MECHANISTIC STUDIES OF FUSOBACTERIUM DEFENSE AND GENETIC SYSTEMS

Ariana Umaña Torres

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biochemistry

Daniel J. Slade

Irving Coy Allen

Kylie D. Allen

Jingsong Zhu

November 17, 2020
Blacksburg, VA

Keywords: *Fusobacterium*, genetics, transformation, electroporation, host-pathogen interaction, molecular biology, colorectal cancer

Copyright © 2020, Ariana Umaña Torres

Mechanistic studies of *Fusobacterium* genetic and defense systems

Ariana Umaña Torres

ABSTRACT

Fusobacterium are Gram-negative anaerobic bacteria that colonize a variety of eukaryotes including cattle and humans. In humans, *Fusobacterium* coordinates the central architecture of the oral biofilm by expressing an abundance of outer membrane adhesins that mediate bridging between early and late colonizing bacteria. While *Fusobacterium* are mostly considered commensal microorganisms, they can also become an opportunistic pathogen that spreads throughout the human body and promote the development of oral and extra-oral infections and diseases including colorectal cancer. Importantly for this work, many *Fusobacterium* species and strains are recalcitrant to genetic manipulation, the majority of which has led to hindrance in the study of their biology, molecular mechanisms, and pathogenesis.

The genetic intractability of *Fusobacterium* is an obstacle for the development of future treatments for diseases associated with these anaerobic bacteria. Therefore, the creation of tools to enhance genome editing in target species is crucial to understand the molecular mechanisms driving *Fusobacterium* infections. This dissertation exploits innate and adaptive defense systems present in *Fusobacterium* for their use as molecular tools for genome editing.

Accordingly, we first investigated restriction-modification systems with a focus on the role of DNA methyltransferases and endonucleases in host defense and genetic recalcitrance in several strains of *Fusobacterium* through bioinformatic and biochemical approaches. Altogether, over 15 DNA methyltransferases were characterized. Most notably, we identified and characterized two type II DNA methyltransferases that are capable of methylating plasmid DNA by treating with

purified enzymes in-vitro and coexpression approaches in *Escherichia coli* strains, enabling an statiscally improved transformation efficiency via electroporation in *F. nucleatum*.

Also contained in this dissertation is the first detailed description of CRISPR-Cas adaptive immunity systems present in *Fusobacterium* strains. Most of the discovered CRISPR-Cas systems in *Fusobacterium* belong to Class 1 systems. Nonetheless we identified Type II-A and Type VI-C Class 2 systems. The discovery of Cas9 and Cas13c effectors respectively from these systems will be crucial in the development of a new generation of genome-editing tools in *Fusobacterium*.

The studies included in this dissertation provide the framework for overcoming *Fusobacterium* genetic recalcitrance by the implementation of host mimicking techniques. By utilizing restriction-modification system enzymes and the adaptive immunity CRISPR-Cas systems, we will gain a better understanding of how *Fusobacterium* modulates infections and diseases, and ultimately explore the potential of novel therapeutic treatments.

Mechanistic studies of *Fusobacterium* genetic and defense systems

Ariana Umaña Torres

GENERAL AUDIENCE ABSTRACT

The oral cavity has one of the most diverse and largest microbial populations, where microorganisms are capable of colonizing hard surfaces of the teeth and the soft tissues of the oral mucosa. A fundamental member of the oral microbiome is *Fusobacterium*, a Gram-negative bacterium which coordinates the oral biofilm formation by interacting with other microorganisms. In recent studies, *Fusobacterium* has been associated with oral and extra-oral infections and diseases including periodontitis, preterm birth, Lemiere syndrome, inflammatory bowel disease and colorectal cancer. Importantly, many *Fusobacterium* species and strains are challenging to study due to their inability to uptake exogenous DNA and lack of genetic tools, which has hindered the study of their biology, molecular mechanisms and pathogenesis.

The challenges in the genetic manipulation of *Fusobacterium* present a significant obstacle for the development of future treatments for diseases associated with these bacteria. Therefore, the creation of tools to expand bacterial transformation of exogenous DNA and genome editing to more than just one *Fusobacterium* species is crucial to understand how *Fusobacterium* is causing these infections. This dissertation explores the presence and utilization of defense systems, which defend bacteria from phage attack, as an alternative to improve *Fusobacterium* genetics.

Accordingly, we first studied a set of over 15 enzymes that recognize a specific DNA pattern and add a methyl group (DNA methyltransferases) to specific nucleotides in several strains of *Fusobacterium*. We discovered that two of these enzymes improve *Fusobacterium's* ability of importing and genomically incorporating exogenous DNA after an electric discharge permeabilizes the bacterial membrane. Furthermore, for the first time we have described the

composition of CRISPR-Cas bacterial defense systems, that detect invading DNA from viruses and provide protection to *Fusobacterium* strains. These systems have previously been successfully used as genetic tools to achieve genome editing. Thus, their further characterization is warranted to create novel molecular tools in *Fusobacterium*. Altogether, these discoveries will lead to a better comprehension of *Fusobacterium* biology in infections and diseases, while exploring novel therapeutic strategies.

To my loving parents and grandparents, my infinite source of inspiration…

AKNOWLEDGMENTS

I would like to thank several people that supported me through my graduate school journey. First, I would like to thank my advisor and mentor, Dr. Daniel Slade, whose passion for science, expertise and most importantly, guidance was invaluable through the years. His support pushed me to be a better scientist everyday by sharpening my thinking and bringing my work to a higher level. I would also like to thank the members of my committee: Dr. Irving Coy Allen, Dr. Kylie Allen, and Dr. Jingsong Zhu. Thank you for your insightful suggestions and for your assistance throughout the years. Additionally, I would like to thank Dr. Pablo Sobrado who gave me the first opportunity to experience research and encouraged me to continue my scientific career at Virginia Tech.

I would like to thank my incredible family, my parents Jose and Rosario, my brother Gabriel, and my grandparents Tita and Chente. Their support and infinite encouragement to follow my Ph.D. dream has definitely made this possible. I would also like to thank my extended family: aunts, uncles and cousins for their words of encouragement and yearly gatherings when I went back home to visit.

I would also like to thank my partner in life and Grad School: Ricardo, your companion and solid support on my crazy ideas, like getting Tango, was essential throughout this process. We did it! I would like to thank all my friends throughout the years, specially Didier who was there every step of the way. Finally, I would also like to thank all the past and present members of The Slade Lab who accompanied me through the years and taught me so much about science and friendship, in particular Blake I couldn't have done it without you!

Table of contents

Chapter 1- Introduction to Fusobacterium: its interactions, virulence factors, role as a pathogen, and therapeutic opportunities 1

Abstract 2

Introduction 3

Fusobacterium interactions in the oral cavity 6

Type 5 secreted autotransporters and virulence protein families in *Fusobacterium* that drive host cell binding and invasion 9

The role of Lipopolysaccharide in infection and *Fusobacterium* survival in blood 19

Fusobacterium interactions with immune cells and *Fusobacterium* intracellular lifestyle 21

Challenges and Therapeutic Opportunities 23

Conclusions 27

References 29

Chapter 2- Utilization of DNA mmethyltransferases for enhaced bacterial genetics in *Fusobacterium* 40

Abstract 41

Introduction 42

Materials and Methods 49

Bacterial strains and plasmids 49

Identification of *Fusobacterium* DNA Methyltransferases 49

Phylogenetic Tree construction 56

Protein sequence similarity networks 56
Cloning of DNA Methyltransferases 57
Purification of Type II Methyltransferases 57
Purification of *F. nucleatum* ATCC 25586 Methyltransferases 58
In vitro treatment of plasmid DNA with *Fusobacterium* DNA Methyltransferases 59
Coexpression of plasmid DNA with *Fusobacterium* DNA Methyltransferases 59
Ultrapure DNA purification 59
Protection Assay 60
Linearization of plasmid DNA 60
Competent *Fusobacterium* and electroporation 61
Results 62
Phylogenetic and protein sequence similarity network analysis of the methyltransferases from six *Fusobacterium* species 63
Experimental validation of the recognition sequences for Type II methyltransferases in *F. nucleatum* ATCC 23726 68
Effect of plasmid methylation on the efficiency of *F. nucleatum* ATCC 23726 transformation and genome incorporation 72
Host-mimicking strategies in *Fusobacterium* strains 73
Linearizing DNA does not increase transformation efficiency in *F. nucleatum* ATCC 23726 76
Deletion of CATG recognition sequence in plasmid DNA increases transformation efficiency 76
Effect of plasmid methylation on an episomal plasmid 82

Effect of plasmid methylation on a Transposon vector 82

Discussion 84

Acknowledgments 90

Author contributions 90

References 91

Chapter 3 – CRISPR-Cas systems in Fusobacterium: An untapped genetic frontier? 95

Abstract 96

Introduction 97

Materials and Methods 101

Use of genomic information for bioinformatic analysis and identification of CRISPR-Cas systems in eight *Fusobacterium* strains. 101

Bioinformatic analysis of CRISPR-Cas Class 2 systems and identification of effector domain 101

Phylogenetic Tree construction 101

Results 103

Comparative Analysis of *Fusobacterium* CRISPR-Cas systems 103

CRISPR-Cas systems in *Fusobacterium* 106

Analysis of the *F. necrophorum* subsp. *funduliforme* 1_1_36S Type II CRISPR-Cas system and identification of putative repeats/spacers/PAM sequence 107

Identification of the *F. necrophorum* subsp. *funduliforme* 1_1_36S and *F. ulcerans* 49185 Type VI CRISPR-Cas system 113

Discussion 114

Acknowledgments 118

Author contributions 118

References 119

Chapter 4- Conclusions and Future Directions 123

References 133

List of Figures

Figure 1-1 *Fusobacterium* infections of humans and livestock.. 5

Figure 1-2 *Fusobacterium* acts as a bridging microorganism between early and late colonizing bacteria in the oral biofilm. 7

Figure 1-3 *Fusobacterium* encodes for an overabundance of T5SS 10

Figure 1-4 *Fusobacterium* promotes the acceleration of colorectal cancer. 12

Figure 1-5 *Fusobacterium* therapeutic alternatives 25

Figure 2-1 Restriction Modification (R-M) systems in this study are classified as Type I, Type II and Type III. 44

Figure 2-2 Host mimicking methylation in *Fusobacterium* to surpass R-M systems in bacteria. 47

Figure 2-3 R-M systems in a total of six *Fusobacterium* strains were identified**.** 64

Figure 2-4 Genome maps detailing R-M systems in six *Fusobacterium* strains. 65

Figure 2-5 Phylogenetic analysis of methyltransferases from four *Fusobacterium* strains..... 67

Figure 2-6 Sequence similarity networks of M.Fnu23I and M.Fnu23II 69

Figure 2-7 Restriction site identification of Type II methyltransferases in *F. nucleatum* 23726..70

Figure 2-8 Transformation efficiency in *F. nucleatum* ATCC 23726 improves when protecting plasmid DNA with methyltransferases.. 71

Figure 2-9 Increased transformation efficiency in *F. nucleatum* 23726 is accomplished by Type II methyltransferase conditioning of plasmid DNA..... 74

Figure 2-10 Transformation efficiency of *F. nucleatum* ATCC 23726 does not increase when linearizing plasmid DNA for homologous recombination before transformation 77

Figure 2-11 Transformation efficieny of *F. nucleatum* ATCC 23726 is influenced by the CATG motif. 78

Figure 2-12 Transformation efficiency of the episomal plasmid pHS30 in *F. nucleatum* ATCC 23726 increased when conditioned with Type II methyltransferases 79

Figure 2-13 Transformation efficiency of the episomal plasmid pHS30 in *F. nucleatum* ATCC 23726 and *F. nucleatum* 10953..... 80

Figure 2-14 Transformation efficiency of pBAMD1-4 in *F. nucleatum* ATCC 23726, *F. nucleatum* ATCC 25586 and *F. nucleatum* 10953 81

Figure 3-1 The CRISPR- Cas system mode of action is categorized in three phases 98

Figure 3-2 CRISPR-Cas systems in *Fusobacterium*.. 105

Figure 3-3 Diversity of Class 2 CRISPR-Cas systems in *Fusobacterium*. 108

Figure 3-4 Structure prediction and phylogenetic analysis of Cas9 from *F. necrophorum* subsp. *funduliforme* 1_1_36S..... 109

Figure 3-5 Characterization of CRISPR-Cas elements from *F. necrophorum* subsp. *funduliforme* 1_1_36S.. 110

Figure 3-6 Phylogenetic analysis of Cas13c from *F. necrophorum* subsp. *fundiliforme* 1_1_36S.. 111

Figure 4-1 Schematic of the utilization of R-M systems to overcome *Fusobacterium* recalcitrance**.** 126

Figure 4-2 The use of native CRISPR-Cas systems as genetic tools could lead to the development of novel genetic systems. 130

List of Tables

Table 2.1 Plasmids utilized in this study. 50

Table 2.2 Strains utilized in this study 51

Table2.3 Primers utilized in this study 52

Table 2.4 Methyltransferase genes identified in *Fusobacterium* strains utilized for host-mimicking techniques 66

Table 3.1 CRISPR-Cas system information of the eight analyzed *Fusobacterium* strains. 104

Chapter 1

Introduction to *Fusobacterium*: its interactions, virulence factors, role as a pathogen and therapeutic opportunities.

Ariana Umaña[1], Daniel J. Slade[1]

1. Virginia Polytechnic Institute and State University, Department of Biochemistry, Blacksburg, VA, USA.

Abstract

Fusobacterium are commensal Gram-negative, anaerobic bacteria that act as bridging microorganisms in the human oral cavity, but can also convert to opportunistic pathogens that contribute to oral and extra-oral infections and diseases. Many *Fusobacterium* strains are associated with distinct diseases distributed throughout the whole human body, including *Fusobacterium nucleatum*, which has been associated with the acceleration and metastasis of colorectal cancer. Modulation of infection is mainly coordinated by bacterial outer membrane proteins that interact with host cells and subsequently elicit adverse immune responses and dysregulation of key signaling pathways at the intersection of health and disease. This introductory chapter presents an exploration of the current state of the *Fusobacterium*-bacteria and *Fusobacterium*-host interactomes, by analyzing the regulation of host signaling pathways during infection and development of disease, while detailing current and future promising bacterial and host targets to combat *Fusobacterium.*

Introduction

Fusobacterium are Gram-negative anaerobic bacteria that colonize a variety of eukaryotes including fish, livestock, and humans [1]. Their cellular morphology varies greatly (coccoid, pleomorphic spherules (*F. necrophorum*) to rod shaped (majority) [2] and many have claimed that the diversity at the sub-species levels warrants reorganizing these bacteria into distinct species [3]. With this phenotypic heterogeneity the phylogenetic tree of the *Fusobacterium* genus is constantly evolving and receiving more complex classifications.

The first studies on *Fusobacterium* divided the genus into 2 main groups based on biochemical and serological properties (Leptotrichiaceae and Fusobacteriaceae) [4]. Currently, the *Fusobacterium* genus includes 13 species (*F. nucleatum*, *F. necrogenes*, *F. necrophorum*, *F. periodonticum*, *F. varium*, *F. ulcerans*, *F. mortiferum*, *F. gonidiaformans*, *F. prausnitzii*, *F equinum*, *F. simiae*, *F. canifelinum* and *F. naviforme*) and conventionally comprises a variety of gram-negative bacilli that either have pointed or fusiform ends [5]. As a genus, *Fusobacterium* is indole positive, encapsulated and abscessogenic as well as sensitive to both kanamycin and colistin and resistant to vancomycin [2,5]. *Fusobacterium* can be differentiated by its bile sensitivity and metabolism of threonine to propionate as well as generating significant amounts of butyric acid at the end of their metabolism [2,5].

F. nucleatum, the most recurrently isolated species in humans, is diverse and presently it comprises five subspecies: subsp. *fusiforme*, subsp. *nucleatum*, subsp. *vincentii*, subsp. *animalis*, and subsp. *polymorphum* [1,5]. *F. necrophorum*, regularly isolated in veterinary medicine is a major pathogen and two subspecies have been characterized; *F. necrophorum* subsp. *necrophorum* and *F. necrophorum* subsp. *funduliforme* [5,6]. Additionally, a novel species, *F. equinum*, that is

closely associated to *F. necrophorum*, has been isolated from horse oral microbiota [5,7]. Another species, isolated from tropical ulcers, *F. ulcerans* is phenotypically similar to gastrointestinal species *F. mortiferum* and *F. varium* [1,5]. Interestingly, bacteria isolated from the human gingival sulcus that are phylogenetically closer to *Eubacterium* were previously known as *F. sulci* [5,8]. In the same way *F. prausnitzii*, which possesses a higher GC% content (47-57) when compared to other *Fusobacterium* (25% GC) is set to be reclassified as *Eubacterium* [5].

Fusobacterium species are implicated in a wide range of human and animal diseases as shown in **Figure1.1**. Lemierre's syndrome, otitis media, bacteremia and sinusitis infections are related to *F. necrophorum* subsp. *funduliforme* and *F. necrophorum* subsp. *necrophorum* that are frequently isolated in animal infections [9]. *F. necrophorum* subsp. *necrophorum* causes cattle and sheep liver abscesses [10], ruminant foot abscesses [10], and sporadic abortions [11]. Meanwhile, *F. nucleatum* is associated with oral infections (periodontitis [12,13] and gingivitis [12–15]), adverse pregnancy outcomes (preterm birth [16,17] and chorioamnionitis [18]), gastrointestinal disorders (colorectal cancer [19–21] and inflammatory bowel disease [22,23]) and other infections (osteomyelitis [24], liver abscess [25], and bacteremia [26]).

This introductory chapter delineates the interactions of *Fusobacterium* in the oral biofilm and its role in the development and acceleration of diseases, with a focus in colorectal cancer; specific topics covered include hematogenous spread through the body, intracellular survival, virulence protein families, interactions with immune cells, and exploring potential therapeutic challenges and opportunities.

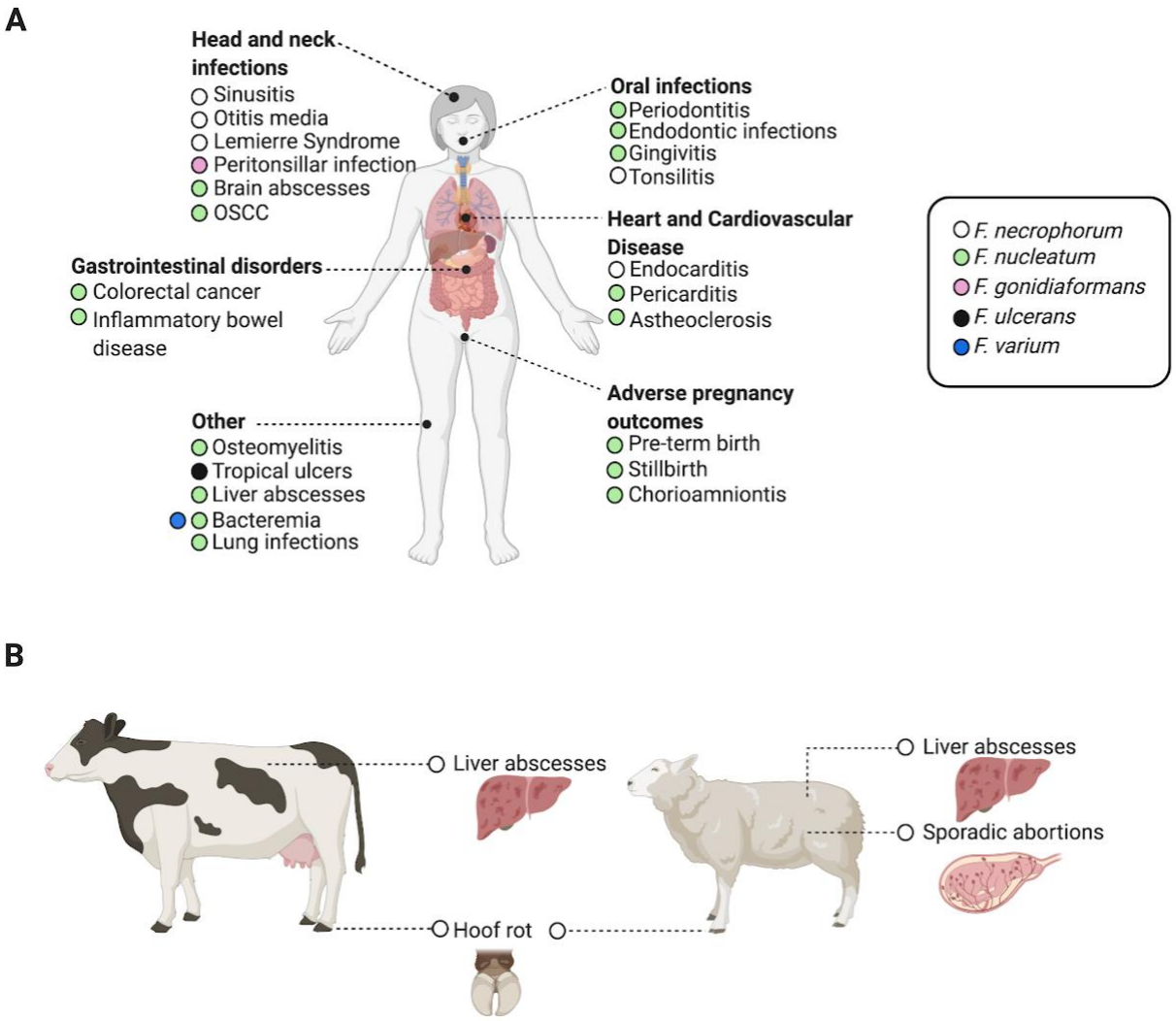


Figure 1-1 *Fusobacterium* infections of humans and livestock. **A.** *Fusobacterium* is a resident of the human oral microbiome. *Fusobacterium* can travel throughout the body through the bloodstream where different species are involved in a wide array of diseases, including chronic inflammatory bowel disease and colorectal cancer. *F. necrophorum* is also a component of the human oral microbiome and can cause fatal infections of the jugular vein. **B.** *F. necrophorum* causes severe hoof rot and lameness in cattle, costing millions of dollars per year to the livestock industry.

Fusobacterium interactions in the oral cavity

Fusobacterium coordinates the central architecture of dental plaque by bridging early and late colonizing bacteria in oral plaque, but frequently migrates from this environment to cause oral and extra-oral infections [27,28]. The composition of the oral biofilm is mediated by the physical interactions of specific bacterial pairs, in a phenomenon known as coaggregation [29]. As a bridging microorganism, *F. nucleatum*, initially coaggregates to primary colonizers such as *Streptococcus* and links them to anaerobic secondary colonizers such as *Aggregatibacter actinomycetemcomitans* and *Porphyromonas gingivalis* [28]**.** *Fusobacterium* interactions in the oral cavity are detailed in **Figure 1.2.**

F. nucleatum also facilitates the ability of obligate anaerobes to survive in aerated environments; a key role of *F. nucleatum* in the oral biofilm structure [29]. Furthermore, *F. nucleatum* can regulate the formation of the oral biofilm and subsequent growth [27]. The biofilm interactions of *Fusobacterium* are mediated by an abundance of outer membrane galactose or arginine inhibitable adhesins that have been well characterized for their bacterial interactions [30]. The Fibroblast activation protein 2 (Fap2) protein which is secreted through the inner membrane and present in the outer membrane and therefore on the surface of *F. nucleatum*, has been identified as a galactose-binding lectin that mediates fusobacterial hemagglutination (which is a particular form of agglutination that involves red blood cells) [31]. In a similar manner to hemagglutination, coaggregation of *F. nucleatum* ATCC 23726 with *P. gingivalis* is also enabled by Fap2 [31,32] and this process is a key factor in periodontitis [31,33]. *F. nucleatum* participates as a mutualist in the oral biofilm by bridging bacteria through its autotransporter protein RadD adhesin [34]. RadD, which is an arginine-inhibitable adhesin, is capable of docking with and aggregating multiple species, assists in the oral biofilm development [34]. RadD is utilized by *Fusobacterium* species

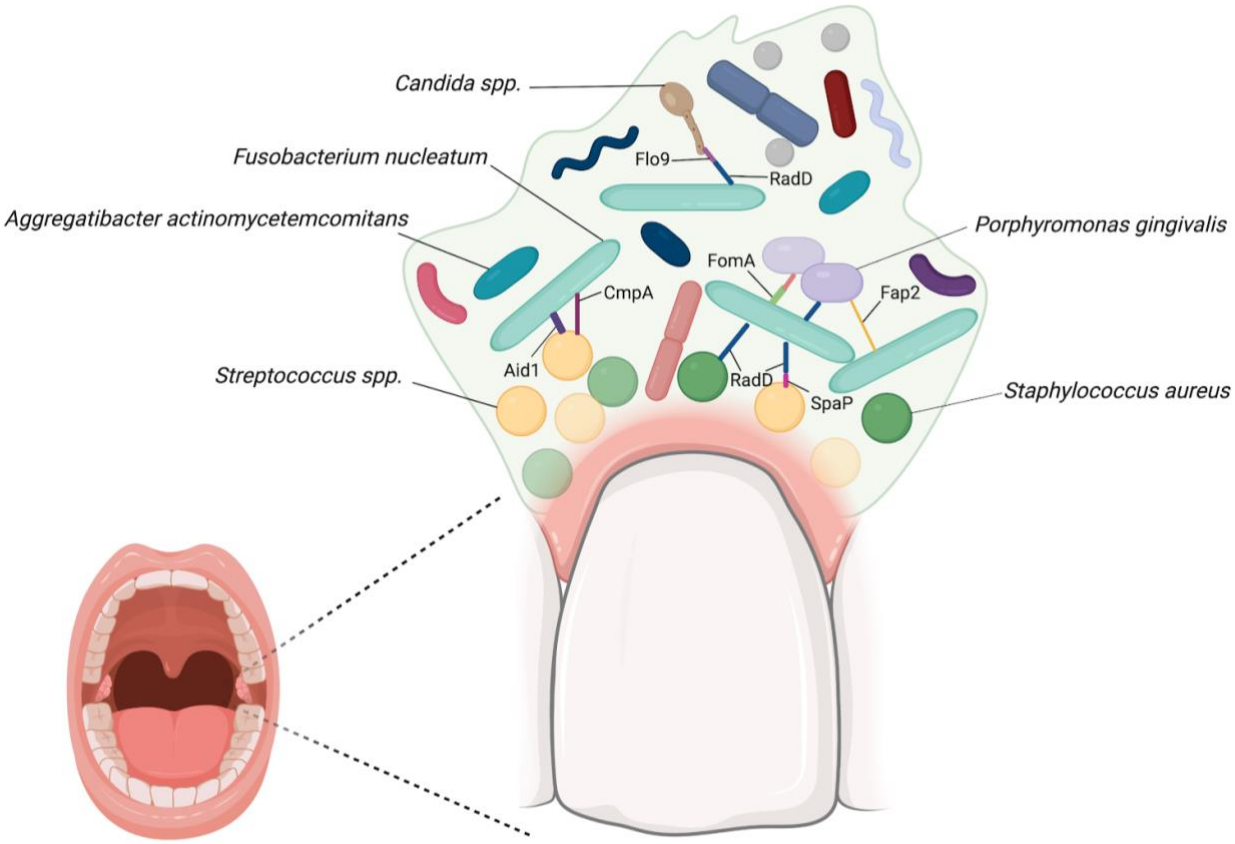


Figure 1-2 *Fusobacterium* acts as a bridging microorganism between early and late colonizing bacteria in the oral biofilm. *F. nucleatum* attaches to primary colonizers *Streptococcus* and *Staphylococcus* aureus through the RadD, Aid1 and CmpA adhesins. RadD interacts closely with SpaP from *Streptococcus*. *Fusobacterium* also coaggregates with anaerobic secondary colonizers such as *Aggregatibacter actinomycetemcomitans*, *Candida spp*. and *Porphyromonas gingivalis* through protein-protein interactions including Flo9 from *Candida spp.* and RaD, FomA as well as Fap2 from *F. nucleatum*.

to physically attach to a number of bacterial species including *P. gingivalis*, *S. sanguinis*, *S. gordonii*, *S. mutans* [30][32], *S. aureus* [35], and the opportunistic pathogenic yeast *Candida albicans* (**Figure 1.2)** [36,37]. More specifically, coaggregation of *S. mutans* with *F. nucleatum ssp. polymorphum* is mediated by a protein-protein interaction; RadD and a major *S. mutans* adhesin SpaP **(Figure 1.2)** [38]. Additionally, RadD is partially involved in aggregate formation during *F. nucleatum-S. aureus* interactions, thereby contributing to biofilm development and leading to upregulation of the *S. aureus* protein, sarA [35].

Much like other streptococcal species interaction, *Fusobacterium* interacts and co-aggregates with yeast, *Candida albicans* through protein-protein interactions via RadD and cell wall protein, Flo9 **(Figure 1.2)** [37].When interacting with *C. albicans*, *Fusobacterium* is capable of inhibiting growth and hyphal morphogenesis via the RadD-Flo9 protein duo [37]. This attenuation of virulence may be essential to maintain balance within the host oral microbiome [37]. As a Gram-negative pathogen, *F. nucleatum* is capable of inducing MCP-1 and TNFα production through LPS, while *C. albicans* seems to diminish this ability as seen in Bor and colleagues studies using a murine macrophage cell line [36]. Thus, along with serving as a bridging organism to bring together bacterial species on the oral surface, RadD also functions along with Flo9 to mutually inhibit virulence by *C. albicans*, promoting long term fitness and commensal survival within the host [36].

Fusobacterium interactions with *Pseudomonas gingivalis* and *Streptococcus sanguis* are a mediated by FomA, a porin protein [40][41] which also binds to the human salivary proteins through a conserved statherin amino-acid region [42]. Moreover, a highly conserved and specific protein across all fusobacterial species, Aid1 another autotransporter, is necessary for recognition and discrimination of Gram-positive species in the oral cavity [43]. The gene for *aid1* was

disrupted and it resulted in a diminished interaction with *Streptococci* and *E. faecalis*. By constrast, overexpression of Aid1 displayed an increased aggregation with *Streptococci* [43]. The presence of multiple adhesins and binding partners in *Fusobacterium* interspecies interaction suggests more complex mechanisms that are yet to be detailed [43].

Specific interactions involved in biofilm integration and binding of bacterial species are complex and involve numerous key players. This complex set of interactions is in part enabled by coaggregation mediating protein A (CmpA) as it was shown that is necessary for coaggregation with *S. gordonii* V288 [39]. CmpA shows a higher coaggregation potential in the late exponential phase of *F. nucleatum* [39]; thus, a combined role for CmpA and RadD was suggested where RadD enables binding in the initial stages and in the latter stages of biofilm formation CmpA takes over [39].

Type 5 secreted autotransporters and virulence protein families in *Fusobacterium* that drive host cell binding and invasion

Bacteria in general secrete a diverse set of proteins on their cell surface that are responsible for signaling modulation, nutrient transport, adhesion and/or virulence [44]. The proteins that coordinate secretion in Gram-negative bacteria can be categorized, from Type 1 through Type 9 secretion systems [45]. These dynamic secretion systems possess distinct structural and mechanistic features that are highly characteristic [45][44]. *Fusobacterium* encodes for an overabundance of Type 5 Secretion Systems (T5SS) and a percentage of them have been shown to be involved in binding [30,31], host cell invasion [46,47] and pathogenesis [48]. The T5SS proteins or autotransporters (AT), are secreted autonomously [49] and are divided into subclasses

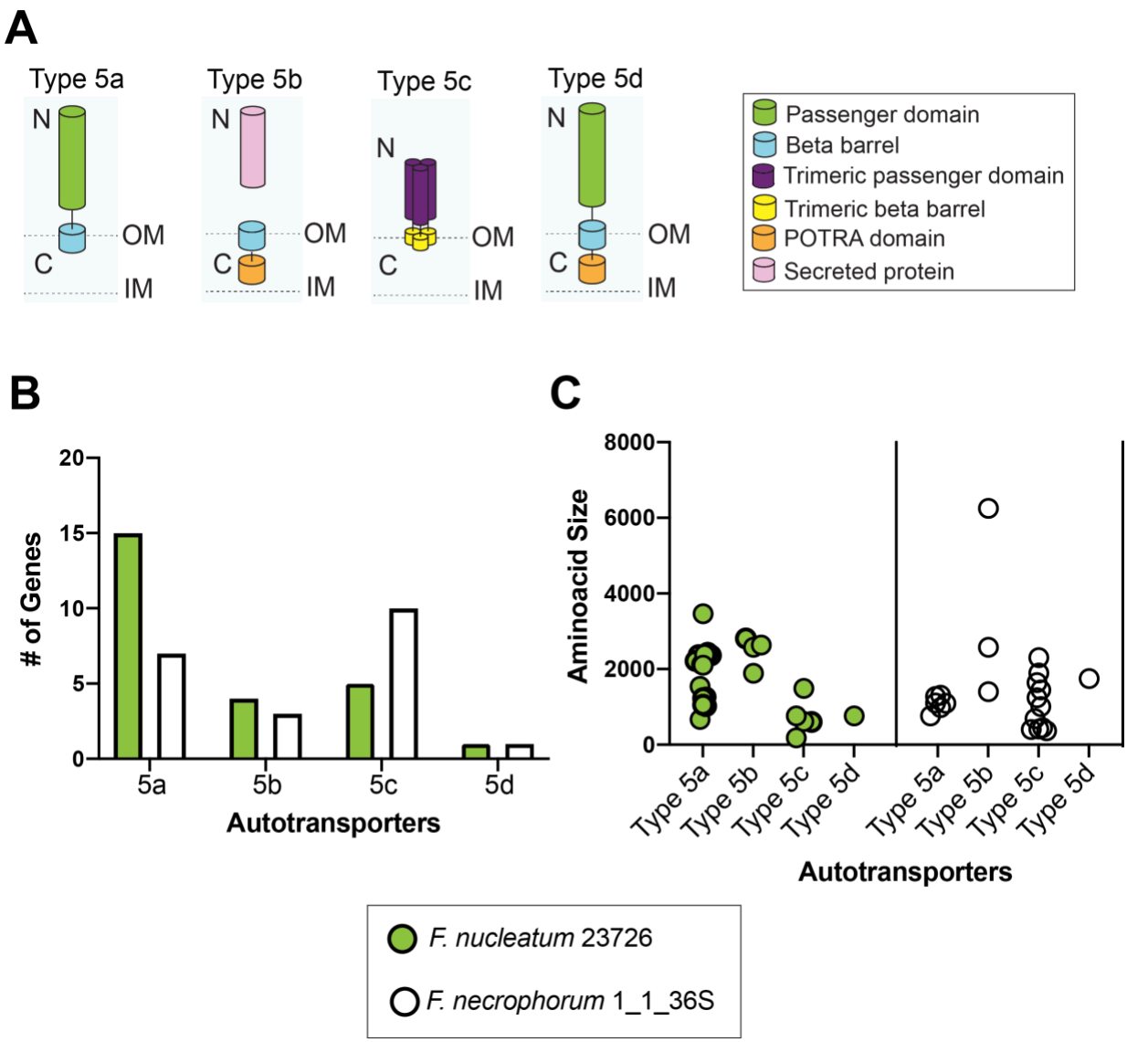


Figure 1-3 Fusobacterium encodes for an overabundance of T5SS A. Schematic of Type 5SS encoded in *Fusobacterium*. **B.** Number of genes encoding for T5SS in *F. nucleatum* 23726 and *F. necrophorum* 1_1_36S. **C.** Aminoacid size of T5SS proteins belonging to Type 5a, 5b, 5c and 5d in *F. nucleatum* 23726 and *F. necrophorum* 1_1_36S. Adapted from Umaña, 2019

(5a, 5b, 5c, 5d and 5e) depending on their structural features and domain organization **(Figure 1.3)** [44]. AT are composed of secreted passenger domains that function in diverse ways (adhesin, enzymes, toxins) and a β-barrel domain that resides in the bacterial outer membrane [44,50]. The T5SS utilizes the SecYEG complex to transport the unfolded autotransporter through the inner membrane to the periplasm which leads to the β-barrel domain to the outer cell membrane, permitting transport of the secreted protein [45]. While 100% of *Fusobacterium* strains encode for T5aSS [51] their role in connection with the variety of diseases associated with this bacterium remains understudied. A detailed description of the contribution of these proteins to colorectal cancer promotion is found in **Figure 1.4.** In this next section, the role of the characterized T5SS and other virulence factors in *Fusobacterium* is reviewed.

Type 5a autotransporters

Fap2 is one of the most widely described Type 5a autotransporter in *F. nucleatum* [31,46,47,52,53]. The ability of *Fusobacterium* to colonize the host placenta is a crucial process for preterm birth and it is facilitated by Fap2's ability to act as an adhesin [31]. In colorectal cancer promotion, Fap2 plays a key role mediating colorectal cancer acceleration through interaction and attachment with Gal-GalNAc, a sugar moieties that is overrepresented in human cancer cells this consequently promotes the internalization of *F. nucleatum* by human epithelial cells and leads to the alteration of key signaling pathways in cancer [46].

An important *Fusobacterium* pathogenicity trait is its capacity to induce the cell death of Jurkat cells (T- lymphocyte cells) via the autotransporters, Fap2 and RadD [47]. The specific mechanism is yet to be described, however it has been suggested that contact-based interaction might be necessary to promote cell death [47]. Additionally, Fap2 was shown to bind to the natural

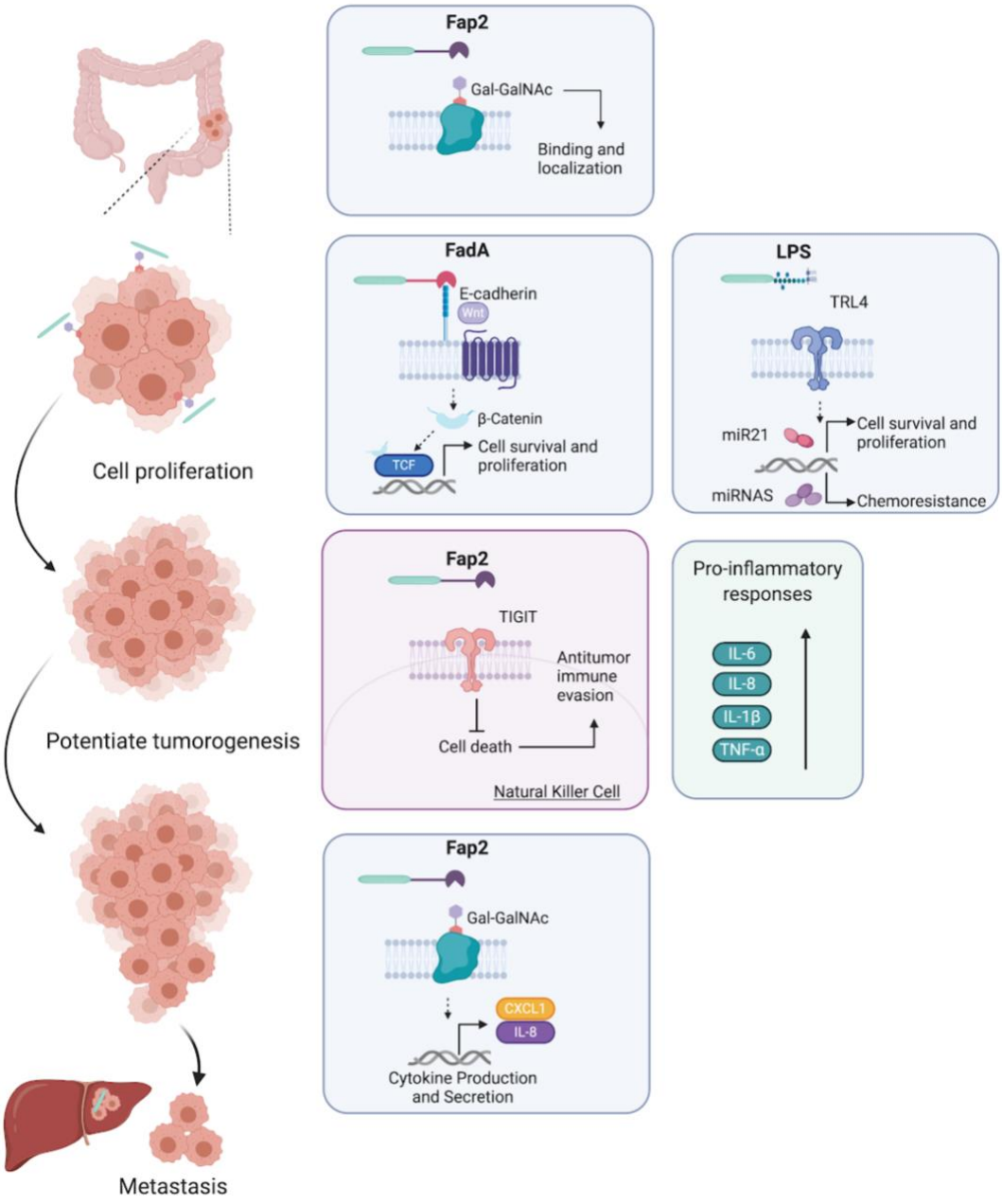


Figure 1-4 *Fusobacterium* promotes the acceleration of colorectal cancer. Fap2 plays a key role mediating colorectal cancer acceleration through interaction and attachment with Gal-GalNAc sugar moieties. *F. nucleatum* contributes to cell proliferation by modulating E-cadherin via FadA which leads to β-catenin signaling activation consequently triggering transcription of Wnt target genes. Additionally, *F. nucleatum LPS* interacts through the LPS with Toll-like receptor 4 (TLR4) which modulates microRNA21 (miR21) expression via the TLR4/MYD88/NFkB pathway in colorectal cancer. The autotransporter Fap2 binds to the natural killer cell receptor TIGIT and promotes immune evasion and metastasis by cytokine secretion.

killer cell receptor TIGIT and effectively block the ability of this immune cell subtype to clear tumor cells in a mouse model [53]. More recently, it has been revealed that Fap2 is required for colorectal cancer cell monolayer (HCT-116) invasion [54] and the infected cells secrete the cytokines CXCL1 and interleukin (IL-8) in a Fap2 dependent manner, promoting the motility of the cells and potentially leading to metastasis **(Figure 1.4)** [54].

Several OMP have been identified as Fap2 homologues for future studies [55,56]. One of them has been characterized as an Apoptosis-inducing Outer Membrane Protein 1(Aim1) [55], it was described through the creation of a deficient mutant strain and it was demonstrated that the Fap2 homolog is capable of inducing apoptosis in Jurkat cells (T- lymphocyte cells) and that most likely several proteins are involved in this process [55].

Autotransporters also act as bacterial serine proteases. Initially their roles in pathogenesis were described in *Neisseria, Shigella* and *Escherichia coli* [57]. These bacterial serine proteases catalyze the hydrolysis of host intracellular and extracellular substrates causing breaks in the cytoskeleton structure, inducing autophagy or impairing innate and adaptive immunity [57,58]. Previous studies have described a Type 5a serine protease autotransporter, Fusolisin, capable of degrading IgA along with host extracellular matrix proteins, including fibronectin and collagen [59]. Additionally, Fusolisin degrades fibrinogen at the C-terminus portion of small aminoacids (Gly, Thr, Ser and Ala) providing essential nutrients for the growth of *F. nucleatum* [58]. However, the mechanism of virulence and potential roles in pathogenesis of Fusolisin remain to be characterized.

Type 5b secreted effectors: Two-partner secretion of toxins and uncharacterized effectors

Most of *Fusobacterium* Type 5b secreted autotransporters remain uncharacterized, except for the leukotoxin (LktA), part of the T5bSS toxin family described in *F. necrophorum* which are secreted by the two-partner secretion pathway. *F. necrophorum* is associated with infections that lead to necrotic diseases in animals [60] and is the originator pathogen of Lemierre's syndrome in humans [9]. Previous research has shown a strong correlation between the presence of the leukotoxin operon which is expressed during the late-log growth phase in *F. necrophorum* and its pathogenesis [10,60–64].

F. necrophorum's leukotoxin (LktA) is distinctive as it doesn't share sequence sequence similarity with any other bacterial leukotoxin and it is uncommonly larger (~300 kDa) in contrast to other bacterial leukotoxins which are usually less than 200 kDa. In addition, the leukotoxin lacks cysteine residues and appears to be highly specific to *F. necrophorum* [63][65]. The leukotoxin operon is encoded by three genes (*lktBAC*), LktB potentially acts as a secretion partner by exporting LktA which corresponds to the leukotoxin protein [66]; whereas LktC has no known function as its amino acid composition is unlike any known determinant [65].

The leukotoxin from *F. necrophorum* is highly toxic to bovine leukocytes as it modulates apoptosis and induces necrosis in a dose-dependent manner [66][48] [60–62]. Although, it displays moderate toxicity to horse polymorphonuclear leukocytes (PMNs) and no toxicity towards swine and rabbit PMNs [67]. To neutralize the effect of the leutkotin, monoclonal antibodies have been produced, however initially they only partially neutralized this protein [68]. More recently, a complete neutralizing effect was achieved by using monoclonal antibodies created from purified leukotoxin from *E. coli*. [63].

Interestingly, whereas *F. necrophorum* subsp. *necrophorum* causes severe infections, *F. necrophorum* subsp. *funduliforme* is less pathogenic; mostly due to leukotoxin activity levels based on differential expression and distinct amino acid sequences [66]. Intriguingly enough, Holm et al determined that all the genomes from human isolates of *F. necrophorum* subsp. *funduliforme* available in GenBank encode the lkt operon[69] and recently, a study demonstrated that 10 strains of *F. equinum* secrete a leukotoxin with toxic effects in equine leukocytes [7]. Nonetheless, only the existence of the *lktA* gene was confirmed and the attempt to identify the cluster was inconclusive [7].

The role of other T5bSS effectors in *Fusobacterium* pathogenesis is yet to be determined. However, it has been shown in other Gram-negative bacteria that these secreted effectors encode for hemagglutinin, adhesin, hemolysin, cytolysin domain [56] and that the main function of the passengers domain is contact-dependent growth inhibition (CDI) [44]. Thus, suggesting involvement of these effectors in colonization, survival and bacterial competition in establishing a niche [56].

Type 5c trimeric autotransporters

A model where *Fusobacterium* interacts and maintains an association with host tissues through trimeric autotransporter adhesins (TAAs) has been proposed [56], as trimeric autotransporters in other Gram-negative bacteria have crucial roles in virulence [70–72]. Recently, the first experimental characterization of multiple Type 5c autotransporters (CbpF, FvcB, FvcC, FvcD, and FvcE) in *F. nucleatum* determined that they potentially contribute to binding and invasion [54]. A previous study, also demonstrated that CbpF, a TAA specific to *F. nucleatum* and *F. vincentii*, interacts with CEACAM1 in gingival cells which results in colonization and invasion.

CbpF contains the CEACAM-binding motif and it possesses high similarity with other pathogenic bacterial binding ligands to an attachment region on the N-domain of CEACAM1 [73].

Type 5d outer membrane lipases

A conserved Type 5d autotransporter in *F. nucleatum* ATCC 25586, was characterized as a phospholipase class A with affinity to for phosphatidylino-sitol 3,5-bisphosphate [PI(3,5) P2], phosphatidyli-nositol 4,5-bisphosphate [PI(4,5)P2], and phosphatidy-linositol 3,4,5-trisphosphate [PI(3,4,5)P3] lipids[74]. Therefore, *Fusobacterium* phospholipase autotransporter (FplA) has high binding preference for host phosphoinositide-signaling lipids, suggesting a potential role in establishment of an intracellular niche. However, a role for FplA in *F. nucleatum* pathogenesis has not been fully characterized[74].

Outer membrane vesicles in Fusobacterium

It has been shown that Gram-negative bacteria produce outer membrane vesicles (OMVs), which deliver virulence factors have the capability to interact with host epithelial cells, because of the presence of adhesion molecules and surface proteins. More recently, computational analysis and experimental studies have confirmed autotransporters, FadA and Membrane Ocuppation and Recognition Nexus 2 (MORN2) domain proteins in OMV [75][76][77].

The OMV from *F. nucleatum* include proteases capable of degrading E-cadherin, which promotes bacterial binding, invasion and modulation of inflammatory responses [58]. OMVs from *F. nucleatum* are also capable of inducing epithelial–mesenchymal transition (EMT), a signaling pathway for development of secondary tumors or metastasis [78]. Interestingly, the OMV in *Fusobacterium* upregulate a list of EMT markers including Snail, integrin-α5, TWIST, bone morphogenetic proteins (BMPs), N-cadherin, ZEB-1, and fibronectin-1[78]. These findings could

be transcendental in the development of a vaccine against *F. nucleatum* since OMV have emerged as powerful vaccine delivery systems [77].

While OMP and OMV have been studied in *F. nucleatum* to some extent the OMP and OMV profile of *F. necrophorum* and their contribution to the initiation and establishment of infection has not been characterized as well. The difference between the OMP profiles within the subspecies could hint to their distinct traits in pathogenicity and preferred environmental niche [79].

FadA family adhesins

Fusobacterium adhesin A (FadA) is highly conserved in oral *Fusobacterium* species but is not present in non-oral species [80]. FadA is a small protein that consists of 129 aminoacids, including an 18 amino acid signal peptide, which makes the form secreted outside of *F. nucleatum*, is 111 amino acids with a calculated 12.6 kDa mass (mature FadA) [80–82]. This protein multimerizes and the initial role of FadA in pathogenesis was characterized when a deletion mutant of FadA prevented binding and colonization to human gingival epithelial cells, human umbilical vein endothelial cells and mice placentas [80,81,83,84].

A constituent of the large cadherin family of surface proteins and a cell-cell junction protein, vascular endothelial V-cadherin was characterized as a receptor for FadA. This receptor was found to be crucial for invasion of endothelial cells [84,85] and for *F. nucleatum* systemic dissemination. Additionally, *F. nucleatum* potentially employs the interaction with this receptor to overcome the placental barrier and stimulate preterm birth [86].

An additional receptor of FadA in epithelial cells is the cell-adhesion protein E-cadherin exposed on the surface on non-colorectal and colorectal cells [87]. *F. nucleatum* modulates E-cadherin via FadA interaction leading to β-catenin signaling activation, consequently triggering

the transcription of Wnt target genes that induce oncogenic and inflammatory responses in colorectal cancer **(Figure 1.4)** [87]. More recently, the identification of FadA2 and FadA3 which are two homologues of FadA was described, as well as the presence of several identical gene copies of the FadA3 (FadA3a, FadA3b, FadA3c) in *F. nucleatum* ATCC 23726 and *F. nucleatum* ATCC 25586 [56]. Further studies are necessary to describe the invasive potential and function of these paralogs.

Another member of the FadA family is FadI, which is capable of inducing the expression of human beta defensin (HBD) by interacting with TLR-1/2 and TLR-2/6 in human epithelial oral cells and it has been shown to induce the highest amount of this antimicrobial peptide compared to other FadA family adhesins [88]. However, differential expression of HBD induced by FadI was shown to be *F. nucleatum* strain dependent and, it has been suggested that this could be due to the variability of the signal peptide and the efficiency to anchor in the outer membrane [89].

FomA family adhesins

Early reports in the late 80's described a 40-42 kDa protein with strong immunobiological activities in a number of in vitro assays [90]**.** Several studies later confirmed that the major outer-membrane protein, FomA, has a role in porin activity [90,91], interbacterial adherence [90,92,93] and stimulation of host immune cells [90,92,93]. Furthermore, FomA porins of *F. nucleatum* strain 12230 also display the ability to bind to Fc fragment of human immunoglobulin G [94], which enables the attachment of antigen-antibody complexes to phagocytic cells or the complement system. Fc-adhesins from other bacteria have been associated with the development of periodontal lesions by modulation of phagocytic activity, downregulation of B-cell proliferation, and complement function [95]. Therefore, FomA could act as a virulence factor in *F. nucleatum* by binding Fc receptors to halt and evade host immune system.

FomA from *F. nucleatum* has immune adjuvant activity in mice, [90] and it was verified that FomA induces TLR2 signaling resulting in IL-8 secretion and NF-kB activation [96]. In *F. necrophorum*, FomA plays a pivotal role in the attachment of bacteria to bovine adrenal gland capillary endothelial cells [97]. More importantly, Lui and colleagues highlighted the potential of FomA as a target for the development of novel treatments to combat periodontal infection and halitosis in humans by confirming the suppression of co-infection and gum swelling through a protein (FomA) targeting vaccine [41].

MORN2 domain proteins

Bioinformatically predicted *Fusobacterium* active invader strains were previously found to encode for an abundant set of proteins termed MORN2 domains [3]. This enrichment of MORN2 domain proteins remains uniquely specific to *Fusobacterium*, except for a few MORN2 genes *Helicobacter bilis* which is involved in several diseases such as colitis, hepatitis and abortion in sheep [98]. Many of the MORN2 domain genes in *Fusobacterium* also colocalize with known virulence factor genes including RadD and FadA genes associated with adhesion and virulence [3].

Recently, the majority of MORN2 proteins were predicted to be extracellular [3,56] and have been found to be in *F. nucleatum* secreted outer membrane vesicles [77]. In addition, they were characterized as part of the predicted 'FusoSecretome' [76]. A cooperative role for the MORN2 proteins along with autotransporters and the FadA family was proposed in host cell docking and invasion [56]. However, if and how this family of proteins impacts the interactions of fusobacteria with host cells and other microorganisms remains uncharacterized [99,100].

The role of Lipopolysaccharide in infection and *Fusobacterium* survival in blood

Lipopolysaccharide (LPS) in Gram-negative bacteria is a considerable constituent of the outer membrane that modulates the innate immune systems by interacting directly with the host receptors [101,102]. Interestingly, LPS from *Fusobacterium* resembles those from aerobic Gram-negative bacteria like *Salmonella* because they have a lipid and a polysaccharide component [103].

A striking detail about *Fusobacterium* LPS composition is the contrast between species, as the principal fatty acids in *F. nucleatum* LPS were identified as β-hydroxymyristic, myristic, and β-hydroxypalmitic acids whereas *F. necrophorum* LPS did not contain β-hydroxypalmitic acid [105]. In addition, while the main sugars in *F. nucleatum* LPS were heptose and galactose, glucose, glucosamine, and rhamnose, glucose and heptose are more common in *F. necrophorum* LPS [104][105].

Several studies have confirmed that *F. nucleatum* interacts through the LPS with Toll-like receptor 4 (TLR4) [106][107,108] and it has been suggested that *F. nucleatum* modulates microRNA21 (miR21) expression via the TLR4/MYD88/NFkB pathway in colorectal cancer **(Figure 1.4)** [109]. The pathogenic role of miR21 in chronic inflammatory diseases and colitis-associated colon cancer has been described [110] and it is widely known that particular miRNAs are associated with colorectal carcinogenesis. Furthermore, miRNAs can be utilized as biomarkers for detection, prognosis, and metastasis prediction in colorectal cancer patients [110–113].

LPS engages the innate immune system TLR4 and MYD88 leading to the triggering of the autophagy pathway and mediation of colorectal cancer chemotherapeutic response by *F. nucleatum* [114][115]. In macrophages, *F. nucleatum* is capable of inducing several proinflammatory cytokines, including TNF-α and IL-1β, by LPS mediation [116][106] and in mice TRL4 activation by LPS resulted in induction of fetal death, indicating the relevance of bacterial LPS in the pathogenesis in intrauterine infection [117].

It is now well accepted that the major route of travel from the native oral cavity for *Fusobacterium* is through the bloodstream; but major questions remain in how a non-motile bacterium can traverse multiple cellular layers to arrive within the blood. It has been suggested that FadA stimulates a signaling cascade that disrupts the junctions in the blood vessels endothelial cells, which permits the entry of *F. nucleatum* and other bacteria into the blood, such as *E. coli* [85].

Additionally, *F. nucleatum* has also been shown to induce neutrophils to expel their cellular components in the form of neutrophil extracellular traps in a phenomenon known as NETosis [119]. This process also renders neutrophils dead, but is critical for controlling the spread of microbial infections [119]. Previous studies have shown that neutrophils are activated and induced through the upregulation by *F. nucleatum* of the receptors: Nucleotide oligomerization domain 1 and 2 (NOD1) (NOD2) [119].

For *F. necrophorum* it has been suggested that LktA is key for bloodstream survival [48] and multiple complement binding proteins have been characterized [118], which enhance bacterial survival in human serum. Importantly, it has been proven that transient bacteremia permits the entry of oral fusobacteria to the circulatory system [46], and Lemierre syndrome caused by *F. necrophorum* is distinguished by antecedent oropharyngeal infection, septic emboli, and the presence of bacteria on the blood or bacteremia [26].

Fusobacterium **interactions with immune cells and** ***Fusobacterium*** **intracellular lifestyle**

Early studies demonstrated that *F. nucleatum* is phagocytized by PMNs and that adherence of *F. nucleatum* resulted in activation of PMNs as well as inhibition of neutrophil chemotaxis [120][121]. More recently, studies suggested that *F. nucleatum* secreted lectin-like proteins

mediate binding or adherence of neutrophils and these interactions can be block by D-galactose, lactose, and N-acetyl-D-galactosamine (GalNAc) [120,122]. However, the exact mechanism by which *F. nucleatum* and PMNs interact is still to be characterized.

F. nucleatum mediates changes in neutrophils transcription [123], it stimulates secretion of proinflammatory cytokines including interleukin IL-1, IL-6, IL-8 as well as tumor necrosis factor a (TNF-a) and reactive oxygen species (ROS) [124–126]. A common feature of periodontal disease as well as gingivitis is the accumulation of neutrophils and tissue destruction caused by reactive ROS [122,124]. Furthermore, several *F. nucleatum* strains have demonstrated variations in superoxide production and in the interaction with TLR2, TLR4 and NF-KB, resulting in different levels of expression of proinflammatory cytokines. Altogether, these effects could be necessary for the onset of periodontitis [127].

F. nucleatum has been coined a facultative intracellular bacterium, yet the mechanisms by which these bacteria invade and survives inside of cells remain an enigma. However, it is widely known that *F. nucleatum* is highly invasive into epithelial and endothelial cells where the invasion appears to occur via a "zipper" mechanism that requires protein synthesis, actins, signal transduction, microtubules, and energy metabolism of the host cell, as well as protein synthesis by *F. nucleatum* [33]. Evidence has shown that *F. nucleatum* is internalized within vesicular structures and once inside of the host cell it colocalizes with endosomes and lysosomes [128].

Interestingly, anaerobic bacterium *F. nucleatum* has the ability to survive and proliferate in oxygenated macrophage cells up to 72 hours and inhibits macrophage apoptosis by modulating the PI3K and ERK pathways [116]. Once Fusobacteria (*F. necrophorum*, *F. mortiferum* along with *F. nucleatum)* invade the cells, they are able to survive and multiply for several hours under aerobic conditions [130], potentially due to *Fusobacterium* oxidative stress response [131]. Nonetheless,

the intracellular number of bacteria rapidly declines after this time (6-8 hours). *F. mortiferum* demonstrated an early capacity to replicate intracellularly, while *F. nucleatum* subsp. *nucleatum* increased its intracellular number after 7 hours incubation. These distinct replication abilities at certain infection timepoints may be significant to the different pathogenic traits of *Fusobacterium* species and strains. Longer periods of survival and replication potentially provide bacteria with enough time to escape from host defense mechanisms, permitting growth in other locations [130].

Despite the fact that *Fusobacterium* are non-motile anaerobes, studies confirm their ability to attach, invade and replicate in epithelial cells, possibly through the reorganization of the host's cytoskeleton. Fusobacteria replication was inhibited by cytocalacin D, which suggests that *Fusobacterium* re-arranges actin filaments potentially creating a barrier-like structure around its invasion vacuole and eventually leading to the release of bacteria into the host cell cytoplasm [130].

Challenges and Therapeutic Opportunities

Combating *Fusobacterium* through novel therapeutic treatments warrants potential responses against periodontitis, breast cancer, esophageal cancer, pancreatic cancer, colorectal cancer, preterm birth and cattle related diseases. Several approaches have been made to tackle these bacteria as described in **Figure 1.5**. The development of alternative treatments in oral infections is necessary as conventional antibacterial therapy is not effective due to the high resistance and resilience of bacterial biofilms [132]. Whereas in colorectal cancer, ongoing therapeutic treatments include removal by surgery, sessions of chemotherapy and radiation therapy, which may not be completely successful. This warrants an approach that specifically targets *Fusobacterium* in biofilms that could lead to a novel therapy to eradicate these diseases [133].

Out of these approaches, a ground-breaking approach recently has been the utilization of bacteriophages to disrupt biofilms in periodontal infection and colon cancer [134]. Bacteriophages are viruses that attack prokaryotes and several of them have been isolated specifically for *Fusobacterium* [132,134]. The first phage discovered to infect *Fusobacterium* was FnpΦ02, a siphovirus that targets three subspecies of *F. nucleatum (nucleatum, vicentii* and *polymorphum)* [134]. Another subspecies of *F. nucleatum* subsp. *animalis* 7_1 presented two bacteriophages ΦFnuI and ΦFnuII after mitomycin C treatment [135]. However, unlike FnpΦ02, the new bacteriophages were non-infective [135]. More recently, the effect of FnuI, another siphovirus, against *F. nucleatum* biofilms was described. FnuI is capable of disrupting *F. nucleatum* biofilm, which enables new possibilities for the treatment and targeted removal of this bacteria [133].

A strategy to utilize whole ultraviolet inactivated bacteria as a vaccine achieved a reduced progression of chronic halitosis associated with abscesses [136]. Studies have traditionally targeted major outer membrane proteins such as FomA and Fap2and a protein based a vaccine against halitosis was created by using purified FomA [137]. A vaccine will be an alternative to antibiotic treatments for chronic halitosis [136] and it will have to prevent the shifting of the microbiome to a pathogenic dysbiosis that results in periodontal diseases. A divalent vaccine targeted tFomA-FlaB and elicited strong immune responses in the mucosa as well as the systemic immune compartments. In addition, it prevented periodontal pathologies caused by live bacterial infections as it inhibited co-aggregation, biofilm formation and host cell interactions [138].

Moreover, Fap2 is also a potential target for therapeutic opportunities as it may result in reduced *Fusobacterium* counts in colorectal cancer. Therefore, the creation of a deficient Fap2 *F. nucleatum* strain is an encouraging option for alternative therapies for bacteria associated with cancer treatment [52]. Another protein targeted was Alkyl hydroperoxide reductase subunit C

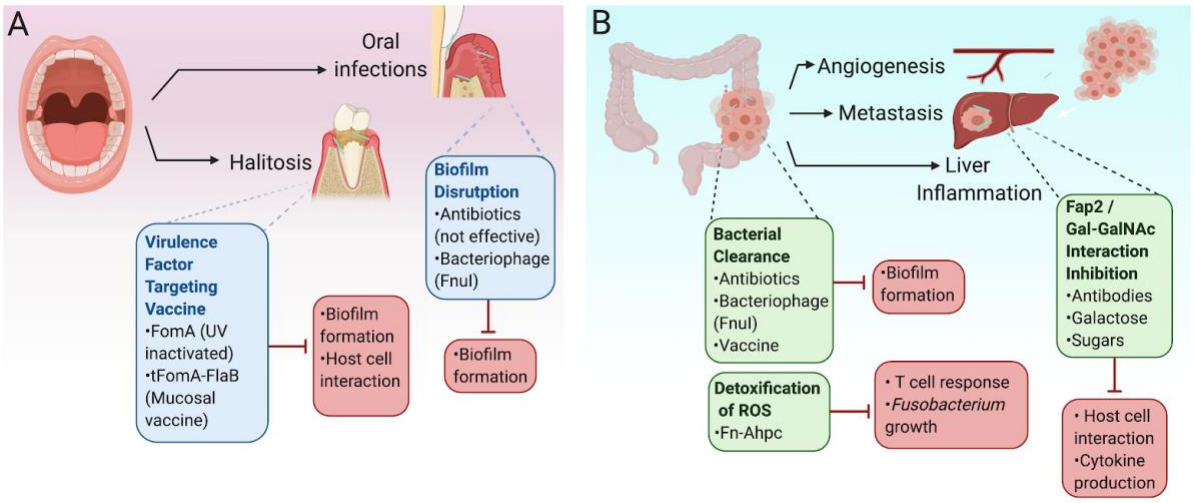


Figure 1-5 *Fusobacterium* therapeutic alternatives. Previous strategies to target *Fusobacterium* in oral diseases and colorectal cancer include virulence factor targeting (FomA, Fap2), biofilm disruption (antibiotics, bacteriophage), detoxification of reactive oxygen species and interaction inhibition.

(AhpC), which enables the detoxification of reactive oxygen species (ROS) that form in microorganisms and induces strong T cell-mediated responses [139,140]. AhpC enables the endurance of pathogens under environmental stresses or during invasion and it was identified in the sera of colorectal cancer patients [139,140]. This tumor vaccine candidate inhibited *F. nucleatum* growth in mice and provided significant protection from infection potentially by inducing higher levels of IgG, IgA and intestinal mucus SIgA [140].

Conclusions

This introductory chapter focuses on the *Fusobacterium* genus, which is a diverse group of bacteria involved in numerous diseases in humans and animals that can shift from an oral commensal microorganism to a pathogen by largely unknown mechanisms. A major emphasis has been in understanding in the pathogenic molecular mechanisms of *F. nucleatum* and *F. necrophorum*. These studies revealed how both species rely on an overabundance of outer membrane proteins to achieve invasion, infection, survival and pathogenesis.

Fusobacterium interacts with a wide array of microorganisms in the mouth through arginine and galactose inhibitable adhesins. These interactions lead to biofilm formation, where *Fusobacterium* plays a central role in coaggregation, regulation and growth. Many of the multi-microbial interactions that *Fusobacterium* mediate in the oral cavity through its adhesins are deemed as essential for the development of disease as *Fusobacterium* is also capable of contributing to the development of infections throughout the body. Nevertheless, many interspecies interactions and host interactions that lead to pathogenesis in the oral cavity remain understudied.

Fusobacterium can exist as an extracellular and intracellular bacterium. As it has been shown, several strains of *Fusobacterium* can adhere to [56], invade [56] and replicate utilizing actin inside of human cells [130]. Nonetheless, a detailed description of how *F. nucleatum* invades and survives still needs to be researched further. In the same way, another mechanism that is yet to be determined is the process by which a non-motile bacterium spreads throughout the human body by surviving passage in the bloodstream.

Additionally, major questions still remain on how *Fusobacterium* triggers immune responses that are detrimental to the human or animal host. An immune suppression of

macrophages, natural killer cells, and T-cells has been shown to contribute to colorectal cancer [141] and the upregulation of the secretion of inflammatory mediators, including NF-κB, IL-6, IL-8, IL-10, CXCL-1 and IL-18 has also shown a key role in cell proliferation and promotion of colorectal cancer [54] [141].

To this date, a few attempts to study *Fusobacterium* through novel approaches have been made. The development of antibiotic therapy, bacteriophages for biofilm disruption, vaccines and targeting of key virulence factors such as FomA and Fap2 has been the priority so far. In addition, the creation of genetic systems in *Fusobacterium* will lead to better understanding and development of successful therapeutic methods that will aid in the control of head-neck and oral infections, cardiovascular diseases, gastrointestinal disorders, and adverse pregnancy outcomes.

References

1. Bennett KW, Eley A. Fusobacteria: new taxonomy and related diseases. J Med Microbiol. 1993;39: 246–254.

2. Garrett WS, Onderdonk AB. 249 - Bacteroides, Prevotella, Porphyromonas, and Fusobacterium Species (and Other Medically Important Anaerobic Gram-Negative Bacilli). In: Bennett JE, Dolin R, Blaser MJ, editors. Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases (Eighth Edition). Philadelphia: Content Repository Only!; 2015. pp. 2773–2780.

3. Manson McGuire A, Cochrane K, Griggs AD, Haas BJ, Abeel T, Zeng Q, et al. Evolution of invasion in a diverse set of Fusobacterium species. MBio. 2014;5: e01864.

4. Spaulding EH, Rettger LF. The Fusobacterium Genus: II. Some Observations on Growth Requirements and Variation. J Bacteriol. 1937;34: 549–563.

5. Citron DM. Update on the taxonomy and clinical aspects of the genus fusobacterium. Clin Infect Dis. 2002;35: S22–7.

6. Brook I. Fusobacterium, Infection and Immunity. In: Delves PJ, editor. Encyclopedia of Immunology (Second Edition). Oxford: Elsevier; 1998. pp. 962–963.

7. Tadepalli S, Stewart GC, Nagaraja TG, Jang SS, Narayanan SK. Fusobacterium equinum possesses a leukotoxin gene and exhibits leukotoxin activity. Vet Microbiol. 2008;127: 89–96.

8. Cato EP, Moore LVH, Moore WEC. Fusobacterium alocis sp. nov. and Fusobacterium sulci sp. nov. from the Human Gingival Sulcus. Int J Syst Bacteriol. 1985;35: 475–477.

9. Riordan T. Human infection with Fusobacterium necrophorum (Necrobacillosis), with a focus on Lemierre's syndrome. Clin Microbiol Rev. 2007;20: 622–659.

10. Tan ZL, Nagaraja TG, Chengappa MM. Fusobacterium necrophorum infections: virulence factors, pathogenic mechanism and control measures. Vet Res Commun. 1996;20: 113–140.

11. Edmondson MA, Roberts JF, Baird AN, Bychawski S, Pugh DG. Theriogenology of sheep and goats. Sheep and goat medicine. Elsevier; 2012. pp. 150–230.

12. Löe H. The role of bacteria in periodontal diseases. Bull World Health Organ. 1981;59: 821–825.

13. Griffen AL, Beall CJ, Campbell JH, Firestone ND, Kumar PS, Yang ZK, et al. Distinct and complex bacterial profiles in human periodontitis and health revealed by 16S pyrosequencing. ISME J. 2012;6: 1176–1185.

14. Kilian M, Chapple ILC, Hannig M, Marsh PD, Meuric V, Pedersen AML, et al. The oral microbiome - an update for oral healthcare professionals. Br Dent J. 2016;221: 657–666.

15. Yang N-Y, Zhang Q, Li J-L, Yang S-H, Shi Q. Progression of periodontal inflammation in adolescents is associated with increased number of Porphyromonas gingivalis, Prevotella intermedia, Tannerella forsythensis, and Fusobacterium nucleatum. Int J Paediatr Dent. 2014;24: 226–233.

16. Han YW, Fardini Y, Chen C, Iacampo KG, Peraino VA, Shamonki JM, et al. Term stillbirth caused by oral Fusobacterium nucleatum. Obstet Gynecol. 2010;115: 442–445.

17. Cahill RJ, Tan S, Dougan G, O'Gaora P, Pickard D, Kennea N, et al. Universal DNA primers amplify bacterial DNA from human fetal membranes and link Fusobacterium nucleatum with prolonged preterm membrane rupture. Mol Hum Reprod. 2005;11: 761–766.

18. Bohrer JC, Kamemoto LE, Almeida PG, Ogasawara KK. Acute chorioamnionitis at term caused by the oral pathogen Fusobacterium nucleatum. Hawaii J Med Public Health. 2012;71: 280.

19. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, et al. Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. Cell Host Microbe. 2013;14: 207–215.

20. Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, et al. Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. Genome Res. 2012;22: 292–298.

21. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, et al. Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. Genome Res. 2012;22: 299–306.

22. Tahara T, Shibata T, Kawamura T, Okubo M, Ichikawa Y, Sumi K, et al. Fusobacterium detected in colonic biopsy and clinicopathological features of ulcerative colitis in Japan. Dig Dis Sci. 2015;60: 205–210.

23. Strauss J, Kaplan GG, Beck PL, Rioux K, Panaccione R, Devinney R, et al. Invasive potential of gut mucosa-derived Fusobacterium nucleatum positively correlates with IBD status of the host. Inflamm Bowel Dis. 2011;17: 1971–1978.

24. Lee MJ, Ha YE, Park HY, Lee JH, Lee YJ, Sung KS, et al. Osteomyelitis of a long bone due to Fusobacterium nucleatum and Actinomyces meyeri in an immunocompetent adult: a case report and literature review. BMC Infect Dis. 2012;12: 161.

25. Jayasimhan D, Wu L, Huggan P. Fusobacterial liver abscess: a case report and review of the literature. BMC Infect Dis. 2017;17: 440.

26. Afra K, Laupland K, Leal J, Lloyd T, Gregson D. Incidence, risk factors, and outcomes of Fusobacterium species bacteremia. BMC Infect Dis. 2013;13: 264.

27. Thurnheer T, Karygianni L, Flury M, Belibasakis GN. Fusobacterium Species and Subspecies Differentially Affect the Composition and Architecture of Supra- and Subgingival Biofilms Models. Front Microbiol. 2019;10: 1716.

28. Kolenbrander PE, Palmer RJ Jr, Periasamy S, Jakubovics NS. Oral multispecies biofilm development and the key role of cell-cell distance. Nat Rev Microbiol. 2010;8: 471–480.

29. Bradshaw DJ, Marsh PD, Watson GK, Allison C. Role of Fusobacterium nucleatum and coaggregation in anaerobe survival in planktonic and biofilm oral microbial communities during aeration. Infect Immun. 1998;66: 4729–4732.

30. Kaplan CW, Lux R, Haake SK, Shi W. The Fusobacterium nucleatum outer membrane protein RadD is an arginine-inhibitable adhesin required for inter-species adherence and the structured architecture of multispecies biofilm. Mol Microbiol. 2009;71: 35–47.

31. Coppenhagen-Glazer S, Sol A, Abed J, Naor R, Zhang X, Han YW, et al. Fap2 of Fusobacterium nucleatum is a galactose-inhibitable adhesin involved in coaggregation, cell adhesion, and preterm birth. Infect Immun. 2015;83: 1104–1113.

32. Park J, Shokeen B, Haake SK, Lux R. Characterization of Fusobacterium nucleatum ATCC 23726 adhesins involved in strain-specific attachment to Porphyromonas gingivalis. Int J Oral Sci. 2016;8: 138.

33. Han YW, Shi W, Huang GT, Kinder Haake S, Park NH, Kuramitsu H, et al. Interactions between periodontal bacteria and human oral epithelial cells: Fusobacterium nucleatum adheres to and invades epithelial cells. Infect Immun. 2000;68: 3140–3146.

34. Kaplan CW, Lux R, Haake SK, Shi W. The Fusobacterium nucleatum outer membrane protein RadD is an arginine-inhibitable adhesin required for inter-species adherence and the structured architecture of multispecies biofilm. Mol Microbiol. 2009;71: 35–47.

35. Lima BP, Hu LI, Vreeman GW, Weibel DB, Lux R. The Oral Bacterium Fusobacterium nucleatum Binds Staphylococcus aureus and Alters Expression of the Staphylococcal Accessory Regulator sarA. Microb Ecol. 2019;78: 336–347.

36. Bor B, Cen L, Agnello M, Shi W, He X. Morphological and physiological changes induced by contact-dependent interaction between Candida albicans and Fusobacterium nucleatum. Sci Rep. 2016;6: 27956.

37. Wu T, Cen L, Kaplan C, Zhou X, Lux R, Shi W, et al. Cellular Components Mediating Coadherence of Candida albicans and Fusobacterium nucleatum. J Dent Res. 2015;94: 1432–1438.

38. Guo L, Shokeen B, He X, Shi W, Lux R. Streptococcus mutans SpaP binds to RadD of Fusobacterium nucleatum ssp. polymorphum. Mol Oral Microbiol. 2017;32: 355–364.

39. Lima BP, Shi W, Lux R. Identification and characterization of a novel Fusobacterium nucleatum adhesin involved in physical interaction and biofilm formation with Streptococcus gordonii. Microbiologyopen. 2017;6. doi:10.1002/mbo3.444

40. Kinder SA, Holt SC. Localization of the Fusobacterium nucleatum T18 adhesin activity mediating coaggregation with Porphyromonas gingivalis T22. J Bacteriol. 1993;175: 840–850.

41. Liu P-F, Shi W, Zhu W, Smith JW, Hsieh S-L, Gallo RL, et al. Vaccination targeting surface FomA of Fusobacterium nucleatum against bacterial co-aggregation: Implication for treatment of periodontal infection and halitosis. Vaccine. 2010;28: 3496–3505.

42. Nakagaki H, Sekine S, Terao Y, Toe M, Tanaka M, Ito H-O, et al. Fusobacterium nucleatum envelope protein FomA is immunogenic and binds to the salivary statherin-derived peptide. Infect Immun. 2010;78: 1185–1192.

43. Kaplan A, Kaplan CW, He X, McHardy I, Shi W, Lux R. Characterization of aid1, a novel gene involved in Fusobacterium nucleatum interspecies interactions. Microb Ecol. 2014;68: 379–387.

44. Meuskens I, Saragliadis A, Leo JC, Linke D. Type V Secretion Systems: An Overview of Passenger Domain Functions. Front Microbiol. 2019;10: 1163.

45. Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, et al. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. Nat Rev Microbiol. 2015;13: 343–359.

46. Abed J, Emgård JEM, Zamir G, Faroja M, Almogy G, Grenov A, et al. Fap2 Mediates Fusobacterium nucleatum Colorectal Adenocarcinoma Enrichment by Binding to Tumor-Expressed Gal-GalNAc. Cell Host Microbe. 2016;20: 215–225.

47. Kaplan CW, Ma X, Paranjpe A, Jewett A, Lux R, Kinder-Haake S, et al. Fusobacterium nucleatum outer membrane proteins Fap2 and RadD induce cell death in human lymphocytes. Infect Immun. 2010;78: 4773–4778.

48. Narayanan S, Stewart GC, Chengappa MM, Willard L, Shuman W, Wilkerson M, et al. Fusobacterium necrophorum leukotoxin induces activation and apoptosis of bovine leukocytes. Infect Immun. 2002;70: 4609–4620.

49. Klauser T, Pohlner J, Meyer TF. The secretion pathway of IgA protease-type proteins in gram-negative bacteria. Bioessays. 1993;15: 799–805.

50. Leo JC, Grin I, Linke D. Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. Philosophical Transactions of the Royal Society B: Biological Sciences. 2012. pp. 1088–1101. doi:10.1098/rstb.2011.0208

51. Abby SS, Cury J, Guglielmini J, Néron B, Touchon M, Rocha EPC. Identification of protein secretion systems in bacterial genomes. Scientific Reports. 2016. doi:10.1038/srep23080

52. Ganesan K, Guo S, Fayyaz S, Zhang G, Xu B. Targeting Programmed Fusobacterium nucleatum Fap2 for Colorectal Cancer Therapy. Cancers . 2019;11. doi:10.3390/cancers11101592

53. Gur C, Ibrahim Y, Isaacson B, Yamin R, Abed J, Gamliel M, et al. Binding of the Fap2 protein of Fusobacterium nucleatum to human inhibitory receptor TIGIT protects tumors from immune cell attack. Immunity. 2015;42: 344–355.

54. Casasanta MA, Yoo CC, Udayasuryan B, Sanders BE, Umaña A, Zhang Y, et al. Fusobacterium nucleatum host-cell binding and invasion induces IL-8 and CXCL1 secretion that drives colorectal cancer cell migration. Sci Signal. 2020;13. doi:10.1126/scisignal.aba9157

55. Kaplan CW, Lux R, Huynh T, Jewett A, Shi W, Haake SK. Fusobacterium nucleatum apoptosis-inducing outer membrane protein. J Dent Res. 2005;84: 700–704.

56. Umana A, Sanders BE, Yoo CC, Casasanta MA, Udayasuryan B, Verbridge SS, et al. Utilizing Whole Fusobacterium Genomes to Identify, Correct, and Characterize Potential Virulence Protein Families. J Bacteriol. 2019. doi:10.1128/JB.00273-19

57. Ruiz-Perez F, Nataro JP. Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence. Cell Mol Life Sci. 2014;71: 745–770.

58. Doron L, Coppenhagen-Glazer S, Ibrahim Y, Eini A, Naor R, Rosen G, et al. Identification and characterization of fusolisin, the Fusobacterium nucleatum autotransporter serine protease. PLoS One. 2014;9: e111329.

59. Bachrach G, Rosen G, Bellalou M, Naor R, Sela MN. Identification of a Fusobacterium nucleatum 65 kDa serine protease. Oral Microbiol Immunol. 2004;19: 155–159.

60. Berg JN, Scanlan CM. Studies of Fusobacterium necrophorum from bovine hepatic abscesses: biotypes, quantitation, virulence, and antibiotic susceptibility. Am J Vet Res. 1982;43: 1580–1586.

61. Scanlan CM, Berg JN, Campbell FF. Biochemical characterization of the leukotoxins of three bovine strains of Fusobacterium necrophorum. Am J Vet Res. 1986;47: 1422–1425.

62. Coyle-Dennis JE, Lauerman LH. Correlations between leukocidin production and virulence of two isolates of Fusobacterium necrophorum. Am J Vet Res. 1979;40: 274–276.

63. Narayanan SK, Nagaraja TG, Chengappa MM, Stewart GC. Cloning, sequencing, and expression of the leukotoxin gene from Fusobacterium necrophorum. Infect Immun. 2001;69: 5447–5455.

64. Tan ZL, Nagaraja TG, Chengappa MM. Factors affecting the leukotoxin activity of Fusobacterium necrophorum. Vet Microbiol. 1992;32: 15–28.

65. Oelke AM, Nagaraja TG, Wilkerson MJ, Stewart GC. The leukotoxin operon of Fusobacterium necrophorum is not present in other species of Fusobacterium. Anaerobe. 2005;11: 123–129.

66. Tadepalli S, Stewart GC, Nagaraja TG, Narayanan SK. Leukotoxin operon and differential expressions of the leukotoxin gene in bovine Fusobacterium necrophorum subspecies. Anaerobe. 2008. pp. 13–18. doi:10.1016/j.anaerobe.2007.09.001

67. Tan ZL, Nagaraja TG, Chengappa MM, Smith JS. Biological and biochemical characterization of Fusobacterium necrophorum leukotoxin. Am J Vet Res. 1994;55: 515–521.

68. Tan ZL, Nagaraja TG, Chengappa MM, Staats JJ. Purification and quantification of Fusobacterium necrophorum leukotoxin by using monoclonal antibodies. Vet Microbiol. 1994;42: 121–133.

69. Holm K, Collin M, Hagelskjær-Kristensen L, Jensen A, Rasmussen M. Three variants of the leukotoxin gene in human isolates of Fusobacterium necrophorum subspecies funduliforme. Anaerobe. 2017;45: 129–132.

70. Łyskowski A, Leo JC, Goldman A. Structure and Biology of Trimeric Autotransporter Adhesins. In: Linke D, Goldman A, editors. Bacterial Adhesion: Chemistry, Biology and Physics. Dordrecht: Springer Netherlands; 2011. pp. 143–158.

71. Nummelin H, Merckel MC, Leo JC, Lankinen H, Skurnik M, Goldman A. The Yersinia adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. EMBO J. 2004;23: 701–711.

72. El Tahir Y, Skurnik M. YadA, the multifaceted Yersinia adhesin. Int J Med Microbiol. 2001;291: 209–218.

73. Brewer ML, Dymock D, Brady RL, Singer BB, Virji M, Hill DJ. Fusobacterium spp. target human CEACAM1 via the trimeric autotransporter adhesin CbpF. J Oral Microbiol. 2019;11: 1565043.

74. Casasanta MA, Yoo CC, Smith HB, Duncan AJ, Cochrane K, Varano AC, et al. A chemical and biological toolbox for Type Vd secretion: Characterization of the phospholipase A1 autotransporter FplA from Fusobacterium nucleatum. J Biol Chem. 2017;292: 20240–20254.

75. Ellis TN, Kuehn MJ. Virulence and immunomodulatory roles of bacterial outer membrane vesicles. Microbiol Mol Biol Rev. 2010;74: 81–94.

76. Zanzoni A, Spinelli L, Braham S, Brun C. Perturbed human sub-networks by Fusobacterium nucleatum candidate virulence proteins. Microbiome. 2017;5: 89.

77. Liu J, Hsieh C-L, Gelincik O, Devolder B, Sei S, Zhang S, et al. Proteomic characterization of outer membrane vesicles from gut mucosa-derived fusobacterium nucleatum. J Proteomics. 2019;195: 125–137.

78. Munshi RM. The association between Fusobacterium nucleatum Outer Membrane Vesicles and Colonic Cancer. 2017 [cited 30 Mar 2020]. Available: https://www.semanticscholar.org/paper/3955fe7ff321fb873854f01032c84b6f3c21d211

79. Kumar A, Peterson G, Nagaraja TG, Narayanan S. Outer membrane proteins of Fusobacterium necrophorum subsp. necrophorum and subsp. funduliforme. J Basic Microbiol. 2014;54: 812–817.

80. Han YW, Ikegami A, Rajanna C, Kawsar HI, Zhou Y, Li M, et al. Identification and characterization of a novel adhesin unique to oral fusobacteria. J Bacteriol. 2005;187: 5330–5340.

81. Xu M, Yamada M, Li M, Liu H, Chen SG, Han YW. FadA from Fusobacterium nucleatum utilizes both secreted and nonsecreted forms for functional oligomerization for attachment and invasion of host cells. J Biol Chem. 2007;282: 25000–25009.

82. Nithianantham S, Xu M, Yamada M, Ikegami A, Shoham M, Han YW. Crystal structure of FadA adhesin from Fusobacterium nucleatum reveals a novel oligomerization motif, the leucine chain. J Biol Chem. 2009;284: 3865–3872.

83. Ikegami A, Chung P, Han YW. Complementation of the fadA mutation in Fusobacterium nucleatum demonstrates that the surface-exposed adhesin promotes cellular invasion and placental colonization. Infect Immun. 2009;77: 3075–3079.

84. Shang F-M, Liu H-L. Fusobacterium nucleatum and colorectal cancer: A review. World J Gastrointest Oncol. 2018;10: 71–81.

85. Fardini Y, Wang X, Témoin S, Nithianantham S, Lee D, Shoham M, et al. Fusobacterium nucleatum adhesin FadA binds vascular endothelial cadherin and alters endothelial integrity. Mol Microbiol. 2011;82: 1468–1480.

86. Vander Haar EL, So J, Gyamfi-Bannerman C, Han YW. Fusobacterium nucleatum and adverse pregnancy outcomes: Epidemiological and mechanistic evidence. Anaerobe. 2018;50: 55–59.

87. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/β-catenin signaling via its FadA adhesin. Cell Host Microbe. 2013;14: 195–206.

88. Gupta S, Ghosh SK, Scott ME, Bainbridge B, Jiang B, Lamont RJ, et al. Fusobacterium nucleatum-associated beta-defensin inducer (FAD-I): identification, isolation, and functional evaluation. J Biol Chem. 2010;285: 36523–36531.

89. Bhattacharyya S, Ghosh SK, Shokeen B, Eapan B, Lux R, Kiselar J, et al. FAD-I, a Fusobacterium nucleatum Cell Wall-Associated Diacylated Lipoprotein That Mediates Human Beta Defensin 2 Induction through Toll-Like Receptor-1/2 (TLR-1/2) and TLR-2/6. Infect Immun. 2016;84: 1446–1456.

90. Takada H, Ogawa T, Yoshimura F, Otsuka K, Kokeguchi S, Kato K, et al. Immunobiological activities of a porin fraction isolated from Fusobacterium nucleatum ATCC 10953. Infect Immun. 1988;56: 855–863.

91. Bakken V, Aarø S, Jensen HB. Purification and partial characterization of a major outer-membrane protein of Fusobacterium nucleatum. J Gen Microbiol. 1989;135: 3253–3262.

92. Haake SK, Wang X. Cloning and expression of FomA, the major outer-membrane protein gene from Fusobacterium nucleatum T18. Arch Oral Biol. 1997;42: 19–24.

93. Kaufman J, DiRienzo JM. Isolation of a corncob (coaggregation) receptor polypeptide from Fusobacterium nucleatum. Infect Immun. 1989;57: 331–337.

94. Guo M, Han YW, Sharma A, De Nardin E. Identification and characterization of human immunoglobulin G Fc receptors of Fusobacterium nucleatum. Oral Microbiol Immunol. 2000;15: 119–123.

95. Tolo K, Helgeland K. Fc-binding components: a virulence factor in Actinobacillus actinomycetemcomitans? Oral Microbiol Immunol. 1991;6: 373–377.

96. Toussi DN, Liu X, Massari P. The FomA porin from Fusobacterium nucleatum is a Toll-like receptor 2 agonist with immune adjuvant activity. Clin Vaccine Immunol. 2012;19: 1093–1101.

97. Menon S, Pillai DK, Narayanan S. Characterization of Fusobacterium necrophorum subsp. necrophorum outer membrane proteins. Anaerobe. 2018;50: 101–105.

98. Gill J, Haydon TG, Rawdon TG, McFadden AMJ, Ha H-J, Shen Z, et al. Helicobacter bilis and Helicobacter trogontum: infectious causes of abortion in sheep. J Vet Diagn Invest. 2016;28: 225–234.

99. Brennan CA, Garrett WS. Fusobacterium nucleatum — symbiont, opportunist and oncobacterium. Nature Reviews Microbiology. 2019. pp. 156–166. doi:10.1038/s41579-018-0129-6

100. Sun C-H, Li B-B, Wang B, Zhao J, Zhang X-Y, Li T-T, et al. The role of Fusobacterium nucleatum in colorectal cancer: from carcinogenesis to clinical management. Chronic Dis Transl Med. 2019;5: 178–187.

101. Maldonado RF, Sá-Correia I, Valvano MA. Lipopolysaccharide modification in Gram-negative bacteria during chronic infection. FEMS Microbiol Rev. 2016;40: 480–493.

102. Raetz CRH, Whitfield C. Lipopolysaccharide endotoxins. Annu Rev Biochem. 2002;71: 635–700.

103. Hase S, Hofstad T, Rietschel ET. Chemical structure of the lipid A component of lipopolysaccharides from Fusobacterium nucleatum. J Bacteriol. 1977;129: 9–14.

104. Dahlén G, Nygren H, Hannsson HA. Immunoelectron microscopic localization of lipopolysaccharides in the cell wall of Bacteroides oralis and Fusobacterium nucleatum. Infect Immun. 1978;19: 265–271.

105. Okahashi N, Koga T, Nishihara T, Fujiwara T, Hamada S. Immunobiological properties of lipopolysaccharides isolated from Fusobacterium nucleatum and F. necrophorum. J Gen Microbiol. 1988;134: 1707–1715.

106. Martinho FC, Leite FRM, Nóbrega LMM, Endo MS, Nascimento GG, Darveau RP, et al. Comparison of Fusobacterium nucleatum and Porphyromonas gingivalis Lipopolysaccharides Clinically Isolated from Root Canal Infection in the Induction of Pro-Inflammatory Cytokines Secretion. Braz Dent J. 2016;27: 202–207.

107. To TT, Gümüş P, Nizam N, Buduneli N, Darveau RP. Subgingival Plaque in Periodontal Health Antagonizes at Toll-Like Receptor 4 and Inhibits E-Selectin Expression on Endothelial Cells. Infect Immun. 2016;84: 120–126.

108. Jia Y-P, Wang K, Zhang Z-J, Tong Y-N, Han D, Hu C-Y, et al. TLR2/TLR4 activation induces Tregs and suppresses intestinal inflammation caused by Fusobacterium nucleatum in vivo. PLoS One. 2017;12: e0186179.

109. Yang Y, Weng W, Peng J, Hong L, Yang L, Toiyama Y, et al. Fusobacterium nucleatum Increases Proliferation of Colorectal Cancer Cells and Tumor Development in Mice by Activating Toll-Like Receptor 4 Signaling to Nuclear Factor-κB, and Up-regulating Expression of MicroRNA-21. Gastroenterology. 2017;152: 851–866.e24.

110. Shi C, Yang Y, Xia Y, Okugawa Y, Yang J, Liang Y, et al. Novel evidence for an oncogenic role of microRNA-21 in colitis-associated colorectal cancer. Gut. 2016;65: 1470–1481.

111. Hur K, Toiyama Y, Okugawa Y, Ide S, Imaoka H, Boland CR, et al. Circulating microRNA-203 predicts prognosis and metastasis in human colorectal cancer. Gut. 2017;66: 654–665.

112. Yamada A, Horimatsu T, Okugawa Y, Nishida N, Honjo H, Ida H, et al. Serum miR-21, miR-29a, and miR-125b Are Promising Biomarkers for the Early Detection of Colorectal Neoplasia. Clin Cancer Res. 2015;21: 4234–4242.

113. Ma Y, Zhang P, Wang F, Zhang H, Yang J, Peng J, et al. miR-150 as a potential biomarker associated with prognosis and therapeutic outcome in colorectal cancer. Gut. 2012;61: 1447–1453.

114. Abreu MT, Peek RM Jr. Gastrointestinal malignancy and the microbiome. Gastroenterology. 2014;146: 1534–1546.e3.

115. Yu T, Guo F, Yu Y, Sun T, Ma D, Han J, et al. Fusobacterium nucleatum Promotes Chemoresistance to Colorectal Cancer by Modulating Autophagy. Cell. 2017;170: 548–563.e16.

116. Xue Y, Xiao H, Guo S, Xu B, Liao Y, Wu Y, et al. Indoleamine 2,3-dioxygenase expression regulates the survival and proliferation of Fusobacterium nucleatum in THP-1-derived macrophages. Cell Death Dis. 2018;9: 355.

117. Liu H, Redline RW, Han YW. Fusobacterium nucleatum induces fetal death in mice via stimulation of TLR4-mediated placental inflammatory response. J Immunol. 2007;179: 2501–2508.

118. Friberg N, Carlson P, Kentala E, Mattila PS, Kuusela P, Meri S, et al. Factor H binding as a complement evasion mechanism for an anaerobic pathogen, Fusobacterium necrophorum. The Journal of Immunology. 2008;181: 8624–8632.

119. Alyami HM, Finoti LS, Teixeira HS, Aljefri A, Kinane DF, Benakanakere MR. Role of NOD1/NOD2 receptors in Fusobacterium nucleatum mediated NETosis. Microb Pathog. 2019;131: 53–64.

120. Mangan DF, Novak MJ, Vora SA, Mourad J, Kriger PS. Lectinlike interactions of Fusobacterium nucleatum with human neutrophils. Infect Immun. 1989;57: 3601–3611.

121. Van Dyke TE, Bartholomew E, Genco RJ, Slots J, Levine MJ. Inhibition of neutrophil chemotaxis by soluble bacterial products. J Periodontol. 1982;53: 502–508.

122. Guo M, Reynolds H, Stinson M, De Nardin E. Isolation and characterization of a human neutrophil aggregation defective mutant of Fusobacterium nucleatum. FEMS Immunol Med Microbiol. 2000;27: 241–246.

123. Wright HJ, Chapple ILC, Matthews JB, Cooper PR. Fusobacterium nucleatum regulation of neutrophil transcription. J Periodontal Res. 2011;46: 1–12.

124. Sheikhi M, Gustafsson A, Jarstrand C. Cytokine, elastase and oxygen radical release by Fusobacterium nucleatum-activated leukocytes: a possible pathogenic factor in periodontitis. J Clin Periodontol. 2000;27: 758–762.

125. Passo SA, Syed SA, Silva J Jr. Neutrophil chemiluminescence in response to Fusobacterium nucleatum. J Periodontal Res. 1982;17: 604–613.

126. Katsuragi H, Ohtake M, Kurasawa I, Saito K. Intracellular production and extracellular release of oxygen radicals by PMNs and oxidative stress on PMNs during phagocytosis of periodontopathic bacteria. Odontology. 2003;91: 13–18.

127. Kurgan Ş, Kansal S, Nguyen D, Stephens D, Koroneos Y, Hasturk H, et al. Strain-Specific Impact of Fusobacterium nucleatum on Neutrophil Function. J Periodontol. 2017;88: 380–389.

128. Ji S, Shin JE, Kim YC, Choi Y. Intracellular degradation of Fusobacterium nucleatum in human gingival epithelial cells. Mol Cells. 2010;30: 519–526.

129. Bui FQ, Johnson L, Roberts J, Hung S-C, Lee J, Atanasova KR, et al. Fusobacterium nucleatum infection of gingival epithelial cells leads to NLRP3 inflammasome-dependent secretion of IL-1β and the danger signals ASC and HMGB1. Cell Microbiol. 2016;18: 970–981.

130. Gursoy UK, Könönen E, Uitto V-J. Intracellular replication of fusobacteria requires new actin filament formation of epithelial cells. APMIS. 2008;116: 1063–1070.

131. Silva VL, Diniz CG, Cara DC, Santos SG, Nicoli JR, Carvalho MAR, et al. Enhanced pathogenicity of Fusobacterium nucleatum adapted to oxidative stress. Microb Pathog. 2005;39: 131–138.

132. Szafrański SP, Winkel A, Stiesch M. The use of bacteriophages to biocontrol oral biofilms. J Biotechnol. 2017;250: 29–44.

133. Kabwe M, Brown TL, Dashper S, Speirs L, Ku H, Petrovski S, et al. Genomic, morphological and functional characterisation of novel bacteriophage FNU1 capable of disrupting Fusobacterium nucleatum biofilms. Sci Rep. 2019;9: 9107.

134. Machuca P, Daille L, Vinés E, Berrocal L, Bittner M. Isolation of a novel bacteriophage specific for the periodontal pathogen Fusobacterium nucleatum. Appl Environ Microbiol. 2010;76: 7243–7250.

135. Cochrane K, McGuire AM, Priest ME, Abouelleil A, Cerqueira GC, Lo R, et al. Complete genome sequences and analysis of the Fusobacterium nucleatum subspecies animalis 7-1 bacteriophage ϕFunu1 and ϕFunu2. Anaerobe. 2016;38: 125–129.

136. Liu P-F, Haake SK, Gallo RL, Huang C-M. A novel vaccine targeting Fusobacterium nucleatum against abscesses and halitosis. Vaccine. 2009;27: 1589–1595.

137. Liu P-F, Huang I-F, Shu C-W, Huang C-M. Halitosis vaccines targeting FomA, a biofilm-bridging protein of fusobacteria nucleatum. Curr Mol Med. 2013;13: 1358–1367.

138. Puth S, Hong SH, Na HS, Lee HH, Lee YS, Kim SY, et al. A built-in adjuvant-engineered mucosal vaccine against dysbiotic periodontal diseases. Mucosal Immunol. 2019;12: 565–579.

139. Wang H-W, Chung C-H, Ma T-Y, Wong H-C. Roles of alkyl hydroperoxide reductase subunit C (AhpC) in viable but nonculturable Vibrio parahaemolyticus. Appl Environ Microbiol. 2013;79: 3734–3743.

140. Guo S-H, Wang H-F, Nian Z-G, Wang Y-D, Zeng Q-Y, Zhang G. Immunization with alkyl hydroperoxide reductase subunit C reduces Fusobacterium nucleatum load in the intestinal tract. Sci Rep. 2017;7: 10566.

141. Wu J, Li Q, Fu X. Fusobacterium nucleatum Contributes to the Carcinogenesis of Colorectal Cancer by Inducing Inflammation and Suppressing Host Immunity. Transl Oncol. 2019;12: 846–851.

Chapter 2

Utilization of DNA methyltransferases for enhanced bacterial genetics in *Fusobacterium*

Ariana Umaña[1], Blake E. Sanders[1], Kevin Williams[1], Bryce F. Wozniak[1], Daniel J. Slade[1]

1. Virginia Polytechnic Institute and State University, Department of Biochemistry, Blacksburg, VA, USA.

Abstract

Fusobacterium are commensal Gram-negative anaerobic bacteria that are known for their genetic recalcitrance. To enhance our understanding in how *Fusobacterium nucleatum* converts to an opportunistic pathogen, this genetic barrier must be overcome to facilitate molecular biology approaches. Herein, we present a comprehensive bioinformatic characterization of restriction modification systems in *Fusobacterium* that greatly enhance the DNA uptake and subsequent incorporation of gene knockout plasmids onto the chromosome by bypassing the restriction modification systems. In total, we analyzed 17 new *Fusobacterium* DNA methyltransferases and present two Type II methyltransferases that are capable of providing endonuclease-protection for genetic modifications. These studies will accelerate our understanding of how these opportunistic oral pathogens contributes to diseases including periodontitis, preterm birth, and colorectal cancer.

Introduction

Bacterial transformation is the process of exogenous DNA introduction into the target bacterium [1]. To achieve bacterial transformation several methods including chemotransformation, electroporation, sonoporation, natural competence, and conjugation have been developed [1]. *F. nucleatum* is composed of a heterogenous group of microorganisms and only a few strains, *F. nucleatum* subsp. *nucleatum ATCC* 23726 [2–5], *F. nucleatum* subsp. *polymorphum* 10953 [5,6], and *F. nucleatum* subsp. *polymorphum* 12230 [7,8] have been successfully transformed by electroporation and sonoporation, respectively. Electroporation was first described in *Bacillus cereus* [9]. This method utilizes an electrical discharge to induce transient permeability of the cell membrane which leads to ion leakage, escape of metabolites and increased uptake of DNA [10,11]. Electroporation is a physical phenomenon that has been used in a broad range of bacteria and is initiated by penetration of water into the lipid bilayer, generating a reorientation of the lipid polar head groups as they interact with these water molecules. This effect leads to the creation of pores which produces a pathway for DNA entry [11].

The genetic competence of a few strains enables *Fusobacterium* to acquire naked DNA and incorporate it into their genome by recombination with homologous sequences or, in the case of episomal plasmids, by establishing a new episome [12]. However, most *Fusobacterium* species are recalcitrant to genetic manipulation and are in most cases considered genetically intractable. Genetic intractability is an elemental obstacle to progress that impedes basic, synthetic, and translational microbiology research by utilizing transformation techniques and development of genetic systems beyond a few model organisms [12]. Previous studies have attempted to achieve bacterial transformation in a wide variety of fusobacterial species, such as *F. necrophorum, F. varium, F. periodonticum,* and *F. mortiferum* [4]. Nonetheless, only *F. nucleatum* has been

successfully transformed, confirming a species and strain specificity of transformation in fusobacteria [4]. The specificity, difficulty to establish transformation methods and discrepancies in transformation efficiencyprove the existence of robust restriction-modification (R-M) systems [4].

The identification of R-M systems in bacteria was determined by the differential ability of bacteriophages to replicate on distinct strains of *E. coli* [13,14]. R-M systems are a type of defense systems that are widespread and are encoded in 90% of all sequenced bacterial genomes [12,15]. R-M systems cleave foreign DNA based on the presence or absence of specific methylation patterns [1]. Thus, encoding for a restriction endonuclease that digests DNA as well as a DNA methyltransferase, which prevents DNA cleavage by the endonuclease [1]. The presence of these cellular defense systems protects bacteria from invading phages and foreign DNA , which maintains chromosomal and genetic integrity[1,16,17].

R-M systems are classified as Type I, II, III or IV according to their molecular structure, subunit composition, cleavage position, restriction site, and cofactor specifications as shown in **Figure 2.1** [1]. Type I cuts exogenous DNA by forming protein complexes and random cleavage usually happens at substantial distances from an asymmetric recognition sequence (400 to 7,000 bp) [18]; while Type II consists of an individual restriction endonuclease and methyltransferase that cleave DNA at symmetrical recognition sites [18]. In a similar way to Type I, Type III forms a protein complex necessary for the restriction enzyme activity; however, the methyltransferase can function independently. DNA cleavage for Type III RM systems takes place 25 to 27 bp 3′ to an asymmetrical recognition sequence that is 5 to 6 bp in length [18]. Furthermore, Type IV shows restriction of heterologously modified DNA[1,19].

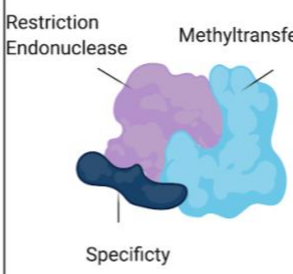
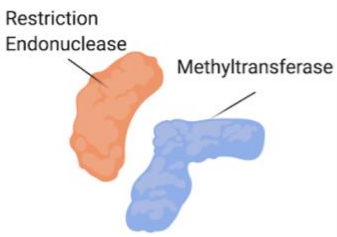
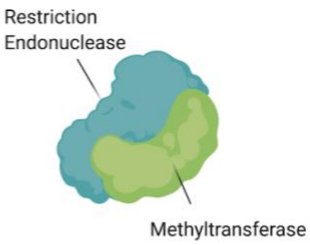
	Type I	Type II	Type III
			
Restriction Machinery	R_2M_2S	RE	R_2M_2
Methylation Machinery	M_2S	MT	M_2
Cleavage site	Variable	Fixed	Variable
Nucleobase	^{6m}A	^{6m}A, ^{5m}C, ^{4m}C	^{6m}A

Figure 2-1 Restriction Modification (R-M) systems in this study are classified as Type I, Type II and Type III. R-M systems are divided according to their molecular structure, subunit composition, cleavage position, restriction site, and cofactor specification.

In addition to R-M systems, bacteria have evolved several approaches to prevent phage infection, including abortive infection (Abi) mechanisms, CRISPR–Cas system and the BREX System [20]. The BREX (Bacteriophage Exclusion) exists in 10% of the sequenced bacterial genomes and phage adsorption occurs in these bacteria [20]. Nonetheless, the BREX system blocks phage DNA replication and lysogeny in infected cells differentiating itself from R-M systems, phage DNA is not cleaved or digested by BREX, which suggests an uncharacterized defense system [20].

DNA methyltransferases are a widespread family of enzymes present in eukaryotic and prokaryotic organisms. The DNA methyltransferases identify sequence-specific double-stranded DNA and relocate a methyl group from the donor S-adenosyl-methionine(SAM), to their target nucleotide [21]. The addition of a methyl group (—CH_3) to DNA, catalyzed by DNA methyltransferases, is the most frequent form of post-replicative nucleotide modification found in genomes of both prokaryotes and eukaryotes [22] and the methylated cytosine or adenine nucleotides of the restriction site are protected from DNA digestion by the cognate restriction endonuclease [1,21]. The DNA methyltransferase enzyme produces methylated bases and S-adenosyl-homocysteine, known inhibitor of methyltransferases. DNA methyltransferases are divided into three classes based on their products [21]. The main distinct product forms are: 5-methylcytosine (5mC), N4-methylcytosine (4mC) and, N6-methyladenine (6mA) [21]. Furthermore, these enzymes are also involved in diverse cellular processes including DNA mismatch repair, DNA replication, and regulation of virulence genes [21].

R-M systems prevent foreign DNA from entering and replicating or genomically altering the bacteria by discerning the methylation restriction sites specific to the host cell DNA and makes any foreign DNA or unrecognized methylation susceptible to digestion by bacterial restriction

endonucleases [18]. Restriction endonucleases possess high sequence and structural diversity and are categorized based on their functional properties rather than their phylogenetic conservation [23]. Initial studies in *Fusobacterium* R-M systems identified the presence of endonucleases; FnuDII and FnuEI were encoded in four *F. nucleatum* isolates, both part of Type II R-M systems with a GGCC and GATC recognition and cleave site, respectively [24,25]. A third endonuclease, Fnu4H, present in *F. nucleatum* 4H was also identified as a Type II enzyme with the recognition and cleavage site GCNGC [26].

R-M systems present a substantial challenge to the genetic manipulation of microorganisms. Approaches attempting to surpass the restriction barrier by performing host mimicking methylation, deletion of methyltransferases or syngenic DNA have successfully achieved higher transformation efficiencies [12,27–29]. The utilization of host-mimicking strategies where exogenous DNA can be conditioned to imitate the methylation pattern of the target bacteria, potentially leads to bypassing the host RM systems and continuing on to the integration into the genome **(Figure 2.2)** [1]. The most common mimicry-by-methylation approaches methylate target sites by utilizing recombinant methyltransferase enzymes achieving *in vitro* methylation, *in vivo* methylation by passaging a plasmid through a related strain or utilizing crude extract of the target bacterium to treat DNA [1,12]. With the current interest and development of gene deletion systems [2,3] in *F. nucleatum* ATCC 23726, a more robust and thorough investigation of R-M systems was warranted. Herein, we report the first comprehensive bioinformatic analysis and identification of R-M genes across several species of *Fusobacterium* as well as an exploration of the role of Type I, Type II and Type III R-M systems in *Fusobacterium* recalcitrance.

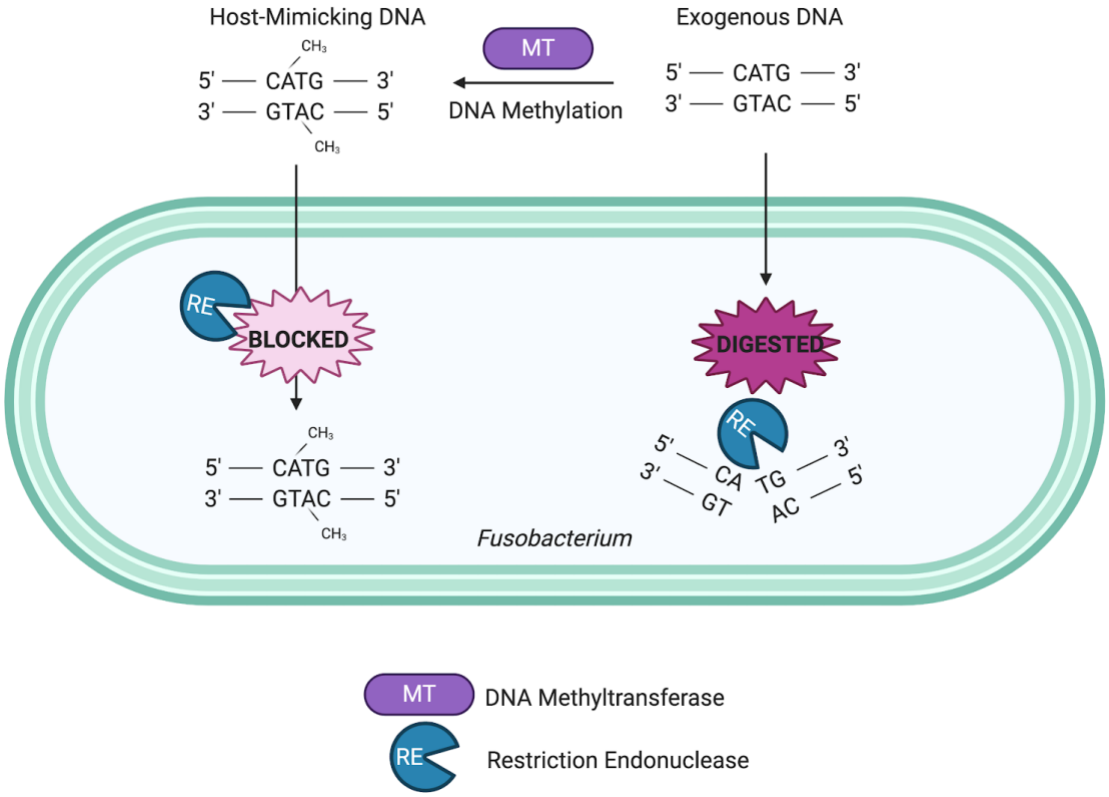


Figure 2-2 Host mimicking methylation in *Fusobacterium* to surpass R-M systems in bacteria. Host-mimicking strategies where exogenous DNA can be conditioned to imitate the methylation pattern of the target bacteria, potentially leads to bypassing the host RM systems and continuing to the integration into the genome. The most common mimicry-by-methylation approach to methylate restriction sites is to utilize recombinant methyltransferase enzymes. Adapted from Suzuki, 2013.

Altogether, we characterized 17 methyltransferases from six *Fusobacterium* strains belonging to Type I, Type II and Type III R-M systems. We identified two Type II methyltransferases from *F. nucleatum* ATCC 23726 that protect plasmid DNA from restriction and consequently, result in increased transformation efficiency in *F. nucleatum* ATCC 23726. Additionally, this novel study describes the development of two different approaches for conditioning plasmids with DNA methyltransferases; *in-vitro* with purified recombinant methyltransferases and a co-expression of vectors with constitutive expression of the methyltransferases utilizing *E. coli* as the host. The success of this method confirms that host-mimicking approaches can overcome R-M systems by statistically improving transformation efficiency.

Materials and Methods

Bacterial strains and plasmids

All *E. coli* strains utilized in these studies were grown aerobically overnight at 37°C in Luria Bertani medium (15g/L NaCl, 15 g/L tryptone, 10 g/L yeast extract). *Fusobacterium* strains were grown on solid agar plates made with Columbia Broth (Gibco), supplemented with hemin (5 μg/mL), menadione (0.5 μg/mL) and resazurin(1 μL/mL) under anaerobic conditions (90% N_2, 5% H_2, 5% CO_{22}) at 37°C. Liquid growths were inoculated from single *F. nucleatum* colonies and grown in CBHK liquid media under identical anaerobic conditions as previously stated. Where necessary, antibiotics were supplemented at the suggested concentrations: gentamicin, 20 μg/mL; carbenicillin, 100 μg/mL; chloramphenicol, 10 or 25 μg/mL; thiamphenicol, 5 μg/mL (CBHK plates); and streptomycin 50 μg/ml (CBHK plates). The plasmids and bacterial strains utilized in these experiments are listed in **Table 2.1** and **Table 2.2**, respectively.

Identification of *Fusobacterium* DNA Methyltransferases

REBASE, a curated database of restriction enzymes, was used to identify the DNA methyltransferases present in the *F. nucleatum* subsp. *nucleatum* ATCC 23726 (GCA_003019875.1), *F. nucleatum* subsp. *nucleatum* ATCC 25586 (GCA_003019295.1), *F. necrophorum* subsp. *funduliforme* 1_1_36S (GCA_003019715.1) found on FusoPortal and *F. nucleatum* subsp. *animalis* 7_1 (GCA_000158275.2), *F. nucleatum* subsp. *animalis* 4_8 *(*GCA_000400875.1*), F. nucleatum* subsp. *polymorphum* 10953 (GCA_000153625.1), and *F. necrophorum* 51357 (GCA_000262225.1) from the NCBI database. Type II and Type III DNA Mtases were further bioinformatically characterized using NIH SMARTBLAST and pHMMER. SMARTBLAST and pHMMER provided conserved domains indicating function of DNA Mtases.

Table 2.1 Plasmids utilized in this study.

Plasmid name	Description	Reference
pDJSVT1	Vector containing homologous regions +/- 1000 bp upstream and downstream of galKT for single crossover Integration in *F. nucleatum* 23726 and 25586 (Cm^r, Tm^r)	[3]
pDJSVT14	Vector containing homologous regions +/- 1000 bp upstream and downstream of galKT for single crossover Integration in *F. nucleatum* 7_1 (Cm^r, Tm^r)	This study
pDJSVT15	Vector containing homologous regions +/- 1000 bp upstream and downstream of galKT for single crossover Integration in *F. nucleatum* 4_8 (Cm^r, Tm^r)	This study
pDJSVT16	Vector containing homologous regions +/- 1000 bp upstream and downstream of galKT for single crossover Integration in *F. nucleatum* 10953 (Cm^r, Tm^r)	This study
pDJSVT17	Vector containing homologous regions +/- 1000 bp upstream and downstream of galKT for single crossover Integration in *F. nucleatum* 1_1_36S (Cm^r, Tm^r)	This study
pDJSVT18	Vector containing homologous regions +/- 1000 bp upstream and downstream of galKT for single crossover Integration in *F. necrophorum* 51357 (Cm^r, Tm^r)	This study
pHS30	pFN1- episomal based shuttle vector(Cm^r, Tm^r)	[4]
pET16b	Lac1, Amp^r	Novagen, EMD Milipore
pBAMD1-4	Broad host range mini-Tn*5* vector($Spec^r$)	[30]
pDJSVT14	pDJSVT13 with all CATG sites removed by OLE-PCR building of the vector (Cm^r, Tm^r)	This study
pDJSVT2	pET16b vector containing M.Fnu23I with Anderson medium promoter (Amp^r)	This study
pDJSVT3	pET16b vector containing M.Fnu23II with Anderson medium promoter (Amp^r)	This study
pDJSVT4	pET16b vector containing	This study

	M.Fnu23I and M.Fnu23II with Anderson medium promoter(Amp^r)	

Cm^r, Chloramphenicol resistance, Tm^r, Thiampenicol resistance
$Spec^r$, Spectinomycin resistance, Amp^r, Ampicillin resistance

Table 2.2 Strains utilized in this study

Strain	Bacteria species	Genotype and characteristics	Reference
Top10	*E. coli*	Δ (mrr-hsdRMS-mcrBC), mcrA, Phi80(del)M15, ΔlacX74, deoR, recA1, araD139, Δ(ara-leu)769, galU, galK, nupG, rpsL (SmR), endA1,	Invitrogen
ArcticExpress (DE3)RIL	*E. coli*	B F–ompT endA Hte [cpn10 cpn60 Gentr] hsdS(r8–m8–)dcm+ Tetr galλ (DE3) [argU ileY leuW Strr]	Agilent
LOBSTR-BL21(DE3)-RIL	*E. coli*	F–ompT hsdSB (rB- mB-) dcm gal (DE3)	[31]
ER2796	*E. coli*	*glnX44*(AS), F-, *fhuA2::IS2*, *λ-*, *e14-*, *trp-31*,*yedZ3069::Tn10*, *dcm-6*, *hisG1*, *rpsL104*, *argG6*, *Δdam-16::KanR*, *xyl-7*, *mtlA2*, *metB1*, *Δ(mcrC-mrr)114:IS10*	CGSC
F. nucleatum subsp. *nucleatum* ATCC 23726	*F. nucleatum*	Wild Type	ATCC[32]
F. nucleatum subsp. *nucleatum* ATCC 25586	*F. nucleatum*	Wild Type	ATCC[32]
F. nucleatum subsp. *animalis* 7_1	*F. nucleatum*	Wild Type	[32]

F. nucleatum subsp. *animalis* 4_8	*F. nucleatum*	Wild Type	[32]
F. nucleatum subsp. *polymorphum* 10953	*F. nucleatum*	Wild Type	[32]
F. necrophorum subsp. *funduliforme* 1_1_36S	*F. necrophorum*	Wild Type	[32]
F. necrophorum subsp. *funduliforme* 51357	*F. necrophorum*	Wild Type	ATCC

Table 2.3 Primers utilized in this study

Primer Name	**Sequence (5'-3')**	**Description**
prAUT91	gatccCCATGGGAATATATATTGATCCTCCATATAATACAGGAAAAG	Forward primer gene M.Fnu25IV to clone from *F. nucleatum* 25586 with NcoI cutsite into 16b
prAUT92	gactccCTCGAGTTAatgatgatgatgatgatgTATACTTCTAATTTCACTTATTCCAACTTG	Reverse primer gene M.Fnu25IV to clone from *F. nucleatum* 25586 with 6XHis-tag and XhoI cutsite into 16b
prAUT93	gactccCCATGGGAGAAAAACTAAATGGAACAAGCATGG	Forward primer gene M.Fnu25V to clone from *F. nucleatum* 25586 with NcoI cutsite into 16b
prAUT94	ggatccCTCGAGTTAatgatgatgatgatgatgTATACTCCTAATTTCATTTATTCCAACTTG	Reverse primer gene M.Fnu25V to clone from *F. nucleatum* 25586 with 6XHis-tag and XhoI cutsite into 16b
prAUT104	gatccCCATGGgaAGTTATAAAGAAAAGATACTAAGTTTATTAAATGAAAATTTG	Forward primer gene M.Fnu25I to clone from *F. nucleatum* 25586 with NcoI cutsite into 16b

prAUT105	gactccCTCGAGTTAatgatgatgatgatgatgATTTTTAATAAAGTCGTTTATTCTTGAATATGC	Reverse primer gene M.Fnu25I to clone from *F. nucleatum* 25586 with 6XHis-tag and XhoI cutsite into 16b
prDJSVT1068	CTGAGATCTAGAtttacagctagctcagtcc	Forward primer to clone 23726 Gene M.Fnu23I from synthetic Anderson-622-Bba into 16b with XbaI cutsite
prDJSVT1069	GACTCTCCATGGaaataataaaaaagccggattaataatctg	Reverse primer to clone 23726 M.Fnu23I(DNA meth trans) from synthetic Anderson-622-Bba into 16b with NcoI cutsite
prDJSVT1070	CTGAGACCATGGtttacagctagctcagtcc	Forward primer to clone 23726 M.Fnu23II(DNA meth trans) from synthetic Anderson-635-Bba into 16b. Has NcoI
prDJSVT1071	GACTCTCATATGaaataataaaaaagccggattaataatctg	Reverse primer to clone 23726 M.Fnu23II(DNA meth trans) from synthetic Anderson-635-Bba into 16b. Has NdeI
prDJSVT1072	CTGAGACATATGtttacagctagctcagtcc	Forward primer to clone 23726 hdsM (DNA meth trans) from synthetic Anderson-1581-Bba into 16b. Has NdeI
prDJSVT1073	GACTCTGGATCCaaataataaaaaagccggattaataatctg	Reverse primer to clone 23726 hdsM (DNA meth trans) from synthetic Anderson-1581-Bba into 16b. Has BamHI
prDJSVT1076	gcactacccatgggaAACTATATCGGCAGCAAATTAAGtTTAAAAAG	Forward primer to clone 23726 M.Fnu23I (DNA meth trans) into pET16b. Has NcoI.
prDJSVT1077	gctagtctcgagTTAgtgatggtgatggtgatgTTTCTTAATCAGGCAATGCAGATATTC	Reverse primer to clone 23726 M.Fnu23I(DNA meth trans) into pET16b. Has 6xHis and XhoI.
prDJSVT1078	gcactacccatgggaGAAAATCTGAAATTGTTCCTCGATG	Forward primer to clone 23726 M.Fnu23II (DNA meth trans) into pET16b. Has NcoI.
prDJSVT1079	gctagtctcgagTTAgtgatggtgatggtgatgCATTTCTTCGGTGTTGTAGTTGGTC	Reverse primer to clone 23726 M.Fnu23II (DNA meth trans) into pET16b. Has 6xHis and XhoI.

prDJSVT1080	gcactacccatgggaAGCAGTGTTTAC GATGTGGTCGATTATAAC	Forward primer to clone 23726 HdsM (DNA meth trans) into pET16b. Has NcoI.
prDJSVT1081	gctagtctcgagTTAgtgatggtgatggtgatg CTCATCATACAGCCCTTCTAAA ATATT	Reverse primer to clone 23726 G HdsM (DNA meth trans) into pET16b. Has 6xHis and XhoI.
prDJSVT1274	GCATCGccatgggaAATAAAAGTA GCTTAAAAATTTTTGCAATAGA AGCAAG	Forward primer to clone 23726 Gene_1553 (BREX pglD) into pET16b anderson (Med) and pET16b. Has NcoI.
prDJSVT1275	CGATGCggatccCTAatgatggtgatggt gatgCCTCTTTTTAAGAATAGTC CCCAATTTTTC	Reverse primer to clone 23726 Gene_1553 (BREX pglD) into pET16b anderson (Med) and pET16b. Has 6xHis and BamHI.
prDJSVT1276	GCATCGccatgggaAGTAGTGTGT ATGATGTGGTTGATTATAAC	Forward primer to clone 23726 Gene_1581_80 (hsdMS) into pET16b anderson (Med) and pET16b. Has NcoI.
prDJSVT1277	CGATGCggatccCTAatgatggtgatggt gatgATCCCCATAGTTCTGAACC TCCATTTC	Reverse primer to clone 23726 Gene_1581_80 (hsdMS) into pET16b anderson (Med) and pET16b. Has 6xHis and BamHI.
prDJSVT1278	GCATCGccatgggaAAAATAATAC ATGGAGATGTTCGAGAGAAAT TATGC	Forward primer to clone 1_1_36S M.Fnn13III into pET16b anderson (Med) and pET16b. Has NcoI.
prDSVT1279	CGATGCggatccTTAatgatggtgatggtg atgTTCTCCATTCAATTTTAATTT ATATTGTGTATTC	Reverse primer to clone 1_1_36S M.Fnn13III into pET16b anderson (Med) and pET16b. Has 6xHis and BamHI.
prDJSVT1280	GCATCGccatgggaGGTTTAATTA TAGATAATTTTGCGGGTGGAG	Forward primer to clone 1_1_36S M.Fnn13II into pET16b anderson (Med) and pET16b. Has NcoI.
prDJSVT1281	CGATGCggatccTTAatgatggtgatggtg atgATTAAAATTTGCTTTTACTA ATGCTTCTGC	Reverse primer to clone 1_1_36S M.Fnn13II into pET16b anderson (Med) and pET16b. Has 6xHis and BamHI.

prDJSVT1282	GCATCGccatgggaGGTAAAAAAC ATTATGTGATTTATACTCCGG	Forward primer to clone 1_1_36S BREX into pET16b anderson (Med) and pET16b. Has NcoI.
prDJSVT1283	CGATGCggatccTCAatgatggtgatggt gatgTATTTTGAAATAATTCTCT ATTTTTTGTTGTCTTTC	Reverse primer to clone 1_1_36S BREX into pET16b anderson (Med) and pET16b. Has 6xHis and BamHI.
prDJSVT1284	GCATCGccatgggaAGGAACATAG AAATTATAAAACTGAATTTAC AAG	Forward primer to clone 1_1_36S M.Fnn13I into pET16b anderson (Med) and pET16b. Has NcoI.
prDJSVT1285	CGATGCggatccTCAatgatggtgatggt gatgGAATTCTTCTTTGATGGCTT CCCAAG	Reverse primer to clone 1_1_36S M.Fnn13II into pET16b anderson (Med) and pET16b. Has 6xHis and BamHI
prDJSVT1286	GCATCGccatgggaGAAATAAATA AAATATTTAAAGGAGATAGTA TTGAAG	Forward primer to clone 7_1 M.Fnu71I into pET16b anderson (Med) and pET16b. Has NcoI.
prDJSVT1287	CGATGCggatccTTAatgatggtgatggtg atgTTTTTCATTATTAAACTCTTT AGCATATTGGTATC	Reverse primer to clone 7_1 M.Fnu71I into pET16b anderson (Med) and pET16b. Has 6xHis and BamHI.
prDJSVT1288	GCATCGccatgggaAATTATATAG GT TCAAAACTTTCTTTGAGACC	Forward primer to clone 7_1 M.Fnu71II into pET16b anderson (Med) and pET16b. Has NcoI.
prDJSVT1289	CGATGCggtaccTTAatgatggtgatggtg atgTTTTTTTATAAGGCAATGTA AATATTCAATTGTTG	Reverse primer to clone 7_1 M.Fnu71II into pET16b anderson (Med). Has 6xHis and KpnI.
prDJSVT1290	CGATGCctcgagTTAatgatggtgatggtg atgTTTTTTTATAAGGCAATGTA AATATTCAATTGTTG	Reverse primer to clone 7_1 M.Fnu71II into pET16b. Has 6xHis and XhoI.
prDJSVT1291	GCATCGccatgggaGATATAAATA AATTTATAAATACAATTATTGA GGGAG	Forward primer to clone 7_1 M.Fnu71III into pET16b anderson (Med) and pET16b. Has NcoI.
prDJSVT1292	CGATGCggatccTTAatgatggtgatggtg atgAAAATAGCTTTTTGTTCCCA TTTTTGAATATAC	Reverse primer to clone 7_1 M.Fnu71III into pET16b anderson (Med) and pET16b. Has 6xHis and BamHI.

prDJSVT1293	GCATCGccatgggaAATGAAATAATAGAATTAAATCTACTTAAAGAAGTAG	Forward primer to clone 7_1 M.Fnu71V into pET16b anderson (Med) and pET16b. Has NcoI.
prDJSVT1294	CGATGCggatccTTAatgatggtgatggtgatgTGTTACATCACCAGCATAGAGTTCATTTTTAAC	Reverse primer to clone 7_1 M.Fnu71V into pET16b anderson (Med) and pET16b. Has 6xHis and BamHI.
prDJSVT1295	GCATCGccatgggaAAAGAGTTAAAAATAATTAATAAAAATATAGATGACATAAAAG	Forward primer to clone 7_1 M.Fnu71VII into pET16b anderson (Med) and pET16b. Has NcoI.
prDJSVT1296	CGATGCggatccTTAatgatggtgatggtgatgATCATCTGTTATGCTATTAAAATCATCTACATC	Reverse primer to clone 7_1 M.Fnu71VII into pET16b anderson (Med) and pET16b. Has 6xHis and BamHI.
prDJSVT1297	GCATCGccatgggaGAACAAGTAAAGAAAGAGATTAAATGTGAG	Forward primer to clone 7_1 M.Fnu71VIII into pET16b anderson (Med) and pET16b. Has NcoI.
prDJSVT1298	CGATGCggatccTTAatgatggtgatggtgatgAAAATTAAAAGTCATTTGTGGGTCTTTATAATTAATC	Reverse primer to clone 7_1 M.Fnu71VIII into pET16b anderson (Med) and pET16b. Has 6xHis and BamHI.

Phylogenetic Tree construction

Full CRISPR-Cas phylogenetic analysis was created in Geneious[33] and adapted in Adobe Illustrator.

Protein sequence similarity networks

Protein networks were generated by utilizing all the identified methyltransferases as described previously. The methyltransferases from six strains additionally examined utilizing the “all-by-all BLAST” and “sequence-BLAST” feature available in the EFI-EST server [34], and completed by creating a hierarchical map in Cytoscape 3.8 [35].

Cloning of DNA Methyltransferases

The primers to clone the DNA methyltransferases were all ordered from Integrated DNA Technologies (IDT). All the plasmids utilized in these studies are described in **Table 2.1** along with the bacterial strains in **Table 2.2** and primers in **Table 2.3**. The genomic DNA used to create all recombinant methyltransferases constructs for expression from *F. nucleatum* ATCC 23726, *F. nucleatum* ATCC 25586, *F. nucleatum* subsp. *animalis* 7_1, and *F. necrophorum* subsp. *funduliforme* 1_1_36S was prepared utilizing the Wizard Genomic DNA Purification Kit (Promega). PCR products were centrifuged and purified utilizing a PCR purification kit (Biobasic) and digested for 2 hours at 37 °C along with pET16b which was used as the expression vector and was obtained through EZ-10 Spin Column Plasmid Miniprep (Biobasic) with the restriction enzymes listed in **Table 2.3**. The vector was then incubated with FastAP (Thermo Fisher Scientific) for 1 hour at 37 °C. Digested products were purified utilizing a spin column and ligated by T4 DNA ligase (New England Biolabs) for 1 hour at room temperature following manufacturer's recommendations. The ligations were utilized to transform into competent Mix&Go! (Zymo Research) Top10 *E. coli* and plated into LB solid agar plates supplemented with 100 μg/mL carbenicillin (ampicillin). Confirmation of positive clones was performed by digestion and if applicable positive clones were then transformed into ARTIC(DE3) RIL and LOBSTR-BL21(DE3) RIL.

Purification of Type II Methyltransferases

Briefly, M.Fnu23I and M.Fnu23II in ARTIC(DE3) RIL *E. coli* cells were grown in LB (15g/L NaCl, 15 g/L tryptone, 10g/L yeast extract) medium at 37°C, 250 rpm shaking until OD=0.6. At OD=0.6. cells were induced with 50 μM Isopropyl β- d-1-thiogalactopyranoside (IPTG)

(GoldBio). Expression was carried out at 8°C and cells were collected at 20 hours after inoculation by centrifugation at 5000×g for 20 min at 4°C. Bacterial pellets were resuspended in a lysis buffer (20 mM Tris, pH 7.5, 400 mM NaCl, 20 mM imidazole). Bacteria were lysed by an EmulsiFlex-C3 homogenizer (Avestin, Mannheim, Germany) for 10 mins at 10,000-15,000 kPa. Afterwards, the unlysed cells and insoluble material was separated by centrifugation at 15,000×g for 20 minutes at 4°C and then discarded. The supernatant containing the methyltransferases His6-tagged was stirred with 6 mL of $NiCl_2$ charged chelating Sepharose beads (GE Healthcare) for 30 minutes at 4°C. The column was washed with 400 mLs of wash buffer (20 mM Tris, pH 7.5, 400 mM NaCl, 40 mM imidazole). After washing, the methyltransferases were eluted in 10 mL of elution buffer (20 mM Tris, pH 7.5, 400 mM NaCl, 250 mM imidazole). The purified protein was then directly put into dialysis in a buffer (20 mM Tris, pH 7.5, 150 mM NaCl, 10% glycerol). Protein concentrations were calculated using a Qubit fluorometer and BCA assays based on the manufacturer's recommended protocol.

Purification of *F. nucleatum* ATCC 25586 Methyltransferases

Briefly, M.Fnu25I, M.Fnu25IV and M.Fnu25V in LOBSTR-BL21(DE3) RIL *E. coli* cells were grown in LB (15g/L NaCl, 15 g/L tryptone, 10g/L yeast extract) medium at 37°C, 250 rpm shaking until 0D=0.6. At OD=0.6. cells were induced with 50 μM Isopropyl β- d-1-thiogalactopyranoside (IPTG) (GoldBio). Expression was carried at 16°C and cells were collected at 20 hours after inoculation by centrifugation at 5000×g for 20 min at 4°C. The purification proceeded as stated in the previous purification protocol for Type II methyltransferases.

In vitro treatment of plasmid DNA with *Fusobacterium* DNA Methyltransferases

Plasmid DNA (35-40 μg), prepared from *E.coli* TOP10 using the EZ-10 Spin column plasmid DNA mini-prep from Biobasic, was combined in a 30 uL reaction with 160 μM SAM (New England Biolabs), 1X cutsmart buffer (New England Biolabs) and 1 μM of M.Fnu23I or M.Fnu23II. The reaction mixes were incubated at 37°C for 2 hours and then extracted by adding 1 volume of Phenol:Chloroform:Isoamyl Alcohol, 25:24:1 Mixture (bioWORLD) vortexed for 20 seconds and centrifuged at 16,000×g for 5 minutes. Plasmid DNA was precipitated with ethanol and dissolved in ultrapure water (bioWORLD).

Coexpression of plasmid DNA with *Fusobacterium* DNA Methyltransferases

Using the expression vector (constitutive activity) pET16b with the DNA methyltransferase under an anderson medium promoter as described in **Table 2.1**, we methylated pDJSVT1 *in-vivo*. Both pET16b (Gene 622 and Gene 635) and pDJSVT13 were transformed into *E.coli* Top10 and grown in LB (15 g/L NaCl, 15 g/L tryptone, 10g/L yeast extract) medium at 37°C, 250 rpm shaking for 24 hours.

Ultrapure DNA purification

Plasmid DNA was purified from overnight expression or co-expression was isolated with an alkaline lysis/column purification technique using the EZ-10 Spin Column Plasmid Miniprep (Biobasic). Plasmid DNA was further purified for use in electroporation using 3M Sodium Acetate (pH 5.5) and glycogen, followed by ethanol precipitation at -80°C. After 3 hours minimum of incubation at -80°C sample was spun at 4°C for 30 minutes at 16,000×g to pellet the DNA and washed five times with 70% ethanol carefully by spinning at 14,000×g for 3 mins. Pellet was then dried at room temperature for 10–15 minutes. Finally, 15μL of ultrapure free water (BioBasic) was

added and incubated at 37°C for 30 minutes-1 hour to solubilize the pellet. Plasmid DNA was quantified using Nanodrop.

Protection Assay

Plasmid DNA (1 μg) prepared from *E.coli* TOP10 strain using the Biobasic mini-preparation procedure, was combined with Cutsmart buffer (New England Biolabs, 50 mM Potassium Acetate, 20 mM Tris-acetate, 10 mM Magnesium Acetate, 100 μg/mL BSA (pH 7.9)), 160 μM SAM (New England Biolabs), and 1 μM of the correspondent methyltransferases. As a control plasmid DNA (1 μg) was mock treated in a reaction without the methyltransferases. All samples were incubated 1 hour at 37°C with the restriction enzymes, single cutters KpnI and MluI or predicted restriction sites NlaIII and SfaNI (New England Biolabs).

Linearization of plasmid DNA

Ultrapure plasmid DNA was digested with restriction enzyme KpnI (single cut in the plasmid) following manufacturer's instructions (New England Biolabs). After two hours at 37 °C the ultrapure DNA underwent phenol chloroform extraction by adding one volume of phenol:chloroform:isoamyl alcohol, 25:24:1 (bioWORLD). The reaction mix was vortexed for 20 seconds. Mixture was then centrifuged at room temperature for 5 minutes at 16,000×g. The upper aqueous phase was removed and transferred to a clean 1.7mL tube for ethanol precipitation at -80°C as described previously for ultrapure DNA purification. The ultrapure DNA digests were run on 1% agarose gel at 105 V for 50 minutes to determine successful digestion before continuing to electroporation into *F. nucleatum* ATCC 23726.

Competent *Fusobacterium* and electroporation

All *Fusobacterium* strains were competently prepared by inoculating and growing a 100-mL anaerobic culture in CBHK media to lag phase (A_{600} = 0.1) followed by centrifugation of bacteria at 3200×g for 10 minutes. The supernatant was removed and the resulting pellet was washed three successive times utilizing 1-mL of ice-cold 20% glycerol in deionized H_2O and 1mM MOPS at 14,000×g for 3 minutes. Bacterial pellet was then resuspended in a final volume of 80 µL of ice-cold 20% glycerol and 1mM MOPS. Bacteria were transferred to cold 1-mm (Lonza) or 2mm (Bio-Rad) electroporation cuvettes, and 500 ng or 3 µg (concentration >300 ng/µL) of pDJSVT1 plasmid was added before electroporating at 2.5 kV/cm, 50 µF, 360Ω, using a BTX Electro Cell Manipulator 600 (Harvard Apparatus). The electroporated cells were promptly transferred by syringe into a sterile, anaerobic tube with 4 mL of recovery medium (CBHK, 1 mM $MgCl_2$) and incubated at 37 °C for 20 h with no shaking in an anaerobic chamber. After the recovery outgrowth, cells were centrifuged at 14,000×g for 3 minutes, supernatant was removed, and pellet cells were resuspended in 0.2 mL of recovery medium. Resuspension was plated on CBHK plates with 5 µg/mL thiamphenicol and incubated in an anaerobic 37 °C incubator for 2 days for colony growth. The transformation efficiency represents the number of thiamphenicol or streptomycin resistant colonies per total number of bacteria per microgram of DNA. Electroporation was conducted in triplicate.

Results

Bioinformatic analysis of DNA methyltransferases in *Fusobacterium*

Since other bacteria have shown that mimicry by methylation improves transformation efficiency we utilized the online databases FusoPortal, NCBI and REBASE to identify R-M systems in *Fusobacterium*. Overall, a total of 35 R-M systems (17.14% Type I, 74.28% Type II and 8.57% Type III) were identified in six strains as shown in **Figure 2.3A.** An average of five R-M systems per strain was determined. However, *F. nucleatum* subsp. *animalis* 7_1 was identified as an outlier as it encodes for a total of 12 R-M systems (two Type I and ten Type II).

The *F. nucleatum* ATCC 23726 genome from FusoPortal encodes four R-M systems as shown in **Figure 2.4**; one Type I and three Type II, when compared to the previous imcomplete build of this genome in which only the Type I R-M system was identified. Most of *F. nucleatum* strains harbor the classical Type I R-M composed of three genes encoding for *hsdR, hsdM* and *hsdS* (*hsdRMS* operon), except for *F. nucleatum* ATCC 25586 that encodes for five R-M systems; three Type II and two Type III as detailed in **Figure 2.4**.

The most common type of R-M systems in the six analyzed strains of *Fusobacterium* is identified as Type II. The majority of these were orthodox Type II R-M systems, which include two independent proteins, a methyltransferase and a restriction endonuclease. In addition, the presence of orphan methyltransferases was characterized in all six strains. The bioinformatic results also confirmed the presence of the Type II BREX system in several *Fusobacterium* strains. The BREX system is generally composed of a 4-8 gene cluster [20] and in *Fusobacterium* is predicted to methylate adenine residues similar to *E. coli* [36]. However, the restriction site is yet to be experimentally characterized.

Only a total of two Type III R-M systems were bioinformatically characterized in *F. nucleatum* ATCC 25586 and one in *F. nucleatum* subsp. *animalis* 4_8 even though they are widespread in most sequenced bacterial genomes [37]. Finally, no Type IV R-M systems were discovered in the six *Fusobacterium* strains analyzed in this study.

A diverse set of methyltransferases **(Table 2.4)** in four strains, *F. nucleatum* ATCC 23726, *F. nucleatum* ATCC 25586, *F. nucleatum* subsp. *animalis* 7_1 and *F. necrophorum* subsp. *funduliforme* 1_1_36S were further characterized. Utilizing REBASE the predicted recognition sites as shown in **Table 2.4** were identified. All of them were bioinformatically characterized to be adenine DNA methyltransferases, where methylation occurs at the N6 position of the adenine (6mA).

Phylogenetic and protein sequence similarity network analysis of the methyltransferases from six *Fusobacterium* species

A phylogenetic analysis of the 17 methyltransferases chosen to treat plasmid DNA *in-vitro* or by coexpression revealed conservation within this group of methyltransferases as shown in **Figure 2.5.** Interestingly, there are two major branches in the phylogenetic tree, where the most distant branch contains to M.Fnu23I which appears to be closely related to M.Fnu25II and M.Fnu71II (**Figure 2.5)** due to the high sequence homology shared between these adenine methyltransferases. The second branch of the phylogenetic tree comprises all of the other Type I, Type II and, Type III methyltransferases. In this branch, the BREX system from *F. necrophorum* subsp. *funduliforme* 1_1_36S and *F. nucleatum* ATCC 23726 show conservation as well as the Type III R-M systems from *F. nucleatum* ATCC 25586 which are closely related as detailed in **Figure 2.5**.

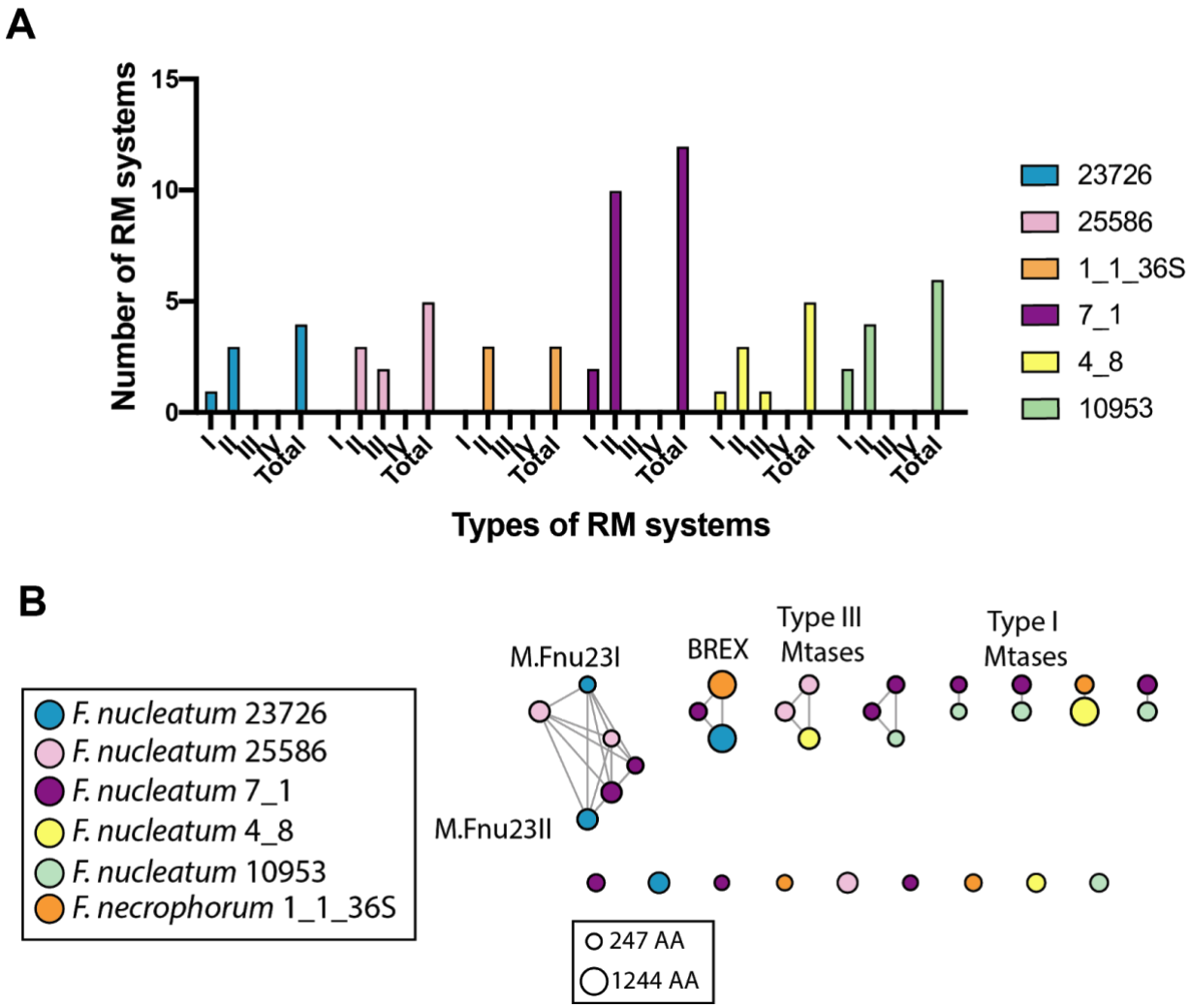


Figure 2-3 R-M systems in a total of six Fusobacterium strains were identified. A. Distribution of R-M systems in six *Fusobacterium* strains. **B.** Sequence similarity network of all methyltransferases identified in the six *Fusobacterium* strains (E value 10^{-125}) as colored by designated genomes. Each node represents a single protein (34 total) and the node size represents the protein size in aminoacids.

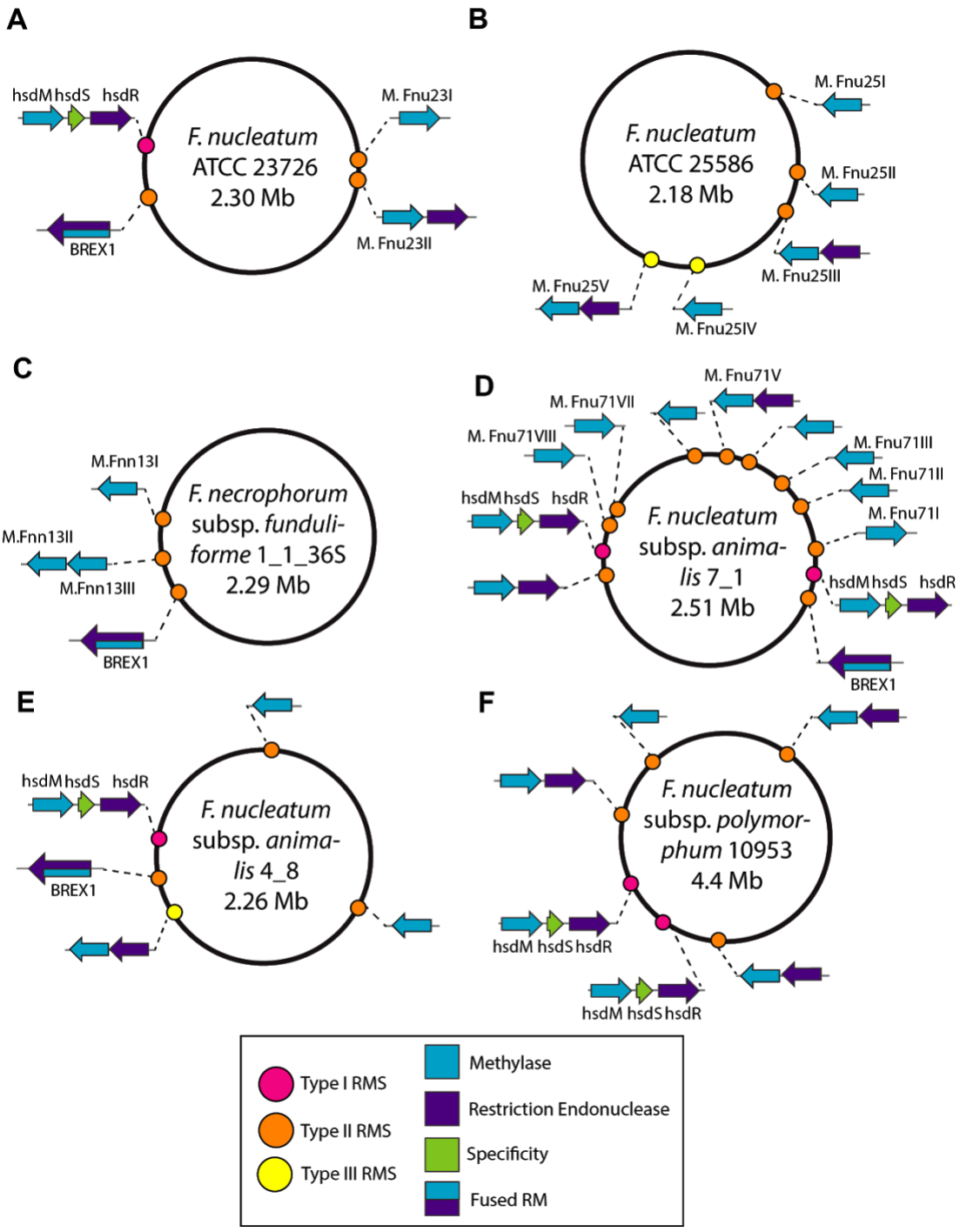


Figure 2-4 Genome maps detailing R-M systems in six *Fusobacterium* strains. A. Genome map of *F. nucleatum* ATCC 23726. **B.** Genome map of F. nucleatum ATCC 25586. **C.** Genome map of *F. necrophorum* subsp. *funduliforme* 1_1_36S. **D.** Genome map of *F. nucleatum* subsp. *animalis* 7_1. **E.** Genome map of *F. nucleatum* subsp. *animalis* 4_8. **F.** Genome map of *F. nucleatum* subsp. *polymorphum* 10953.

Table 2.4 Methyltransferase genes identified in *Fusobacterium* strains utilized for host-mimicking techniques

Methyltransferase name	*Fusobacterium* strain	FusoPortal ID: Size (AA)	Type R-M system	Predicted restriction site
hdSMR	23726	Gene_1581: 725	I	ATTYNNNNNCTC
M.Fnu23I	23726	Gene_622: 333	II	CATG
M.Fnu23II	23726	Gene_635: 670	II	GCATC
BREX	23726	Gene_1553: 1215	II	-
M.Fnu25I	25586	Gene_458: 690	I	CCNNNNNNNGG
M.Fnu25IV	25586	Gene_967: 532	III	CAGCAG
M.Fnu25V	25586	Gene_972: 641	III	CAGCAG
M.Fnn13I	1 1 36S	Gene_587: 438	II	CAAAAA
M.Fnn13II	1_1_36S	Gene_1324: 416	II	GCTAGC
M.Fnn13III	1_1_36S	Gene_1325: 311	II	CCAG
BREX	1_1_36S	Gene_1629: 481	II	-
M.Fnu71I	7_1	Gene_58: 357	II	-
M.Fnu71II	7_1	Gene_400: 333	II	CATG
M.Fnu71II	7_1	Gene_434: 370	II	GANTC
M.Fnu71V	7_1	Gene_801: 438	II	-
M.Fnu71VII	7_1	Gene_2128: 441	II	-
M.Fnu71VIII	7_1	Gene_2137: 261	II	-

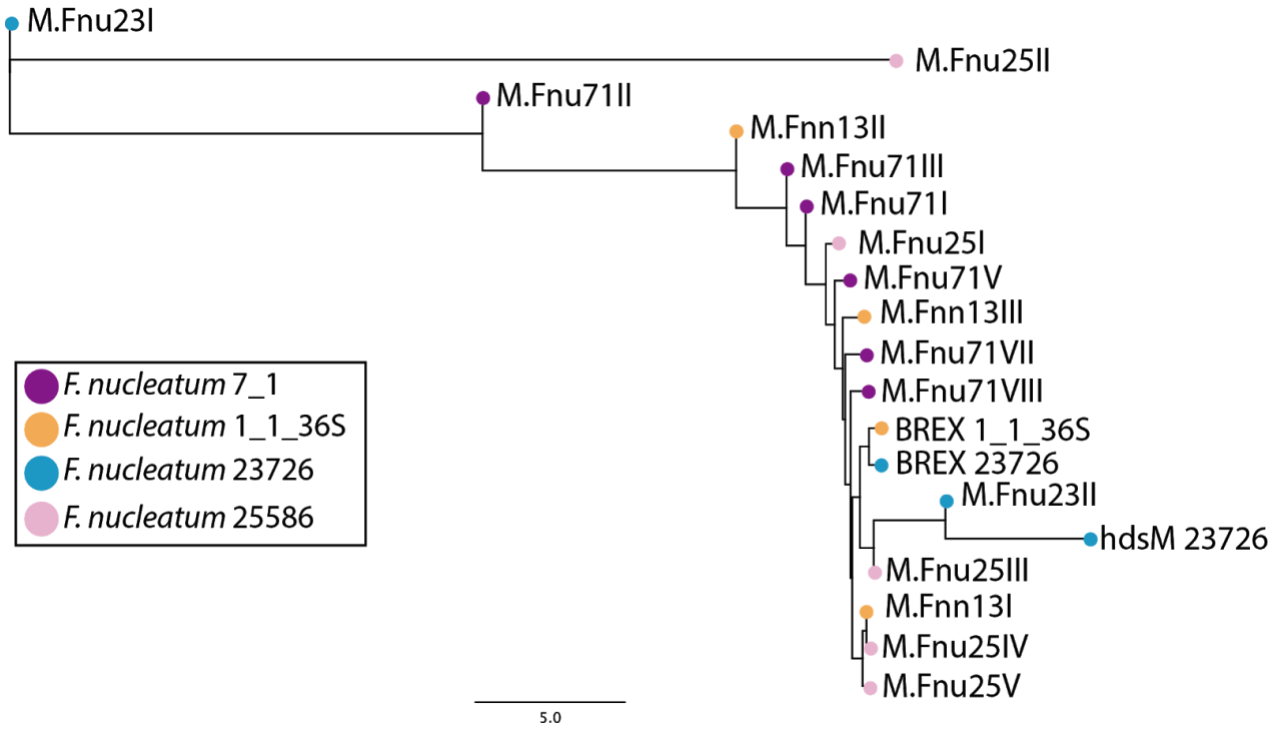


Figure 2-5 Phylogenetic analysis of methyltransferases from four *Fusobacterium* strains. The analyzed methyltransferases were utilized for host-mimicking strategies and phylogenetic tree was constructed in Geneious and adapted in Adobe Illustrator.

A closer analysis of all the methyltransferases from the six *Fusobacterium* strains in a protein sequence similarity network revealed distinct subsets of predicted methyltransferases as shown in **Figure 2.3B**. A Type II R-M system subset was identified with the characterized M.Fnu23I and M.Fnu23II along with their closely related homologs from *F. nucleatum* ATCC 25586 and *F. nucleatum* subsp. *animalis* 7_1. Moreover, the BREX system was also identified again as a subset, along with a different set of Type II methyltransferases yet to be characterized. The Type III methyltransferases also comprised a subset including the two identified in *F. nucleatum* ATCC 25586 and one in *F. nucleatum* subsp. *animalis* 4_8.

Finally, to investigate whether M.Fnu23I and M.Fnu23II were conserved in the *Fusobacterium* genus we created a protein sequence similarity network utilizing the sequence-BLAST feature in EFI-EST [34]. The implementation of this search resulted in the identification of the closest homologues to M.Fnu23I and M.Fnu23II. M.Fnu23I is highly conserved in the *Fusobacterium* genus and related to bacteria in the *Clostridium* genus as shown in **Figure 2.6A** whereas M.Fnu23II shows less conservation and multiple homologues were identified in Gram-negative and Gram-positive microorganisms **(Figure 2.6B)**.

Experimental validation of the recognition sequences for Type II methyltransferases in *F. nucleatum* ATCC 23726

The Type II R-M methyltransferases from *F. nucleatum* ATCC 23726 were experimentally investigated and recombinantly expressed and purified **(Figure 2.7A)**. Recombinant proteins were tested for their methyltransferase activity by performing a methyltransferase protection assay. We showed that the shuttle vector is completely sensitive to digestion by the restriction endonuclease NlaIII. Meanwhile, mock treated plasmid DNA with M.Fnu23I, with a bioinformatic predicted

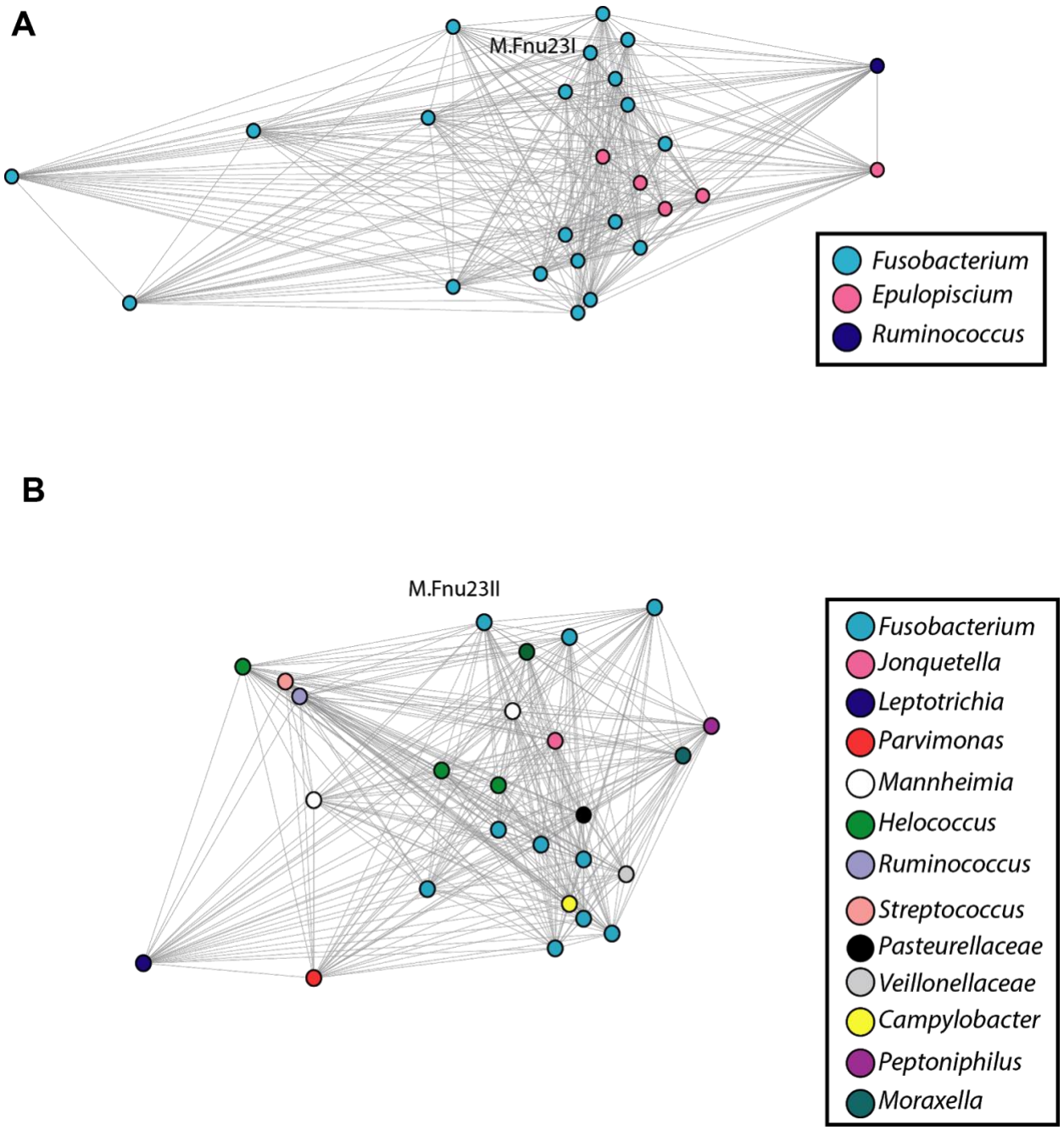


Figure 2-6 Sequence similarity networks of M.Fnu23I and M.Fnu23II. A. Sequence similarity network of M.Fnu23I (E value 10^{-125}) as colored by designated genomes. Each node represents a single protein (25 total). **B.** Sequence similarity network of M.Fnu23II (E value 10^{-125}) as colored by designated genomes. Each node represents a single protein (25 total).

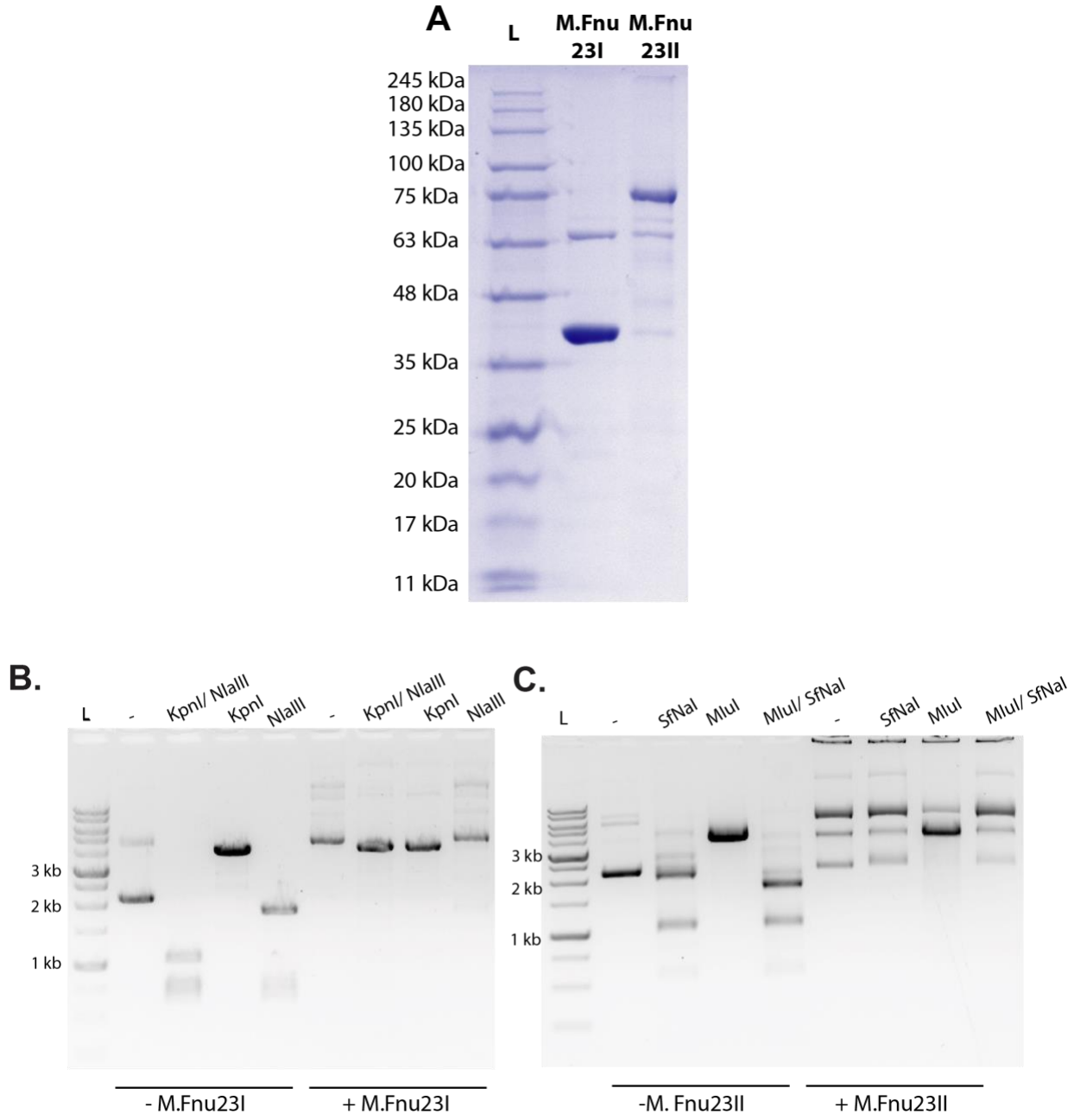


Figure 2-7 Restriction site identification of Type II methyltransferases in *F. nucleatum* 23726. A. SDS-polyacrylamide gel of purified M.Fnu23I and M.Fnu23II**. B.** M.Fnu23I methylates CATG. Plasmid DNA is cleaved when digested with KpnI and NlaIII. When conditioned with M.Fnu23I plasmid is protected from digestion. **C.** M.Fnu23II methylates GCATC. Plasmid DNA is cleaved when digested with MluI and SfaNI. When conditioned with M.Fnu23II plasmid is protected from digestion.

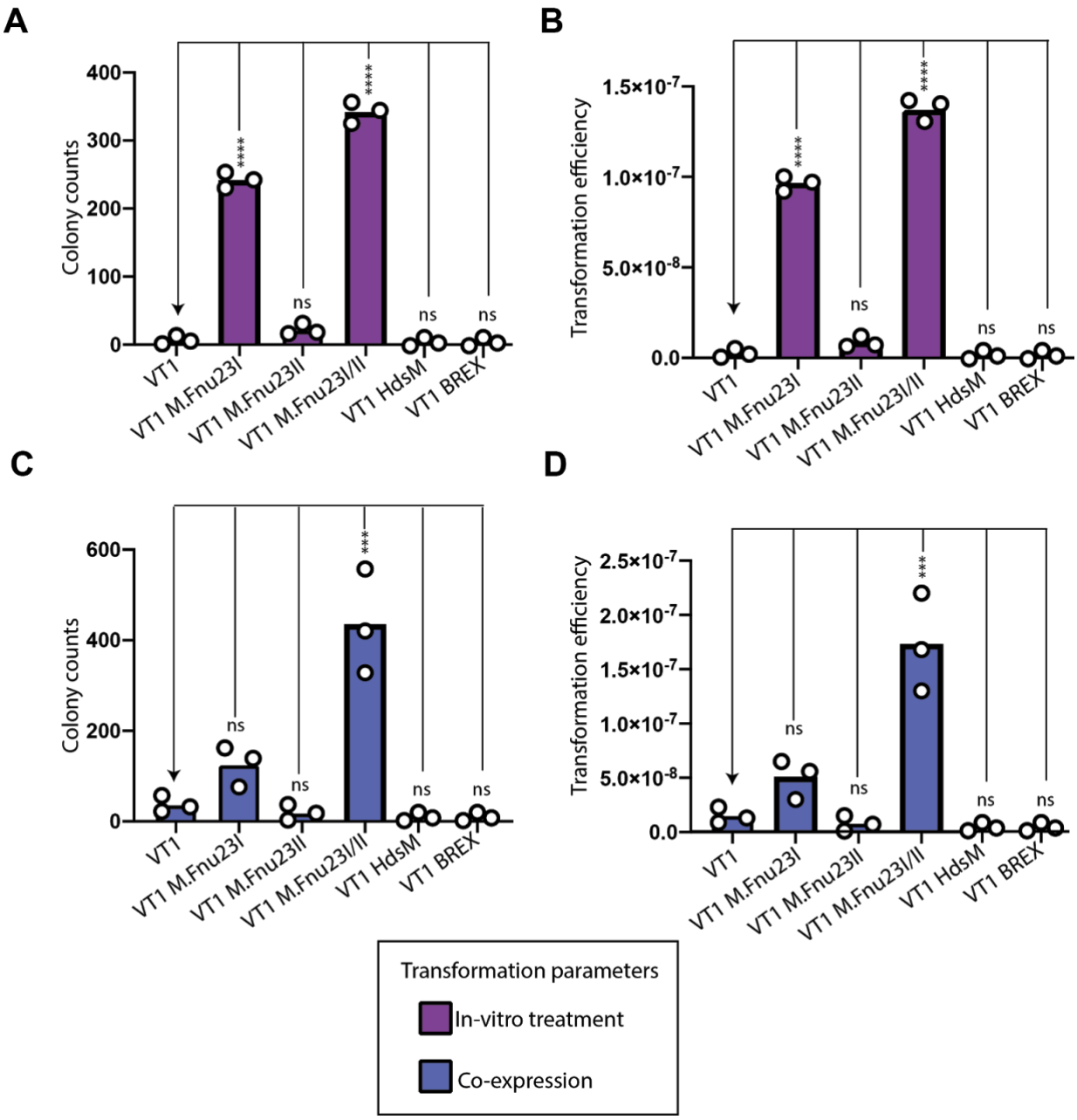


Figure 2-8 Transformation efficiency in *F. nucleatum* ATCC 23726 improves when protecting plasmid DNA with methyltransferases. A. Chloramphenicol resistant transformants were counted from electroporation with 3 μg pDNA treated with either 1μM M.Fnu23I, 1μM M.Fnu23II, 1μM M.Fnu23I/M.Fnu23II, 1μM HdsM or 1μM BREX system. **B.** Transformation efficiency calculated from colony counts of electroporation with in vitro treated pDNA with either 1μM M.Fnu23I, 1μM M.Fnu23II, 1μM M.Fnu23I/M.Fnu23II, 1μM HdsM or 1μM BREX system. **C.** Chloramphenicol resistant transformants were counted from electroporation with 3 μg pDNA coexpressed with either M.Fnu23I, M.Fnu23II, M.Fnu23I/M.Fnu23II, HdsM or BREX system. **D.** Transformation efficiency calculated from colony counts of electroporation with coexpressed pDNA with either M.Fnu23I, M.Fnu23II, M.Fnu23I/M.Fnu23II, HdsM or BREX system.

restriction site of CATG, is resistant to NlaIII digestion, which recognizes the sequence CATG (**Figure 2.7B**).

In a similar manner, the plasmid DNA is sensitive to digestion by SfaNI, a restriction enzyme that recognizes GCATCN (5), before conditioning with M.Fnu23II, which has a bioinformatically GCATC predicted restriction site. After M.Fnu23II treatment, SfaNI digestion is inhibited **(Figure 2.7C)**. Altogether these results confirm the recognition sequences for M.Fnu23I as CATG, and GCATC for M.Fnu23II. Additionally, we identified that the methyl donor S-adenosyl-L-methione (SAM) was a requirement for DNA methylation. However, it is possible that the recognition sequence is a subset (eg CATGN) or variation of each of these sites as we did not perform a methylome analysis.

Effect of plasmid methylation on the efficiency of *F. nucleatum* ATCC 23726 transformation and genome incorporation

The previous results show that plasmid DNA was successfully methylated *in vitro* and protected from cleavage. Therefore, we expanded our experiments to in *vitro* host mimicry as we anticipated that the conditioned plasmids would acquire the same pattern of methylation as genomic DNA characteristic to *F. nucleatum* ATCC 27326. As *F. nucleatum* ATCC 23726 has been previously transformed with a lower efficiency [2–5] our goal was to increase the transformation efficiency by bypassing R-M systems using mimicry by methylation.

For that reason, an *in vitro* and a coexpression method for conditioning plasmid DNA that homologously recombines with methyltransferases were created. Utilizing both approaches in *F. nucleatum* ATCC 23726 the results show a statistically significant increase transformation efficiency when plasmid DNA is conditioned with M.Fnu23I and M.Fnu23II. However, when

plasmid DNA is treated with the hsdM and BREX system enzymes there was no difference in the transformation efficiency of this strain **(Figure 2.8**). The combination of the methyltransferases M.Fnu23I and M.Fnu23II for conditioning plasmid DNA in *F. nucleatum* ATCC 23726 demonstrated and additive effect that helps circumvent of R-M systems and this is validated by the statistically significant transformation efficiency as a 50-fold improvement is achieved when compared to untreated plasmid **(Figure 2.8**).

To demonstrate that the increased transformation efficiency of plasmid DNA when treated with the methyltransferases was related to the specific protection provided by M.Fnu23I and M.Fnu23II against endonucleases, we transformed *E.coli* ER2796 a DNA-methyltransferase deficient strain with our coexpression plasmid DNA. Transformation with the coexpressed M.Fnu23I methyltransferase both in *E.coli* TOP10 (Dam and Dcm positive) and ER2796 (Dam and Dcm negative) results in a higher number of colonies **(Figure 2.9)** when compared to untreated plasmid DNA. Moreover, because *E.coli* TOP10 codes for its own Type II methyltransferases we also verified whether its native methylation would affect the efficiency of plasmid DNA transformation. However, we show that the difference between transformation efficiencies between *E.coli* TOP10 and a deficient methyltransferases strain ER2796 is not statistically significant **(Figure 9).**

Host-mimicking strategies in *Fusobacterium* strains

Since the transformation efficiency for *F. nucleatum* ATCC 23726 was improved by conditioning plasmid previous electroporation, the next step was to expand this set of molecular tools to the other five *Fusobacterium* strains included in this study. Initially we attempted to genetically manipulate *F. nucleatum* ATCC 25586, a recalcitrant strain, utilizing M.Fnu23I and M.Fnu23II which are homlogus to M.Fnu25II and M.Fnu25III, respectively. Utilizing our *in vitro* and

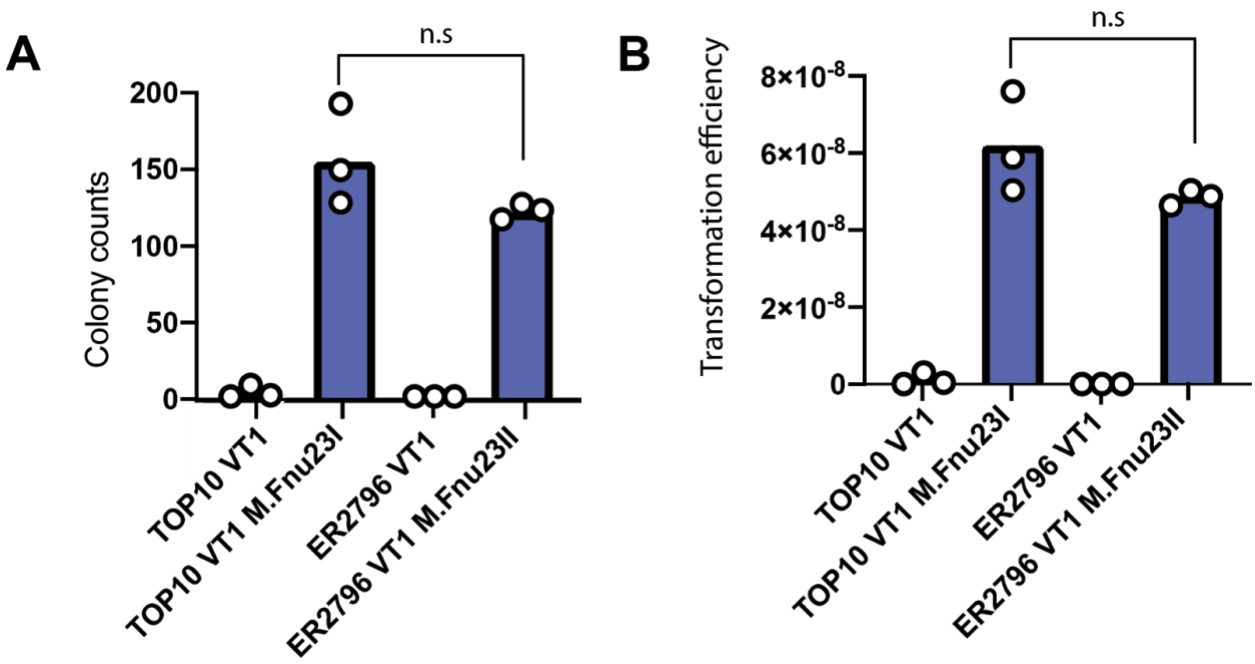


Figure 2-9 Increased transformation efficiency in *F. nucleatum* ATCC 23726 is accomplished by Type II methyltransferase conditioning of plasmid DNA. A. Chloramphenicol resistant transformants were counted from electroporation with 3 µg pDNA coexpressed with M.Fnu23I in either *E.coli* TOP10 or E.coli ER2796. **B.** Transformation efficiency calculated from colony counts of electroporation with coexpressed pDNA with M.Fnu23I in *E.coli* TOP10 or E.coli ER2796.

coexpression approach, the conditioning of plasmid DNA with these enzymes yielded no thiamphenicol resistant transformants in *F. nucleatum* 25586. To further achieve plasmid DNA protection from native restriction endonuclease, the purified Type II methyltransferases M.Fnu23I, M.Fnu23II and M.Fnu25I were utilized to condition plasmid DNA, but again no transformants were recovered. For this reason, we purified the Type III methyltransferases M.Fnu25I, M.Fnu25IV and M.Fnu25V to *in vitro* treat plasmid DNA with all the R-M systems encoded by *F. nucleatum* ATCC 25586. Nonetheless, plasmid DNA treated with all predicted DNA methyltransferases from R-M systems characterized in *F. nucleatum* ATCC 25586 yielded no colonies and no improvement in its transformation efficiency.

Utilizing the coexpression approach we aimed to condition plasmid DNA for *F. nucleatum* subsp. *animalis* 7_1 and *F. necrophorum* subsp. *funduliforme* 1_1_36S with their corresponding enzymes, listed in **Table 2.4**. Additionally, we extended these *in vitro* experiments to *F. nucleatum* subsp. *animalis* 4_8 and *F. nucleatum* subsp. *polymorphum* 10953, treating the pDNA with *F. nucleatum* subsp. *animalis* 7_1 and *F. nucleatum* ATCC 232726 methyltransferases. Another *Fusobacterium* species was also included in these studies, we attempted to trespass the physical barrier of *F. necrophorum* subsp. *funduliforme* 51357 by electroporating with pDNA conditioned with *F. necrophorum* subsp. *funduliforme* 1_1_36S methyltransferases. Once again the plasmid conditioning with methyltransferases didn't yield colonies for either strain. Finally, to try to bypass R-M systems in *Fusobacterium* strains we prepared crude extract of the target bacterium for DNA methylation. However, no thiamphenicol transformants were recovered by this method either. These results show that we still have a lot to learn about how to bypass R-M systems to enhance genetic modification in *Fusobacterium.*

Linearizing DNA does not increase transformation efficiency in *F. nucleatum* ATCC 23726

To identify if linearized plasmid DNA increased the transformation efficiency in *F. nucleatum* ATCC 23726, the non-treated plasmid DNA was digested utilizing a single cut restriction enzyme in non-conditioned plasmid DNA. This assay was carried out utilizing naked plasmid and a diminished number of colonies was obtained when compared to methyltransferase treated plasmid. It is noteworthy that no transformants were recovered in two independent transformations with different times of incubation of plasmid DNA and competent *F. nucleatum* ATCC 23726 (0 minutes and 30 minutes) with linearized plasmid **(Figure 2.10)**.

Deletion of CATG recognition sequence in plasmid DNA increases transformation efficiency

To further characterize how *F. nucleatum* ATCC 23726 discerns methylated plasmid DNA, we successfully deleted all four CATG sites present in pDJSVT1 by the quick change directed mutagenesis. method We identified that electroporated ΔCATG plasmid DNA yields a significantly higher number of thiamphenicol transformants when compared to untreated plasmid DNA as shown in **Figure 2.11.** As previously shown, when plasmid DNA is treated with M.Fnu23I a higher transformation efficiency in *F. nucleatum* ATCC 23726 is achieved by protecting the CATG motif. However, when the CATG motif is deleted a similar transformation efficiency is observed when the ΔCATG plasmid DNA is non-treated or treated with either M.Fnu23I, M.Fnu23II and M.Fnu23I/M.Fnu23II **(Figure 2.11)**. Altogether, these results evidence differential protein expression levels for the coexpression approach when utilizing M.Fnu23I and M.Fnu23II, as the transformation efficiency is enhanced when utilizing both enzymes but not individually.

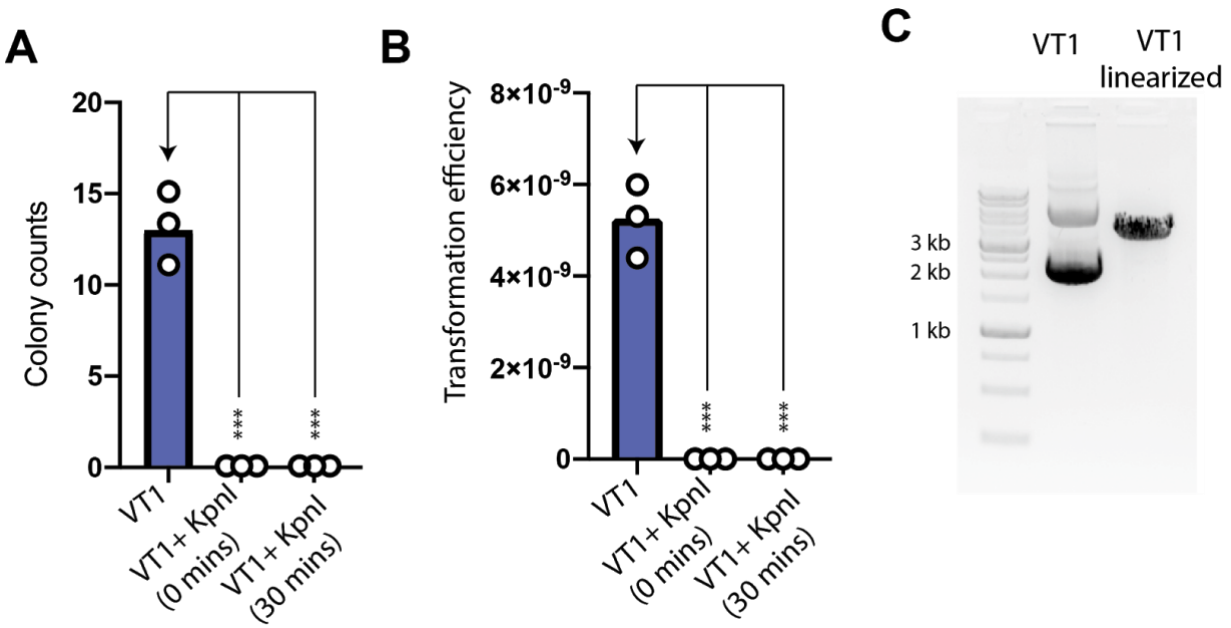


Figure 2-10 Transformation efficiency of *F. nucleatum* ATCC 23726 does not increase when linearizing plasmid DNA for homologous recombination before transformation. A. Chloramphenicol resistant transformants were counted from electroporation with 3 µg pDNA linearized and incubated for different periods before electroporation. **B.** Transformation efficiency calculated from colony counts of electroporation with 3 µg pDNA linearized and incubated for different periods before electroporation. **C.** Circular pDNA and digested VT1 with restriction endonuclease KpnI before electroporation.

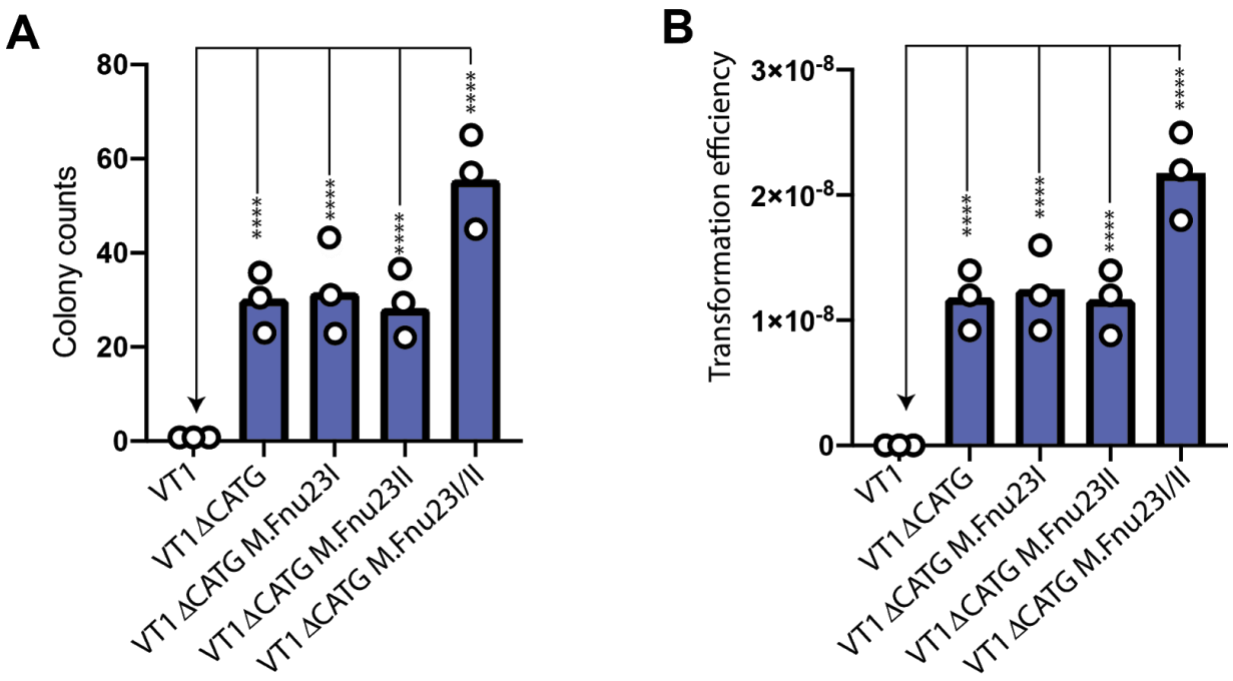


Figure 2-11 Transformation efficiency of *F. nucleatum* ATCC 23726 is influenced by the CATG motif. A. Chloramphenicol resistant transformants were counted from electroporation with 3 µg ΔCATGpDNA coexpressed with either M.Fnu23I, M.Fnu23II and M.Fnu23I/M.Fnu23II. **B.** Transformation efficiency calculated from colony counts of electroporation with 3 µg ΔCATGpDNA coexpressed with either M.Fnu23I, M.Fnu23II and M.Fnu23I/M.Fnu23II.

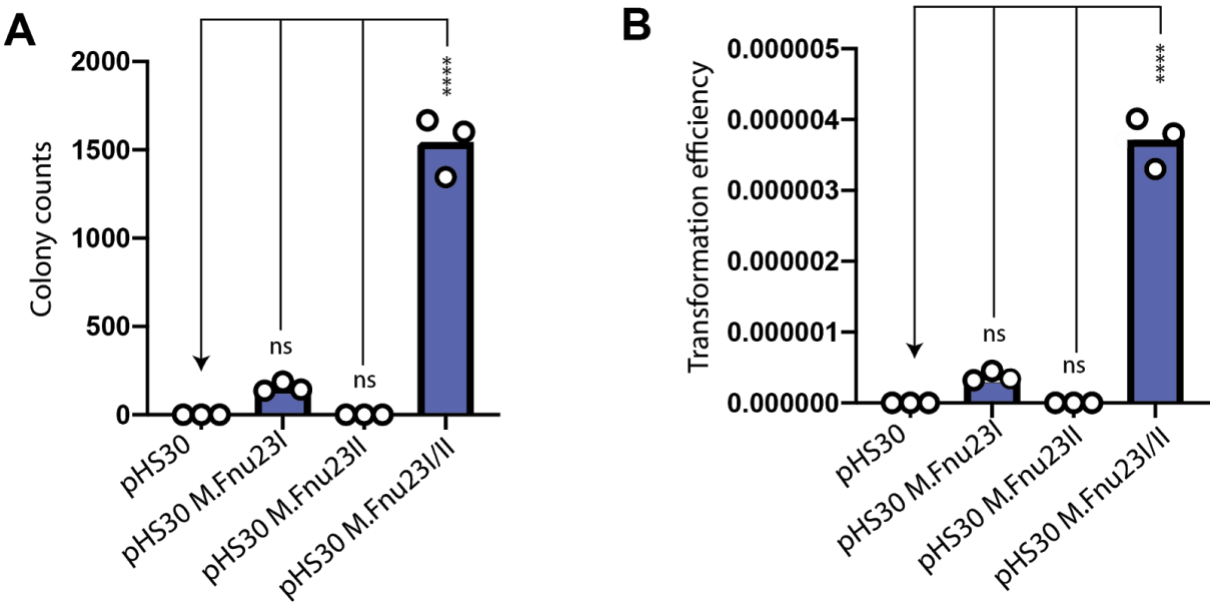


Figure 2-12 Transformation efficiency of the episomal plasmid pHS30 in F. nucleatum ATCC 23726 increased when conditioned with Type II methyltransferases. A. Chloramphenicol resistant transformants were counted from electroporation with 500 ng pHS30 coexpressed with either M.Fnu23I, M.Fnu23II and M.Fnu23I/M.Fnu23II. **B.** Transformation efficiency calculated from colony counts of electroporation with 500 ng pHS30 coexpressed with either M.Fnu23I, M.Fnu23II and M.Fnu23I/M.Fnu23II.

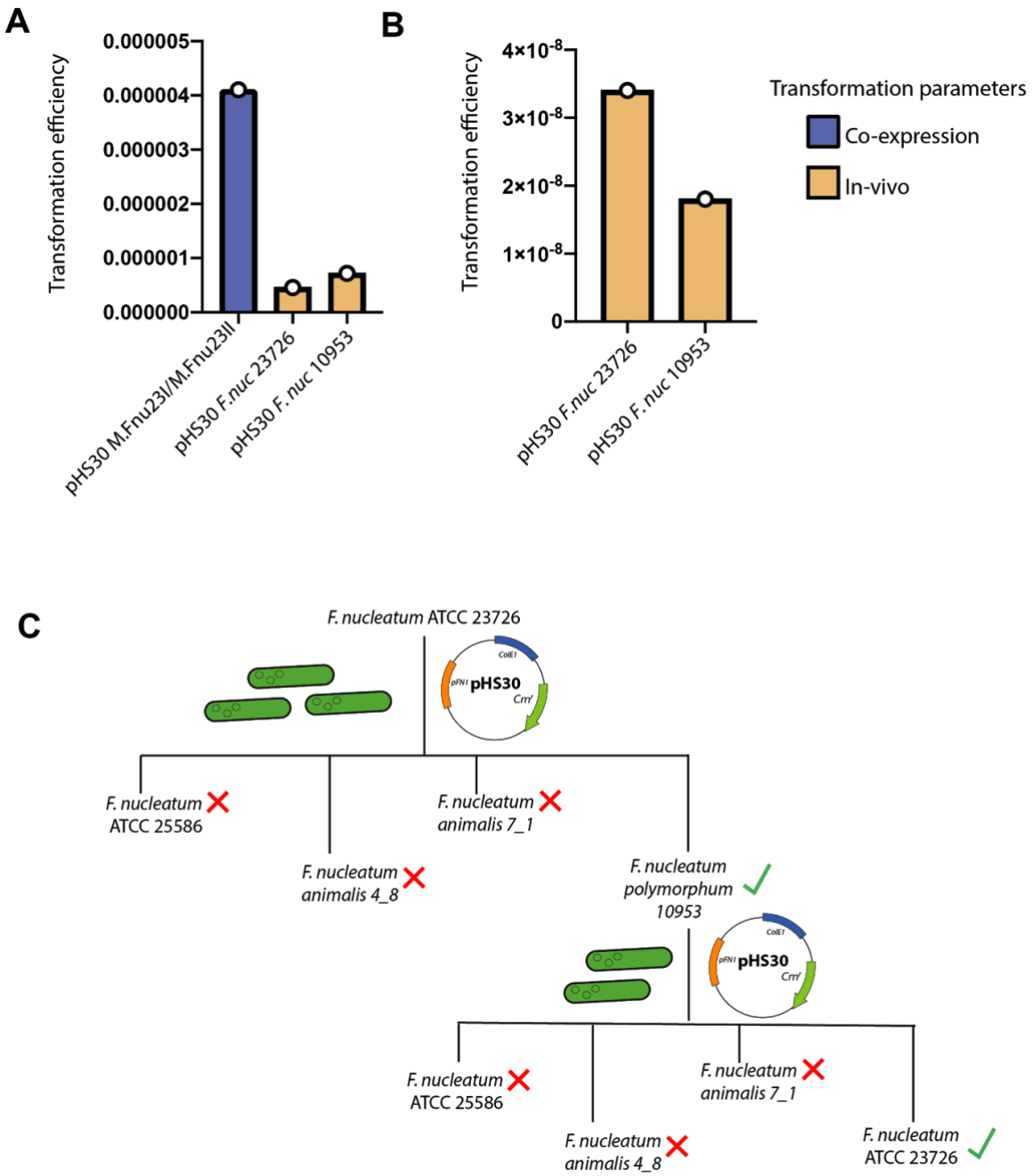


Figure 2-13 Transformation efficiency of the episomal plasmid pHS30 in *F. nucleatum* ATCC 23726 and *F. nucleatum* 10953. A Transformation efficiency calculated from colony counts of electroporation with 500 ng pHS30 coexpressed with M.Fnu23I, M.Fnu23II and transformed into *F. nucleatum* 23726. **B.** Transformation efficiency calculated from colony counts of electroporation with 500 ng pHS30 expressed in *F. nucleatum* 10953. **C.** Diagram of electroporation of native-methylated pHS30 in *F. nucleatum* strains.

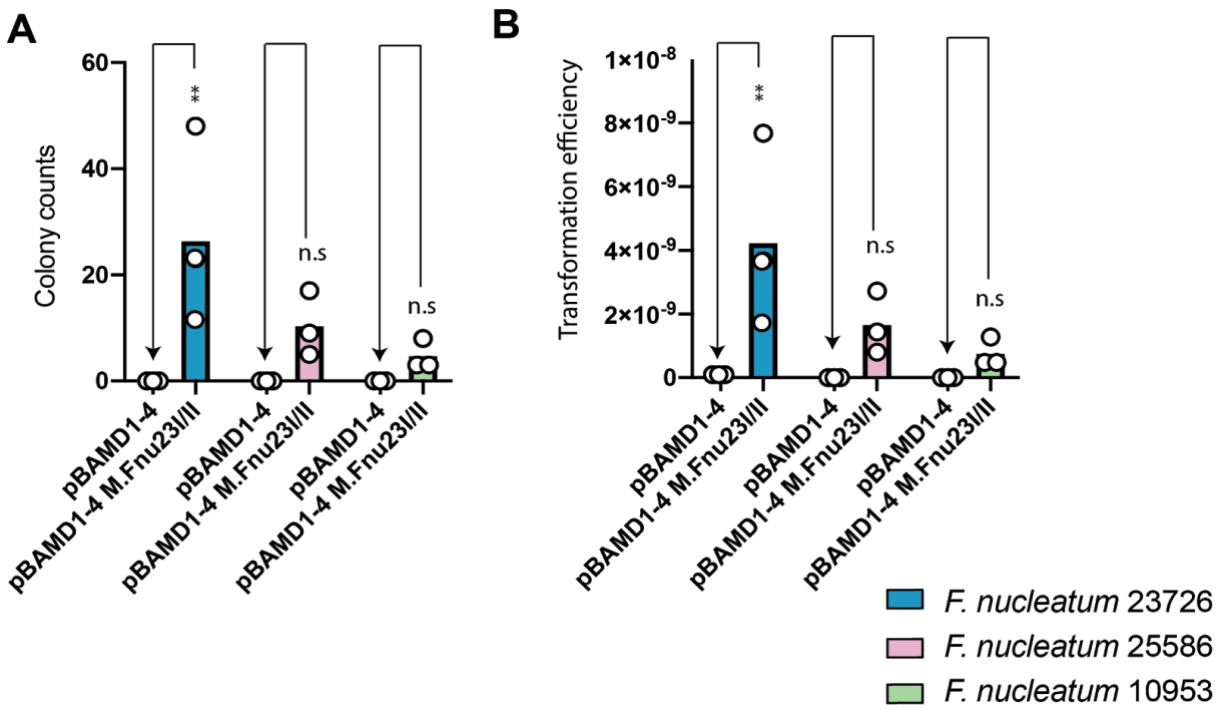


Figure 2-14 Transformation efficiency of pBAMD1-4 in *F. nucleatum* ATCC 23726, *F. nucleatum* ATCC 25586 and *F. nucleatum* 10953. A. Spectinomycin resistant transformants were counted from electroporation of *F. nucleatum* ATCC 23726, *F. nucleatum* ATCC 25586 and *F. nucleatum* 10953 with 3 μg pBAMD1-4 coexpressed with M.Fnu23I/M.Fnu23II. **B.** Transformation efficiency calculated from colony counts of electroporation with 3 μg pBAMD1-4 coexpressed with M.Fnu23I/M.Fnu23II.

Effect of plasmid methylation on an episomal plasmid

We examined whether in vitro methylation of episomal plasmid pHS30 increased transformation efficiency as shown previously with the suicide plasmid pDJSVT1. Therefore, we compared the transformation efficiency of non-conditioned pHS30 and M.Fnu23I/M.Fnu23II treated pHS30. A higher transformation efficiency was observed when pHS30 was treated with both Type II methyltransferases **(Figure 2.12).**

Our previous attempts to introduce pHS30 treated with M.Fnu23I and M.Fnu23II in *F. nucleatum* subsp. *polymorphum* 10953 were not successful. However, when *in vivo* methylated pHS30 from *F. nucleatum* ATCC 23726 is electroporated, thiamphenicol-transformants are achieved **(Figure 2.13)**. Similarly, native-methylated and purified plasmid from *F. nucleatum* subsp. *polymorphum* 10953 yielded *F. nucleatum* ATCC 23726 transformants. Interestingly, the transformation efficiency in *F. nucleatum* ATCC 23726 was higher when pHS30 was treated *in vitro* with M.Fnu23I and M.Fnu23II **(Figure2.13A)**. Interestingly, no other *F. nucleatum* strain as shown in **Figure 2.13C** was capable of transformation with the episomal pHS30.

Effect of plasmid methylation on a Transposon vector

To further elucidate the potential of host mimicking strategies we conditioned pBAMD1-4, a broad-host-range mini-Tn5-vector for Gram-negative bacteria [30] for the six strains of *Fusobacterium* with M.Fnu23I and M.Fnu23II. The results showed that *F. nucleatum* ATCC 23726 was transformed with higher efficiency than *F. nucleatum* subsp. *polymorphum* 10953 and *F. nucleatum* ATCC 25586 **(Figure 2.14)**. We provide the first experimental evidence that *F. nucleatum* ATCC 25586 was transformed with exogenous plasmid DNA. Additionally, the

pBAMD1-4 vector provides resistance to the streptomycin antibiotic, which is the third known functional antibiotic cassette (first, erythromycin and second, chloramphenicol) in *Fusobacterium*.

Discussion

The transformation efficiency of *Fusobacterium* by plasmids is highly diverse, with most of the strains displaying nearly absolute resistance to genetic transformation by current methods [4,6]. The wide variety of DNA R-M systems between strains has been suggested as the possible cause of variability in transformation [4,6]. There have been several R-M systems identified in *F. nucleatum* [25,26], yet we present for the first time a strategy to utilize the methyltransferases to improve transformation efficiency. We identified two major methyltransferases that are capable of increasing the transformation efficiency in *F. nucleatum* ATCC 23726. Our experiments show that the restriction barriers in *F. nucleatum* ATCC 23726 can be overcome by host-mimicking methylation, or deletion of methylation sites of plasmid DNA *in vitro* and coexpression utilizing M.Fnu23I and M.Fnu23II. Nevertheless, the use of these enzymes didn't improve the transformation efficiency in other *Fusobacterium* strains mostly because *Fusobacterium* strains differ in the composition and number of R-M systems.

The genome size of the *Fusobacterium* strains used in this study range between 2.19-4.4 Mb as shown in **Figure 2.4** and the number of R-M systems ranges between 3-12 per strain. The quantity of R-M systems usually increases with the genome size, and bacteria with a genome size of 2 to 3 Mbp encode for an average of 3 R-M systems [16]. However, *Fusobacterium* is unique as it possesses on average more R-M systems despite having such small genomes [16].

The Type I R-M identified in *F. nucleatum* ATCC 23726 displays a classical conformation with three essential components of a protein complex, HsdR required for restriction, HsdM, and HsdS both necessary for methyltransferase activity [38] with a potential ATCNNNNNNCTC or ATTYNNNNNCTC restriction site as it shares 77% identity with CdiBI9 from *Clostridium*

difficile and SmoLIII from *Streptobacillus moniliformis*. However, further experimental studies will be necessary to determine the specific restriction site for Type I in *F. nucleatum* ATCC 23726.

In contrast to Type I R-M systems, Type II R-M systems contain independent restriction endonuclease and methyltransferase enzymes. In *F. nucleatum* ATCC 23726 the orphan methyltransferase M.Fnu23I shares a high identity percentage and the same predicted methylation pattern, CATG, as M.NlaIII from *Helicobacter pylori*. In a similar fashion, the restriction barrier to plasmid transformation in *H. pylori* was as well surpassed by DNA methylation in vitro [27]. Whereas M.Fnu23II is highly similar to M.SfaNI, a protein of 695 amino acids with a HTH domain and a different DNA-strand recognition and modification domain, encoded by *S. faecalis* with a restriction site GCATC; and it is cognate endonuclease disrupts the DNA five and nine nucleotides downstream from respective the methylation pattern [39].

The BREX system from *F. nucleatum* ATCC 23726 seems to be highly conserved in the *Fusobacterium* genus as 42% is the highest identity percentage shared with the BREX system from *Carboxydothermus hydrogenoformans,* an anaerobic gram positive. The BREX system contains 2-8 conserved genes, some of the core genes share homology with genes from the previously characterized Pgl system [40]. The Pgl system consists of four genes *pglWXYZ;* PlgW acts as a kinase, PlgX as a SAM dependent methyltransferase, PlgY an adenosine triphosphatase (ATPase) and PglZ an alkaline phosphatase [36]. Previous analyses of *pglZ*-domain genes revealed that this domain is enriched in defense islands of bacteria and archaea [41]. In *F. nucleatum* ATCC 23726 a *plgZ* gene was identified four genes downstream from the BREX-*pglX* gene. Additionally, a *pgly* gene was found five genes upstream from the BREX gene identified. Neverheless, further studies are necessary to characterize the complex BREX system encoded by *F. nucleatum* ATCC 23726

and the to fully decipher its mechanism of action while comprehending the role of each of its proteins in phage defense.

Furthermore, we studied Type III R-M systems in *Fusobacterium*, which are the most poorly characterized among all Types of R-M systems [42]. Both methyltransferases exhibit a high identity percentage to EcoP15 from *E.coli* which led us to identify a potential CAGCAG recognition site. However, when a protection assay was performed the DNA was susceptible to digestion by the cognate EcoP15 endonuclease (data not shown). We utilized the methyltransferase proteins from these systems to protect the plasmid DNA from native endonucleases, potentially leading to genetic modification of *F. nucleatum* ATCC 25586, however the R-M and physical barrier of this strain was not surpassed when utilizing purified proteins and *in vitro* approaches for the incorporation of suicide plasmid.

The main goal of this study was to create a toolkit to surpass *Fusobacterium* R-M systems and increase transformation efficiency, since they cleave exogenous DNA after discerning it from endogenous DNA based on host-specific methylation [1]. We successfully utilized DNA methyltransferases in *F. nucleatum* ATCC 23726 to mimic these patterns to evade digestion and actively tried to expand this toolkit to other *Fusobacterium* strains. Since we selected and cloned a number of methyltransferases from *F. nucleatum* subsp. *animalis* 7_1 and *F. necrophorum* subsp. *funduliforme* 1_1_36S and subsequently treated the plasmid DNA, we identified that the R-M barrier was not surpassed in *F. nucleatum* subsp. *animalis* 7_1, *F. nucleatum* subsp. *polymorphum* 10953, *F. nucleatum subsp. animalis* 4_8, *F. necrophorum* subsp. *funduliforme* 1_1_36S and *F. necrophorum* subsp. *funduliforme* 51357 even when pooling the methyltransferases. This could suggest that there is a difference in expression levels of methyltransferases and cognate restriction

endonuclease. In addition, the conditions established for electroporation may have to be optimized for each strain.

Another strategy for host mimicking methylation is the use of crude extract prepared from the bacteria. However, when used in all six *Fusobacterium* strains it showed no improvement of transformation efficiency potentially indicating that the methyltransferase activity in crude extract is low under our experimental conditions and potentially the DNA could be degraded by nucleases [1].

To attain *in vivo* methylation we passaged pHS30, the episomal plasmid through *F. nucleatum* ATCC 23726 which is a related strain to the *F. nucleatum* strains electroporated with pHS30. Only *F. nucleatum* subsp. *polymorphum* 10953 was capable of being transformed with in vivo methylated pHS30 from *F. nucleatum* ATCC 23726; while F. *nucleatum* ATCC 23726 was also transformed with the episomal plasmid derived from *F. nucleatum* subsp. *polymorphum* 10953. This phenomenon could be explained by the fact that both strains are one of the only few of *F. nucleatum* strains that have been previously transformed [4].

Furthermore, as transposon mutagenesis represents an opportunity to identify new virulence factors in an unbiased manner, we conditioned pBAMD1-4 with M.Fnu23I and M.Fnu23II. To this date, only a couple of transposon libraries have been developed for *F. nucleatum* ATCC 23726 and *F. nucleatum* subsp. *polymorphum* 10953 [2,5] and all have been constructed utilizing the pMOD-CatP plasmid, which confers thiamphenicol resistance. A total of approximately 1,200 mutants were initially obtained utilizing this method [5], however when optimized by Wu and colleagues in 2018, 24,000 colonies resulted from transformation. While the novel transposon method developed in this dissertation is yet to be optimized to achieve a higher transformation efficiency, the fact that streptomycin resistant colonies were obtained from *F.*

nucleatum subsp. *polymorphum* 10953, *F. nucleatum* ATCC 23726, and *F. nucleatum* ATCC 25586 reveals a promising future in genetic manipulation.

Additionally, roles for the R-M systems have been proposed in the adaptation to hostile environments, virulence of the organism, regulation of genes and genetic exchange in bacteria [16]. Another crucial purpose of the R-M systems is to regulate genetic variation and therefore control the rate of evolution [43,44]. Moreover, organisms with methyltransferases that methylate N6-adenine like *Fusobacterium* seem to take advantage of this exocyclic methylation and also employ the methyl group addition in the regulation of transcription of housekeeping genes and virulence genes, cell cycle control, DNA replication initiation and mismatch repair, post-transcriptional regulation and, nucleoid reorganization [16]. Further experiments are necessary to confirm if *Fusobacterium* methyltransferases have a role in these cellular mechanisms.

Little is known about the role of DNA-methylation in phase variation of the pathogen *Fusobacterium*. Complete analysis of the methylome will be key in understanding the pathobiology of *Fusobacterium* [45]. In addition, methylome studies will contribute to investigating the role of DNA methyltransferases in *Fusobacterium* physiology [45]. However, experimental data is crucial to determine if the R-M systems mediate phase variation through mutation of simple tandem repeats since the formation of DNA methylation patterns are frequent epigenetic mechanisms for transcriptional regulation [46,47]. In the human pathogens *Neisseria meningitidis*, *N. gonorrhoeae* and *Haemophilus influenzae*, N6-adenine methylation by DNA methyltransferases has been demonstrated to modulate gene transcription as a consequence of enzyme inactivation by mutation, and the loci under methyltransferase control includes genes related to virulence, envelope structure, and stress responses [46,47].

This current study is the first identification and characterization of host mimicking strategies utilizing R-M systems in *F. nucleatum*. Although we characterizeR-M systems and their role in genetic recalcitrance, further studies are required to utilize methyltransferases to efficiently circumvent R-M systems in more clinically relevant *Fusobacterium* strains. A thorough understanding of methyltransferases, and their potential role in phase-variable control, is also necessary to decode *Fusobacterium* pathogenesis and potentially develop new treatments and vaccines.

Acknowledgments

The Fralin Life Science Institute at Virginia Tech (Slade), and the USDA National Institute of Food and Agriculture (Slade).

Author contributions

AU and BES designed, performed, and analyzed experiments. KW and BFW performed experiments. AU and DJS wrote the manuscript.

Competing Interests

The authors declare that they have no conflicts of interest with the contents of this article.

References

1. Suzuki H. Host-Mimicking Strategies in DNA Methylation for Improved Bacterial Transformation. In: Dricu A, editor. Methylation - From DNA, RNA and Histones to Diseases and Treatment. InTech; 2012.

2. Wu C, Al Mamun AAM, Luong TT, Hu B, Gu J, Lee JH, et al. Forward Genetic Dissection of Biofilm Development by Fusobacterium nucleatum: Novel Functions of Cell Division Proteins FtsX and EnvC. MBio. 2018;9. doi:10.1128/mBio.00360-18

3. Casasanta MA, Yoo CC, Udayasuryan B, Sanders BE, Umaña A, Zhang Y, et al. Fusobacterium nucleatum host-cell binding and invasion induces IL-8 and CXCL1 secretion that drives colorectal cancer cell migration. Sci Signal. 2020;13. doi:10.1126/scisignal.aba9157

4. Kinder Haake S, Yoder S, Gerardo SH. Efficient gene transfer and targeted mutagenesis in Fusobacterium nucleatum. Plasmid. 2006;55: 27–38.

5. Coppenhagen-Glazer S, Sol A, Abed J, Naor R, Zhang X, Han YW, et al. Fap2 of Fusobacterium nucleatum is a galactose-inhibitable adhesin involved in coaggregation, cell adhesion, and preterm birth. Infect Immun. 2015;83: 1104–1113.

6. Haake SK, Yoder SC, Attarian G, Podkaminer K. Native plasmids of Fusobacterium nucleatum: characterization and use in development of genetic systems. J Bacteriol. 2000;182: 1176–1180.

7. Ikegami A, Chung P, Han YW. Complementation of the fadA mutation in Fusobacterium nucleatum demonstrates that the surface-exposed adhesin promotes cellular invasion and placental colonization. Infect Immun. 2009;77: 3075–3079.

8. Han YW, Ikegami A, Rajanna C, Kawsar HI, Zhou Y, Li M, et al. Identification and characterization of a novel adhesin unique to oral fusobacteria. J Bacteriol. 2005;187: 5330–5340.

9. Hashimoto H, Morikawa H, Yamada Y, Kimura A. A novel method for transformation of intact yeast cells by electroinjection of plasmid DNA. Appl Microbiol Biotechnol. 1985;21: 336–339.

10. Miller JF, Dower WJ, Tompkins LS. High-voltage electroporation of bacteria: genetic transformation of Campylobacter jejuni with plasmid DNA. Proc Natl Acad Sci U S A. 1988;85: 856–860.

11. Shi J, Ma Y, Zhu J, Chen Y, Sun Y, Yao Y, et al. A Review on Electroporation-Based Intracellular Delivery. Molecules. 2018;23. doi:10.3390/molecules23113044

12. Johnston CD, Cotton SL, Rittling SR, Starr JR, Borisy GG, Dewhirst FE, et al. Systematic evasion of the restriction-modification barrier in bacteria. Proc Natl Acad Sci U S A. 2019;116: 11454–11459.

13. Arber W, Hattman S, Dussoix D. ON THE HOST-CONTROLLED MODIFICATION OF BACTERIOPHAGE LAMBDA. Virology. 1963;21: 30–35.

14. Arber W, Dussoix D. Host specificity of DNA produced by Escherichia coli. I. Host controlled modification of bacteriophage lambda. J Mol Biol. 1962;5: 18–36.

15. Official REBASE Homepage | The Restriction Enzyme Database | NEB. [cited 26 Mar 2020]. Available: http://rebase.neb.com/rebase/rebase.html

16. Vasu K, Nagaraja V. Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense. Microbiology and Molecular Biology Reviews. 2013. pp. 53–72. doi:10.1128/mmbr.00044-12

17. Dybvig K, Sitaraman R, French CT. A family of phase-variable restriction enzymes with differing specificities generated by high-frequency gene rearrangements. Proc Natl Acad Sci U S A. 1998;95: 13923–13928.

18. Donahue JP, Peek RM. Restriction and modification systems, chapter 24. Helicobacter pylori: physiology and genetics ASM Press, Washington, DC [Google Scholar]. 2001.

19. Sitaraman R. The Role of DNA Restriction-Modification Systems in the Biology of Bacillus anthracis. Front Microbiol. 2016;7: 11.

20. Goldfarb T, Sberro H, Weinstock E, Cohen O, Doron S, Charpak-Amikam Y, et al. BREX is a novel phage resistance system widespread in microbial genomes. EMBO J. 2015;34: 169–183.

21. Buryanov Y, Shevchuk T. The use of prokaryotic DNA methyltransferases as experimental and analytical tools in modern biology. Anal Biochem. 2005;338: 1–11.

22. Jensen TØ, Tellgren-Roth C, Redl S, Maury J, Jacobsen SAB, Pedersen LE, et al. Genome-wide systematic identification of methyltransferase recognition and modification patterns. Nat Commun. 2019;10: 3311.

23. Orlowski J, Bujnicki JM. Structural and evolutionary classification of Type II restriction enzymes based on theoretical and experimental analyses. Nucleic Acids Res. 2008;36: 3552–3569.

24. Cloning the FnuDI, Nael, N coI and XbaI restriction-modification systems. Gene. 1988;74: 55–59.

25. Lui AC, McBride BC, Vovis GF, Smith M. Site specific endonuclease from Fusobacterium nucleatum. Nucleic Acids Res. 1979;6: 1–15.

26. Leung DW, Lui AC, Merilees H, McBride BC, Smith M. A restriction enzyme from Fusobacterium nucleatum 4H which recognizes GCNGC. Nucleic Acids Res. 1979;6: 17–25.

27. Donahue JP, Israel DA, Peek RM, Blaser MJ, Miller GG. Overcoming the restriction barrier to plasmid transformation of Helicobacter pylori. Mol Microbiol. 2000;37: 1066–1074.

28. Yang X, Xu M, Yang S-T. Restriction modification system analysis and development of in vivo methylation for the transformation of Clostridium cellulovorans. Appl Microbiol Biotechnol. 2016;100: 2289–2299.

29. Costa SK, Donegan NP, Corvaglia A-R, François P, Cheung AL. Bypassing the Restriction System To Improve Transformation of Staphylococcus epidermidis. J Bacteriol. 2017;199. doi:10.1128/JB.00271-17

30. Martínez-García E, Aparicio T, de Lorenzo V, Nikel PI. New transposon tools tailored for metabolic engineering of gram-negative microbial cell factories. Front Bioeng Biotechnol. 2014;2: 46.

31. Andersen KR, Leksa NC, Schwartz TU. Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. Proteins. 2013;81: 1857–1861.

32. Manson McGuire A, Cochrane K, Griggs AD, Haas BJ, Abeel T, Zeng Q, et al. Evolution of invasion in a diverse set of Fusobacterium species. MBio. 2014;5: e01864.

33. Geneious. [cited 2 Sep 2020]. Available: https://www.geneious.com/

34. Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, et al. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. Biochim Biophys Acta. 2015;1854: 1019–1037.

35. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13: 2498–2504.

36. Gordeeva J, Morozova N, Sierro N, Isaev A, Sinkunas T, Tsvetkova K, et al. BREX system of Escherichia coli distinguishes self from non-self by methylation of a specific DNA site. Nucleic Acids Res. 2019;47: 253–265.

37. Rao DN, Dryden DTF, Bheemanaik S. Type III restriction-modification enzymes: a historical perspective. Nucleic Acids Res. 2014;42: 45–55.

38. Murray NE. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiol Mol Biol Rev. 2000;64: 412–434.

39. Furmanek-Blaszk B, Sektas M. The SfaNI restriction-modification system from Enterococcus faecalis NEB215 is located on a putative mobile genetic element. FEMS Microbiol Lett. 2015;362: fnv028.

40. Sumby P, Smith MCM. Genetics of the phage growth limitation (Pgl) system of Streptomyces coelicolor A3(2). Mol Microbiol. 2002;44: 489–500.

41. Makarova KS, Wolf YI, Snir S, Koonin EV. Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. J Bacteriol. 2011;193: 6039–6056.

42. Kwiatek A, Mrozek A, Bacal P, Piekarowicz A, Adamczyk-Popławska M. Type III Methyltransferase M.NgoAX from Neisseria gonorrhoeae FA1090 Regulates Biofilm Formation and Interactions with Human Cells. Front Microbiol. 2015;6: 1426.

43. Arber W. Genetic variation: molecular mechanisms and impact on microbial evolution. FEMS Microbiol Rev. 2000;24: 1–7.

44. Arber W. Genetically encoded generators of genetic variants. J Proteomics. 2009;72: 836–837.

45. Atack JM, Tan A, Bakaletz LO, Jennings MP, Seib KL. Phasevarions of Bacterial Pathogens: Methylomics Sheds New Light on Old Enemies. Trends Microbiol. 2018;26: 715–726.

46. Casadesús J, Low DA. Programmed heterogeneity: epigenetic mechanisms in bacteria. J Biol Chem. 2013;288: 13929–13935.

47. Srikhanta YN, Fox KL, Jennings MP. The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. Nature Reviews Microbiology. 2010. pp. 196–206. doi:10.1038/nrmicro2283

Chapter 3

CRISPR-Cas systems in *Fusobacterium*: An untapped genetic frontier?

Ariana Umaña[1], Daniel J. Slade[1]

1. Virginia Polytechnic Institute and State University, Department of Biochemistry, Blacksburg, VA, USA.

Abstract

Fusobacterium nucleatum has recently received significant attention for its strong connection with the acceleration and gravity of colorectal cancer, yet understanding the molecular mechanisms that drive infection by this 'oncomicrobe' have been hindered by a deficiency of genetic tools. Using newly completed and highly accurate *Fusobacterium* genomes for bioinformatic analysis, herein we report the first full characterization of multiple *Fusobacterium* CRISPR-Cas adaptive immune systems. We have identified multiple CRISPR-Cas effectors including Cas9 and the novel Cas13c. Additionally, we analyzed the CRISPR-array spacer/repeat loci, *trans*-activating CRISPR RNA (tracrRNA), and CRISPR RNA (cRNA) from a Type II-A CRISPR-Cas9 system. Since *Fusobacterium* are genetically recalcitrant, this CRISPR-Cas bioinformatic roadmap could be used to build new genome editing and transcriptional tuning tools to characterize an increasingly important genus of human opportunistic-pathogens.

Introduction

Fusobacteria are obligate anaerobic Gram-negative bacteria and that include human as well as animal pathogens. Due to robust restriction modification systems that hinder genome manipulation, *Fusobacterium* genetics have classically been considered difficult. Nevertheless, in recent times genetic tool advances in *Fusobacterium* expanded and continue to increase [1–3]. Genetic systems incorporating homologous recombination for gene deletion are frequent among many microorganisms, including *Fusobacterium.* The most common counterselection marker is galactose kinase (GalK) and this system was validated in a single strain for *F. nucleatum* subsp. *nucleatum* ATCC 23726 [1,2].

Recent advances in our understanding of Clustered regularly interspaced short palindromic repeats-(CRISPR)-associated proteins (CRISPR-Cas) systems which are adaptive immune mechanisms that protect bacteria from invading viruses or Mobile Genetic Elements (MGE), have been exploited by molecular biologists to allow for efficient and robust genome editing in both prokaryotic and eukaryotic systems [4–6]. The CRISPR-Cas system mode of action is categorized into three steps as detailed in **Figure 3.1**: (1) adaptation or spacer acquisition, (2) processing of the primary transcript of the CRISPR locus (pre-crRNA) and maturation of the crRNA which encompasses the spacer and variable regions corresponding to 5′ and 3′ fragments of CRISPR repeats, and (3) DNA (or RNA) interference [7].

CRISPR and Cas proteins comprise the basis of adaptive immunity in bacteria and archaea [8]. CRISPR-Cas systems identify and cleave non-host DNA and RNA (viruses and plasmids), and deposit a ~30 bp segment into its host genome at defined CRISPR loci (host DNA repeats and foreign protospacer, or ‘spacer’ DNA) to provide organism- and sequence-specific defenses [9].

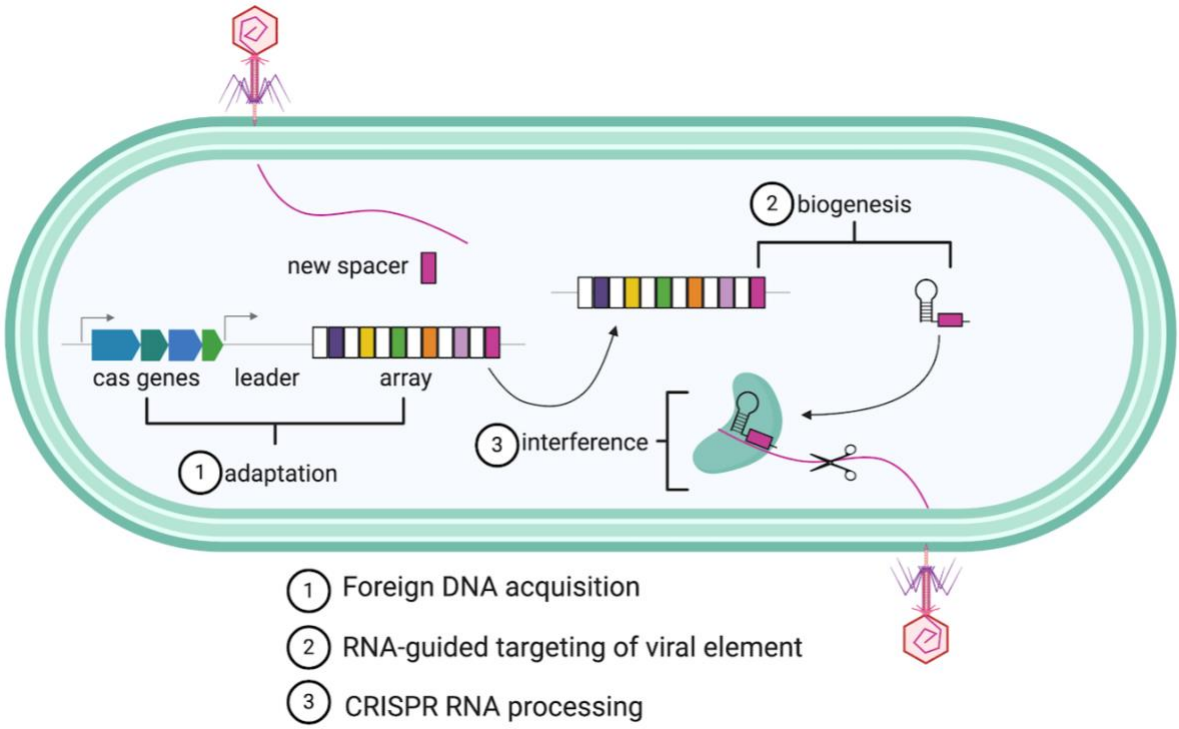


Figure 3-1The CRISPR- Cas system mode of action is categorized in three phases: (1) adaptation or foreign spacer acquisition, **(2)** processing of the primary transcript of the CRISPR locus (pre-crRNA) and maturation of the crRNA which encompasses the spacer and variable regions corresponding to 5′ and 3′ fragments of CRISPR repeats, and **(3)** DNA (or RNA) interference.

During repeat infections, CRISPR-Cas complexes recognize these specific sequences and rapidly clear the invading DNA through active nucleotide cleavage [10]. CRISPR systems have been characterized in 87% of archaea and 47% of bacteria and have been proposed to proliferate through horizontal gene transfer across bacteria as a basis of adaptive immunity [9]. CRISPR arrays are divided into two main categories (Class 1-2) and sixteen distinct subcategories based on the cas genes encoded within bacteria and archaea [11].

CRISPR-Cas Class 1 systems are characterized by an uneven stoichiometry of Cas protein subunits [12]. Although CRISPR-Cas Class 1 complexes are the most widely distributed CRISPR systems in nature (90%), their genome engineering applications are yet to be explored [13]. Type I CRISPR–Cas systems utilize the signature Cas3 nuclease–helicase to degrade invading DNA, and are further categorized into three Types (Type I, Type III, and Type IV) and eight subtypes (I-A to I-G and I-U) based on accessory cas genes [13].

Class 2 systems are only found in bacteria, they are more simple as their effector module consists of a single, multidomain protein whos most recent classification encompasses six major Types (I–VI) [7,12]. This characteristic makes Class 2 CRISPR–Cas systems a sought-after option for the development of a new generation of genome-editing tools [12]. The prevailing and best characterized Class 2 system is the Type II effector, Cas9. Cas9 is a crRNA dependent endonuclease consisted of two different nuclease domains, HNH and RuvC. Both domains are essential for cleavage of the superseded (non-target) and target DNA, in the crRNA–target DNA complex [12]. In addition, Type II CRISPR–cas system includes a trans-acting CRISPR RNA (tracrRNA) that potentially evolved from the corresponding CRISPR and that is required for pre-crRNA processing and target identification in Type II systems [12,14].

Among the Class 2 systems the Type VI system is comprised of the subtypes VI-A (effector Cas13a), VI-B (effector Cas13b), VI-C (effector Cas13c), and VI-D (effector Cas13d) depending on its domain composition [15]. The Cas13 family enzyme of the Type VI category is unique as it is the single Class 2 Cas enzyme that has been characterized to exclusively cleave single-stranded RNA [16]. RNA targeting and cleavage is moderated by two higher eukaryotes and prokaryotes nucleotide (HEPN)-binding domains (HEPN-1 is comprised of ten α-helices while HEPN-2 domain fold into eleven α-helices) and their characteristic motif R-X4-H [17]. Commonly Cas13 effectors display collateral nonspecific RNase activity triggered by binding of target RNA sequence [16]. Furthermore, Cas13 enzymes have been utilized as RNA-binding proteins to target as well as to edit RNA [16].

The majority of genome editing studies have used the Cas9 protein from *S. pyogenes* [11]. While the use of *S. pyogenes* Cas9 in prokaryotic and eukaryotic systems has been achieved, our goal was to characterize *Fusobacterium* CRISPR-Cas proteins to determine if these proteins will be more efficient for future gene modulation studies in an AT rich bacterium. The discovery and characterization of native Class 2 effectors in *Fusobacterium* will potentially enhance the application and development of CRISPR systems for genome engineering. In summary, we report the first full characterization of multiple *Fusobacterium* CRISPR-Cas adaptive immune systems, focusing on Cas9 and the newly described Cas13c present in *F. necrophorum* subsp. *funduliforme* 1_1_36S. The development of new genetic tools from CRISPR-Cas enzymes hast the potential to result in the improvement of our understanding of fusobacterial physiology and pathogenesis.

Materials and Methods

Use of genomic information for bioinformatic analysis and identification of CRISPRCas systems in eight *Fusobacterium* strains.

Fusobacterium genomes (*F. nucleatum* subsp. *nucleatum* ATCC 23726 (GCA_003019875.1), *F. nucleatum* subsp. *nucleatum* ATCC 25586 (GCA_003019295.1), *F. necrophorum* subsp. *funduliforme* 1_1_36S (GCA_003019715.1), *F. varium 27725* (GCA_003019655.1), *F. ulcerans 49185* (GCA_003019675.1), *F. mortiferum* 9817 (GCA_003019315.1), *F. gonidiaformans* 25563 (GCA_003019695.1), and *F. periodonticum* 2_1_31 (GCA_003019755.1) were used to extract all sequences to analyze using the CRISPROne web server [18] and CRISPRCasFinder [19]. CRISPROne and CRISPRCasFinder were used to predict all CRISPR-Cas associated elements as well as the repeat arrays.

Bioinformatic analysis of CRISPR-Cas Class 2 systems and identification of effector domain

The genome sequence from *F. necrophorum* subsp. *funduliforme* 1_1_36S (GCA_003019715.1) was used to predict the open reading frame for the CRISPR-Cas systems. An open reading frame of 1337 aminoacids for Cas9 and 1121 aminoacids for Cas13c were identified using the biosequence analyzer profiler web server pHMMER [20] and the stand alone HMMER [20] software package, respectively.

Phylogenetic Tree construction

Full CRISPR-Cas phylogenetic analysis was created using the Micropan [22] plugin in Rstudio by using only CRISPR associated proteins to build the tree. The gene maps in **Figure 3.2B** were created using the gggenes [23] plugin in Rstudio. Additional phylogenetic trees for Cas9 and

Cas13c were created in Geneious [21] utilizing the multiple sequence alignment from SMARTBLAST [24] and adapted in Affinity Designer.

Structure Prediction

The structure prediction using the amino acid sequence of Cas9 from *Fusobacterium necrophorum* subsp. *necrophorum* 1_1_36S was performed in the web-based suite Phyre2 [25]. RNA structure prediction by NUPACK [26].

Results

Comparative Analysis of *Fusobacterium* CRISPR-Cas systems

A total of eight *Fusobacterium* genome sequences were utilized for bioinformatic comparative analysis and the composition, occurrence, and diversity of CRISPR-Cas systems. Presently, there are two classes and six types of CRISPR-Cas systems based in the organization of their effector modules. A majority of the identified CRISPR-Cas systems in *Fusobacterium* belong to Class 1 (77% of CRISPR-Cas systems in the eight genomes) and interestingly, only *F. necrophorum* susbp. *funduliforme* 1_1_36S and *F. ulcerans* 49185 encode for Class 2 systems (23% of CRISPR-Cas systems in the eight genomes). The overall analysis of CRISPR-Cas systems is shown in **Table 3.1**.

To depict the correlation among *Fusobacterium* species and strains, a CRISPR-Cas system-based phylogenetic tree was constructed. As shown in Figure **3.2A**, the eight *Fusobacterium* strains are divided into five distinct clades according to their predicted CRISPR-Cas systems. Interestingly, *F. nucleatum* ATCC 23726 (Lineage 1) clades with *F. gonidiaformans* 25563 (Lineage 3) and *F. periodonticum* 2_1_31 (Lineage 1) clades with *F. necrophorum* susbp. *funduliforme* 1_1_36S (Lineage 3).

The rapid evolution of CRISPR systems explains the diversity of systems from highly similar strains. For instance, *F. nucleatum* ATCC 23726 and *F. nucleatum* ATCC 25586 are the closest phylogenetically when using whole genome phylogenetic analysis, yet as seen in **Figure 3.2A**, they are far more divergent in their CRISPR loci. Additionally, the analysis revealed that *F. mortiferum* 9817 CRISPR-Cas systems have less similarities when compared to the other *Fusobacterium* strains.

Table 3.1 CRISPR-Cas system information of the eight analyzed *Fusobacterium* strains.

Species	Strain	CRISPR-Cas system Type	Subtype	Number of spacers	Conservation
F. nucleatum	25586	I	B1	15	0%
F. nucleatum	23726	I	B1	23	0%
F. periodonticum	2_1_31	I, III	B1, A2	L1:42, L2:13, L3:13, L4: 28	L1:0%, L2:0%, L3:0%, L4:0%
F. varium	27725	I	B1	L1: 13, L2:13, L3:24, L4:18	L1: 9.75%, L2:0%, L3:0%, L4:2.04%
F. ulcerans	49185	I, VI	B1,C	L1:17, L2:23, L3:7	L1:0%, L2:0%, L3:0%
F. mortiferum	9817	III	A1	L1:4, L2:6	L1:0%, L2:0%
F. gonifdiaformans	25563	I	B1	83	0%
F. necrophorum	1_1_36S	III,VI	A1,C	L1:13, L2:9	L1:0%, L2:0%

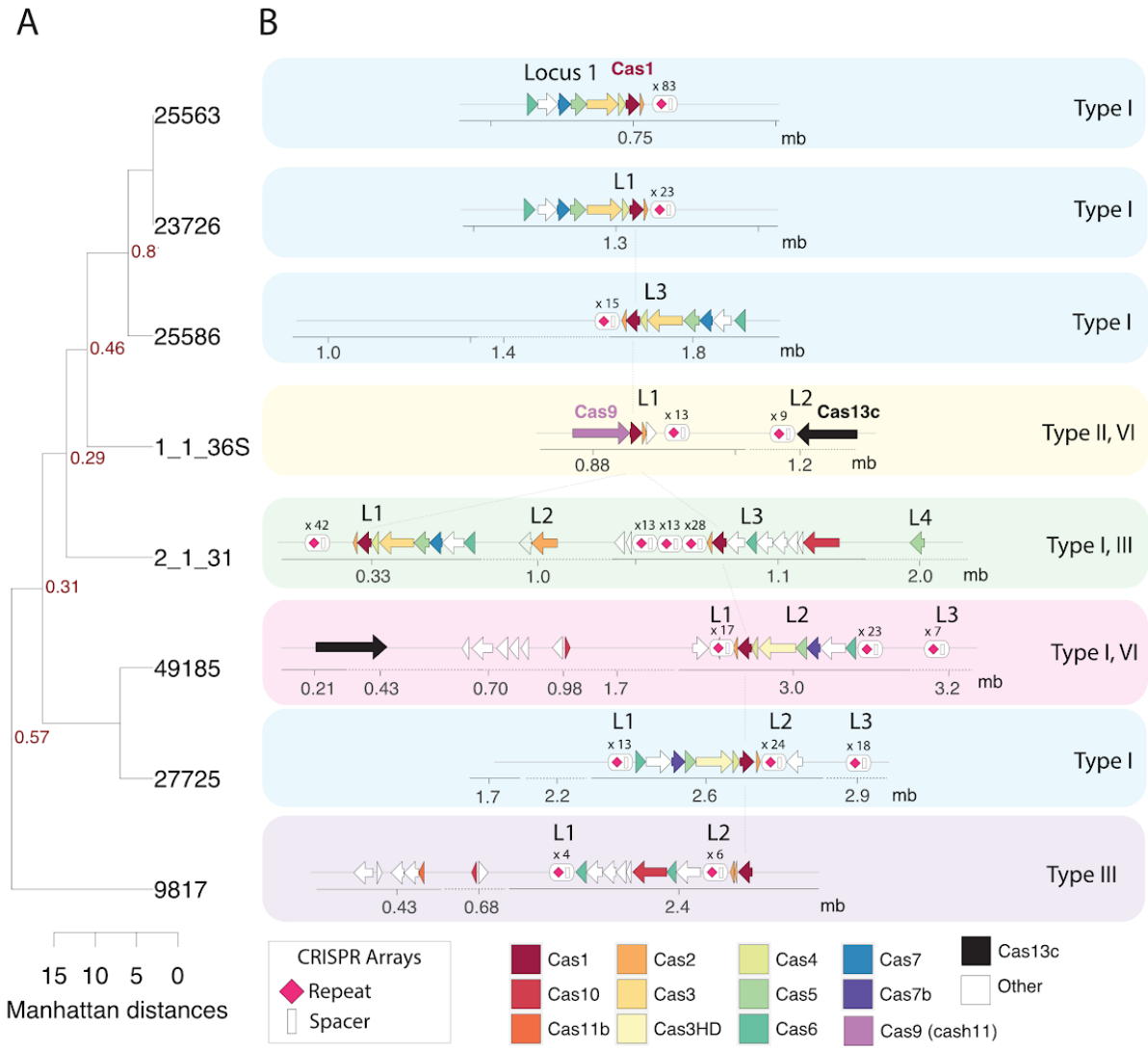


Figure 3-2 CRISPR-Cas systems in *Fusobacterium*. A. CRISPR-Cas phylogenetic tree of eight *Fusobacterium* strains. **B.** A total of eight *Fusobacterium* strains were utilized for bioinformatic analysis and identification of CRISPR-Cas systems. A majority of the identified CRISPR-Cas systems in *Fusobacterium* belong to Class 1 and only *F. necrophorum* susbp. *funduliforme* 1_1_36S and *F. ulcerans* 49185 encode for Class 2 systems.

Estimation of spacer numbers and sequences among *Fusobacterium* species exhibited substantial diversity with 0% (0/348) of the spacers shared by eight strains. As shown in **Table 3.1**, the quantity of predicted spacers varies among the strains, ranging from 4 to 83, demonstrating the biological heterogeneity of *Fusobacterium* strains that encode an adaptive defense mechanism. Interestingly, Type I-B presented a higher number of spacers when compared to the other CRISPR-CAS system types.

CRISPR-Cas systems in *Fusobacterium*

To determine the classification and assortment of CRISPR-Cas systems across the eight *Fusobacterium* strains, we utilized the CRISPRCasFinder and CRISPROne web server that enable the discovery of direct repeat consensus sequences within a genome as well as the characterization of related spacers and cas genes [18,19]. This analysis detected a total of 13 CRISPR-Cas systems belonging to both Class 1 (Type I, Type III and Type IV,) and Class 2 (Type II, Type VI): six Type I (54.54%), one Type II (9.09%), two Type III (18.18%), and two Type VI (18.18%) while no Type IV or V was identified. Therefore, we conclude that Type I is the most dominant CRISPR-Cas system among the studied *Fusobacterium* strains. Detailed illustrations of the CRISPR-Cas systems identified among the strains analyzed in this study are summarized in **Figure 3.2B**.

F. nucleatum ATCC 23726, *F. nucleatum* ATCC 25586, *F. gonidiaformans* 25563 and *F. periodonticum* 2_1_31 Type I-B is a Class 1 CRISPR-Cas loci that includes full gene sets to achieve adaptive immunity: *cas1*, *cas2* and *cas4* for spacer integration or adaptation, cas6 for crRNA-processing and *cas3*, *cas5*, *cas7* and *cas8* for interference. Both *F. ulcerans* 49185 and *F. varium* 27725 include a nearly complete Type I loci, but they do not encode for the *cas8* gene. Additionally, most CRISPRs of Type I-B CRISPR-Cas systems discovered in this analysis were encoded downstream of the *cas2* gene and/or upstream of the *cas6* gene**.** We highlight that only *F.*

nucleatum ATCC 23726 and *F. gonidiaformans* 25563 presented an inverted loci composition of the Type I-B CRISPR-Cas systems identified **(Figure 3.2)**.

Furthermore, in *F. periodonticum* 2_1_31 Type I and Type III appear to combine as they often coexist [27]. Type III systems have been divided into four major subtypes (III-A to D) [28][29]. A complete operon of *cas* genes associated with adaptation, crRNA-processing and interference steps was identified in *F. periodonticum* 2_1_31. Additionally, Type III systems are found to frequently encode for accessory genes near the core cas gene cluster.

Analysis of the *F. necrophorum* subsp. *funduliforme* 1_1_36S Type II CRISPR-Cas system and identification of putative repeats/spacers/PAM sequence

The only Type II-A CRISPR-Cas system identified in *Fusobacterium* belonged to *F. necrophorum* subsp. *funduliforme* 1_1_36S. Type II-A is a Class 2 system that contains four adaptation genes *cas1, cas2, csn2* and *cas9*. The Cas9 additionally acts as a nuclease responsible for invader DNA cleavage **(Figure 3.3)** [32]. The CRISPR-Cas loci was located downstream of *c*as9. The Type II CRISPR-Cas systems function through the assembly of an effector module consisting of a single Cas9 protein, a crRNA and a tracrRNA.

From the complete genome on the Fusoportal server Cas9 is 1337 amino acids in size. A previous draft genome assembly of *F. necrophorum* subsp. *funduliforme* 1_1_36S had this protein as 329 aminoacids, proving that complete and accurate genome sequences can be for identifying these proteins [33]. The predicted structure of Cas9 resembles the characteristic bi-lobed architecture from previously characterized Cas9 proteins **(Figure 3.4A)** and it contains the RuvC and HNH domains for nucleic acid cleavage **(Figure 3.3).**

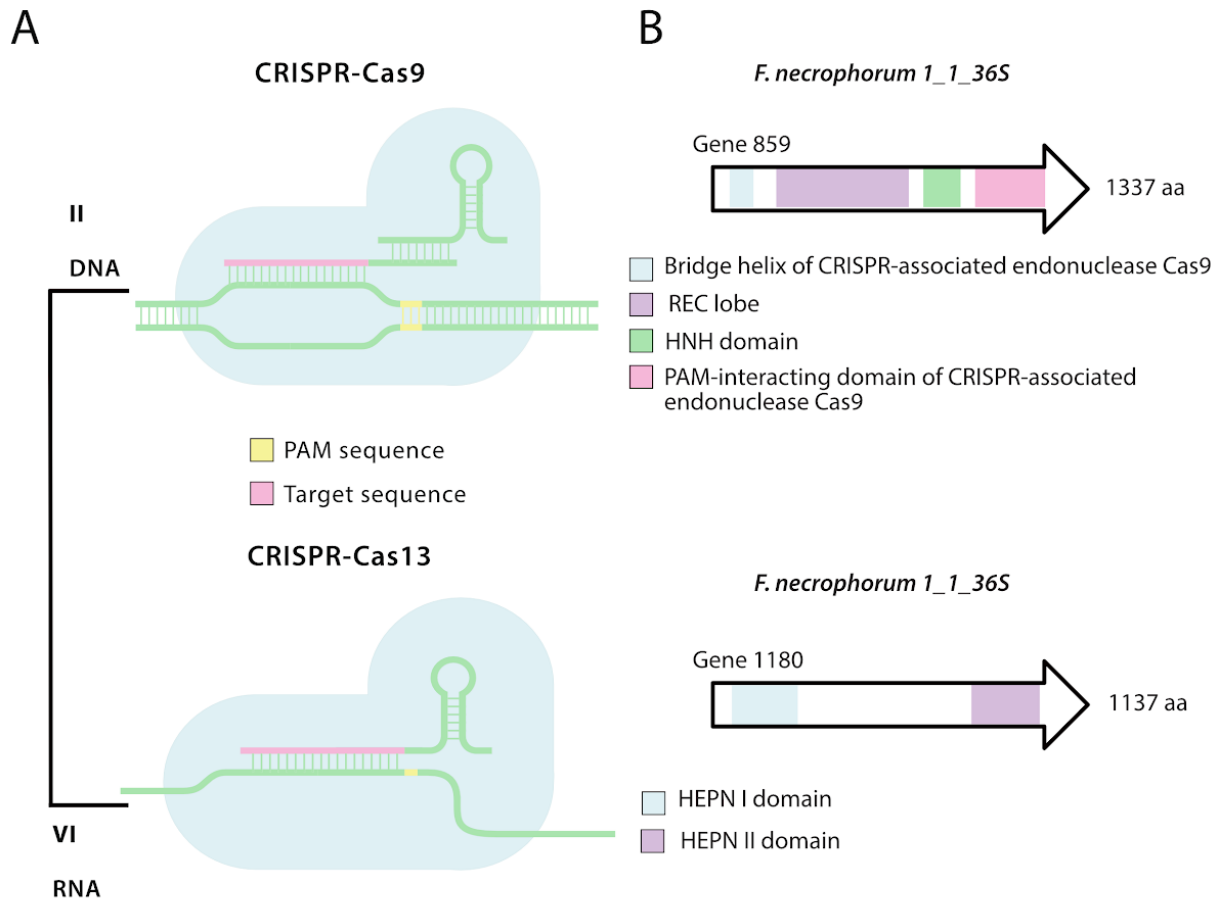


Figure 3-3 Diversity of Class 2 CRISPR-Cas systems in *Fusobacterium*. A. Schematic of CRISPR-Cas9 and CRISPR-Cas13 in *F. necrophorum* subsp. *funduliforme* 1_1_36S. **B.** Effector domain organization in Cas9 and Cas13c. Nuclease domains are detailed.

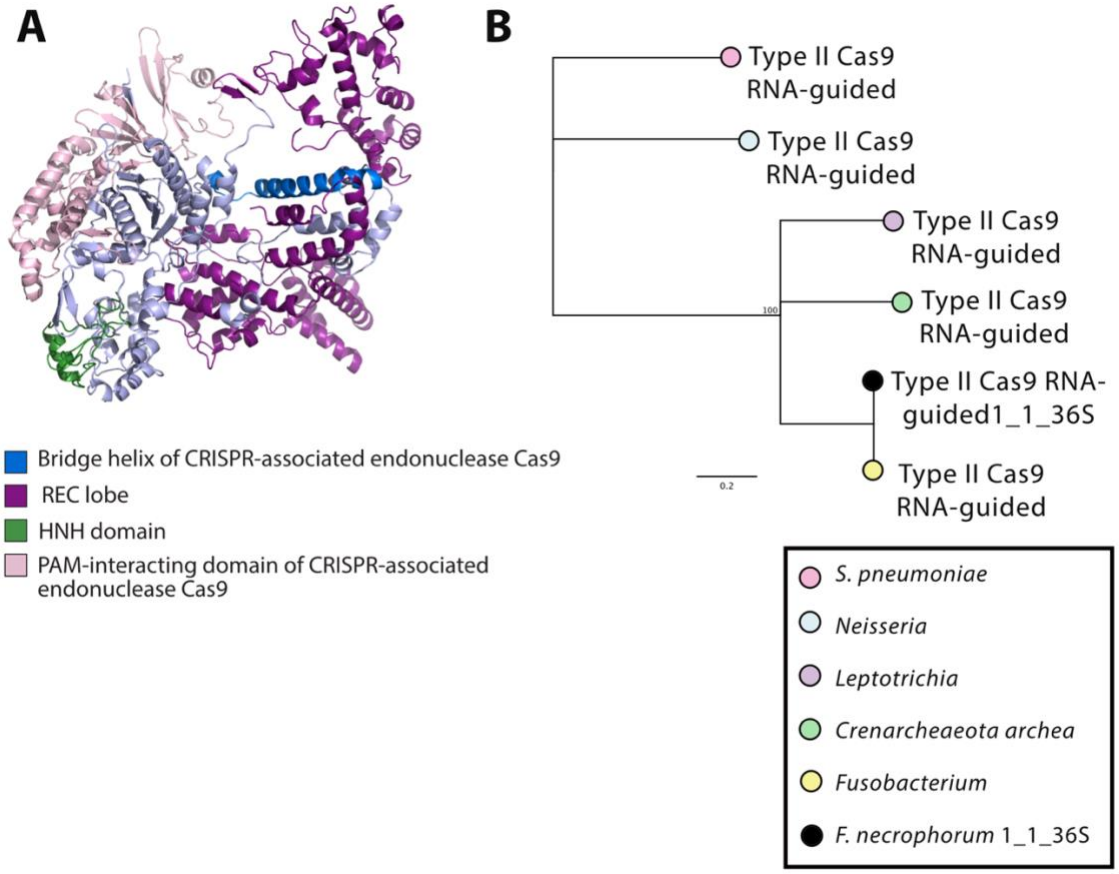


Figure 3-4 Structure prediction and phylogenetic analysis of Cas9 from *F. necrophorum* subsp.*funduliforme* 1_1_36S. **A.** Predicted structure of Cas9 is based on the crystal structure of Cas9 from S. pyogenes (PDB:4OO8). **B.** SMART-BLAST phylogenetic analysis of Cas9 from *F. necrophorum* subsp. *funduliforme* 1_1_36S.

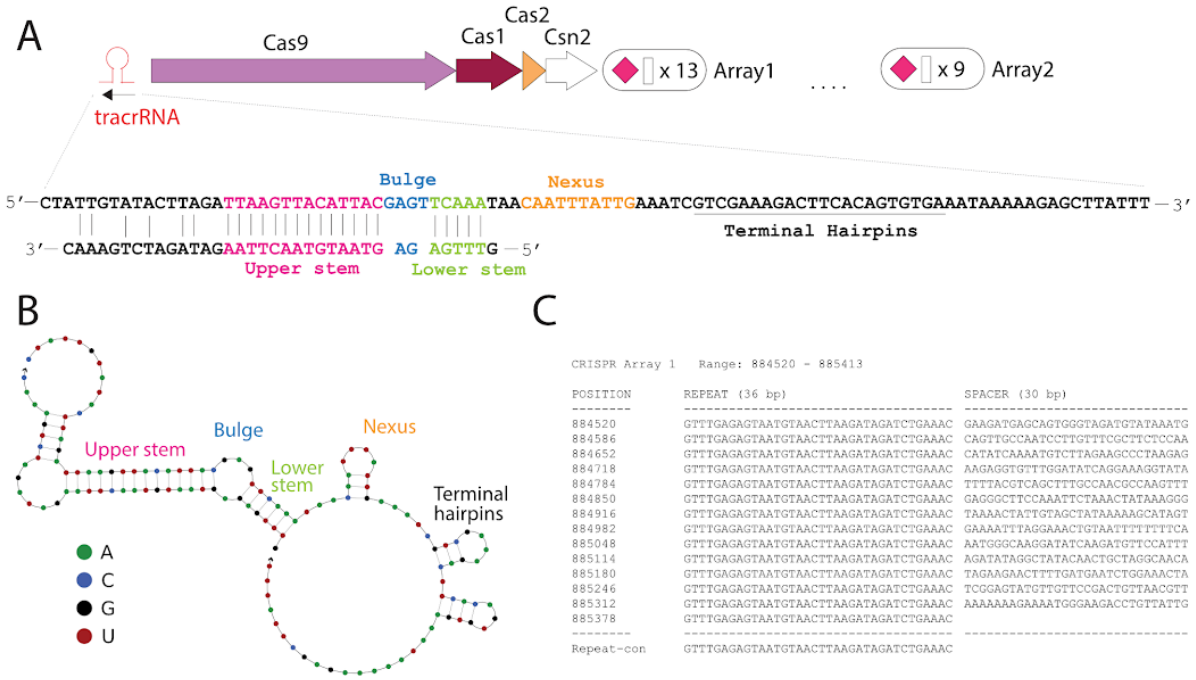


Figure 3-5 Characterization of CRISPR-Cas elements from *F. necrophorum* subsp. *funduliforme* 1_1_36S.A. Predicting the upstream Trans-activating CRISPR RNA (tracrRNA) of Cas9, coupled with the repeat region complementary to the tracrRNA. **B.** Prediction of the secondary structure of Cas9 tracrRNA. **C.** Overview of CRISPR repeats (36 bp) and spacers (30 bp) for the Cas9 locus of F. *necrophorum* subsp. *fundulliforme* 1_1_36S.

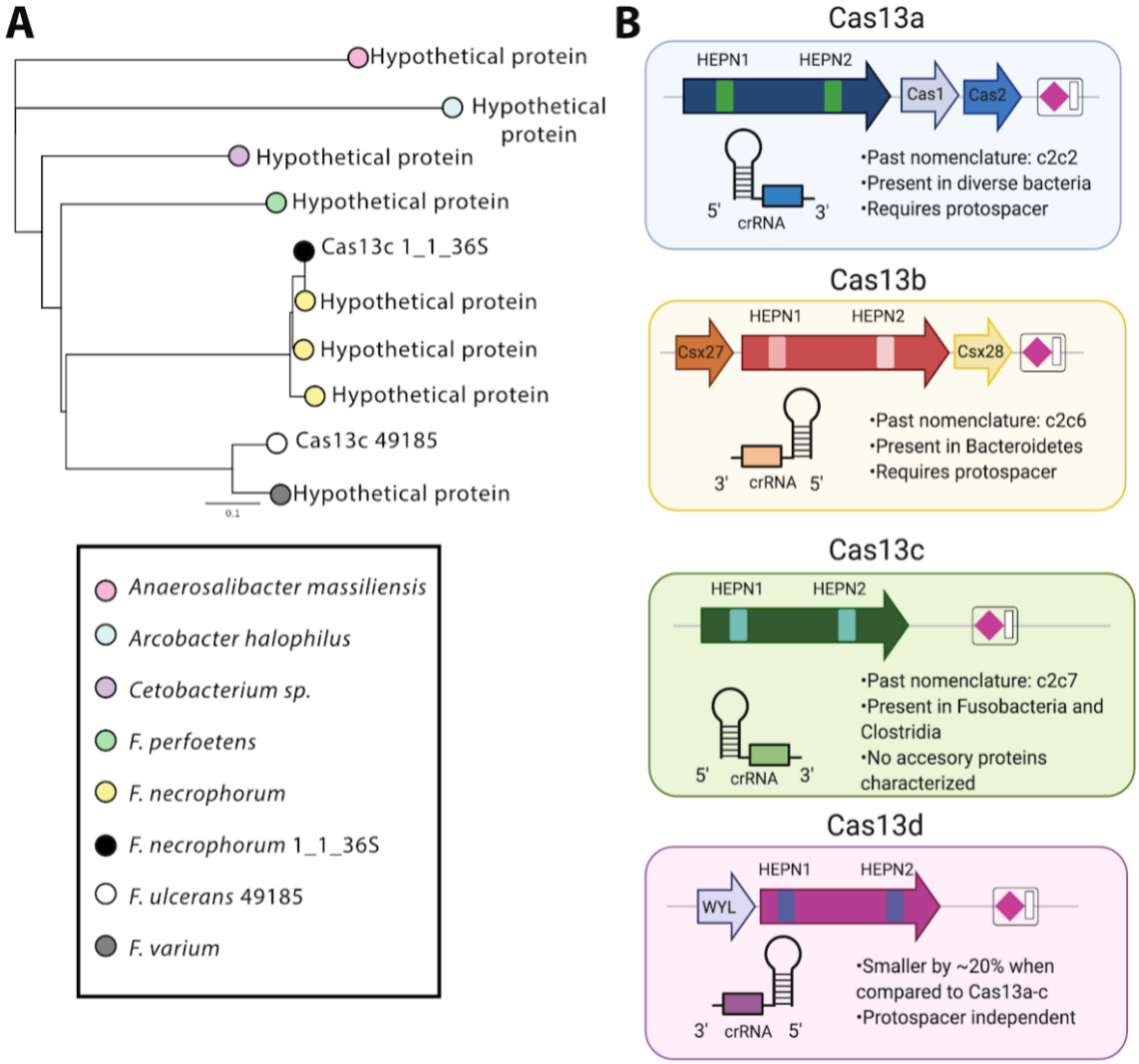


Figure 3-6 Phylogenetic analysis of Cas13c *from F. necrophorum* subsp. *fundiliforme* 1_1_36S. A. SMART-BLAST phylogenetic analysis of Cas13c from *F. necrophorum* susbp. *funduliforme* 1_1_36S. **B.** Differences in type VIA-VID Cas13 effectors.

In addition to Cas9, recognition and targeting of DNA is dependent on the crRNA and tracrRNA complex[34]. For this reason, we characterized the tracrRNA element of the Type II CRISPR-Cas system in *F. necrophorum* subsp. *funduliforme* 1_1_36S **(Figure 3.5)**. The first nucleotides predicted in the nexus include adenine residues that are known to be highly conserved among tracrRNAs of Type II systems as shown in **Figure 3.5** [34] and the bulge interacts through a specific side-chain with the Rec1 domain or lobe through specific side-chain [35]. Both bulge and nexus in tracrRNA are required for DNA cleavage and the spacer sequence detailed in **Figure 3.5B** will determine the location of Cas9 endonucleolytic cleavage [34]. Finally, the stems interact with Cas9 mainly through sequence-independent interactions with the phosphate backbone [34].

The phylogenetic analysis of Cas9 from *F. necrophorum* subsp. *funduliforme* 1_1_36S revealed conservation within the *Fusobacterium* genus **(Figure 3.4B)**. *F. necrophorum* Cas9 has a distant relationship with *Crenarchaeota archaea* which comprise a significant number of microorganisms in the marine environment [36] as well as with the Gram-positive *Streptococcus* and Gram-negative *Neisseria* **(Figure 3.4B)**.

However, while this has not been confirmed, the predicted PAM sequence for this *F. necrophorum* subsp. *funduliforme* 1_1_36S Cas9 is NGG [37], the same as used for the commonly used *S. pyogenes* Cas9 Interestingly enough, the predicted PAM sequence for the Type I-B CRISPR-Cas systems from *F. nucleatum* ATCC 23726 is YNW [37]. Nevertheless, no *Fusobacterium* PAM sites have been confirmed experimentally.

Overall, a total of 3,427 spacers have been identified in Fusobacteria with only 629 matching viral, intergenic or ORFs [38]. Utilizing viruSite [39], the integrated database for viral genomics, most of the 13 spacers of the CRISPR array from *F. necrophorum* subsp. *funduliforme*

1_1_36S were identified as a potentially part of a *Synechococcus*, *Bacillus*, *Enterobacteria*, *Aeromonas* phage genome.

Identification of the *F. necrophorum* subsp. *funduliforme* 1_1_36S and *F. ulcerans* 49185 Type VI CRISPR-Cas system

Finally, Type VI CRISPR-Cas systems are divided into four types due to their low sequence similarity: VI-A, VI-B, VI-C and VI-D. However, they all share the catalytic motif of the HEPN domain [12]. The R-X4-H motifs which are characteristic of HEPN-binding domains in both Cas13c that potentially mediate RNA cleavage are highlighted in **Figure 3.3B** [15]. *F. necrophorum* subsp. *funduliforme* 1_1_36S and *F. ulcerans* 49185 share a 40.51% and 40.64% of identity when compared to the previously identified Cas13c from *F. perfoetens* [40]. The Cas13c from *F. ulcerans* is composed of 1117 aminoacids while the Cas13c from *F. necrophorum* is 1137 aminoacids, which is consistent with the average size for Cas13c of approximately 1120 aminoacids [16].

A phylogenetic analysis of bioinformatically predicted Cas13c uncovered the close phylogenetic relationship between Type VI effectors in *Fusobacterium* **(Figure 3.6A)**. However, it revealed a more distant relationship with the hypothetical proteins encoded by the Gram-negative *Cetobacterium sp.* from the Fusobacteriaceae family as well as *Arcobacter halophilus* from the *Arcobacter* genus, and *Anerosalibacter massilensis* the Clostridieae family **(Figure 3.6A)**. Finally, when compared to other Type VI effectors Cas13c are the least studied and no accessory proteins have been identified like the other recently discovered Type VI systems **(Figure 3.6B)**.

Discussion

As CRISPR loci consist of host repeats and segments of foreign DNA extracted from invading organisms and mobile genetic elements, these repetitive stretches of DNA have proven difficult to resolve using short read DNA technologies that are commonly used to sequence bacterial genomes. The emergence of long-read DNA technologies from Oxford Nanopore [41] and Pacific Biosciences [42] has opened the door for the completion of fragmented genomes; many of which are likely incomplete due to the inability to resolve repeat regions. Previous attempts to determine the conformation, quantity, and diversity of CRISPR-Cas systems in *Fusobacterium* were exceptionally challenging because genome assemblies were incomplete and, in many cases, misannotated. Complete and accurate genome sequences such as the ones hosted by FusoPortal [43] are crucial for the analysis of these systems, which will facilitates the recombinant production of Cas proteins and expansion of our *Fusobacterium* genetic toolkit [43,44].

A large percentage of bacteria harbor adaptive defense mechanisms known as CRISPR-Cas systems, our analysis of CRISPR-Cas systems in *Fusobacterium* uncovered Class 1 and Class 2 systems. Class 2 systems are less common within prokaryotes when compared to Class 1 systems, at most 10% of CRISPR–Cas systems within sequenced genomes belong to Class 2 systems [13]. For that reason, most of the identified CRISPR-Cas systems in our annotated *Fusobacterium* species belong to Class 1 and interestingly, only *F. necrophorum* subsp. *funduliforme* 1_1_36S and *F. ulcerans* 49185 encode for Class 2 systems which is similar to the CRISPR-Cas systems identified in *Leptotrichia*, which belongs to *Fusobacterium* genus, where they are unevenly distributed [45].

Type I CRISPR-Cas systems were found to be the most abundant in the eight *Fusobacterium* strains. They often coexist with Type III CRISPR-Cas systems as shown in *F.*

periodonticum 2_1_31, and it is suspected that Type I systems complement Type III systems functionally [31]. Whereas, Type I target dsDNA, Type III systems target ssDNA, dsDNA, and ssRNA [31]. Particularly, most of *Fusobacterium* genomes encode a complete operon of Type I-B CRISPR-Cas systems which they use to achieve adaptive immunity against viruses. First, the adaption step by Cas1, Cas2 and Cas4 requires the acquisition and inclusion of short segments of exogenous nucleotides into CRISPR loci; then crRNA-processing by Cas6 where the crRNA expression and maturation take place, and the last step is interference by Cas3, Cas5, Cas7 and Cas8 which recognize the crRNA and target nucleic acid [46]. Generally, the Type I-B CRISPR-Cas system utilizes a cascade to target foreign DNA by base pairing CRISPR RNA (crRNA) to protospacers [47]. In the case of *F. ulcerans* 49185 and *F. varium* 27725 which lack the signature *cas8* gene, the interference step could be defective as it has been demonstrated that Cas8 plays a key role in targeting invading DNA [47]. The majority of Type I-B CRISPR-Cas systems in the eight strains analyzed were encoded downstream of the *cas2* gene and/or upstream of the cas6 gene which is in accordance to a CRISPR-Cas system present in the anaerobic rod *Clostridium thermocellum* [48].

Type III-A CRISPR-Cas adaptive immune systems are composed of a protein-RNA complex [29]. The complete loci in *F. periodonticum* 2_1_31 is contains five conserved proteins, and this strain harbors additional proteins that contain the Csm targeting complex which is made up of Cas10/ Csm1, Csm2–Csm5 and the Csm6 ribonuclease. *F. mortiferum* 9817 lacks Csm1 and Csm6 which could result in a non-functional CRISPR-Cas system. The role of these accessory genes may be related to modulation, complementation or extension of these systems [28,30,31].

To determine how CRISPR-Cas system function the Class, Type, and PAM recognition site of the endonucleases needs to be characterized. For the Type II-A CRISPR-Cas system in *F.*

necrophorum subsp. *necrophorum* 1_1_36S, the PAM sequence was previously predicted as NGG and we further characterized its tracrRNA and spacer composition [37]. This system could be utilized for genome modification in *Fusobacterium* strains. However, as shown in this study distinct strains of an organism encode diverse elements in their CRISPR systems and consequently, may have unique requirements for genetic editing.

The most conserved trait in Type VI CRISPR-Cas is the HEPN domain [12]. Type III systems possess RNase activity through the HEPN domain contained in the Csm6 and Csx1, although their role in this specific system remains to be studied [49]. The first report of a Type VI system bioinformatically identified in *Fusobacterium* belonged to *F. perfoetens* [40], a strain that was found in children with diarrhea, nursing infants and in feces of a piglet [50]. Additionally, CRISPR-CAS V1-C systems were only found in *Clostridium* and *Fusobacterium* [12].

The subtypes in the VI systems are based in the HEPN domains which are highly variable, with the most widely studied being Type VI-A [11]. The HEPN domains are found frequently in various defense systems [12], in Cas13d proteins, this domain has been found to orchestrate CRISPR RNA maturation and target cleavage [15] which could suggest a potential role for the two HEPN domains identified in Cas13c from *F. necrophorum* subsp. *funduliforme* 1_1_36S and *F. ulcerans* 49185.

Additionally, CRISPR-Cas spacers represent a successful defense mechanism from past intruders, and the conserved spacers reflect the evolutionary phylogeny or conservation of species and strains [45]. In the studied *Fusobacterium* strains, the spacers were not similar within the spacer-repeat loci and the number of spacers reveals the biological heterogeneity of *Fusobacterium* CRISPR-Cas systems [45]. A higher number of spacers was observed in Type I-B which coincides

with the fact that Type I-B, II-C and III-like CRISPR-Cas systems overall contain an increased number of spacers when compared to other CRISPR-Cas systems [45].

In conclusion, the *Fusobacterium* strains analyzed have endogenous CRISPR-Cas systems that could be exploited as genetic tools for genomic engineering. The development of a CRISPR-Cas gene editing tool in *Fusobacterium* requires the description and detailed examination of the Type/Class of CRISPR array and identifying the PAM recognition sequence. CRISPR-Cas systems are a good candidate for genetic engineering since they depend on the endogenous systems that provide existing components and it is necessary to expand the genetic toolbox as *Fusobacterium* only has a few genetic systems described [1,2].

The utilization of bioinformatics has contributed to the continued discovery of novel CRISPR-Cas systems, such as Type V and Type VI CRISPR-Cas systems. As this bioinformatic study validates the presence of diverse CRISPR-Cas systems in *Fusobacterium,* we hypothesize that this technology could be used to more efficiently introduce gene interruptions, point mutations, and transcriptional control through CRISPR interference (CRISPRi) and activation (CRISPRa) [51]. Furthermore, the CRISPR-Cas technologies and applications that have been developed are rapidly evolving and could enable diverse novel applications [40], which is of significant interest.

Acknowledgments

We thank Dr. Kira Marakova from the computational biology branch at NIH for the insightful information regarding the Cas13c and HEPN domain composition.

Author contributions

AU designed, performed, and analyzed experiments. AU and DJS wrote the manuscript.

Competing Interests

The authors declare that they have no conflicts of interest with the contents of this article.

References

1. Wu C, Al Mamun AAM, Luong TT, Hu B, Gu J, Lee JH, et al. Forward Genetic Dissection of Biofilm Development by Fusobacterium nucleatum: Novel Functions of Cell Division Proteins FtsX and EnvC. MBio. 2018;9. doi:10.1128/mBio.00360-18

2. Casasanta MA, Yoo CC, Udayasuryan B, Sanders BE, Umaña A, Zhang Y, et al. Fusobacterium nucleatum host-cell binding and invasion induces IL-8 and CXCL1 secretion that drives colorectal cancer cell migration. Sci Signal. 2020;13. doi:10.1126/scisignal.aba9157

3. Kinder Haake S, Yoder S, Gerardo SH. Efficient gene transfer and targeted mutagenesis in Fusobacterium nucleatum. Plasmid. 2006;55: 27–38.

4. Zhang Z-T, Jiménez-Bonilla P, Seo S-O, Lu T, Jin Y-S, Blaschek HP, et al. Bacterial Genome Editing with CRISPR-Cas9: Taking Clostridium beijerinckii as an Example. Methods Mol Biol. 2018;1772: 297–325.

5. Nowak CM, Lawson S, Zerez M, Bleris L. Guide RNA engineering for versatile Cas9 functionality. Nucleic Acids Res. 2016;44: 9555–9564.

6. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014;346: 1258096.

7. Makarova KS, Koonin EV. Annotation and Classification of CRISPR-Cas Systems. Methods Mol Biol. 2015;1311: 47–75.

8. Barrangou R, Marraffini LA. CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. Mol Cell. 2014;54: 234–244.

9. Karginov FV, Hannon GJ. The CRISPR system: small RNA-guided defense in bacteria and archaea. Mol Cell. 2010;37: 7–19.

10. Rath D, Amlinger L, Rath A, Lundgren M. The CRISPR-Cas immune system: biology, mechanisms and applications. Biochimie. 2015;117: 119–128.

11. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, et al. An updated evolutionary classification of CRISPR–Cas systems. Nat Rev Microbiol. 2015;13: 722.

12. Shmakov S, Smargon A, Scott D, Cox D, Pyzocha N, Yan W, et al. Diversity and evolution of class 2 CRISPR-Cas systems. Nat Rev Microbiol. 2017;15: 169–182.

13. Pickar-Oliver A, Black JB, Lewis MM, Mutchnick KJ, Klann TS, Gilcrest KA, et al. Targeted transcriptional modulation with type I CRISPR–Cas systems in human cells. Nat Biotechnol. 2019;37: 1493–1501.

14. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. Science. 2010;327: 167–170.

15. Zhang B, Ye Y, Ye W, Perčulija V, Jiang H, Chen Y, et al. Two HEPN domains dictate CRISPR RNA maturation and target cleavage in Cas13d. Nat Commun. 2019;10: 2544.

16. Cox DBT, Gootenberg JS, Abudayyeh OO, Franklin B, Kellner MJ, Joung J, et al. RNA editing with CRISPR-Cas13. Science. 2017;358: 1019–1027.

17. Anantharaman V, Makarova KS, Burroughs AM, Koonin EV, Aravind L. Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. Biol Direct. 2013;8: 15.

18. Zhang Q, Ye Y. Not all predicted CRISPR-Cas systems are equal: isolated cas genes and classes of CRISPR like elements. BMC Bioinformatics. 2017;18: 92.

19. Couvin D, Bernheim A, Toffano-Nioche C, Touchon M, Michalik J, Néron B, et al. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res. 2018;46: W246–W251.

20. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 2011;39: W29–37.

21. Geneious. [cited 2 Sep 2020]. Available: https://www.geneious.com/

22. Snipen L, Liland KH. micropan: an R-package for microbial pan-genomics. BMC Bioinformatics. 2015;16: 79.

23. Wilkins D. Draw Gene Arrow Maps in "ggplot2" [R package gggenes version 0.4.0]. [cited 27 Jul 2020]. Available: https://CRAN.R-project.org/package=gggenes

24. NCBI Staff. NCBI Insights : SmartBLAST: Faster BLASTp search results in a graphical view. 29 Jul 2015 [cited 17 Sep 2020]. Available: https://ncbiinsights.ncbi.nlm.nih.gov/2015/07/29/smartblast/

25. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015;10: 845–858.

26. Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, et al. NUPACK: Analysis and design of nucleic acid systems. J Comput Chem. 2011;32: 170–173.

27. Silas S, Lucas-Elio P, Jackson SA, Aroca-Crevillén A, Hansen LL, Fineran PC, et al. Type III CRISPR-Cas systems can provide redundancy to counteract viral escape from type I systems. Elife. 2017;6. doi:10.7554/eLife.27601

28. Vestergaard G, Garrett RA, Shah SA. CRISPR adaptive immune systems of Archaea. RNA Biol. 2014;11: 156–167.

29. Liu TY, Liu J-J, Aditham AJ, Nogales E, Doudna JA. Target preference of Type III-A CRISPR-Cas complexes at the transcription bubble. Nat Commun. 2019;10: 3001.

30. Haft DH, Selengut J, Mongodin EF, Nelson KE. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS Comput Biol. 2005;1: e60.

31. Shah SA, Alkhnbashi OS, Behler J, Han W, She Q, Hess WR, et al. Comprehensive search for accessory proteins encoded with archaeal and bacterial type III CRISPR-cas gene cassettes reveals 39 new cas gene families. RNA Biol. 2019;16: 530–542.

32. Wei Y, Terns RM, Terns MP. Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. Genes Dev. 2015;29: 356–361.

33. Fuso_necr_1_1_36S_V1 - Genome - Assembly - NCBI. [cited 2 Sep 2020]. Available: https://www.ncbi.nlm.nih.gov/assembly/GCA_000242215.1

34. Briner AE, Donohoue PD, Gomaa AA, Selle K, Slorach EM, Nye CH, et al. Guide RNA functional modules direct Cas9 activity and orthogonality. Mol Cell. 2014;56: 333–339.

35. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science. 2014;343: 1247997.

36. Gribaldo S, Brochier-Armanet C. The origin and evolution of Archaea: a state of the art. Philos Trans R Soc Lond B Biol Sci. 2006;361: 1007–1022.

37. Mendoza BJ, Trinh CT. In Silico Processing of the Complete CRISPR-Cas Spacer Space for Identification of PAM Sequences. Biotechnol J. 2018;13: e1700595.

38. Shmakov SA, Sitnik V, Makarova KS, Wolf YI, Severinov KV, Koonin EV. The CRISPR Spacer Space Is Dominated by Sequences from Species-Specific Mobilomes. MBio. 2017;8. doi:10.1128/mBio.01397-17

39. Stano M, Beke G, Klucar L. viruSITE-integrated database for viral genomics. Database . 2016;2016. doi:10.1093/database/baw162

40. Wang J, Zhang C, Feng B. The rapidly advancing Class 2 CRISPR-Cas technologies: A customizable toolbox for molecular manipulations. J Cell Mol Med. 2020;24: 3256–3270.

41. Oxford Nanopore Technologies. [cited 3 Sep 2020]. Available: https://nanoporetech.com/

42. Tallon L, Schmutz J, Garimella K, Steven Wiley H. Home - PacBio - Sequence with Confidence. [cited 3 Sep 2020]. Available: https://www.pacb.com/

43. Sanders BE, Umana A, Lemkul JA, Slade DJ. FusoPortal: an Interactive Repository of Hybrid MinION-Sequenced Fusobacterium Genomes Improves Gene Identification and Characterization. mSphere. 2018;3: e00228–18.

44. Michelle Todd S, Settlage RE, Lahmers KK, Slade DJ. Fusobacterium Genomics Using MinION and Illumina Sequencing Enables Genome Completion and Correction. mSphere. 2018;3: e00269–18.

45. Watanabe S, Cui B, Kiga K, Aiba Y, Tan X-E, Sato'o Y, et al. Composition and Diversity of CRISPR-Cas13a Systems in the Genus Leptotrichia. Front Microbiol. 2019;10: 2838.

46. Bhaya D, Davison M, Barrangou R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annu Rev Genet. 2011;45: 273–297.

47. Cass SDB, Haas KA, Stoll B, Alkhnbashi OS, Sharma K, Urlaub H, et al. The role of Cas8 in type I CRISPR interference. Biosci Rep. 2015;35. doi:10.1042/BSR20150043

48. Richter H, Zoephel J, Schermuly J, Maticzka D, Backofen R, Randau L. Characterization of CRISPR RNA processing in Clostridium thermocellum and Methanococcus maripaludis. Nucleic Acids Res. 2012;40: 9887–9896.

49. Niewoehner O, Jinek M. Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. RNA. 2016;22: 318–329.

50. Van Assche PF, Wilssens AT. Fusobacterium perfoetens (Tissier) Moore and Holdeman 1973: Description and Proposed Neotype Strain. Int J Syst Bacteriol. 1977;27: 1–5.

51. Horlbeck MA, Gilbert LA, Villalta JE, Adamson B, Pak RA, Chen Y, et al. Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. Elife. 2016;5. doi:10.7554/eLife.19760

Chapter 4

Conclusions and Future Directions

The human microbial community comprises a diverse set of microorganisms distributed throughout the entirety of the body [1]. In the oral cavity, *Fusobacterium* is a common member of the oral biofilm that engages in a variety of interactions with other microorganisms via outer membrane proteins and secreted adhesins [2]. *Fusobacterium* acts as a bridging microorganism, as it coaggregates with early and late colonizers. As bridging bacteria *Fusobacterium* are a key component in biofilm regulation and growth [3].

Most notably, *Fusobacterium* is a heterogeneous genus that is highly unique as it encodes for an overabundance of Type 5 secreted autotransporters that mediate invasion, pathogenesis, and survival [4]. *Fusobacterium* is also known to be a commensal microorganism that is implicated in a wide assortment of human and animal diseases [2]. The development of new technologies such as metagenomics and metatranscriptomics has permitted an assessment of the role of *Fusobacterium* in several diseases. More recently, *F. nucleatum* has been associated with the promotion of cancer; colorectal cancer [5,6], breast cancer [7] and esophageal cancer [8]. The role of bacteria in cancer has been explored since the discovery that *Helicobacter pylori can* lead to the development of gastric cancer [9].

F. nucleatum is a multifaceted microorganism that is capable of invading host cells and modulating host signaling pathways. Briefly, *F. nucleatum* accelerates the development of colorectal cancer by attaching through its large secreted autotransporter adhesin Fap2 to host-cell exposed Gal-GalNac, which is an overexpressed sugar residue in colorectal cancer tissues [10]. Next, it promotes cell proliferation via interaction of FadA with E-cadherin that leads to transcription of Wnt-gene by β-catenin [11]. Additionally, LPS engages with TLR4 regulating

miR21 and promoting uncontrolled cell growth and chemoresistance by interacting with other miRNAS [12]. In the later stages of colorectal cancer, Fap2 potentiates tumorigenesis by interacting with TIGIT and promoting NK cell death and metastasis progression by upregulating cytokine (CXCL-1 and IL-8) productions [13,14]. Finally, new data shows that F. nucleatum can induce double stranded DNA break which could signal this bacterium to initiate mutations that cause cancer [15,16].

Although the *Fusobacterium* interactome in colorectal cancer and other diseases is being rapidly studied many questions remain unanswered as to the specific role of outer membrane proteins and Type 5 secreted autotransporter in *Fusobacterium*-bacteria and *Fusobacterium*-host interactions in disease. A critical hindrance in the *Fusobacterium* field is the genetic recalcitrance of most strains in the *Fusobacterium* genus. Recent breakthroughs have permitted the development of multiple new genetic systems to genetically modified *F. nucleatum* ATCC 23726 [13,17]. However, efforts to achieve genetic manipulation in *Fusobacterium* started in the early 2000's [18,19] with little success within the group as a whole. To this date only electroporation and sonoporation methods have been used to introduce exogenous DNA to a total of 3 *F. nucleatum* strains; *F. nucleatum* subsp. *nucleatum* ATCC 23726 [13,17,18,20], *F. nucleatum* subsp. *polymorphum* 10953 [19], and *F. nucleatum* subsp. *polymorphum* 12230 [21,22].

There is a major gap in knowledge as to the reason why only a few strains are able to acquire target DNA and achieve exogenous incorporation into their genome by recombination or establishment of a new episome. A potential answer is the presence of robust R-M systems in *Fusobacterium*, as discussed previously in this dissertation and shown in **Figure 4.1**. R-M systems are defense mechanisms that recognize foreign DNA by the methylation patterns that are different

than its own. The development of FusoPortal [23], an online database of *Fusobacterium* complete and highly accurate genomes allowed for a bioinformatic characterization of R-M systems.

We identified that the genome size of the strains does not correlate with the amount of R-M systems and that *Fusobacterium* encodes for a higher quantity of R-M when compared to other bacteria. The enrichment of R-M systems likely impedes the uptake of DNA as its restriction endonucleases discern the exogenous DNA from host DNA. This has limited the genetic manipulation of *Fusobacterium* to a few model strains. Moreover, R-M systems in *Fusobacterium* are composed of DNA methyltransferases that are predicted to have diverse recognition sequences for that reason, a mixture of purified methyltransferases for all *Fusobacterium* strains might not be efficient.

A technique that could elucidate the methylome in all strains of *Fusobacterium* is single-molecule, real-time SMRT DNA sequencing. The Pacifc Biosciences (PacBio) Single Molecule Real Time (SMRT) sequencing platform has the ability of resolving complete bacterial genomes [24,25] while defining the methylation status of every sequenced nucleotide[26] Additionally, it could reveal the activity. specificity, and fidelity of DNA methyltransferases. The unification of PacBio SMRT sequencing and biochemical characterization of DNA methyltransferases could lead into a solution to tackle *Fusobacterium*'s genetic recalcitrance. Another proposed method to surpass R-M systems is RM-silent syngenicDNA which removes recognition sequences from plasmids before transformation [27]. This method achieved a significant improvement in the transformation efficiency of the clinically relevant *S. aureus* USA300[27]. However, the efficacy of this method is yet to be explored in *Fusobacterium* strains.

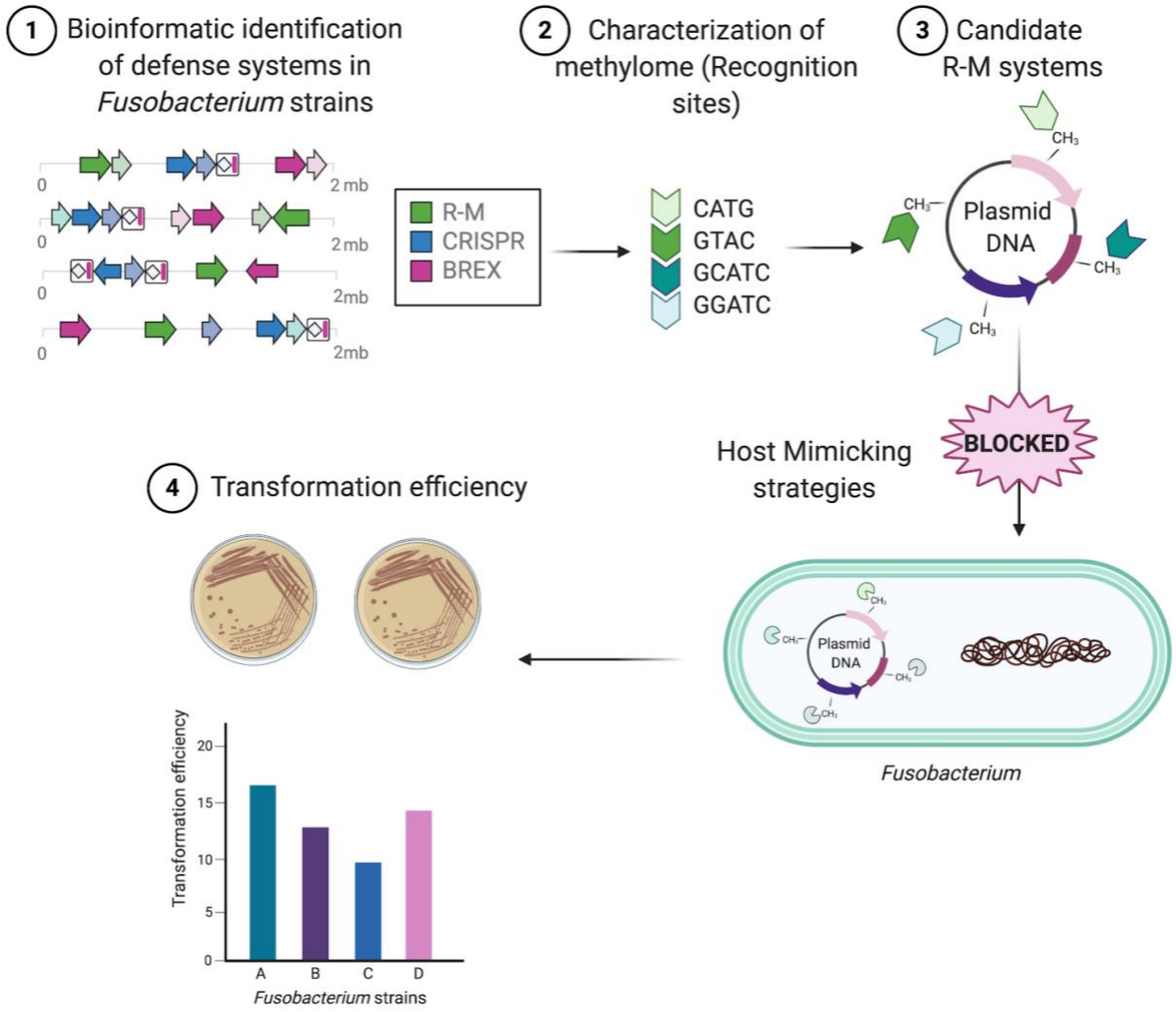


Figure 4-1 Schematic of the utilization of R-M systems to overcome Fusobacterium recalcitrance. The unification of bioinformatic study of defense systems, PacBio SMRT sequencing and host-mimicking techniques could lead into a solution to tackle *Fusobacterium* genetic recalcitrance.

The most widely accepted method to deceit R-M systems is the use of host-mimicking strategies **(Figure 4.1)**. Most commonly, recombinant purified enzymes are utilized to achieve methylation of plasmid DNA in the same recognition sequences as the target bacterium. This method achieved a 50 fold improvement in the transformation efficiency/ug in *F. nucleatum* ATCC 23726 when compared to non-treated suicide pDNA. It is important to denote that the utilization of the purified methyltransferases, M.Fnu23I and M.Fnu23II from *F. nucleatum* ATCC 23726 only protected suicide pDNA for the same strain. Both methyltransferases were utilized through the coexpression and *in vitro* approach to attempt to condition suicide plasmids for *F. nucleatum* ATCC 25586, *F. necrophorum* subsp. *funduliforme* 1_1_36S, *F. nucleatum* subsp. *animalis* 7_1, *F. nucleatum* subsp. *animalis* 4_8 and, *F. nucleatum* subsp. *polymorphum* 10953. Nevertheless, all of these produced negative results with no integration onto the chromosome. However, utilizing an episomal plasmid and a transposon plasmid conditioned with M.Fnu23I and M.Fnu23II we were able to expand bacterial transformation in *F. nucleatum* subsp. *polymorphum* 10953 and *F. nucleatum* ATCC 25586.

For this reason, the next step was to utilize the coexpression or in *vitro* method with a selected number of methyltransferases from *F. nucleatum* ATCC 25586, *F. necrophorum* subsp. *funduliforme* 1_1_36S and, *F. nucleatum subsp. animalis* 7_1. In total, we attempted to transform six strains by combining 17 methyltransferases belonging to type I, type II and type III R-M systems. To further characterize the impact of DNA methylation in transformation, we created a crude-extract method, the different strains of *Fusobacterium* were lysed and exogenous naked pDNA was added along with SAM and cutsmart buffer. However, the combination of electroporation and these methods resulted in no transformation.

Most of the identified DNA methyltransferases in *Fusobacterium* are adenine methyltransferases and interestingly the addition of a methyl group (—CH_3) to the adenine nucleotide is exclusive for prokaryotes. Adenine methyltransferases are known to regulate virulence in bacterial pathogens such as *Aeromonas hydrophila* [28], *Photorhabdus luminescens* [29], *Neisseria gonorrhoeae* [30], *N. meningitidis* [30] and *Mycobacterium tuberculosis* [31]. The overexpression of DNA adenine methyltransferase in several bacterial species has led to the differential expression of genes, likely through phase variation [30].

Fusobacterium is a commensal oral microorganism that can turn on its pathogenicity by uncharacterized mechanisms. A potential pathway to activate its pathogenesis is phase variation. Phase variation is yet to be characterized in *Fusobacterium*, however the methyltransferases present could be altering global DNA methylation patterns leading to phase variation of autotransporters as it has been suggested previously[32]. Autotransporters are known virulence factors that are essential for host-cell interactions and necessary for infectious diseases development[33].

Altogether, further studies will be necessary to understand the role of DNA methyltransferases in phase variation and regulation of virulence in *Fusobacterium*. Integration of functional studies, optimization of host mimicking strategies and physical transformation could lead to better understanding of *Fusobacterium* bacterial pathogenesis. These studies could disrupt the genetic barrier and achieve an unlimited catalog of *Fusobacterium* model strains including clinically relevant strains.

The development of new genetic tools to produce gene interruptions, point mutations, and regulate transcription must also be considered. CRISPR-Cas systems are an alternative and are increasingly emerging as powerful tools for genetic manipulation and transcriptional control in

prokaryotic and eukaryotic systems **(Figure 4.2)**. As *Fusobacterium* is a genus of opportunistic pathogens for which there are no broadly available genetic systems, native CRISPR-Cas systems could be used to fill this methodological gap. The characterization of these systems as a result of accurate genome sequences has enabled the identification of all CRISPR components, which will guide future molecular biology studies. These technologies have the potential to be controlled through chemical induction, allowing for the study of essential genes through the CRISPRi, or the overexpression of target proteins using CRISPRa.

In *Clostridium spp*., to surpass the toxic repercussions of *S. pyogenes* Cas9 or *Acidaminococcus sp.* Cpf1 endonucleases, endogenous CRISPR-Cas (Cas9, dCas9/CRISPRi and Cpf1/Cas12a) tools were developed [34]. Recently, these tools have resulted in an enhancement of the genetic manipulation and study of *Clostridium*'s biology [34]. In a similar way to this study, the first step was to identify the presence and composition of CRISPR systems within that organism. In the case of *Fusobacterium,* the analyzed strains encode for slightly different CRISPR-Cas components, which may result in different requirements for editing. The experimental identification and validation of the PAM sequence is the next step as it has been bioinformatically predicted for *F. necrophorum* subsp. *fundiliforme* 1_1_36S as NGG.

Although the CRISPR-Cas systems developed in *Clostridium* are highly efficient there is no universal system as all promoters and plasmids may not be compatible with every strain [34]. In *Fusobacterium*, the same issue may be encountered but could be optimized by modifying the preexisting system of a closely related species or strain [34]. Additionally, the utilization of endogenous CRISPR-Cas systems from the organism of interest is a feasible alternative [34]. The bioinformatic validation of the presence of CRISPR-Cas systems in *Fusobacterium* will broaden the tools for genetic manipulation as only counterselection markers have been previously

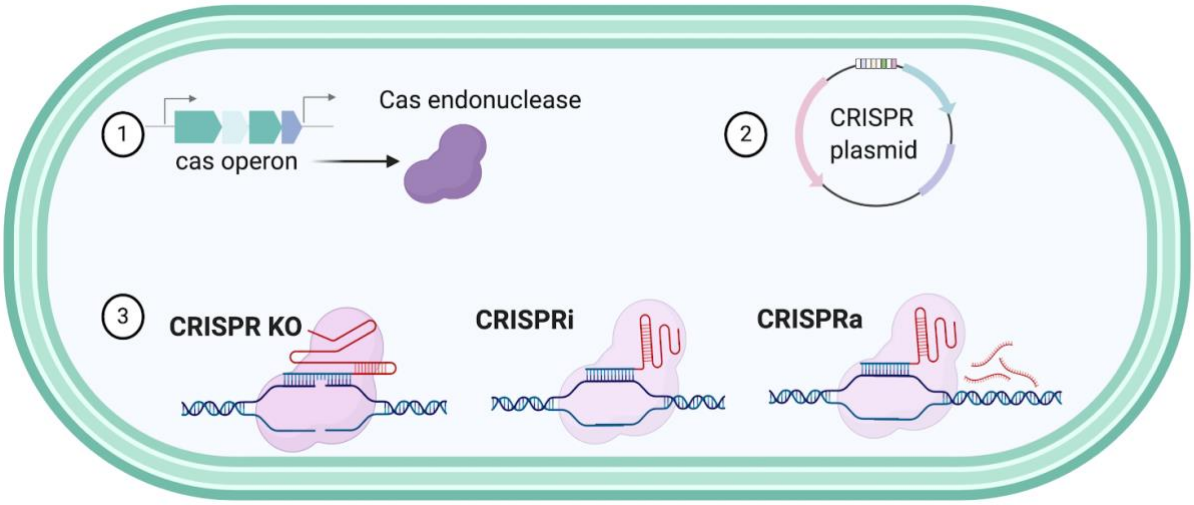


Figure 4-2 The use of native CRISPR-Cas systems as genetic tools could lead to the development of novel genetic systems. These tools could improve genome editing of *Fusobacterium* by creating deletion systems, characterization of essential genes through the CRISPRi, or the overexpression proteins using CRISPRa.

developed [13,17]. and the utilization of native CRISPR-Cas proteins will likely result in a more efficient system. Only by continuing to improve the genetic manipulation of *Fusobacterium* we will be able to understand its biology and pathogenesis, as the characterization of these microbes has been delayed by the lack of genetic tools.

Moreover, the novel Type VI CRISPR-Cas systems discovered in *F. necrophorum subsp. fundiliforme* 1_1_36S and *F. ulcerans* 49185 consisted of a Cas13 protein that is characterized as a single-effector RNA-guided ribonuclease could enable numerous novel applications; such as nucleic acid detection [35], RNA editing [36]; and knockdown of target RNA [37]. Nevertheless, mechanistic studies will be necessary to further comprehend CRISPR-Cas systems and their potential applications in the *Fusobacterium* field.

The unification of the molecular tools presented in this dissertation will aid in the study of *Fusobacterium* basic biology. Furthermore, a closer study of *Fusobacterium* associated diseases may reveal a causal relationship. For this reason, it is necessary to characterize its *Fusobacterium*-bacteria and *Fusobacterium*-host interactome to determine how to address this anaerobic microorganism in infections.

The importance of addressing the knowledge gaps in *Fusobacterium* is fundamental to understanding the effect of these bacteria in disease, as patients with higher bacterial loads show poorer prognosis in colorectal cancer and periodontal disease. A major emphasis must be taken to identify virulence factors, the T5SS and its secreted and exposed of outer membrane proteins. Only a comprehensive understanding of *Fusobacterium* will allow the development of precise treatment.

This dissertation describes the exploitation of *Fusobacterium* defense systems as biochemical tools to improve *F. nucleatum* transformation efficiency and expand the catalog genetic systems available. The first characterization of R-M systems in *Fusobacterium*

encompassing six strains and over 15 DNA methyltransferases lead to the development of two methods to utilize M.Fnu23I and M.Fnu23II to circumvent restriction endonucleases that recognize exogenous DNA. Moreover, we bioinformatically described in detail the composition of CRISPR-Cas systems which lead to major advances in the creation of new genetic systems and tools for the *Fusobacterium* field. Altogether, the knowledge gathered from these studies can overcome the genetic recalcitrance of many *Fusobacterium* strains, this will accelerate our comprehension of how this opportunistic pathogen contributes to diseases and potentially explore new therapeutic treatments.

References

1. Malla MA, Dubey A, Kumar A, Yadav S, Hashem A, Abd Allah EF. Exploring the Human Microbiome: The Potential Future Role of Next-Generation Sequencing in Disease Diagnosis and Treatment. Front Immunol. 2018;9: 2868.

2. Brennan CA, Garrett WS. Fusobacterium nucleatum — symbiont, opportunist and oncobacterium. Nature Reviews Microbiology. 2019. pp. 156–166. doi:10.1038/s41579-018-0129-6

3. Bradshaw DJ, Marsh PD, Watson GK, Allison C. Role of Fusobacterium nucleatum and coaggregation in anaerobe survival in planktonic and biofilm oral microbial communities during aeration. Infect Immun. 1998;66: 4729–4732.

4. Umana A, Sanders BE, Yoo CC, Casasanta MA, Udayasuryan B, Verbridge SS, et al. Utilizing Whole Fusobacterium Genomes to Identify, Correct, and Characterize Potential Virulence Protein Families. J Bacteriol. 2019. doi:10.1128/JB.00273-19

5. Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, et al. Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. Genome Res. 2012;22: 292–298.

6. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, et al. Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. Genome Res. 2012;22: 299–306.

7. Parhi L, Alon-Maimon T, Sol A, Nejman D, Shhadeh A, Fainsod-Levi T, et al. Breast cancer colonization by Fusobacterium nucleatum accelerates tumor growth and metastatic progression. Nat Commun. 2020;11: 3259.

8. Yamamura K, Baba Y, Nakagawa S, Mima K, Miyake K, Nakamura K, et al. Human Microbiome Fusobacterium Nucleatum in Esophageal Cancer Tissue Is Associated with Prognosis. Clin Cancer Res. 2016;22: 5574–5581.

9. Ishaq S, Nunn L. Helicobacter pylori and gastric cancer: a state of the art review. Gastroenterol Hepatol Bed Bench. 2015;8: S6–S14.

10. Abed J, Emgård JEM, Zamir G, Faroja M, Almogy G, Grenov A, et al. Fap2 Mediates Fusobacterium nucleatum Colorectal Adenocarcinoma Enrichment by Binding to Tumor-Expressed Gal-GalNAc. Cell Host Microbe. 2016;20: 215–225.

11. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/β-catenin signaling via its FadA adhesin. Cell Host Microbe. 2013;14: 195–206.

12. Yang Y, Weng W, Peng J, Hong L, Yang L, Toiyama Y, et al. Fusobacterium nucleatum Increases Proliferation of Colorectal Cancer Cells and Tumor Development in Mice by

Activating Toll-Like Receptor 4 Signaling to Nuclear Factor−κB, and Up-regulating Expression of MicroRNA-21. Gastroenterology. 2017. pp. 851–866.e24. doi:10.1053/j.gastro.2016.11.018

13. Casasanta MA, Yoo CC, Udayasuryan B, Sanders BE, Umaña A, Zhang Y, et al. Fusobacterium nucleatum host-cell binding and invasion induces IL-8 and CXCL1 secretion that drives colorectal cancer cell migration. Sci Signal. 2020;13. doi:10.1126/scisignal.aba9157

14. Gur C, Ibrahim Y, Isaacson B, Yamin R, Abed J, Gamliel M, et al. Binding of the Fap2 protein of Fusobacterium nucleatum to human inhibitory receptor TIGIT protects tumors from immune cell attack. Immunity. 2015;42: 344–355.

15. Sayed IM, Chakraborty A, Abd El-Hafeez AA, Sharma A, Sahan AZ, Huang WJM, et al. The DNA Glycosylase NEIL2 Suppresses Fusobacterium-Infection-Induced Inflammation and DNA Damage in Colonic Epithelial Cells. Cells. 2020;9. doi:10.3390/cells9091980

16. Geng F, Zhang Y, Lu Z, Zhang S, Pan Y. Fusobacterium nucleatum Caused DNA Damage and Promoted Cell Proliferation by the Ku70/p53 Pathway in Oral Cancer Cells. DNA Cell Biol. 2020;39: 144–151.

17. Wu C, Al Mamun AAM, Luong TT, Hu B, Gu J, Lee JH, et al. Forward Genetic Dissection of Biofilm Development by Fusobacterium nucleatum: Novel Functions of Cell Division Proteins FtsX and EnvC. MBio. 2018;9. doi:10.1128/mBio.00360-18

18. Kinder Haake S, Yoder S, Gerardo SH. Efficient gene transfer and targeted mutagenesis in Fusobacterium nucleatum. Plasmid. 2006;55: 27–38.

19. Haake SK, Yoder SC, Attarian G, Podkaminer K. Native plasmids of Fusobacterium nucleatum: characterization and use in development of genetic systems. J Bacteriol. 2000;182: 1176–1180.

20. Coppenhagen-Glazer S, Sol A, Abed J, Naor R, Zhang X, Han YW, et al. Fap2 of Fusobacterium nucleatum is a galactose-inhibitable adhesin involved in coaggregation, cell adhesion, and preterm birth. Infect Immun. 2015;83: 1104–1113.

21. Ikegami A, Chung P, Han YW. Complementation of the fadA mutation in Fusobacterium nucleatum demonstrates that the surface-exposed adhesin promotes cellular invasion and placental colonization. Infect Immun. 2009;77: 3075–3079.

22. Han YW, Ikegami A, Rajanna C, Kawsar HI, Zhou Y, Li M, et al. Identification and characterization of a novel adhesin unique to oral fusobacteria. J Bacteriol. 2005;187: 5330–5340.

23. Sanders BE, Umana A, Lemkul JA, Slade DJ. FusoPortal: an Interactive Repository of Hybrid MinION-Sequenced Fusobacterium Genomes Improves Gene Identification and Characterization. mSphere. 2018;3. doi:10.1128/mSphere.00228-18

24. Forde BM, McAllister LJ, Paton JC, Paton AW, Beatson SA. SMRT sequencing reveals differential patterns of methylation in two O111:H- STEC isolates from a hemolytic uremic syndrome outbreak in Australia. Sci Rep. 2019;9: 9436.

25. Tallon L, Schmutz J, Garimella K, Steven Wiley H. Home - PacBio - Sequence with Confidence. [cited 3 Sep 2020]. Available: https://www.pacb.com/

26. Clark TA, Murray IA, Morgan RD, Kislyuk AO, Spittle KE, Boitano M, et al. Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing. Nucleic Acids Res. 2012;40: e29.

27. Johnston CD, Cotton SL, Rittling SR, Starr JR, Borisy GG, Dewhirst FE, et al. Systematic evasion of the restriction-modification barrier in bacteria. Proc Natl Acad Sci U S A. 2019;116: 11454–11459.

28. Erova TE, Pillai L, Fadl AA, Sha J, Wang S, Galindo CL, et al. DNA adenine methyltransferase influences the virulence of Aeromonas hydrophila. Infect Immun. 2006;74: 410–424.

29. Payelleville A, Lanois A, Gislard M, Dubois E, Roche D, Cruveiller S, et al. DNA Adenine Methyltransferase (Dam) Overexpression Impairs Photorhabdus luminescens Motility and Virulence. Front Microbiol. 2017;8: 1671.

30. Seib KL, Jen FE-C, Scott AL, Tan A, Jennings MP. Phase variation of DNA methyltransferases and the regulation of virulence and immune evasion in the pathogenic Neisseria. Pathog Dis. 2017;75. doi:10.1093/femspd/ftx080

31. Grover S, Gangwar R, Jamal S, Ali S, Nisaa K, Ehtesham NZ, et al. Mycobacterial Methyltransferases: Significance in Pathogenesis and Virulence. In: Hasnain SE, Ehtesham NZ, Grover S, editors. Mycobacterium Tuberculosis: Molecular Infection Biology, Pathogenesis, Diagnostics and New Interventions. Singapore: Springer Singapore; 2019. pp. 103–122.

32. Desvaux M, Khan A, Beatson SA, Scott-Tucker A, Henderson IR. Protein secretion systems in Fusobacterium nucleatum: genomic identification of Type 4 piliation and complete Type V pathways brings new insight into mechanisms of pathogenesis. Biochim Biophys Acta. 2005;1713: 92–112.

33. Nishimura K, Tajima N, Yoon Y-H, Park S-Y, Tame JRH. Autotransporter passenger proteins: virulence factors with common structural themes. J Mol Med . 2010;88: 451–458.

34. McAllister KN, Sorg JA. CRISPR Genome Editing Systems in the Genus Clostridium: a Timely Advancement. J Bacteriol. 2019;201. doi:10.1128/JB.00219-19

35. Gootenberg JS, Abudayyeh OO, Lee JW, Essletzbichler P, Dy AJ, Joung J, et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. Science. 2017;356: 438–442.

36. Cox DBT, Gootenberg JS, Abudayyeh OO, Franklin B, Kellner MJ, Joung J, et al. RNA editing with CRISPR-Cas13. Science. 2017;358: 1019–1027.

37. Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, Cox DBT, et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science. 2016;353: aaf5573.