for hundreds of thousands of probes and facilitate conversion between different data formats, ii) choice of software that would summarize probe intensity data for each gene, and iii) choice of method and software to determine "cut-off" values that would separate absent or highly divergent genes from present genes. It is now widely accepted that the empirically determined "constant" or "naive" cut-offs used in the early days of bacterial CGH are not appropriate, especially if less phylogenetically related strains are used.

As it is difficult to infer the relationship between sequence identity and hybridisation signal, the accuracy of classification of genes into present/absent categories will always depend on the choice of the cut-off determination method. It is inevitable that a dataset of several thousand genes will contain a number of false-positive and false-negative results. The dynamic cut-off determination method GACK (Genomotyping Analysis by Charles Kim, Kim *et al*. 2002) chosen for our CGH data analysis overcomes the problem of constant cut-offs by assigning two cut-off values for each individual hybridisation, based on statistical parameters of the hybridisation distribution. These cut-offs mark the transition from the absent to intermediate and from intermediate to present categories. Our interpretation of the trinary system is that genes classified as absent or present are classified with high accuracy, whereas genes in the intermediate category are genes with a degree of sequence divergence that could not be assigned a status with high accuracy. By opting for GACK and the trinary classification which includes an "intermediate" gene category, we attempted to minimise the number of falsely assigned genes. The GACK method was criticised by some authors for introducing complacency by creating a third category of genes with not enough evidence to decide (Carter *et al*. 2008). We believe, however, that this method was extremely useful for our CGH experiment involving a diverse bacterial population that included five different *Cronobacter* species.

Excluding our study, there has not been published any whole-genome comparative study examining *Cronobacter* spp. using the microarray technology. However, microarray-based comparative genome indexing (CGI) with an array covering 10% of the *C. sakazakii* BAA-894 genome (276 CDS) was used to compare 78 *Cronobacter* isolates obtained from food, environmental and clinical sources (Healy

et al. 2009). The authors used constant cut-offs for classification of *Cronobacter* genes and trinary classification of genes into present/unknown/absent categories.

4.2. Bacterial species, genome plasticity and mobile genetic elements

The CGH data distribution confirmed the taxonomic revision of *Cronobacter* spp. which separated the organisms into five different species (Iversen *et al*. 2008). *C. turicensis* LMG 23827^T, *C. malonaticus* LMG 23826^T, *C. dublinensis* LMG 23823^T and *C. muytjensii* ATCC 51329^T clearly displayed a greater level of sequence divergence from *C. sakazakii* BAA-894 compared to the other strains classified as *C. sakazakii* species, which was in accordance with the DNA-DNA hybridization data presented by Iversen *et al*. (2008). Of the 4,382 unique annotated gene sequences represented on the microarray 43.3% (1,899) were classified as present in all ten *Cronobacter* strains including five *Cronobacter* species, and represented the core genome defined by the most stringent criteria. The majority of these genes are predicted to encode cellular essential functions such as energy metabolism, biosynthesis, DNA, RNA and protein synthesis, cell division and membrane transport. 80.9% (3,547) genes were classified as present or intermediate in all strains. In a microarray-based comparative genome indexing study by Healy *et al*. (2009), the authors reported that 72.5% (200/276) of the CDS represented on the microarray were present in all *Cronobacter* and defined a core set of *Cronobacter* genes. The species core gene set of 72.5% is indeed consistent with the generally accepted notion that members of the same bacterial species should share at least 70% of their genome content (Stackebrandt *et al*. 2002). However, the value obtained by Healy and colleagues does strictly depend on the choice of the 10% of the genome content used for the CGI study, as well as the choice of the constant cut-offs used for classification of the genes.

Whole-genome clustering based on the gene category assignment by GACK correlates with the taxonomic reclassification of *Cronobacter* proposed by Iversen *et al*. (2008). *Cronobacter* strains formed two distinct phylogenetic clusters, with all *C. sakazakii* strains clustering together, and all other *Cronobacter* species *C. malonaticus, C. turicensis, C. dublinensis* and *C. malonaticus* forming a second cluster. Within the *C. sakazakii* cluster, strains 701 and 767 were the most closely related and clustered together with strain 20. Previously, strains 701 and 767 were shown to belong to the same pulsed-field gel electrophoresis restriction digestion type (Caubilla-Barron *et al*. 2007). Although the clinical details of the source of *C. sakazakii* strain 20 are unknown, the strain belongs to MLST sequence type 4 (as do 701 and 767), which is a stable clone of *C. sakazakii* isolated from both powdered infant formula and clinical sources (Baldwin *et al*. 2009). *C. sakazakii* strain ATCC 29544^{T} (species type strain) formed a separate branch within the *C. sakazakii* cluster. This correlated well with our in-house observations of different growth and heat resistance behaviour of this strain compared to other *C. sakazakii* strains. In future, using more strains of each species, as well as using representatives of diverse origin would be advantageous. Our results also correlated well with the clustering based on the CGI analysis by Healy *et al*. (2009) who showed that the *C. sakazakii* strains formed a large cluster, separate from the other *Cronobacter* species. However, Healy and colleagues reported that three strains identified as *C. malonaticus* based on phenotypic criteria described in Iversen *et al*. (2008) clustered together with the *C. sakazakii* strains, which shows the close genetic relatedness of *C. sakazakii* and *C. malonaticus* that may causes difficulties in distinguishing between these two species, or potentially the danger of misclassification of *C. sakazakii/C. malonaticus* by the current phenotypic criteria.

There are defined criteria to distinguish between bacterial species, such as 70% threshold of DNA–DNA relatedness and the requirement that each species have a distinct phenotype that would allow its differentiation (Stackebrandt *et al*. 2002). However, no theoretical framework that would define species as phylogenetic entities that would truthfully reflect the bacterial diversity exists (Rossello-Mora & Amann 2001, Konstantinidis *et al*. 2006). As formulated by Staley (2009) the current definition of species has several drawbacks, such as its dependence on an artificial

threshold that has been defined based on phenotypic properties, the fact that it does not take evolutionary concepts into account, and its dependence on a single type strain that may undergo mutational modification during passaging. Classification of bacterial species is further complicated by the vast amount of mobile DNA present in bacterial genomes, which blurs the species concept and makes it challenging to understand the evolutionary relationship between strains and species. The CGH analysis revealed that a significant part of the variable gene pool of *Cronobacter* was due to gene acquisition via mobile DNA elements. Typing of such mobile elements, or islands, which frequently encode virulence determinants, might be in fact more important than trying to identify different species of *Cronobacter*.

Extensive genomic diversity and genomic plasticity is not uncommon among both Gram-negative and Gram-positive bacteria and other CGH studies examining the diversity of a particular species or genus report similar results to our study. Previous studies examining the genomes *E. coli* and *Salmonella* suggest that gene acquisition and gene loss play a major role in the evolution of the bacterial genomes (McClelland *et al*. 2000, Anjum *et al*. 2003). A comparative genomics study of pathogenic *S. aureus* strains revealed that ca 22% of the genome was composed of accessory genes, which were grouped in large chromosomal regions of difference, were often associated with phage genes, integrases or transposases, and most of them contained putative virulence factors. Dorrell *et al*. (2001) observed that at least 21% of the genes present in the sequenced *C. jejuni* strain appeared dispensable in other *C. jejuni* strains tested by CGH.

The CGH data also revealed a surprising genetic diversity among the strains of the same species, *C. sakazakii.* An interesting observation made by Konstantinidis & Tiedje (2004) was that organisms with larger genomes tend to show larger intra-species genomic differences compared to bacteria with smaller genomes. The authors hypothesized that species with larger genomes are more metabolically versatile and therefore able to exploit a larger number of niches, which in turn shapes the content of the organisms' genomes. This is certainly true for *C. sakazakii,* which is an ubiquitous organism successfully colonising various different environments, from water, soil,

178

plant material, foods, herbs and spices, to insects and larvae, livestock and finally, human hosts.

Fifteen clusters of "accessory" genes that were present in less than half of the *Cronobacter* strains were identified. At least five of these regions were of putative prophage origin and included potentially active prophages, as well as putative prophage fragments. The variable genomic regions also included: i) four T6SS gene clusters, ii) the O-antigen gene locus used for serotyping of *Cronobacter* (Mullane *et al*. 2008b), iii) a tellurite resistance cluster, iv) a fimbriae cluster, v) a copper resistance cluster, vi) regions with hypothetical proteins or uncharacterised conserved proteins with no homologs with known function. Some of these "accessory" regions have probably been acquired by HGT. For example, the tellurite cluster was entirely absent from all other *Cronobacter* strains except the sequenced strain, but was found to be carried on plasmid pK29 of *Klebsiella pneumoniae* strain NK29, plasmid pEC-IMPQ of *Enterobacter cloacae*, plasmid R478 of *Serratia marcescens* and plasmid pAPEC-O1-R of *Escherichia coli*. Moreover, the cluster encoded two putative transposases, which is further evidence of horizontal gene transfer. The sequenced *C. sakazakii* strain probably acquired the tellurite resistance cluster via a plasmid integration.

According to Staley (2009) "the bacterial species are in a state of dynamic genetic flux". Bacteria must be well adapted to the lifestyle in their niche, but at the same time, be equipped to move into another niche should the conditions change (Staley 2009). This means that genomes that readily undergo modifications such as gene acquisition and gene loss resulting in new combinations of genes and potentially new metabolic capabilities or increased fitness have a considerable advantage in rapidly changing environments. In whole genome alignments, the vast majority of the non-homologous sequences are typically attributed to mobile DNA that includes integrative plasmids, transposons and prophages (Canchaya *et al*. 2004). This highlights the importance of genome plasticity in general, and the genetic mechanisms underlying it, for the survival and propagation of bacterial species.

With the advent of genome sequencing and comparative genomics, the role of mobile genome elements (MGE) in the plasticity of bacterial genomes has been highlighted,

and it has been accepted that MGEs have more importance in evolution than previously acknowledged. Action of MGEs like conjugative plasmids or prophages can indeed turn a harmless bacterium into a deadly pathogen, as in the cases of cholera (Karaolis *et al*. 1999) and diphtheria (Uchida *et al*. 1971, Karaolis *et al*. 1999). While the role of MGE in genome plasticity is undisputable, the contribution of horizontal and vertical elements in evolution of bacteria is the subject of vivid scientific debate. It is likely that on a shorter timescale of years and decades needed for emergence of hyper-virulent epidemic strains, HGT has much more impact on the evolution of pathogenicity than vertical evolution, which operates on a longer time-scale.

4.3. Virulence factors in the variable genomic regions and other virulence determinants

The abundance of putative and proven virulence factors in the accessory genomic elements has been well documented. Proteases, adherence proteins, iron uptake, intracellular survival, capsule formation, antiphagocytic genes, as well as type III, IV and VI secretion system genes are frequently associated with these regions (Ho Sui *et al*. 2009). Type III and IV secretion systems, toxins and adherence factors including the corresponding effector proteins have been found to be over-represented in genomic islands (Ho Sui *et al*. 2009). Virulence factors associated with variable genomic regions are important targets for typing of virulent strains, as well as hotspots of genome plasticity that can clarify the mechanisms pathogenesis of a particular bacterium. Most of the described accessory regions contain suspected or proven virulence factors. The genes in GR1, GR9 and GR15 are involved in a type VI secretion system, a newly described mechanism for protein transport across the cell envelope of Gram-negative bacteria that can increase adherence to epithelial cells (Schlieker *et al*. 2005, Mougous *et al*. 2006). GR3 contains four genes (ESA_00614–

ESA_00617) that are homologous to a restriction-modification gene cluster (api49–api52) in the *Yersinia pseudotuberculosis* pathogenicity island (YAPI) (Collyn *et al*. 2004). As these genes were present in *Cronobacter* strains 701 and 767 isolated from two neonates that died as a result of infection by *Cronobacter* during an outbreak in France (Caubilla-Barron *et al*. 2007) and are absent in all other strains tested by CGH, these genes may be important virulence factors contributing to the pathogenicity of *Cronobacter*.

Bacterial lipopolysaccharide (LPS) is a conserved structure essential for virtually all Gram-negative organisms playing an important role in the activation of the innate immune system. Consequently, the structure and and biosynthesis of LPS has been a focus of intense research. Characterisation of LPS structure and consequently O-antigen can be important in developing identification schemes based on serotyping. Overall, the biochemical synthesis of lipid A is a highly conserved process; however, investigation of the lipid A structures of various organisms shows an impressive amount of variation. The LPS is one of the few structural features of *Cronobacter* which has been investigated and it is known that it varies across the *Cronobacter* spp. In *C. sakazakii* and *C. malonaticus*, the LPS are composed of various branched polymers, whereas they are unbranched in *C. muytjensii*. In *C. sakazakii* BAA-894 (MacLean *et al*. 2009a) LPS is a branched polymer of pentasaccharide units, in *C. sakazakii* strain 767 it is also a branched polymer but of a repeating heptasaccharides (Czerwicka *et al*. 2010). In *C. malonaticus* LPS is also a branched pentasaccharide unit (MacLean *et al*. 2009c) whereas in *C. muytjensii* LPS is a linear unbranched pentasccharide polymer (MacLean *et al*. 2009b). These considerable differences correspond with the lack of sequence conservation in GR5 as revealed in the microarray analysis. The individual genes encoding these differences in enzymology have yet to be assigned. The CGH analysis identified the region GR 5 (ESA_01179–ESA_01189) encoding the O-antigen and lipopolysaccharide (LPS) *rfb* gene cluster as one of the regions that varied significantly between *Cronobacter* strains. The genetic architecture of the O-antigen cluster in the sequenced *C. sakazakii* strain corresponds to the serotype O:1 as defined by (Mullane *et al*. 2008b). According to our CGH data genes *gal*F (ESA_01177) and *rfb*B (ESA_01178) are conserved among all

Cronobacter strains, whereas the rest of the genes from the O-antigen locus are highly divergent. This correlates with the findings of Mullane and colleagues, who showed that both serotypes O:1 and O:2 had the two genes *gal*F and *rfb*B in common, whereas the rest of the genetic content differed between the two serotypes. As the genes located in the variable regions of the *rfb* cluster were predicted to encode glycosyltransferases and sugar modifying enzymes, the genetic heterogeneity of this cluster may cause the variations in the LPS structure of *Cronobacter* strains. Homologues of genes required for the processing and assembly of the O-antigen from O-subunits, *wzy* (ESA_01185) and *wzx* (ESA_01183) were found in the sequenced *C. sakazakii* BAA-894 but were not classified as present in any other *Cronobacter* strains including five *C. sakazakii* strains according to our CGH data. Conversely, Healy *et al*. (2009) reported that 23/60 *C. sakazakii* strains shared the *wzy* and *wzx* genes. The difference in the results might be simply due to the fact that a limited collection of *C. sakazakii* strains was used in our study compared to Healy *et al*. (2009). Also, as it is known that the genes involved in the O-antigen synthesis are highly variable between strains and species, but the general structure of the O-antigen is conserved between Gram-negative bacteria, the variability in the presence of *wzy* and *wzx* might be a result of the gene sequence divergence rather than of the gene absence. Mullane *et al*. (2008b) suggested that the *wehC* and *wehI* loci are polymorphic and specific to the serotypes O:1 and O:2, respectively. A homolog of *wehI* was not found in *C. sakazakii* BAA-894 (O:1) which is in line with the finding of (Mullane *et al*. 2008b) that this locus is specific to serotype O:1. However, despite the fact that Mullane suggested using the *wehC* locus for distinguishing between the two serotypes O:1 and O:2, a homolog of *wehC* was not found in *C. sakazakii* BAA-894 by BLASTP, only a protein with a match on 33% of the length (ESA_01187) was found. Because of the key role of the LPS in pathologenesis of *Cronobacter* disease, further examination of the LPS biosynthetic genes in a diverse population of *Cronobacter* strains and species is warranted.

Type VI secretion systems (T6SS) are recently described secretion systems in bacteria. Although T6SS are implicated in bacterial virulence, they do not correspond

to the classical definition of virulence factors, as they may be equally present in commensal non-pathogenic organisms. Five putative T6SS clusters were identified in the genome of *C. sakazakii* BAA-894 with complex patterns of presence in other *Cronobacter* strains. The longest and most complete cluster 4 of 48 genes (ESA_03899 – ESA_03948) encodes all the components of T6SS typically conserved among different T6SS systems. Interestingly, ESA_pESA3p05494 from cluster 5 encodes DotU-like protein with a C-terminal extension with similarity to OmpA, a protein strongly associated with virulence properties of *Cronobacter* (Mittal *et al*. 2009b, Nair *et al*. 2009, Kim *et al*. 2010). It is also of interest that the lipoprotein ESA_00145 from Cluster 1 was detected in the most virulent strains *C. sakazakii* 701, 767 and 696, whilst the other genes from Cluster 1 were absent from all other *Cronobacter* strains. It would be highly speculative to make any conclusions about the functionality of the putative T6SS clusters in *Cronobacter* based on the sequence and CGH data alone. In other bacteria T6SS involve different genomic organisation and various combinations of genes. No consensus in the model for the assembly of the T6SS has been defined yet and very little is known about the protein interactions of the various components of the system (Cascales 2008, Filloux *et al*. 2008). The experimental study of T6SS is challenging, as T6SS genes are usually not induced in laboratory conditions and upregulation of most T6SS genes clusters depends on contact with the host cell or growth inside the host cell.

An interesting theory states that the variability in composition and sequence of T6SS is due to their specialisation for different purposes and different cellular targets. Many of the bacterial species with multiple T6SS clusters that are not paralogous, have multiple hosts or diverse environments (Jani & Cotter 2010). The presence of diverse T6SS loci in *Cronobacter* would be in line with the organism's ubiquitous presence, and its lifestyle adapted to survival in many different plant and animal species, including humans. Some evidence also suggests that T6SS can limit bacterial replication of virulence by changing the state of virulence to commensal or mutualistic relationship. It seems that in *Salmonella*, T6SS induction limits the bacterial load in host cells and intracellular replication, which leads to bacterial

persistence rather than macrophage death and damage to the host (Parsons & Heffron 2005). It would be extremely interesting to obtain similar data for *Cronobacter*.

As the route of infection is probably through attachment and invasion of the intestinal cells by *Cronobacter*, surface appendages such as fimbriae are of interest. There is some evidence of fimbriae production in *Cronobacter,* although this important virulence trait has not been studied extensively. Adegbola & Old (1983) reported production of a mannose-sensitive hemagglutinin associated with type I fimbriae and presence of fimbriae in two strains (n=4) *Cronobacter.* By combining the CGH analysis with searching for known genes associated with fimbriae assembly, we identified four putative type I chaperone-usher fimbriae clusters in the genome of *C. sakazakii* BAA-894. Three of the four putative fimbrial clusters (Cluster 2, 3 and 4) were specific to *C. sakazakii* strains, and genes belonging to Cluster 1 were only present in 1/6 *C. sakazakii* strains. None of the putative fimbrial genes were detected in the other *Cronobacter* species. This suggests that there is considerable intra- and inter-species variability of fimbrial operons, which may account for some differences in attachment of *Cronobacter* to eukaryotic targets. Our results were in accordance with the results of Healy *et al.* (2009) who reported that Cluster 2 and two genes ESA_04070 and ESA_04071 from Cluster 4 were only present in *C. sakazakii* isolates, and that genes ESA_01970 and ESA_01976 from Cluster 1 were only present in 14/60 *C. sakazakii* strains. Similar to our study, in several microarray genomic comparisons of *Salmonella* strains, a great degree of variation was observed for fimbrial operons, phage genomes and pathogenicity islands (Porwollik *et al.* 2004, Anjum *et al.* 2005, Scaria *et al.* 2008). There was not enough evidence for a robust annotation of all components of functional type-I fimbriae. This is because the proteins FimA, FimG and FimH shared at least 30% amino acid identity with each other and with the corresponding proteins in the model organism for the type-I pili, *E. coli*. Hence we can only speculate whether the putative fimbrial clusters encode for a complete set of proteins necessary for functional fimbriae. The presence of functional fimbriae in *Cronobacter* as well as their involvement in attachment and invasion of human brain cells needs to be investigated experimentally.

The ability of bacteria to sequester copper from the environment and to tightly regulate copper homeostasis due to the high toxicity of free copper has been recognised. It has been proposed that copper efflux is critical for virulence in some pathogenic strains of *E. coli*, but very little is still known about the role of copper uptake systems in pathogenicity. The CGH microarray data highlighted two clusters of genes related to copper resistance in the *Cronobacter* genomes. Firstly, the copper and silver resistance efflux system *cusCFBA* and its regulatory two-component signal transduction system *cusRS*. These were present in strains associated with neonatal infections (*C. sakazakii* ATCC 29544^T, 701, 767, 696, *C. turicensis* and *C. malonaticus*), and absent in the other tested strains, which had no association with clinical cases of *Cronobacter*-induced illness. *CusC* was found to be homologous to a gene *ibeB* required for invasion of the blood-brain barrier in *E. coli* K1 (Huang *et al*. 1999), which made *cusC* and potentially the whole *cus* operon strong candidates for virulence factors implicated in meningitis. Secondly, the putative *cop* operon was identified as one of the genomic regions absent in more than half of the *Cronobacter* strains tested by CGH. As the *cop* operon was not present in the highly invasive fatal strains *C. sakazakii* 707 and 767, we hypothesized that the *cus* operon and *cusC* in particular has a more important role in the virulence of *Cronobacter*. An interesting question is whether the increased invasive properties of the strains possessing *cusC* are due to the regulation of copper uptake itself or whether the outer membrane factor *cusC* facilitates penetration of HBMEC by a functionality not-related to copper uptake.

To test the hypothesis that the *cus* operon is implicated in the invasiveness of *Cronobacter*, we attempted to delete the genes *cusA, cusB* and *cusC* from the *C. sakazakii* BAA-894 strain using λ-Red targeted gene deletion developed by Lee *et al*. (2009). This method has been designed for use in pathogenic strains of *E. coli* and was not successful in generating deletion mutants in *Cronobacter*. The possible reasons include: i) low recombination efficiency during arabinose induction or recombination events induced at other loci, ii) insufficient arabinose induction and therefore cleavage of donor plasmid, iii) too short overhangs used as the regions of homology, iv) action of lambda phages. The most likely explanation is that a problem in the cleavage of the donor plasmid occurred, which led to a lack of the linear

substrate needed for homologous recombination. This could be explained by improper arabinose induction or a malfunction of the I-SceI endonuclease.

Two thirds of all gamma-proteobacteria and low GC Gram-negative bacteria harbour prophages (Brussow *et al*. 2004) and there is an increasing body of evidence that phages play a pivotal role in the diversification of bacterial species. Some phages can carry additional cargo genes, which are not required for the phage cycle and are suspected or proven virulence factors (Brussow *et al*. 2004). In this study, three putative prophage genomes and three putative prophage fragments were identified in the genome of *C. sakazakii* BAA-894 and the presence of the prophage genes was evaluated in ten other *Cronobacter* strains. These putative prophages contain genes which are not similar to any other known prophage genes. Given the lifestyle constraints of phages, the genes probably encode proteins necessary for the phage life cycle or increasing their fitness. It is possible that the unknown genes encode novel proteins or virulence-related factors. So where do we look for putative virulence factors that have been brought to the bacterial genomes by prophages that could have a significant impact on the organism's pathogenicity ? As fixation of whole prophages does not make much evolutionary sense for bacteria in the long term, and prophages tend to be lost from the bacterial genomes as fast as acquired, the prophage genes surrounding the virulence factor are likely to be deleted (Canchaya 2004). The regions carrying virulence factors or fitness factors that increase the survival of the bacterium in its host are therefore most likely isolated phage genes or phage fragments conserved in virulent strains of the species. Further research into these regions is warranted.

There were a number of remarkable observations about the putative *Cronobacter* prophage genomes that might be linked with virulence. Firstly it is the region in Prophage 3 that was homologous to genes of the *S. enterica* serovar Typhi Vi type II phage E1 which uses virulence-associated capsular antigen as entry. This region was present in *C. turicensis* and partially present in *C. sakazakii* strains 707 and 767 associated with a fatal outbreak. Secondly, A group of restriction endonucleases

homologous to genes api49, *api50, api51* and *api52*, from the *Yersinia pseudotuberculosis* adhesion pathogenicity island (Collyn *et al*. 2004) were also present in the hypervirulent *C. sakazakii* strains 701 and 767 and absent in other strains. Thirdly, a cluster of three O-antigen conversion genes that may alter the bacterium's serotype was found in putative prophage 3.

It has been suggested that prophage genomes consist of combinations of non-homologous regions and functional modules which have a considerable variation in the population (Aubrey *et al*. 1992). This mosaic structure results from "recombinational promiscuity" – a term used by (Frost *et al*. 2005) - which describes frequent recombination with other prophages and other mobile elements within the same bacterial host. This is consistent with our microarray data, which show the conservation of distinct phage genetic modules between the *Cronobacter* strains, rather than conservation of complete phages. The prophage mosaicism causes difficulties in the classification of phages, which often share very little homology, although some conservation in the structure of modules can be observed. For that reason, the putative prophages 2 and 3 were classified as lambda-like due to the presence of regulatory genes typical for lambda phages, but we made no attempts to further classify the *Cronobacter* phages. The prophages that were unique to the sequenced strain (prophage 3, prophage fragments 2 and 3) are likely to be recent acquisitions. It is probable that the genome of *C. sakazakii* BAA-894 contains more prophages or prophages fragments than are described here. The annotation of prophage genes is complicated by the low sequence similarity of equivalent phage functions (Frost *et al*. 2005), especially if these prophage-encoded genes are integrated in other bacterial genomes. In many of the variable regions including prophages, genes with no homology to known proteins were found. These "ORFan genes" atypically abundant in MGEs, could be rapidly evolving genes potentially implicated in virulence. On the other hand, many of them are probably pseudogenes with no physiological function. It is important to note that the simple presence of any gene in a bacterial genome does not mean expression of a functional protein, and that includes phage-encoded factors. It has been documented that in lambdoid phages, most genes are not transcribed and constitute a passive DNA (Canchaya *et al*. 2004).

Expression of phage genes may be also regulated by environmental factors. For example, Shiga-like toxins in *E.coli* 0157 are expressed when iron becomes a growth limiting factor, signalling that the bacterium is in its host (Wagner & Waldor 2002); and upregulation of prophage genes has been observed in *P. aeruginosa* upon change from planktonic to biofilm growth (Whiteley *et al*. 2001). This highlights the importance of experimental testing of any putative virulence factors discovered by genome studies. The putative prophages identified in the genomes of *Cronobacter* strains via analysis of the CGH data may contribute to the spread and pathogenicity of this species by harbouring virulence of fitness factors. It is important to bear in mind, however, that the sequenced *C. sakazakii* BAA-894 strain that served as a template for the microarray design, was isolated from a can of formula associated with a fatal outbreak of meningitis and not a clinical isolate. Although tissue culture and *in vitro* assays showed that *C. sakazakii* BAA-894 is more virulent than other *Cronobacter* strains including the *C. sakazakii* type strain, other, more virulent *Cronobacter* strains indeed exist. It would be therefore extremely useful to obtain sequence information of the prophages and prophage fragments carried by those strains, which would harbour virulence factors with a higher probability.

An attempt to identify additional virulence factors in *Cronobacter* was made by comparing the genomes of the three strains associated with *C. sakazakii* outbreaks in NICUs (BAA-894, 701 & 767) with the *C. sakazakii* species type strain ATCC 29544[T], which showed decreased virulence properties in tissue culture studies (Townsend *et al*. 2008a). Among the one hundred and forty-four genes present in the three NICU strains and absent in the type strain, genes encoding proteins associated with resistance to different forms of stress including multidrug efflux systems, proteins involved in resistance to oxidative stress, and those with a putative function in resistance to metals were identified. Along with the putative virulence factors identified as part of prophages, fimbriae and T6SS, these genes will serve as potential targets for virulence studies of *Cronobacter*, that will help us to understand the organism's complex role in pathogenicity of human disease. We acknowledge that a major limitation of our study is the fact that the *Cronobacter* microarray was based on

a sequence of *Cronobacter* strain that was not a clinical isolate directly associated with a death, and that other potentially more virulent strains of *Cronobacter* exist. To our defence, at the time of the design of this study, *C. sakazakii* BAA-894 was the only *Cronobacter* strain sequenced. Please note that in the course of this work, the genome of another *Cronobacter* strains (*C. turicensis* z3032) was published. For a selection of genes (fimbrial clusters, iron acquisition systems, some prophage genes), our CGH results were verified by BLAST analysis and correlated well with the published sequence data. Just before this thesis was submitted the following paper was released (Stephan *et al*. 2010). This short paper compliments our previous studies and does not conflict with any statements in the thesis.

Genomic regions that are specific to a particular species can reflect acquisition of new metabolic potential or adaptation of the species to a particular niche. Note that due to the availability of a single sequence of *Cronobacter* at the time of the design of the microarray, only genes from the sequenced *C. sakazakii* BAA-894 were present on the microarray. We could not therefore access the status of additional genes that are present in other *Cronobacter* strains. However, thirteen regions differentially distributed between the *Cronobacter* species (regions that were present in *C. sakazakii* BAA-894 and some but not all *Cronobacter* species) could be determined. As there were unique combinations of these regions for each species, these regions could be used for molecular typing of *Cronobacter* species after thorough verification of the method in a larger collection of strains. The analysis highlighted three genomic regions related to mannose metabolism SR8, SR10 and SR11. Clusters SR10 and SR11, which were only present in *C. sakazakii* strains, encode some of the component of the mannose degradation pathway, such as alpha-mannosidase, phosphomannose isomerase, N-acetylmannosamine kinase and acetylneuraminate lyase. As these genes encode enzymes from the beginning of the mannose degradation pathway, their absence (or absence of close homologs) may result in the inability of *Cronobacter* species other that *C. sakazakii* to process mannose. Earlier studies did report production of acid from mannose in *Cronobacter*. However, these studies were done with "*E. sakazakii*" strains, which were likely to contain a mixture of *Cronobacter* species, with *C. sakazakii* being the most prevalent species. The analysis also

highlighted a homologue of a gene encoding ecotin (ESA_02129), which was intermediate in *C. sakazakii* strains and in *C. malonaticus*, but absent in *C. muytjensii, C. dublinensis* and *C. turicensis.* It has been shown that ecotins from species that come into contact with the mammalian immune system like *E. coli, Y. pestis* and *P. aeruginosa* protect bacteria against the effects of neutrophil elastase (Eggers *et al.* 2004). ESA_02129 might encode a virulence factor that promotes resistance of *Cronobacter* during active infection of the host. Other regions of interest identified in this analysis included fimbriae clusters, ABC-type multidrug efflux system, fatty acid desaturase cluster and a toxin-antitoxin RelE/RelB system.

4.4. Concluding remarks

It is extremely important to experimentally verify the results of any sequence analysis data. The CGH technology, as well as homology searches in gene databases do not take into account the possible role of adaptive mutation such as occurrence of SNPs or other events that may render a gene non-functional, whilst preserving the sequence homology to its functional counterparts. It is also important to note that although genome sequencing and CGH provides information about the gene content, it cannot reveal the true phenotypic potential of the organism as it does not provide enough evidence about the expression of a particular trait. Further experiments, such as in-depth analysis of sequence data to reveal possible mutations leading to inactivation of genes, gene expression studies as well as laboratory verification of expression of virulence factors are warranted.

It should be also noted that CGH technology cannot reliably distinguish between absence of sequence and sequence divergence. For this reason, it could have been useful to include another sequenced bacterium closely related to *Cronobacter* on the microarray to define the relationship between sequence identity and fluorescence signal. However, the most closely related sequenced bacterium to *Cronobacter* was at the time when the microarray was designed *Citrobacter koseri*, and including a more

distantly related organism on the microarray might present some more challenges in the data analysis. Another limitation of the CGH technology is the DNA information contained on the array itself. Typically the array is designed using the genome sequenced of a single strain, or in case of a pan-array, available sequences from multiple strains. In any case, the useful information that we can obtain from this type of experimental design does not include potentially interesting genes present in other, perhaps more virulent strains. The *Cronobacter* microarray was designed when a single complete genomic sequence of *Cronobacter* was available, originating from a can of powdered formula linked to the Tenessee fatal outbreak (Himelright *et al*. 2002). However, the strain was not a clinical isolate of the baby that died, and whether it contained the complete genetic arsenal that caused the infant's death is highly speculative. Another obstacle for CGH data interpretation is the fact that it is uncertain how much sequence conservation (and the corresponding hybridisation signal) is required for preservation of gene function. In some instances even a difference of a single nucleotide in the coding sequence of a gene can lead to inactivation of the gene product. On the other hand, a considerable advantage of the *Cronobacter* array was the use of overlapping probes spanning the whole *C. sakazakii* BAA-894 genome. To classify a gene into the present/absent/intermediate category, a median of often more than 50 probes was calculated, which decreased the influence of any aberrant intensity signals. Despite the possible limitations, comparative genomic hybridisation remains a powerful tool for examination of the gene content on the whole genome level, which would be otherwise impossible or very costly.

Cronobacter is a two-faced organism. It can turn from a harmless, commensal organism into a deadly pathogen, causing life-threatening disease with lifetime impairments. Is its virulence a direct result of the expression of its virulence genes or is it the interaction with its host and activation of the host defence system that triggers the final form of the disease? Killing the host by a severe form of sepsis or meningitis within a few days after infection with limited spread of the pathogen, is clearly not the best survival strategy for any pathogenic organism. *Cronobacter*-induced disease is a biological mistake, occurring when an ideal commensal relationship gets out of hand. It only occurs in an unfortunate combination of conditions, such as when the lack of

competing gut flora and low acidity in a neonatal stomach promote the growth of *Cronobacter* that would have otherwise been eliminated, or when the protection of the immune system is weakened in the immune-compromised or elderly. By identifying the regions of difference between *Cronobacter* strains and species, highlighting putative virulence factors and compiling comparative genomic information about key bacterial structures and virulence mechanisms, our study provides ground for a better understanding of this remarkable accidental pathogen.

5. GENERAL CONCLUSIONS

Cronobacter species are dangerous pathogens that can cause life-threatening disease including necrotizing enterocolitis, septicaemia and meningitis, with premature neonates and infants at particular risk. Several outbreaks of *Cronobacter* infection have been linked to contamination of powdered infant formula at neonatal intensive care units, although it is becoming increasingly clear that there may be other equally important sources of contamination. The phenotypical diversity of *Cronobacter* sp. (formerly '*Enterobacter sakazakii'*) as well as the vast variation in virulence potential of *Cronobacter* isolates have been known to researchers for several decades. However, the genetic basis of the diversity of *Cronobacter* sp. as well as the basis of the variation in their virulence potential have not been known. Sequencing of the first *Cronobacter* strain, *C. sakazakii* BAA-894, uncovered the organism's full genetic arsenal and opened promising prospects in bacterial pathogenomics.

The first part of this project was dedicated to the investigation of growth behaviour of *Cronobacter* species and other selected *Enterobacteriaceae* in powdered infant formula products. We determined the generation times of 28 bacterial strains across 21 bacterial species in casein- and whey-dominant infant formula at temperatures 21, 27, 37, 41 and 44°C, including 11 *Cronobacter* strains representing five *Cronobacter* species, and lag times of the same organisms at 21^{0}C. The data revealed considerable diversity in growth profiles of the *Cronobacter* genus, as well as remarkably fast growth of the organisms at a wide range of temperatures including 44°C which is inhibitiory for most *Enterobacteriaceae*. This data was useful in the context of control of *Cronobacter* infection in neonates through providing scientific basis for revision of the current formula reconstitution guidelines.

In the second part of this project, we used the sequenced *C. sakazakii* BAA-894 genome to construct the first whole-genome *Cronobacter* microarray. The comparative analysis of isolates representing all five *Cronobacter* species using a

193

whole-genome 384,030 oligonucleotide tiling microarray revealed that the main genetic features that distinguished the *Cronobacter* strains were putative prophages and other variable gene clusters, most of which have been probably acquired by horizontal gene transfer. This highlighted the role of horizontal gene acquisition in the plasticity of the bacterial geenomes, as well as in evolution of pathogenicity, as most of these clusters contained proved or putative virulence factors. The genomic regions described in this work such as the type VI secretion systems, fimbriae, LPS, O-antigen loci and metal uptake systems deserve further in-depth analysis. The putative virulence factors identified in these regions or elsewhere in the *Cronobacter*'s genome require further study and experimental verification, including targeted gene deletions, expression analysis and tissue culture assays. The comparison of ten *Cronobacter* strains including strains from a fatal outbreak in France (Caubilla-Barron *et al.* 2007) revealed many features of *Cronobacter* that may contribute to the organism's pathogenicity such as a copper/silver resistance cluster associated with invasion of BMEC, efflux pumps or adhesins, which were unique to the three *Cronobacter* species associated with neonatal infections.

During the course of the project the diversity of the *Cronobacter* genus was further investigated by multiple-locus sequence typing (Baldwin *et al.* 2009), which confirmed the taxonomic revision by Iversen *et al.* (2008) and highlighted oversights in classification of *Cronobacter* strains based solely on biotypes. The *Cronobacter* genus is still not well defined, and further work including MLST and CGH analysis with a broader selection of strains representing *Cronobacter* species of different origin and virulence potential should considerably improve the definition of the *Cronobacter* genus, as well as classification methods. Although *Cronobacter* is an ubiquitous organism present in a wide range of environments, MLST data suggest that some clinical isolates are highly clonal. More research needs to be done to clarify the clonal distribution of clinical isolates, as well as to correlate genetic markers with the actual virulence potential of a particular strain. For this, typing of pathogenicity islands or other virulence determinants described in this work and serotyping could prove particularly useful.

The ultimate goal of this study is to increase the knowledge base that will lead to a better understanding of this remarkable organism, which is ubiquitous and mostly harmless, and yet capable to cause fatal infectios in neonates and major infections in adults. This should lead to improved recognition of this emergent pathogen, and ultimately to better identification methods that would be able to identify the strains with the greatest virulence potential. To date, there is no available sequence of a hypervirulent *Cronobacter* strain available. This limits any studies using the *C. sakazakii* BAA-894 sequence information for the discovery of virulence factors of *Cronobacter*. The future offers an exciting prospect of sequencing more *Cronobacter* strains of each *Cronobacter* species, as well as important hypervirulent clinical strains, which will provide material for comparative genomics on a larger scale, and help us gain insight into *Cronobacter*'s fascinating genome dynamics, the role of mobile genome elements on the evolution of virulence, its possible plant origin, and many more mysteries that are currently left to our speculation.

REFERENCES

Aarts, M.G., Keijzer, C.J., Stiekema, W.J. & Pereira, A. 1995, "Molecular characterization of the CER1 gene of arabidopsis involved in epicuticular wax biosynthesis and pollen fertility", *The Plant Cell,* vol. 7, no. 12, pp. 2115-2127.

Adamson, H. & Rogers, J.R. 1981, "*Enterobacter sakazakii* meningitis with sepsis", *Clin. Microbiol. Newsl,* vol. 3, pp. 19-20.

Adegbola, R.A. & Old, D.C. 1983, "Fimbrial haemagglutinins in *Enterobacter* species", *Journal of general microbiology,* vol. 129, no. 7, pp. 2175-2180.

Aisen, P., Leibman, A. & Zweier, J. 1978, "Stoichiometric and site characteristics of the binding of iron to human transferrin", *The Journal of Biological Chemistry,* vol. 253, no. 6, pp. 1930-1937.

Alakomi, H.L., Skytta, E., Saarela, M., Mattila-Sandholm, T., Latva-Kala, K. & Helander, I.M. 2000, "Lactic acid permeabilizes gram-negative bacteria by disrupting the outer membrane", *Applied and Environmental Microbiology,* vol. 66, no. 5, pp. 2001-2005.

Al-Holy, M.A., Castro, L.F. & Al-Qadiri, H.M. 2010, "Inactivation of *Cronobacter* spp. (*Enterobacter sakazakii*) in infant formula using lactic acid, copper sulfate and monolaurin", *Letters in Applied Microbiology,* vol. 50, no. 3, pp. 246-251.

Al-Nabulsi, A.A., Osaili, T.M., Shaker, R.R., Olaimat, A.N., Ayyash, M.M. & Holley, R.A. 2009, "Survival of *Cronobacter* species in reconstituted herbal infant teas and their sensitivity to bovine lactoferrin", *Journal of Food Science,* vol. 74, no. 9, pp. M479-84.

Amalaradjou, M.A., Hoagland, T.A. & Venkitanarayanan, K. 2009, "Inactivation of *Enterobacter sakazakii* in reconstituted infant formula by trans-cinnamaldehyde", *International Journal of Food Microbiology,* vol. 129, no. 2, pp. 146-149.

Anderson, T.F. 1949, *In the Nature of the Bacterial Surface,* Oxford University Press, Oxford, UK.

Anjum, M.F., Lucchini, S., Thompson, A., Hinton, J.C. & Woodward, M.J. 2003, "Comparative genomic indexing reveals the phylogenomics of *Escherichia coli* pathogens", *Infection and Immunity,* vol. 71, no. 8, pp. 4674-4683.

Anjum, M.F., Marooney, C., Fookes, M., Baker, S., Dougan, G., Ivens, A. & Woodward, M.J. 2005, "Identification of core and variable components of the *Salmonella enterica* subspecies I genome by microarray", *Infection and Immunity,* vol. 73, no. 12, pp. 7894-7905.

Arbatsky, N.P., Wang, M., Shashkov, A.S., Chizhov, A.O., Feng, L., Knirel, Y.A. & Wang, L. 2010, "Structure of the O-polysaccharide of *Cronobacter sakazakii* O2 with a randomly O-acetylated l-rhamnose residue", *Carbohydrate research,* vol. 345, no. 14, pp. 2090-2094.

Arseni, A., Malamou-Ladas, E., Koutsia, C., Xanthou, M. & Trikka, E. 1987, "Outbreak of colonization of neonates with *Enterobacter sakazakii*", *The Journal of Hospital Infection,* vol. 9, no. 2, pp. 143-150.

Arthington, B.A., Bennett, L.G., Skatrud, P.L., Guynn, C.J., Barbuch, R.J., Ulbright, C.E. & Bard, M. 1991, "Cloning, disruption and sequence of the gene encoding yeast C-5 sterol desaturase", *Gene,* vol. 102, no. 1, pp. 39-44.

Aubrey, K.L., Casjens, S.R. & Thomas, G.J.,Jr 1992, "Secondary structure and interactions of the packaged dsDNA genome of bacteriophage P22 investigated by Raman difference spectroscopy", *Biochemistry,* vol. 31, no. 47, pp. 11835-11842.

Baldwin, A., Loughlin, M., Caubilla-Barron, J., Kucerova, E., Manning, G., Dowson, C. & Forsythe, S. 2009, "Multilocus sequence typing of *Cronobacter sakazakii* and *Cronobacter malonaticus* reveals stable clonal structures with clinical significance which do not correlate with biotypes", *BMC Microbiology,* vol. 9, pp. 223.

Baranyi, J. & Roberts, T.A. 1994, "A dynamic approach to predicting bacterial growth in food", *International Journal of Food Microbiology,* vol. 23, no. 3-4, pp. 277-294.

Barker, J.R., Chong, A., Wehrly, T.D., Yu, J.J., Rodriguez, S.A., Liu, J., Celli, J., Arulanandam, B.P. & Klose, K.E. 2009, "The *Francisella tularensis* pathogenicity island encodes a secretion system that is required for phagosome escape and virulence", *Molecular Microbiology,* vol. 74, no. 6, pp. 1459-1470.

Barondess, J.J. & Beckwith, J. 1990, "A bacterial virulence determinant encoded by lysogenic coliphage lambda", *Nature,* vol. 346, no. 6287, pp. 871-874.

Bar-Oz, B., Preminger, A., Peleg, O., Block, C. & Arad, I. 2001, "*Enterobacter sakazakii* infection in the newborn", *Acta Paediatrica (Oslo, Norway : 1992),* vol. 90, no. 3, pp. 356-358.

Barron, J.C. & Forsythe, S.J. 2007, "Dry stress and survival time of *Enterobacter sakazakii* and other *Enterobacteriaceae* in dehydrated powdered infant formula", *Journal of Food Protection,* vol. 70, no. 9, pp. 2111-2117.

Baty, F. & Delignette-Muller, M.L. 2004, "Estimating the bacterial lag time: which model, which precision?", *International Journal of Food Microbiology,* vol. 91, no. 3, pp. 261-277.

Bensing, B.A., Rubens, C.E. & Sullam, P.M. 2001a, "Genetic loci of *Streptococcus mitis* that mediate binding to human platelets", *Infection and Immunity,* vol. 69, no. 3, pp. 1373-1380.

Bensing, B.A., Siboo, I.R. & Sullam, P.M. 2001b, "Proteins PblA and PblB of *Streptococcus mitis,* which promote binding to human platelets, are encoded within a lysogenic bacteriophage", *Infection and Immunity,* vol. 69, no. 10, pp. 6186-6192.

Betley, M.J. & Mekalanos, J.J. 1985, "Staphylococcal enterotoxin A is encoded by phage", *Science (New York, N.Y.),* vol. 229, no. 4709, pp. 185-187.

Beuchat, L.R., Kim, H., Gurtler, J.B., Lin, L.C., Ryu, J.H. & Richards, G.M. 2009, "*Cronobacter sakazakii* in foods and factors affecting its survival, growth, and inactivation", *International Journal of Food Microbiology, .*

Biering, G., Karlsson, S., Clark, N.C., Jonsdottir, K.E., Ludvigsson, P. & Steingrimsson, O. 1989, "Three cases of neonatal meningitis caused by *Enterobacter sakazakii* in powdered milk", *Journal of Clinical Microbiology,* vol. 27, no. 9, pp. 2054-2056.

Bladergroen, M.R., Badelt, K. & Spaink, H.P. 2003, "Infection-blocking genes of a symbiotic *Rhizobium leguminosarum* strain that are involved in temperature-dependent protein secretion", *Molecular Plant-microbe Interactions : MPMI,* vol. 16, no. 1, pp. 53-64.

Block, C., Peleg, O., Minster, N., Bar-Oz, B., Simhon, A., Arad, I. & Shapiro, M. 2002, "Cluster of neonatal infections in Jerusalem due to unusual biochemical variant of *Enterobacter sakazakii*", *European Journal of Clinical Microbiology & Infectious Diseases : official publication of the European Society of Clinical Microbiology,* vol. 21, no. 8, pp. 613-616.

Bowen, A.B. & Braden, C.R. 2006, "Invasive *Enterobacter sakazakii* disease in infants", *Emerging Infectious Diseases,* vol. 12, no. 8, pp. 1185-1189.

Boyer, F., Fichant, G., Berthod, J., Vandenbrouck, Y. & Attree, I. 2009, "Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources?", *BMC Genomics,* vol. 10, pp. 104.

Braun, V. & Killmann, H. 1999, "Bacterial solutions to the iron-supply problem", *Trends in Biochemical Sciences,* vol. 24, no. 3, pp. 104-109.

Breeuwer, P., Lardeau, A., Peterz, M. & Joosten, H.M. 2003, "Desiccation and heat tolerance of *Enterobacter sakazakii*", *Journal of Applied Microbiology,* vol. 95, no. 5, pp. 967-973.

Breitbart, M. & Rohwer, F. 2005, "Here a virus, there a virus, everywhere the same virus?", *Trends in Microbiology,* vol. 13, no. 6, pp. 278-284.

Brenner, D.J. 1974, "DNA reassociation for the clinical differentiation of enteric bacteria", *Public Health Laboratory,* vol. 32, no. 4, pp. 118-130.

Brussow, H. & Hendrix, R.W. 2002, "Phage genomics: small is beautiful", *Cell,* vol. 108, no. 1, pp. 13-16.

Brussow, H., Canchaya, C. & Hardt, W.D. 2004, "Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion", *Microbiol Mol Biol Rev,* vol. 68, no. 3, pp. 560-602.

Buchanan, R.L. & Klawitter, L.A. 1991, "Effect of temperature history on the growth of *Listeria monocytogenes* Scott A at refrigeration temperatures", *International Journal of Food Microbiology,* vol. 12, no. 2-3, pp. 235-245.

Burdette, J.H. & Santos, C. 2000, "*Enterobacter sakazakii* brain abscess in the neonate: the importance of neuroradiologic imaging", *Pediatric Radiology,* vol. 30, no. 1, pp. 33-34.

Campellone, K.G. & Leong, J.M. 2003, "Tails of two Tirs: actin pedestal formation by enteropathogenic *E. coli* and enterohemorrhagic E. coli O157:H7", *Current Opinion in Microbiology,* vol. 6, no. 1, pp. 82-90.

Campos, L.C., Lobianco, L.F., Seki, L.M., Santos, R.M. & Asensi, M.D. 2007, "Outbreak of *Enterobacter hormaechei* septicaemia in newborns caused by contaminated parenteral nutrition in Brazil", *The Journal of Hospital Infection,* vol. 66, no. 1, pp. 95-97.

Canchaya, C., Proux, C., Fournous, G., Bruttin, A. & Brussow, H. 2003, "Prophage genomics", *Microbiol Mol Biol Rev,* vol. 67, no. 2, pp. 238-76, table of contents.

Canchaya, C., Fournous, G. & Brussow, H. 2004, "The impact of prophages on bacterial chromosomes", *Molecular Microbiology,* vol. 53, no. 1, pp. 9-18.

Carsiotis, M., Stocker, B.A., Weinstein, D.L. & O'Brien, A.D. 1989, "A *Salmonella typhimurium* virulence gene linked to flg", *Infection and Iimmunity,* vol. 57, no. 11, pp. 3276-3280.

Carter, B., Wu, G., Woodward, M.J. & Anjum, M.F. 2008, "A process for analysis of microarray comparative genomics hybridisation studies for bacterial genomes", *BMC Genomics,* vol. 9, pp. 53.

Cascales, E. 2008, "The type VI secretion toolkit", *EMBO rports,* vol. 9, no. 8, pp. 735-741.

Casjens, S. 2003, "Prophages and bacterial genomics: what have we learned so far?", *Molecular Microbiology,* vol. 49, no. 2, pp. 277-300.

Casman, E.P. 1965, "Staphylococcal enterotoxin", *Annals of the New York Academy of Sciences,* vol. 128, no. 1, pp. 124-131.

Caubilla-Barron, J., Hurrell, E., Townsend, S., Cheetham, P., Loc-Carrillo, C., Fayet, O., Prere, M.F. & Forsythe, S.J. 2007, "Genotypic and phenotypic analysis of *Enterobacter sakazakii* strains from an outbreak resulting in fatalities in a neonatal intensive care unit in France", *Journal of Clinical Microbiology,* vol. 45, no. 12, pp. 3979-3985.

Caubilla-Barron, J., Kucerova, E., Loughlin, M. & Forsythe, S.J. 2009, "Bacteriocidal preparation of powdered infant formula. Food Standards Agency http://www.foodbase.org.uk/results.php?f_category_id=&f_report_id=395".

Centers for Disease Control (CDC) 2002, "*Enterobacter sakazakii* infections associated with the use of powdered infant formula—Tennessee, 2001.", *Morbid. Mortal. Weekly Rep.,* vol. 51, pp. 297-300.

Cha, J.S. & Cooksey, D.A. 1991, "Copper resistance in *Pseudomonas syringae* mediated by periplasmic and outer membrane proteins", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 88, no. 20, pp. 8915-8919.

Chang, C.H., Chiang, M.L. & Chou, C.C. 2010, "The effect of heat shock on the response of *Cronobacter sakazakii* to subsequent lethal stresses", *Foodborne pathogens and disease,* vol. 7, no. 1, pp. 71-76.

Chap, J., Jackson, P., Siqueira, R., Gaspar, N., Quintas, C., Park, J., Osaili, T., Shaker, R., Jaradat, Z., Hartantyo, S.H., Abdullah Sani, N., Estuningsih, S. & Forsythe, S.J. 2009, "International survey of *Cronobacter sakazakii* and other *Cronobacter* spp. in follow up formulas and infant foods", *International Journal of Food microbiology,* vol. 136, no. 2, pp. 185-188.

Chow, J. & Mazmanian, S.K. 2010, "A pathobiont of the microbiota balances host colonization and intestinal inflammation", *Cell Host & Microbe,* vol. 7, no. 4, pp. 265-276.

Clark, N.C., Hill, B.C., O'Hara, C.M., Steingrimsson, O. & Cooksey, R.C. 1990, "Epidemiologic typing of *Enterobacter sakazakii* in two neonatal nosocomial outbreaks", *Diagnostic Microbiology and Infectious Disease,* vol. 13, no. 6, pp. 467-472.

Cleveland, W.S. 1979, " Robust locally weighted regression and smoothing scatterplots", *Journal of the American Statistical Association,* vol. 74, pp. 829-836.

Codex Alimentarius 1981, *CODEX STAN 72 – 19. www.codexalimentarius.net/download/standards/288/CXS_072e.pdf.*

Codex Alimentarius Commission (CAC) 2008, Code of hygienic practice for powdered formulae for infants and young children.

Coignard, B., Vaillant, V., Vincent, J., Leflèche, A., Mariani, P., Bernet, C., L'Hériteau, F., Sénéchall, H., Grimont, P., Bingen, E. & Desenclos, J. 2006, " Infections sévères à *Enterobacter sakazakii* chez des nouveau-nés ayant consommé une préparation en poudre pour nourrissons, France, octobre-décembre 2004 ", *Bulletin Epidémiologique Hebdomadaire (INVS),* , pp. 10-13.

Coimbra, R.S., Grimont, F., Lenormand, P., Burguiere, P., Beutin, L. & Grimont, P.A. 2000, "Identification of *Escherichia coli* O-serogroups by restriction of the amplified O-antigen gene cluster (rfb-RFLP)", *Research in Microbiology,* vol. 151, no. 8, pp. 639-654.

Coleman, D.C., Sullivan, D.J., Russell, R.J., Arbuthnott, J.P., Carey, B.F. & Pomeroy, H.M. 1989, "*Staphylococcus aureus* bacteriophages mediating the simultaneous lysogenic conversion of beta-lysin, staphylokinase and enterotoxin A: molecular mechanism of triple conversion", *Journal of General Microbiology,* vol. 135, no. 6, pp. 1679-1697.

Collyn, F., Billault, A., Mullet, C., Simonet, M. & Marceau, M. 2004, "YAPI, a new *Yersinia pseudotuberculosis* pathogenicity island", *Infection and Immunity,* vol. 72, no. 8, pp. 4784-4790.

Comstock, L.E. & Kasper, D.L. 2006, "Bacterial glycans: key mediators of diverse host immune responses", *Cell,* vol. 126, no. 5, pp. 847-850.

Cossart, P. & Sansonetti, P.J. 2004, "Bacterial invasion: the paradigms of enteroinvasive pathogens", *Science (New York, N.Y.),* vol. 304, no. 5668, pp. 242-248.

Cox, A.J., Hunt, M.L., Ruffolo, C.G. & Adler, B. 2000, "Cloning and characterisation of the *Pasteurella multocida* ahpA gene responsible for a haemolytic phenotype in Escherichia coli", *Veterinary Microbiology,* vol. 72, no. 1-2, pp. 135-152.

Crosa, J.H., Mey, A.R. & Payne, S.M. 2004, "Iron transport in bacteria", *ASM press* .

Czerwicka, M., Forsythe, S.J., Bychowska, A., Dziadziuszko, H., Kunikowska, D., Stepnowski, P. & Kaczynski, Z. 2010, "Structure of the O-polysaccharide isolated from *Cronobacter sakazakii* 767", *Carbohydrate Research,* vol. 345, no. 7, pp. 908-913.

Dancer, G.I., Mah, J.H., Rhee, M.S., Hwang, I.G. & Kang, D.H. 2009, "Resistance of *Enterobacter sakazakii* (*Cronobacter* spp.) to environmental stresses", *Journal of Applied Microbiology,* vol. 107, no. 5, pp. 1606-1614.

Datsenko, K.A. & Wanner, B.L. 2000, "One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 97, no. 12, pp. 6640-6645.

de Lorenzo, V., Bindereif, A., Paw, B.H. & Neilands, J.B. 1986, "Aerobactin biosynthesis and transport genes of plasmid ColV-K30 in *Escherichia coli* K-12", *Journal of Bacteriology,* vol. 165, no. 2, pp. 570-578.

Dinh, T., Paulsen, I.T. & Saier, M.H.,Jr 1994, "A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of gram-negative bacteria", *Journal of Bacteriology,* vol. 176, no. 13, pp. 3825-3831.

Dorrell, N., Mangan, J.A., Laing, K.G., Hinds, J., Linton, D., Al-Ghusein, H., Barrell, B.G., Parkhill, J., Stoker, N.G., Karlyshev, A.V., Butcher, P.D. & Wren, B.W. 2001, "Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity", *Genome Research,* vol. 11, no. 10, pp. 1706-1715.

Dorrell, N., Hinchliffe, S.J. & Wren, B.W. 2005, "Comparative phylogenomics of pathogenic bacteria by microarray analysis", *Current Opinion in Microbiology,* vol. 8, no. 5, pp. 620-626.

Drudy, D., Mullane, N.R., Quinn, T., Wall, P.G. & Fanning, S. 2006, "*Enterobacter sakazakii:* an emerging pathogen in powdered infant formula", *Clin Infect Dis,* vol. 42, no. 7, pp. 996-1002.

Dudley, E.G., Thomson, N.R., Parkhill, J., Morin, N.P. & Nataro, J.P. 2006, "Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative *Escherichia coli", Molecular Microbiology,* vol. 61, no. 5, pp. 1267-1282.

Duguid, J.P., Smith, I.W., Dempster, G. & Edmunds, P.N. 1955, "Non-flagellar filamentous appendages (fimbriae) and haemagglutinating activity in Bacterium coli", *The Journal of Pathology and Bacteriology,* vol. 70, no. 2, pp. 335-348.

Edelson-Mammel, S.G. & Buchanan, R.L. 2004, "Thermal inactivation of *Enterobacter sakazakii* in rehydrated infant formula", *Journal of Food Protection,* vol. 67, no. 1, pp. 60-63.

Eden, R. & and Eden, G. 1984, Impedance Microbiology. *Research Studies Press Ltd.* Research Studies Press Ltd., Herts, UK.

EFSA 2004, "Opinion of the Scientific Panel on Biological Hazards on a request from the Commission related to the microbiological risks in infant formulae and follow-on formulae", *The EFSA Journal,* vol. 113, pp. 1-35.

EFSA 2007, "Opinion of the Scientific Panel on Biological Hazards on the request from the Commission for review of the opinion on microbiological risks in infant formulae and follow-on formulae with regard to *Enterobacteriaceae* as indicators", *The EFSA Journal,* vol. 444, pp. 1-14.

Eggers, C.T., Murray, I.A., Delmar, V.A., Day, A.G. & Craik, C.S. 2004, "The periplasmic serine protease inhibitor ecotin protects bacteria against neutrophil elastase", *The Biochemical Journal,* vol. 379, no. Pt 1, pp. 107-118.

Eisen, M.B., Spellman, P.T., Brown, P.O. & Botstein, D. 1998, "Cluster analysis and display of genome-wide expression patterns", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 95, no. 25, pp. 14863-14868.

FAO/WHO 2004, "Workshop on *Enterobacter sakazakii* and Other Microorganisms in Powdered Infant Formula, Geneva, 2-5 February 2004", http://www.who.int/foodsafety/micro/jemra/meetings/feb2004/en/index.html.

FAO/WHO 2006, "Expert meeting on *Enterobacter sakazakii* and *Salmonella* in powdered infant formula, Rome, 16-20 January 2006", http://www.who.int/foodsafety/micro/jemra/meetings/jan2006/en/index.html".

FAO/WHO 2008, "*Enterobacter sakazakii* (*Cronobacter* spp.) in powdered follow-up formulae. Microbiological Risk Assessment Series No. 15. Rome. 90pp",

Farber, J. & Forsythe, S.J. 2008, "*Enterobacter sakazakiI", ASM press* .

Farmer, J.J. 1980, "*Enterobacter sakazakii:* a new species of "*Enterobacteriaceae*" isolated from clinical specimens", *Intl J System Bacteriol,* vol. 30, pp. 569-584.

Feng, L., Senchenkova, S.N., Wang, W., Shashkov, A.S., Liu, B., Shevelev, S.D., Liu, D., Knirel, Y.A. & Wang, L. 2005, "Structural and genetic characterization of the *Shigella boydii* type 18 O antigen", *Gene,* vol. 355, pp. 79-86.

Figueroa-Bossi, N. & Bossi, L. 1999, "Inducible prophages contribute to *Salmonella* virulence in mice", *Molecular Microbiology,* vol. 33, no. 1, pp. 167-176.

Filloux, A., Hachani, A. & Bleves, S. 2008, "The bacterial type VI secretion machine: yet another player for protein transport across membranes", *Microbiology (Reading, England),* vol. 154, no. Pt 6, pp. 1570-1583.

Fischetti, V.A., Novick, R.P., Ferretti, J.J., Portnoy, D.A. & Rood, J.I. (eds) 2000, Gram-positive pathogens, ASM Press, Washington, D.C.

Fitzgerald, C., Gheesling, L., Collins, M. & Fields, P.I. 2006, "Sequence analysis of the rfb loci, encoding proteins involved in the biosynthesis of the *Salmonella enterica* O17 and O18 antigens: serogroup-specific identification by PCR", *Applied and Environmental Microbiology,* vol. 72, no. 12, pp. 7949-7953.

Folkesson, A., Lofdahl, S. & Normark, S. 2002, "The *Salmonella enterica* subspecies I specific centisome 7 genomic island encodes novel protein families present in bacteria living in close contact with eukaryotic cells", *Research in Microbiology,* vol. 153, no. 8, pp. 537-545.

Forsythe, S.J. 2005, "*Enterobacter sakazakii* and other bacteria in powdered infant milk formula", *Maternal & Child Nutrition,* vol. 1, no. 1, pp. 44-50.

Franco, A.V., Liu, D. & Reeves, P.R. 1996, "A Wzz (Cld) protein determines the chain length of K lipopolysaccharide in Escherichia coli O8 and O9 strains", *Journal of Bacteriology,* vol. 178, no. 7, pp. 1903-1907.

Franke, S., Grass, G., Rensing, C. & Nies, D.H. 2003, "Molecular analysis of the copper-transporting efflux system CusCFBA of Escherichia coli", *Journal of Bacteriology,* vol. 185, no. 13, pp. 3804-3812.

Fronzes, R., Remaut, H. & Waksman, G. 2008, "Architectures and biogenesis of non-flagellar protein appendages in Gram-negative bacteria", *The EMBO journal,* vol. 27, no. 17, pp. 2271-2280.

Frost, L.S., Leplae, R., Summers, A.O. & Toussaint, A. 2005, "Mobile genetic elements: the agents of open source evolution", *Nature reviews,* vol. 3, no. 9, pp. 722-732.

Fujii, N., Oguma, K., Yokosawa, N., Kimura, K. & Tsuzuki, K. 1988, "Characterization of bacteriophage nucleic acids obtained from Clostridium botulinum types C and D", *Applied and Environmental Microbiology,* vol. 54, no. 1, pp. 69-73.

Galvani, C., Terry, J. & Ishiguro, E.E. 2001, "Purification of the RelB and RelE proteins of Escherichia coli: RelE binds to RelB and to ribosomes", *Journal of Bacteriology,* vol. 183, no. 8, pp. 2700-2703.

Garmendia, J., Beuzon, C.R., Ruiz-Albert, J. & Holden, D.W. 2003, "The roles of SsrA-SsrB and OmpR-EnvZ in the regulation of genes encoding the Salmonella typhimurium SPI-2 type III secretion system", *Microbiology* (Reading, England), vol. 149, no. Pt 9, pp. 2385-2396.

Gay, P., Le Coq, D., Steinmetz, M., Berkelman, T. & Kado, C.I. 1985, "Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria", *Journal of Bacteriology*, vol. 164, pp. 918–921.

Goshorn, S.C. & Schlievert, P.M. 1989, "Bacteriophage association of streptococcal pyrogenic exotoxin type C", *Journal of Bacteriology,* vol. 171, no. 6, pp. 3068-3073.

Guan, S., Bastin, D.A. & Verma, N.K. 1999, "Functional analysis of the O antigen glucosylation gene cluster of *Shigella flexneri* bacteriophage SfX", *Microbiology* (Reading, England), vol. 145 (Pt 5), no. Pt 5, pp. 1263-1273.

Gurtler, J.B., Kornacki, J.L. & Beuchat, L.R. 2005, "*Enterobacter sakazakii*: a coliform of increased concern to infant health", *International Journal of Food Microbiology,* vol. 104, no. 1, pp. 1-34.

Gurtler, J.B. & Beuchat, L.R. 2007a, "Growth of *Enterobacter sakazakii* in reconstituted infant formula as affected by composition and temperature", *Journal of Food Protection,* vol. 70, no. 9, pp. 2095-2103.

Gurtler, J.B. & Beuchat, L.R. 2007b, "Survival of *Enterobacter sakazakii* in powdered infant formula as affected by composition, water activity, and temperature", *Journal of Food Protection,* vol. 70, no. 7, pp. 1579-1586.

Gurtler, J.B. & Beuchat, L.R. 2007c, "Inhibition of growth of *Enterobacter sakazakii* in reconstituted infant formula by the lactoperoxidase system", *Journal of Food Protection,* vol. 70, no. 9, pp. 2104-2110.

Hamilton, J.V., Lehane, M.J. & Braig, H.R. 2003, "Isolation of *Enterobacter sakazakii* from midgut of *Stomoxys calcitrans*", *Emerging Infectious Diseases,* vol. 9, no. 10, pp. 1355-1356.

Hardt, W.D., Chen, L.M., Schuebel, K.E., Bustelo, X.R. & Galan, J.E. 1998, "S. typhimurium encodes an activator of Rho GTPases that induces membrane ruffling and nuclear responses in host cells", *Cell,* vol. 93, no. 5, pp. 815-826.

Harris, L.S. & Oriel, P.J. 1989, "Heteropolysaccharide produced by *Enterobacter sakazakii*. United States Patent number 4,806,636.",

Healy, B., Huynh, S., Mullane, N., O'Brien, S., Iversen, C., Lehner, A., Stephan, R., Parker, C.T. & Fanning, S. 2009, "Microarray-based comparative genomic indexing of the *Cronobacter* genus (*Enterobacter sakazakii*)", *International Journal of Food Microbiology,* vol. 136, no. 2, pp. 159-164.

Henderson, D.P. & Payne, S.M. 1994, "Vibrio cholerae iron transport systems: roles of heme and siderophore iron transport in virulence and identification of a gene associated with multiple iron transport systems", *Infection and Immunity,* vol. 62, no. 11, pp. 5120-5125.

Herring, C.D., Glasner, J.D. & Blattner, F.R. 2003, "Gene replacement without selection: regulated suppression of amber mutations in *Escherichia coli",* *Gene,* vol. 311, pp. 153-163.

Hillen, W. & Berens, C. 1994, "Mechanisms underlying expression of Tn10 encoded tetracycline resistance", *Annual Review of Microbiology,* vol. 48, pp. 345-369.

Himelright, I., Harris, E., Lorch, V. & Anderson, M. 2002, "*Enterobacter sakazakii* infections associated with the use of powdered infant formula - Tennessee, 2001", *J Am Med Assoc,* vol. 287, pp. 2204â€"2205.

Ho Sui, S.J., Fedynak, A., Hsiao, W.W., Langille, M.G. & Brinkman, F.S. 2009, "The association of virulence factors with genomic islands", *PloS one,* vol. 4, no. 12, pp. e8094.

Holloway, B.W. & Cooper, G.N. 1962, "Lysogenic conversion in *Pseudomonas aeruginosa*", *Journal of Bacteriology,* vol. 84, pp. 1321-1324.

Holmes, R.K. & Barksdale, L. 1969, "Genetic analysis of tox+ and tox- bacteriophages of *Corynebacterium diphtheriae",* *Journal of Virology,* vol. 3, no. 6, pp. 586-598.

Houwink, A.L. & van Iterson, W. 1950, "Electron microscopical observations on bacterial cytology; a study on flagellation", *Biochimica et Biophysica Acta,* vol. 5, no. 1, pp. 10-44.

Hsiao, W., Wan, I., Jones, S.J. & Brinkman, F.S. 2003, "IslandPath: aiding detection of genomic islands in prokaryotes", *Bioinformatics (Oxford, England),* vol. 19, no. 3, pp. 418-420.

Huang, A., de Grandis, S., Friesen, J., Karmali, M., Petric, M., Congi, R. & Brunton, J.L. 1986, "Cloning and expression of the genes specifying Shiga-like toxin production in *Escherichia coli* H19", *Journal of Bacteriology,* vol. 166, no. 2, pp. 375-379.

Huang, S.H., Chen, Y.H., Fu, Q., Stins, M., Wang, Y., Wass, C. & Kim, K.S. 1999, "Identification and characterization of an *Escherichia coli* invasion gene locus, ibeB, required for penetration of brain microvascular endothelial cells", *Infection and Immunity,* vol. 67, no. 5, pp. 2103-2109.

Huang, S.H., Wan, Z.S., Chen, Y.H., Jong, A.Y. & Kim, K.S. 2001, "Further characterization of *Escherichia coli* brain microvascular endothelial cell invasion gene ibeA by deletion, complementation, and protein expression", *The Journal of Infectious Diseases,* vol. 183, no. 7, pp. 1071-1078.

Humphries, A.D., Raffatellu, M., Winter, S., Weening, E.H., Kingsley, R.A., Droleskey, R., Zhang, S., Figueiredo, J., Khare, S., Nunes, J., Adams, L.G., Tsolis, R.M. & Baumler, A.J. 2003, "The use of flow cytometry to detect expression of subunits encoded by 11 *Salmonella enterica* serotype Typhimurium fimbrial operons", *Molecular Microbiology,* vol. 48, no. 5, pp. 1357-1376.

Hunter, C.J., Singamsetty, V.K., Chokshi, N.K., Boyle, P., Camerini, V., Grishin, A.V., Upperman, J.S., Ford, H.R. & Prasadarao, N.V. 2008, "*Enterobacter sakazakii* enhances epithelial cell injury by inducing apoptosis in a rat model of necrotizing enterocolitis", *The Journal of Infectious Diseases,* vol. 198, no. 4, pp. 586-593.

Hunter, C.J., Williams, M., Petrosyan, M., Guner, Y., Mittal, R., Mock, D., Upperman, J.S., Ford, H.R. & Prasadarao, N.V. 2009, "Lactobacillus bulgaricus prevents intestinal epithelial cell injury caused by *Enterobacter sakazakii*-induced nitric oxide both in vitro and in the newborn rat model of necrotizing enterocolitis", *Infection and Immunity,* vol. 77, no. 3, pp. 1031-1043.

Hurrell, E., Kucerova, E., Loughlin, M., Caubilla-Barron, J. & Forsythe, S.J. 2009a, "Biofilm formation on enteral feeding tubes by *Cronobacter sakazakii, Salmonella* serovars and other *Enterobacteriaceae*", *International Journal of Food Microbiology,* vol. 136, no. 2, pp. 227-231.

Hurrell, E., Kucerova, E., Loughlin, M., Caubilla-Barron, J., Hilton, A., Armstrong, R., Smith, C., Grant, J., Shoo, S. & Forsythe, S. 2009b, "Neonatal enteral feeding tubes as loci for colonisation by members of the *Enterobacteriaceae*", *BMC Infectious Diseases,* vol. 9, pp. 146.

Huyghe, A., Francois, P. & Schrenzel, J. 2009, "Characterization of microbial pathogens by DNA microarrays", *Infection, genetics and evolution : journal of*

molecular epidemiology and evolutionary genetics in infectious diseases, vol. 9, no. 5, pp. 987-995.

Hynes, W.L. & Ferretti, J.J. 1989, "Sequence analysis and expression in *Escherichia coli* of the hyaluronidase gene of Streptococcus pyogenes bacteriophage H4489A", *Infection and Immunity,* vol. 57, no. 2, pp. 533-539.

Iversen, C. & Forsythe, S.J. 2003, "Risk profile of *Enterobacter sakazakii*, an emergent pathogen associated with infant milk formula", *Trends in Food Science & Technology,* vol. 14, no. 11, pp. 443-454.

Iversen, C. & Forsythe, S. 2004a, "Isolation of *Enterobacter sakazakii* and other *Enterobacteriaceae* from powdered infant formula milk and related products", *Food Microbiology,* vol. 21, no. 6, pp. 771-777.

Iversen, C., Lane, M. & Forsythe, S.J. 2004b, "The growth profile, thermotolerance and biofilm formation of *Enterobacter sakazakii* grown in infant formula milk", *Letters in Applied Microbiology,* vol. 38, no. 5, pp. 378-382.

Iversen, C., Waddington, M., On, S.L. & Forsythe, S. 2004c, "Identification and phylogeny of *Enterobacter sakazakii* relative to *Enterobacter* and *Citrobacter* Species", *Journal of Clinical Microbiology,* vol. 42, no. 11, pp. 5368-5370.

Iversen, C., Waddington, M., Farmer, J.J & Forsythe, S.J. 2006, "The biochemical differentiation of *Enterobacter sakazakii* genotypes", BMC Microbiology, vol.6, pp.94.

Iversen, C. & Forsythe, S.J. 2007, "Comparison of media for the isolation of Enterobacter sakazakii", *Applied and Environmental Microbiology,* vol. 73, no. 1, pp. 48-52.

Iversen, C., Lehner, A., Mullane, N., Bidlas, E., Cleenwerck, I., Marugg, J., Fanning, S., Stephan, R. & Joosten, H. 2007, "The taxonomy of *Enterobacter sakazakii*: proposal of a new genus *Cronobacter* gen. nov. and descriptions of *Cronobacter* sakazakii comb. nov. *Cronobacter sakazakii* subsp. sakazakii, comb. nov., *Cronobacter sakazakii* subsp. malonaticus subsp. nov., *Cronobacter turicensis* sp. nov., *Cronobacter muytjensii* sp. nov., Cronobacter dublinensis sp. nov. and *Cronobacter* genomospecies 1", *BMC evolutionary biology,* vol. 7, pp. 64.

Iversen, C., Mullane, N., McCardell, B., Tall, B.D., Lehner, A., Fanning, S., Stephan, R. & Joosten, H. 2008, "*Cronobacter* gen. nov., a new genus to accommodate the biogroups of *Enterobacter sakazakii,* and proposal of *Cronobacter* sakazakii gen. nov., comb. nov., Cronobacter malonaticus sp. nov., Cronobacter turicensis sp. nov., *Cronobacter muytjensii* sp. nov., *Cronobacter dublinensis* sp. nov., *Cronobacter* genomospecies 1, and of three subspecies, *Cronobacter dublinensis* subsp. *dublinensis* subsp. nov., *Cronobacter dublinensis* subsp. lausannensis subsp. nov. and *Cronobacter dublinensis* subsp. lactaridi subsp. nov", *International Journal of Systematic and Evolutionary Microbiology,* vol. 58, pp. 1442-1447.

Izard, D., Richar, C. & Leclerc, H. 1983, "DNA relatedness between *Enterobacter sakazakii and* other members of the genus *Enterobacter*", *Annales de l'Institut Pasteur.Microbiologie,* vol. 134, no. 3, Supplement 1, pp. 241-245.

Jani, A.J. & Cotter, P.A. 2010, "Type VI secretion: not just for pathogenesis anymore", *Cell host & microbe,* vol. 8, no. 1, pp. 2-6.

Jaradat, Z.W., Ababneh, Q.O., Saadoun, I.M., Samara, N.A. & Rashdan, A.M. 2009, "Isolation of *Cronobacter* spp. (formerly *Enterobacter sakazakii*) from infant food, herbs and environmental samples and the subsequent identification and confirmation of the isolates using biochemical, chromogenic assays, PCR and 16S rRNA sequencing", *BMC Microbiology,* vol. 9, pp. 225.

Jasin, M. & Schimmel, P. 1984, "Deletion of an essential gene in *Escherichia coli* by site-specific recombination with linear DNA fragments", *Journal of Bacteriology,* vol. 159, no. 2, pp. 783-786.

Johnson, L.P. & Schlievert, P.M. 1984, "Group A streptococcal phage T12 carries the structural gene for pyrogenic exotoxin type A", *Molecular & general genetics : MGG,* vol. 194, no. 1-2, pp. 52-56.

Joosten, H., Marugg, J., Stephan, R., Klijn, A., Jackson, T. & Iversen, C. 2008, "A rapid and reliable alternative to ISO 21528-1:2004 for detection of *Enterobacteriaceae*", *International Journal of Food Microbiology,* vol. 125, no. 3, pp. 344-346.

Jourdan, N., Le Hello, S., Delmas, G., Clouzeau, J., Manteau, C., Desaubliaux, B., Chagnon, V., Thierry-Bled, F., Demare, N., Weill, F. & de Valk, H. 2008, "Nationwide outbreak of *Salmonella enterica* serotype gives infections in infants in France, linked to infant milk formula, September 2008", *Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin,* vol. 13, no. 39, pp. 18994.

Kado, C.I. & Liu, S.T. 1981, "Rapid procedure for detection and isolation of large and small plasmids", *Journal of Bacteriology,* vol. 145, no. 3, pp. 1365-1373.

Kandhai, M.C., Reij, M.W., Gorris, L.G., Guillaume-Gentil, O. & van Schothorst, M. 2004, "Occurrence of *Enterobacter sakazakii* in food production environments and households", *Lancet,* vol. 363, no. 9402, pp. 39-40.

Kandhai, M.C., Reij, M.W., Grognou, C., van Schothorst, M., Gorris, L.G. & Zwietering, M.H. 2006, "Effects of preculturing conditions on lag time and specific growth rate of *Enterobacter sakazakii* in reconstituted powdered infant formula", *Applied and Environmental Microbiology,* vol. 72, no. 4, pp. 2721-2729.

Kaneko, J., Muramoto, K. & Kamio, Y. 1997, "Gene of LukF-PV-like component of Panton-Valentine leukocidin in *Staphylococcus aureus* P83 is linked with lukM", *Bioscience, Biotechnology, and Biochemistry,* vol. 61, no. 3, pp. 541-544.

Karaolis, D.K., Somara, S., Maneval, D.R.,Jr, Johnson, J.A. & Kaper, J.B. 1999, "A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria", *Nature,* vol. 399, no. 6734, pp. 375-379.

Kempf, B. & Bremer, E. 1998, "Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments", *Archives of Microbiology,* vol. 170, no. 5, pp. 319-330.

Kim, C.C., Joyce, E.A., Chan, K. & Falkow, S. 2002, "Improved analytical methods for microarray-based genome-composition analysis", *Genome Biology,* vol. 3, no. 11, pp. RESEARCH0065.

Kim, H., Ryu, J.H. & Beuchat, L.R. 2006, "Attachment of and biofilm formation by *Enterobacter sakazakii* on stainless steel and enteral feeding tubes", *Applied and Environmental Microbiology,* vol. 72, no. 9, pp. 5846-5856.

Kim, H., Ryu, J.H. & Beuchat, L.R. 2007a, "Effectiveness of disinfectants in killing *Enterobacter sakazakii* in suspension, dried on the surface of stainless steel, and in a biofilm", *Applied and Environmental Microbiology,* vol. 73, no. 4, pp. 1256-1265.

Kim, S.H. & Park, J.H. 2007b, "Thermal resistance and inactivation of *Enterobacter sakazakii* isolates during rehydration of powdered infant formula", *Journal of Microbiology and Biotechnology,* vol. 17, no. 2, pp. 364-368.

Kim, J.B., Cho, S.H., Park, Y.B., Lee, J.B., Kim, J.C., Lee, B.K., Lee, H.K. & Chae, H.S. 2008a, "Surveillance of stool samples for the presence of *Enterobacter sakazakii* among Korean people", *Yonsei Medical Journal,* vol. 49, no. 6, pp. 1017-1022.

Kim, K.P. & Loessner, M.J. 2008b, "*Enterobacter sakazakii* invasion in human intestinal Caco-2 cells requires the host cell cytoskeleton and is enhanced by disruption of tight junction", *Infection and Immunity,* vol. 76, no. 2, pp. 562-570.

Kim, K., Kim, K.P., Choi, J., Lim, J.A., Lee, J., Hwang, S. & Ryu, S. 2010, "Outer membrane proteins A (OmpA) and X (OmpX) are essential for basolateral invasion of *Cronobacter* sakazakii", *Applied and Environmental Microbiology,* vol. 76, no. 15, pp. 5188-5198.

Kline, M.W. 1988, "*Citrobacter* meningitis and brain abscess in infancy: Epidemiology, pathogenesis, and treatment", *The Journal of Pediatrics,* vol. 113, no. 3, pp. 430-434.

Kline, K.A., Falker, S., Dahlberg, S., Normark, S. & Henriques-Normark, B. 2009, "Bacterial adhesins in host-microbe interactions", *Cell Host & Microbe,* vol. 5, no. 6, pp. 580-592.

Kline, K.A., Dodson, K.W., Caparon, M.G. & Hultgren, S.J. 2010, "A tale of two pili: assembly and function of pili in bacteria", *Trends in Microbiology,* vol. 18, no. 5, pp. 224-232.

Konstantinidis, K.T. & Tiedje, J.M. 2004, "Trends between gene content and genome size in prokaryotic species with larger genomes", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 101, no. 9, pp. 3160-3165.

Konstantinidis, K.T., Ramette, A. & Tiedje, J.M. 2006, "The bacterial species definition in the genomic era", *Philosophical transactions of the Royal Society of London.Series B, Biological sciences,* vol. 361, no. 1475, pp. 1929-1940.

Kothary, M.H., McCardell, B.A., Frazar, C.D., Deer, D. & Tall, B.D. 2007, "Characterization of the zinc-containing metalloprotease encoded by zpx and development of a species-specific detection method for *Enterobacter sakazakii*", *Applied and Environmental Microbiology,* vol. 73, no. 13, pp. 4142-4151.

Kretchmar, S.A., Reyes, Z.E. & Raymond, K.N. 1988, "The spectroelectrochemical determination of the reduction potential of diferric serum transferrin", *Biochimica et Biophysica Acta,* vol. 956, no. 1, pp. 85-94.

Kucerova, E., Clifton, S.W., Xia, X.Q., Long, F., Porwollik, S., Fulton, L., Fronick, C., Minx, P., Kyung, K., Warren, W., Fulton, R., Feng, D., Wollam, A., Shah, N., Bhonagiri, V., Nash, W.E., Hallsworth-Pepin, K., Wilson, R.K., McClelland, M. & Forsythe, S.J. 2010, "Genome sequence of *Cronobacter sakazakii* BAA-894 and comparative genomic hybridization analysis with other *Cronobacter* species", *PloS One,* vol. 5, no. 3, pp. e9556.

Kuhnert, P., Korczak, B.M., Stephan, R., Joosten, H. & Iversen, C. 2009, "Phylogeny and prediction of genetic similarity of *Cronobacter* and related taxa by multilocus sequence analysis (MLSA)", *International Journal of Food Microbiology,* vol. 136, no. 2, pp. 152-158.

Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A., Mizutani-Ui, Y., Takahashi, N.K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. & Hiramatsu, K. 2001, "Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*", *The Lancet,* vol. 357, no. 9264, pp. 1225-1240.

Lafont, J.P., Dho, M., D'Hauteville, H.M., Bree, A. & Sansonetti, P.J. 1987, "Presence and expression of aerobactin genes in virulent avian strains of *Escherichia coli*", *Infection and Immunity,* vol. 55, no. 1, pp. 193-197.

Lai, K.K. 2001, "*Enterobacter sakazakii* infections among neonates, infants, children, and adults. Case reports and a review of the literature", *Medicine,* vol. 80, no. 2, pp. 113-122.

Langille, M.G., Hsiao, W.W. & Brinkman, F.S. 2008, "Evaluation of genomic island predictors using a comparative genomics approach", *BMC Bioinformatics,* vol. 9, pp. 329.

Langille, M.G. & Brinkman, F.S. 2009, "IslandViewer: an integrated interface for computational identification and visualization of genomic islands", *Bioinformatics (Oxford, England),* vol. 25, no. 5, pp. 664-665.

Lazdunski, C.J. 1988, "Pore-forming colicins: synthesis, extracellular release, mode of action, immunity", *Biochimie,* vol. 70, no. 9, pp. 1291-1296.

Lee, D.J., Bingle, L.E., Heurlier, K., Pallen, M.J., Penn, C.W., Busby, S.J. & Hobman, J.L. 2009, "Gene doctoring: a method for recombineering in laboratory and pathogenic *Escherichia coli* strains", *BMC Microbiology,* vol. 9, pp. 252.

Lee, S.Y. & Jin, H.H. 2008, "Inhibitory activity of natural antimicrobial compounds alone or in combination with nisin against *Enterobacter sakazakii*", *Letters in Applied Microbiology,* vol. 47, no. 4, pp. 315-321.

Lehner, A., Riedel, K., Eberl, L., Breeuwer, P., Diep, B. & Stephan, R. 2005, "Biofilm formation, extracellular polysaccharide production, and cell-to-cell signaling in various *Enterobacter sakazakii* strains: aspects promoting environmental persistence", *Journal of Food Protection,* vol. 68, no. 11, pp. 2287-2294.

Lehner, A., Grimm, M., Rattei, T., Ruepp, A., Frishman, D., Manzardo, G.G. & Stephan, R. 2006, "Cloning and characterization of *Enterobacter sakazakii* pigment genes and in situ spectroscopic analysis of the pigment", *FEMS Microbiology Letters,* vol. 265, no. 2, pp. 244-248.

Leiman, P.G., Basler, M., Ramagopal, U.A., Bonanno, J.B., Sauder, J.M., Pukatzki, S., Burley, S.K., Almo, S.C. & Mekalanos, J.J. 2009, "Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 106, no. 11, pp. 4154-4159.

Lenati, R.F., O'Connor, D.L., Hebert, K.C., Farber, J.M. & Pagotto, F.J. 2008, "Growth and survival of *Enterobacter sakazakii* in human breast milk with and without fortifiers as compared to powdered infant formula", *International Journal of Food Microbiology,* vol. 122, no. 1-2, pp. 171-179.

Lima-Mendez, G., Van Helden, J., Toussaint, A. & Leplae, R. 2008, "Prophinder: a computational tool for prophage prediction in prokaryotic genomes", *Bioinformatics (Oxford, England),* vol. 24, no. 6, pp. 863-865.

Lin, L.C. & Beuchat, L.R. 2007, "Survival of *Enterobacter sakazakii* in infant cereal as affected by composition, water activity, and temperature", *Food Microbiology,* vol. 24, no. 7-8, pp. 767-777.

Lindsay, J.A., Ruzin, A., Ross, H.F., Kurepina, N. & Novick, R.P. 1998, "The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus*", *Molecular Microbiology,* vol. 29, no. 2, pp. 527-543.

Lucas, A. & Cole, T.J. 1990, "Breast milk and neonatal necrotising enterocolitis", *Lancet,* vol. 336, no. 8730, pp. 1519-1523.

MacLean, L.L., Pagotto, F., Farber, J.M. & Perry, M.B. 2009a, "Structure of the antigenic repeating pentasaccharide unit of the LPS O-polysaccharide of *Cronobacter sakazakii* implicated in the Tennessee outbreak", *Biochemistry and Cell Biology = Biochimie et biologie cellulaire,* vol. 87, no. 2, pp. 459-465.

MacLean, L.L., Pagotto, F., Farber, J.M. & Perry, M.B. 2009b, "The structure of the O-antigen in the endotoxin of the emerging food pathogen *Cronobacter (Enterobacter) muytjensii* strain 3270", *Carbohydrate Research,* vol. 344, no. 5, pp. 667-671.

MacLean, L.L., Vinogradov, E., Pagotto, F., Farber, J.M. & Perry, M.B. 2009c, "Characterization of the O-antigen in the lipopolysaccharide of *Cronobacter*

(Enterobacter) malonaticus 3267", *Biochemistry and cell biology = Biochimie et Biologie Cellulaire,* vol. 87, no. 6, pp. 927-932.

Marshall, K.C. 1992, "Biofilms: an overview of bacterial adhesion activity and control at surfaces", *American Society for Microbiology News,* vol. 58, pp. 202-207.

Martinez, J.L., Herrero, M. & de Lorenzo, V. 1994, "The organization of intercistronic regions of the aerobactin operon of pColV-K30 may account for the differential expression of the iucABCD iutA genes", *Journal of Molecular Biology,* vol. 238, no. 2, pp. 288-293.

Mavris, M., Manning, P.A. & Morona, R. 1997, "Mechanism of bacteriophage SfII-mediated serotype conversion in *Shigella flexneri*", *Molecular Microbiology,* vol. 26, no. 5, pp. 939-950.

McClelland, M., Florea, L., Sanderson, K., Clifton, S.W., Parkhill, J., Churcher, C., Dougan, G., Wilson, R.K. & Miller, W. 2000, "Comparison of the *Escherichia coli K*-12 genome with sampled genomes of a *Klebsiella pneumoniae* and three salmonella enterica serovars, Typhimurium, Typhi and Paratyphi", *Nucleic Acids Research,* vol. 28, no. 24, pp. 4974-4986.

McDonough, M.A. & Butterton, J.R. 1999, "Spontaneous tandem amplification and deletion of the shiga toxin operon in *Shigella dysenteriae* 1", *Molecular Microbiology,* vol. 34, no. 5, pp. 1058-1069.

Miled-Bennour, R., Ells, T.C., Pagotto, F.J., Farber, J.M., Kerouanton, A., Meheut, T., Colin, P., Joosten, H., Leclercq, A. & Besse, N.G. 2010, "Genotypic and phenotypic characterisation of a collection of *Cronobacter (Enterobacter sakazakii)* isolates", *International Journal of Food Microbiology,* vol. 139, no. 1-2, pp. 116-125.

Miles, A.A. & Misra, S.S. 1938, "The estimation of the bactericidal power of the blood", *J. Hygiene,* vol. 38, pp. 732-749.

Mirold, S., Rabsch, W., Rohde, M., Stender, S., Tschape, H., Russmann, H., Igwe, E. & Hardt, W.D. 1999, "Isolation of a temperate bacteriophage encoding the type III effector protein SopE from an epidemic *Salmonella typhimurium* strain", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 96, no. 17, pp. 9845-9850.

Mittal, R., Bulgheresi, S., Emami, C. & Prasadarao, N.V. 2009a, "*Enterobacter sakazakii* targets DC-SIGN to induce immunosuppressive responses in dendritic cells by modulating MAPKs", *Journal of Immunology (Baltimore, Md.: 1950),* vol. 183, no. 10, pp. 6588-6599.

Mittal, R., Wang, Y., Hunter, C.J., Gonzalez-Gomez, I. & Prasadarao, N.V. 2009b, "Brain damage in newborn rat model of meningitis by *Enterobacter sakazakii*: a role for outer membrane protein A", *Laboratory investigation; a Journal of Technical Methods and Pathology,* vol. 89, no. 3, pp. 263-277.

Mohan Nair, M.K. & Venkitanarayanan, K.S. 2006, "Cloning and sequencing of the ompA gene of *Enterobacter sakazakii* and development of an ompA-targeted

PCR for rapid detection of *Enterobacter sakazakii* in infant formula", *Applied and Environmental Microbiology,* vol. 72, no. 4, pp. 2539-2546.

Molloy, C., Cagney, C., Fanning, S. & Duffy, G. 2010, "Survival characteristics of *Cronobacter* spp. in model bovine gut and in the environment", *Foodborne Pathogens and Disease,* vol. 7, no. 6, pp. 671-675.

Molloy, C., Cagney, C., O'Brien, S., Iversen, C., Fanning, S. & Duffy, G. 2009, "Surveillance and characterisation by pulsed-field gel electrophoresis of *Cronobacter* spp. in farming and domestic environments, food production animals and retail foods", *International Journal of Food Microbiology,* vol. 136, no. 2, pp. 198-203.

Mougous, J.D., Senaratne, R.H., Petzold, C.J., Jain, M., Lee, D.H., Schelle, M.W., Leavell, M.D., Cox, J.S., Leary, J.A., Riley, L.W. & Bertozzi, C.R. 2006, "A sulfated metabolite produced by stf3 negatively regulates the virulence of *Mycobacterium tuberculosis*", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 103, no. 11, pp. 4258-4263.

Mougous, J.D., Gifford, C.A., Ramsdell, T.L. & Mekalanos, J.J. 2007, "Threonine phosphorylation post-translationally regulates protein secretion in *Pseudomonas aeruginosa*", *Nature Cell Biology,* vol. 9, no. 7, pp. 797-803.

Mukhopadhyay, S., Kapatral, V., Xu, W. & Chakrabarty, A.M. 1999, "Characterization of a Hank's type serine/threonine kinase and serine/threonine phosphoprotein phosphatase in *Pseudomonas aeruginosa*", *Journal of Bacteriology,* vol. 181, no. 21, pp. 6615-6622.

Mullane, N., Healy, B., Meade, J., Whyte, P., Wall, P.G. & Fanning, S. 2008a, "Dissemination of *Cronobacter* spp. (*Enterobacter sakazakii)* in a powdered milk protein manufacturing facility", *Applied and Environmental Microbiology,* vol. 74, no. 19, pp. 5913-5917.

Mullane, N., O'Gaora, P., Nally, J.E., Iversen, C., Whyte, P., Wall, P.G. & Fanning, S. 2008b, "Molecular analysis of the *Enterobacter sakazakii* O-antigen gene locus", *Applied and Environmental Microbiology,* vol. 74, no. 12, pp. 3783-3794.

Munson, G.P., Lam, D.L., Outten, F.W. & O'Halloran, T.V. 2000, "Identification of a copper-responsive two-component system on the chromosome of *Escherichia coli* K-12", *Journal of Bacteriology,* vol. 182, no. 20, pp. 5864-5871.

Murphy, K.C. 1998, "Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*", *Journal of Bacteriology,* vol. 180, no. 8, pp. 2063-2071.

Muytjens, H.L., Zanen, H.C., Sonderkamp, H.J., Kollee, L.A., Wachsmuth, I.K. & Farmer, J.J.,3rd 1983, "Analysis of eight cases of neonatal meningitis and sepsis due to *Enterobacter sakazakii*", *Journal of Clinical Microbiology,* vol. 18, no. 1, pp. 115-120.

Muytjens, H.L., Roelofs-Willemse, H. & Jaspar, G.H. 1988, "Quality of powdered substitutes for breast milk with regard to members of the family *Enterobacteriaceae*", *Journal of Clinical Microbiology,* vol. 26, no. 4, pp. 743-746.

Muytjens, H.L. & Kollee, L.A. 1990, "*Enterobacter sakazakii* meningitis in neonates: causative role of formula?", *The Pediatric Infectious Disease Journal,* vol. 9, no. 5, pp. 372-373.

Nair, M.K., Joy, J. & Venkitanarayanan, K.S. 2004, "Inactivation of Enterobacter sakazakii in reconstituted infant formula by monocaprylin", *Journal of food protection,* vol. 67, no. 12, pp. 2815-2819.

Nair, M.K., Venkitanarayanan, K., Silbart, L.K. & Kim, K.S. 2009, "Outer membrane protein A (OmpA) of *Cronobacter sakazakii* binds fibronectin and contributes to invasion of human brain microvascular endothelial cells", *Foodborne Pathogens and Disease,* vol. 6, no. 4, pp. 495-501.

Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y., Okahashi, N., Kawabata, S., Yamazaki, K., Shiba, T., Yasunaga, T., Hayashi, H., Hattori, M. & Hamada, S. 2003, "Genome sequence of an M3 strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution", *Genome Research,* vol. 13, no. 6A, pp. 1042-1055.

Nazarowec-White, M. & Farber, J.M. 1999, "Phenotypic and genotypic typing of food and clinical isolates of *Enterobacter sakazakii*", *Journal of Medical Microbiology,* vol. 48, no. 6, pp. 559-567.

Newland, J.W., Strockbine, N.A., Miller, S.F., O'Brien, A.D. & Holmes, R.K. 1985, "Cloning of Shiga-like toxin structural genes from a toxin converting phage of *Escherichia coli*", *Science (New York, N.Y.),* vol. 230, no. 4722, pp. 179-181.

Nishiyama, M., Horst, R., Eidam, O., Herrmann, T., Ignatov, O., Vetsch, M., Bettendorff, P., Jelesarov, I., Grutter, M.G., Wuthrich, K., Glockshuber, R. & Capitani, G. 2005, "Structural basis of chaperone-subunit complex recognition by the type 1 pilus assembly platform FimD", *The EMBO journal,* vol. 24, no. 12, pp. 2075-2086.

Nuccio, S.P. & Baumler, A.J. 2007, "Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek", *Microbiology and Molecular Biology Reviews : MMBR,* vol. 71, no. 4, pp. 551-575.

Nystrom, T. & Neidhardt, F.C. 1994, "Expression and role of the universal stress protein, UspA, of *Escherichia coli* during growth arrest", *Molecular Microbiology,* vol. 11, no. 3, pp. 537-544.

Ohnishi, M., Kurokawa, K. & Hayashi, T. 2001, "Diversification of *Escherichia coli* genomes: are bacteriophages the major contributors?", *Trends Microbiol,* vol. 9, pp. 481–485.

Okeke, I.N., Scaletsky, I.C., Soars, E.H., Macfarlane, L.R. & Torres, A.G. 2004, "Molecular epidemiology of the iron utilization genes of enteroaggregative *Escherichia coli*", *Journal of Clinical Microbiology,* vol. 42, no. 1, pp. 36-44.

Osaili, T.M., Al-Nabulsi, A.A., Shaker, R.R., Ayyash, M.M., Olaimat, A.N., Al-Hasan, A.S., Kadora, K.M. & Holley, R.A. 2008, "Effects of extended dry storage of powdered infant milk formula on susceptibility of *Enterobacter sakazakii* to hot water and ionizing radiation", *Journal of Food Protection,* vol. 71, no. 5, pp. 934-939.

Osaili, T. & Forsythe, S. 2009, "Desiccation resistance and persistence of *Cronobacter* species in infant formula", *International Journal of Food Microbiology,* vol. 136, no. 2, pp. 214-220.

Pagan, R., Condon, S. & Sala, F.J. 1997, "Effects of several factors on the heat-shock-induced thermotolerance of *Listeria monocytogenes*", *Applied and Environmental Microbiology,* vol. 63, no. 8, pp. 3225-3232.

Pagotto, F.J., Nazarowec-White, M., Bidawid, S. & Farber, J.M. 2003, "*Enterobacter sakazakii*: infectivity and enterotoxin production in vitro and in vivo", *Journal of Food Protection,* vol. 66, no. 3, pp. 370-375.

Pan, W. & Spratt, B.G. 1994, "Regulation of the permeability of the gonococcal cell envelope by the mtr system", *Molecular Microbiology,* vol. 11, no. 4, pp. 769-775.

Pangalos, G. 1929, "Sur un bacille chromogene isole par hemoculture.", *C.R. Soc. Biol. (Comptes Rendus Seances Soc. Biol.),* vol. 1000, pp. 1097.

Parsons, D.A. & Heffron, F. 2005, "sciS, an icmF homolog in *Salmonella enterica* serovar Typhimurium, limits intracellular replication and decreases virulence", *Infection and immunity,* vol. 73, no. 7, pp. 4338-4345.

Pedersen, K., Christensen, S.K. & Gerdes, K. 2002, "Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins", *Molecular Microbiology,* vol. 45, no. 2, pp. 501-510.

Peter, C.S., Feuerhahn, M., Bohnhorst, B., Schlaud, M., Ziesing, S., von der Hardt, H. & Poets, C.F. 1999, "Necrotising enterocolitis: is there a relationship to specific pathogens?", *European Journal of Pediatrics,* vol. 158, no. 1, pp. 67-70.

Pitout, J.D., Moland, E.S., Sanders, C.C., Thomson, K.S. & Fitzsimmons, S.R. 1997, "Beta-lactamases and detection of beta-lactam resistance in *Enterobacter* spp", *Antimicrobial Agents and Chemotherapy,* vol. 41, no. 1, pp. 35-39.

Porwollik, S., Wong, R.M. & McClelland, M. 2002, "Evolutionary genomics of *Salmonella:* gene acquisitions revealed by microarray analysis", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 99, no. 13, pp. 8956-8961.

Porwollik, S., Boyd, E.F., Choy, C., Cheng, P., Florea, L., Proctor, E. & McClelland, M. 2004, "Characterization of *Salmonella enterica* subspecies I genovars by use of microarrays", *Journal of Bacteriology,* vol. 186, no. 17, pp. 5883-5898.

Prasadarao, N.V., Wass, C.A., Weiser, J.N., Stins, M.F., Huang, S.H. & Kim, K.S. 1996, "Outer membrane protein A of *Escherichia coli* contributes to invasion of brain microvascular endothelial cells", *Infection and Immunity,* vol. 64, no. 1, pp. 146-153.

Proft, T. & Baker, E.N. 2009, "Pili in Gram-negative and Gram-positive bacteria - structure, assembly and their role in disease", *Cellular and Molecular Life Sciences : CMLS,* vol. 66, no. 4, pp. 613-635.

Pukatzki, S., Ma, A.T., Revel, A.T., Sturtevant, D. & Mekalanos, J.J. 2007, "Type VI secretion system translocates a phage tail spike-like protein into target cells

where it cross-links actin", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 104, no. 39, pp. 15508-15513.

Raymond, K.N., Dertz, E.A. & Kim, S.S. 2003, "Enterobactin: an archetype for microbial iron transport", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 100, no. 7, pp. 3584-3588.

Reeve, J.N. & Shaw, J.E. 1979, "Lambda encodes an outer membrane protein: the lom gene", *Molecular & General Genetics : MGG,* vol. 172, no. 3, pp. 243-248.

Reeves, P. 1993, "Evolution of Salmonella O antigen variation by interspecific gene transfer on a large scale", *Trends in Genetics : TIG,* vol. 9, no. 1, pp. 17-22.

Riedel, K. & Lehner, A. 2007, "Identification of proteins involved in osmotic stress response in *Enterobacter sakazakii* by proteomics", *Proteomics,* vol. 7, no. 8, pp. 1217-1231.

Robbins, P.W. & Uchida, T. 1962, "Determinants of specificity in *Salmonella:* changes in antigenic structure mediated by bacteriophage", *Federation Proceedings,* vol. 21, pp. 702-710.

Rossello-Mora, R. & Amann, R. 2001, "The species concept for prokaryotes", *FEMS Microbiology Reviews,* vol. 25, no. 1, pp. 39-67.

Rossmann, M.G., Mesyanzhinov, V.V., Arisaka, F. & Leiman, P.G. 2004, "The bacteriophage T4 DNA injection machine", *Current Opinion in Structural Biology,* vol. 14, no. 2, pp. 171-180.

Saier, M.H.,Jr, Tam, R., Reizer, A. & Reizer, J. 1994, "Two novel families of bacterial membrane proteins concerned with nodulation, cell division and transport", *Molecular Microbiology,* vol. 11, no. 5, pp. 841-847.

Sako, T., Sawaki, S., Sakurai, T., Ito, S., Yoshizawa, Y. & Kondo, I. 1983, "Cloning and expression of the staphylokinase gene of *Staphylococcus aureus* in *Escherichia coli*", *Molecular & General Genetics : MGG,* vol. 190, no. 2, pp. 271-277.

Sanders, W.E.,Jr & Sanders, C.C. 1997, "*Enterobacter* spp.: pathogens poised to flourish at the turn of the century", *Clinical Microbiology Reviews,* vol. 10, no. 2, pp. 220-241.

Scaria, J., Palaniappan, R.U., Chiu, D., Phan, J.A., Ponnala, L., McDonough, P., Grohn, Y.T., Porwollik, S., McClelland, M., Chiou, C.S., Chu, C. & Chang, Y.F. 2008, "Microarray for molecular typing of *Salmonella enterica* serovars", *Molecular and Cellular Probes,* vol. 22, no. 4, pp. 238-243.

Scheepe-Leberkuhne, M. & Wagner, F. 1986, "Optimization and preliminary characterization of an exopolysaccharide synthezised by *Enterobacter sakazakii",* *Biotechnology letters,* vol. 8, pp. 695-700.

Schlieker, C., Zentgraf, H., Dersch, P. & Mogk, A. 2005, "ClpV, a unique Hsp100/Clp member of pathogenic proteobacteria", *Biological Chemistry,* vol. 386, no. 11, pp. 1115-1127.

Schmid, M., Iversen, C., Gontia, I., Stephan, R., Hofmann, A., Hartmann, A., Jha, B., Eberl, L., Riedel, K. & Lehner, A. 2009, "Evidence for a plant-associated natural habitat for *Cronobacter* spp", *Research in Microbiology,* vol. 160, no. 8, pp. 608-614.

Schnaitman, C.A. & Klena, J.D. 1993, "Genetics of lipopolysaccharide biosynthesis in enteric bacteria", *Microbiological Reviews,* vol. 57, no. 3, pp. 655-682.

Shaker, R.R., Osaili, T.M., Abu Al-Hasan, A.S., Ayyash, M.M. & Forsythe, S.J. 2008, "Effect of desiccation, starvation, heat, and cold stresses on the thermal resistance of *Enterobacter sakazakii* in rehydrated infant milk formula", *Journal of Food Science,* vol. 73, no. 7, pp. M354-9.

Shalom, G., Shaw, J.G. & Thomas, M.S. 2007, "In vivo expression technology identifies a type VI secretion system locus in *Burkholderia pseudomallei* that is induced upon invasion of macrophages", *Microbiology (Reading, England),* vol. 153, no. Pt 8, pp. 2689-2699.

Sherlock, O., Schembri, M.A., Reisner, A. & Klemm, P. 2004, "Novel roles for the AIDA adhesin from diarrheagenic *Escherichia coli:* cell aggregation and biofilm formation", *Journal of Bacteriology,* vol. 186, no. 23, pp. 8058-8065.

Silley, P. & Forsythe, S. 1996, "Impedance microbiology--a rapid change for microbiologists", *The Journal of Applied Bacteriology,* vol. 80, no. 3, pp. 233-243.

Simmons, B.P., Gelfand, M.S., Haas, M., Metts, L. & Ferguson, J. 1989, "*Enterobacter sakazakii* infections in neonates associated with intrinsic contamination of a powdered infant formula", *Infection Control and Hospital Epidemiology : the official journal of the Society of Hospital Epidemiologists of America,* vol. 10, no. 9, pp. 398-401.

Singamsetty, V.K., Wang, Y., Shimada, H. & Prasadarao, N.V. 2008, "Outer membrane protein A expression in *Enterobacter sakazakii* is required to induce microtubule condensation in human brain microvascular endothelial cells for invasion", *Microbial Pathogenesis,* vol. 45, no. 3, pp. 181-191.

Soares, T.A., Straatsma, T.P. & Lins, R.D. 2008, "Influence of the B-band O-antigen Chain in the Structure and electrostatics of the Lipopolysaccharide Membrane of *Pseudomonas aeruginosa.*", *Journal of the Brazilian Chemical Society,* vol. 2, pp. 312-320.

Soler, P., Herrera, S., Rodriguez, J., Cascante, J., Cabral, R., Echeita-Sarriondia, A., Mateo, S. & National Surveillance Network of Spain 2008, "Nationwide outbreak of *Salmonella enterica* serotype Kedougou infection in infants linked to infant formula milk, Spain, 2008", *Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin,* vol. 13, no. 35, pp. 18963.

Spanier, J.G. & Cleary, P.P. 1980, "Bacteriophage control of antiphagocytic determinants in group A streptococci", *The Journal of Experimental Medicine,* vol. 152, no. 5, pp. 1393-1406.

Stahl, F.W., Kobayashi, I. & Stahl, M.M. 1985, "In phage lambda, cos is a recombinator in the red pathway", *Journal of Molecular Biology,* vol. 181, no. 2, pp. 199-209.

Staley, J.T. 2009, "Universal species concept: pipe dream or a step toward unifying biology?", *Journal of Industrial Microbiology & Biotechnology,* vol. 36, no. 11, pp. 1331-1336.

Stanley, T.L., Ellermeier, C.D. & Slauch, J.M. 2000, "Tissue-specific gene expression identifies a gene in the lysogenic phage Gifsy-1 that affects *Salmonella enterica* serovar typhimurium survival in Peyer's patches", *Journal of Bacteriology,* vol. 182, no. 16, pp. 4406-4413.

Stenutz, R., Weintraub, A. & Widmalm, G. 2006, "The structures of *Escherichia coli* O-polysaccharide antigens", *FEMS Microbiology Reviews,* vol. 30, no. 3, pp. 382-403.

Stephan, R., Lehner, A., Tischler, P. & Rattei, T. 2010, "Complete Genome Sequence of *Cronobacter turicensis* LMG 23827, a foodborne pathogen causing deaths in neonates", *Journal of Bacteriology,* Epub ahead of print.

Stevenson, G., Andrianopoulos, K., Hobbs, M. & Reeves, P.R. 1996, "Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid", *Journal of Bacteriology,* vol. 178, no. 16, pp. 4885-4893.

Stock, I. & Wiedemann, B. 2002, "Natural antibiotic susceptibility of *Enterobacter amnigenus, Enterobacter cancerogenus, Enterobacter gergoviae* and *Enterobacter sakazakii strains*", *Clinical Microbiology and Infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases,* vol. 8, no. 9, pp. 564-578.

Suzuki, T., Murai, T., Fukuda, I., Tobe, T., Yoshikawa, M. & Sasakawa, C. 1994, "Identification and characterization of a chromosomal virulence gene, vacJ, required for intercellular spreading of *Shigella flexneri*", *Molecular Microbiology,* vol. 11, no. 1, pp. 31-41.

Swinnen, I.A., Bernaerts, K., Dens, E.J., Geeraerd, A.H. & Van Impe, J.F. 2004, "Predictive modelling of the microbial lag phase: a review", *International Journal of Food Microbiology,* vol. 94, no. 2, pp. 137-159.

Teplyakov, A., Obmolova, G., Badet-Denisot, M.A., Badet, B. & Polikarpov, I. 1998, "Involvement of the C terminus in intramolecular nitrogen channeling in glucosamine 6-phosphate synthase: evidence from a 1.6 A crystal structure of the isomerase domain", *Structure (London, England : 1993),* vol. 6, no. 8, pp. 1047-1055.

Terragno, R., Salve, A., Pichel, M., Epszteyn, S., Brengi, S. & Binsztein, N. 2009, "Characterization and subtyping of *Cronobacter* spp. from imported powdered infant formulae in Argentina", *International Journal of Food Microbiology,* vol. 136, no. 2, pp. 193-197.

Thanassi, D.G., Saulino, E.T. & Hultgren, S.J. 1998, "The chaperone/usher pathway: a major terminal branch of the general secretory pathway", *Current Opinion in Microbiology,* vol. 1, no. 2, pp. 223-231.

Townsend, S.M., Hurrell, E., Gonzalez-Gomez, I., Lowe, J., Frye, J.G., Forsythe, S. & Badger, J.L. 2007a, "*Enterobacter sakazakii* invades brain capillary endothelial cells, persists in human macrophages influencing cytokine secretion and induces severe brain pathology in the neonatal rat", *Microbiology (Reading, England),* vol. 153, pp. 3538-3547.

Townsend, S., Caubilla Barron, J., Loc-Carrillo, C. & Forsythe, S. 2007b, "The presence of endotoxin in powdered infant formula milk and the influence of endotoxin and *Enterobacter sakazakii* on bacterial translocation in the infant rat", *Food Microbiology,* vol. 24, no. 1, pp. 67-74.

Townsend, S., Hurrell, E. & Forsythe, S. 2008a, "Virulence studies of *Enterobacter sakazakii* isolates associated with a neonatal intensive care unit outbreak", *BMC Microbiology,* vol. 8, pp. 64.

Townsend, S.M., Hurrell, E., Caubilla-Barron, J., Loc-Carrillo, C. & Forsythe, S.J. 2008b, "Characterization of an extended-spectrum beta-lactamase *Enterobacter hormaechei* nosocomial outbreak, and other *Enterobacter hormaechei* misidentified as *Cronobacter* (*Enterobacter)* sakazakii", *Microbiology (Reading, England),* vol. 154, no. Pt 12, pp. 3659-3667.

Trent, M.S., Stead, C.M., Tran, A.X. & Hankins, J.V. 2006, "Diversity of endotoxin and its impact on pathogenesis", *Journal of Endotoxin Research,* vol. 12, no. 4, pp. 205-223.

Uchida, T., Gill, D.M. & Pappenheimer, A.M.,Jr 1971, "Mutation in the structural gene for diphtheria toxin carried by temperate phage", *Nature: New biology,* vol. 233, no. 35, pp. 8-11.

Urmenyi, A.M. & Franklin, A.W. 1961, "Neonatal death from pigmented coliform infection", *Lancet,* vol. 1, no. 7172, pp. 313-315.

van Acker, J., de Smet, F., Muyldermans, G., Bougatef, A., Naessens, A. & Lauwers, S. 2001, "Outbreak of necrotizing enterocolitis associated with *Enterobacter sakazakii* in powdered milk formula", *Journal of Clinical Microbiology,* vol. 39, no. 1, pp. 293-297.

Waack, S., Keller, O., Asper, R., Brodag, T., Damm, C., Fricke, W.F., Surovcik, K., Meinicke, P. & Merkl, R. 2006, "Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models", *BMC Bioinformatics,* vol. 7, pp. 142.

Wagner, P.L. & Waldor, M.K. 2002, "Bacteriophage control of bacterial virulence", *Infection and Immunity,* vol. 70, no. 8, pp. 3985-3993.

Wakabayashi, H., Yamauchi, K. & Takase, M. 2008, "Inhibitory effects of bovine lactoferrin and lactoferricin B on *Enterobacter sakazakii*", *Biocontrol Science,* vol. 13, no. 1, pp. 29-32.

Waldor, M.K. & Mekalanos, J.J. 1996, "Lysogenic conversion by a filamentous phage encoding cholera toxin", *Science (New York, N.Y.),* vol. 272, no. 5270, pp. 1910-1914.

Wang, Y., Huang, S.H., Wass, C.A., Stins, M.F. & Kim, K.S. 1999, "The gene locus yijP contributes to *Escherichia coli* K1 invasion of brain microvascular endothelial cells", *Infection and Immunity,* vol. 67, no. 9, pp. 4751-4756.

Wang, Y. & Kim, K.S. 2002, "Role of OmpA and IbeB in *Escherichia coli* K1 invasion of brain microvascular endothelial cells in vitro and in vivo", *Pediatric Research,* vol. 51, no. 5, pp. 559-563.

Warner, P.J., Williams, P.H., Bindereif, A. & Neilands, J.B. 1981, "ColV plasmid-specific aerobactin synthesis by invasive strains of *Escherichia coli*", *Infection and Immunity,* vol. 33, no. 2, pp. 540-545.

Weeks, C.R. & Ferretti, J.J. 1984, "The gene for type A streptococcal exotoxin (erythrogenic toxin) is located in bacteriophage T12", *Infection and Immunity,* vol. 46, no. 2, pp. 531-536.

Weir, E. 2002, "Powdered infant formula and fatal infection with *Enterobacter sakazakii*", *CMAJ : Canadian Medical Association journal = journal de l'Association Medicale Canadienne,* vol. 166, no. 12, pp. 1570.

Wenger, J.D., Hightower, A.W., Facklam, R.R., Gaventa, S. & Broome, C.V. 1990, "Bacterial meningitis in the United States, 1986: report of a multistate surveillance study. The Bacterial Meningitis Study Group", *The Journal of Infectious Diseases,* vol. 162, no. 6, pp. 1316-1323.

Whiteley, M., Bangera, M.G., Bumgarner, R.E., Parsek, M.R., Teitzel, G.M., Lory, S. & Greenberg, E.P. 2001, "Gene expression in *Pseudomonas aeruginosa* biofilms", *Nature,* vol. 413, no. 6858, pp. 860-864.

WHO 2007, "*Enterobacter sakazakii* and other microorganisms in powdered infant formula",

Willis, J. & Robinson, J.E. 1988, "*Enterobacter sakazakii* meningitis in neonates", *The Pediatric Infectious Disease Journal,* vol. 7, no. 3, pp. 196-199.

Willshaw, G.A., Smith, H.R., Scotland, S.M. & Rowe, B. 1985, "Cloning of genes determining the production of vero cytotoxin by *Escherichia coli*", *Journal of General Microbiology,* vol. 131, no. 11, pp. 3047-3053.

Winans, S.C., Elledge, S.J., Krueger, J.H. & Walker, G.C. 1985, "Site-directed insertion and deletion mutagenesis with cloned fragments in *Escherichia coli", J Bacteriol,* vol. 161, pp. 1219-1221.

Workman, C., Jensen, L.J., Jarmer, H., Berka, R., Gautier, L., Nielser, H.B., Saxild, H.H., Nielsen, C., Brunak, S. & Knudsen, S. 2002, "A new non-linear normalization method for reducing variability in DNA microarray experiments", *Genome Biology,* vol. 3, no. 9, pp. research0048.

Worley, M.J., Ching, K.H. & Heffron, F. 2000, "*Salmonella* SsrB activates a global regulon of horizontally acquired genes", *Molecular Microbiology,* vol. 36, no. 3, pp. 749-761.

Wright, A. 1971, "Mechanism of conversion of the salmonella O antigen by bacteriophageepsilon 34", *Journal of Bacteriology,* vol. 105, no. 3, pp. 927-936.

Yamaguchi, T., Hayashi, T., Takami, H., Nakasone, K., Ohnishi, M., Nakayama, K., Yamada, S., Komatsuzawa, H. & Sugai, M. 2000, "Phage conversion of exfoliative toxin A production in *Staphylococcus aureus*", *Molecular Microbiology,* vol. 38, no. 4, pp. 694-705.

Yeo, H.J. & Waksman, G. 2004, "Unveiling molecular scaffolds of the type IV secretion system", *Journal of Bacteriology,* vol. 186, no. 7, pp. 1919-1926.

Zhao, G., Zhang, Q., Yao, L. & Chen, Y. 2008, "Isolation and characterization of bacteriophages of *Enterobacter sakazakii*", *Wei sheng wu xue bao = Acta microbiologica Sinica,* vol. 48, no. 10, pp. 1373-1377