Comparative genome analysis of pathogenic and apathogenic strains of *Carnobacterium maltaromaticum*

By

Tahira Tayyaba

NUST 00000275349

Master of Science in Industrial Biotechnology

Supervisor

Dr. Abdur Rahman

Atta-Ur-Rahman School of Applied Biosciences (ASAB),

National University of Sciences & Technology (NUST),

Islamabad, Pakistan.

2021

Comparative genome analysis of pathogenic and apathogenic strains of *Carnobacterium maltaromaticum*

A thesis submitted as a final year project as a requirement of MS

In

Industrial Biotechnology

By

Tahira Tayyaba

NUST 00000275349

Supervisor

Dr. Abdur Rahman

Atta-Ur-Rahman School of Applied Biosciences (ASAB),

National University of Sciences & Technology (NUST),

Islamabad, Pakistan.

2021

FORM TH-4

National University of Sciences & Technology

MS THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by: **Tahira Tayyaba Reg No. NUST 00000275349 Titled: "Comparative genome analysis of pathogenic and apathogenic strains of *Carnobacterium maltaromaticum*"** be accepted in partial fulfillment of the requirements for the award of **MS in Industrial Biotechnology** degree with (A grade).

Examination Committee Members

1. Name: **Dr. Saadia Andleeb** Signature: _____

2. Name: **Dr. Amjad Ali** Signature: _____

3. Name: **Dr. Muhammad Imran** Signature: _____

Supervisor's name: **Dr. Abdur Rahman** Signature: _____

Date: 15/10/2021

15-10-2021
Date

Head of Department

COUNTERSIGNED

15-10-2021
Date

Dr. Hussnain A. Janjua
Principal
Atta-ur-Rahman School of
Applied Biosciences (ASAB)
NUST, Islamabad

Dean/Principal

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by **Ms. Tahira Tayyaba** Registration No. **NUST 00000275349** of **ASAB** has been vetted by undersigned, found complete in all aspects as per NUST Statutes/ Regulation, is free of plagiarism, errors and mistakes and is accepted as partial fulfilment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature with Stamp: _____

Dr. Abdur Rahman
Assistant Professor
Atta Ur Rahman School of
Applied Biosciences (ASAB)
NUST Islamabad

Name of Supervisor: **Dr. Abdur Rahman**

Date: 15-10-2021

Signature of HOD with Stamp: _____

Dr. Saadia Andleeb
Head of Department (HoD)
Deptt of Industrial Biotechnology
Atta-ur-Rahman School of Applied
Biosciences (ASAB), NUST Islamabad

Date: 15-10-2021

Countersign By

Dr. Hussnain A. Janjua
Principal
Atta-ur-Rahman School of
Applied Biosciences (ASAB)
NUST, Islamabad

Signature (Dean/Principal): _____

Date: 15-10-2021

CERTIFICATE FOR PLAGIARISM

It is to confirm that MS thesis entitled **"Comparative genome analysis of pathogenic and apathogenic strains of *Carnobacterium maltaromaticum*"** of **Ms. Tahira Tayyaba** Reg No. **00000275349** has been examined by me. I undertake that,

1. Thesis has significant new work/ knowledge as compared to already elsewhere. No sentence, table, equation, diagram, paragraph or section has copied verbatim from previous work except when placed under quotation marks and duly referenced.
2. The work presented is original and own work of author i.e. there is no plagiarism. No idea, results or work of others have been presented as author's own work.
3. There is no fabrication of data or results such that the research is not accurately represented in the records. The thesis has been checked using Turnitin, A copy of the originality report attached and focused within the limits as per HEC plagiarism policy and instruction based from time to time.

(Signature of Student)
Tahira Tayyaba
Reg No. 00000275349

Dr. Abdur Rahman
Assistant Professor
Atta Ur Rahman School of
Applied Biosciences (ASAB)
NUST Islamabad

(Supervisor)
Dr. Abdur Rahman
Assistant Professor
ASAB, NUST

DECLARATION

I certify that this research work titled **"Comparative genome analysis of pathogenic and apathogenic strains of *Carnobacterium maltaromaticum*"** is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources has been properly acknowledged/ referred.

_____

Tahira Tayyaba

Dedicated To

My Parents & Siblings

ACKNOWLEDGMENTS

All praise is to Allah Almighty, the Bestowed and the most Affectionate, whose countless blessings are always upon me and whose blessings enable me to complete my thesis. He blessed me with such a strength and endurance that I was able to fulfil my task. All admiration and regards to the Holy Prophet, Hazrat Muhammad (P.B.U.H) for enlightening our conscience with the essence of faith in Allah who is forever a torch of guidance and knowledge for all seekers.

It's my immense pleasure to thank and express my profound regards and gratitude to my respected supervisor, Dr. Abdur Rahman, who has been a great support and source of encouragement during research. He always encouraged me to do better and put my best efforts. Without his moral support, I would not be able to complete my research work and thesis. May Allah always bless him (Ameen).

Besides my supervisor, I am also thankful to my GEC committee members Dr. Saadia Andleeb, Dr. Amjad Ali & Dr. Muhammad Imran for mentoring me and providing insightful comments and encouragement. Their assistance and guidance helped me during my research work.

Also thanks to my seniors, M. Abu Bakar Siddique, Hamza Zafar, Noor Sajid, Sara Azim for their guidance and help. It's time to pay my gratitude to all my lab-partners who had been a great help and support. Thanks to Roomana Ali & Azka Tauqeer for helping me during my work and supporting morally during hardtimes.

Finally, I am obliged to mention my dear friends Zaara Ishaq, Atia Sandhu, Kashaf Khalid, Anam Rabail & Ayesha Rehman whose continuous support and assistance has been a great help during my research work. Thanks for always pumping me up by appreciating my work.

Tahira Tayyaba

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AAI	Average Amino acid Identity
ANI	Average Nucleotide Identity
BRIG	Blast Ring Image Generator
CARD	Comprehensive Antibiotic Resistance Database
CDS	Coding Sequences
CFU	Colony Forming Unit
COG	Cluster of Orthologous Group
CTAS	Cresol Red Thallium Acetate Sucrose
CTSI	Cresol Red Thallium Acetate Sucrose Inulin Agar
DNA	Deoxyribonucleic Acid
EDGAR	Efficient Database framework for comparative Genome Analyses using BLAST score Ratios
EFSA	European Food Safety Agency
GC	Guanine Cytosine
GEI	Genomic Island
GIPSY	Genomic Island Prediction Software
GIT	Gastrointestinal Tract
GRAS	Generally Recognized as Safe
KEGG	Kyoto Encyclopedia of Genes and Genomes
LAB	Lactic Acid Bacteria
MAP	Modified Atmosphere Packaging
MRS	De Man, Rogosa and Sharpe agar
NCBI	National Center for Biotechnology Information
PAI	Pathogenicity Island
PATRIC	Pathosystems Resource Integration Center
PHASTER	**PHA**ge **S**earch **T**ool **E**nhanced **R**elease
TS-YE	Tryptone Soy Yeast Extract Agar
VFDB	Virulence Factor Data Base

ABSTRACT

Carnobacterium maltaromaticum is a diverse lactic acid bacteria (LAB) that is found and isolated from various environments and habitats. It is present in a variety of foods that includes dairy as well as non-dairy products such as meat and fish where they are capable of producing flavors in milk and effective bacteriocins against *Listeria monocytogenes* in processed-food products, respectively. Apart from technological potential in food products, *C. maltaromaticum* is a known pathogen in fish and due to this reason, they are not considered safe by EFSA. This study aims to differentiate between genomes of fish pathogenic strains and non-pathogenic food strains of *C. maltaromaticum*. Therefore, to study the genomic differences in this diversely found specie, the phylogenetic & pan-core genome analysis, as well as safety assessment on the basis of virulence factors, antibiotic resistance and pathogenicity island acquisition was determined using in silico approach. Results show that increase in pseudogenes of SK series of pathogenic strains suggests the transition of pathogenic strains to host-adapted (fish) lifestyle. Moreover, the safety assessment revealed that apathogenic strains cannot be termed completely safe due to the presence of virulence factors and pathogenicity islands and thus, safety might be a strain-specific feature. The genomic variations and safety determined in this study will provide insight into the potential application of dairy strains as probiotics. However, further in-vivo evaluations are required to confirm their potential and safety.

INTRODUCTION

Carnobacterium is a gram positive and facultatively anaerobic, ubiquitous lactic acid bacteria (LAB), which ferment glucose to produce lactic acid (Collins et al., 1987). At present, 12 species of *Carnobacterium* are known. Out of these 12 species, the most common species isolated from natural environment & foods is *C. maltaromaticum*. *C. maltaromaticum* is found in both cold and temperate environments (Jørgen J Leisner et al., 2007).

Among the genomes of *Carnobacterium* that have been sequenced, *C. maltaromaticum* exhibit the largest and variable genomes ranging approximately 3.3 ± 0.7 Mbp, that depicts that the strains have lost as well as gained large fragment of DNA (Iskandar et al., 2017). Also, *C. maltaromaticum* is the only specie that is found in different environments and habitats. Various environments include dairy products, processed food as well as the gut of teleost species (Jørgen J Leisner et al., 2007).

C. maltaromaticum strains are present in a wide variety of foods that include dairy & non-dairy products such as meat and fish. Several studies are being carried out to explore the potential of *Carnobacterium maltaromaticum* in product development due to it's organoleptic as well as its antimicrobial properties against food-borne pathogens. In food products, *C. maltaromaticum* produce certain flavors and may play a significant role in ripening. Moreover, it also acts as a biopreservative, as it inhibits the growth of *Listeria monocytogenes,* which is a well-known pathogen and also a spoilage causing micro-organism (Afzal et al., 2010). In contrast, it is also found in processed food products such as vacuum packaged fish and meat products, where it is responsible for spoilage (Jørgen J Leisner et al., 2007). Moreover, the probiotic potential of *C. maltaromaticum* in fish has also been explored (D.-H. Kim & Austin, 2006; D. Kim et al., 2007;

Robertson et al., 2000). Although, currently there is no commercially applied protective culture of *C. maltaromaticum* (Jørgen J Leisner et al., 2007).

According to United States Food and Drug Administration (USFDA) *C. maltaromaticum* is considered GRAS which means "generally recognized as safe". However, because of its virulent nature in fish, *C. maltaromaticum* is not included in the Qualified Presumption of Safety (QPS) of European Food Safety Agency (Bourdichon et al., 2012; J J Leisner et al., 2012; Ramia et al., 2019)(EFSA, 2012). Indeed, *C. maltaromaticum* has been found to cause infections in fishes, such as teleosts and cartilaginous fish. Such cases of infections by *C. maltaromaticum* has been reported worldwide, including United States, Europe and Africa (Bruno et al., 2013; Hiu et al., 1984; J J Leisner et al., 2012; Loch et al., 2008, 2011; Mohamed et al., 2017; Ramia et al., 2019; Starliper et al., 1992). In rainbow trout, many cases of pseudo-kidney diseases caused by *C. maltaromaticum* have been reported (Loch et al., 2008, 2011). In 2011, Loch et al. studied that some infections caused by *C. maltaromaticum* in different species of fish such as feral and captive rainbow trout as well as some species of salmon leads to mortality. Moreover, in recent studies, more cases of *C. maltaromaticum* infection in sharks, particularly meningitis was reported (Schaffer et al., 2013; Steele et al., 2019). Also, *C. maltaromaticum* strains have been isolated from different organs of infected fish, thus giving insights into the pathogenicity of *C. maltaromaticum* (Roh et al., 2020).

In 1984, first cases of *C. maltaromaticum* (previously known as *C. piscicola)* infection was reported in rainbow trout suffering from damaged kidney (Hiu et al., 1984) as well as from cutthroat trout (Michel et al., 1986). Further in early 90s, the mortality cases of cultured striped bass, channel fish and rainbow trout were reported (Baya et al., 1991; Toranzo et al., 1993).

Hence, *C. maltaromaticum* is potentially virulent in fish, however it also has a significant potential as probiotic (Jørgen J Leisner et al., 2007).

For a long time, the relevance of *C. maltaromaticum* has been underrated as it is unable to grow on media that contains acetate, for example MRS and Rogosa medium. Due to this inability, *C. maltaromaticum* was identified late, initially as *Lactobacillus maltaromaticus* in 1974 (Miller III et al., 1974). However, later the re-classification of this bacterium named it as *C. maltaromaticum* (Mora et al., 2003). *C. maltaromaticum* has the ability to grow and survive at temperatures as low as 2-4°C such as chilled-meat products & cheese, and alkaline pH up to 9.6 (Catherine Cailliez-Grimal et al., 2007; Sakala et al., 2002). This ability of *C. maltaromaticum* to grow as well as survive alkaline environment (high pH) allows it to colonize in soft cheeses, whose pH increase after ripening. Moreover, although limited data is available, but studies describes that *Carnobacterium* has the ability to survive in the mammalian gut.

At the time of this study, there are 33 publicly available genomes of *Carnobacterium maltaromaticum* species listed in National Center for Biotechnology Information (NCBI). Out of these, only 2 genomes are complete, 1 chromosome available, whereas 13 are at scaffold assembly level and 17 are at contigs assembly level. These strains have been isolated from dairy, non-dairy, processed food such as vacuum packaged meat & beef and diseased fish. Comparative genome analysis allows us to discriminate pathogenic species from apathogenic (non-pathogenic) species found in food products and to find the niche-specific genomic features and genetic relatedness in *C. maltaromaticum* strains. The aim of this study is to analyze and compare the pathogenic and apathogenic strains of *C. maltaromaticum* by focusing on evolutionary relationship, pan and core genome, safety assessment and genome plasticity. Following are the objectives of this study:

OBJECTIVES:

1. Evaluation of genomic differences between fish pathogenic strains and apathogenic food strains of *C. maltaromaticum*.
2. Inter and intra-strain comparative genomic analysis for safety assessment of *C. maltaromaticum* strains.

LITERATURE REVIEW

2.1 Taxonomy

In 1974, *C. maltaromaticum* was discovered relatively late, initially as *Lactobacillus maltaromaticus* (Miller III et al., 1974). However, in 2003, it was reclassified as *C. maltaromaticum* (Mora et al., 2003). The non-typical heterofermentative species isolated from processed food-products such as meat, fish & chicken, previously classified as *Lactobacillus*, were suggested to be re-classified as the genus *Carnobacterium*, due to its inability to grow on agar containing acetate (Collins et al., 1987; WALTER P Hammes & Hertel, 2006). On the basis of niche and habitat, *C. maltaromaticum* was categorized into two groups: Group I which includes species originated from animals and their products; Group II includes species which occur in natural and cold environment (W P Hammes & Hertel, 2009).

Carnobacterium is a rod-shaped LAB bacterium that ferment glucose and results in the production of $_{L}$ (+)-lactic acid. Their cell wall composition includes meso-diaminopimelic acid, which is a characteristic of *Carnobacterium*. Sequencing of 16S rDNA shows that in a phylogenetic analysis, *Carnobacterium* forms an evolutionary distinct group and significantly differ from other lactic acid bacteria (Wallbanks et al., 1990).

Initially, *Carnobacterium* contained only four strains belonging to these species: *C. divergens, C. piscicola, C. gallinarum* and *C. mobile. C. piscicola* and *Lactobacillus maltaromaticus* whose isolation source is malty-flavored milk are heterotypic synonyms (Collins et al., 1987; Miller III et al., 1974; Mora et al., 2003). However, this bacterium has not been commercialized as non-starter LAB. The basic properties of *C. maltaromaticum* has been demonstrated (Afzal et al., 2010).

2.2 Characteristics of *Carnobacterium*

Currently, the genus *Carnobacterium* contain 12 species that had been found from multiple environments such as cold and temperate environment, as well as from animals GIT (gastrointestinal tract), from fish and meat associated food products and dairy products.

Among the genus *Carnobacterium*, *C. maltaromaticum* and *C. divergens* are the most common species isolated from diverse sources (Jørgen J Leisner et al., 2007). Technological potential *C. maltaromaticum* and *C. divergens* in dairy products is reported (Laursen et al., 2005; Millière et al., 1994; Morea et al., 1999). Also, they have been isolated from habitats that are associated with animals. In meat & fish-associated food products, these species were dominantly found among the bacterial communities (Chaillou et al., 2015; Duan et al., 2016; Fougy et al., 2016; Jääskeläinen et al., 2016). Due to their ability of adaptation in various environments, it is necessary to evaluate the genomic traits in these strains.

C. maltaromaticum possess a large genome, ranging in size 3.3 Mbp ± 0.7. The large genome size suggests the reason behind adaptation of this specie in diverse environments (Jørgen J Leisner et al., 2007). Conversely, the size of the genome does not always predict the colonization ability in various environments as *C. divergens* possess a relatively smaller genome (approximately 2.7 Mbp), but still found in diverse environments (Catherine Cailliez-Grimal et al., 2013; Remenant et al., 2016; Sun et al., 2015).

They are mesophilic (i.e they can grow at moderate temperatures), whereas few species are also able to grow and survive at temperatures as low as 0°C. Moreover, few species are halotolerant and can grow with 8% NaCl. Whereas, some of the species are alkaliphilic that can grow at pH 9.5 (C Cailliez-Grimal et al., 2014; Pikuta & Hoover, 2014).

2.3 Isolation and enumeration of *C. maltaromaticum*

Food & dairy products-associated bacterial strains isolation and identification seem to be a challenging task in microbial taxonomy. Several different types of culturing media for the isolation of *Carnobacterium*, includes non-selective, semi-selective and selective media (WALTER P Hammes & Hertel, 2006).

To detect and selectively isolate of *Carnobacterium*, CTAS medium abbreviated as Cresol Red Thallium Acetate Sucrose was recommended. However, this media may allow the growth of some other gram-positive bacteria such as *Leuconostoc* and *Enterococcus.* The culture is incubated at temperature 30°C for 24-48 hours, or at a bit lower temperature, i.e 25°C for 3-4 days (Holzapfel, 1992). Other media include EBRER medium, which require relatively low temperature, i.e. 7°C for incubation for 10 days, and Cresol Red Thallium Acetate Sucrose Inulin Agar (CTSI), which has certain limitations such as its red color, prolonged time of incubation, that is 4 days and complexity (Wasney et al., 2001).

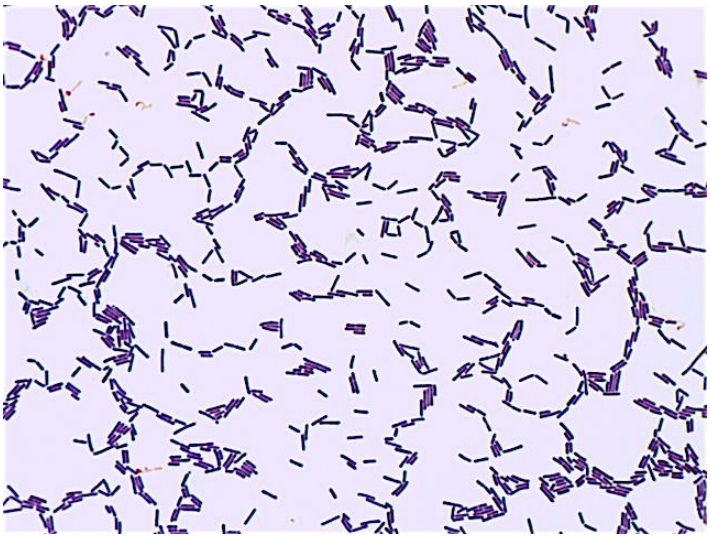


Figure 1. *C. maltaromaticum* cells on the medium agar plate incubation at 30°C for 3 days.

(Miller III et al., 1974; Mora et al., 2003)

As *C. maltaromaticum* is able to grow on pH upto 9.6, it can be isolated on a medium based on pH and pattern of antibiotic resistance. Therefore, a new selective medium, known as CM, was suggested, which is based on TS-YE agar and supplemented with antibiotics including vancomycin, gentamycin and nalidixic acid with a pH of 8.8. It is highly selective for dairy strains of *C. maltaromaticum* after incubation at 25°C for 36-48 hours (Hélène Carole Edima et al., 2007). Thus, this selective medium was found suitable for the *C. maltaromaticum*.

For enumeration, methods such as flow cytometry, epiflourescent microscopy and *in situ* hybridization were used for *C. divergens* and *C. maltaromaticum*. These techniques allow rapid and selective enumeration of mixed bacterial cultures, when coupled together (Connil et al., 1998).

2.4 Occurence of different species of *Carnobacterium* in diverse environments

Several genera and various species of Lactic Acid Bacteria (LAB) are explored for many years due to their ability to ferment food and exert many benefits on health, particularly as probiotics. LAB are able to grow and colonize in animals-originated foods, such as meat & beef, fish as well as dairy products. Several LAB are also the part of microbiota in gut and vagina (Douglas & Klaenhammer, 2010; Douillard & De Vos, 2014).

Lactic acid bacteria can be classified as either specialist or generalist bacteria. The specialist bacteria are those that exert low diversity at genetic level and are capable of being used as starter cultures for some fermented food products (Delorme et al., 2010). These bacteria have undergone massive loss of genes, particularly those responsible for biosynthetic pathways (Douglas & Klaenhammer, 2010). Alternatively, the functions of lost genes were executed by other acquired genes responsible for transportation of amino acids and carbohydrates that consequently makes the bacteria able to grow in nutrients-rich environment of fermentation (Lorca et al., 2007). Such

changes of gain and loss of genes were associated with the specialization towards food environment, particularly illustrated in dairy strains. Different LAB exerts certain genomic changes according to their ecological niche specialization (Iskandar et al., 2017)

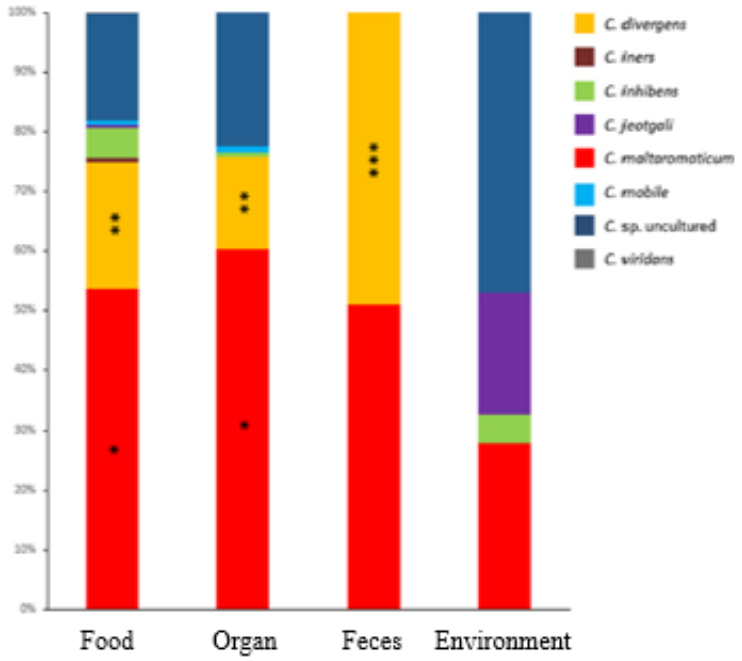


Figure 2. Relative abundance of each *Carnobacterium* species obtained from 16S metagenomic analyses of 681 samples originating from food. One asterisk indicates significant difference ($P < 0.0001$) compared to all other *Carnobacterium* sp, two asterisks indicate significant difference ($P < 0.01$) compared to *C. iners, C. inhibens, C. jeotgali, C. mobile* and *C. viridans*. Three asterisks indicate significant difference ($P < 0.0001$) compared to all *Carnobacterium* sp. Except *C. maltaromaticum* (Iskandar et al., 2017).

Iskander et al., in 2017, conducted a study and found that *C. maltaromaticum* is the most dominant species in animal-derived foods, for example, dairy products, meat and beef, fish products as well as in the microbiota of GIT tract, followed by *C. divergens*. The evolutionary relationship between these two species adapted in animal, suggests that these species are more

closely related to each other as they belong to one of the two deeply branched lineages. Whereas, the species rarely associated with animals belong to the other lineage. Genomic analysis of these species suggests that the lineage species of animal origin exhibit large secretome as compared to species belonging to another lineage. Also, the proteome of *C. maltaromaticum* and *C. divergens* is highly diverse & include genes responsible for adaptation in animal environment (Iskandar et al., 2017).

2.5 *Carnobacterium maltaromaticum* in dairy products

The presence of *Carnobacterium* in dairy, i.e milk associated, products as well as other food products is not reported frequently. The reason is the inability of *Carnobacterium* to grow on media containing acetate, such as MRS or Rogosa agar, which are commonly used for the enumeration of LAB. Acetate inhibit the growth of *Carnobacterium* and thus these acetate-containing media reduce concentrations of *Carnobacterium* in food (Chenoll et al., 2007; WALTER P Hammes & Hertel, 2006; Leblanc et al., 1997; Sakala et al., 2002; Susiluoto et al., 2003).

As far as dairy products are concerned, *C. maltaromaticum* is found in soft cheeses made from pasteurized and raw milk of ewe, goat or cow milk. They are also able to survive at high alkaline pH in cheese during ripening (Catherine Cailliez-Grimal et al., 2007; Millière et al., 1994).

High genetic diversity exists between different dairy isolated strains of *C. maltaromaticum* & starter Lactic acid bacteria, for example *Streptococcus thermophiles* and *Lactococcus lactis* (Delorme et al., 2010; Passerini et al., 2010; Rahman et al., 2014) which increase the potential for *C. maltaromaticum* for innovation in dairy industry. *C. maltaromaticum* has the potential for culture innovation (Iskandar et al., 2016). Furthermore, this specie is able to survive the process of acidification without interfering with the commercial starters (H C Edima et al., 2008). *C.*

maltaromaticum has the potential in dairy industry, where it can be employed as a adjunct culture because:

1. It has the ability to colonize cheeses and can reach up to the concentration of 10^9 CFU/g suggesting that it can survive during ripening of cheese and do not compete with other starter LAB (Afzal et al., 2010).
2. In soft cheeses, this bacterium does not produce significant amount of harmful compounds, for instance tyramine and histamine (Afzal et al., 2010)
3. It can produce different flavor compounds in dairy products such as 3-methylbutanal.
4. It is able to produce bacteriocins that can prevent the growth of *Listeria monocytogenes*, which is a well-known food-borne pathogen.
5. It possess insignificant or slight anti-inflammatory properties in comparison to other probiotic lactic acid bacteria (Rahman, Gleinser, et al., 2014).
6. This bacterium is rarely related to any human disease and no case of infection in human has been reported or linked directly to using *C. maltaromaticum*-containing dairy products (Afzal et al., 2010).

C. maltaromaticum is widely found in different foods, as shown in Table 1. *C. maltaromaticum* was firstly isolated from malty flavored-milk, in which aldehydes, i.e. 3-methylbutanal, 2-methylbutanal, 2-methylpropanal were responsible for flavor (Miller III et al., 1974). Moreover, studies have suggested that the industrialization of dairy manufacturing resulted in reduction in diversity (Bachmann et al., 2011)

Table 1. *Carnobacterium maltaromaticum* in food products

Product	Sensory impact	Reference
Dairy		
Skimmed Milk	Malty flavor	Miller et al., 1974
Soft cheeses	No off-flavours	Millière and Lefebvre 1994; Millière et al., 1994; Cailliez-Grimal et al., 2007
Mozzarella	No off-flavours	Morea et al., 1999
Fish and shellfish		
MAP Salmon	No off-flavours	Emborg et al., 2002
Cold-smoked Salmon	Malt culture medium	Leroi et al., 1998
	Slightly sour, sweet/ nauseous (vacuum) Oxydized, bitter and fishy/malty (MAP)	Paludan-Muller et al., 1998
	No spoilage odour detected	Duffes et al., 1999a
Meat		
MAP Beef muscles	VOC Aldehydes, lactones, sulphur compounds	Ercolini et al., 2009
MAP Broilers	Not determined	Vihavainen et al., 2007
Marinated pork	No spoilage odour detected	Schirmer et al., 2009

MAP: Modified atmosphere packed.
VOC: Volatile Organic Compounds.

Also, the key factor considered essential to be a part of gut colonization is the ability to hydrolyze bile (Kleerebezem et al., 2010; Seedorf et al., 2014). In a study, it was reported that *C. maltaromaticum* LMA 28, which has been isolated from soft ripened cheese, possess genes that are responsible for the survivability of this bacterium in the gastrointestinal (GIT) tract of mouse (Rahman, Gleinser, et al., 2014; Sun et al., 2015).

2.5.1 Cheese ripening

In 1994, for the first time, *C. maltaromaticum* was reported to be found in Brie cheese, a type of French ripened soft cheese (Millière et al., 1994; Millière & Lefebvre, 1994) and later in 2007, it was confirmed (Catherine Cailliez-Grimal et al., 2007). Analysis of 30 samples of French ripened cheese, 10 was found to have *C. maltaromaticum* and 3 isolated from cheese exhibited anti-bacterial activity against *Listeria*. When cheese samples were analyzed, *C. maltaromaticum* constituted the flora able to survive at low temperature at higher pH levels (Catherine Cailliez-Grimal et al., 2007). It also plays a role in ripening soft cheese (H C Edima et al., 2008). Another study reported that this species also plays a role in citrate fermentation in the Mozzarella cheese and constitutes 70% of the curd (Morea et al., 1999).

Some examples of cheeses in which *C. maltaromaticum* is found: Camembert, Epoisse, Brie, Picodon, Pérail and Petit Munster & Petit Livarot (Catherine Cailliez-Grimal et al., 2007; Millière et al., 1994).

C. maltaromaticum is a slow acidification species in comparison to other commercialized lactic acid bacterial strains that are used as starter culture, for instance *Lactococcus lactis* and *Streptococcus thermophilus*. However, it has been found that the psychrotrophic *C. maltaromaticum* LMA 28 survive acidic conditions (low pH) when co-cultured with *L. lactis* DSMZ 20481 or *S. thermophilus* INRA 302 (H C Edima et al., 2008). Due to this property of slow-acidification, *Carnobacterium* strains are not usable as starter culture, and hence they can be used as non-starter LAB (NSLAB). However, they can play a role in cheese ripening (Briggiler-Marcó et al., 2007) as well as in the development of aroma in soft cheeses.

2.5.2 Aroma and flavor compounds

In fact, *C. maltaromaticum* is significantly known for producing aroma and flavor compounds, for example 3-methylbutanal as a result of catabolism of leucine. These aldehydes are reported to be responsible for malty flavor in milk, which is a characteristic feature of *C. maltaromaticum* (Larrouture-Thiveyrat & Montel, 2003; Laursen et al., 2006; Miller III et al., 1974; Smit et al., 2009).

Table 2. Compounds produced by catabolism of branched chain amino-acids by *C. maltaromaticum*

Branched-chain amino acids	**Compounds**
Leucine	α-ketoisocaproicacid,3-methylbutanal and 3-methylbutanol
Isoleucine	3-methylbutanoicacid, 2-methylbutanal
Valine	2-methylbutanol,2- methylpropanal and 2-methylpropanol

The alcohols produced by this catabolism reaction, that is 2-methylbutanol, 3-methylbutanol and 2-methylpropanol are responsible for malty, alcoholic and fruity aroma (Thierry & Maillard, 2002). Whereas, the other products of catabolism, i.e. branched-chain acids, that is, 2-methylbutanoic acid, 3-methylbutanoic acid and 2-methylpropanoic acid) produce rotten-fruit like, rancid and putrid ester flavors (Marilley & Casey, 2004; Thierry & Maillard, 2002). Thus, *C. maltaromaticum* has the potential to play a role in developing the taste and texture during cheese ripening (Millière et al., 1994; Millière & Lefebvre, 1994).

2.5.3 Lactose metabolism

Lactose is the main source of carbon in milk. Iskander et al., in 2016, investigated the genes for lactose metabolism in *C. maltaromaticum*. Results showed that *C. maltaromaticum* possess genes

that are involved in lactose and galactose metabolism pathways, namely Leloir and Tagatose-6-phosphate (Tagatose-6P) pathways. Also, the strains may have genes associated with one or both pathways. Moreover, the study suggests that lactose and galactose metabolism genes, i.e *lac* and *gal* genes, evolved as a result of horizontal gene transfer, duplication and translocation, which consequently contribute to the adaptation of *C. maltaromaticum* in wide range of environment and variable properties related to technological potential (Iskandar et al., 2016).

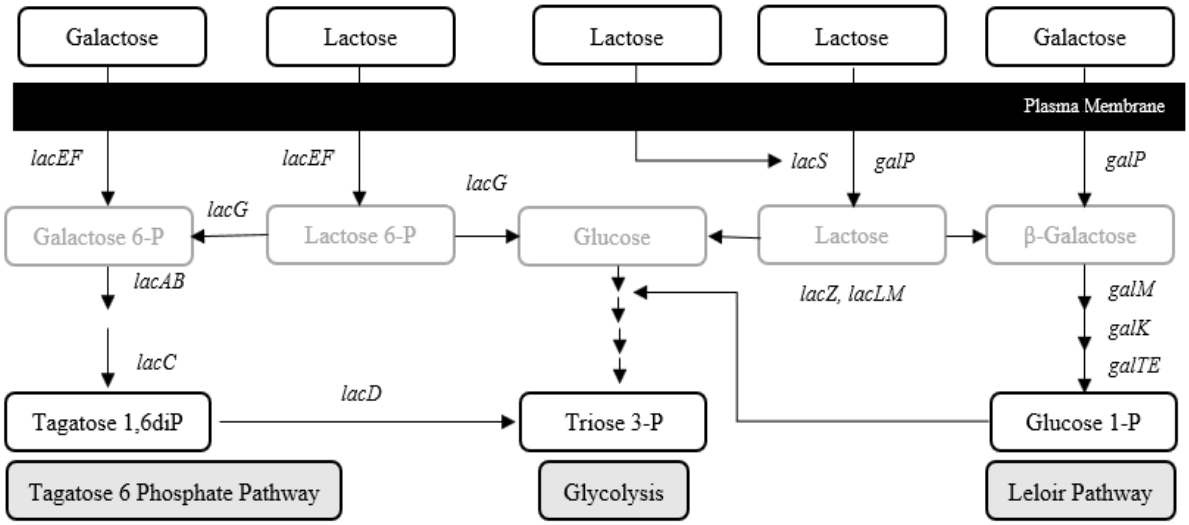


Figure 3. Tagatose 6-Phosphate and Leloir pathway (Iskandar et al., 2016)

2.6 *Carnobacterium maltaromaticum* in non-dairy products

Two closely related animal-associated strains of *Carnobacterium,* i.e. *C. divergens* and *C. maltaromaticum* has the ability of growing in meat products at low temperatures ranging from 2 to -1.5°C (Jones, 2004; McMullen & Stiles, 1993; Sakala et al., 2002). *C. divergens* and *C. maltaromaticum* predominates the gram + and other lactic acid bacteria in the raw meat environment such as beef, lamb, poultry and pork. These two species are found and survive irrespective of the environmental condition in the meat products, such as aerobic environment,

vacuum-packaged, MAP conditions with concentration of CO_2/N_2 ranging from 10/90 to 80/20 (Barakat et al., 2000; Björkroth et al., 2005; GRANT & PATTERSON, 1991; Jones, 2004; Laursen et al., 2005; McMullen & Stiles, 1993; Sakala et al., 2002; Shaw & Harding, 1984; Susiluoto et al., 2003; Vihavainen & Björkroth, 2007).

C. maltaromaticum found in processed food such as vacuum packaged meat and beef product, are of significant importance as they are responsible for certain antibacterial activities, as they produce bacteriocins and organic acids. Zhang et al., in 2019, demonstrated that *Carnobacterium* produce organic acids, such as formate and acetate, which are associated with anti-bacterial activity of *Carnobacterium* in conditions of vacuum-packaged meat products. However, lactate was previously believed to be of importance, for its inhibitory activity on chilled meat products (Zhang et al., 2019).

Several studies demonstrate the preservation of processed ready-to-eat food products such as cooked chicken, minced meat and refrigerated food, for example cold-smoked salmon to protect them from *Listeria monocytogenes*. Two ways of preservation are used: Firstly, the bacteriocin, produced by Carnobacterium is added to the food. Secondly, the bacteriocin-producing *Carnobacterium* is directly added to the food (Afzal et al., 2010).

Recently, a study carried out by Bergem et al., in 2020 reported that *C. maltaromaticum* EBP3019 and *C. maltaromaticum* SF668 are highly effective in inhibiting *Listeria monocytogenes.* The study provides insights into 16kDa unmodified bacteriocin and five different bacteriocins produced by *C. maltaromaticum* EBP3019 and *C. maltaromaticum* SF668, respectively (Begrem et al., 2020).

In 2011, Casaburi et al. compared multiple *C. maltaromaticum* strains activity of spoilage in both vacuum-packaged & air-stored beef. Their study suggested that volatile organic compounds with low sensory effect are produced by *C. maltaromaticum*, independent of the strain. The spoilage effect of these compounds on meat is negligible. Moreover, *Carnobacteria* is being explored as protective culture for inhibition of pathogens and organisms responsible for spoilage, in dairy and non-dairy (meat, seafood) products (Casaburi et al., 2011).

C. maltaromaticum is known to be part of fish intestinal microbiota and isolated from different fishes such as cod, trout, charr and salmon (Jørgen J Leisner et al., 2007). Contrary to that, *C. maltaromaticum* is rarely found in the mammalian intestine. Rahman et al., in 2014, demonstrated that in addition to fish, *Carnobacterium* might exert certain immunomodulatory properties in mammals as well (Rahman, Gleinser, et al., 2014).

2.6.1 Bacteriocin production

Carnobacteria produce antibacterial compounds, such as bacteriocins and organic acids, in order to adapt in conditions of meat microbiota. Bacteriocins are antimicrobial peptides, synthesized by the ribosomes. Depending on the structural properties, bacteriocins are classified into three major groups: Class I, II and III (Alvarez-Sieiro et al., 2016; Cotter et al., 2005, 2013).

Class I bacteriocins are peptides of small size (<5kDa) that have undergone post-translational modification. Example include lantibiotics, that contain an unusual amino acid lanthionine (Lan). Their activity varies according to their structure (McAuliffe et al., 2001). Class II bacteriocins are unmodified peptides of small size as Class I (<5kDa) but are permeable through bacterial membranes. On the basis of structure, they are further classified into four groups: IIa, IIb, IIc and IId. Class I & II bacteriocins exhibit thermostable properties. Class III bacteriocins are relatively

larger (>30kDa), and generally comprises bacteriolysins. They are thermoliable (Cotter et al., 2005, 2013).

Class IIa bacteriocin, namely Carnobacteriocin, inhibit *Listeria* by antimicrobial activity which is carried out by creation of pore, disruption in potential of membrane and also by exudation of substances that have low molecular weight present inside (Drider et al., 2006; Suzuki et al., 2005).

Carnobacteriocins belonging to class II bacteriocins transport with the help of ABC transport system (ATP-binding cassette) (Drider et al., 2006). Whereas, transport of divergicin A is based on cellular secretion pathway (sec pathway) and does not depend on secretion protein (Worobo et al., 1994).

Divercin V41 and Divergicin M35 (Class IIa bacteriocins), and Divergicin A (Class IId bacteriocins) have been isolated from *C. divergens* strains (Metivier et al., 1998; Tahiri et al., 2004; Worobo et al., 1994). However, in comparison to *C. divergens*, *C. maltaromaticum* produce a large number of bacteriocins. Examples include Carnolysin (Class I), Carnobacteiocin B2 and BM1, Piscicolin 126 and CS526, and maltaricin CPN belonging to class IIa, and Carnobacteriocin X (Class IId) (Bhugaloo-Vial et al., 1996; Hammi et al., 2016; Herbin et al., 1997; Jack et al., 1996; Martin-Visscher et al., 2008; Quadri et al., 1994; Tulini et al., 2014; Yamazaki et al., 2005).

Table 3. Bacteriocins produced by *C. maltaromaticum* (Jørgen J Leisner et al., 2007)

Bacteriocin (Class)	Gene	Location of gene	Reference
Carnobacteriocin A (IIc)/ Piscicolin 61	*cbnA*	Plasmid	Worobo et al. (1994) Holck et al. (1994)
Carnobacteriocin BM1 (IIa)/ Carnobacteriocin B1/ Piscicosin V1b/ Carnocin CP51	*cbnBM1*	Chromosome	Quadri et al. (1994) Bhugaloo-Vial et al. (1996) Herbin et al. (1997)
Carnobacteriocin B2 (IIa)	*cbnB2*	Plasmid	Quadri et al. (1994) Wang et al. (1999)
Carnocin CP52/ A9b			Herbin et al. (1997) Nilsson et al. (2002)
Piscicolin 126 (IIa)/ Piscicocin V1a	*pisA*	Chromosome	Jack et al. (1996) Bhugaloo-Vial et al. (1996) Gursky et al. (2006)
Piscicocin CS526 (IIa)	ND	ND	Yamazaki et al. (2005)

2.6.2 Preservation of food

Since *Carnobacterium* is significantly important for producing bacteriocins (Jørgen J Leisner et al., 2007). Different types of bacteriocins are produced by different strains of *Carnobacterium*. These bacteriocins have the ability to inhibit the spoilage and pathogenic micro-organism such as *Listeria*. *Carnobacteria* and *Listeria* exhibit same behaviors toward temperature and pH and are psychrophilic. Therefore, due to bacteriocin producing ability, the use of *Carnobacterium* isolated from food such as cheese and other, is very well reported. In several refrigerated foods, bacteriocin produced by *Carnobacterium* results in inhibition of *Listeria* during critical phases (Buchanan & Klawitter, 1991; Catherine Cailliez-Grimal et al., 2007; Herbin et al., 1997).

2.7 *Carnobacterium maltaromaticum* as pathogen (in diseased fish)

C. maltaromaticum is a bacterium of technological interest as it possesses anti-bacterial properties, inhabit the intestines of fish, for example GIT of teleost, and might has a potential to play a role as probiotic. Besides its technological potential, *C. maltaromaticum* is a known fish pathogen, and infect fishes in cold water (teleosts) under stress conditions (such as spawning and handling events) (Loch et al., 2008, 2011).

C. maltaromaticum has been known to cause infections in fish. Fish infected with *C. maltaromaticum* are reported to suffer from septicemia, visceral congestion, pseudokidney disease, splenomegaly, internal hemorrhages, thickened swim-bladder walls and muscular abscesses.

Not long ago, *C, maltaromaticum* have been found & isolated from organs of diseased fish, that is from brain and ear of *Alopias vulpunis*, commonly known as common thresher and *Lamna ditropis*, which is salmon shark. While strandings of thresher shark are recent, the strandings of salmon sharks have long been known, that suggest a long-term relation of *C. maltaromaticum* with sharks. More interesting is the fact that some strains of *C. maltaromaticum* have applications in food industry, due to their probiotic properties in fish, antimicrobial properties in vacuum packaged products and flavoring properties in dairy products (Steele et al., 2019).

Common thresher and salmon shark are both top predators that belongs to Lamniformes order and migrate between Northeast Pacific Ocean coast and offshore waters and also use the continental shelf as a nursery during juveniles (Cartamil, 2009). Severe meningitis, as well as acute otitis, have been reported in the brain and ear of infected/diseased threshers and salmon sharks, that are infected by *C. maltaromaticum* (Schaffer et al., 2013). In a recent study, Steele et al., sequenced 9 strains of *C. maltaromaticum* (SK_AV1, SK_AV2, SK_AV3, SK_AV4,

SK_AV5, SK_AV6, SK_LD1, SK_LD2, SK_LD3) that have been originated from the brain and ear tissues of diseased fish and found high degree of pseudogenization, depicting the transition of host and lifestyle of bacterium (Steele et al., 2019).

Roh et al., in 2020, evaluated the pathogenicity of *C. maltaromaticum* strains found in diseased fish. It was found that, those strains that originated from diseased fish were able to cause high histopathological changes and mortality. The *C. maltaromaticum* strains studied were found to exhibit more than 90% of KEGG orthologs id, however the genes, that is *wecC* and *xtmA* were specifically present only in strains isolated from diseased fish. Also, only the strains from infected fish possessed two paralogs of *wecC*, which is involved in the production of a major component of teichuronic acid (a virulence factor), that is D-mannosaminuronic acid. Their study also describes the potential of *wecC* gene as virulence factor.

In 2018, Korea, *C maltaromaticum* strain was isolated from infected rainbow trout. The body surface of trout was infected with ulcerative lesions. And confirmed that *C. maltaromaticum* 18ISCm was able to cause mortality when infected in healthy rainbow trout, thus confirming the pathogenicity of this strain. Altogether, their study suggests that *C. maltaromaticum* strains isolated from infected fish differ from the food-associated strains depending upon pathogenicity and presence/absence of certain virulence factors (Roh et al., 2020).

In another study by Leisner et al., in 2012 suggested that *C. maltaromaticum* ATCC35586 carry genes that might be virulent and encode products that are responsible for adhesion to fibronectin and collagen, synthesis of capsule, modification of cell wall, mechanisms involved in iron scavenging, invasion, hemolysis and resistance to toxic compounds, that might cause infection in fish. Furthermore, the potent virulence factors were found to have homology with genes in *Listeria* (J J Leisner et al., 2012).

A recent study tested the pathogenicity of *C. maltaromaticum* strains, i.e. 18ISCm, ATCC35586 & DSM 20342 (isolated from dairy origin) and found the highest mortality (almost 80%) in fish infected with 18ISCm strain, 60% with ATCC35586 and 0% from DSM 20342 (Roh et al., 2020).

The study described that only the fish that were exposed with pathogenic strains: *C. maltaromaticum* ATCC35586 and 18ISCm exhibited infectious conditions such as reddish swim bladders, abdominal distension with clear ascites and abnormal behavior, for instance loss of equilibrium and erratic swimming. The histopathology of the organs of fish infected with *C. maltaromaticum* ATCC 35586 showed that kidney, swim bladder and spleen were mainly affected by the infection. Moreover, only the diseased fish isolated strains: *C. maltaromaticum* 18ISCm, *C. maltaromaticum* ATCC 35586 and *C. maltaromaticum* SK_AV1 were, particularly, found to contain three genes particularly: *wecC*, *panE* and *xtmA*.

In this study, they found that *wecC* gene encoding the UDP-N-acetyl-D-mannosamine dehydrogenase was absent in most *C. maltaromaticum* strains. However, it was present in the diseased fish associated strains. Thus, this indicate that this gene might be involved in the pathway responsible for pathogenicity. On the basis of comparative genomics, genes related to D-mannosaminuronic acid, precursor for synthesis of a virulence factor, namely Teichouronic acid, were found particularly in strains isolated from diseased and infectious fish.

According to this study by Roh et al., in 2020, only the fish infected with *C. maltaromaticum* 18ISCm and *C. maltaromaticum* ATCC 35886 showed symptoms of diseases and even led to death of rainbow trout. Conversely, the fish to which high dose of *C. maltaromaticum* DSM 20342 (derived from dairy products) was administered, neither show any symptoms of the diseases nor mortality. This study illustrates significant difference between the pathogenic & apathogenic strains of *C. maltaromaticum* (Roh et al., 2020).

2.8 Safety assessment

Despite bacteriocins production and other probiotic-related properties, there are limitations in the applications of *C. maltaromaticum* in several food products as bioprotective culture. This is due to the ability of *C. maltaromaticum* to break tyrosine and produce tyramine. However, *C. maltaromaticum* LMA28 isolated from soft cheese didn't produce tyramine and histamine (Hélène Carole Edima et al., 2007). Thus, the concentration of tyramine that is produced, depends on the type of strain & the type of food (Jørgen J Leisner et al., 2007; Masson et al., 1996). One strain, *C. maltaromaticum* CB1 is considered GRAS (Generally recognized as Safe, GRN 00159), for applications in ready-to-eat products of meat, since 2005.

C. maltaromaticum is a known pathogen in fish but has not been known to cause infection in human and is therefore, not regarded as opportunistic pathogen. However, one case has been reported in which a strain *C. maltaromaticum* has been found from pus of human (Chmelař et al., 2002).

MATERIALS & METHODS

3.1 Genomes, genome validation and properties

3.1.1 Sequence retrieval from NCBI

At the time of this study, all publicly available genomes of *Carnobacterium maltaromaticum* were obtained and genome sequence assemblies (Genbank and Fasta) were retrieved from National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/genome/browse/). A total of 33 genomes were available. Only 2 were complete genomes, one was at chromosome level, and the other 30 were draft genomes assembled at scaffold and contig level. The isolation source of retrieved genomes was from: dairy products, processed food, fish products and diseased fish. The properties, accession numbers and origin of isolation of the retrieved *C. maltaromaticum* strains that are analyzed in this study are described in Table 4.

Table 4. *Carnobacterium maltaromaticum* genomes retrieved from NCBI for analysis

Food associated strains						
S. No.	Accession No.	Strain	Assembly Level	No. of Contigs	Isolation source	
1	HE999757.1	LMA 28	Complete	1	Ripened soft cheese	Dairy isolates
2	JQMX00000000	DSM 20342 MX5	Contig	5	Milk with malty flavor	
3	JQBG01000000	DSM 20342	Scaffold	111	Milk with malty flavor	
4	CP016844 Plasmids (4): CP016845-48	TMW 2.1581	Complete	5	Poultry meat-spoilage	Non-Dairy isolates
5	NRQC01000000	A7	Contig	20	Vacuum packaged beef	
6	NRPU01000000	A15	Contig	23	Vacuum-packaged beef	
7	NRPV01000000	A14	Contig	23	Vacuum-packaged beef	
8	NRPT01000000	A16	Contig	21	Vacuum-packaged beef	
9	NRQE01000000	A5	Contig	22	Vacuum-packaged beef	
10	NRPR01000000	A18	Contig	22	Vacuum-packaged beef	
11	NRQD01000000	A6	Contig	23	Vacuum-packaged beef	
12	NRQG01000000	A3	Contig	28	Vacuum-packaged beef	
13	NRQI01000000	A1	Contig	29	Vacuum-packaged beef	
14	NRPS01000000	A17	Contig	21	Vacuum packaged meat	
15	JQBU01000000	DSM 20722	Scaffold	35	Vacuum packaged meat	
16	WNJR01000000	SF668	Contig	88	Fish product	
17	WNJS01000000	EBP3019	Scaffold	121	Fish product	
18	CVNA00000000	ML_1_97	Contig	229	Fresh Salmon	
19	CVMZ00000000	3-18	Contig	160	Pork meat product	

Diseased fish associated						
S. No.	Accession No.	Strain	Assembly Level	No. of Contigs	Isolation source	
20	Chromosome: CP045040 Plasmids (4): CP045041-44	18ISCm	Chromosome	5	Rainbow trout	Diseased Fish
21	AGNS00000000	ATCC 35586	Contig	74	Diseased salmon	
22	BJOJ01000000	NBRC 15685	Contig	108	Diseased adult cutthroat trout	
23	JQBV01000000	DSM 20730	Scaffold	56	Diseased rainbow trout	
24	PKFM01000000	SK_AV1	Scaffold	18	*Alopias vulpinus* (Common thresher) – Brain	
25	PKFL01000000	SK_AV2	Scaffold	19	*Alopias vulpinus* (Common thresher) – Ear	
26	PKFK01000000	SK_AV3	Scaffold	18	*Alopias vulpinus* (Common thresher) – Brain	
27	PKFJ01000000	SK_AV4	Scaffold	18	*Alopias vulpinus* (Common thresher) – Ear	
28	PKFI01000000	SK_AV5	Scaffold	33	*Alopias vulpinus* (Common thresher) – Brain	
29	PKFH01000000	SK_AV6	Scaffold	19	*Alopias vulpinus* (Common thresher) – Ear	
30	PKFG01000000	SK_LD1	Scaffold	18	*Lamna ditropis* (Salmon shark)- Brain	
31	PKFF01000000	SK_LD2	Scaffold	17	*Lamna ditropis* (Salmon shark)- Ear	
32	PKFE01000000	SK_LD3	Scaffold	19	*Lamna ditropis* (Salmon shark)- Brain	
33	JUUF01000000	757_CMAL	Contig	514		

3.1.2 Genome synteny

Draft genomes were aligned and reordered to refine genome assemblies according to the reference genome. Thus, genome synteny, that is orthologous gene set having same local organization in species (Iskandar et al., 2017), was determined using MAUVE v. 20150226, which is a multiple alignment tool. Draft genome sequences (Genbank format) were aligned with ProgressiveMauve according to the reference genome and the output was obtained in FASTA format (Darling et al., 2004).

3.1.3. Validation of genomes

To confirm the validity of the genomes, average nucleotide identity (ANI) was determined. Using *C. maltaromaticum* LMA28 as reference, ANI values were determined using OrthoANItool v 0.93.1 (OAT) (Lee et al., 2016).

3.1.4. Genome annotation

On the basis of ANI results, 32 valid strains were used for further analysis. The valid genomes were annotated using PATRIC (Pathosystems Resource Integration Center) available at https://www.patricbrc.org/ (Brettin et al., 2015) which is a free, online webserver tool that provide various analysis including genome annotation. It is a microbial bioinformatics tool that provides various services including comparative genome and proteome analysis. The size, GC%, CDS, tRNA and rRNA were determined. PATRIC provide annotation of genomes using RAST (Rapid annotation using subsystem technology). Output is obtained in various formats (Genbank, Fasta, embl etc).

3.2 Phylogenetic & Proteome analysis

For analysis of evolutionary relationship and variation in genomes, all 32 strains of *C. maltaromaticum* were analyzed and 16S rRNA having 1400-1700 nucleotides were retrieved and selected for phylogenetic evaluation (Ali et al., 2012).

The 16S rRNA phylogenetic tree was constructed using Neighbor joining method (Saitou & Nei, 1987). 16S rRNA sequences from all strains were first aligned and then phylogenetic tree was constructed, with 1000 bootstrap replications, using MEGAx (Kumar et al., 2018). *Enterococcus faecalis* V583, which is a closely related specie was used as an outgroup, for construction of phylogenetic tree.

For pair-wise comparison of genetic relatedness, AAI (Average amino acid identity) was determined and AAI matrix was generated which shows the % amino acid identity among the 32 strains. AAI analysis was performed under authorized projects using EDGAR (Efficient Database framework for comparative Genome Analyses using BLAST score Ratios) (Blom et al., 2016).

3.3 Pan-core genome analysis

3.3.1 Pan-core genome plot

To estimate conservation in genome and to better understand the phylogenetic relationship, pan-core genome analysis was performed. Pan genome represents the global gene repertoire of the species whereas, core genome represents the genes that are shared and conserved among all the strains. Dispensable, also known as accessory genome includes genes which are shared by two or more strains, but not all. And unique (Singleton) genes are present in only one strain.

For comparison of 32 strains of *C. maltaromaticum*, pan-core genome analysis was performed using EDGAR (Efficient Database framework for comparative Genome Analyses using BLAST score Ratios) (Blom et al., 2016). The plot of pan-core genome analysis was also constructed on the basis of total number of distinct gene families and shared genes, respectively, with the subsequent addition of a genomes.

3.3.2. Core Phylogeny

Another phylogenetic analysis was performed in which core genome-based tree was constructed including 32 species of *C. maltaromaticum*. This tree was also constructed using EDGAR (Blom et al., 2016).

3.4 Functional annotations

COG (Cluster of Orthologous Group) and KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis was also performed (Kanehisa et al., 2012). Graphical representation of accessory, core and unique genes distribution, on the basis of COG and KEGG function, was obtained using BPGA Software (Chaudhari et al., 2016).

3.5 Safety assessment

By determination and assessment of following factors, the pathogenic nature of *C. maltaromaticum* species was explored:

- Virulence factors
- Antibiotic resistance

3.5.1 Virulence Factors

Virulence factors refers to the properties of pathogenesis that are responsible and enhance its potential for causing diseases. To identify the virulence factors in *C. maltaromaticum* strains, an online database VFDB (Virulence Factor Database) was used (Available online at: http://www.mgc.ac.cn/VFs/). VFDB identify virulence factors and their functional features that allow bacteria to cause pathogenicity (Chen et al., 2016).

Sequences in Genbank format were uploaded in Regular BLAST. Blastn and DNA sequences from VFDB full dataset (Set B) was chosen. E-value was set at 0.00001 and BLOSUM62 matrix was chosen. To obtain precise and accurate results, only factors having score more than 90 were considered. Results were manually exported to tabular form, along with the virulence factor and related genes and a heatmap was generated on the basis of presence and absence of virulence factors, according to the bitscore.

3.5.2. Antibiotic resistance genes

To analyze the antibiotic resistancepresent in the genomes of *C. maltaromaticum*, antibiotic resistance genes were determined using RGI (Resistance Gene Identifier)- CARD (Comprehensive Antibiotic Resistance Database) (https://card.mcmaster.ca/analyze/rgi). The CARD provides curated reference sequences and SNPs organized via the Antibiotic Resistance Ontology ("ARO"). RGI 5.1.1 Web Portal predicts ORF using Prodigal, homolog detection by Diamond and strict significance by CARD bit score cut-offs (Alcock et al., 2020).

DNA sequences in FASTA format were uploaded and perfect, strict and loose hits criteria was selected under high quality/coverage. Results with >50% matching identity were recorded and the output was compiled in tabular format.

3.6 Genome plasticity

The ability of *C. maltaromaticum* to adapt in diverse environment can be illustrated by determining the genetic events. For the identification of following horizontally acquired genes, genome plasticity was analyzed:

- Genomic Islands
- Pathogenic Islands
- Prophage Sequences

3.6.1. Genomic islands (GEIs)

Genomic islands are the region of genes that are acquired and transferred through horizontal gene transfer. Virulence & antibiotic resistance genes are also present in genomic islands, particularly called pathogenicity islands and resistance islands, respectively.

The genomic islands in *C. maltaromaticum* strains were determined using IslandViewer4 (http://www.pathogenomics.sfu.ca/islandviewer/) which is a computational tool that employs four different genomic island prediction methods: IslandPick, IslandPath-DIMOB, SIGI-HMM, and Islander (Bertelli et al., 2017). The .gbff file of each *C. maltaromaticum* strain retrieved from NCBI, after manually removing comments, was used as an input, individually. The results provided the image, sequences and tab delimited file of the predicted genomic islands. The percentage of genomic islands in each strain was also calculated using following formula:

$$\% \text{ of genomic islands} = \frac{\text{Total no. of bp in genomic islands}}{\text{Total no. of bp in genome of strain}} \times 100$$

The % genomic islands were visualized in the form of a graph, whereas the other results were manually curated and further visualized using BRIG (Blast Ring Image Generator) in the form of circular image.

3.6.2. Pathogenicity islands (PAIs)

Pathogenicity islands comprise of virulence factors responsible for causing disease. They are present in pathogenic species, while generally absent from non-pathogenic species. PAIs exhibit certain characteristics such as G+C content and codon usage deviation, transposases and integrases, virulence factors, direct repeats, flanking tRNAs & frequently present mobile genetic elements.

To predict the putative pathogenicity islands, GIPSy v 1.1.2. (Genomic Island Prediction Software) was used. Genbank sequences obtained from IslandViewer4 (.gbk) were used as input and *C. maltaromaticum* LMA 28 was used as reference. GIPSY generates a list of putative pathogenicity islands on the basis of GC content, codon usage, transposase, virulence factors, tRNAs and provide associated files of the results (Soares et al., 2016).

For visualization of these PAIs, BRIG (Blast Ring Image Generator) was used for easy comparison among *C. maltaromaticum* strains.

3.6.3. Prophage sequences

To predict the integrated prophage sequences in *C. maltaromaticum* strains, online webserver tool PHASTER (**PHA**ge **S**earch **T**ool **E**nhanced **R**elease), which is an updated version of PHAST, was employed. It is available at https://phaster.ca/. It is an efficient and user-friendly webserver for fast detection & annotation of prophages in the genomes and plasmids of bacteria and provide accurate and precise results (Arndt et al., 2016).

The Genbank file, obtained from IslandViewer4 were used as an input in PHASTER. The results were exported in tabular format and only the intact prophage sequences were manually curated in the form of tab delimited file to further visualize in BRIG (Blast Ring Image Generator).

3.6.4 Circular comparison map of genomes

BRIG (Blast Ring Image Generator) was employed for visualization of genome plasticity including the predicted genomic islands, pathogenicity islands and prophage sequences. It is a free-cross platform application that display comparisons among large genomes in the form of circular genome map. The software is available at http://brig.sourceforge.net/. The similarity (% identity) of reference genome with other strains is performed using Blast (Alikhan et al., 2011).

The Genbank files obtained from IslandViewer4 were used to visualize the genomes and genomic islands, pathogenicity islands and prophage sequences were added as custom features in the form of tab-delimited file. Image size were adjusted according to the requirement. All other features were set as default. Results were generated in image format.

For all food associated strains (both dairy & non-dairy strains), only strong pathogenicity islands and intact prophage sequences were plotted by keeping *C. maltaromaticum* LMA28 as a reference and the others as query sequences.

For all the other pathogenic strains, only strong pathogenicity islands and intact prophages were visualized due to large dataset again by keeping *C. maltaromaticum* LMA28 as a reference.

RESULTS

4.1 Genomes, genome validation and properties

4.1.1 Sequence retrieval from NCBI

Sequences of 33 publicly available genomes (Table 4) were retrieved from NCBI in Genbank and FASTA format. At the time of this study, among these 33 strains, 2 genomes were complete, i.e. *C. maltaromaticum* LMA28 (isolated from ripened cheese) and *C. maltaromaticum* TMW 2.1581 (isolated from spoilage poultry meat). One strain isolated from diseased rainbow trout, *C. maltaromaticum* 18ISCm was at chromosome level. Whereas 13 genomes were assembled at scaffold and 17 at contig level.

Only two strains: *C. maltaromaticum* TMW 2.1581 and *C. maltaromaticum* 18ISCm exhibited 4 plasmids, each. On the basis of their source of isolation, they were grouped as follows:

Table 5. Grouping of *C. maltaromaticum* genomes available on NCBI, on the basis of source of isolation

S. No	Source of isolation	Number of genomes available on NCBI	*C. maltaromaticum* Strains
1	Dairy (Milk and cheese)	3	LMA 28, DSM 20342 MX5, DSM 20342
2	Processed-food products (Non-Dairy)	16	TMW 2.1581, A1, A3, A5, A6, A6, A14, A15, A16, A17, A18, DSM 20722, EBP 3019, SF668, ML_1_97, 3_18
3	Diseased Fish	13	18ISCm, ATCC 35586, NBRC 15685, DSM 20730, SK_AV1, SK_AV2, SK_AV3, SK_AV4, SK_AV5, SK_AV6, SK_LD1, SK_LD2, SK_LD3
4	Unknown	1	757_CMAL

4.1.2 Validation of genomes

To ensure the validation of genomes, in-silico ANI were determined. Average nucleotide identity (ANI) was calculated with *C. maltaromaticum* LMA28 as reference and only strains with ANI > 95% were considered valid. Only one strain *C. maltaromaticum* 757_CMAL had 67.71% ANI, whereas all the other strains had >95% ANI. This might be due to the reason that this genome is not completely assembled yet and contains 514 contigs in genome assembly, with size of 1.8 MBp, which is much less than other genomes.

Due to the divergence of *C. maltaromaticum* 757_CMAL from other strains, this strain was excluded from further analysis and all the other 32 strains were considered valid.

Table 6. ANI of *C. maltaromaticum* strains

Strains	**ANI**
C. maltaromaticum LMA28	
C. maltaromaticum TMW2.1581	99.56
C. maltaromaticum DSM 20342 MX5	99.65
C. maltaromaticum DSM 20342	99.66
C. maltaromaticum NBRC 15685	99.19
C. maltaromaticum DSM20722	99.17
C. maltaromaticum A6	99.15
C. maltaromaticum A5	99.14
C. maltaromaticum A17	99.14
C. maltaromaticum A15	99.13
C. maltaromaticum A16	99.13
C. maltaromaticum A18	99.13
C. maltaromaticum A14	99.11
C. maltaromaticum A7	99.07
C. maltaromaticum EBP3019	98.98
C. maltaromaticum SF668	98.92
C. maltaromaticum ATCC 35586	98.9
C. maltaromaticum ML_1_97	98.89
C. maltaromaticum DSM20730	98.86
C. maltaromaticum 3_18	98.83
C. maltaromaticum A1	98.51
C. maltaromaticum A3	98.49

Strains	ANI
C. maltaromaticum 18ISCm	98.12
C. maltaromaticum SK_AV1	96.69
C. maltaromaticum SK_AV2	96.69
C. maltaromaticum SK_AV3	96.7
C. maltaromaticum SK_AV4	96.7
C. maltaromaticum SK_AV5	96.71
C. maltaromaticum SK_AV6	96.71
C. maltaromaticum SK_LD1	96.69
C. maltaromaticum SK_LD2	96.72
C. maltaromaticum SK_LD3	96.71
C. maltaromaticum 757_CMAL	67.71

4.1.3 Genome Annotation

Average genome size of all *C. maltaromaticum* strains is about 3.5 Mb with an average GC content 34.4%. The largest genomes were acquired by *C. maltaromaticum* 18ISCm (4.06 MBp) followed by dairy isolated strains (*C. maltaromaticum* DSM 20342 MX5) probably due to the reason of better adaptability in environment, followed by food-associated strains. Whereas, it can be seen that the diseased fish isolated strains (*C. maltaromaticum* SK series of strains) possessed smaller genomes (3.30 Mbp). Moreover, there is no substantial difference in average genome size of pathogenic (3.41 Mbp) and non-pathogenic strains (3.56 MBp). This difference in genome size is generally considered normal among strains of same species.

Furthermore, the average CDS found in the genomes of *C. maltaromaticum* were 3417. Highest number of Coding Sequences (CDS), 4036, were found in *C. maltaromaticum* DSM 18ISCm, and lowest CDS, 3214, were found in *C. maltaromaticum* NBRC 15685.

It can be seen that very high number of pseudogenes were present in SK series of strains isolated from common thresher and salmon shark, which depicts that the strains have transitioned and adapted according to lifestyle of their host. However, *C. maltaromaticum* ML_1_97 also

possessed a relatively higher number of pseudogenes, which can possibly be due to the exceptionally high number of contigs in genome assembly. The genomic features are shown in Table 7.

Table 7. General genomic features of apathogenic and pathogenic *C. maltaromaticum* strains

Apathogenic food strains						
Origin	Strains	Size	Pseudogenes	CDS	tRNA	rRNA
Dairy	*C. maltaromaticum* LMA 28	3.65	120	3585	59	19
	C. maltaromaticum DSM 20342 MX5	3.88	108	3779	64	19
	C. maltaromaticum DSM 20342	3.75	77	3676	60	4
Non-Dairy/ Processed Food	*C. maltaromaticum* TMW 2.1581	3.70	90	3541	64	19
	C. maltaromaticum A7	3.50	65	3360	62	9
	C. maltaromaticum A15	3.53	50	3367	55	9
	C. maltaromaticum A14	3.53	50	3366	55	9
	C. maltaromaticum A16	3.53	50	3362	47	9
	C. maltaromaticum A5	3.53	50	3367	48	9
	C. maltaromaticum A18	3.53	52	3362	60	9
	C. maltaromaticum A6	3.53	50	3366	54	9
	C. maltaromaticum A3	3.50	67	3391	61	8
	C. maltaromaticum A1	3.50	68	3366	55	9
	C. maltaromaticum A17	3.53	50	3360	48	9
	C. maltaromaticum DSM 20722	3.58	58	3412	65	4
	C. maltaromaticum SF668	3.64	48	3514	62	9
	C. maltaromaticum EBP3019	3.48	63	3354	62	8
	C. maltaromaticum ML_1_97	3.31	68	3384	37	4
	C. maltaromaticum 3_18	3.56	60	3485	59	5

Fish pathogenic strains						
Origin	Strains	Size	Pseudogenes	CDS	tRNA	rRNA
Diseased Fish	*C. maltaromaticum* 18ISCm	4.07	34	4036	66	19
	C. maltaromaticum ATCC 35586	3.54	53	3454	61	9
	C. maltaromaticum NBRC 15685	3.43	49	3214	30	3
	C. maltaromaticum DSM 20730	3.54	58	3433	47	3
	C. maltaromaticum SK_AV1	3.30	243	3306	49	4
	C. maltaromaticum SK_AV2	3.30	244	3296	49	4
	C. maltaromaticum SK_AV3	3.30	248	3312	49	4
	C. maltaromaticum SK_AV4	3.30	247	3308	49	4
	C. maltaromaticum SK_AV5	3.38	252	3390	45	5
	C. maltaromaticum SK_AV6	3.30	242	3303	36	7
	C. maltaromaticum SK_LD1	3.29	246	3304	43	4
	C. maltaromaticum SK_LD2	3.29	247	3304	49	4
	C. maltaromaticum SK_LD3	3.30	243	3299	47	6

4.2 Phylogenetic and proteome analysis

Phylogenetic analysis and evolutionary relationship between *C. maltaromaticum* strains was studied and a tree was constructed on the basis of 16S rRNA (Fig. 4). *Enterococcus faecalis* V583 was used as an outgroup for the tree. The dairy-isolated strains are marked with blue dot, food-associated strains in green dot, and diseased fish-isolated strains in red.

All the 3 dairy isolated strains were closely related in the phylogenetic tree and belong to the same ancestor. The vacuum-packaged products formed a monophyletic clade in the tree, except *C. maltaromaticum* A1 and *C. maltaromaticum* A3, which formed a separate clade. The SK series of strains isolated from diseased fish also shared a monophyletic clade suggesting the same common ancestor. Other pathogenic strains, *C. maltaromaticum* ATCC 35586, *C. maltaromaticum* DSM 20730, *C. maltaromaticum* NBRC 15685 and *C. maltaromaticum* 18ISCm were rather more closely related to apathogenic strains. This could be due to the reason that genetic makeup of the strains has probably not gone complete adaptation and is still undergoing. Thus, on the basis of 16S rRNA, it is difficult to clearly differentiate between pathogenic and apathogenic strains.

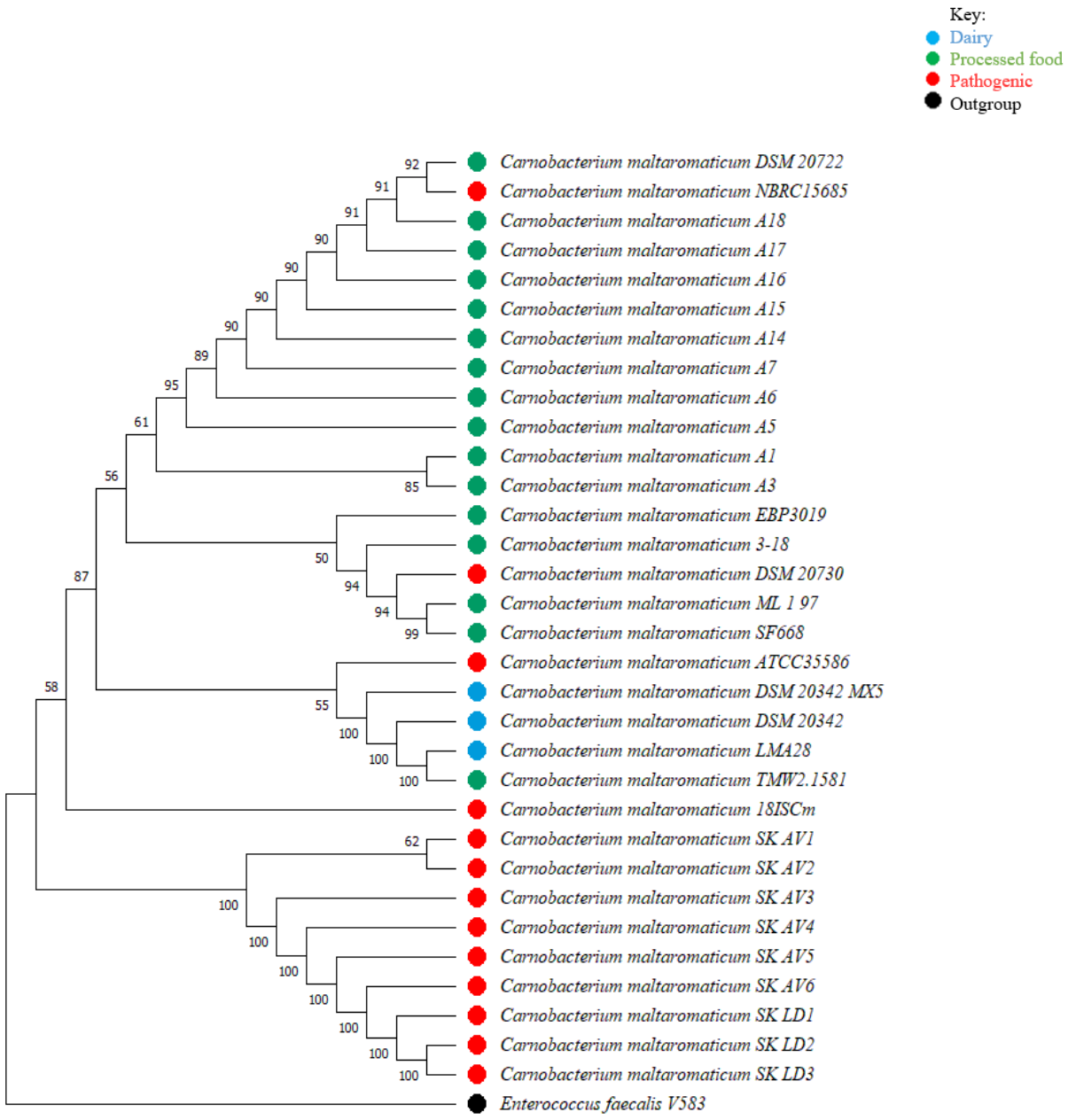


Figure 4. 16S rRNA phylogenetic analysis of *C. maltaromaticum* strains woth 1000 bootstrap replications

After phylogenetic analysis, proteome of the strains was analyzed by evaluating AAI (Amino Acid identity) and a color matrix was constructed. Darker color represents maximum % identity, followed by lighter colors.

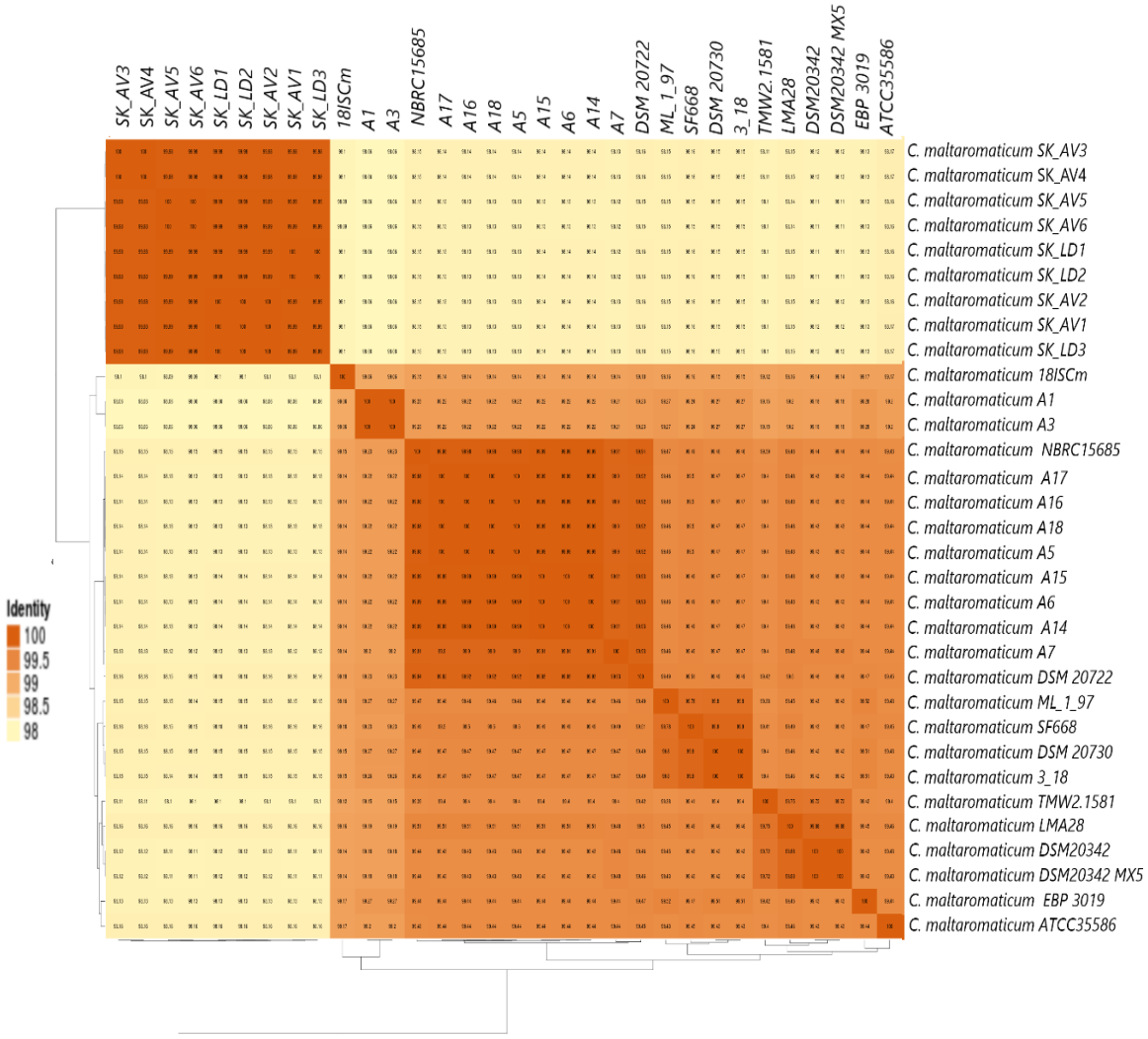


Figure 5. AAI matrix showing pair-wise genetic relatedness between 32 strains of *C. maltaromaticum*

The 9 SK series of strains, i.e. SK_AV1, SK_AV2, SK_AV3, SK_AV4, SK_AV5, SK_AV6, SK_LD1, SK_LD2, SK_LD3 isolated from diseased fish showed maximum similarity (>99%-100%) among themselves and least (<99%) with the other strains of *C. maltaromaticum*. However, all the other strains had an AAI of >99% among themselves and <99% with SK series of strains. *C. maltaromaticum* A1 and A3 were more closely related to each other rather other vacuum-packaged beef isolated strains. Also, the pathogenic strain *C. maltaromaticum* NBRC

15685 shared a high AAI with vacuum-packaged beef originated strains, as compared to other pathogenic strains. Another pathogenic strain, *C. maltaromaticum* DSM 20730 shared 100% AAI with *C. maltaromaticum* 3_18 and had >99% AAI with the other dairy and food-associated strains. *C. maltaromaticum* 18ISCm and *C. maltaromaticum* ATCC35586 also showed a similar behavior by having >99% identity with all the strains of *C. maltaromaticum,* except 9 SK series of strains.

4.3 Pan-Core Genome Analysis

4.3.1 Pan-Core Genome Plot

Pan core genome analysis results are shown in Table 8 and Fig 6. With the subsequent addition of genomes, the pan genome increased which suggests that it is an open-pan genome. While the core genome decreased with the addition of genomes until the end, where it stabilized. Initially, complete genomes were added so that no unusual increase or decrease could happen due to the draft genomes.

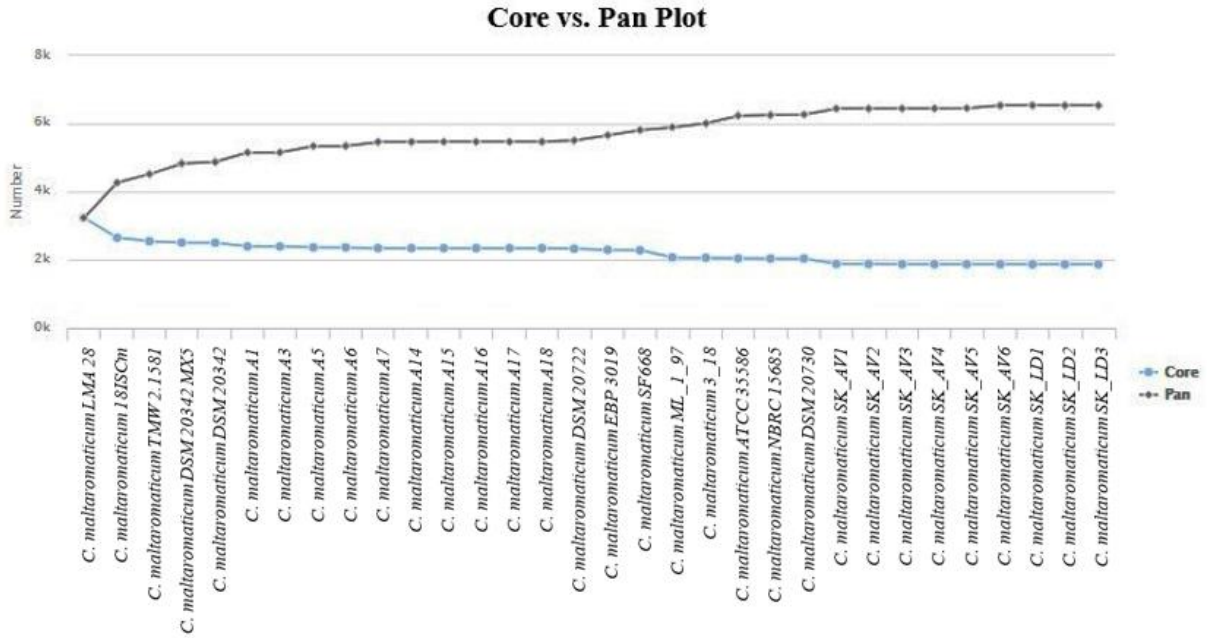


Figure 6. Core vs. pan plot between 32 strains of *C. maltaromaticum*

Table 8. Pan and core genome of *C. maltaromaticum* strains

Organism name	**Core genome**	**Pan genome**
C. maltaromaticum LMA28	3235	3235
C. maltaromaticum 18ISCm	2651	4270
C. maltaromaticum TMW 2.1581	2546	4521
C. maltaromaticum DSM 20342 MX5	2507	4832
C. maltaromaticum DSM 20342	2502	4874
C. maltaromaticum A1	2393	5155
C. maltaromaticum A3	2393	5158
C. maltaromaticum A5	2367	5344
C. maltaromaticum A6	2367	5344
C. maltaromaticum A7	2344	5467
C. maltaromaticum A14	2344	5467
C. maltaromaticum A15	2344	5468
C. maltaromaticum A16	2344	5469
C. maltaromaticum A17	2344	5469
C. maltaromaticum A18	2344	5471
C. maltaromaticum DSM 20722	2329	5508
C. maltaromaticum EBP 3019	2295	5663
C. maltaromaticum SF668	2084	5816
C. maltaromaticum ML_1_97	2074	5898
C. maltaromaticum 3-18	2063	6008
C. maltaromaticum ATCC 35586	2044	6235
C. maltaromaticum NBRC 15685	2040	6252
C. maltaromaticum DSM 20730	2040	6269
C. maltaromaticum SK_AV1	1879	6443
C. maltaromaticum SK_AV2	1878	6445
C. maltaromaticum SK_AV3	1873	6448
C. maltaromaticum SK_AV4	1873	6450
C. maltaromaticum SK_AV5	1872	6529
C. maltaromaticum SK_AV6	1872	6533
C. maltaromaticum SK_LD1	1869	6537
C. maltaromaticum SK_LD2	1868	6538
C. maltaromaticum SK_LD3	1867	6535

Across all 32 genomes, *C. maltaromaticum* pan genome constitutes 6535 genes and shared 1867 core genes. 2957 genes were found as dispensable genome. However, the number of unique genes was quite variable among the species. Maximum unique/ singleton genes were present in *C. maltaromaticum* ATCC 35586 (n=138) followed by *C. maltaromaticum* EBP 3019 (n=86) and *C. maltaromaticum* A7 (n=84).

4.3.2. Core Phylogeny

For further investigation on phylogenetic analysis, a core gene tree was constructed on the basis of core genome (Fig. 7).

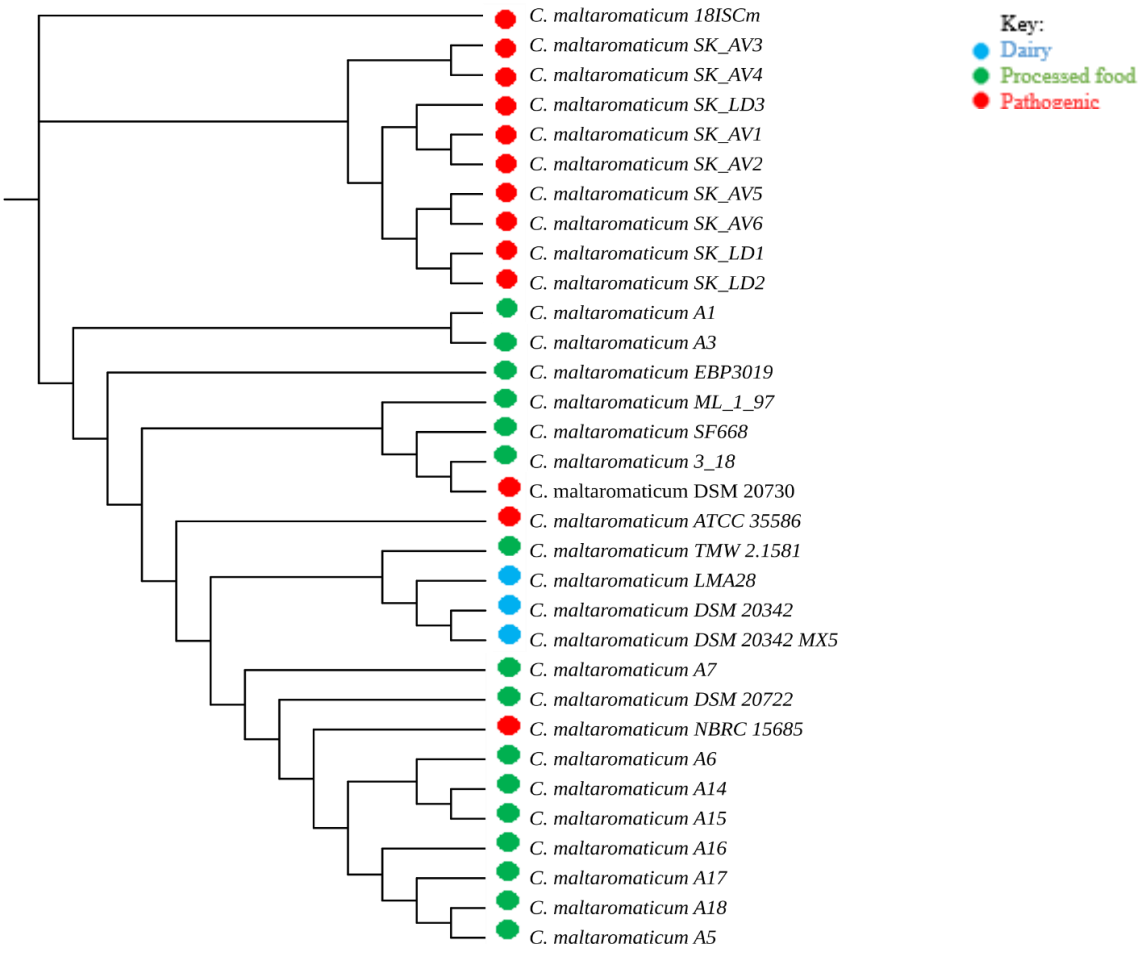


Figure 7. Core phylogenetic tree of C. maltaromaticum strains

C. maltaromaticum 18ISCm and 9 strains of *C. maltaromaticum* (SK series of strains) that have been isolated from common thresher & salmon shark formed separate clades in the tree apart from other strains which are placed closely together. Like 16S rRNA tree, the pathogenic strains, *C. maltaromaticum* ATCC 35586, NBRC 15685 and DSM 20730 were more closely related to dairy, processed food and food product isolated strains, respectively.

4.4 Functional annotation

Identification of functional categories of COG (Clusters of orthologous groups) in *C. maltaromaticum* strains (Fig. 8) suggested that the core genome was highly associated with metabolism and information storage and processing. Similar results were observed with accessory genes. Also, some accessory genes were also associated with cellular processing and signaling. Moreover, most of the information storage and processing were related to unique genes.

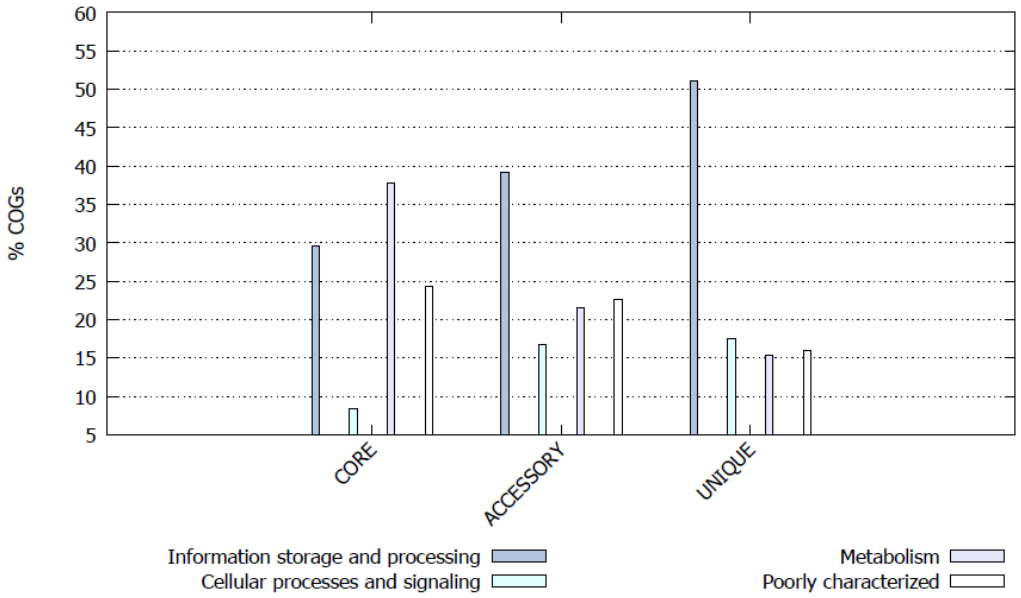


Figure 8. COG-based functional analysis *of C. maltaromaticum* strains

In the KEGG (Kyoto Encyclopedia of Genes and Genomes) functional annotation (Fig. 8), significantly higher number of core, accessory and unique genes were associated with metabolism. However, few were associated with environmental and genetic information processing.

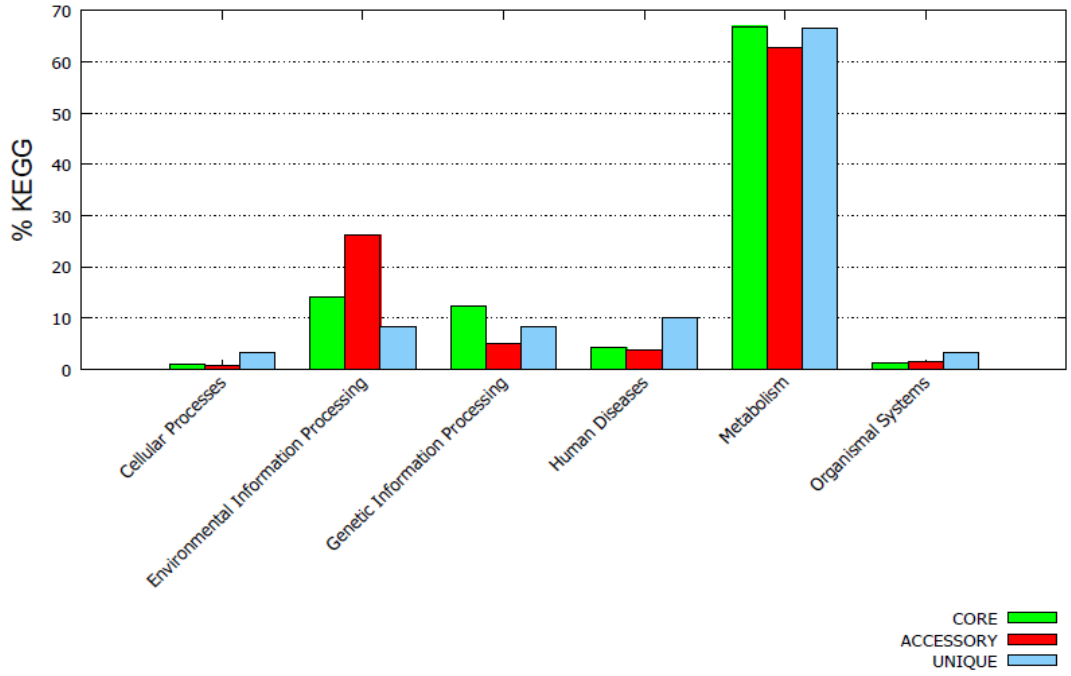


Figure 9. KEGG-based functional analysis of *C. maltaromaticum* strains

4.5 Safety assessment

4.5.1 Virulence factors

In this study, virulence factor database (VFDB) was employed to identify the potential virulence factors among the 32 strains of *C. maltaromaticum* (Fig. 10). The results were compared with 1 closely related probiotic *Lactobacillus sakei subsp. 23K*. The most common virulence factors determined in most of the strains were related to adherence and stress. Other virulence factors found were related to immune evasion, secretion, anti-phagocytosis, serum resistance, enzyme, bile resistance, regulation and toxin.

Highest number of virulence factors were found in *C. maltaromaticum* ATCC 35586, SK_AV1, SK_AV2, SK_AV3, SK_AV4 and SK_AV6 and 3 food-associated strains *C. maltaromaticum* DSM 20342, DSM 20722 and ML_1_97. Moreover, it was found that immune evasion related polyglutamic acid capsule genes (CapE, cap5E and cap8E) were absent in vacuum packaged food strains, along with antiphagocytosis related gene (capG) and serum resistance and immune evasion genes (wbtE and wbtF), except A1 and A3 where wbtF is present. Interestingly, toxin related virulence factor cnf1 was found unique to *C. maltaromaticum* SK_AV5 only and was absent from all other strains.

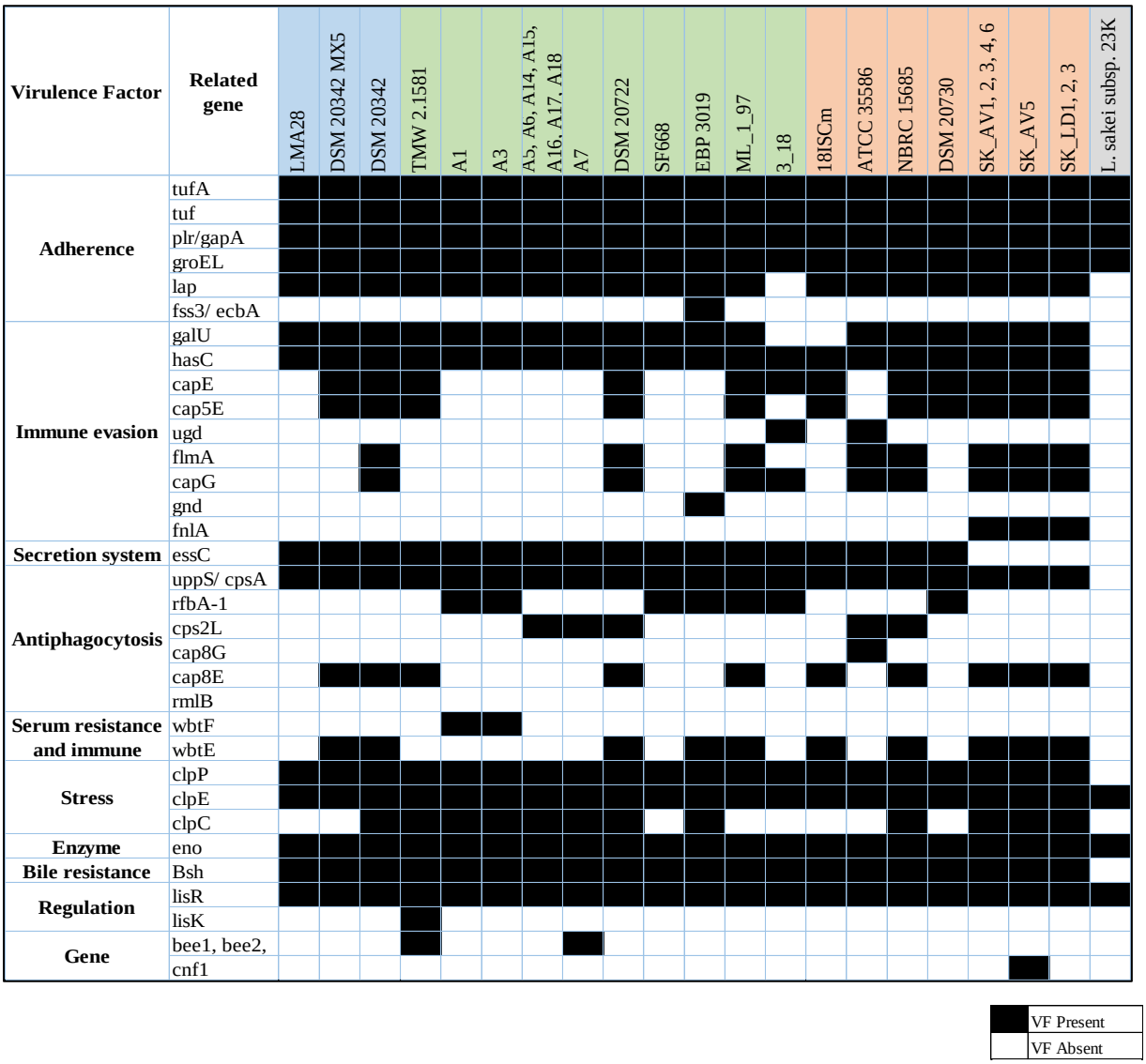


Figure 10. Distribution of virulence factors in *C. maltaromaticum* strains and closely related probiotic *Lactobacillus* species.

To compare the dairy strains, virulence factors in a closely related probiotic strain were also determined. It was observed that *Lactobacillus sakei subsp.* 23K also possess some virulence factors similar to *C. maltaromaticum* LMA 28. Some VFs were common in both closely related probiotic and *C. maltaromaticum* strains. However, our *C. maltaromaticum* strains have relatively more VFs in both pathogenic and apathogenic strains.

4.5.2 Antibiotic Resistance

In *C. maltaromaticum* A1 and A3, *tetM* gene for tetracycline resistance was found with 100% identity (Perfect hit). However, all the other resistance genes were detected under strict and loose hits. All the strains possessed almost similar antibiotic resistance pattern. However, some exceptions include streptogramin resistance gene *vatA* unique to *C. maltaromaticum* 18ISCm, penam resistance gene CBP-1 (CBP-β lactamase) was only found in *C. maltaromaticum* ATCC 35586 and rifamycin resistance gene arr-5 present only in *C. maltaromaticum* ML_1_97. Other resistance genes that were present in few but not all strains include tetracycline, pulvomycin, Fosfomycin and tunicamycin resistance genes.

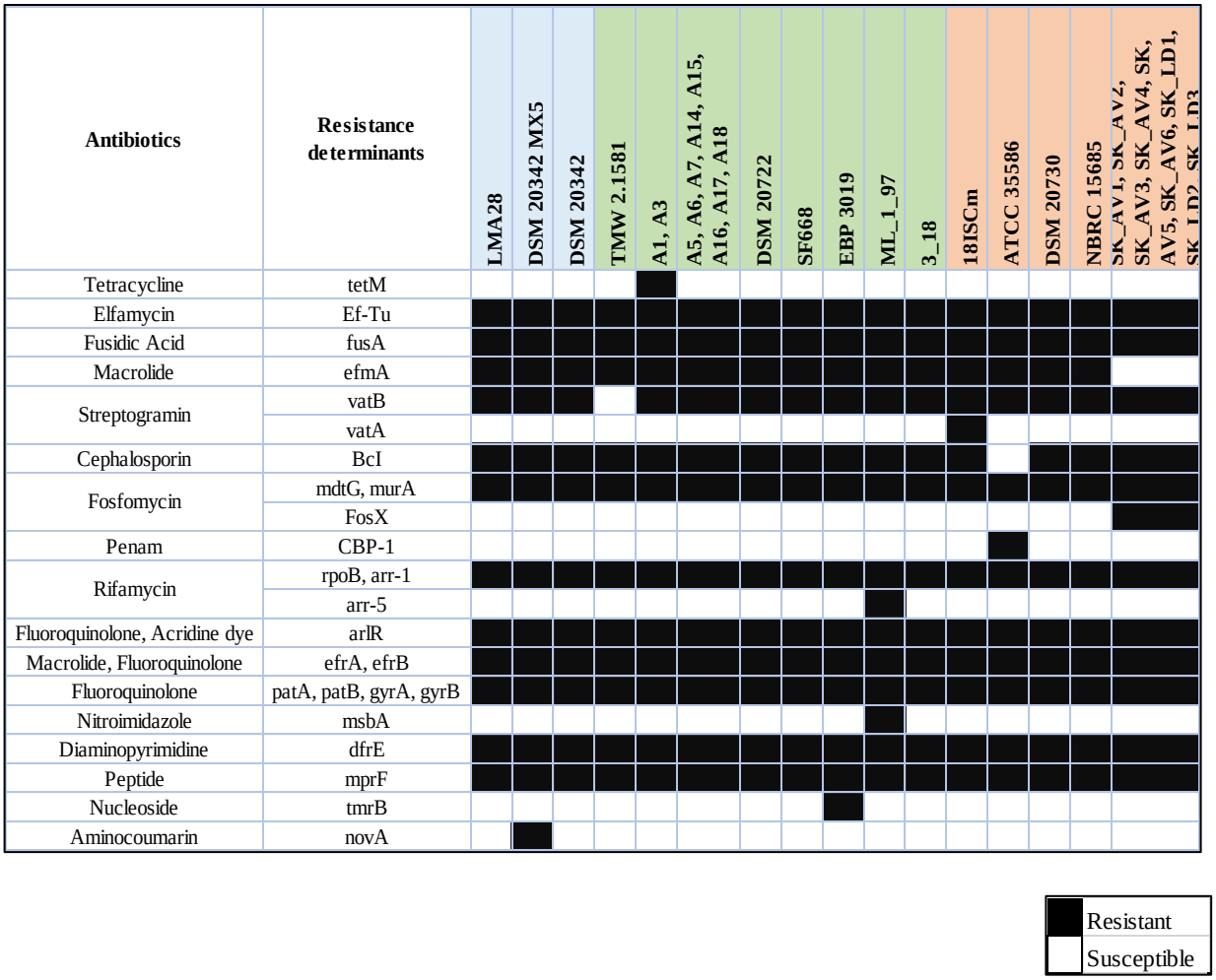


Figure 11. Resistance genes determination in *C. maltaromaticum* strains under perfect, strict and loose hits

4.6 Genome plasticity

Genome plasticity is referred as the intermittent presence of gain or loss of genetic information and is responsible for the niche-adaptation in a particular environment. For genome plasticity analysis, genomic and pathogenic islands, prophage sequences were predicted in *C. maltaromaticum* strains.

4.6.1 Genomic Islands

Maximum number of genomic islands (i.e. 20) were found in *C. maltaromaticum* DSM 20342 MX5 and *C. maltaromaticum* 18ISCm, followed by *C. maltaromaticum* LMA 28, TMW 2.1581, DSM 20342, EBP 3019 and 3_18. For prediction of highest acquisition of genomic islands, the percentage of acquired genomic islands was calculated. On the basis of the GEI %, a graph was created, which depicts that highest % of genomic islands were acquired by dairy isolated strains (*C. maltaromaticum* DSM 20342, DSM 20342 MX5 and LMA 28), food-derived strains (*C. maltaromaticum* A7, EBP3019, 3_18, SF668) and one pathogenic specie (*C. maltaromaticum* SK_AV5). The higher % of GEI in these strains suggests acquisition of large gene sets to adapt in their environment. Graphical representation of % GEI of *C. maltaromaticum* is shown in Fig. 12.

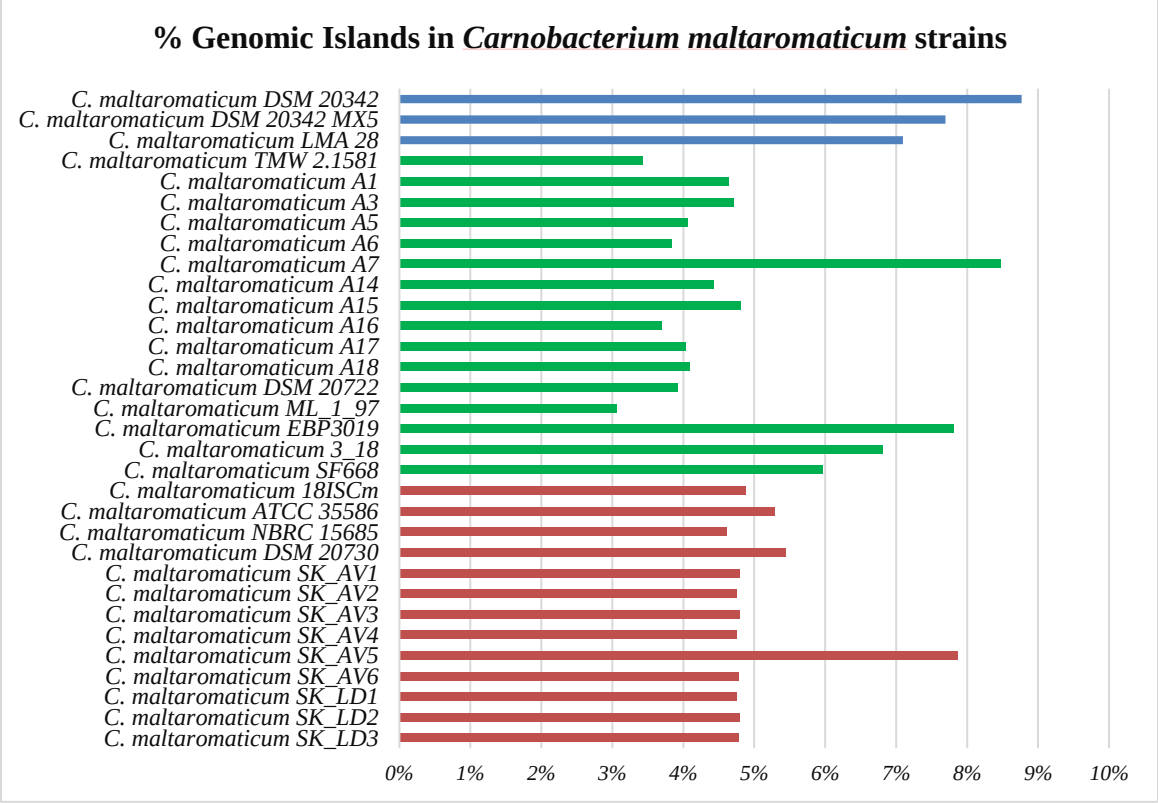


Figure 12. % GEI acquisition in *C. maltaromaticum* strains

Furthermore, the specific genes in these genomic islands were also determined. it was found that some food strains of *C. maltaromaticum* possess genes for metabolism of sugar, including lactose, mannitol, sucrose and galactose, as shown in Table 9.

Table 9. Genes in genomic islands of *C. maltaromaticum*

Genes in GEI	***C. maltaromaticum* Strains**
Cell Surface Protein	
LPXTG/ Sortase	LMA28, DSM 20342 MX5, DSM 20342 EBP3019, TMW 2.1581, 18ISCm
Sugar Metabolism	
Lactose	DSM 20342 MX5, DSM 20342, A5, A6, A7, A14, A15, A16, A17, A18, SF668, ATCC 35586
Mannitol, Lactose	EBP 3019
Mannose	NBRC 15685
Sucrose, Galactose	TMW 2.1581
Virulence family protein	
Isochorismatase	ATCC 35586, EBP 3019

4.6.2 Pathogenicity Islands

Pathogenicity islands exhibit certain features such as virulence factors, insertion sequences, integrases, transposases, phages, direct repeats and flanking tRNAs. GIPSy (Genomic Island Prediction Software) predicted putative pathogenicity islands in *C. maltaromaticum strains* with strong, normal and weak prediction.

Highest number of strong PAIs were predicted in 9 *C. maltaromaticum* SK series of strains isolated from diseased fish, ATCC 35586, followed by food-associated strains. Pathogenic strains *C. maltaromaticum* NBRC 15685 and DSM 20730 possess 1 and 0 strong pathogenicity island. Few food-associated strains were also predicted to have no strong pathogenicity island.

Table 10. Putative pathogenicity islands (PAI) in *C. maltaromaticum*

Strains	**PAI**	**Strong**	**Normal**	**Weak**
C. maltaromaticum LMA 28				
C. maltaromaticum SK_AV1	5	3	1	1
C. maltaromaticum SK_AV2	5	3	1	1
C. maltaromaticum SK_AV3	5	3	1	1
C. maltaromaticum SK_AV4	5	3	1	1
C. maltaromaticum SK_AV5	5	3	1	1
C. maltaromaticum SK_AV6	5	3	1	1
C. maltaromaticum SK_LD1	5	3	1	1
C. maltaromaticum SK_LD2	5	3	1	1
C. maltaromaticum SK_LD3	5	3	1	1
C. maltaromaticum ATCC 35586	3	3	0	0
C. maltaromaticum SF668	8	2	6	0
C. maltaromaticum A5	5	2	2	1
C. maltaromaticum A6	5	2	2	1
C. maltaromaticum A14	5	2	2	1
C. maltaromaticum A15	5	2	2	1
C. maltaromaticum A16	5	2	2	1
C. maltaromaticum A17	5	2	2	1
C. maltaromaticum A18	5	2	2	1
C. maltaromaticum DSM 20342 MX5	2	2	0	0
C. maltaromaticum ML_1_97	6	1	4	1
C. maltaromaticum EBP3019	5	1	3	1
C. maltaromaticum NBRC 15685	4	1	2	1
C. maltaromaticum A7	4	1	3	0
C. maltaromaticum A3	4	1	3	0
C. maltaromaticum A1	4	1	3	0
C. maltaromaticum 18ISCm	7	0	7	0
C. maltaromaticum DSM 20730	7	0	6	1
C. maltaromaticum DSM 20722	6	0	6	0
C. maltaromaticum TMW 2.1581	4	0	4	0
C. maltaromaticum DSM 20342	4	0	4	0
C. maltaromaticum 3_18	4	0	4	0

4.6.3 Prophage sequences

Prophage sequences were determined using PHASTER. Intact (score>90), questionable (score between 70-90) and incomplete (score<70) prophage sequences were predicted in *C. maltaromaticum* strains. The maximum number of prophages (n=3) were found in *C. maltaromaicum* 18ISCm. Intact prophage sequences were found in all dairy and pathogenic, except *C. maltaromaticum* NBRC 15685. Whereas, in food-associated strains, intact sequences were found in *C. maltaromaticum* A1, A3, A7, ML_1_97, SF668, EBP 3019, 3_18.

Table 11. Prophage sequences in *C. maltaromaticum* strains

Strains	***No. of prophages**	**Intact prophage proteins**
LMA 28	2 (0,3)	AttL, endolysin, Ig-like virion protein, putative RNA polymerase sigma 24, XRE like HTH transcriptional regulator, Cro/CI family XRE family transcriptional regulator, attR, putative integrase, putative peptidase, Mg2+/CO2+ transport protein CorA
		AttL, autolytic lysozyme, N-acetylmuramoyl-L- alanine amidase, holin, endolysin, terminase, large subunit, ArpU family transcriptional regulator, single-stranded DNA binding protein, YopX protein, DNA-methylase, D replication protein DC, zinc-ribbon domain containing protein, bifunctional S24 family peptidase/transcriptional regulator, attR, site-specific integrase
DSM 20342	1 (1,4)	attL, lysin, holin, putative short tail fiber, putative tail endopeptidase, tail-related proteins, head-tail joining protein, capsid, portal protein, cI-like repressor, DNA replication initiation control protein YabA, stage O sporulation protein YaaT, putative DNA polymerase III, , protein from Nitrogen regulatory protein P-II (GLNB) family, putative thymidylate kinase, attR,
DSM 20342 MX5	2 (0,4)	Autolytic lysozyme, attL, N-acetylmuramoyl-L- alanine amidase, endolysin, terminase large subunit, putative transposase A, transposase insF, sugar phosphate nucleotidyltransferase, anti-represser kilAC domain protein, cI-like repressor, site-specific integrase, attR
		Lysin, holin, membrane protein, putative short tail fiber, putative tail endopeptidase, tail-related proteins, head-tail joining protein, capsid & portal protein, putative transposase A, putative anti-repressor, transposase insF, cI-like repressor, attR, DNA replication initiation control protein, YabA, stage O sporulation protein YaaT, putative DNA polymerase III, protein from Nitrogen regulatory protein P-II (GLNB) family, putative thymidylate kinase

Strains	*No. of prophages	Intact prophage proteins
TMW 2.1581	0 (1,2)	-
A1	1 (0,2)	Tail-related proteins; prohead protease; putative prohead protease; Portal protein, RegA, Terminase large subunit; Terminase small subunit; endonuclease; conserved phage protein, DNA polymerase B region protein; sliding clamp DNA polymerase, membrane-associated initiation of head vertex; helicase; VRR-NUC domain-containing protein; Rep protein;, DUF955 domain containing protein; integrase, tRNA, attR, clpP
A3	1 (0,2)	PlyB054, holin, minor structural protein 4, endopeptidase, putative tail component, tail tape measure protein, major tail protein; prohead protease; putative prohead protease; Portal protein, RegA, Terminase large subunit; Terminase small subunit; endonuclease; conserved phage protein; gp34, DNA polymerase B region protein; sliding clamp DNA polymerase, membrane-associated initiation of head vertex; helicase; VRR-NUC domain-containing protein; Rep protein;, DUF955 domain containing protein; integrase, tRNA, attR, clpP
A5 A6	0 (0,1)	-
A7	1 (2,2)	site-specific recombinase; DNA primase subunit, holin, tail-related protein; phage head-tail adapter protein;, phage gp6-like head-tail connector protein; major capsid protein; membrane-associated initiation of head vertex; putative portal protein, transferase; S-adenosylmethionine synthetase; terminase; putative class I holin; helicase; putative primase; ribonucleotide reductase A subunit; polymerase;
A14 A15 A16 A17 A18	0 (0,1)	-
DSM 20722	0 (1,3)	-
SF668	1(0,2)	putative DNA polymerase III; carbonic anhydrase family protein; attL, lysin, holin, glycerophosphoryl diester phosphodiesterase;, putative phage tail protein; clamp-loader subunit; putative major tail shaft protein; major capsid protein b; scaffold; portal; terminase large subunit; Phage terminase small subunit; ArpU family transcriptional regulator; DUF1642 domain containing protein; baseplate wedge subunit; DNA cytosine methyltransferase; putative DNA replication & excisionase protein, attR, integrase, pyrroline-5-carboxylate reductase; deoxycytidylate deaminase
EBP 3019	1 (1,3)	N-acetylmuramoyl-L-alanine amidase; tail-related proteins; structure, head & capsid proteins, head maturation protease; portal protein; putative phage terminase large subunit

Strains	*No. of prophages	Intact prophage proteins
ML_1_97	1 (0,0)	DNA topoisomerase II large subunit; N-acetylmuramoyl-L- alanine amidase, tail protein, glycerophosphoryl diester phosphodiesterase, putative phagerelated tail tape measure protein, putative structure & head-tail joining protein, HK97 family phage major capsid, Clp protease-like protein, portal protein, putative phage terminase large subunit, , ArpU family transcriptional regulator, DNA repair protein recN, D replication protein DC, phage protein,
3_18	1 (1,4)	N-acetylmuramoyl-L- alanine amidase , holin, glycerophosphoryl diester phosphodiesterase, tail-related proteins, structure protein, head-tail joining protein, HK97 family phage major capsid protein, ClpP, portal protein, putative phage terminase-large subunit, RNA polymerase sigma-70, phage portal protein HK97 family , portal protein, putative single stranded DNA binding protein, RNA polymerase binding protein, , D replication protein DC, putative DNA replication protein
18ISCm	3 (1,3)	attL, integrase, late promoter transcription, anti-repressor protein, putative DnaB-like helicase, Phage related protein, terminase, phage portal protein, phage prohead, phage major capsid protein, aminopeptidase, tail-related proteins, putative minor structural protein 2
		attl, dihydrofolate reductase, N-acetylmuramoyl-L-alanine amidase, putative host interaction protein, tail component, Ig-like virion protein, major capsid protein gpP, portal, TerL, terminase small subunit, DUF3850 domain-containing protein, gp51, putative HNH endonuclease, putative SSB protein, zinc finger protein, Phage replication initiation, putative excisionase protein, Rep protein, putative regulatory protein, integrase, attR, Pyrroline-5-carboxylate reductase, deoxycytidylate deaminase
		attL, XRE family transcriptional regulator, repressor, putative rep protein, integrase, attL, DUF3800 domain-containing protein, N-acetylmuramoyl-L-alanine amidase, putative minor structural protein 2, tail-related proteins, aminopeptidase, phage major capsid protein, phage prohead, phage portal protein, terminase, Phage related protein, ABC-type transport system, ATP-binding protein, putative DnaB-like helicase, recombination endonuclease subunit, anti-repressor protein, CI phage repressor protein, putative cI-like repressor, metallo-prtoeinase motif, integrase, attR, SSU ribosomal protein S9p, LSU ribosomal protein L13p, baseplate hub subunit
ATCC 35586	1 (2,6)	attL, site-specific integrase, anti-repressor kilAC domain protein, putative phage replisome organizer, zinc finger protein, phage terminase small subunit, portal, minor capsid protein, tail protein, holin, Nacetylmuramoyl-L- alanine amidase, attR

Strains	*No. of prophages	Intact prophage proteins
DSM 20730	1 (1,3)	AttL,virion associated hydrolase; tail protein, capsid proteins, portal, TerL, t RNA, positive control sigma-like factor, putative single stranded DNA binding protein, RNA polymerase binding protein & replication, phage replication initiation, putative phage replisome organizer, repressor, site-specific integrase, attR
NBRC 15685	0 (0,2)	-
SK_AV1, SK_AV2, SK_AV3, SK_AV4, SK_AV5, SK_AV6, SK_LD1, SK_LD2, SK_LD3	1(0,3)	attL, replication and recombination DNA helicase; putative structural protein; tail fiber; capsid and scaffold protein; clamp-loader subunit; putative major tail shaft protein; DNA polymerase; major capsid protein; scaffold; portal; terminase large subunit; terminase small subunit; tRNA; ArpU family transcriptional regulator; single-stranded DNA binding protein; DNA replication protein; anti-repressor KilAC domain protein; repressor; site-specific integrase; attR

* No. of prophages: Intact (Questionable, Incomplete)

4.6.4 Circular genome visualization

The genomic islands, pathogenicity islands and prophage sequences were visualized using BRIG (Blast Ring Image Generator).

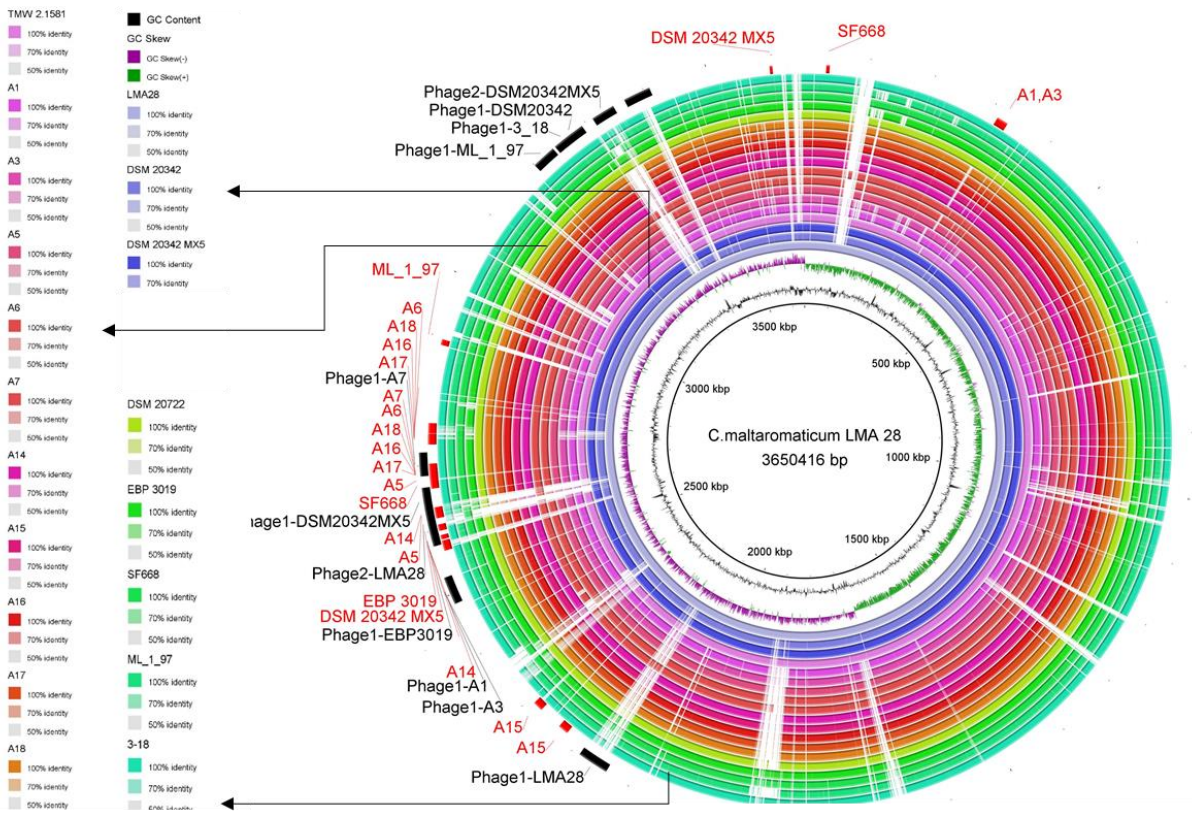


Figure 13. Comparison of food associated strains with *C. maltaromaticum* LMA 28 as reference genome. Each ring represents a genome as depicted in the legends. Pathogenicity islands and prophages are shown outside the rings in red and black, respectively.

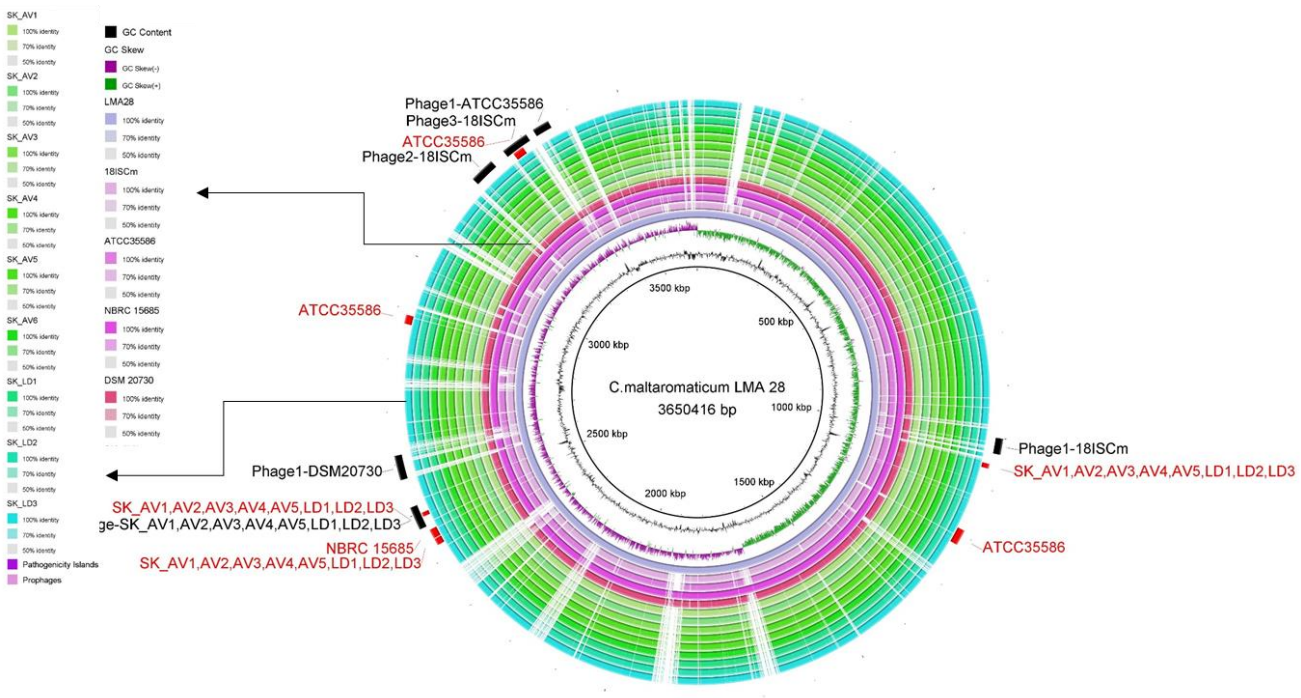


Figure 14. Comparison of pathogenic strains with *C. maltaromaticum* LMA 28 as reference genome. Each ring represents a genome as depicted in the legends. Pathogenicity islands and prophages are shown outside the rings in red and black, respectively.

DISCUSSION

Carnobacterium is a rod-shaped, gram positive, facultatively anaerobic, ubiquitous LAB, which ferment glucose and results in the production of L (+)-lactic acid. It had been isolated from multiple environments such as cold and temperate environment, as well as the gastrointestinal tract of animals, from fish and meat associated food products and dairy products.

C. maltaromaticum exhibit diverse environment (Jørgen J Leisner et al., 2007). Technological potential of *C. maltaromaticum* has been reported in milk-associated (dairy) products (Laursen et al., 2006; Millière et al., 1994; Morea et al., 1999). They have the potential to be used as adjunct culture (Afzal et al., 2010). Also, in meat and fish-associated food products, this specie is dominantly found among the bacterial communities (Chaillou et al., 2015; Duan et al., 2016; Fougy et al., 2016; Jääskeläinen et al., 2016). *C. maltaromaticum* produce anti-bacterial compounds that inhibit *Listeria monocytogenes.* Besides, *C. maltaromaticum* is known to cause diseases and infections in fish and its pathogenic strains have been described in some studies (Loch et al., 2008, 2011; Roh et al., 2020).

The purpose of this study was to compare and get insights of the fundamental genomic differences between fish pathogenic strains and non-pathogenic food strains of *Carnobacterium maltaromaticum* at genome level. In this study, sequences of publicly available genomes of *C. maltaromaticum* were obtained from NCBI which were then validated, followed by annotation, phylogenetic analysis, pan-core genome analysis, evaluation of safety on the basis of virulence factors and antibiotic resistance, and determination of genome plasticity by prediction of genomic islands, pathogenic islands and prophage sequences.

Among the 33 publicly available *C. maltaromaticum* strains, 3 strains belonged to dairy products, 16 from vacuum-packaged meat (processed food) and fish products, 13 from diseased fish and 1

strain had ambiguous source. After alignment of the genomes, the 33 genomes were validated on the basis of ANI (Awan et al., 2018) which excluded 1 strain, that is *C. maltaromaticum* 757_CMAL from further analysis. This divergence might be due to the reason that this genoms is not completely assembled yet and contains high number of contigs in it's genome assembly, with a size of only 1.8 MBp. The rest of the genomes were annotated.

Larger genomes were acquired by dairy isolated strains and *C. maltaromaticum* 18ISCm, followed by food-associated strains and relatively smaller genomes were exhibited by pathogenic strains, with few exceptions. However, there was no substantial difference between the average genome size of apathogenic and pathogenic strains. Also, a high number of pseudogenes were present in pathogenic strains. Studies have shown that high degree of pseudogenization suggests genome decay and indicate reduction in genome often observed in pathogens, which means that the strains might have lost and shed genes and adapted due to the transition according to the host lifestyle (i.e fish) (Moran, 2002; Steele et al., 2019). Moreover, the strains exhibit low GC content (34.4%). Studies have reported that recently acquired genes tend to be more AT-rich than recipient's average which depicts horizontally transferred DNA have lower GC content than host chromosome in bacteria. Ultimately this can be associated with evolution and environmental adaptation of *C. maltaromaticum* strains (Ravenall, et al., 2015; Nishida, 2012).

Phylogenetic analysis on the basis of 16S rRNA revealed that dairy strains belong to same common ancestor. Some other pathogenic and apathogenic strains form a clade which differentiate these strains. Previous studies have also reported shared evolution of the SK series of strains and their divergence from other known isolates (Steele et al., 2019). However, some pathogenic strains such as NBRC 15685, DSM 20730 and ATCC 35586 are defying this rule, and they are rather more closely related to apathogenic strains. This could be due to the reason

that genetic makeup of the strains has probably not gone complete adaptation and is still going on. So, on the basis of 16S rRNA, we can't differentiate between pathogenic and apathogenic strains. Further, pairwise-genetic relatedness on the basis of AAI among *C. maltaromaticum* strains were evaluated. All the strains (except SK series of strains) had an AAI of >99% among themselves and <99% with SK series of strains. Interestingly, the 9 SK series of strains i.e. isolated from diseased fish showed maximum similarity (>99%-100%) among themselves and least (<99%) with the other strains of *C. maltaromaticum*. These genomic variations in these strains depicts a certain level of host-adaptation in fish pathogenic strains, and they are different at genomic level.

Pan & core genome studies are carried out to determine the conservation in genetic structures present among different individual bacteria strains of a species. Previous studies have reported that this variation in the proteins might be correlated with the geographical isolation and lifestyles (Caputo et al., 2015; Francis et al., 2013). Pan-core genome analysis revealed that *C. maltaromaticum* shared core genome of 55% of the average CDS among the strains. These set of genes conserved between the strains contain essential genes that are required for the independent survival of a specie. Pan-core plot suggests that the strains exhibited an open pan and the strains are still harboring genetic exchange. However, variable numb er of unique genes were found. Highest number of unique genes were found in *C. maltaromaticum* ATCC 35586. Leisner et al., has reported that *C. maltaromaticum* ATCC 35586 possess certain virulent genes that might be responsible for pathogenicity in fish. Thus, these unique genes in *C. maltaromaticum* ATCC 35586 might be the virulence genes. Furthermore, a core-based phylogenetic tree was also constructed. The tree revealed that the dairy strains were closely related, like in 16S rRNA tree. Pathogenic strains *C. maltaromaticum* ATCC 35586, DSM 20730 and NBRC 15685 were more

closely related to dairy, *C. maltaromaticum* 3_18 and vacuum packaged meat isolated strains, respectively. Thus, phylogenetic analysis, on the basis of 16S rRNA and core phylogeny does not clearly differentiate between pathogenic and apathogenic strains of *C. maltaromaticum*.

The functional annotation using COG and KEGG is quite significant as it categorizes the functions into families and super families. The COG results suggested that the core genome was highly associated with metabolism and information storage and processing. Similar results were observed with accessory genes. Also, some accessory genes were also associated with cellular processing and signaling. Moreover, most of the information storage and processing were related to unique genes. In the KEGG functional annotation, significantly higher number of core, accessory and unique genes were associated with metabolism. This is due to the reason that KEGG includes more metabolism-related genes. However, few genes were associated with environmental and genetic information processing.

C. maltaromaticum LMA 28, isolated from soft cheese, possess certain probiotic-related features. Rahman et al., studied *C. maltaromaticum* LMA 28 ability to survive in the gut of mammals and found that it exhibits genes for adaptation in the mammalian gut (Rahman, Gleinser, et al., 2014). Conversely, *C. maltaromaticum* strains, such as 18ISCm and ATCC 35586 have been reported as fish pathogen (Roh et al., 2020). Therefore, there is a need for determination of the safety of *C. maltaromaticum* strains if they are to be used for their technological potential in further applications and evaluate whether is safety of *C. maltaromaticum* strains is ecosystem-specific or strain-specifc.

Virulence factors were determined in *C. maltaromaticum* strains and compared with a closely related probiotic specie. All the strains, that is dairy, non-dairy and diseased fish isolates, possessed some virulence factors. Virulence factors related to adherence, stress, regulation and

enzymes were found in nearly all strains as well as in *Lactobacillus sakei subsp.* 23K. Previous studies have reported that some virulence factors sometimes might be present in non-pathogenic commensal bacteria which are either suppressed by other genes or tended to be involved in general host interactions that contributes to adhesion and protection. Remarkably, many probiotic-related traits could be associated to virulence factors. Thus, VFs don't clearly differentiate pathogenic strains from non-pathogenic strains. Consequently, each virulence factor has its specific role in the particular bacterial species that encode them. detection of virulence genes may not always be sufficient to determine pathogenicity.

Antibiotic resistance was also investigated in *C. maltaromaticum* strains. Only *tetM* gene for tetracycline resistance was found with 100% identity in *C. maltaromaticum* A1 and A3. Previous research has shown *C. maltaromaticum* to be phenotypically resistant to ampicillin, vancomycin, kanamycin, gentamycin and nalidixic acid and *C. maltaromaticum* ATCC 35586, a fish pathogen, has been reported to be phenotypically resistant to aquaculture antibiotics such as chlortetracycline, nitrofurans and sulfonamides (J J Leisner et al., 2012). Thus, genotypic antibiotic resistance must be verified by phenotypic methods such as disc diffusion method using CLSI guideline. Moreover, further in vitro and in vivo studies are needed to confirm if these strains are safe to use.

Genome plasticity was predicted in order to understand the gain or loss of genes, through horizontal gene transfer. This plays a significant role in the adaptation of bacteria in particular environment (Dobrindt et al., 2010). In this study, genomic islands, pathogenicity islands and prophage sequences were predicted in *C. maltaromaticum* strains. Dairy isolated strains possessed a high % of genomic island (GEI). The other strains acquiring relatively higher % of GEI include food associated strains *C. maltaromaticum* A7, EBP 3019, 3_18, SF668 and a

pathogenic strain *C. maltaromaticum* SK_AV5. Studies have shown that large size of *C. maltaromaticum* may be the reason of its well-adaptation to environmental challenges. Hence, large % of acquired GEIs in some food and pathogenic strains depicts that all the strains are adapting towards their specific niche or ecosystem.

Pathogenicity islands are subsets of genomic islands, acquired through horizontal gene transfer and comprise of virulence factors that are generally absent from non-pathogenic strains of the same or closely related species (Gal-Mor & Finlay, 2006). Highest number of strong PAIs were predicted in nine *C. maltaromaticum* SK series strains isolated from diseased fish, ATCC 35586, followed by food-associated strains. Few food-associated strains were also predicted to have no strong pathogenicity island. As revealed in this study, few strong PAIs were found in both pathogenic and apathogenic strains thus there is a certain degree of probability that some apathogenic strains may not be safe due to the presence of PAI.

Prophages are of high importance as they are involved in horizontal gene transfer and contributes to the diversity in bacteria (Li et al., 2017). In this study, the maximum number of prophages (n=3) were found in *C. maltaromaticum* 18ISCm. Intact prophage sequences were found in all dairy and pathogenic strains, except *C. maltaromaticum* NBRC 15685. Whereas, in food-associated strains, intact sequences were only found in *C. maltaromaticum* A1, A3, A7, ML_1_97, SF668, EBP 3019, 3_18.

Hence, variation between pathogenic and apathogenic strains of *C. maltaromaticum* that are found in diverse environments, exist in pseudogenes with pathogenic strains exhibiting higher number of pseudogenes suggesting their transition towards host-adapted lifestyle. Also, acquisition of genomic islands in both pathogenic and non-pathogenic *C. maltaromaticum* strains suggests that these strains are on the coarse of adaptation. Furthermore, apathogenic strains

cannot be termed as completely safe due to the presence of VFs, in comparison to reference probiotic and PAIs. And the safety assessment revealed that safety might be a strain-specific feature and due to high diversity, virulence factors are widespread in apathogenic strains, as well.

CONCLUSION & FUTURE PROSPECTS

The genome size and acquisition of genomic islands in *C. maltaromaticum* strains depicts better adaptability of these strains in their particular environment. Moreover, . Low GC content in *C. maltaromaticum* might be associated to its evolution and adaptation in environment. High number of pseudogenes in pathogenic strains depicts a transition of *C. maltaromaticum* SK series of strains to host (fish) -adapted lifestyle. Phylogenetic and AAI analysis reveals that some pathogenic strains are more closely related to food-associated strains rather pathogenic strains and that SK series of strains are distantly-related compared to other pathogenic and apathogenic strains. Therefore, no clear differentiation can be seen between fish pathogenic and apathogenic food strains at genomic level. Also, *C. maltaromaticum* strains exhibit open pan genome. Regarding safety assessment, adherence and stress related virulence factors, and enzymes including *bsh* were found to be present in almost all strains of *C. maltaromaticum* as well as closely related probiotic specie, *Lacobacillus sakei subsp.* 23K as well. Thus, the safety of *C. maltaromaticum* strains is hard to determine on the basis of VFs as there is no clear variation in virulence factors of pathogenic and apathogenic strains. Moreover, the genotypic antibiotic resistance in C. maltaromaticum strains needs to be phenotypically verified. Furthermore, strong pathogenicity islands were also found in all *C. maltaromaticum* strains, except 6 strains (2 pathogenic, 1 dairy and 3 food-associated), along with intact prophage sequences. Thus, apathogenic strains cannot be termed as completely safe due to the presence of VFs, in comparison to reference probiotic and PAIs. Thus, safety of *C. maltaromaticum* might be a strain-specific feature and due to high diversity, virulence factors are widespread in apathogenic strains, as well.

The Dairy isolated strains as well as some food associated strains may have the potential to be used a probiotic as they exhibit certain genes such as sugar (lactose) metabolism, bile tolerance genes. However, in vitro studies are needed to confirm the presence and functionality of these probiotic-related genes and validate if these strains have technological potential, so that they can be employed for further industrial applications. Furthermore, in vivo trials are required to evaluate and confirm the safety of dairy and food associated strains, before using at larger scale. Moreover, a large number of *C. maltaromaticum* strains and its data from respective ecosystem is required to better understand the genome dynamics, safety and technological potential of the strains.

REFERENCES

Afzal, M. I., Jacquet, T., Delaunay, S., Borges, F., Millière, J. B., Revol-Junelles, A. M., & Cailliez-Grimal, C. (2010). Carnobacterium maltaromaticum: identification, isolation tools, ecology and technological aspects in dairy products. *Food microbiology*, *27*(5), 573-579.

Alcock, B. P., Raphenya, A. R., Lau, T. T., Tsang, K. K., Bouchard, M., Edalatmand, A., ... & Min, S. Y. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic acids research, 48(D1), D517-D525.

Ali, A., Soares, S. C., Santos, A. R., Guimarães, L. C., Barbosa, E., Almeida, S. S., ... & Hassan, S. S. (2012). Campylobacter fetus subspecies: comparative genomics and prediction of potential virulence targets. Gene, 508(2), 145-156.

Alikhan, N. F., Petty, N. K., Zakour, N. L. B., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC genomics, 12(1), 402.

Alvarez-Sieiro, P., Montalbán-López, M., Mu, D., & Kuipers, O. P. (2016). Bacteriocins of lactic acid bacteria: extending the family. Applied microbiology and biotechnology, 100(7), 2939-2951.

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. Nucleic acids research, 44(W1), W16-W21.

Bachmann, H. P., Fröhlich-Wyder, M. T., Jakob, E., Roth, E., Wechsler, D., Beuvier, E., & Buchin, S. (2011). Cheese| Raw Milk Cheeses.

Barakat, R. K., Griffiths, M. W., & Harris, L. J. (2000). Isolation and characterization of Carnobacterium, Lactococcus, and Enterococcus spp. from cooked, modified atmosphere packaged, refrigerated, poultry meat. International journal of food microbiology, 62(1-2), 83-94.

Baya, A. M., Toranzo, A. E., Lupiani, B., Li, T., Roberson, B. S., & Hetrick, F. M. (1991). Biochemical and serological characterization of Carnobacterium spp. isolated from farmed and natural populations of striped bass and catfish. Applied and Environmental Microbiology, 57(11), 3114-3120.

Begrem, S., Ivaniuk, F., Gigout-Chevalier, F., Kolypczuk, L., Bonnetot, S., Leroi, F., ... & Passerini, D. (2020). New Insight into Antimicrobial Compounds from Food and Marine-Sourced Carnobacterium Species through Phenotype and Genome Analyses. Microorganisms, 8(7), 1093.

Bertelli, C., Laird, M. R., Williams, K. P., Simon Fraser University Research Computing Group, Lau, B. Y., Hoad, G., ... & Brinkman, F. S. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic acids research, 45(W1), W30-W35.

Bhugaloo-Vial, P., Dousset, X., Metivier, A., Sorokine, O., Anglade, P., Boyaval, P., & Marion, D. (1996). Purification and amino acid sequences of piscicocins V1a and V1b, two class IIa bacteriocins secreted by Carnobacterium piscicola V1 that display significantly different levels of specific inhibitory activity. Applied and Environmental Microbiology, 62(12), 4410-4416.

Björkroth, J., Ristiniemi, M., Vandamme, P., & Korkeala, H. (2005). Enterococcus species dominating in fresh modified-atmosphere-packaged, marinated broiler legs are overgrown by Carnobacterium and Lactobacillus species during storage at 6 C. International Journal of Food Microbiology, 97(3), 267-276.

Blom, J., Kreis, J., Spänig, S., Juhre, T., Bertelli, C., Ernst, C., & Goesmann, A. (2016). EDGAR 2.0: an enhanced software platform for comparative gene content analyses. Nucleic acids research, 44(W1), W22-W28.

Bourdichon, F., Casaregola, S., Farrokh, C., Frisvad, J. C., Gerds, M. L., Hammes, W. P., ... & Powell, I. B. (2012). Food fermentations: microorganisms with technological beneficial use. International journal of food microbiology, 154(3), 87-97.

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., ... & Shukla, M. (2015). RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Scientific reports, 5, 8365.

Briggiler-Marcó, M., Capra, M. L., Quiberoni, A., Vinderola, G., Reinheimer, J. A., & Hynes, E. (2007). Nonstarter Lactobacillus strains as adjunct cultures for cheese making: in vitro characterization and performance in two model cheeses. Journal of Dairy Science, 90(10), 4532-4542.

Bruno, D. W., Noguera, P. A., & Poppe, T. T. (2013). A colour atlas of salmonid diseases (Vol. 91). Springer Science & Business Media.

Buchanan, R. L., & Klawitter, L. A. (1991). EFFECTIVENESS OF CARNOBACTERIUM PISCICOLA LK5 FOR CONTROLLING THE GROWTH OF LISTERIA MONOCYTOGENES SCOTT A IN REFRIGERATED FOODS 1. Journal of food safety, 12(3), 219-236.

Cailliez-Grimal, C., Afzal, M. I., & Revol-Junelles, A. M. (2014). Carnobacterium.

Cailliez-Grimal, C., Chaillou, S., Anba-Mondoloni, J., Loux, V., Afzal, M. I., Rahman, A., ... & Leisner, J. J. (2013). Complete chromosome sequence of Carnobacterium maltaromaticum LMA 28. Genome announcements, 1(1).

Cailliez-Grimal, C., Edima, H. C., Revol-Junelles, A. M., & Millière, J. B. (2007). Carnobacterium maltaromaticum: the only Carnobacterium species in French ripened soft cheeses as revealed by polymerase chain reaction detection. Journal of dairy science, 90(3), 1133-1138.

Cartamil, D. P. (2009). Movement patterns, habitat preferences, and fisheries biology of the common thresher shark (Alopias vulpinus) in the Southern California Bight.

Casaburi, A., Nasi, A., Ferrocino, I., Di Monaco, R., Mauriello, G., Villani, F., & Ercolini, D. (2011). Spoilage-related activity of Carnobacterium maltaromaticum strains in air-stored and vacuum-packed meat. Applied and environmental microbiology, 77(20), 7382-7393.

Chaillou, S., Chaulot-Talmon, A., Caekebeke, H., Cardinal, M., Christieans, S., Denis, C., ... & Joffraud, J. J. (2015). Origin and ecological selection of core and food-specific bacterial communities associated with meat and seafood spoilage. The ISME journal, 9(5), 1105-1118.

Chaudhari, N. M., Gupta, V. K., & Dutta, C. (2016). BPGA-an ultra-fast pan-genome analysis pipeline. Scientific reports, 6(1), 1-10.

Chen, L., Zheng, D., Liu, B., Yang, J., & Jin, Q. (2016). VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. Nucleic acids research, 44(D1), D694-D697.

Chenoll, E., Macian, M. C., Elizaquivel, P., & Aznar, R. (2007). Lactic acid bacteria associated with vacuum-packed cooked meat product spoilage: population analysis by rDNA-based methods. Journal of Applied Microbiology, 102(2), 498-508.

Chmelař, D., Matušek, A., Korger, J., Durnová, E., Steffen, M., & Chmelařová, E. (2002). Isolation of Carnobacterium piscicola from human Pus—Case report. Folia microbiologica, 47(4), 455-457.

Collins, M. D., Farrow, J. A. E., Phillips, B. A., Ferusu, S., & Jones, D. (1987). Classification of Lactobacillus divergens, Lactobacillus piscicola, and some catalase-negative, asporogenous, rod-shaped bacteria from poultry in a new genus, Carnobacterium. International Journal of Systematic and Evolutionary Microbiology, 37(4), 310-316.

Cosentino, S., Larsen, M. V., Aarestrup, F. M., & Lund, O. (2013). PathogenFinder-distinguishing friend from foe using bacterial whole genome sequence data. PloS one, 8(10), e77302.

Cotter, P. D., Hill, C., & Ross, R. P. (2005). Bacteriocins: developing innate immunity for food. Nature Reviews Microbiology, 3(10), 777-788.

Cotter, P. D., Ross, R. P., & Hill, C. (2013). Bacteriocins—a viable alternative to antibiotics?. Nature Reviews Microbiology, 11(2), 95-105.

Darling, A. C., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome research, 14(7), 1394-1403.

Douglas, G. L., & Klaenhammer, T. R. (2010). Genomic evolution of domesticated microorganisms. Annual Review of Food Science and Technology, 1, 397-414.

Douillard, F. P., & De Vos, W. M. (2014, August). Functional genomics of lactic acid bacteria: from food to health. In Microbial cell factories (Vol. 13, No. S1, p. S8). BioMed Central.

Drider, D., Fimland, G., Héchard, Y., McMullen, L. M., & Prévost, H. (2006). The continuing story of class IIa bacteriocins. Microbiology and molecular biology reviews, 70(2), 564-582.

Duan, S., Hu, X., Li, M., Miao, J., Du, J., & Wu, R. (2016). Composition and metabolic activities of the bacterial community in shrimp sauce at the flavor-forming stage of fermentation as revealed by metatranscriptome and 16S rRNA gene sequencings. Journal of agricultural and food chemistry, 64(12), 2591-2603.

Duffes, F., Leroi, F., Boyaval, P., & Dousset, X. (1999). Inhibition of Listeria monocytogenes by Carnobacterium spp. strains in a simulated cold smoked fish system stored at 4 C. International journal of food microbiology, 47(1-2), 33-42.

Edima, H. C., Cailliez-Grimal, C., Revol-Junelles, A. M., Rondags, E., & Millière, J. B. (2008). Impact of pH and Temperature on the Acidifying Activity of Carnobacterium maltaromaticum. Journal of dairy science, 91(10), 3806-3813.

Edima, H. C., Cailliez-Grimal, C., Revol-Junelles, A. M., Tonti, L., Linder, M., & Millière, J. B. (2007). A selective enumeration medium for Carnobacterium maltaromaticum. Journal of microbiological methods, 68(3), 516-521.

EFSA Scientific Committee. (2012). Guidance on selected default values to be used by the EFSA Scientific Committee, Scientific Panels and Units in the absence of actual measured data. EFSA journal, 10(3), 2579.

Emborg, J., Laursen, B. G., Rathjen, T., & Dalgaard, P. (2002). Microbial spoilage and formation of biogenic amines in fresh and thawed modified atmosphere-packed salmon (Salmo salar) at 2° C. Journal of Applied Microbiology, 92(4), 790-799.

Ercolini, D., Russo, F., Nasi, A., Ferranti, P., & Villani, F. (2009). Mesophilic and psychrotrophic bacteria from meat and their spoilage potential in vitro and in beef. Applied and environmental microbiology, 75(7), 1990-2001.

Fougy, L., Desmonts, M. H., Coeuret, G., Fassel, C., Hamon, E., Hézard, B., ... & Chaillou, S. (2016). Reducing salt in raw pork sausages increases spoilage and correlates with reduced bacterial diversity. Applied and environmental microbiology, 82(13), 3928-3939.

Grant, I. R., & Patterson, M. F. (1991). Effect of irradiation and modified atmosphere packaging on the microbiological safety of minced pork stored under temperature abuse conditions. International journal of food science & technology, 26(5), 521-533.

Hammes, W. P., & Hertel, C. (2009). Genus I. Lactobacillus Beijerinck 1901, 212AL. Bergey's manual of systematic bacteriology, 3, 465-511.

Hammes, W. P., & Hertel, C. H. R. I. S. T. I. A. N. (2006). The genera lactobacillus and carnobacterium. Prokaryotes, 4, 320-403.

Hammi, I., Delalande, F., Belkhou, R., Marchioni, E., Cianferani, S., & Ennahar, S. (2016). Maltaricin CPN, a new class II a bacteriocin produced by Carnobacterium maltaromaticum CPN isolated from mould-ripened cheese. Journal of applied microbiology, 121(5), 1268-1274.

Herbin, S., Mathieu, F., Brulé, F., Branlant, C., Lefebvre, G., & Lebrihi, A. (1997). Characteristics and genetic determinants of bacteriocin activities produced by Carnobacterium piscicola CP5 isolated from cheese. Current microbiology, 35(6), 319-326.

Hiu, S. F., Holt, R. A., Sriranganathan, N., Seidler, R. J., & Fryer, J. L. (1984). Lactobacillus piscicola, a new species from salmonid fish. International Journal of Systematic and Evolutionary Microbiology, 34(4), 393-400.

Holzapfel, W. H. (1992). Culture media for non-sporulating Gram-positive food spoilage bacteria. International journal of food microbiology, 17(2), 113-133.

Iskandar, C. F., Borges, F., Taminiau, B., Daube, G., Zagorec, M., Remenant, B., ... & Cailliez-Grimal, C. (2017). Comparative genomic analysis reveals ecological differentiation in the genus Carnobacterium. Frontiers in microbiology, 8, 357.

Iskandar, C. F., Cailliez-Grimal, C., Rahman, A., Rondags, E., Remenant, B., Zagorec, M., ... & Revol-Junelles, A. M. (2016). Genes associated to lactose metabolism illustrate the high diversity of Carnobacterium maltaromaticum. Food microbiology, 58, 79-86.

Jääskeläinen, E., Hultman, J., Parshintsev, J., Riekkola, M. L., & Björkroth, J. (2016). Development of spoilage bacterial community and volatile compounds in chilled beef under vacuum or high oxygen atmospheres. International Journal of Food Microbiology, 223, 25-32.

Jack, R. W., Wan, J., Gordon, J., Harmark, K., Davidson, B. E., Hillier, A. J., ... & Coventry, M. J. (1996). Characterization of the chemical and antimicrobial properties of piscicolin 126, a bacteriocin produced by Carnobacterium piscicola JG126. Applied and Environmental Microbiology, 62(8), 2897-2903.

Jones, R. J. (2004). Observations on the succession dynamics of lactic acid bacteria populations in chill-stored vacuum-packaged beef. International journal of food microbiology, 90(3), 273-282.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., & Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. Nucleic acids research, 40(D1), D109-D114.

Kim, D. H., & Austin, B. (2006). Innate immune responses in rainbow trout (Oncorhynchus mykiss, Walbaum) induced by probiotics. Fish & shellfish immunology, 21(5), 513-524.

Kim, D. H., Brunt, J., & Austin, B. (2007). Microbial diversity of intestinal contents and mucus in rainbow trout (Oncorhynchus mykiss). Journal of applied microbiology, 102(6), 1654-1664.

Kleerebezem, M., Hols, P., Bernard, E., Rolain, T., Zhou, M., Siezen, R. J., & Bron, P. A. (2010). The extracellular biology of the lactobacilli. FEMS microbiology reviews, 34(2), 199-230.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular biology and evolution, 35(6), 1547-1549.

Larrouture-Thiveyrat, C., & Montel, M. C. (2003). Effects of environmental factors on leucine catabolism by Carnobacterium piscicola. International journal of food microbiology, 81(3), 177-184.

Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R. L., ... & Lund, O. (2012). Multilocus sequence typing of total-genome-sequenced bacteria. Journal of clinical microbiology, 50(4), 1355-1361.

Laursen, B. G., Bay, L., Cleenwerck, I., Vancanneyt, M., Swings, J., Dalgaard, P., & Leisner, J. J. (2005). Carnobacterium divergens and Carnobacterium maltaromaticum as spoilers or protective cultures in meat and seafood: phenotypic and genotypic characterization. Systematic and applied microbiology, 28(2), 151-164.

Laursen, B. G., Leisner, J. J., & Dalgaard, P. (2006). Carnobacterium species: effect of metabolic activity and interaction with Brochothrix thermosphacta on sensory characteristics of modified atmosphere packed shrimp. Journal of Agricultural and Food Chemistry, 54(10), 3604-3611.

Leblanc, E. L., Stiles, M. E., McMullen, L. M., & Leisner, J. J. (1997). Characterizationof the prevailing flora of Carnobacterium spp. growing on CO2 modified atmosphere packaged rainbow trout (Salmo gairdner) fillets. In Seafood from Producer to Consumer, Integrated Approach to Quality, Seafood from producer to consumer, integrated approach to quality: proceedings of the international seafood conference on the occasion of the 25. anniversary of the WEFTA, Noordwijkerhout, 1995 (pp. 445-454). Elsevier.

Lee, I., Kim, Y. O., Park, S. C., & Chun, J. (2016). OrthoANI: an improved algorithm and software for calculating average nucleotide identity. International journal of systematic and evolutionary microbiology, 66(2), 1100-1103.

Leisner, J. J., Hansen, M. A., Larsen, M. H., Hansen, L., Ingmer, H., & Sørensen, S. J. (2012). The genome sequence of the lactic acid bacterium, Carnobacterium maltaromaticum ATCC 35586 encodes potential virulence factors. International journal of food microbiology, 152(3), 107-115.

Leisner, J. J., Laursen, B. G., Prévost, H., Drider, D., & Dalgaard, P. (2007). Carnobacterium: positive and negative effects in the environment and in foods. FEMS microbiology reviews, 31(5), 592-613.

Leroi, F., Joffraud, J. J., Chevalier, F., & Cardinal, M. (1998). Study of the microbial ecology of cold-smoked salmon during storage at 8 C. International journal of food microbiology, 39(1-2), 111-121.

Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic acids research, 44(W1), W242-W245.

Liu, M., Nauta, A., Francke, C., & Siezen, R. J. (2008). Comparative genomics of enzymes in flavor-forming pathways from amino acids in lactic acid bacteria. Applied and environmental microbiology, 74(15), 4590-4600.

Loch, T. P., Kumar, R., Xu, W., & Faisal, M. (2011). Carnobacterium maltaromaticum infections in feral Oncorhynchus spp.(Family Salmonidae) in Michigan. The Journal of Microbiology, 49(5), 703-713.

Loch, T. P., Xu, W., Fitzgerald, S. M., & Faisal, M. (2008). Isolation of a Carnobacterium maltaromaticum-like bacterium from systemically infected lake whitefish (Coregonus clupeaformis). FEMS microbiology letters, 288(1), 76-84.

Lorca, G. L., Barabote, R. D., Zlotopolski, V., Tran, C., Winnen, B., Hvorup, R. N., ... & Saier Jr, M. H. (2007). Transport capabilities of eleven gram-positive bacteria: comparative genomic analyses. Biochimica et Biophysica Acta (BBA)-Biomembranes, 1768(6), 1342-1366.

Marilley, L., & Casey, M. G. (2004). Flavours of cheese products: metabolic pathways, analytical tools and identification of producing strains. International journal of food microbiology, 90(2), 139-159.

Martin-Visscher, L. A., van Belkum, M. J., Garneau-Tsodikova, S., Whittal, R. M., Zheng, J., McMullen, L. M., & Vederas, J. C. (2008). Isolation and characterization of carnocyclin A, a novel circular bacteriocin produced by Carnobacterium maltaromaticum UAL307. Applied and environmental microbiology, 74(15), 4756-4763.

Masson, F., Talon, R., & Montel, M. C. (1996). Histamine and tyramine production by bacteria from meat products. International Journal of Food Microbiology, 32(1-2), 199-207.

McAuliffe, O., Ross, R. P., & Hill, C. (2001). Lantibiotics: structure, biosynthesis and mode of action. FEMS microbiology reviews, 25(3), 285-308.

McMullen, L. M., & Stiles, M. E. (1993). Microbial ecology of fresh pork stored under modified atmosphere at− 1, 4.4 and 10° C. International journal of food microbiology, 18(1), 1-14.

Metivier, A., Pilet, M. F., Dousset, X., Sorokine, O., Anglade, P., Zagorec, M., ... & Fremaux, C. (1998). Divercin V41, a new bacteriocin with two disulphide bonds produced by Carnobacterium divergens V41: primary structure and genomic organization. Microbiology, 144(10), 2837-2844.

Michel, C., Faivre, B., & Kerouault, B. (1986). Biochemical identification of Lactobacillus piscicola strains from France and Belgium. Diseases of aquatic organisms, 2(1), 27-30.

Miller III, A., Morgan, M. E., & Libbey, L. M. (1974). Lactobacillus maltaromicus, a new species producing a malty aroma1. International Journal of Systematic and Evolutionary Microbiology, 24(3), 346-354.

Millière, J. B., & Lefebvre, G. (1994). Carnobacterium piscicola, a common species of French soft cheeses from cow's raw milk. Nederlands melk en Zuiveltijdschrift, 48(1), 19-30.

Millière, J. B., Michel, M., Mathieu, F., & Lefebvre, G. (1994). Presence of Carnobacterium spp. in French surface mould-ripened soft-cheese. Journal of applied bacteriology, 76(3), 264-269.

Mohamed, M. H., Abogabal, I. I., & Einas, H. (2017). Isolation and Molecular Characterization of β-hemolytic Carnobacterium maltaromaticum from Naturally Infected Longnose Parrot Fish (Hipposcarus harid) in Hurghada Red Sea, Egypt. International Journal of Innovative Studies in Aquatic Biology and Fisheries (IJISABF), 3(1), 1-8.

Mora, D., Scarpellini, M., Franzetti, L., Colombo, S., & Galli, A. (2003). Reclassification of Lactobacillus maltaromicus (Miller et al. 1974) DSM 20342T and DSM 20344 and Carnobacterium piscicola (Collins et al. 1987) DSM 20730T and DSM 20722 as Carnobacterium maltaromaticum comb. nov. International Journal of Systematic and Evolutionary Microbiology, 53(3), 675-678.

Morea, M., Baruzzi, F., & Cocconcelli, P. S. (1999). Molecular and physiological characterization of dominant bacterial populations in traditional Mozzarella cheese processing. Journal of Applied Microbiology, 87(4), 574-582.

Paludan-Müller, C., Dalgaard, P., Huss, H. H., & Gram, L. (1998). Evaluation of the role of Carnobacterium piscicola in spoilage of vacuum-and modified-atmosphere-packed cold-smoked salmon stored at 5 C. International journal of food microbiology, 39(3), 155-166.

Pikuta, E. V., & Hoover, R. B. (2014). The genus Carnobacterium. Lactic Acid Bacteria: Biodiversity and Taxonomy, eds EV Pikuta and RB Hoover (Chichester: John Wiley & Sons, Ltd.), 109-123.

Quadri, L. E., Sailer, M., Roy, K. L., Vederas, J. C., & Stiles, M. E. (1994). Chemical and genetic characterization of bacteriocins produced by Carnobacterium piscicola LV17B. Journal of Biological Chemistry, 269(16), 12204-12211.

Rahman, A., Cailliez-Grimal, C., Bontemps, C., Payot, S., Chaillou, S., Revol-Junelles, A. M., & Borges, F. (2014). High genetic diversity among strains of the unindustrialized lactic acid bacterium Carnobacterium maltaromaticum in dairy products as revealed by multilocus sequence typing. Applied and environmental microbiology, 80(13), 3920-3929.

Rahman, A., Gleinser, M., Lanhers, M. C., Riedel, C. U., Foligné, B., Hanse, M., ... & Mangavel, C. (2014). Adaptation of the lactic acid bacterium Carnobacterium maltaromaticum LMA 28 to the mammalian gastrointestinal tract: from survival in mice to interaction with human cells. International dairy journal, 34(1), 93-99.

Ramia, N. E., El Kheir, S. M., Taha, S., Mangavel, C., Revol-Junelles, A. M., & Borges, F. (2019). Multilocus sequence typing of Carnobacterium maltaromaticum strains associated with fish disease and dairy products. Journal of applied microbiology, 126(2), 377-387.

Remenant, B., Borges, F., Cailliez-Grimal, C., Revol-Junelles, A. M., Marché, L., Lajus, A., ... & Zagorec, M. (2016). Draft genome sequence of Carnobacterium divergens V41, a bacteriocin-producing strain. Genome announcements, 4(5).

Robertson, P. A. W., O'Dowd, C., Burrells, C., Williams, P., & Austin, B. (2000). Use of Carnobacterium sp. as a probiotic for Atlantic salmon (Salmo salar L.) and rainbow trout (Oncorhynchus mykiss, Walbaum). Aquaculture, 185(3-4), 235-243.

Roh, H., Kim, B. S., Lee, M. K., Park, C. I., & Kim, D. H. (2020). Genome-wide comparison of Carnobacterium maltaromaticum derived from diseased fish harbouring important virulence-related genes. Journal of Fish Diseases, 43(9), 1029-1037.

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution, 4(4), 406-425.

Sakala, R. M., Hayashidani, H., Kato, Y., Hirata, T., Makino, Y., Fukushima, A., ... & Ogawa, M. (2002). Change in the composition of the microflora on vacuum-packaged beef during chiller storage. International journal of food microbiology, 74(1-2), 87-99.

Schaffer, P. A., Lifland, B., Sommeran, S. V., Casper, D. R., & Davis, C. R. (2013). Meningoencephalitis associated with Carnobacterium maltaromaticum–like bacteria in stranded juvenile salmon sharks (Lamna ditropis). Veterinary pathology, 50(3), 412-417.

Schirmer, B. C., Heir, E., & Langsrud, S. (2009). Characterization of the bacterial spoilage flora in marinated pork products. Journal of Applied Microbiology, 106(6), 2106-2116.

Seedorf, H., Griffin, N. W., Ridaura, V. K., Reyes, A., Cheng, J., Rey, F. E., ... & Spormann, A. M. (2014). Bacteria from diverse habitats colonize and compete in the mouse gut. Cell, 159(2), 253-266.

Shaw, B. G., & Harding, C. D. (1984). A numerical taxonomic study of lactic acid bacteria from vacuum-packed beef, pork, lamb and bacon. Journal of Applied Bacteriology, 56(1), 25-40.

Smit, B. A., Engels, W. J., & Smit, G. (2009). Branched chain aldehydes: production and breakdown pathways and relevance for flavour in foods. Applied Microbiology and Biotechnology, 81(6), 987-999.

Soares, S. C., Geyik, H., Ramos, R. T., de Sá, P. H., Barbosa, E. G., Baumbach, J., ... & Azevedo, V. (2016). GIPSy: genomic island prediction software. Journal of biotechnology, 232, 2-11.

Starliper, C. E., Shotts, E. B., & Brown, J. (1992). Isolation of Carnobacterium piscicola and an unidentified Gram-positive bacillus from sexually mature and post-spawning rainbow trout Oncorhynchus mykiss. Diseases of aquatic organisms, 13(3), 181-187.

Steele, L. M., Lowe, C. G., Okihiro, M. S., Dillon, J. G., & Berlemont, R. (2019). Pseudogenization, genome streamlining and specific gene repertoire landmark the genomes of Carnobacterium maltaromaticum isolated from diseased sharks. BioRxiv, 600684.

Sun, Z., Harris, H. M., McCann, A., Guo, C., Argimón, S., Zhang, W., ... & Liu, W. (2015). Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. Nature communications, 6(1), 1-13.

Susiluoto, T., Korkeala, H., & Björkroth, K. J. (2003). Leuconostoc gasicomitatum is the dominating lactic acid bacterium in retail modified-atmosphere-packaged marinated broiler meat strips on sell-by-day. International journal of food microbiology, 80(1), 89-97.

Suzuki, M., Yamamoto, T., Kawai, Y., Inoue, N., & Yamazaki, K. (2005). Mode of action of piscicocin CS526 produced by Carnobacterium piscicola CS526. Journal of applied microbiology, 98(5), 1146-1151.

Tahiri, I., Desbiens, M., Benech, R., Kheadr, E., Lacroix, C., Thibault, S., ... & Fliss, I. (2004). Purification, characterization and amino acid sequencing of divergicin M35: a novel class IIa bacteriocin produced by Carnobacterium divergens M35. International journal of food microbiology, 97(2), 123-136.

Thierry, A., & Maillard, M. B. (2002). Production of cheese flavour compounds derived from amino acid catabolism by Propionibacterium freudenreichii. Le Lait, 82(1), 17-32.

Toranzo, A. E., Romalde, J. L., Nunez, S., Figueras Huerta, A., & Barja, J. L. (1993). An epizootic in farmed, market-sized rainbow trout in Spain caused by a strain of Carnobactenum piscicola of unusual virulence.

Tulini, F. L., Lohans, C. T., Bordon, K. C., Zheng, J., Arantes, E. C., Vederas, J. C., & De Martinis, E. C. (2014). Purification and characterization of antimicrobial peptides from fish

isolate Carnobacterium maltaromaticum C2: Carnobacteriocin X and carnolysins A1 and A2. International journal of food microbiology, 173, 81-88.

van Heel, A. J., de Jong, A., Song, C., Viel, J. H., Kok, J., & Kuipers, O. P. (2018). BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. Nucleic acids research, 46(W1), W278-W281.

Vihavainen, E. J., & Björkroth, K. J. (2007). Spoilage of value-added, high-oxygen modified-atmosphere packaged raw beef steaks by Leuconostoc gasicomitatum and Leuconostoc gelidum. International journal of food microbiology, 119(3), 340-345.

Wallbanks, S., Martinez-Murcia, A. J., Fryer, J. L., Phillips, B. A., & Collins, M. D. (1990). 16S rRNA sequence determination for members of the genus Carnobacterium and related lactic acid bacteria and description of Vagococcus salmoninarum sp. nov. International Journal of Systematic and Evolutionary Microbiology, 40(3), 224-230.

Wasney, M. A., Holley, R. A., & Jayas, D. S. (2001). Cresol red thallium acetate sucrose inulin (CTSI) agar for the selective recovery of Carnobacterium spp. International journal of food microbiology, 64(1-2), 167-174.

Worobo, R. W., Van Belkum, M. J., Sailer, M., Roy, K. L., Vederas, J. C., & Stiles, M. E. (1995). A signal peptide secretion-dependent bacteriocin from Carnobacterium divergens. Journal of Bacteriology, 177(11), 3143-3149.

Yamazaki, K., Suzuki, M., Kawai, Y., Inoue, N., & Montville, T. J. (2005). Purification and characterization of a novel class IIa bacteriocin, piscicocin CS526, from surimi-associated Carnobacterium piscicola CS526. Applied and environmental microbiology, 71(1), 554-557.

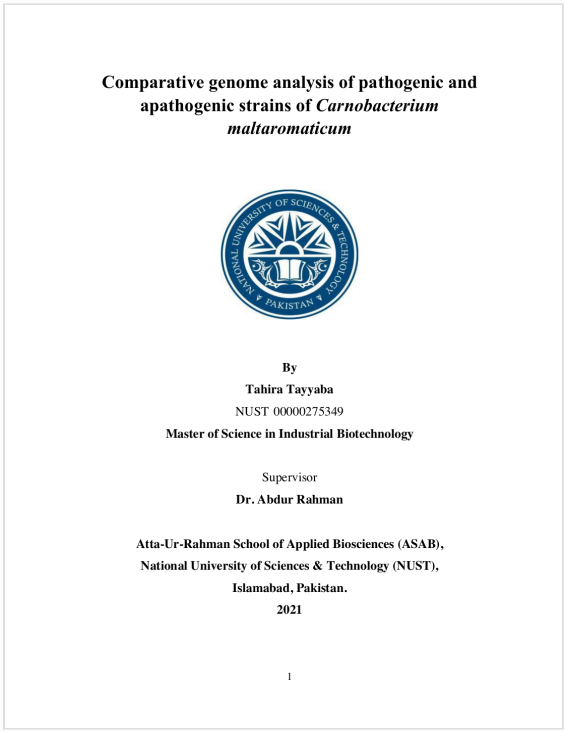
Zhang, P., Gänzle, M., & Yang, X. (2019). Complementary antibacterial effects of bacteriocins and organic acids as revealed by comparative analysis of Carnobacterium spp. from meat. Applied and environmental microbiology, 85(20), e01227-19.

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Ms. Tahira Tayyaba
Assignment title:	Animals_Siddique 2021
Submission title:	MS_Thesis Tahir Tayyaba_2021
File name:	Tahira_Thesis_Plagiarism_file.docx
File size:	4.15M
Page count:	66
Word count:	13,614
Character count:	78,434
Submission date:	02-Nov-2021 11:52PM (UTC-0700)
Submission ID:	1691804726



Comparative genome analysis of pathogenic and apathogenic strains of *Carnobacterium maltaromaticum*

By

Tahira Tayyaba

NUST 00000275349

Master of Science in Industrial Biotechnology

Supervisor

Dr. Abdur Rahman

Atta-Ur-Rahman School of Applied Biosciences (ASAB),
National University of Sciences & Technology (NUST),
Islamabad, Pakistan.
2021

1

MS_Thesis Tahir Tayyaba_2021

ORIGINALITY REPORT

10%	4%	9%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Muhammad Inam Afzal, Thibaut Jacquet, Stéphane Delaunay, Frédéric Borges et al. "Carnobacterium maltaromaticum: Identification, isolation tools, ecology and technological aspects in dairy products", Food Microbiology, 2010 Publication	1%
2	"The Prokaryotes", Springer Nature, 2014 Publication	1%
3	HyeongJin Roh, Bo Seong Kim, Mu Kun Lee, Chan - Il Park, Do - Hyung Kim. " Genome - wide comparison of derived from diseased fish harbouring important virulence - related genes ", Journal of Fish Diseases, 2020 Publication	1%
4	curis.ku.dk Internet Source	1%
5	Furqan Awan, Yuhao Dong, Jin Liu, Nannan Wang, Muhammad Hassan Mushtaq, Chengping Lu, Yongjie Liu. "Comparative genome analysis provides deep insights into	<1%

Aeromonas hydrophila taxonomy and virulence-related factors", BMC Genomics, 2018
Publication

6	Jørgen J. Leisner, Birgit Groth Laursen, Hervé Prévost, Djamel Drider, Paw Dalgaard. " positive and negative effects in the environment and in foods ", FEMS Microbiology Reviews, 2007 Publication	<1 %
7	www.ncbi.nlm.nih.gov Internet Source	<1 %
8	Christelle F. Iskandar, Catherine Cailliez-Grimal, Abdur Rahman, Emmanuel Rondags et al. "Genes associated to lactose metabolism illustrate the high diversity of Carnobacterium maltaromaticum", Food Microbiology, 2016 Publication	<1 %
9	Karl-Heinz Schleifer. "Phylum XIII. Firmicutes Gibbons and Murray 1978, 5 (Firmacutes [sic] Gibbons and Murray 1978, 5)", Systematic Bacteriology, 2009 Publication	<1 %
10	www.frontiersin.org Internet Source	<1 %

11	Chrizelle W. Beukes, Marike Palmer, Puseletso Manyaka, Wai Y. Chan et al. "Genome Data Provides High Support for Generic Boundaries in Burkholderia Sensu Lato", Frontiers in Microbiology, 2017 Publication	<1 %
12	Laura Martinez Steele, Christopher G Lowe, Mark S Okihiro, Jesse G. Dillon, Renaud Berlemont. " Pseudogenization, genome streamlining and specific gene repertoire landmark the genomes of isolated from diseased sharks ", Cold Spring Harbor Laboratory, 2019 Publication	<1 %
13	Abdur Rahman, Marita Gleinser, Marie-Claire Lanhers, Christian U. Riedel et al. "Adaptation of the lactic acid bacterium Carnobacterium maltaromaticum LMA 28 to the mammalian gastrointestinal tract: From survival in mice to interaction with human cells", International Dairy Journal, 2014 Publication	<1 %
14	Simon Begrem, Flora Ivaniuk, Frédérique Gigout-Chevalier, Laetitia Kolypczuk et al. "New Insight into Antimicrobial Compounds from Food and Marine-Sourced Carnobacterium Species through Phenotype	<1 %

and Genome Analyses", Microorganisms, 2020
Publication

15	"The Prokaryotes", Springer Science and Business Media LLC, 2006 Publication	<1 %
16	Abdur Rahman, Sara M. El Kheir, Alexandre Back, Cécile Mangavel, Anne-Marie Revol-Junelles, Frédéric Borges. "Repeat-based Sequence Typing of Carnobacterium maltaromaticum", International Journal of Food Microbiology, 2016 Publication	<1 %
17	"The Genera of Lactic Acid Bacteria", Springer Science and Business Media LLC, 1995 Publication	<1 %
18	Submitted to Higher Education Commission Pakistan Student Paper	<1 %
19	J.J. Leisner, M.A. Hansen, M.H. Larsen, L. Hansen, H. Ingmer, S.J. Sørensen. "The genome sequence of the lactic acid bacterium, Carnobacterium maltaromaticum ATCC 35586 encodes potential virulence factors", International Journal of Food Microbiology, 2012 Publication	<1 %

20	Internet Source	<1%
21	pregi.bi.up.ac.za Internet Source	<1%
22	www.pathogenomics.sfu.ca Internet Source	<1%
23	Huys, Geert, Jørgen Leisner, and Johanna Björkroth. "The Lesser LAB Gods", Lactic Acid Bacteria Microbiological and Functional Aspects Fourth Edition, 2011. Publication	<1%
24	Nancy E Ramia, Sara M. El Kheir, Samir Taha, Cécile Mangavel, Anne-Marie Revol-Junelles, Frédéric Borges. " MultiLocus Sequence Typing of strains associated with fish disease and dairy products ", Journal of Applied Microbiology, 2018 Publication	<1%
25	hal.archives-ouvertes.fr Internet Source	<1%
26	bioinformatics.psb.ugent.be Internet Source	<1%
27	journals.plos.org Internet Source	<1%
28	bmcoralhealth.biomedcentral.com Internet Source	<1%

29	www.cidrap.umn.edu Internet Source	<1 %
30	www.mdpi.com Internet Source	<1 %
31	www.researcherslinks.com Internet Source	<1 %
32	Ravindran, Aravind, Neha Jalan, Joshua S. Yuan, Nian Wang, and Dennis C. Gross. "Comparative genomics of Pseudomonas syringae pv. syringae strains B301D and HS191 and insights into intrapathovar traits associated with plant pathogenesis", MicrobiologyOpen, 2015. Publication	<1 %
33	openaccessebooks.com Internet Source	<1 %
34	Ali, Amjad, Siomar C. Soares, Anderson R. Santos, Luis C. GuimarÃ£es, Eudes Barbosa, Sintia S. Almeida, VinÃcius A.C. Abreu, Adriana R. Carneiro, Rommel T.J. Ramos, Syeda M. Bakhtiar, Syed S. Hassan, David W. Ussery, Stephen On, Artur Silva, Maria P. Schneider, Andrey P. Lage, Anderson Miyoshi, and Vasco Azevedo. "Campylobacter fetus subspecies: Comparative genomics and prediction of potential virulence targets", Gene, 2012. Publication	<1 %

35	Bjørn Christian Schirmer, Solveig Langsrud. "A dissolving CO2 headspace combined with organic acids prolongs the shelf-life of fresh pork", Meat Science, 2010 Publication	<1 %
36	Marisa E. Garcés, Nelda L. Olivera, Melania Fernández, Carla Riva Rossi, Cynthia Sequeiros. " Antimicrobial activity of bacteriocin - producing spp. isolated from healthy Patagonian trout and their potential for use in aquaculture ", Aquaculture Research, 2020 Publication	<1 %
37	Xiu-Zhong Zhang, Chang-Wei Lei, Jin-Xin Zeng, Yan-Peng Chen et al. "An IncX1 plasmid isolated from Salmonella enterica subsp. enterica serovar Pullorum carrying blaTEM-1B, sul2, arsenic resistant operons", Plasmid, 2018 Publication	<1 %
38	Yoonjee Chang, Ju-Hoon Lee, Hakdong Shin, Sunggi Heu, Sangryeol Ryu. "Characterization and complete genome sequence analysis of Staphylococcus aureus bacteriophage SA12", Virus Genes, 2013 Publication	<1 %
39	"The Springer Index of Viruses", Springer Science and Business Media LLC, 2002	<1 %

Publication

40 Amr T. M. Saeb, Khalid A. Al-Rubeaan, Mohamed Abouelhoda, Manojkumar Selvaraju, Hamsa T. Tayeb. "Genome sequencing and analysis of the first spontaneous Nanosilver resistant bacterium Proteus mirabilis strain SCDR1", Antimicrobial Resistance & Infection Control, 2017 <1 %
Publication

41 Camila Carlos, Letícia Bianca Pereira, Laura Maria Mariscal Ottoboni. "Comparative genomics of Paracoccus sp. SM22M-07 isolated from coral mucus: insights into bacteria-host interactions", Current Genetics, 2016 <1 %
Publication

42 Camilo Gómez-Garzón, Alejandra Hernández-Santana, Jenny Dussán. "Comparative genomics reveals Lysinibacillus sphaericus group comprises a novel species", BMC Genomics, 2016 <1 %
Publication

43 Che, Dongsheng, Mohammad Hasan, and Bernard Chen. "Identifying Pathogenicity Islands in Bacterial Pathogenomics Using Computational Approaches", Pathogens, 2014. <1 %
Publication

44	Vasundhara Paliwal, Sajan C. Raju, Arnab Modak, Prashant S. Phale, Hemant J. Purohit. "Pseudomonas putida CSV86: A Candidate Genome for Genetic Bioaugmentation", PLoS ONE, 2014 Publication	<1 %
45	www.ulntc.ru Internet Source	<1 %
46	Fanny George, Séverine Mahieux, Catherine Daniel, Marie Titécat et al. "Assessment of Pb(II), Cd(II), and Al(III) Removal Capacity of Bacteria from Food and Gut Ecological Niches: Insights into Biodiversity to Limit Intestinal Biodisponibility of Toxic Metals", Microorganisms, 2021 Publication	<1 %
47	Luís Cláudio Lima de Jesus, Thiago de Jesus Sousa, Nina Dias Coelho-Rocha, Rodrigo Profeta et al. "Safety Evaluation of Lactobacillus delbrueckii subsp. lactis CIDCA 133: a Health-Promoting Bacteria", Probiotics and Antimicrobial Proteins, 2021 Publication	<1 %
48	dspace.mit.edu Internet Source	<1 %
49	gutpathogens.biomedcentral.com Internet Source	<1 %

50	lookformedical.com Internet Source	<1 %
51	prodata.swmed.edu Internet Source	<1 %
52	www.ashg.org Internet Source	<1 %
53	Amjad Ali, Siomar C. Soares, Anderson R. Santos, Luis C. Guimarães et al. "Campylobacter fetus subspecies: Comparative genomics and prediction of potential virulence targets", Gene, 2012 Publication	<1 %
54	Pallavi Baliga, Malathi Shekar, Girisha Shivani Kallappa. "Genome-Wide Identification and Analysis of Chromosomally Integrated Putative Prophages Associated with Clinical Klebsiella pneumoniae Strains", Current Microbiology, 2021 Publication	<1 %

Exclude quotes	On	Exclude matches	< 7 words
Exclude bibliography	On		