



UNIVERSIDADE CATÓLICA PORTUGUESA | INSTITUTO DE CIÊNCIAS DA SAÚDE

MESTRADO INTEGRADO EM MEDICINA DENTÁRIA

**O Proteoma do Microbioma Oral Humano – que funções exercem as proteínas microbianas?**

*Dissertação apresentada à Universidade Católica Portuguesa para obtenção do grau de Mestre em Medicina Dentária*

Por:

Jenifer Amorim Dos Santos

Viseu, 2014





UNIVERSIDADE  
CATÓLICA PORTUGUESA | INSTITUTO DE  
CIÊNCIAS DA SAÚDE

## **O Proteoma do Microbioma Oral Humano – que funções exercem as proteínas microbianas?**

*Dissertação apresentada à Universidade Católica Portuguesa para obtenção do grau de Mestre em Medicina Dentária*

Orientadora: Professora Doutora Maria José Correia

Co-Orientadora: Professora Doutora Marlene Barros

Por:

Jenifer Amorim Dos Santos

Viseu, 2014



*“A winner is a dreamer who never gives up”*  
Nelson Mandela



Dedico este trabalho aos meus pais, Anabela e Alberto, que sempre acreditaram em mim e sem os quais nada teria sido possível.



## Agradecimentos

À Professora Doutora Maria José Correia,

Pela sua total disponibilidade desde o primeiro ano. Não consigo encontrar as palavras para agradecer as horas que me atendeu a acompanhar o meu trabalho.

A Professora Doutora Marlene Barros,

Por todo o apoio, pela disponibilidade e por sempre me ter ajudado em tudo.

Ao Professor Doutor Nuno Rosa,

Por toda a disponibilidade, pela extrema paciência e por ter escutado sempre as minhas dúvidas.

Aos meus pais,

Pela motivação, pelo carinho incondicional e por serem exemplos de coragem.

Aos meus avós,

Apesar da distância sabem que nunca me esqueço de vocês.

As minhas primas Ana Paula e Anabela,

Pelo apoio dado durante estes últimos cinco anos.

Ao J<sup>2</sup>,

Por todos os bons momentos passados que ficarão gravados na minha memória.

Aos meus amigos,

Por todos os bons momentos passados e pela vossa amizade.



# Índice

Agradecimentos .....	v
Lista de Figuras.....	ix
Lista de Tabelas.....	xii
Anexos .....	xii
Lista de abreviaturas .....	xiii
Resumo.....	xv
Abstract.....	xvii
Nota prévia .....	xix
1. INTRODUÇÃO.....	21
1.1 O Microbioma Humano .....	25
1.1.1 O microbioma oral .....	27
1.1.1.1 Quorum sensing .....	30
1.1.1.2 Alterações do microbioma oral .....	32
1.2 Os fatores de virulência bacterianos .....	33
1.2.1 Definição de um fator de virulência.....	33
1.2.2 Principais mecanismos de virulência .....	34
1.2.2.1 Adesão .....	36
1.2.2.2 Invasão.....	36
1.2.2.3 Mecanismos de sobrevivência intracelular .....	37
1.2.2.4 As toxinas.....	38
1.2.2.5 Regulação dos fatores de virulência .....	40
1.3 Ferramentas bioinformáticas.....	41
1.3.1 Aplicação da bioinformática.....	41
1.3.2 UniprotKB .....	42
1.3.3 OralOma e OralCard.....	43

1.3.4 Gene Ontology.....	43
2. OBJETIVOS .....	49
3. MATERIAIS E METÓDOS .....	55
3.1 Utilização das proteínas catalogadas de origem microbiana depositadas no OralCard .....	57
3.2 Análise dos processos biológicos em que participam as proteínas bacterianas .....	58
3.3 Identificação das proteínas bacterianas no microbioma oral que interagem com outros microrganismos e com o hospedeiro .....	66
3.4 Classificação ontológica das proteínas que interagem com outros microrganismos e com o hospedeiro .....	67
4. RESULTADOS E DISCUSSÃO .....	73
4.1 Catalogação das proteínas de origem microbiana .....	75
4.3 Análise da classificação ontológica das proteínas microbianas do microbioma oral que interagem com outros microrganismos e com o hospedeiro .....	81
4.4 A função das proteínas bacterianas envolvidas na patogénese .....	87
4.5 Associação entre os mecanismos moleculares da patogénese e as principais patologias da cavidade oral .....	94
5. CONCLUSÃO .....	95
6. BIBLIOGRAFIA .....	95
7. ANEXOS.....	95

## **Lista de Figuras**

### **Introdução**

**Figura 1** - As ontologias são uma rede complexa..... 45

**Figura 2** - Representação em árvore dos termos antecedentes e descendentes do “GO:0044403 symbiosis encompassing mutualismo through parasitism”. . 47

### **Objetivos**

**Figura 3**- Fluxograma que ilustra a estratégia de trabalho usada para a realização deste trabalho. As setas indicam ações, os quadrados objetivos e os losângulos as ferramentas bioinformáticas e as bases de dados usadas ..... 53

### **Materiais e Métodos**

**Figura 4**- Figura retirada do portal Web da ferramenta bioinformática Agbase (<http://www.agbase.msstate.edu/>). Para obter as ferramentas bioinformáticas “GoRetriever” ou “GoSlim Viewer” deve-se selecionar o campo “Tools”, e, posteriormente, uma dessas ferramentas. .... 59

**Figura 5**- Página inicial da ferramenta informática GoRetriever obtida através do site da Agbase (<http://www.agbase.msstate.edu/>). .... 59

**Figura 6**- Descarregar o ficheiro “GO Summary file (for use in GOSlimViewer)” para usar na ferramenta “SlimViewer” ..... 60

**Figura 7**-Página inicial da ferramenta informática GoSlimViewer obtida através do site da Agbase. Seleção da opção “TIGR Prokaryote (BP only)” na seção “select GoSlim Set”. .... 61

**Figura 8**- Selecionar a opção “View Biological Process”. De seguida, surgem os processos biológicos em que estão envolvidas as proteínas bacterianas do microbioma oral. Ao lado do nome do processo está indicado o número de proteínas envolvidas em cada um destes. .... 62

**Figura 9**- Selecionar a opção “View accesions for each slim id” na página dos resultados da ferramenta bioinformática GOSlimViewer..... 64

<b>Figura 10</b> - Identificação dos microrganismos associados a cada proteína nas 829 proteínas com pouca informação acerca dos processos biológicos em que estão envolvidas.....	65
<b>Figura 11</b> - Figura do STRAP onde estão representadas, por um triângulo verde, as proteínas com mais impacto no hospedeiro na subcategoria “Interaction with cells and organisms”. .....	67
<b>Figura 12</b> - Inserir na caixa “UniProt identifiers” do “Retriever” as 298 proteínas envolvidas na patogénese. Após isto, selecionar “UniProtKB (298)”, e, por fim, através da opção “download”, obtêm-se os resultados numa folha da Microsoft Excel® 2007.....	69
<b>Figura 13</b> - Resultados obtidos através da ferramenta bioinformática “Retriever” do portal Web UniProtKB. ....	70
<b>Figura 14</b> - Figura retirada e adaptada do portal Web UniprotKB. Um dos códigos associados ao nome da proteína é inserido na caixa de texto “Query”, visto ter sido considerado que proteínas com o mesmo nome têm funções idênticas. ....	71

## **Resultados**

<b>Figura 15</b> - Quantidade de proteínas microbianas do microbioma oral identificadas em amostras de saliva de paciente com patologias. ....	75
<b>Figura 16</b> – Classificação quantitativa ontológica de acordo com o Processo Biológico (“ <i>Biological Process</i> ”) das proteínas de origem bacteriana do microbioma oral. Esta classificação foi obtida recorrendo ao programa Agbase ( <a href="http://www.agbase.msstate.edu/">http://www.agbase.msstate.edu/</a> ).....	77
<b>Figura 17</b> - Análise detalhada dos grupos “biological process” e “metabolic process” do gráfico da figura 16. ....	80
<b>Figura 18</b> - Classificação ontológica de acordo com o Processo Biológico (Biological Process) das proteínas de origem bacteriana do microbioma oral que interagem com outros microrganismos e com o hospedeiro. Esta classificação foi obtida com recurso ao programa Agbase ( <a href="http://www.agbase.msstate.edu/">http://www.agbase.msstate.edu/</a> ).....	83

<b>Figura 19</b> - Análise detalhada do grupo “pathogenesis” do gráfico da figura 17. .....	86
<b>Figura 20</b> - Mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível da proteína.....	93

## Lista de Tabelas

### Resultados e Discussão

<b>Tabela 1</b> - Bactérias produtoras das proteínas analisadas neste trabalho e associação com patologias da cavidade oral. ....	94
---	----

### Anexos

<b>Tabela I</b> - Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível da proteína. ....	113
<b>Tabela II</b> - Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene.....	118
<b>Tabela III</b> - Proteínas para as quais não há informação na bibliografia acerca dos seus mecanismos associados a patogénese. ....	132

## **Lista de abreviaturas:**

**VF-** Fator de virulência

**GO-***Gene Ontology*

**STRAP-** *Software Tool for Researching Annotations of Proteins*

**HP-** proteínas hipotéticas

**MP-** *Metabolic Process*

**BP-** *Biological Process*

**QS-** *Quorum sensing*

**Cdt-** *Cytolethal distending toxin*

**HBL-** Hemolisina BL

***S.typhimurium***-*Salmonella typhimurium*

***A. actinomycetemcomitans*** - *Aggregatibacter actinomycetemcomitans*



## **Resumo:**

A base de dados OralOme, consultável na interface OralCard, reúne toda a informação atualizada, resultante dos estudos de proteômica de tecidos orais e estudos *in vitro*. Além de proteínas do hospedeiro inclui proteínas produzidas pelo microbioma oral sendo possível conhecer as proteínas produzidas por cada género e/ou espécie. Esta ferramenta bioinformática pretende ajudar no desenvolvimento de novos métodos diagnósticos, prognósticos ou na pesquisa de agentes terapêuticos.

**Objetivo:** De acordo com os dados de proteômica de tecidos orais e estudos *in vitro* depositados no OralCard, foi realizada uma análise crítica do papel funcional desempenhado pelas proteínas de origem bacteriana dando mais ênfase àquelas envolvidas no processo biológico da patogénese.

**Materiais e métodos:** As proteínas do microbioma oral foram analisadas quanto aos processos biológicos em que participam, sendo feita a análise funcional das proteínas classificadas com o GO "*pathogenesis*".

**Resultados:** Nas 10428 proteínas bacterianas analisadas, 298 foram classificadas com o GO "*pathogenesis*". Destas, apenas 44 estão caracterizadas a nível da proteína sendo que as restantes têm evidência a nível do gene. A análise revela que destas 298 apenas 11 foram identificadas em amostras da cavidade oral sendo as restantes identificadas *in vitro*. Os mecanismos específicos de virulência com mais número de proteínas associadas são a sobrevivência do microrganismo dentro da célula do hospedeiro e a lise da célula hospedeira.

**Conclusão:** Os mecanismos específicos da patogénese para a maioria das proteínas do microbioma oral encontram-se bem relatados. No entanto, quase todas as proteínas foram descritas em estudos *in vitro* e não em amostras da cavidade oral. Seria assim pertinente em trabalhos futuros, comprovar a existência destas proteínas em amostras salivares de pacientes com as patologias orais para as quais os microrganismos responsáveis pela sua produção estão envolvidos.

**Palavras-chave:**

Microbioma oral, proteoma microbiano, fatores de virulência e patogênese

**Abstract:**

In the OralOme database, available through the OralCard interface, information resulting from proteomics studies of oral tissues and *in vitro* studies is gathered and updated. Besides the host proteins, proteins produced by the oral microbiome are included and it is possible to know which are produced by each genus and / or species. This bioinformatics tool facilitates the development of new diagnostics, prognostics and the search for new therapeutic agents.

**Objective:** Using the data deposited in the OralCard relative to the proteins produced by the oral microbiome, a critical analysis of the functional role performed by the proteins of bacterial origin was conducted. Emphasis was given to proteins involved in the pathogenesis biological process.

**Materials and methods:** The oral microbiome proteins were analyzed as to the biological processes in which they participate. A functional analysis of the proteins classified with the GO "*pathogenesis*" was performed using public databases.

**Results:** Of the 10 428 bacterial proteins analyzed, 298 were classified with the GO "*pathogenesis*". Of these, only 44 are characterized at the protein level and the remaining, have evidence at the gene level. The analysis revealed that only 11 of the 298 were identified in samples from the oral cavity and the remaining identified *in vitro*. The specific mechanisms of virulence with more number of associated proteins are the survival of the organism within the host cell and its lysis.

**Conclusion:** The specific mechanisms of the pathogenesis for most oral microbiome proteins are well described. However, almost all proteins have been described *in vitro* studies. Therefore, it would be desirable that future studies, reveal the existence of these proteins in saliva samples from patients.

**Keywords:**

Oral microbiome, microbial proteome, virulence factors and pathogenesis



## **Nota prévia**

Ao longo deste trabalho aparecem expressões em língua Inglesa, escritas em itálico, sendo exemplo disso os termos de classificação de ontologias e nomes de proteínas. Optou-se por não traduzir essas expressões, pois poderia implicar a alteração do seu significado original, e, assim, a interpretação imprecisa do conceito biológico.



# **1.INTRODUÇÃO**



A colonização microbiana da cavidade oral inicia-se no momento do parto. Algumas horas após o nascimento, a cavidade oral estéril torna-se colonizada por um número pequeno de bactérias, principalmente facultativas e aeróbias. A partir do segundo dia, bactérias anaeróbias podem ser detetadas na cavidade oral edêntula da criança. No final da infância, há um aumento no número e na diversidade bacteriana oral devido à erupção dentária, e, conseqüentemente, há mais áreas para adesão e retenção bacteriana (1).

De todos os sítios do corpo humano, a cavidade bucal é aquela que apresenta o maior índice de diversidade microbiológica a seguir ao intestino. No ser humano, a flora oral é extremamente complexa e variada. Na verdade, a variabilidade supracitada encontra-se diretamente relacionada com os diferentes *habitats* que compõem a cavidade oral: dentes, sulco gengival, língua, bochecha, lábio, palato duro e palato mole (2, 3).

A saliva presente na cavidade oral é composta por proteínas provenientes das glândulas *major* e *minor*, do plasma sanguíneo, do fluido crevicular e do microbioma oral. Estas proteínas são os elementos moleculares que mais se aproximam das características funcionais observáveis, e, por isso, torna-se claro que o esclarecimento do proteoma revele particular importância na caracterização funcional de um dado sistema biológico, sendo este o principal objetivo da proteômica (4).

Considerando que o conhecimento dos microrganismos presentes na cavidade oral é de extrema relevância para a compreensão da biologia e patologia oral, bem como o das proteínas por eles produzidos e das interações que estas estabelecem com o hospedeiro, têm sido desenvolvidos esforços no sentido do esclarecimento do microbioma oral (5).

As funções que ocorrem na cavidade oral, e em especial a função da saliva, são um reflexo das proteínas humanas e microbianas. Algumas destas funções, nomeadamente a capacidade de degradar proteínas presentes na saliva, dependem do conjunto de protéases, sejam estas de origem microbiana ou humana (4, 6).

O proteoma da cavidade oral tem sido objeto de estudo, como tal, foi desenvolvida a ferramenta *OralCard- Web information System for Oral Health*. (<http://http://bioinformatics.ua.pt/OralCard/>), pela Universidade Católica Portuguesa, em parceria com a Universidade de Aveiro. Esta ferramenta compreende a catalogação do proteoma oral humano e do microbioma oral, estabelecendo as bases para a definição do fisioma da cavidade Oral-OralOma. Por intermédio da caracterização do OralOma humano, esta ferramenta facilita a análise dos resultados da proteómica da saliva, permitindo assim o desenvolvimento de novos métodos de diagnóstico, prognóstico e de auxílio à pesquisa de novos agentes terapêuticos (4, 7). Através desta ferramenta bioinformática, torna-se possível identificar quais as proteínas produzidas por cada género e/ou espécie de microrganismo. Na base de dados Oral Card, versão de 15-12-2013, existem 11 110 proteínas de origem bacteriana.

## 1.1 O Microbioma Humano

O número de bactérias que habita o corpo humano é cerca de dez vezes superior ao número de células humanas no corpo; a maioria dessas bactérias é encontrada no trato gastrointestinal humano (8).

Foi Anton van Leeuwenhoek que conseguiu com os seus microscópios rudimentares, pela primeira vez, descrever microrganismos da água, saliva e fezes, chamando-lhes “animáculos”. Além disso, viu e desenhou, a partir das suas observações, bactérias, protozoários e outros microrganismos (9, 10).

Alguns anos depois, surgiu a introdução do termo “microbioma” por Joshua Lederberg, que traduz a “comunidade ecológica de microrganismos comensal, simbiótica e patogénica que, literalmente, partilha o nosso corpo e que funciona como um determinante da saúde e da doença” (11). Ou seja, no sentido lato, o microbioma humano refere-se a uma enorme comunidade de microrganismos que ocupam os diversos habitats do corpo humano. Sabe-se que existem microrganismos em todas as partes do corpo, e que estes desempenham papéis fundamentais na saúde humana e nas doenças (12).

A maioria dos microrganismos não causa doenças, mas tem funções importantes nos seus hospedeiros: ajuda a digestão dos alimentos, induz a produção de vitaminas (por exemplo, a vitamina K) e protege o hospedeiro contra a colonização por parte de microrganismos patogénicos (12-15). Adicionalmente, o hospedeiro e os fatores ambientais determinam os constituintes, a diversidade e a estabilidade do microbioma. Estes fatores prendem-se com: temperatura, humidade, *stress* oxidativo, nutrição, osmolaridade, recetores do hospedeiro, competição com outros microrganismos e resistência aos produtos de secreção dos mesmos, e, por fim, a resposta imunitária inata e específica contra os microrganismos (9, 16).

Cada microbioma ocupa, assim, um único nicho ecológico específico, sendo que cada um tem vários fatores ambientais que se unem para apoiar ou limitar a seleção dos seus membros, desmarcando desta forma os seus limites. Por exemplo, o microbioma de nasofaringe em crianças varia sazonalmente, e

o microbioma vaginal varia de acordo com a menstruação (9, 12). Posto isto, são encontradas diferentes comunidades microbianas em cada um dos variados ambientes da anatomia humana. Esta diversidade também existe entre indivíduos, estando provavelmente relacionada com fatores genéticos do hospedeiro, com fatores ambientais, ou com a exposição microbiana prévia (2). Alguns autores chegam mesmo a considerar o microbioma humano como o nosso segundo genoma (17). No entanto, para cada pessoa, o microbioma pode alterar ao longo do tempo, como no caso do desenvolvimento progressivo de atrofia gástrica, que, com a idade, altera o microbioma gástrico (9).

As diferentes comunidades microbianas encontradas em cada um dos variados ambientes da anatomia humana têm impacto na saúde do ser humano (13). Deste modo, uma compreensão da sua complexidade, poderá contribuir enquanto ferramenta de diagnóstico médico e capacitar à manipulação do microbioma, conduzindo a uma otimização da saúde do indivíduo. Estes objetivos deram origem ao Projeto Microbioma Humano, iniciado pelo *National Institutes of Health* em 2007 (18). No âmbito deste projeto, foram sequenciadas e analisadas amostras de 242 adultos saudáveis: 129 homens e 113 mulheres. Foram recolhidas amostras individuais, pelo menos 3 vezes durante 22 meses, de 18 partes distintas do corpo (9 de distintas zonas da cavidade oral, 5 da pele, 1 das fezes e 3 da vagina). No total, foram mais de 11 000 amostras nas quais foi usado o sequenciamento do gene 16sRNA como marcador taxonómico para identificar os membros das comunidades microbianas. Para obter amostras suficientemente diversificadas e serem, assim, representativas da população dos Estados Unidos da América, tentou-se a recolha de amostras relativamente diversificadas no que diz respeito à raça, à etnia e outras características demográficas (10, 11, 18, 19).

Depois de 5 anos de estudo, chegou-se à conclusão que a diversidade encontrada foi enorme e estima-se que no corpo humano existam mais de 10 000 espécies bacterianas diferentes. A maior diversidade microbiana encontra-se no trato intestinal e na boca (20).

Diversos estudos descrevem as variações da população microbiana que aparecem não só nos estados de doença, mas também nas alterações

microbianas temporárias durante o curso de uma doença (11). A vaginose é, por exemplo, a consequência da perturbação da flora vaginal: a flora da vagina é habitualmente composta por *Lactobacillus*, e uma diminuição desta bactéria, acompanhada pela proliferação de uma flora dominada por espécies anaeróbias, produz no meio vaginal aminas voláteis, malcheirosas e alergénicas que explicam a síndrome clínica. No caso da doença de Crohn, a redução da diversidade bacteriana anaeróbia, especialmente as *Firmicutes*, tem um papel relevante no início e na cronicidade da doença (9, 21).

### **1.1.1 O microbioma oral**

No ser humano, a flora oral é altamente complexa e diversificada, sendo que a cavidade oral de indivíduos saudáveis contém centenas de diferentes espécies bacterianas, virais e fúngicas (3). Muitas destas podem associar-se para formar biofilmes, que, por sua vez são resistentes à tensão mecânica ou tratamento antibiótico. A maioria delas também são espécies comensais, mas podem tornar-se patogénicas em resposta às mudanças do ambiente na cavidade oral, nomeadamente à qualidade da higiene oral do indivíduo (2, 3, 22).

Os dentes, a gengiva, a língua, a garganta e a mucosa bucal (bochechas) proporcionam diferentes superfícies para a colonização microbiana (23). A constante produção de saliva e o fornecimento intermitente de açúcares e aminoácidos da comida ingerida fornecem nutrientes para o crescimento microbiano (8).

Atualmente, sabe-se que a maioria dos microrganismos existentes na cavidade oral estabelece relações simbióticas baseadas em benefícios mútuos com o hospedeiro. As populações comensais podem impedir que as espécies patogénicas adiram às superfícies mucosas. Assim, as bactérias oportunistas não conseguem tornar-se patogénicas, provocando infeções e doenças até romperem a barreira de defesa dos comensais (24, 25).

Os microrganismos da cavidade oral existem, principalmente, como biofilmes sobre as superfícies dos dentes, próteses (quando presentes), gengiva, língua e restaurações. Os biofilmes dentários consistem numa comunidade microbiana organizada numa matriz extracelular complexa, constituída por produtos extracelulares microbianos e compostos salivares (8, 23). É provável que muitos microrganismos destas comunidades não possam sobreviver fora desta matriz, uma vez que o seu metabolismo é interdependente, o que justifica em grande medida a incapacidade da sua cultura em laboratório (24). Assim, sabe-se hoje que nem todas as espécies que residem no microbioma oral foram identificadas, dado que vários fatores impedem a identificação deste vasto número de espécies. Além de muitas das espécies não serem cultiváveis com as técnicas de laboratório atuais, as semelhanças entre os genomas também dificultam a identificação específica através de porções fragmentadas de ADN (23, 26). No entanto, com o desenvolvimento recente de ferramentas de meta-genómica, a complexidade do microbioma oral está a ser caracterizada (27). O projeto HMP usou esta técnica e revelou que a cavidade oral tem a maior diversidade microbiana a seguir ao intestino, na qual predominam os géneros *Streptococcus*, *Haemophilus*, *Actinomyces* e *Prevotella* (11, 19, 27).

A fixação de bactérias pioneiras é a etapa inicial na formação do biofilme dentário. A adesão inicial envolve a ligação das bactérias aos componentes da saliva, que são adsorvidos nas superfícies da cavidade oral (película adquirida) (25). Os componentes específicos adsorvidos dependem da composição da superfície, o que significa que diferentes superfícies da cavidade oral apresentam diferentes recetores salivares (3). Mais de 60% dos colonizadores primários no microbioma oral para as superfícies mucosas e dentes são *Streptococcus*. Outros géneros de bactérias que se encontram entre os primeiros colonizadores incluem: *Actinomyces*, *Veillonella* e *Neisseria* (8). As adesinas microbianas permitem a ligação à película salivar, às células do hospedeiro e à dentina radicular exposta. Uma das mais importantes adesinas de *Streptococcus* são os polipéptidos da família dos antigénios I/II, estes que se ligam à gp340, à fibronectina e ao colagénio (2, 8).

A adesão é um alvo potencial para novas terapias concebidas para inibir a colonização de bactérias. Após a ligação descrita anteriormente, os organismos pioneiros proporcionam uma nova superfície e outros sinais para a colonização, esta por parte de uma sucessão de outros organismos (colonização secundária) (23, 25). Os colonizadores secundários são na sua maioria bactérias Gram-negativas incapazes de ligar à película aderida mas capazes de se prender à superfície dos colonizadores primários num fenómeno denominado de co-agregação. Esta co-agregação pode ser homotípica, isto é, entre bactérias de uma mesma espécie, ou heterotípica, entre bactérias de género ou de espécies diferentes (3). Os *Streptococcus* e os *Actinomyces* são ambos exemplos desta ligação, visto que as fímbrias do tipo 2 dos *Actinomyces* se ligam aos receptores de polissacarídeos nos *Streptococcus*. Após algumas horas, novas espécies fixam-se sobre as bactérias já instaladas, aumentando assim a diversidade do biofilme. Essas espécies bacterianas são, por vezes, denominadas de “colonizadores tardios” (3, 23).

Deste modo, o biofilme dentário desenvolve-se em espessura, na consequência da divisão celular e da co-aderência de novos microrganismos, constituindo, assim, um biofilme complexo e maduro. Com o aumento das camadas de bactérias, aparecem novas condições ambientais, a quantidade de oxigénio diminui e as bactérias anaeróbias desenvolvem-se (2, 3, 8).

Quando há homeostase microbiana, estado de equilíbrio e de estabilidade que depende das interações bactérias-bactérias por um lado e das interações bactérias-hospedeiro por outro, esta flora comensal é uma flora de barreira que se opõe à colonização por eventuais patógenos exógenos ou ao desenvolvimento de microrganismos endógenos patógenos oportunistas. A rutura desse equilíbrio dá origem a patologias como cáries dentárias ou doenças periodontais (23, 24, 28). Este biofilme *clímax* atinge um equilíbrio dinâmico, onde os diversos microrganismos estão em constante adaptação às alterações ambientais (29). O biofilme dentário é um biofilme dinâmico, parcialmente destruído aquando da escovagem dos dentes, e que é reconstruído rapidamente sobre a película aderida, após 4 a 6 horas (3).

É de realçar que dentro do microbioma oral existe uma grande diversidade de microbiomas que dependem das alterações das condições ambientais, sendo estas: a temperatura, a hidrometria, o pH, o potencial redox e a disponibilidade em nutrientes. Todas estas condições modificam a composição de espécies presentes nos biofilmes (2, 23, 30). Uma vez que as condições ecológicas diferem entre as diferentes superfícies da cavidade oral, a diversidade nos microrganismos está diretamente relacionada com os diferentes *habitats* desta. De entre estes diferentes habitats destacam-se: os dentes, o sulco gengival, a língua, a bochecha, o lábio, o palato duro e o palato mole (3, 23). O dente pode ser considerado um *habitat*; no entanto, seria mais correto considerá-lo como um conjunto de várias superfícies, cada uma delas favorecendo uma ecologia que lhe é própria (3, 8). Os sulcos das faces oclusais protegem a microflora do ambiente, mas permanecem sensíveis às influências nutricionais, encontrando-se, assim, maioritariamente, Gram positivos, aeróbios ou anaeróbios aerotolerantes, e, em particular, os *Streptococcus*. As faces interproximais, dado o seu baixo potencial redox, abrigam diversas comunidades bacterianas que contêm grandes proporções de bactérias anaeróbias estritas (3). Cada *habitat* da cavidade oral permite, pois, a certas espécies, dominar um local e ser menos prevalentes noutra (3, 23, 31).

#### 1.1.1.1 Quorum sensing

Dentro do biofilme, os produtos metabólicos e as moléculas sinalizadoras produzidas por uma espécie podem influenciar as bactérias vizinhas. Por vezes, o resultado é a concorrência interespecies; no entanto, muitas interações entre bactérias orais são mutuamente benéficas (32, 33)

Um sistema de sinalização usado pelas bactérias é o *quorum sensing*, (QS) e refere-se a moléculas secretadas pelas próprias bactérias. Vários sistemas QS foram identificados em bactérias e em fungos, sendo os mais prevalentes na cavidade oral os auto-indutores peptídicos. O sistema QS controla a expressão de genes em resposta a mudanças na densidade celular e é considerado um mecanismo de regulação na adaptação ecológica e na

patogenicidade (32). Há indícios de que os mecanismos de *quorum sensing* controlam a produção de fatores de virulência nalgumas espécies de bactérias (34). Por exemplo, a *Porphyromona gingivalis* tem sido desde há muito tempo considerada um dos principais patógenos periodontais, tendo um papel importante na destruição do tecido ósseo. Esta, está ausente num estado de saúde; no entanto, quando em doença, atinge uma parte relevante da população total e tem a capacidade de produzir um grande número de fatores de virulência. Foi provado que no caso de *Porphyromona gingivalis* a produção de AI-2 pode desempenhar um papel de alguma importância no controlo da expressão de seus fatores de virulência (35).

Para o desenvolvimento de uma infeção bacteriana, é necessário que um agente patogénico invasor atinja a densidade necessária para superar as defesas do hospedeiro e, assim, estabelecer a infeção. O facto de controlar a expressão de fatores de virulência, de acordo com a densidade celular pelo *quorum sensing*, confere uma vantagem significativa para a sobrevivência do patogénico, de tal forma que o hospedeiro está sobrecarregado antes que uma resposta de defesa seja totalmente iniciada. Inibidores de tais sistemas de comunicação bacteriana podem, assim, proporcionar uma nova ferramenta terapêutica contra infeções bacterianas (32, 34)

Além disso, o sistema QS parece estar envolvido em todas as fases da formação do biofilme (32). Este sistema regula a densidade populacional e a atividade metabólica dentro do biofilme maduro para atender às exigências nutricionais e aos recursos disponíveis. Análises comparativas dos proteomas de bactérias patogénicas presentes no biofilme e modos de crescimento plantónicos foram reportados. Estas investigações revelaram mudanças na abundância de numerosas proteínas, estas que podem estar relacionadas com a adaptação ao modo de crescimento no biofilme e com a variação da expressão de genes que codificam para a virulência (3, 36).

### 1.1.1.2 Alterações do microbioma oral

Para determinar se um microrganismo específico é causador da doença, são usados os postulados de Koch (13). Estes dizem que o microrganismo causador da doença deve ser encontrado em todos os casos da doença examinada, preparado e mantido em cultura pura, sendo capaz de produzir a infecção original em animais modelo e recuperado a partir de um hospedeiro infetado. Estes postulados não são válidos em situações de doenças orais, visto que muitas destas doenças não são causadas por um único microrganismo, mas são polimicrobianas, ou seja, dependem da presença de várias espécies bacterianas que interagem umas com as outras para se multiplicar ou para expressarem fatores de virulência (23).

Uma hipótese proposta para o desenvolvimento do biofilme, apelidada de “hipótese da placa ecológica”, fornece um argumento para explicar as doenças mediadas pela placa bacteriana: estas podem ser a consequência de desequilíbrios na microflora residente, esta que por sua vez resulta de um aumento dos “patógenos orais” dentro do biofilme. Ou seja, estes microrganismos patogénicos podem estar presentes em situações de saúde, mas em menor quantidade (23). Assim, a homeostase microbiana pode ser quebrada devido à alteração de um parâmetro que é fundamental para a manutenção da estabilidade ecológica. Estes parâmetros importantes que regulam a homeostase na boca incluem a integridade das defesas do hospedeiro (tal como o fluxo de saliva) e a composição da dieta. Os indivíduos que consomem regularmente componentes da dieta com alto teor de açúcar fermentável têm maiores proporções de bactérias acidogénicas e de ácido tolerante, especialmente *Streptococcus mutans* e *Lactobacillus*, sendo que estes são os microrganismos associados à cárie dentária. Estes microrganismos são pouco competitivos a um pH neutro e estão presentes em pequena proporção no biofilme saudável; todavia, são mais adaptados a um ambiente com pH baixo, sobreexpressando os genes que protegem contra o stresse ácido (23-25, 30, 36).

## 1.2 Os fatores de virulência bacterianos

### 1.2.1 Definição de um fator de virulência

Os fatores de virulência (VF) podem ser definidos como moléculas produzidas pelos patógenos que aumentam a capacidade dos agentes patogênicos de causar doença. Os VFs estão envolvidos na invasão, na transmissão da doença, na toxicidade, na colonização e no combate das defesas do hospedeiro. Os VFs incluem: as proteínas que permitem aos micróbios ligarem-se ao hospedeiro (adesinas), os efetores que suprimem as defesas do hospedeiro, as enzimas que degradam os tecidos do hospedeiro (invasinas), as proteínas do sistema de secreção (estão atualmente descritos 5 sistemas compostos por proteínas diferentes em que alguns são compartilhados por vários gêneros), as cápsulas de polissacarídeos que protegem patógenos das defesas do hospedeiro e moléculas que atuam na captação de ferro (sideróforos). As moléculas de *housekeeping* que estão implicadas na sobrevivência e crescimento do microrganismo não são considerados por alguns autores como sendo VF (37, 38).

A forma como se classificam ou organizam os fatores de virulência é variável com o autor ou com a fonte considerada. A organização pode ser feita tendo em conta: as propriedades das moléculas envolvidas (39), os processos em questão (40), ou a localização das moléculas na célula bacteriana (41). Independentemente da forma de classificação, o mais importante é poder compreender os mecanismos de virulência - quer nas suas especificidades, quer naquilo que têm de comum - por forma a desenhar novas medidas terapêuticas de combate à infeção.

### 1.2.2 Principais mecanismos de virulência

Compreender e descobrir fatores de virulência de bactérias patogénicas é uma chave para a compreensão da patogénese, para a identificação de alvos para novos medicamentos e para o desenho de novas vacinas (42). A associação do aumento da disponibilidade do sequenciamento dos genomas completos de bactérias patogénicas e da bioinformática levaram a avanços significativos em direção a esse objetivo, existindo já bases de dados dedicadas a este assunto, das quais destacamos o *Virulence Factors of Pathogenic Bacteria* (<http://www.mgc.ac.cn/VFs/>) (40, 41, 43).

Nesta revisão, é proposta uma classificação que organiza o conhecimento sobre os fatores de virulência bacterianos em 4 grandes grupos que pretendem sumariar as interações hospedeiro-bactéria. Esta classificação, embora tenha por base grupos idênticos aos estabelecidos na base de dados *Virulence Factor Database for Pathogenic Bacteria* (40), foi reorganizada de uma forma considerada mais intuitiva que separa processos e tipo de molécula. Assim distinguimos 4 grandes grupos de proteínas associadas a virulência que passamos a descrever.

- 1) Os fatores de virulência que são específicos e desenhados para “atacar” o hospedeiro quer a nível tecidual, quer a nível celular. A maior parte destes mecanismos destina-se a permitir à célula bacteriana a invasão das células do hospedeiro. Por vezes as bactérias invadem o tecido e o espaço intercelular, mas esses processos são sobretudo mediados por moléculas descritas no grupo 3 de fatores de virulência, pois muitas das enzimas produzidas pelas bactérias para a sua nutrição estão incluídas nesta categoria. Neste primeiro grupo de fatores de virulência, numa sequência temporal de contacto com o hospedeiro, incluímos os processos a seguir listados. É importante realçar que nos vários casos estudados há microrganismos que não realizam todos os processos listados mas apenas alguns deles.
  - a. Adesão (à matriz extracelular ou à célula);
  - b. Sistemas de secreção;

- c. Invasão celular do hospedeiro;
  - d. Mobilidade inter e intracelular baseada em reorganização do citoesqueleto;
  - e. Lise da célula do hospedeiro;
  - f. Apoptose da célula do hospedeiro;
- 2) Os fatores de virulência que pressupõem uma defesa do microrganismo dos ataques do hospedeiro. Algumas das moléculas aqui apresentadas poderiam ser incluídas no ponto seguinte pois muitas delas fazem parte dos atributos metabólicos que a bactéria possui para se nutrir.
- a. Defesas antifagocíticas;
  - b. Defesas antiprotelíticas;
  - c. Metabolismo celular;
  - d. Resistência a substâncias do soro;
  - e. Proteases para Igs do hospedeiro;
  - f. Proteases para o sistema de complemento do hospedeiro;
  - g. Sobrevivência intracelular;
  - h. Proteínas de stress;
- 3) Os fatores de virulência não específicos, ou seja, que derivam de processos metabólicos/estruturas que a bactéria tem e que expressa independentemente de estabelecer relações com o hospedeiro:
- a. Sistemas de captação de ferro;
  - b. Sistemas de captação de magnésio;
  - c. Exoenzimas;
  - d. Formação de biofilme;
- 4) Este quarto grupo inclui os fatores de regulação dos fatores de virulência que são considerados por muitos autores como sendo eles mesmos fatores de virulência.

Serão de seguida descritos os mecanismos de virulência e as proteínas bacterianas responsáveis pelos mesmos para os quais há mais evidência experimental.

#### **1.2.2.1 Adesão**

Um passo fundamental na interação patógeno-hospedeiro é a adesão do patógeno às superfícies do hospedeiro, uma vez que numerosas forças mecânicas produzidas pelo hospedeiro, como por exemplo a secreção salivar, dificultam esta ligação. Ao ficar ligado ou aderido a uma superfície específica do hospedeiro, o agente patogénico é capaz de iniciar os processos bioquímicos específicos que irão dar origem à doença, nomeadamente a proliferação, a secreção de toxina, a invasão de células do hospedeiro e a ativação de cascatas de sinalização na célula colonizada (43, 44).

Os fatores de adesão microbiana têm o nome de adesinas e estas podem ser proteínas ou hidratos de carbono. As adesinas proteicas podem estar ou não localizadas nas fimbrias (fimbriais ou afimbriais). As fimbrias são apêndices que se projetam a partir da superfície bacteriana e desempenham um papel fundamental na colonização de superfícies. Há no entanto adesinas que não estão nas fimbrias mas sim à superfície da célula, mediando um contato mais íntimo com a célula hospedeira. Já as adesinas compostas por hidratos de carbono fazem parte da membrana celular bacteriana, parede celular ou cápsula (39, 44).

#### **1.2.2.2 Invasão**

A adesão da bactéria patogénica à célula do hospedeiro ativa sinais na célula do hospedeiro que facilitam a entrada da bactéria. Assim sendo, a invasão bacteriana é um evento ativo dependente das funções desempenhadas pelo hospedeiro (43).

Muitas bactérias patogênicas invadem células fagocíticas, bem como células não fagocíticas (como por exemplo células epiteliais), durante o estabelecimento de infecções no hospedeiro. A internalização de bactérias por fagócitos faz parte da resposta imune inata do hospedeiro. Em contrapartida, de modo a contrariar a ação imunitária realizada pelos fagócitos, muitas das bactérias patogênicas intracelulares evoluíram de modo a sobreviver, e até proliferar, no interior dos mesmos (macrófagos, neutrófilos e células dendríticas) (45, 46).

As bactérias patogênicas intracelulares, para invadir as células não-fagocitárias, tais como as células epiteliais, usam os mecanismos “*zipper*” ou “*trigger*”. As bactérias que usam o mecanismo “*zipper*” possuem proteínas na sua superfície (adesinas) que se ligam a receptores existentes na membrana das células do hospedeiro, induzindo cascatas de sinalização que reorganizam a actina do citoesqueleto de maneira a internalizar a bactéria (47).

Já as bactérias patogênicas que se servem do mecanismo “*trigger*” usam os sistemas de secreção III ou IV para injetarem proteínas através da membrana plasmática do hospedeiro. Estas proteínas interagem diretamente com os componentes celulares, estes que, por sua vez, regulam a dinâmica da actina alterando o citoesqueleto da célula infectada (47, 48)

### **1.2.2.3 Mecanismos de sobrevivência intracelular**

Após a invasão das células do hospedeiro, a capacidade de sobreviver dentro destas é essencial para as várias bactérias patogênicas (49). Tentam, portanto, contornar a imunidade inata do hospedeiro, tal como o sistema do complemento (50).

As bactérias patogênicas intracelulares (que já estão, portanto, dentro das células do hospedeiro) encontram-se dentro de um compartimento vacuolar, ligado à membrana, designado fagossoma. Em circunstâncias normais, o vacúolo acidifica progressivamente e, conseqüentemente, termina na formação de um fagolisossoma (fusão de fagossoma com lisossoma).

Devido à ação do baixo valor de pH e de enzimas digestivas, ocorre, pois, a morte da bactéria (45, 47).

Como método de sobrevivência dentro das células do hospedeiro, algumas bactérias têm a capacidade de escapar aos fagossomas, continuando o seu ciclo de vida dentro do citosol. Possuem, ainda, a capacidade de impedir a formação do fagolisossoma, evitando assim a digestão lisossomal referida acima (45, 49).

Assim, de modo a escapar à degradação fagolisossomal, algumas bactérias produzem enzimas, ativadas pelo ambiente ácido do vacúolo, que promovem a lise do fagossoma, permitindo assim a colonização do citosol. O mecanismo de lise do vacúolo pela *Listeria* é um dos mais documentados, sendo mediado pelas enzimas *listeriolysin O* (LLO) , *phosphatidylinositol-specific phospholipase C* (PI-PLC) e por uma larga gama de *phosphatidylcholine-specific phospholipase C* (PC-PLC). A enzima LLO, ativada em condições ácidas, insere-se na membrana do fagossoma, formando nela poros que levam à sua rutura. As enzimas PI-PLC e LLO são necessárias para ir do fagossoma até ao citosol, enquanto que a PC-PLC contribui para a rutura do vacúolo de dupla membrana durante a disseminação que ocorre de célula a célula (49). No que diz respeito a outras bactérias patogénicas como a *Salmonella* e a *Mycobacterium* as mesmas param a maturação do fagossoma em fases específicas da via fagolisossómica (45, 50).

#### **1.2.2.4 As toxinas**

A presença de bactérias é importante no processo infeccioso, mas não é suficiente para explicar os sintomas da patologia. Os sintomas são essencialmente devidos a compostos tóxicos produzidos pela bactéria patogénica e libertos no sítio da infeção, de modo a destruir ou danificar a célula hospedeira (30, 51). As toxinas podem ser de dois tipos: endotoxinas (moléculas não proteicas) ou exotoxinas (moléculas proteicas) (43, 51).

As endotoxinas devem o nome ao facto de serem libertadas aquando da lise do microrganismo (30). Por exemplo, o lipídeo A dos lipopolissacarídeos das bactérias Gram-negativas é considerado como sendo uma poderosa endotoxina (51).

Quanto às exotoxinas bacterianas, estas são geralmente enzimas que são secretadas para o meio extracelular, ou diretamente injetadas para este, por via do sistema de secreção tipo III, entre outros. Estas podem ser classificadas, com base na sua composição em aminoácidos e função, da seguinte forma: (1) toxinas AB (constituídas por duas subunidades - A e B); (2) toxinas proteolíticas; (3) as toxinas formadoras de poros, e (4) outras toxinas que incluem, por exemplo, a serina protease ou a hemaglutinina (39, 43).

Um exemplo de toxinas proteolíticas são as proteases bacterianas que são secretadas para o meio extracelular e/ou expressas nas superfícies das células bacterianas (41, 43). Tais proteases podem ser classificadas em quatro grupos, que são os seguintes (tendo em conta o resíduo catalítico essencial no seu sítio ativo): proteases de serina, proteases de cisteína (também chamados tioproteases), proteases de aspartato, e as metaloproteases. Estas proteases têm o potencial de destruir a integridade do tecido, efetuando a lise das proteínas do hospedeiro, e podem ativar as formas latentes (zimogénios) das proteases do hospedeiro (por exemplo, metaloproteinases da matriz) ou inativar os inibidores naturais das proteases do hospedeiro. Possuem, ainda, a capacidade de inativar proteínas do sistema imune deste. Algumas das proteases mais estudadas foram as proteases da cisteína, nomeadamente as Arg-gingipain de *Porphyromonas gingivalis*, que comprovaram ser importantes fatores de virulência em doenças periodontais (52-54).

Há mais toxinas bacterianas que, aquando da sua inserção na membrana da célula do hospedeiro, provocam a formação de poros que conduzem à lise da célula hospedeira. Uma grande família das referidas toxinas são as RTX (“*repeats in toxins*”, nome devido às repetições de certos péptidos na estrutura da proteína) que, por sua vez, são encontradas em muitos organismos patogénicos Gram-negativos (43).

O nome das toxinas AB deve-se às estruturas que as constituem, uma vez que são compostas por: uma subunidade B, responsável pela ligação ao recetor da célula do hospedeiro, e uma subunidade A, possuidora da atividade enzimática e que atua na célula do hospedeiro (a atividade enzimática da subunidade A varia desde a atividade proteolítica até a atividade de ribosilação do ADP) (39, 43).

#### **1.2.2.5 Regulação dos fatores de virulência**

Em resposta às mudanças das condições ambientais durante a infecção decorrente no hospedeiro (incluindo mudanças de temperatura, de pH, da pressão osmótica e da disponibilidade em oxigénio ou em nutriente), a bactéria regula a expressão dos seus fatores de virulência, permitindo uma rápida adaptação desta ao novo ambiente em questão (43).

Os principais mecanismos usados pelo patógeno para controlar a expressão dos genes associados a virulência são: os fatores sigma alternativos e os sistemas reguladores de dois componentes (39).

Os fatores sigma são subunidades da proteína ARN polimerase bacteriana e controlam o início da transcrição na sequência promotora. Vários patógenos usam fatores sigma alternativos para regular a transcrição de genes associados à virulência. Por exemplo, foi mostrado que o fator sigma RpoS regula a expressão de genes numa série de agentes patogénicos bacterianos, incluindo *Salmonella typhimurium*, *Escherichia coli* e *Pseudomonas aeruginosa*, em resposta à fase estacionária (fase na qual há falta de nutrientes e na qual os produtos tóxicos se estão a tornar mais abundantes) e à privação de nutrientes, bem como resposta ao estresse osmótico e oxidativo (39, 43, 55).

Os sistemas reguladores de dois componentes são constituídos por duas proteínas envolvidas na expressão da virulência: uma proteína sensor, que é incorporada na membrana bacteriana e que é “sensível” às diferentes condições fisiológicas da célula bacteriana, e a proteína reguladora de resposta, esta que se liga à região promotora de um gene para ativar ou

reprimir a transcrição deste (39, 55). Estes sistemas foram identificados em numerosas bactérias e estão envolvidos, entre outros casos, na regulação de ferro, na de fosfato, na de nitrogénio e na de carbono (43).

O *quorum sensing* é também considerado por alguns autores como um fator de virulência (42, 56) -tópico desenvolvido na secção anterior.

## **1.3 Ferramentas bioinformáticas**

### **1.3.1 Aplicação da bioinformática**

A bioinformática pode ser definida como a união da biologia e da informática: consiste, portanto, na aplicação da tecnologia que usa computadores para o armazenamento, a recuperação, a manipulação e a distribuição de informações relacionadas com macromoléculas biológicas, como o ADN, ARN e as proteínas. Devido à enorme quantidade de dados, a maioria das tarefas de análise de dados genómicos torna o uso de computadores e algoritmos indispensável. A bioinformática incide em dois campos: em primeiro lugar, foca-se no desenvolvimento de ferramentas computacionais e de bancos de dados, e, em segundo, incide sobre a aplicação destas ferramentas e bases de dados na criação de conhecimento biológico para uma melhor compreensão dos seres vivos (57).

Para um melhor conhecimento do OralOma humano é, assim, necessário o uso de ferramentas bioinformáticas para a análise de proteínas, quer de origem humana, quer de origem microbiana. Há numerosas ferramentas bioinformáticas que visam o armazenamento, a catalogação e a consulta de dados sobre proteínas (como por exemplo o Uniprot e o OralCard), e outras que colmatam a necessidade de interpretar esses dados de forma sistemática e uniforme (é o caso das ontologias génicas ou do *Gene Ontologies*, do Strap ou do AgBase). Apesar de haver várias ferramentas e

várias bases de dados com abrangências comuns, na introdução que a seguir se apresenta serão descritas, de forma breve, apenas as que foram usadas no âmbito deste trabalho.

### 1.3.2 UniprotKB

O Uniprot (<http://www.uniprot.org/>) é um repositório de sequências de proteínas com uma cobertura abrangente e com uma abordagem sistemática para a anotação de proteínas. Integra, interpreta e padroniza os dados de literatura e de outros numerosos recursos para alcançar a mais abrangente base de dados sobre as proteínas. Foi criado pelo Uniprot Consortium, resultando da colaboração do *European Bioinformatics Institutes (EBI)*, do *Swiss Institute of Bioinformatics (SIB)* e do *Protein Information Resource (PIR)*. É composto por quatro componentes otimizados para diferentes usos: *Uniprot Knowledgebase* (UniprotKB), *Uniprot Archive* (Uniparc), *Uniprot Reference Cluster* (Uniref) e *Uniprot Metagenomic and Environmental Sequences* (UniMES) (58).

A ferramenta bioinformática UniprotKB (*Uniprot Knowledgebase*) foi utilizada no âmbito deste trabalho e é composta por duas secções: a UniprotKB/Swiss-Prot, na qual estão presentes os dados que foram manualmente anotados e revistos, extraindo a informação biológica da literatura, de modo a fornecer todas as informações relevantes sobre cada proteína, e ainda a secção UniProtKB / TrEMBL, um banco de dados de sequências de proteínas anotadas por computador. Juntas, estas duas secções cobrem todas as proteínas caracterizadas ou inferidas a partir das sequências de nucleótidos de forma atualizada (59, 60).

### 1.3.3 OralOma e OralCard

O proteoma da cavidade oral tem sido objeto de estudo e o desenvolvimento de técnicas relativas à proteômica permitiu um aumento exponencial na identificação de proteínas na saliva (5).

A compilação de proteínas orais levou à necessidade de criar uma ferramenta que permitisse armazenar, integrar e visualizar os dados de forma interpretativa. Visando este objetivo, foi desenvolvida a ferramenta *OralCard-Web information System for Oral Health*. (<http://bioinformatics.ua.pt/OralCard/>) pela Universidade Católica Portuguesa, em parceria com a Universidade de Aveiro. O Oral Card apresenta a catalogação do proteoma oral humano e do microbioma oral, estabelecendo as bases para a definição do fisioma da cavidade Oral- OralOma. Este, está organizado segundo três pontos de vista diferentes: pelo nome da proteína ou respetivo Código UniProt, pelo nome da doença (código OMIM ou termo MeSH) e pelo organismo (7).

Os dados, presentes no OralCard, têm como objetivo integrar os diferentes resultados gerados pelas ciências “ómicas”, dedicados à cavidade oral. Desta forma, permite a caracterização do OralOma humano, facilitando a análise dos resultados de proteômica da saliva, permitindo, assim, o desenvolvimento de novos métodos de diagnóstico, prognóstico e de auxílio à pesquisa de novos agentes terapêuticos (4).

### 1.3.4 Gene Ontology

O crescimento exponencial do volume de informações biológicas levou à necessidade de anotação da informação molecular sobre os genes e sobre os seus produtos (entende-se por produtos de genes, proteínas ou RNA funcional). O projeto *Gene Ontology (GO)*, criado em 1998, teve, assim, como objetivo proporcionar uma linguagem universal e adequada para a anotação de

características moleculares nos organismos com o intuito de descrever os produtos dos genes de qualquer organismo (61). O GO tem dois objetivos: o primeiro, é o de facilitar a comunicação entre investigadores, sendo o segundo melhorar a interoperabilidade entre ferramentas bioinformáticas. O Gene Ontology está a ser constantemente ampliado e revisto (62).

As ontologias podem ser definidas como sendo vocabulários estruturados que representam uma rede na qual cada termo pode ser *child* ou *parent*. Assim, a relação dos produtos de genes com o processo biológico, com a função molecular e com o componente celular, será de um para muitos, dado que estes podem estar envolvidos em vários processos biológicos, uma vez que contêm domínios que realizam várias funções moleculares, ocorrendo em um ou em mais locais celulares (63). As relações dos *child terms* com os *parent terms* podem ser do tipo "is a" ou do tipo "part of". O relacionamento "is a" refere-se quando um *child term* é um exemplo do pai. Relativamente a "part of", este é usado quando um *child term* é componente do *parent term*. É de salientar, ainda, que existem *child terms* que podem ter relacionamentos diferentes com *parent terms*, também estes diferentes (Figura 1) (62).

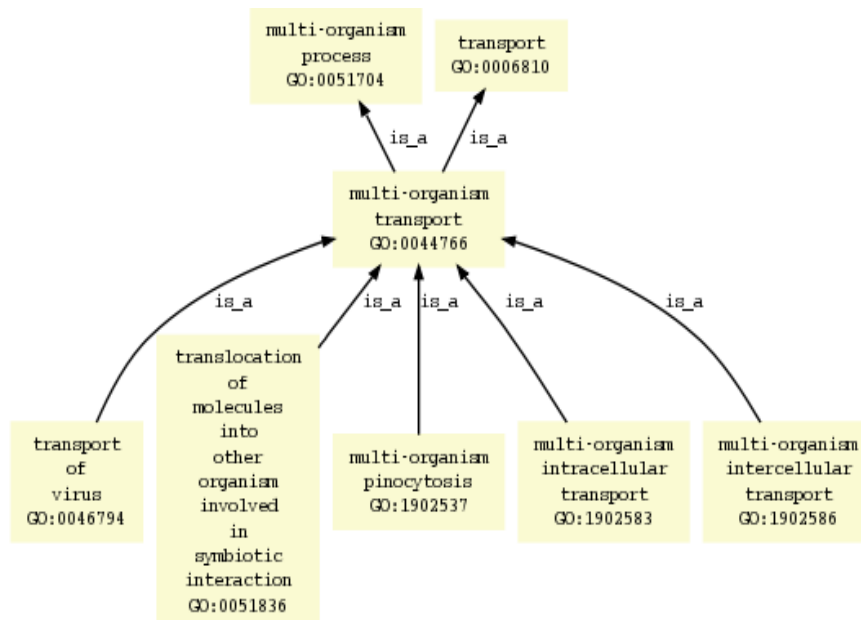
O Consórcio GO desenvolveu, assim, três ontologias: função molecular (*Molecular Function*), processo biológico (*Biological Process*) e componente celular (*Cellular Component*) que são usadas para descrever os produtos de genes. Estes três termos foram escolhidos para serem ontologias porque representam conjuntos de informações que são comuns a todas as formas de vida, tornando-os assim fundamentais para a anotação de informações sobre genes e produtos de genes.

A ontologia *Molecular Function* (função molecular) descreve o que um produto do gene faz ao nível bioquímico, descrevendo somente o que este faz sem especificar onde e quando a função do gene em causa ocorre. Alguns exemplos de *parent terms* que descrevem a função molecular são: "*protein transporter activity*" ou "*cytoskeletal protein binding*". Outros exemplos de um dos seus *child terms* associados são "*bacteriocin transmembrane transporter*" e "*actin binding*".

O “*Biological Process*” (processo biológico) descreve os objetivos biológicos realizados por um ou mais conjuntos ordenados de funções moleculares. Os exemplos de termos antecedentes associados ao processo biológico são os seguintes: “*transport*” e “*response to stress*”, sendo os seus termos de nível mais inferior, respetivamente, “*metal ion transport*” e “*stress activated protein kinase signaling cascade*”.

Por fim, a ontologia “*Cellular Component*” (componente celular) refere-se ao lugar na célula onde um produto do gene é encontrado (por exemplo, “*membrane*”, que tem o “*child term*”, “*organelle membrane*”, entre outros) (61, 64, 65).

Acrescenta-se, também, que cada termo GO tem uma definição, sendo que para a maioria delas foi usado o *Oxford Dictionary of Molecular Biology* (1997) para as definir (61, 64).



**Figura 1** - As ontologias são uma rede complexa. A figura mostra que as ontologias podem ser visualizadas como uma rede em que cada termo tem vários *child terms*, podendo estar associados a vários *parent terms*. É o caso do termo ontológico “*multi-organism transport*” que tem 5 *child terms* e é ele mesmo descendente dos processos biológicos “*multi-organism*

*process*” e *“transport”*. Imagem adaptada de <http://www.ebi.ac.uk/ontology-lookup/browse.do?ontName=GO>.

Desde 2004, o consórcio *Plant-Microbe Associated Gene Ontology* (PAMGO) (<http://pamgo.vbi.vt.edu/>) expandiu o GO, incluindo os termos que descrevem os processos envolvidos na interação entre micróbios e os seus respectivos hospedeiros. Assim, foram adicionados mais de 700 termos de processos biológicos. A implementação de novos termos começou com a criação do termo GO *“symbiosis”*, introduzindo a sua definição como “um contínuo que abrange interações que vão desde o mutualismo, passando pelo comensalismo, até ao parasitismo”. Foi, assim, introduzido o novo termo ontológico *“symbiosis, encompassing mutualismo through parasitism”* (66).

Todos os termos GO designados para os processos envolvidos na interação micróbio-hospedeiro (independentemente da associação ser mutualista, comensal ou parasitária, ou se os hospedeiros são plantas ou animais) são colocados em *“GO: 0044403 symbiosis, encompassing mutualismo through parasitism”*, estando este ligado ao termo do nível superior *“GO: 0008150 biological process”* através de termos intermediários como *“GO: 0044419 interespecies interaction between organism”* e *“GO: 0051704 multiorganism process”*. Nesta estrutura, o termo *“pathogenesis”* é também um *child term* de *“symbiosis, encompassing mutualismo through parasitism”*, uma vez que é um tipo de interação simbiótica. Assim sendo, um produto de gene envolvido na patogénese é anotado, quer no processo específico (como por exemplo *“adhesion to host”*), quer para o termo *“pathogenesis”*. Porém, se está apenas envolvido na *“adhesion to host”* e não está associado à doença, então só é anotado para *“adhesion to host”* (Figura 2) (66, 67).

**A**

- GO:0008150 : biological\_process [340066 gene products]
- GO:0051704 : multi-organism process [10891 gene products]
- GO:0044419 : interspecies interaction between organisms [2846 gene products]
- GO:0044403 : symbiosis, encompassing mutualism through parasitism [2355 gene products]**
  - GO:0051816 : acquisition of nutrients from other organism during symbiotic interaction [8 gene products]
  - GO:0051825 : adhesion to other organism involved in symbiotic interaction [166 gene products]
  - GO:0075071 : autophagy involved in symbiotic interaction [0 gene products]
  - GO:0085031 : commensalism [0 gene products]
  - GO:0044111 : development involved in symbiotic interaction [15 gene products]
  - GO:0051821 : dissemination or transmission of organism from other organism involved in symbiotic interaction [3 gene products]
  - GO:0044110 : growth involved in symbiotic interaction [53 gene products]
  - GO:0051701 : interaction with host [1813 gene products]
  - GO:0052047 : interaction with other organism via secreted substance involved in symbiotic interaction [178 gene products]
  - GO:0051702 : interaction with symbiont [59 gene products]
  - GO:0051708 : intracellular protein transport in other organism involved in symbiotic interaction [11 gene products]
  - GO:0051817 : modification of morphology or physiology of other organism involved in symbiotic interaction [517 gene products]
  - GO:0052192 : movement in environment of other organism involved in symbiotic interaction [846 gene products]
  - GO:0044399 : multi-species biofilm formation [6 gene products]
  - GO:0085030 : mutualism [0 gene products]
  - GO:0009877 : nodulation [15 gene products]
  - GO:0051824 : recognition of other organism involved in symbiotic interaction [5 gene products]
  - GO:0043903 : regulation of symbiosis, encompassing mutualism through parasitism [71 gene products]
  - GO:0052173 : response to defenses of other organism involved in symbiotic interaction [680 gene products]
  - GO:0051836 : translocation of molecules into other organism involved in symbiotic interaction [0 gene products]

**Figura 2** - Representação em árvore dos termos antecedentes e descendentes do “GO:0044403 symbiosis encompassing mutualismo through parasitism”. Imagem adaptada de <http://amigo.geneontology.org>.

As ferramentas bioinformáticas Agbase (<http://www.agbase.msstate.edu/>) e STRAP (*Software Tool for Researching Annotations of Proteins*) que foram usadas na realização deste trabalho fornecem a anotação ontológica GO de proteínas de origem bacteriana (68, 69).



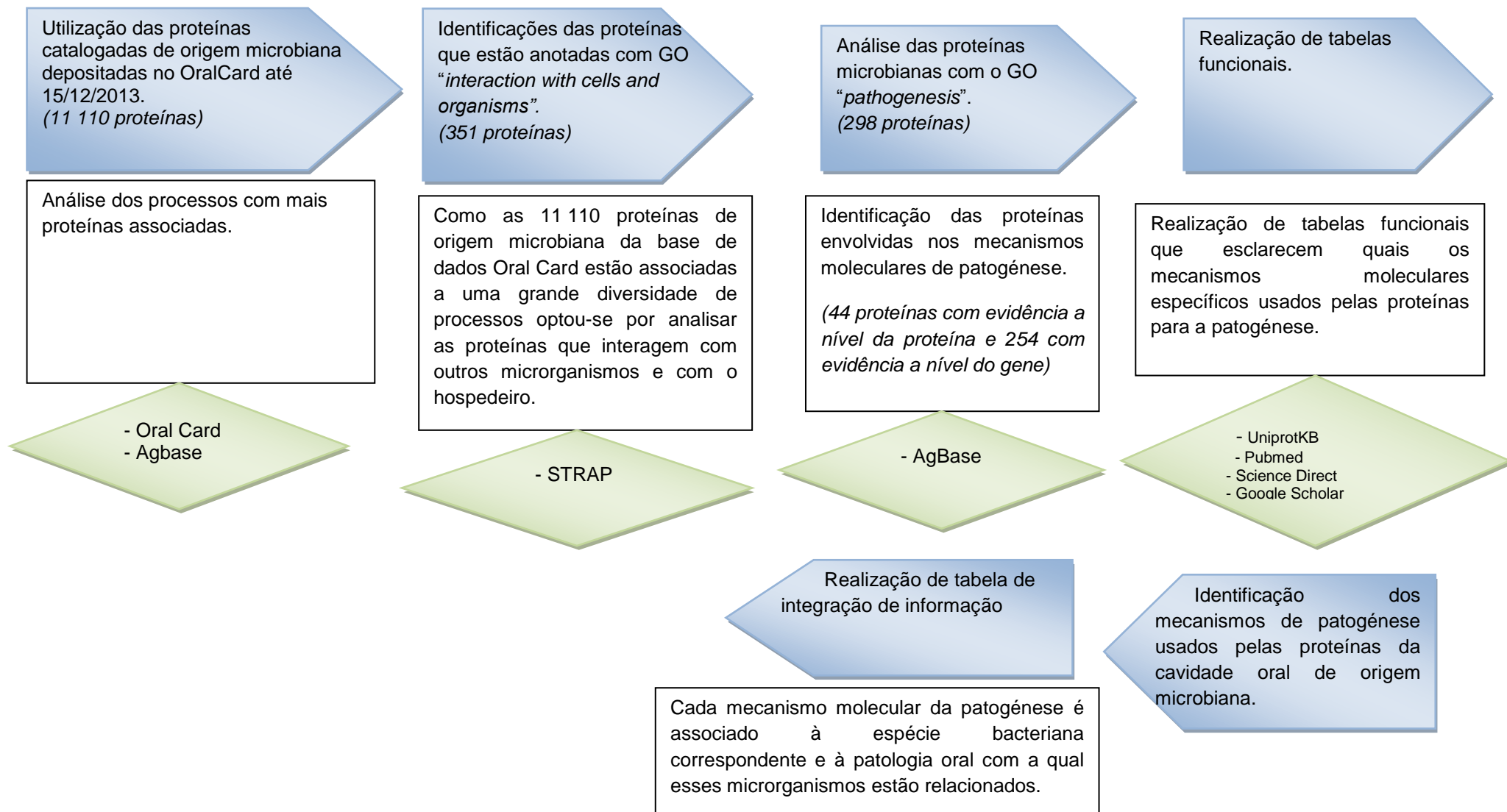
## **2. OBJETIVOS**



De acordo com os dados de proteômica de tecidos orais e estudos *in vitro* depositados no OralCard, relativamente às proteínas produzidas pelo microbioma oral, este trabalho tem como objetivo identificar quais as proteínas do microbioma oral envolvidas no processo biológico de patogénese, e mais precisamente quais os mecanismos específicos associados a cada proteína.

A figura 3 apresenta a estratégia que foi usada na realização deste trabalho de investigação.





**Figura 3-** Fluxograma que ilustra a estratégia de trabalho usada para a realização deste trabalho. As setas indicam ações, os quadrados objetivos e os losângulos as ferramentas bioinformáticas e as bases de dados usadas



### **3.MATERIAIS E METÓDOS**



### 3.1 Utilização das proteínas catalogadas de origem microbiana depositadas no OralCard

Na base de dados do OralCard, o OralOme foi consultado usando o Microsoft Excel 2007<sup>®</sup>, versão de 15-12-2013, o qual permitiu que, com a aplicação de um filtro na coluna “*organisms*”, fosse possível obter apenas proteínas de origem microbianas, sendo excluído os microrganismos de origem vírica e os fungos: *Carnation etched ring vírus (CERV)*, *Encephalomyocarditis vírus (strain emc-b nondiabetogenic)*, *Encephalomyocarditis vírus (strain emc-d diabetogenic)*, *Equine herpesvirus 1 (strain Ab4P) (EHV-1) (Equine abortion vírus)*, *Acanthamoeba polyphaga mimivirus (APMV)*, *Infectious salmon anemia vírus (Isolate Atlantic salmon/ Norway/810/9/99) (ISAV)*, *Bovine herpesvirus type 1.2*, *Candida Albicans (Yeast)*, e *Candida Albicans (strain SC 5314/ATCC MYA-2876)*. De seguida, aplicou-se um filtro na coluna “*proteins*” e eliminaram-se as proteínas não caracterizadas - “*uncharacterized proteins*”. Foram também eliminadas as proteínas duplicadas, visto que na base de dados OralCard as proteínas que foram caracterizadas *in vivo* e *in vitro* estão catalogadas duas vezes, uma para cada categoria.

Após ter obtido as proteínas de origem bacteriana e com o intuito de saber quantas estavam associadas a cada doença e a que tipo de doença, usou-se o código de identificação do Mesh, no qual cada patologia é associada a um número. Assim, após pesquisa no *site* da Pubmed na secção “*mesh database*” (<http://www.ncbi.nlm.nih.gov/mesh>), pôde-se verificar a que patologias da cavidade oral estavam associadas esses códigos.

Aplicaram-se filtros no Microsoft Excel nas colunas “*in vitro*” com o objetivo de saber quantas foram identificadas através dos dados de estudos *in vitro* e quantas foram obtidas a partir de estudos de proteómica de amostras da cavidade oral.

## 3.2 Análise dos processos biológicos em que participam as proteínas bacterianas

A identificação das proteínas bacterianas no OralCard, excluindo as proteínas não caracterizadas, revelou que existem 10428 proteínas de origem bacteriana.

No sentido de esclarecer a importância funcional dessas proteínas presentes no ecossistema da cavidade oral essas mesmas foram classificadas na ferramenta bioinformática Agbase (<http://www.agbase.msstate.edu/>), de acordo com a ontologia processo biológico “*biological process*”. O objetivo foi o de saber em que processo celular estão envolvidas as proteínas. Foi usada a ferramenta bioinformática “Go Retriever”, a qual permite saber rapidamente as anotações GO existentes para um conjunto de dados experimentais (Figura 4) (68, 70).

Para isso, o código Uniprot dessas proteínas foi exportado para um ficheiro com extensão “.txt”, sendo de seguida feito o *upload* desse ficheiro na ferramenta informática “Go Retriever” da AgBase, havendo o cuidado de escolher a opção “UniProt ID” na seção “Accession Type” (Figura 5).

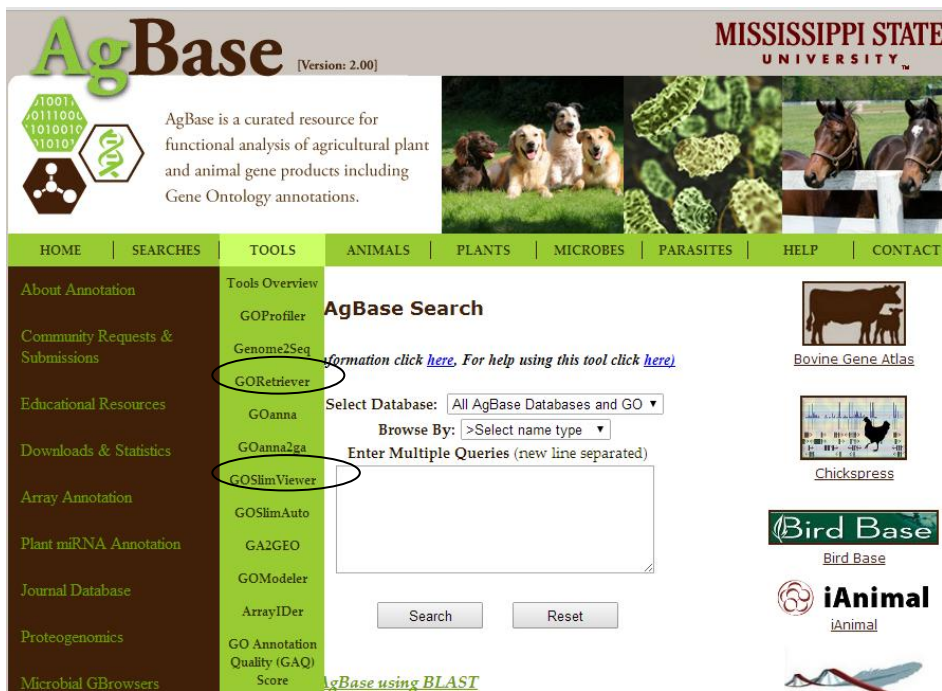


Figura 4- Figura retirada do portal Web da ferramenta bioinformática Agbase (<http://www.agbase.msstate.edu/>). Para obter as ferramentas bioinformáticas “GoRetriever” ou “GoSlim Viewer” deve-se selecionar o campo “Tools”, e, posteriormente, uma dessas ferramentas.

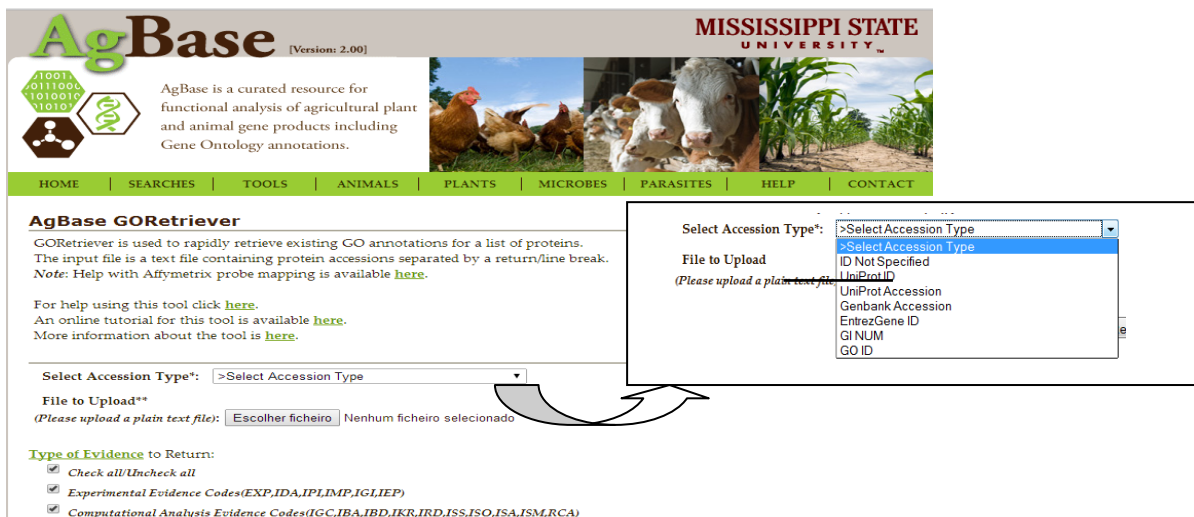
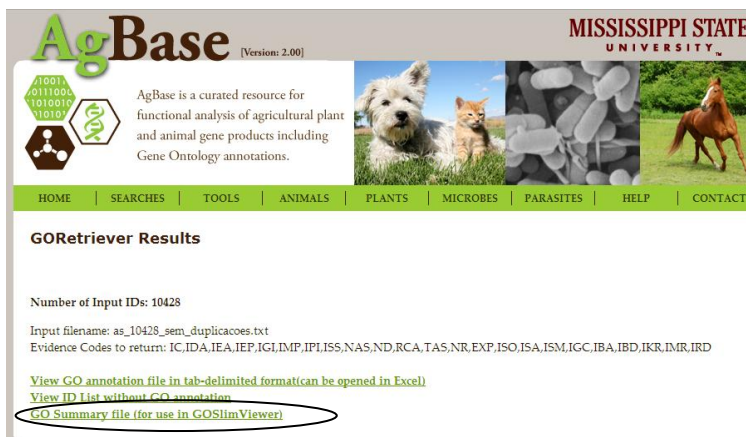


Figura 5- Página inicial da ferramenta informática GoRetriever obtida através do site da <http://www.agbase.msstate.edu/>.

Numa fase posterior, descarregou-se o ficheiro “GO Summary file (for use in GOSlim Viewer) ” e realizou-se o *upload* desse ficheiro na ferramenta “SlimViewer” da Agbase selecionando a opção “TIGR procaryote (BP only) ” na opção “Select GOSlim Set” (Figura 6 e 7). Este ficheiro apresenta, de uma forma simplificada, as ontologias, contendo três colunas separadas: o código Uniprot da proteína, o ID da ontologia e a classificação ontológica a que estão associadas - “P” (processo biológico), “C” (componente celular) ou “M” (função molecular) (Figura 7).

O “SlimViewer” é usado para fornecer um resumo das funções para um conjunto de dados usando “GO slims”. Os “GO Slims” são uma ontologia criada para agrupar termos ontológicos de modo a que estes integrem conjuntos mais amplos. O objetivo foi evitar ter uma lista muito extensa dos termos supracitados, o que dificultava a interpretação. Como tal, a partir desta ferramenta bioinformática, é possível consultar os níveis mais elevados das ontologias génicas das proteínas de interesse (*parent terms*) na coluna “Slim\_GO\_Name”, e respetivas ontologias *child* (na coluna “Input Go Name”) (68, 70).



The screenshot shows the AgBase website interface. At the top, there is a header with the AgBase logo (Version: 2.00) and the Mississippi State University logo. Below the header, there is a navigation menu with links for HOME, SEARCHES, TOOLS, ANIMALS, PLANTS, MICROBES, PARASITES, HELP, and CONTACT. The main content area displays "GORetriever Results" for a search with 10428 input IDs. It provides the input filename and evidence codes. Three links are listed: "View GO annotation file in tab-delimited format (can be opened in Excel)", "View ID List without GO annotation", and "GO Summary file (for use in GOSlimViewer)", which is circled in red.

**Figura 6-** Descarregar o ficheiro “GO Summary file (for use in GOSlimViewer) ” para usar na ferramenta “SlimViewer”.

AgBase [Version: 2.00]

MISSISSIPPI UNIVERSITY

AgBase is a curated resource for functional analysis of agricultural plant and animal gene products including Gene Ontology annotations.

HOME | SEARCHES | TOOLS | ANIMALS | PLANTS | MICROBES

GOSlimViewer

GOSlimViewer is used to provide a high level summary of functions for a dataset. The output includes publication quality figures. Note that records without annotation are not analyzed by GOSlimViewer.

An example of an input file is given [here](#).  
 For help using the tool click [here](#).  
 An online tutorial for this tool is available [here](#).  
 More information about the tool can be found [here](#).  
 A standalone version of the tool can be found [here](#).

Upload GO Summary File:  Nenhum ficheiro selecionado  
 ("Please upload text file!")

Select GOSlim Set:

Search

Number of Input IDs: 10428  
 Input filename: ag\_10428\_sum\_duplicates.txt  
 Evidence Codes to return: IC,IDA,IEA,IEP,IPL,ISS,NAS,ND,BCA,TAS,NR,DP,ISO,ISA,ISM,IGC,IBI

View GO annotation file in tab-delimited format (can be opened in Excel)  
 View ID List without GO annotation  
 GO Summary file (for use in GOSlimViewer)

	A	B	C	D	E	F	G	H	I	J
1	Accessions	GO: IDs	Aspect							
2	AOL319	GO:0001166	F							
3	AOL319	GO:0003677	F							
4	AOL319	GO:0003697	F							
5	AOL319	GO:0005524	F							
6	AOL319	GO:0005737	C							
7	AOL319	GO:0006260	P							
8	AOL319	GO:0006281	P							
9	AOL319	GO:0006974	P							
10	AOL319	GO:0009432	P							
11	AOLCK8	GO:0005622	C							
12	AOLCK8	GO:0005886	C							
13	AOLCK8	GO:0006605	P							
14	AOLCK8	GO:0006810	P							
15	AOLCK8	GO:0006886	P							
16	AOLCK8	GO:0015031	P							
17	AOLCK8	GO:0015450	F							
18	AOLCK8	GO:0016020	C							
19	AOLCK8	GO:0016021	C							

**Figura 7**—Página inicial da ferramenta informática GoSlimViewer obtida através do site da Agbase <http://www.agbase.msstate.edu/>. Seleção da opção “TIGR Prokaryote (BP only)” na seção “select GoSlim Set”. Figura adaptada da ferramenta bioinformática GOSlimViewer consultável no portal Web Agbase.

Deste modo, após a escolha da opção “*View Biological Process*”, obtêm-se os processos biológicos em que estão envolvidas as proteínas bacterianas do microbioma oral e o número destas (que estão associadas a estes processos) (Figura 8).

**AgBase** [Version: 2.00]

AgBase is a curated resource for functional analysis of agricultural plant and animal gene products including Gene Ontology annotations.

HOME | SEARCHES | TOOLS | ANIMALS | PLANTS | MICROBES | FA

**GOSlimViewer Results**  
To download, right click a file and select Save As

To view the detail for cellular components, molecular components, and biological processes click the following link:

[View Cellular Component](#)  
[View Molecular Function](#)  
[View Biological Process](#)  
[View accessions for each slim id](#)

GO:0000160	phosphorelay signal transduction system	68
GO:0000270	peptidoglycan metabolic process	9
GO:0000271	polysaccharide biosynthetic process	50
GO:0000272	polysaccharide catabolic process	6
GO:0000746	conjugation	4
GO:0005975	carbohydrate metabolic process	511
GO:0005976	polysaccharide metabolic process	50
GO:0006040	amino sugar metabolic process	49
GO:0006086	acetyl-CoA biosynthetic process from pyruvate	1
GO:0006091	generation of precursor metabolites and energy	25
GO:0006094	gluconeogenesis	58
GO:0006096	glycolytic process	181
GO:0006098	pentose-phosphate shunt	52
GO:0006099	tricarboxylic acid cycle	47
GO:0006113	fermentation	3
GO:0006139	nucleobase-containing compound metabolic process	1341
GO:0006163	purine nucleotide metabolic process	706
GO:0006220	pyrimidine nucleotide metabolic process	106
GO:0006259	DNA metabolic process	174
GO:0006260	DNA replication	178
GO:0006281	DNA repair	146
GO:0006308	DNA catabolic process	49
GO:0006310	DNA recombination	93
GO:0006351	transcription, DNA-templated	337
GO:0006355	regulation of transcription, DNA-templated	343
GO:0006396	RNA processing	184
GO:0006401	RNA catabolic process	24
GO:0006412	translation	845
GO:0006417	regulation of translation	70
GO:0006457	protein folding	227
GO:0006464	cellular protein modification process	243
GO:0006520	cellular amino acid metabolic process	419
GO:0006595	polyamine metabolic process	5
GO:0006596	cellular biosynthetic process	2

**Figura 8-** Selecionar a opção “View Biological Process”. De seguida, surgem os processos biológicos em que estão envolvidas as proteínas bacterianas do microbioma oral. Ao lado do nome do processo está indicado o número de proteínas envolvidas em cada um destes. Figura adaptada da ferramenta bioinformática GOSlimViewer consultável no portal Web Agbase.

Essas informações foram transportadas para uma folha do Microsoft Excel 2007<sup>®</sup> e elaboraram-se gráficos que permitiram um estudo quantitativo. Para facilitar a interpretação do gráfico, as ontologias que possuem um número de proteínas abaixo de 100 não foram representadas.

A análise do gráfico revela que os processos com mais proteínas associadas são “*biological process*” e “*metabolic process*”. Assim, com o objetivo de realizar uma análise desta classificação ontológica, selecionou-se a opção “*View accessions for each slim id*”; conseqüentemente, descarregou-se o ficheiro “*View accessions for each slim id text only*”, e, por fim, abriu-se esse ficheiro numa folha da *Microsoft Excel® 2007* (Figura 9).

Após ter aplicado um filtro na coluna “Slim Go Name” e escolhido “*biological process*”, foi constatado que, para 3154 proteínas das 3277 associadas a este grupo, além de haver na coluna “Input GO Name” informação mais detalhada sobre o processo em que estas intervêm, também estes grupos estão representados no gráfico da classificação ontológica dos resultados. Apenas 123 proteínas das 3277 não estão associadas aos outros processos biológicos.

A mesma análise foi realizada para o processo biológico “*metabolic process*”, sendo que neste grupo há 792 proteínas que não estão associadas aos outros processos biológicos.

Após isto, eliminaram-se os códigos Uniprot repetidos com a função “remover duplicados” para as 915 proteínas dos grupos “*biological process*” e “*metabolic process*” com pouca informação acerca dos processos biológicos em que estão envolvidas. Assim, restam 829 proteínas (Figura 9).

Finalmente, com a ajuda da função “PROCV” do Microsoft Excel 2007<sup>®</sup> e com o auxílio da base de dados Oral Card (versão de 15-12-2013), realizou-se um gráfico para perceber a que microrganismos estão associados as 829 proteínas. Para isso, selecionou-se a coluna “bacteria”, de seguida escolheu-se a opção “inserir”, e, posteriormente, “tabela dinâmica”, a fim de conseguir saber quais as diferentes estirpes de bactérias que estavam representadas. Por fim, usou-se a função “CONTAR.SE” do Microsoft Excel 2007<sup>®</sup> com o intuito de averiguar quantas proteínas eram produzidas por cada estirpe. Finalmente, para facilitar a interpretação dos gráficos, foram considerados apenas os géneros, realçando que estes só foram representados quando tinham mais de 20 proteínas associadas (Figura 10).

Além disso, notou-se que nos 142 processos biológicos nos quais estão envolvidas as proteínas do microbioma oral, a maioria faz parte do mecanismo de “*housekeeping*”. Com o intuito de saber quantas proteínas estão envolvidas neste mecanismo, aplicou-se um filtro na coluna “Slim\_GO\_Name”, excluindo assim os seguintes processos biológicos: “*pathogenesis*”, “*killing of cells of others organisms*”, “*biofilm formation*”, “*symbiosis, encompassing mutualismo through parasitism*” e “*multiorganism process*”. Com a função “remover duplicados”, eliminaram-se as proteínas repetidas.

**AgBase** [Version: 2.00]

AgBase is a curated resource for functional analysis of agricultural plants and animal gene products including Gene Ontology annotations.

HOME | SEARCHES | TOOLS | ANIMALS

**GOSlimViewer Results**  
To download, right click a file and select Save As...

To view the detail for cellular components, click on the following links:

- [View Cellular Component](#)
- [View Molecular Function](#)
- [View Biological Process](#)
- [View accessions for each slim id](#)

**GOSlimViewer accession detail for each slim id**

Index

- Cellular Component Accessions
- Molecular Function Accessions
- Biological Process Accessions

**Cellular Component Accessions** Back to top

	A	B	C	D	E	F
1	GO_Typ	Slim_GO_ID	Slim_GO_Name	Input_Accession	Input_GOID	Input_GO_Name
17024	P	GO:0008150	biological_process	A8AUH5	GO:0006066	alcohol metabolic process
17025	P	GO:0008150	biological_process	C7ND57	GO:0006066	alcohol metabolic process
17026	P	GO:0008150	biological_process	Q1RFM3	GO:0006066	alcohol metabolic process
17027	P	GO:0008150	biological_process	Q818A4	GO:0006066	alcohol metabolic process
17028	P	GO:0008150	biological_process	Q8DWB9	GO:0006066	alcohol metabolic process
17029	P	GO:0008150	biological_process	Q8ZP45	GO:0006066	alcohol metabolic process
17030	P	GO:0008150	biological_process	Q97NK0	GO:0006066	alcohol metabolic process
17031	P	GO:0008150	biological_process	P16682	GO:0042916	alkylphosphonate transport
17032	P	GO:0008150	biological_process	Q8DU65	GO:0018364	peptidyl-glutamine methylation
17033	P	GO:0008150	biological_process	A6VDT2	GO:0044237	cellular metabolic process
17034	P	GO:0008150	biological_process	A8AWR5	GO:0044237	cellular metabolic process
17035	P	GO:0008150	biological_process	A8AW20	GO:0044237	cellular metabolic process
17036	P	GO:0008150	biological_process	A8AWZ1	GO:0044237	cellular metabolic process
17037	P	GO:0008150	biological_process	A8AZ66	GO:0044237	cellular metabolic process
17038	P	GO:0008150	biological_process	A8AZH0	GO:0044237	cellular metabolic process
17039	P	GO:0008150	biological_process	A8AZQ4	GO:0044237	cellular metabolic process
17040	P	GO:0008150	biological_process	A8AZR0	GO:0044237	cellular metabolic process
17041	P	GO:0008150	biological_process	B2RUH2	GO:0044237	cellular metabolic process
17042	P	GO:0008150	biological_process	B2RUQ2	GO:0044237	cellular metabolic process
17043	P	GO:0008150	biological_process	B2RKJ4	GO:0044237	cellular metabolic process
17044	P	GO:0008150	biological_process	B2RL25	GO:0044237	cellular metabolic process
17045	P	GO:0008150	biological_process	B2RLB1	GO:0044237	cellular metabolic process
17046	P	GO:0008150	biological_process	D1Y7MA4	GO:0044237	cellular metabolic process

**Figura 9-** Selecionar a opção “View accesions for each slim id” na página dos resultados da ferramenta bioinformática GOSlimViewer. De seguida, o ficheiro obtido do GoSlimViewer “View accessions for each slim id text only” foi descarregado e aberto numa folha da Microsoft Excel® 2007. Posteriormente, aplicou-se um filtro na coluna “Slim\_GO\_Name”, selecionando-se apenas o processo “*biological process*”. Desta forma, obtiveram-se na coluna “Input\_Accession” apenas os códigos Uniprot das proteínas envolvidas neste processo. É de enfatizar que na coluna “Input GO Name” existem informações mais detalhadas sobre o processo em que estas intervêm. Figuras adaptadas da ferramenta bioinformática GOSlimViewer consultável no portal Web Agbase e do Microsoft Excel® 2007.



### **3.3 Identificação das proteínas bacterianas no microbioma oral que interagem com outros microrganismos e com o hospedeiro**

Foi realizada a análise das 10428 proteínas de origem bacterianas do microbioma oral no STRAP. Para este estudo, transferiram-se os códigos Uniprot dessas proteínas para um ficheiro com extensão “.txt” para de seguida as importar para o programa STRAP.

Esta ferramenta, STRAP, é capaz de anotar as ontologias de grandes listas de proteínas. Para cada código Uniprot obtêm-se as seguintes informações: nome da proteína- “protein name”; nome do gene primário- “primary name gene”; taxonomia- “taxonomy”; comprimento- “length”; função- “function”; atividade catalítica- “catalytic activity” e vários termos GO.

Os termos GO estão organizados em três categorias principais: Processo Biológico, Componente Celular e Função Molecular, sendo que cada uma tem diferentes símbolos representativos. Além disso, cada categoria é subdividida em subcategorias que se sustentam numa versão simplificada das categorias originalmente publicadas pelo Consórcio Gene Ontology (69).

Ao analisar a ontologia “*biological process*”, existe uma subcategoria “*Interaction with cells and organisms*” (interação com células e organismos) em que estão identificadas, com um triângulo verde, as proteínas que lhe estão associadas, sendo essas que interagem com outros microrganismos e com o hospedeiro. Foram identificadas 351 proteínas nessa subcategoria (Figura 11).

The screenshot shows the STRAP software interface with a table of proteins. The table has columns for Length, Function, Catalytic Activity, GO: Biological Process, GO: Cellular Component, GO: Molecular Function, Cellular process, Growth, Immune system process, Interaction with cells and organisms (highlighted with a green triangle), and Localization. The table lists various proteins such as SC5314, strain Chalis, and others, with their respective GO terms and impact indicators.

Protein	Length	Function	Catalytic Activity	GO: Biological Process	GO: Cellular Component	GO: Molecular Function	Cellular process	Growth	Immune system process	Interaction with cells and organisms	Localization
SC5314 / ATCC MYA-2876	568		5-amino-1-(5-phos...	GO:0006189 cellular pr...		GO:0034023 (catalytic ...	▲			▲	
(strain Chalis / ATCC 35105 / CH1 ...	694			GO:0007155 interacto...	GO:0005618 (other, cell...					▲	
	2978			GO:0009405 interacto...		GO:0004197 (catalytic ...				▲	
(strain 342)	344			GO:0007155 interacto...	GO:0009289 (other, cell...					▲	
	865			GO:0009405 interacto...	GO:0005576 (extracellul...	GO:0005509 (binding, i...				▲	
us ATCC 4796	3517			GO:0009405 interacto...	GO:0005618 (other, cell...					▲	
17748	3531			GO:0009405 interacto...	GO:0019867 (other, out...					▲	
17748	2778			GO:0009405 interacto...	GO:0019867 (other, out...					▲	
us (strain Lc 705)	1794			GO:0007155 interacto...		GO:0030246 (binding, c...				▲	
A9635	231			GO:0009405 interacto...	GO:0005576 (extracellul...					▲	
ATCC 49175	1750			GO:0007155 interacto...	GO:0005576 (extracellul...	GO:0033926 (catalytic ...				▲	
703	781			GO:0009405 interacto...	GO:0019867 (other, out...					▲	
703	363			GO:0009405 interacto...	GO:0019867 (other, out...					▲	
virus (strain emc-b nondiabetogenic)	2292	RNA-direct... NTP + H <sub>2</sub> O = N...	GO:0006811 localizati...	GO:0044162 (extracellul...	GO:0005524 (binding, c...		▲			▲	▲
virus (strain emc-d diabetogenic)	2292	RNA-direct... NTP + H <sub>2</sub> O = N...	GO:0006811 localizati...	GO:0044162 (extracellul...	GO:0005524 (binding, c...		▲			▲	▲
	159			GO:0007155 interacto...	GO:0009289 (other, cell...					▲	
olosis	464	Appears to ... Protein N16)-dihy...	GO:0045454 regulatio...	GO:0005829 (cytoplasm...	GO:0016209 (antioxi...		▲	▲		▲	
serotype O3:K5 (strain RIMD 22106...	445	Required fo...		GO:0044781 cellular c...	GO:0009424 (other intra...		▲			▲	
	153			GO:0007155 interacto...	GO:0009289 (other, cell...					▲	
es serotype M12 (strain MGAS2096)	1161			GO:0007155 interacto...	GO:0005618 (other, cell...					▲	
(strain USA300)	300	Contributes ...		GO:0009405 interacto...						▲	
(strain MR-1)	374	Involved in ... Protein L-glutama...	GO:0006935 localizati...	GO:0005737 (cytoplasm...	GO:0000156 (molecul...					▲	▲
riae serotype 4 (strain ATCC BAA-33...	434	Catalyzes th... 2-phospho-D-glyc...	GO:0006096 cellular pr...	GO:0009986 (cell surfa...	GO:0000287 (binding, i...		▲			▲	
in J99 / ATCC 700824)	685	Required fo...		GO:0007155 interacto...	GO:0009424 (other intra...					▲	

**Figura 11-** Figura do STRAP onde estão representadas, por um triângulo verde, as proteínas com mais impacto no hospedeiro na subcategoria “Interaction with cells and organisms”.

### 3.4 Classificação ontológica das proteínas que interagem com outros microrganismos e com o hospedeiro

As 351 proteínas que interagem com outros microrganismos e com o hospedeiro foram exportadas para um ficheiro “.txt”, tendo-se realizado a classificação ontológica através do programa Agbase, usando os mesmos métodos anteriormente descritos para as 10428 proteínas de origem bacterianas do Oral Card.

Da mesma forma, através da opção “View Biological Process”, obtemos os processos biológicos em que estão envolvidas as proteínas bacterianas do microbioma oral e observamos quantas estão associadas a estes processos. Os dados foram, posteriormente, exportados para uma folha do Microsoft Excel 2007<sup>®</sup> e elaboraram-se gráficos com o intuito de realizar um estudo quantitativo

do número de proteínas associadas a cada ontologia. Optou-se por fazer a análise na ferramenta bioinformática Agbase por ser mais informativa a nível da classificação ontológica em comparação com a ferramenta STRAP.

### **3.5 Pesquisa bibliográfica sobre a função das proteínas bacterianas envolvidas na patogénese**

Após a análise das 351 proteínas que interagem com outros microrganismos e com o hospedeiro na ferramenta Agbase, concluiu-se que 298 destas estão envolvidas na patogénese. Os códigos Uniprot destas proteínas foram exportados para uma folha do Microsoft Excel 2007<sup>®</sup>, tendo-se usado a função “PROCV” para associar a cada código o nome da proteína e o microrganismo que a produz.

O estudo do nível de evidência das proteínas em questão, na ferramenta bioinformática “Retriever”, consultável no portal web UniprotKB, revelou que 44 destas 298 proteínas têm evidência a nível da proteína e que 254 têm evidência a nível do gene. Para isto, os códigos Uniprot das 298 proteínas envolvidas na patogénese foram inseridos na caixa “UniProt identifiers” do “Retriever”, seleccionou-se “UniProtKB (298)”, e, por fim, através da opção “download”, os resultados foram transportados para uma folha do *Microsoft Excel 2007*<sup>®</sup> na qual a coluna “Protein evidence” indica o tipo de evidência que suporta a existência da proteína (Figuras 12 e 13).

The image shows the UniProtKB web interface. On the left, under 'UniProt identifiers', there is a list of identifiers and a section for downloading 298 entries. The 'UniProtKB (298)' option is selected. An arrow points from this selection to the main search results page. The search query is 'job:2014041272D33MZQP8'. The results show a list of proteins, with 'Biofilm regulatory protein A' and 'C5a peptidase' highlighted. A 'Download' button is circled in the top right corner of the results panel.

**Figura 12-** Inserir na caixa “UniProt identifiers” do “Retriever” as 298 proteínas envolvidas na patogênese. Após isto, selecionar “UniProtKB (298)”, e, por fim, através da opção “download”, obtêm-se os resultados numa folha da Microsoft Excel® 2007. Figuras adaptadas do portal Web UniProtKB.

Na base de dados UniProtKB as proteínas estudadas neste trabalho apresentam 4 tipos de provas para à existência da proteína;

1. Evidência a nível da proteína- indica que há evidências experimentais claras quanto à existência da proteína;
2. Evidência a nível da transcrição- indica que a existência de uma proteína não foi rigorosamente provada mas que há provas da existência de um transcrito;
3. Inferida a partir de homologia- a existência da proteína é provável porque existem proteínas ortólogas em espécies estreitamente relacionadas;
4. Prevista- não há provas da existência quer da proteína quer de transcritos da mesma, ou provas a nível da homologia.

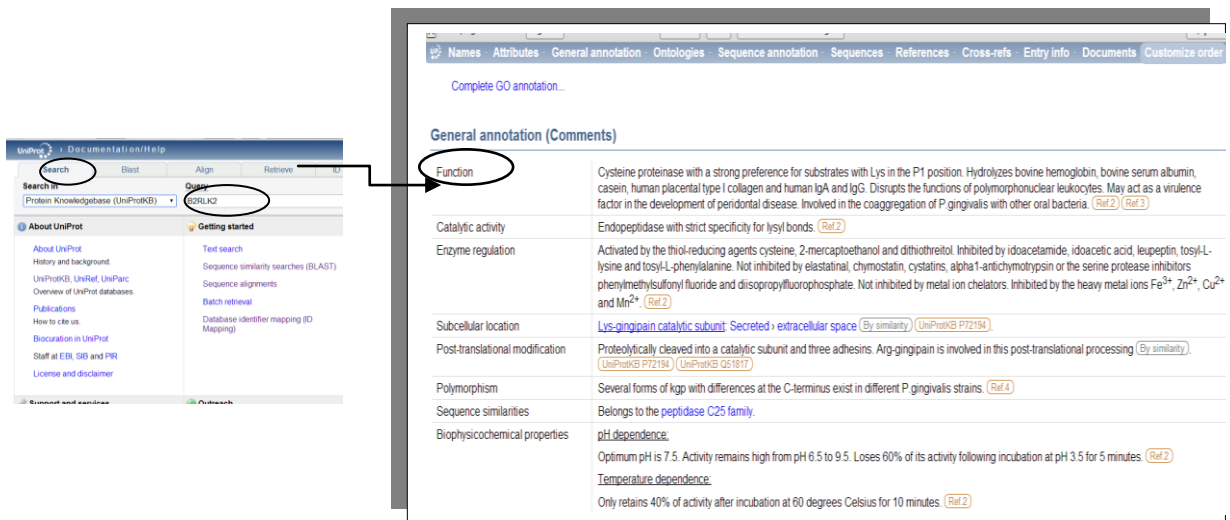
Entry	Entry n	Status	Protein names	Organism	General annotation (function)
1	Q8DVRO	BRPA_STR reviewed	Biofilm regulatory protein A	Streptococcus mutans serotype c (strain AT)	FUNCTION: Involved in biofilm formation
2	Q87120	CDTA_AG reviewed	Cytolysin toxin subunit A (CDT A)	Aggregatibacter actinomycetemcomitans (ATCC 49619)	FUNCTION: CDTs are cytotoxins which induce cell death
3	P28784	CPG1_POF reviewed	Gingipain R1 (EC 3.4.22.37) (Arg-gingipain) (Gingipain 1) (RGP)	Porphyromonas gingivalis (strain ATCC 35410)	FUNCTION: Thiol protease which is believed to be involved in the invasion of epithelial cells
4	P95493	CPG2_POF reviewed	Gingipain R2 (EC 3.4.22.37) (Arg-gingipain) (Gingipain 2) (RGP)	Porphyromonas gingivalis (strain ATCC 35410)	FUNCTION: Thiol protease which is believed to be involved in the invasion of epithelial cells
5	P66004	DLDH_MYI reviewed	Dihydrolipooyl dehydrogenase (LPD) (EC 1.8.1.4) (Component 1)	Mycobacterium tuberculosis (strain H37Rv)	FUNCTION: Lipamide dehydrogenase is involved in the synthesis of lipids
6	Q80T59	ENO_STRN reviewed	Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-)	Streptococcus mutans serotype c (strain AT)	FUNCTION: Catalyzes the reversible conversion of 2-phosphoglycerate to phosphoenolpyruvate
7	Q97Q52	ENO_STRF reviewed	Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-)	Streptococcus pneumoniae serotype 4 (strain ATCC 49619)	FUNCTION: Catalyzes the reversible conversion of 2-phosphoglycerate to phosphoenolpyruvate
8	Q80P50	ENO_STRF reviewed	Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-)	Streptococcus pneumoniae (strain ATCC 49619)	FUNCTION: Catalyzes the reversible conversion of 2-phosphoglycerate to phosphoenolpyruvate
9	B2RH54	FIMA1_PC reviewed	Major fimbrial subunit protein type-1 (Fimbriillin) (Fimbriillin)	Porphyromonas gingivalis (strain ATCC 33277)	FUNCTION: Fimbriillin is the structural subunit of the major fimbriae
10	P59915	HAGA1_P reviewed	Hemagglutinin A	Porphyromonas gingivalis (strain ATCC 35410)	FUNCTION: Agglutinates erythrocytes by binding to sialic acid
11	P35672	INVG_SAL reviewed	Protein InvG	Salmonella typhimurium (strain LT2 / SGSC1)	FUNCTION: Involved in the invasion of epithelial cells
12	B2RLK2	KGP_POR reviewed	Lys-gingipain (EC 3.4.22.47) (Lysine-specific cysteine proteinase)	Porphyromonas gingivalis (strain ATCC 33277)	FUNCTION: Cysteine proteinase with a stannous ion requirement
13	P72194	KGP38_PC reviewed	Lys-gingipain 381 (EC 3.4.22.47) [Cleaved into: Lys-gingipain c]	Porphyromonas gingivalis (strain ATCC 33277)	FUNCTION: Cysteine proteinase with a stannous ion requirement
14	Q51817	KGP83_PC reviewed	Lys-gingipain W83 (EC 3.4.22.47) (Lysine specific cysteine proteinase)	Porphyromonas gingivalis (strain ATCC 33277)	FUNCTION: Cysteine proteinase with a stannous ion requirement
15	Q82R19	PAGN_SAI reviewed	Outer membrane protein PagN (Adhesin/Invasin protein PagI)	Salmonella typhimurium (strain LT2 / SGSC1)	FUNCTION: Haemagglutinin that facilitates adherence to epithelial cells
16	P0DM78	PHOP_SAI reviewed	Virulence transcriptional regulatory protein PhoP (Acid shock protein)	Salmonella typhimurium (strain LT2 / SGSC1)	FUNCTION: Member of the two-component regulatory system
17	Q8ZMM8	PIP2_SAI reviewed	Secreted effector protein pipB2 (Type III effector pipB2)	Salmonella typhimurium (strain LT2 / SGSC1)	FUNCTION: Effector proteins function to inhibit the host immune response
18	P41783	PRGH_SAI reviewed	Protein PrgH	Salmonella typhimurium (strain LT2 / SGSC1)	FUNCTION: Required for invasion of epithelial cells
19	P41784	PRGI_SAI reviewed	Protein PrgI	Salmonella typhimurium (strain LT2 / SGSC1)	FUNCTION: Required for invasion of epithelial cells
20	Q56061	SIFA_SALT reviewed	Secreted effector protein Sifa	Salmonella typhimurium (strain LT2 / SGSC1)	FUNCTION: Effector proteins function to inhibit the host immune response
21	E1WAC6	SIPA_SALT reviewed	Cell invasion protein sipA (Effector protein sipA)	Salmonella typhimurium (strain SL1344)	FUNCTION: Actin-binding protein that inhibits the host immune response
22	Q56019	SIPB_SALT reviewed	Cell invasion protein SipB (Effector protein SipB)	Salmonella typhimurium (strain LT2 / SGSC1)	FUNCTION: Required for entry into the host cell
23	POCL52	SIPA_SALT reviewed	Cell invasion protein SipA (Effector protein SipA)	Salmonella typhimurium (strain LT2 / SGSC1)	FUNCTION: Actin-binding protein that inhibits the host immune response
24	POCL47	SIPC_SALT reviewed	Cell invasion protein SipC (Effector protein SipC)	Salmonella typhimurium (strain LT2 / SGSC1)	FUNCTION: Actin-binding protein that inhibits the host immune response

**Figura 13-** Resultados obtidos através da ferramenta bioinformática “Retriever” do portal Web UniProtKB. Na coluna “protein existence”, está indicado o nível de evidência das 298 proteínas envolvidas na patogênese. A aplicação de um filtro nesta coluna revela que 44 destas proteínas têm evidência a nível da proteína.

A aplicação de um filtro na coluna “Status” mostrou que 65 dessas 298 proteínas foram revistas (“Reviewed”), ou seja, foram anotadas manualmente através de dados obtidos por literatura (Figura 13). As restantes estão anotadas como “Unreviewed”, o que corresponde a registos analisados computacionalmente por algoritmos que aguardam uma revisão manual.

Deste modo, e com a finalidade de saber de que forma estas 65 proteínas podem contribuir para a virulência do microrganismo, recorreu-se à base de dados UniprotKB (<http://www.uniprot.org/>), mais particularmente à ferramenta bioinformática “search”, com o intuito de compreender de que modo as proteínas atuam. Esta base de dados fornece informação funcional e estrutural sobre cada proteína, incluindo um nome recomendado, a classificação taxonómica do organismo em que a proteína está presente, referências bibliográficas e mesmo a sequência da proteína. Desta forma, conhecemos a função da proteína na secção “general annotation (function)” (Figura 14). Para as proteínas que não foram revistas exploraram-se os arquivos digitais de artigos científicos PubMed ou Sciencedirect com a seguinte fórmula de pesquisa: (“nome da proteína”) AND (“género específico da

bactéria”) OR (“gênero da bactéria”). Caso não houvesse resultados nestes optava-se, por último, pela pesquisa de artigos no Google Scholar com as mesmas palavras-chaves.



**Figura 14-** Figura retirada e adaptada do portal Web UniprotKB. Um dos códigos associados ao nome da proteína é inserido na caixa de texto “Query”, visto ter sido considerado que proteínas com o mesmo nome têm funções idênticas. Desta forma, após ter lançado a pesquisa, surge o resultado da função da proteína na secção “*general annotation (function)*”. Aqui, foi feita a simulação para a proteína cujo código Uniprot é B2RLK2 (Lys-gingipain).



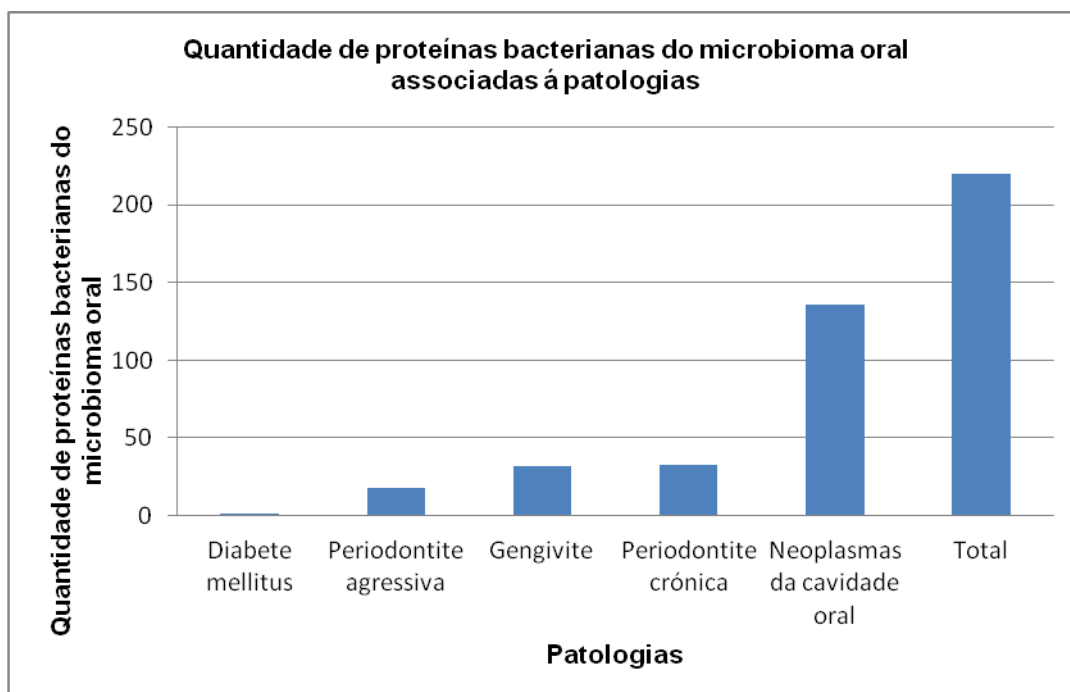
## **4.RESULTADOS E DISCUSSÃO**



## 4.1 Catalogação das proteínas de origem microbiana

Os resultados obtidos da análise da base dados do OralCard, versão de 15-12-2013, revelaram que 11110 proteínas são de origem microbiana e que 46 dessas proteínas são propostas como biomarcadores de patologias orais. Dessas 11110 proteínas microbianas, 9765 foram obtidas através dos dados de estudos *in vitro*, sendo que as restantes foram obtidas através de estudos de proteômica de amostras da cavidade oral (71, 72). Das amostras de tecidos da cavidade oral, verifica-se que há 220 proteínas que foram identificadas em amostras de pacientes com patologias da cavidade oral, nomeadamente gengivite, periodontite crónica, periodontite agressiva, diabete mellitus e neoplasmas.

A quantidade de proteínas identificadas para cada patologia é apresentada na figura 15.



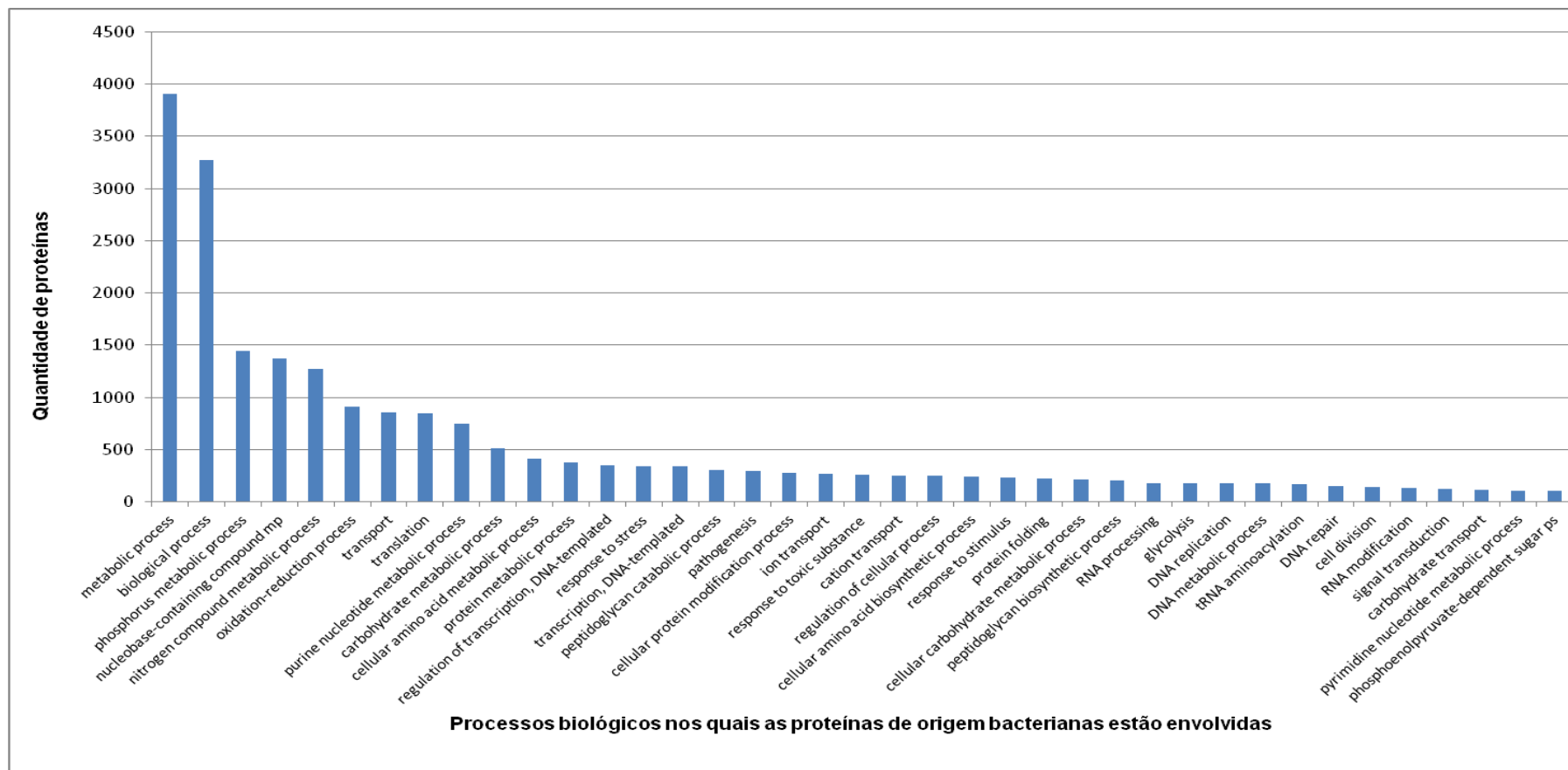
**Figura 15-** Quantidade de proteínas microbianas do microbioma oral identificadas em amostras de saliva de paciente com patologias.

O gráfico revela que existem mais proteínas bacterianas do microbioma oral associadas aos neoplasmas da cavidade oral, enquanto que há um número reduzido associado a diabetes mellitus. No entanto, a base de dados OralCard mostra que existe apenas um estudo que identifica proteínas do neoplasma, ao contrário da periodontite crónica ou diabetes mellitus, as quais têm vários estudos. Esta disparidade pode ter como causa as técnicas, sendo que a espectrometria de massa (que foi usada na identificação de proteínas bacterianas associadas aos neoplasmas) costuma ser mais sensível e dar origem a maior quantidade de identificações.

## **4.2 Análise da classificação ontológica das proteínas de origem microbianas**

Com o objetivo de saber em que processos biológicos estão envolvidas cada uma das proteínas bacterianas presentes no ecossistema oral, usou-se a ferramenta bioinformática Agbase, isto após eliminação das proteínas não caracterizadas (*“uncharacterized protein”*). A exclusão destas proteínas deve-se à pouca quantidade de informação existente, o que impede uma análise completa. Desta forma, foram obtidas 10428 proteínas de origem bacteriana.

Para facilitar a interpretação e a análise do gráfico, e visto que as proteínas bacterianas no microbioma oral estão associadas a 142 processos, só foram representados no gráfico os processos que possuem um número de proteínas acima de 100 (Figura 16).



**Figura 16** – Classificação quantitativa ontológica de acordo com o Processo Biológico (“*Biological Process*”) das proteínas de origem bacteriana do microbioma oral. Esta classificação foi obtida recorrendo ao programa Agbase (<http://www.agbase.msstate.edu>). O gráfico apresenta a quantidade de proteínas associadas aos processos biológicos nos quais estas intervêm. Com o objectivo de facilitar a visualização do gráfico, usaram-se as abreviaturas “mp” e “ps” que correspondem, respetivamente, a “*metabolic process*” e a “*phosphotransferase system*” nos seguintes processos: “*nucleobase - containing compound metabolic process*” e “*phosphoenolpyruvate-dependent sugar phosphotransferase system*”.

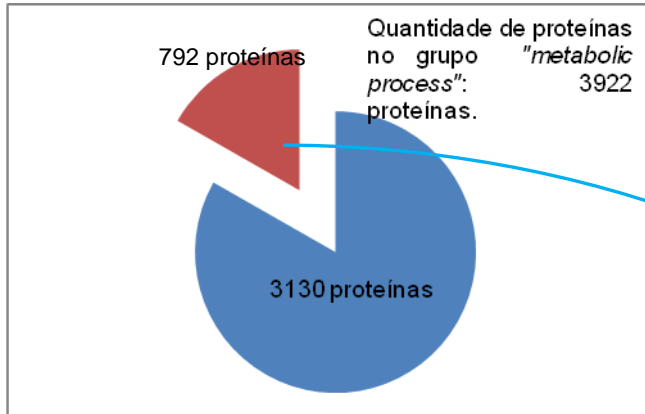
Na figura 16 são apresentados os resultados referentes aos processos biológicos em que as proteínas de origem bacteriana do microbioma oral intervêm. Da análise do gráfico, verifica-se que as categorias “*biological process*” e “*metabolic process*” são aquelas que reúnem o maior número de proteínas.

No sentido de verificar se para as proteínas incluídas nos processos “*biological process*” e “*metabolic process*” existe outra anotação adicional, foi feito o estudo detalhado das proteínas associadas a esses processos (Figura 17). Nesta figura, verificamos que no caso das 3277 proteínas associadas ao grupo “*biological process*”, apenas 123 destas não estão associadas aos outros processos, estando apenas alguns apresentados na figura 16. A mesma análise para o processo “*metabolic process*” mostrou que 792 das 3922 proteínas não estão envolvidas nos outros processos apresentados na figura 16. Assim sendo, para 829 das 10428 proteínas do microbioma oral de origem bacteriana (ou seja, cerca de 8%), não se consegue ter informação ontológica suficiente em termos de processo biológico.

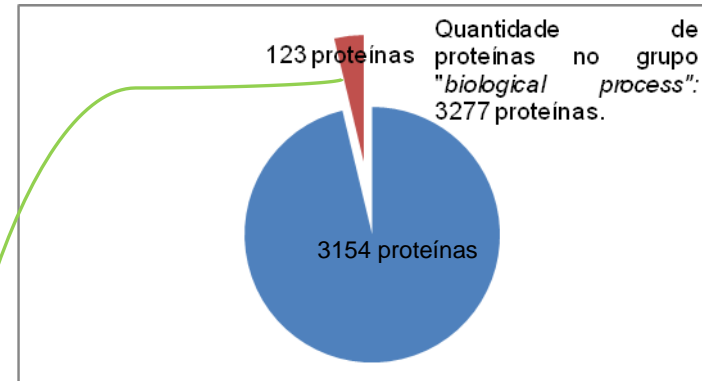
A maioria das bactérias associadas a estas proteínas fazem parte das mais estudadas para a cavidade oral: *Streptococcus mutans*, *Porphyromona gingivalis*, *Enterococcus faecalis*, *Streptococcus gordonii* e *Bacillus cereus* (Figura 17). Este resultado mostra que mesmo para estas bactérias, bastante estudadas, a anotação total do seu proteoma ainda não existe, isto apesar de já ser conhecido o genoma destas espécies (73-77).

Da análise da figura 16 é possível concluir ainda que a maioria das proteínas bacterianas está anotada como participando em processos “*housekeeping*” da célula (à exceção do processo “*pathogenesis*” em que participam 298 proteínas). Destacamos ainda o facto de haver um número relativamente grande de proteínas dedicadas ao transporte de forma genérica (857 em “*transport*”) e ao transporte de hidratos de carbono (117 em “*carbohydrate transport*”). Este facto pode ser explicado pela dependência que a célula bacteriana tem dos sistemas de transporte membranar para conseguir nutrientes para o seu metabolismo. Um exemplo reside nos sistemas de transporte PTS (*phosphotransferase system*), os quais são específicos das

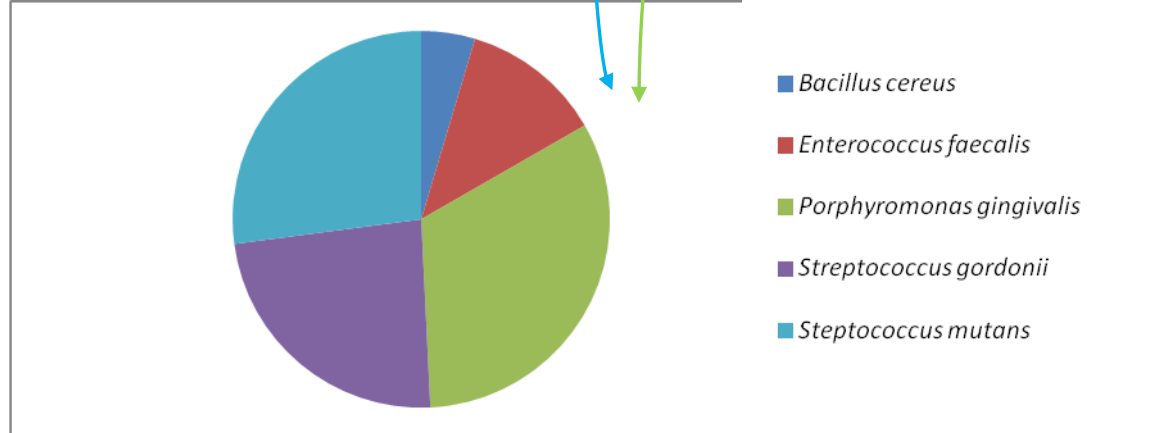
células bacterianas e que envolvem um grande número de proteínas transmembranares.



**Diagrama 1.** Está representada a azul a quantidade de proteínas que estão no grupo "metabolic process", mas que também estão associadas a outros processos biológicos, alguns dos quais estão no gráfico da **figura 16**. O vermelho representa a quantidade de proteínas do grupo "metabolic process" que não estão associadas a outros processos, alguns dos quais estão no gráfico da **figura 16**.



**Diagrama 2.** Está representado a azul a quantidade de proteínas que estão no grupo "biological process" e que se encontram também associadas aos outros processos, alguns dos quais estão no gráfico da **figura 16**. O vermelho representa a quantidade de proteínas do grupo "biological process" que não estão associadas a outros processos, alguns dos quais estão no gráfico da **figura 16**.



**Diagrama 3.** Espécies com maior representação nas proteínas dos grupos "biological process" e "metabolic process", que não estão associadas aos outros processos biológicos, alguns dos quais estão representados na **figura 16**. Para facilitar a interpretação do gráfico, apenas estão representados os géneros específicos das bactérias e os géneros que possuíam um número de proteínas associadas superior a 20.

**Figura 17 -** Análise detalhada dos grupos "biological process" e "metabolic process" do gráfico da **figura 16**.

### 4.3 Análise da classificação ontológica das proteínas microbianas do microbioma oral que interagem com outros microrganismos e com o hospedeiro

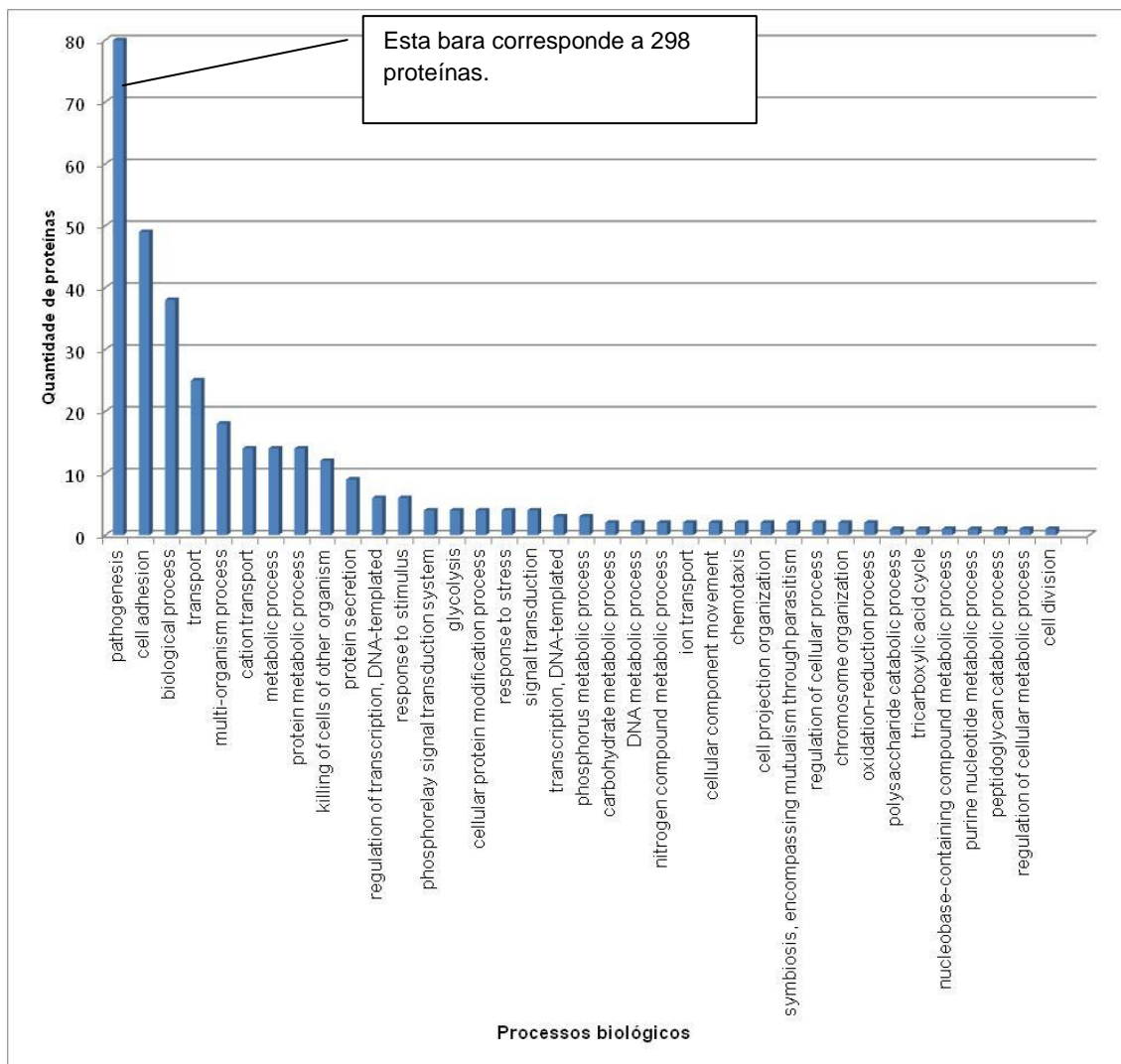
Como seria de esperar, a análise das proteínas de origem bacteriana da cavidade oral revela uma grande quantidade de processos biológicos presentes que seriam impossíveis de analisar em detalhe. No âmbito desta tese, optou-se por se analisar apenas as proteínas identificadas que interagem com células e organismos. Esta escolha decorre do interesse em conhecer as relações que se estabelecem entre os vários microrganismos da cavidade oral e entre estes e o hospedeiro.

Usando as ferramentas descritas na secção “Material e Métodos”, foram usadas as 351 proteínas que, segundo a classificação ontológica “*biological process*”, estão na categoria “*Interaction with cells and organisms*”, e, portanto, estabelecem interações com outras células.

Na figura 18 verifica-se que quase todas as proteínas estão associadas à patogénese, isto é, 298 das 351 proteínas (85%). No entanto, estas apenas estão associadas à ontologia génica “*pathogenesis*” sem *child term* associado.

Como foi referido na introdução, o *PAMGO Plant-Microbe Associated Gene Ontology* (<http://pamgo.vbi.vt.edu/>) introduziu novos termos GO para descrever processos envolvidos na interação entre os micróbios e os seus hospedeiros. No entanto, alguns autores defendem que muitos dos termos que a PAMGO adicionou não são específicos para fatores de virulência e que são mais adaptados aos genes que facilitam mutualismo. Além disso, estes não incluem os efeitos dos fatores de virulência nos sistemas funcionais do hospedeiro. Adicionalmente, muitos dos processos gerais envolvidos na patogénese são *child terms* do grupo “*symbiosis, encompassing mutualismo through parasitism*”, como é o caso de “*killing of cells of other organism*”. Consequentemente, estes autores alertam para a necessidade de criar um vocabulário controlado para os fatores de virulência sem incluir termos que

fazem referência a mecanismos de “*housekeeping*”. Estes também indicam que a ontologia “*molecular function*” tem termos que descrevem atividades moleculares que se aplicam a VF, podendo assim ser usados para descrever o local onde o VF atua (37, 67). O termo que melhor faz referência ao VF é o GO “*pathogenesis*”. Este termo geral não dá conta do universo abrangente dos VFs existentes, pelo que, cremos, deveria servir de base para o estabelecimento de outros *child terms* que dessem conta dos mecanismos específicos envolvidos na patogênese.



**Figura 18** - Classificação ontológica de acordo com o Processo Biológico (Biological Process) das proteínas de origem bacteriana do microbioma oral que interagem com outros microrganismos e com o hospedeiro. Esta classificação foi obtida com recurso ao programa Agbase (<http://www.agbase.msstate.edu/>). O gráfico apresenta o número de proteínas associadas aos processos biológicos nos quais estas intervêm. Com o intuito de visualizar melhor o número de proteínas correspondentes aos processos com pouca quantidade de proteínas, a barra correspondente ao processo biológico “*pathogenesis*” não corresponde ao valor indicado no eixo dos Y mas sim a 298 proteínas. No total encontram-se representadas 351 proteínas.

A figura 19 revela que das 298 proteínas apenas 53 (15%) estão associadas a outros processos biológicos, e que alguns desses processos (“*cell adhesion*”, “*killing of cells of other organism*” e “*celular protein modification*”

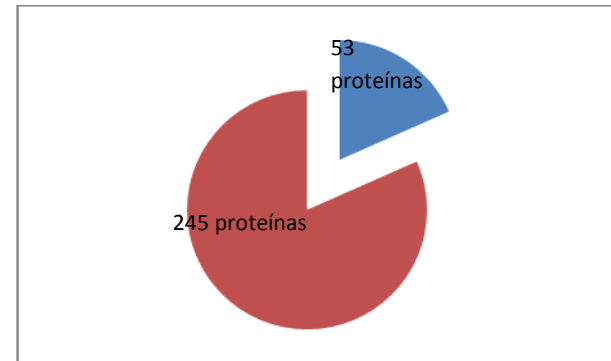
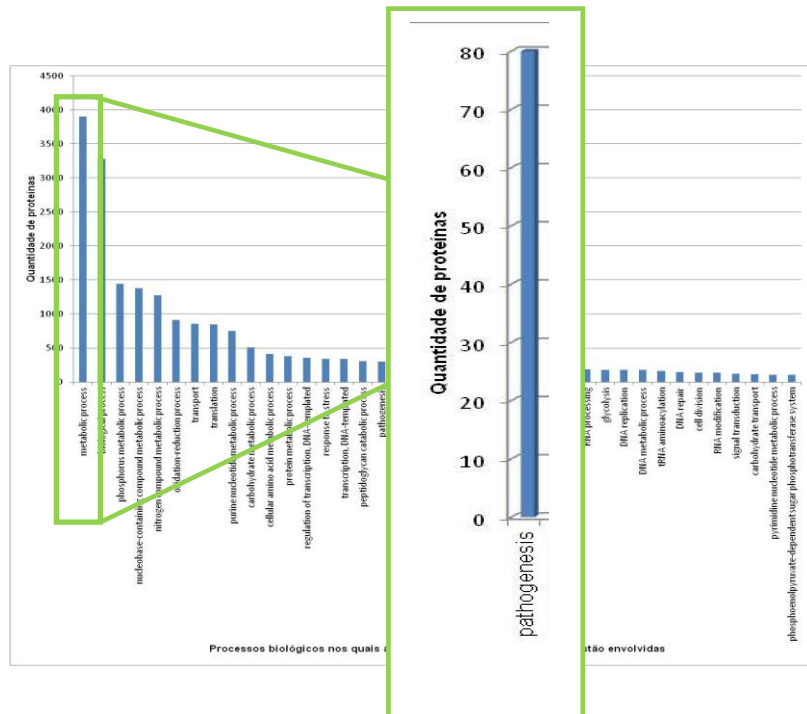
process”) poderiam ser *child terms* da ontologia “*pathogenesis*”. Para as restantes 245 proteínas apenas se sabe que estão envolvidas na patogénese sem mais descrição em termos de processo biológico. A análise da ontologia “*molecular function*” mostrou que algumas das funções poderiam ser *child terms* do termo GO “*pathogenesis*” para descrever fatores de virulência como: “*peptidase activity*”, “*lyase activity*”, ligadas à atividade hidrolítica; “*cytoskeletal protein binding*”, “*lipid binding*” e “*ion binding*” que podem estar ligadas à adesão das bactérias às células do hospedeiro; “*transferase activity, transferring glycosyl groups*” que fazem parte de modificações pós-tradução e que, através da adição do grupo “*glycosyl*”, podem regular a atividade das proteínas humanas. Adicionalmente, estes *child terms* do termo GO “*pathogenesis*” estão também associados a outras ontologias.

Os microrganismos produtores de 245 das 298 proteínas, que não estão associados aos outros processos além da patogénese, são: *Bacillus cereus*, *Aggregatibacter actinomycetemcomitans*, *Aggregatibacter aphrophilus* e *Salmonella typhimurium* (Figura 19). Estes microrganismos fazem parte de grupos de bactérias muito estudados, sendo que o *Bacillus cereus* e a *Salmonella typhimurium* estão associados, respetivamente, à gastroenterite e à febre tifoide. Quanto à *Aggregatibacter actinomycetemcomitans*, é uma das principais bactérias ligadas à etiologia da Periodontite. Desta forma, pode-se concluir que, apesar de serem microrganismos muito estudados pelas suas relações com patologias orais ou sistémicas, ainda há muito por esclarecer sobre a função específica do seu proteoma uma vez que não estão esclarecidos os mecanismos pelos quais algumas das proteínas têm acções patogénicas.

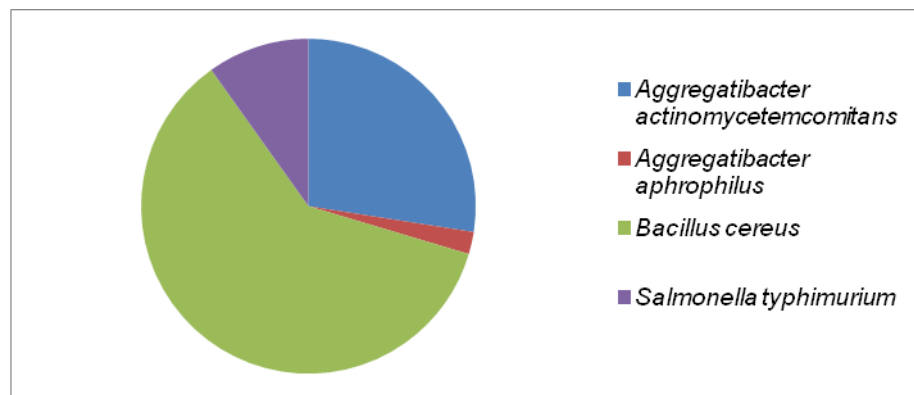
Quanto às espécies associadas às 53 proteínas envolvidas noutros processos além da patogénese, encontram-se, entre outros, a *Salmonella typhimurium*, o *Streptococcus mutans*, o *Streptococcus pyogenes* e o *Streptococcus gordonii* (Figura 19). O género mais representado é o *Streptococcus*, o que era previsível, pois para além de ser o género mais abundante na cavidade oral, está também associado à patologia oral mais prevalente na cavidade oral - a cárie dentária. Os papéis das proteínas deste género foram ainda explorados nos estudos de proteómica, provavelmente por

serem colonizadores primários e terem assim um papel importante na formação e no desenvolvimento do biofilme (72).

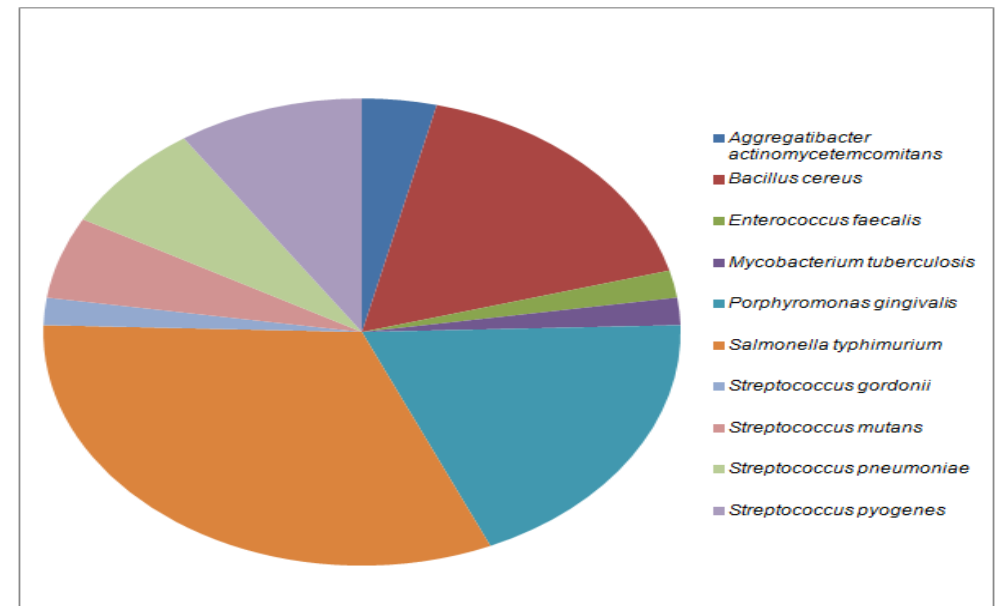
O cruzamento dos diagramas 2 e 3 da figura 19 revelam que o *Bacillus cereus* é o microrganismo que possui mais proteínas virulentas na cavidade oral. Isso pode ser explicado não por ser o microrganismo mais virulento da cavidade oral, mas sim por ser um dos mais estudados (provavelmente pela sua importância na indústria alimentar, uma vez que é uma das causas mais importante de intoxicação alimentar (78)).



**Diagrama 1** - Está representada a azul a quantidade de proteínas que estão no grupo “pathogenesis”, mas que também estão associadas aos outros processos representados no gráfico da **figura 18**. O vermelho representa a quantidade de proteínas do grupo “pathogenesis” que não estão associadas aos outros processos representados no gráfico da **figura 18**.



**Diagrama 2** - Espécies com maior representação nas proteínas dos grupos “pathogenesis” que não estão associadas aos outros processos biológicos, representados na **figura 18**. Para facilitar a interpretação e análise do gráfico, apenas estão representados os géneros específicos das bactérias e os géneros que possuíam um número de proteínas associadas superior a 10.



**Diagrama 3** - Espécies representadas nas proteínas do grupo “pathogenesis” que estão associadas aos outros processos biológicos, representados na **figura 18**,

**Figura 19** - Análise detalhada do grupo “pathogenesis” do gráfico da figura 17.

## 4.4 A função das proteínas bacterianas envolvidas na patogénese

As 298 proteínas brevemente classificadas na secção anterior (grupo que apresenta a ontologia “*biological process*” com a descrição “*pathogenesis*”) foram analisadas a nível funcional. Os resultados desta análise estão nas tabelas I e II em anexo, onde são apresentadas as 44 proteínas com evidência a nível da proteína e as 254 proteínas com grau de evidência a nível do gene.

Nestas, pode ser consultado o código Uniprot, o nome das proteínas, as bactérias que as produzem, os mecanismos específicos associados a patogénese, e, por fim, a referência bibliográfica que associa a proteína ao determinado mecanismo. Além disso, para facilitar a compreensão das tabelas, foram agrupadas as proteínas com as mesmas características, baseando-se esta junção na classificação referida na secção da introdução (tabelas em anexo).

As proteínas com grau de evidência, a nível de genes, são proteínas hipotéticas (HP) sobre as quais não existe qualquer prova da sua existência enquanto proteína. Uma proteína é considerada hipotética quando é prevista a partir das sequências de ácidos nucleicos e quando a sua existência não é comprovada através de uma evidência experimental. Optou-se, pois, por não excluir do estudo estas HPs, visto que este grupo é de extrema importância para completar a informação genómica e proteómica. A identificação génica de novas proteínas hipotéticas não só oferece pertinência para o seu isolamento e caracterização, como também elucida sobre novas e possíveis funções que possam ser realizadas por estas proteínas que ainda não foram identificadas (79).

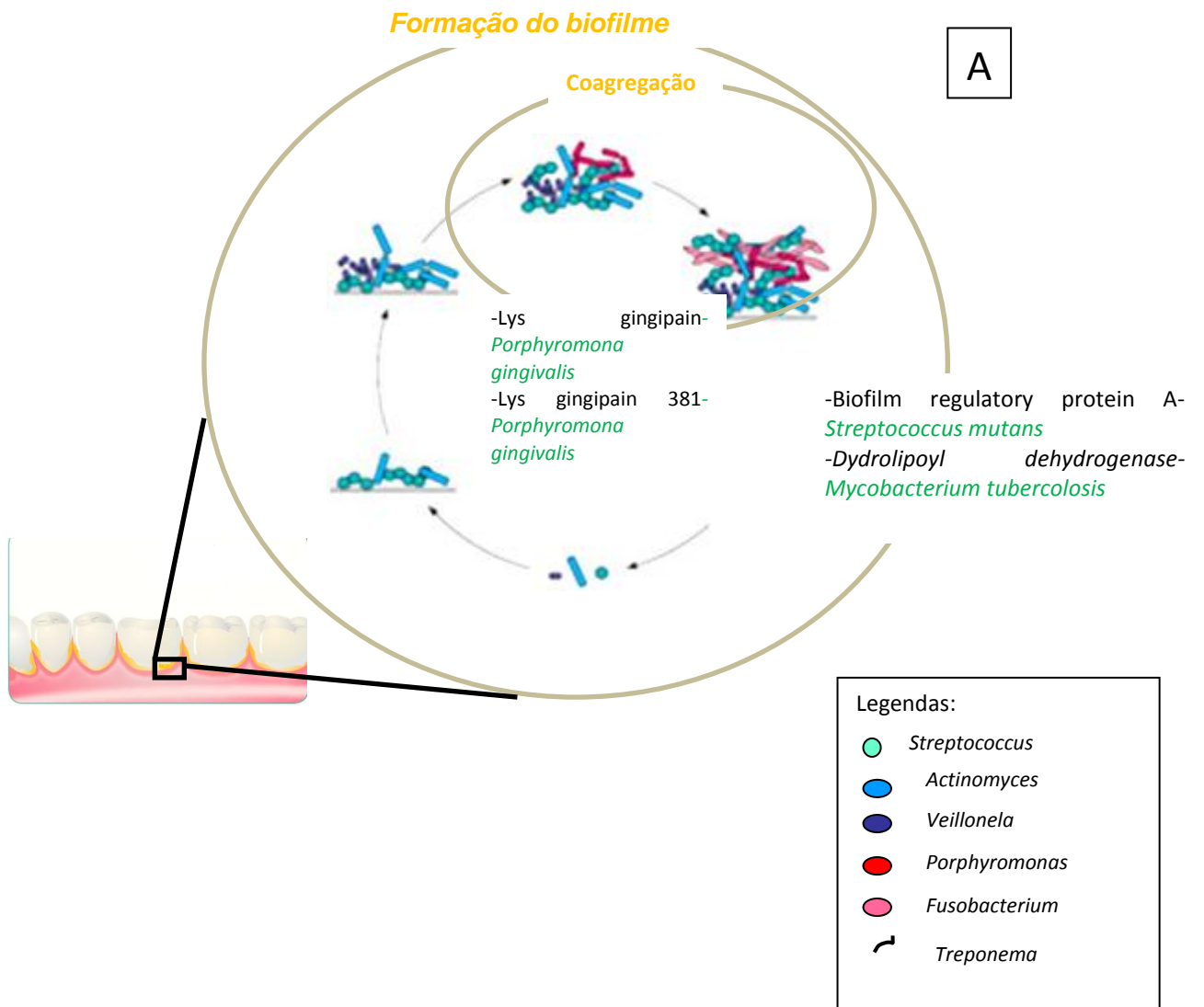
No microbioma oral, para as proteínas envolvidas no processo biológico da patogénese, existem quase cinco vezes mais HPs do que proteínas com evidências experimentais claras quanto à sua existência. Assim, torna-se importante que sejam utilizadas técnicas experimentais para comprovar a sua existência enquanto proteínas. Essa importância deriva de estas poderem ter um papel fundamental na associação com patologias orais.

Da análise das tabelas I e II (em anexo) ressaltam quatro aspectos fundamentais:

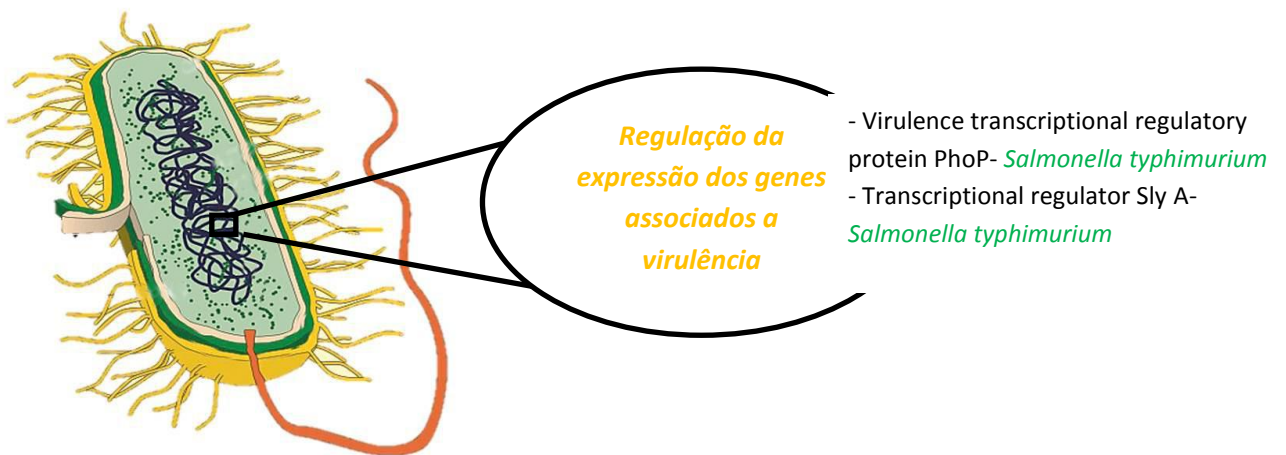
- 1) Atualmente, a maioria dos fatores de virulência descritos baseia-se em poucas espécies cuja forma de colonizar e de infectar o hospedeiro está bem estudada, como é o caso da *Salmonella* (tabela I e II). Já pelo contrário, para algumas proteínas com grau de evidência a nível do gene existe pouca informação sobre os mecanismos específicos dos fatores de virulência (tabela II).
- 2) Alguns dos fatores de virulência são conjuntos de proteínas que atuam de forma orquestrada para o efeito tóxico. As proteínas *Sip A*, *Sip B*, *Sip C* e *Sip D* de *Salmonella*, as enterotoxinas não hemolíticas A, B e C, bem como as hemolisinas BL (HBL) B, hBLL1 e hBLL2 são exemplos apresentados.
- 3) Pode ainda constatar-se que muitas das proteínas têm mais que uma forma de intervir na virulência da bactéria. Por exemplo as gingipains de *Porphyromonas gingivalis* encontram-se várias vezes na tabela I e II. Aparecem como fatores não específicos pois são proteases que a bactéria usa para se nutrir, mas que têm também o efeito de degradar a matriz extracelular dos tecidos do hospedeiro. Finalmente, ressalta a grande quantidade de proteínas associadas à sobrevivência intracelular das bactérias capazes de invadir as células do hospedeiro para as proteínas caracterizadas a nível da proteína. Quanto as proteínas caracterizadas a nível do gene, a maioria estão envolvidas na lise das células do hospedeiro pela formação de poros.
- 4) É de acrescentar que para algumas proteínas da tabela II (citadas na tabela III) não foi encontrada qualquer informação na literatura sobre os mecanismos específicos associados a patogénese com as palavras-chave definidas.

Estudos relativos à leucotoxina da *Aggregatibacter actinomycetemcomitans* revelam que uma estirpe desta natureza é altamente tóxica e capaz de produzir grandes quantidades de leucotoxina, enquanto outra estirpe é pouco ou minimamente tóxica e sua produção de leucotoxina é muito pequena. Além disso, a estirpe minimamente tóxica não faz a lise dos neutrófilos, ao contrário da estirpe que lisa os mesmos. Verifica-se assim que há estirpes de *Aggregatibacter actinomycetemcomitnas* mais virulentas do que outras (80). Este facto pode constituir, portanto, um limite a este trabalho, visto que no nosso estudo não foram consideradas as estirpes mas apenas as espécies. No entanto, identificar quais os géneros específicos e quais os fatores associados a patologias orais presentes é, de qualquer modo, um primeiro passo.

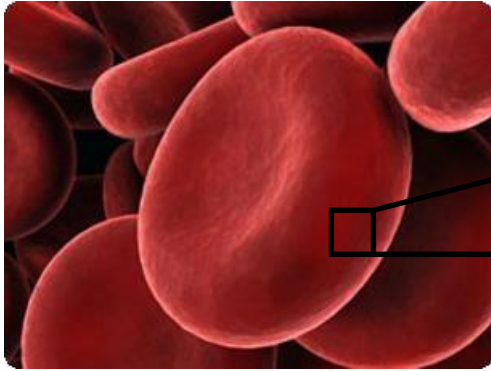
Para facilitar a compreensão das tabelas em anexo, a figura 20 apresenta sob a forma de esquema os mecanismos específicos associados à patogénese para as proteínas cuja evidência existe a nível da proteína.



B



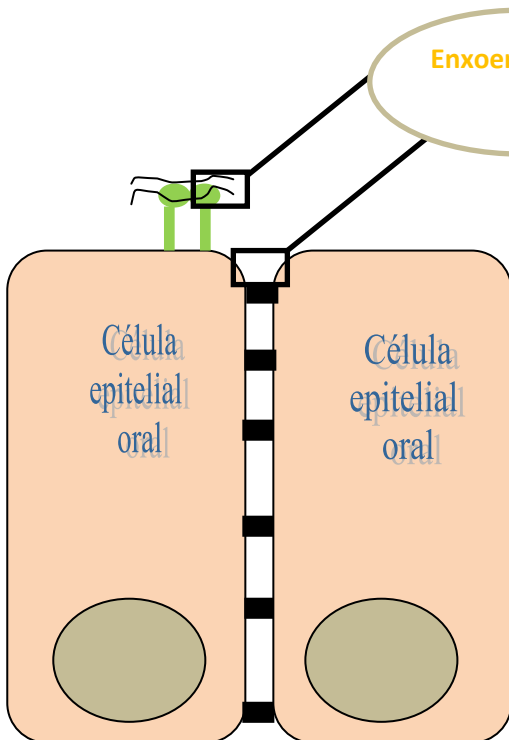
C



Captação de ferro por lise

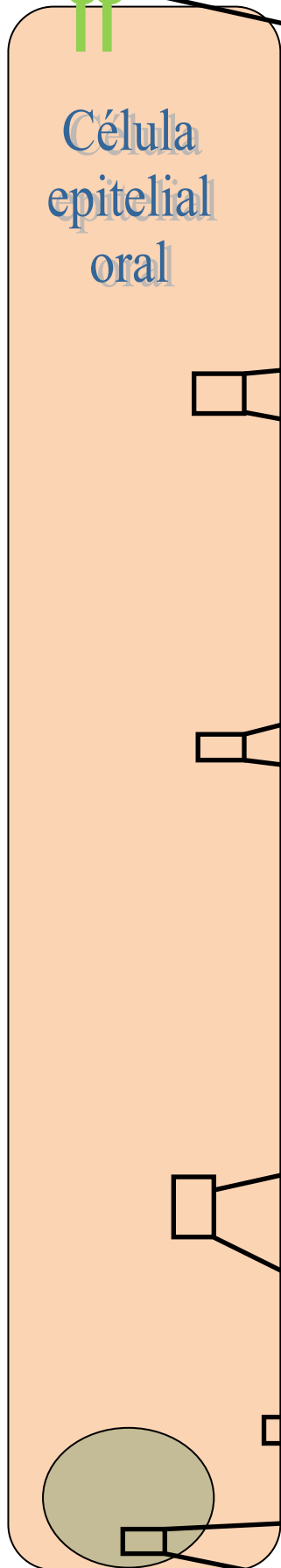
- Gingipain R1 e R2- *P.gingivalis*
- Lys-gingipain- *P.gingivalis*
- Lys-gingipain 381- *P.gingivalis*
- Lys-gingipain W83- *P.gingivalis*

D



Enxoenzimas

- Enolase-*Streptococcus pneumoniae*
- Enolase-*Streptococcus mutans*



**Adesão à matriz extracelular ou à célula**

- Outer membrane protein PagN- *S. typhimurium*
- Major fimbrial subunit protein type 1- *P.gingivalis*
- Lys-gingipain 381- *P.gingivalis* - Lys-gingipain W83- *P.gingivalis*
- Gingipain R1- *P.gingivalis*

**Invasão**

- Protein InvG- *S.typhimurium*
- Secreted effector protein Sop D- *S.typhimurium*
- Cell invasion protein Sip A, Sip B, Sip Ce Sip D-*S.typhimurium*
- Outer membrane protein PagN- *Salmonella typhimurium*
- Inositol phosphate phosphatase SopB- *S.typhimurium*
- Secreted effector kinase SteC- *S. typhimurium*
- Streptotain-Streptococcus pyogenes*

**Sobrevivência intracelular**

- Secreted effector protein Sif A- *S. typhimurium*
- Secreted effector protein pipB2- *Salmonella typhimurium*
- Inositol phosphate phosphatase SopB- *Salmonella typhimurium*
- Secreted effector protein Sop D e D2- *Salmonella typhimurium*
- Secreted effector protein SptP- *S. typhimurium*
- Secreted effector protein SseB- *S. typhimurium*
- Secreted effector protein SseI- *S. typhimurium*
- Secreted effector kinase SteC- *S. typhimurium*
- Protein PrgH- *S. typhimurium*
- Protein PrgI- *S. typhimurium*
- Virulence protein STM3117- *S. typhimurium*

**Interfere com a via de ubiquitinação**

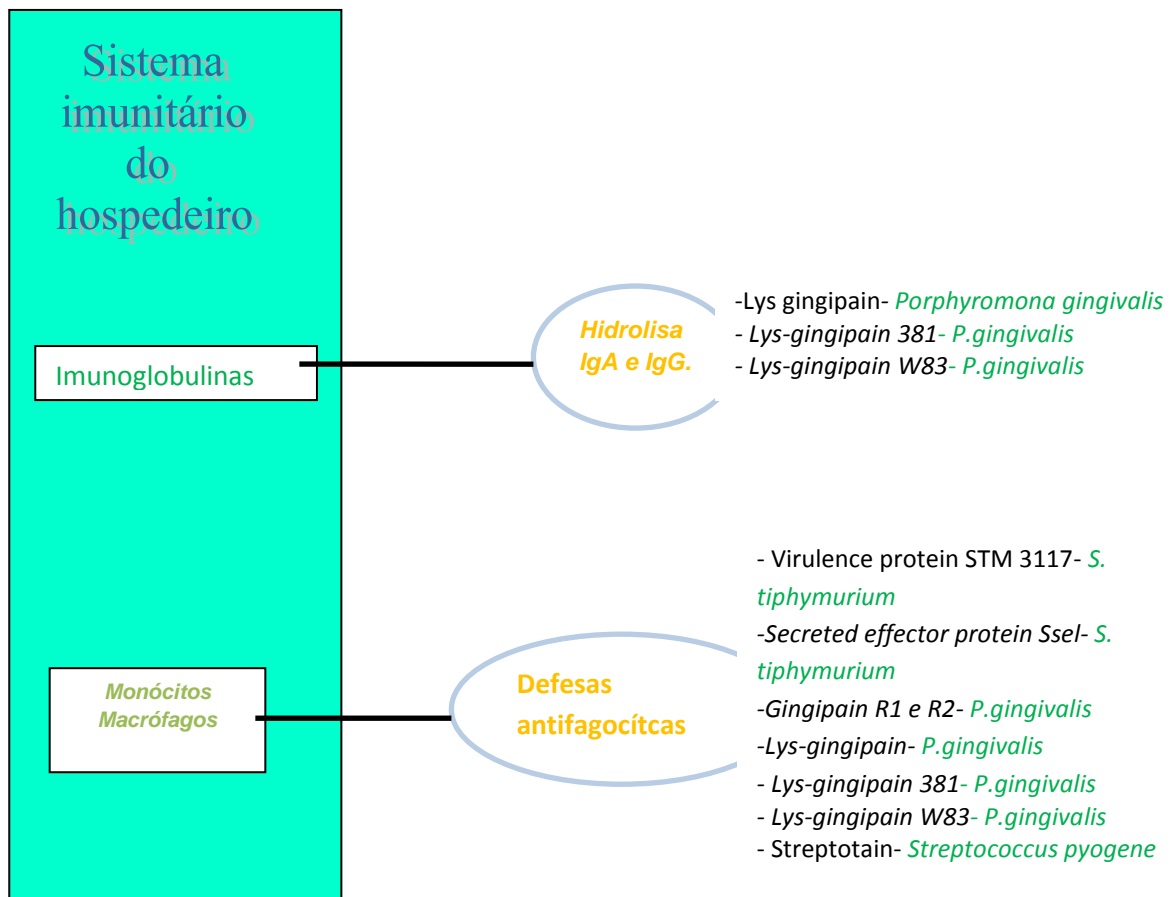
- E3 ubiquitin- protein ligase SlrP- *Salmonella typhimurium*
- E3 ubiquitin- protein ligase SspH2- *Salmonella typhimurium*

**Lise- Formação de poros**

- Non-hemolytic enterotoxina A- *Bacillus cereus*

**Apoptose**

- Cytolethal distending protein A, B e C- *Aggregatibacter actinomycetemcomitans*



**Figura 20** - Mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível da proteína: (A) proteínas envolvidas na formação do biofilme, (B) proteínas que regulam a expressão de genes associados à virulência, (C) proteínas que promovem a lise dos eritrócitos para captação de ferro necessário para a nutrição da *P. gingivalis*, (D) proteínas que promovem a transmigração da bactéria nos tecidos do hospedeiro, (E) proteínas com ação ao nível da célula epitelial oral, e (F) proteínas que atuam ao nível do sistema imunitário do hospedeiro.

A cor de círculo corresponde ao tipo de mecanismo: o azul corresponde a mecanismos ofensivos, o roxo a mecanismos defensivos, o cinzento a mecanismos não específicos e o preto a regulação dos genes de virulência.

## 4.5 Associação entre os mecanismos moleculares da patogénese e as principais patologias da cavidade oral

A tabela 1 apresenta os microrganismos que produzem as proteínas estudadas e a patologia oral ao qual já foi associado. Nesta tabela verifica-se que a maioria dos microrganismos incluídos estão associados a doença periodontal, uma das patologias orais em que os estudos de proteómica são mais abundantes.

**Tabela 1** - Bactérias produtoras das proteínas analisadas neste trabalho e associação com patologias da cavidade oral.

<b>Espécies bacterianas</b>	<b>Patologias orais associadas</b>
<i>Aggregatibacter actinomycetemcomitans</i>	Periodontite crónica/Periodontite agressiva (81, 82)
<i>Aggregatibacter aphrophilus</i>	Não está associada a patologias orais
<i>Bacillus cereus</i>	Não está associada a patologias orais
<i>Lactobacillus acidophilus</i>	Cárie dentária (83)
<i>Mycobacterium tuberculosis</i>	Neoplasma da cavidade oral (84)
<i>Porphyromona gingivalis</i>	Gengivite/Periodontite crónica/Periodontite agressiva (85, 86)
<i>Salmonella typhimurium</i>	Não está associada a patologias orais
<i>Stafilococcus aureus</i>	Não está associada a patologias orais
<i>Streptococcus gordonii</i>	Cárie dentária (77)
<i>Streptococcus mutans</i>	Cárie dentária (77)
<i>Streptococcus pneumoniae</i>	Não está associada a patologias orais
<i>Streptococcus pyogenes</i>	Periodontite crónica (85)

O microrganismo *Bacillus cereus* não faz parte da microflora oral saudável, nem existe bibliografia relativamente à sua associação com patologias orais. No entanto, sabe-se que está presente em alimentos, principalmente cereais, daí resultando provavelmente a sua presença na cavidade oral. A bactéria *Bacillus cereus* produz as toxinas tripartidas “*non-hemolytic enterotoxin*” e “*hemolysin BL*” assim como as toxinas “*citotoxin K*”, “*perfringolysin*” e “*hemolysin II*”. Todas estas toxinas intervêm no mecanismo patogénico através da formação de poros nas células epiteliais do hospedeiro. De forma idêntica, as bactérias *Salmonella typhimurium*, *Streptococcus pneumoniae*, *Aggregatibacter aphrophilus*, e *Staphylococcus aureus* não estão associadas com patologias orais, no entanto possuem proteínas envolvidas na patogénese. Estas bactérias estão também muito bem estudadas a nível dos seus mecanismos de invasão nas células epiteliais do intestino humano ou do sistema respiratório (87-93), e, hipoteticamente poderão expressar os mesmos fatores de virulência quando colonizam a cavidade oral.

A identificação de novos determinantes de virulência é uma chave para compreender a interação das bactérias patogénicas com os seus hospedeiros, de modo a produzir a doença clínica (94).

No microbioma oral, o conhecimento dos processos patológicos em que as proteínas de origem bacteriana presentes podem participar, fornece informações valiosas, as quais podem ser utilizadas na prevenção de infeções orais e no tratamento das principais patologias associadas à cavidade oral. Um aspeto que deve ser abordado para o desenvolvimento de vacinas ou de terapêutica contra doenças orais é que estas são polimicrobianas. Esta multiplicidade de microrganismos patogénicos indica que o desenho da vacina ou fármaco irá ser muito complexo.

É de relevar ainda que para as 298 proteínas envolvidas no processo biológico da patogénese, a grande maioria foram identificadas *in vitro*. Apenas 11 foram identificadas em amostras de pacientes (“*virulence transcriptional regulatory protein PhoP*” (Q2FH47), “*dehydrolipoyl dehydrogenase*” (P66004), “*enolase*”, “*hemagglutinin family protein (fragment)*” (D3A1B7 e D3A8N8), “*enolase*” (Q97QS2), “*exotoxin 15*” (C8MGI5), “*RTX-toxin family pore-forming*

*toxin*" (C2DNR2), "putative toxin B" (B3WLE9), "hep/hag repeat protein" (C4FMC5, C4FMC7) e "surface protein" (C2HR32)). Seria assim oportuno, em trabalhos futuros, identificar estas proteínas em amostras salivares de pacientes com as patologias orais para as quais os microrganismos responsáveis pela sua produção estão envolvidos (tabela 1).

Como foi referido na secção da introdução, a base de dados de referência para fatores de virulência bacteriana *Virulence Factors of Pathogenic Bacteria* (VFDB), criada em 2004 (<http://www.mgc.ac.cn/VFs/>), tem servido como um repositório global dos fatores de virulência de bactérias. Este banco de dados contém informações para 25 géneros bacterianos representativos dos patógenos mais importantes: os genes associados à virulência e seus mecanismos associados, com as respetivas referências bibliográficas (38, 40). Dos 25 géneros bacterianos representados nesta base de dados não estão incluídas algumas das principais espécies envolvidas nas patologias orais, como por exemplo *Agregatibacter* ou *Porphyromonas*. Através das tabelas criadas neste trabalho, é possível completar as informações desta base de dados relativamente aos fatores de virulência de microrganismos da cavidade oral, e assim ajudar na pesquisa para a prevenção e tratamento de patologias orais.

## **5.CONCLUSÃO**



Através deste trabalho chegou-se às seguintes conclusões:

1. Das 11428 proteínas bacterianas estudadas e que integram o microbioma oral:
  - a. A maioria (7659) participa em processos de “*housekeeping*”.
  - b. Apenas 298 estão implicadas em mecanismos de patogênese.
2. Das 298 proteínas associadas a patogênese, 44 proteínas estão caracterizadas a nível da proteína e 254 apenas a nível do gene. Assim, na cavidade oral existem cinco vezes mais proteínas com evidência a nível do gene. Torna-se desta forma, importante comprovar a existência destas proteínas por métodos experimentais.
3. Das proteínas com o GO “*pathogenesis*” depositadas no OralCard, os mecanismos específicos com mais número de proteínas associadas são o de sobrevivência do microrganismo dentro da célula do hospedeiro (para as proteínas caracterizadas a nível da proteína) e de lise através da formação de poros nas células do hospedeiro (para as proteínas caracterizadas a nível do gene).
4. A maioria das proteínas bacterianas envolvidas no processo biológico da patogênese foi descrita em estudos *in vitro* e não em amostras da cavidade oral. Seria assim pertinente, em trabalhos futuros, comprovar a existência dessas em amostras salivares de pacientes com as patologias orais para as quais os microrganismos responsáveis pela sua produção estão envolvidos.
5. Algumas das proteínas apresentadas pertencem a microrganismos que não estão associadas a patologias orais, no entanto, possuem proteínas bem estudadas que intervêm no processo da patogênese noutros locais (por exemplo o intestino) e que aparentemente têm muitas semelhanças com proteínas produzidas por microrganismos da cavidade oral. Assim, será interessante explorar as funções de proteínas homólogas

produzidas pelo microbioma oral de forma a compreender os mecanismos patofisiológicos que decorrem no biofilme oral.

6. Em termos ontológicos, notámos alguns limites a nível do vocabulário disponível para definir fatores de virulência; por exemplo as proteínas associadas ao processo biológico da patogénese não têm *child terms* associados. Sugere-se assim que os seguintes termos, além de serem *child terms* dos processos aos quais estão atualmente associados, sejam também associados ao termo ontológico "*pathogenesis*": "*cell adesion*", "*killing of cells of other organism*", "*celular protein modification proces*s", "*peptidase activity*", "*lyase activity*", "*cytoskeletal protein binding*", "*lipid binding*", "*ion binding*", e, "*transferase activity - transferring glycosyl groups*".

## **6. BIBLIOGRAFIA**



1. Newman MG, Takei H, Klokkevold PR, Carranza FA. Carranza's clinical periodontology. Elsevier health sciences; 2011. p. 100-275.
2. Avila M, Ojcius DM, Yilmaz O. The oral microbiota: living with a permanent guest. *DNA Cell Biol.* 2009;28(8):405-11.
3. Chardin H, Barsotti O, Bonnaure-Mallet M. Microbiologie en odontostomatologie. Maloine; 2006. p. 141-64.
4. Rosa N, Correia MJ, Arrais JP, Lopes P, Melo J, Oliveira JL, et al. From the salivary proteome to the OralOme: comprehensive molecular oral biology. *Arch Oral Biol.* 2012;57(7):853-64.
5. Do Proteoma Salivar ao OralOma: Tese apresentada à Universidade Católica Portuguesa para obtenção do grau de doutor em Ciências da saúde; 2011.
6. N R. Do Proteoma Salivar ao OralOma: Tese apresentada à Universidade Católica Portuguesa para obtenção do grau de doutor em Ciências da saúde; 2011.
7. Arrais JP, Rosa N, Melo J, Coelho ED, Amaral D, Correia MJ, et al. OralCard: a bioinformatic tool for the study of oral proteome. *Arch Oral Biol.* 2013;58(7):762-72.
8. Jenkinson HF, Lamont RJ. Oral microbial communities in sickness and in health. *Trends Microbiol.* 2005;13(12):589-95.
9. Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. *Nature Rev Genet.* 2012;13(4):260-70.
10. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Ligget C, Knight R, J. G. The human microbiome project: exploring the microbial part of ourselves in changing world. *Nature.* 2007;449(7164):804-10.
11. Consortium HMP. A framework for human microbiome research. *Nature.* 2012;486(7402):215-21.
12. Murray PR. *Microbiologia medica.* Guanabara Koogan; 2000. p. 72-5.
13. Salyers AA, Whitt DD. *Microbiology: diversity, disease, and the environment.* Fitzgerald Science Press, Inc; 2001. p. 289-314.
14. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, et al. Microbial co-occurrence relationships in the human microbiome. *PLoS Computat Biol.* 2012;8(7):e1002606.
15. Weber TK, Polanco I. Gastrointestinal microbiota and some children diseases: a review. *Gastrol Res Pract.* 2012;2012:676585.
16. Gevers D, Pop M, Schloss PD, Huttenhower C. Bioinformatics for the Human Microbiome Project. *PLoS Comput Biol.* 2012;8(11):e1002779.
17. Grice EA, Segre JA. The human microbiome: our second genome. *Annu Rev Genomics Human Genet.* 2012;13:151-70.
18. Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, et al. The NIH Human Microbiome Project. *Genome Res.* 2009;19(12):2317-23.

19. HMP C. Structure, function and diversity of the healthy human microbiome. *Nature*. 2012;486(7402):207-14.
20. Nelson KE, Weinstock GM, Highlander SK, Worley KC, Creasy HH, Wortman JR, et al. A catalog of reference genomes from the human microbiome. *Science*. 2010;328(5981):994-9.
21. Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, et al. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut*. 2006;55(2):205-11.
22. Zainal-Abidin Z, Veith PD, Dashper SG, Zhu Y, Catmull DV, Chen YY, et al. Differential proteomic analysis of a polymicrobial biofilm. *J Proteom Res*. 2012;11(9):4449-64.
23. Lamont RJ, Jenkinson HF. *Oral microbiology at a glance*. John Wiley & Sons; 2010. p. 29-54.
24. Sbordone L, Bortolaia C. Oral microbial biofilms and plaque-related diseases: microbial communities and their role in the shift from oral health to disease. *Clinical Oral Investig*. 2003;7(4):181-8.
25. Jakubovics NS, Kolenbrander PE. The road to ruin: the formation of disease-associated oral biofilms. *Oral Dis*. 2010;16(8):729-39.
26. Janoff EN, Gustafson C, Frank DN. The world within: living with our microbial guests and guides. *Transl Res*. 2012;160(4):239-45.
27. Chen T, Yu WH, Izard J, Baranova OV, Lakshmanan A, Dewhirst FE. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database (Oxford)*. 2010;2010:baq013.
28. Grover HS, Kapoor S, Saksena N. Periodontal proteomics: wonders never cease! *Int J Proteomics*. 2013;2013:850235.
29. Marsh P. Are dental diseases examples of ecological catastrophes? *Microbiology*. 2003;149(2):279-94.
30. Martinko JM, Madigan M. *Brock biology of microorganisms*. Englewood Cliffs, NJ: Prentice Hall; 2005. p. 773-800.
31. Samaranayake LP. *Essential microbiology for dentistry*: Elsevier Health Sciences; 2006.
32. Asad S, Opal SM. Bench-to-bedside review: Quorum sensing and the role of cell-to-cell communication during invasive bacterial infection. *Crit Care*. 2008;12(6):236.
33. Miller MB, Bassler BL. Quorum sensing in bacteria. *Annu Rev Microbiol*. 2001;55:165-99.
34. Williams P, Camara M, Hardman A, Swift S, Milton D, Hope VJ, et al. Quorum sensing and the population-dependent control of virulence. *Philos Trans R Soc Lond B Biol Sci*. 2000;355(1397):667-80.
35. Taga ME, Bassler BL. Chemical communication among bacteria. *Proc Natl Acad Sci USA*. 2003;100 Suppl 2:14549-54.
36. Marsh PD. Dental plaque as a biofilm and a microbial community—implications for health and disease. *BMC Oral Health*. 2006;6(Suppl 1):S14.

37. Korves T, Colosimo ME. Controlled vocabularies for microbial virulence factors. *Trends Microbiol.* 2009;17(7):279-85.
38. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, et al. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res.* 2005;33(suppl 1):D325-D8.
39. Finlay BB, Falkow S. Common themes in microbial pathogenicity revisited. *Microbiol Mol Biol Rev.* 1997;61(2):136-69.
40. Chen L, Xiong Z, Sun L, Yang J, Jin Q. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic Acids Res.* 2012;40(Database issue):D641-5.
41. Wu HJ, Wang AH, Jennings MP. Discovery of virulence factors of pathogenic bacteria. *Curr Opin Chem Biol.* 2008;12(1):93-101.
42. Allen RC, Popat R, Diggle SP, Brown SP. Targeting virulence: can we make evolution-proof drugs? *Nature Rev Microbiol.* 2014;12(4):300-8.
43. Wilson J, Schurr M, LeBlanc C, Ramamurthy R, Buchanan K, Nickerson C. Mechanisms of bacterial pathogenicity. *Postgrad Med J.* 2002;78(918):216-24.
44. Henderson B, Wilson M, McNab R, Lax AJ. *Bacteria-Host Interaction in Health and Disease.* Wiley; 1999. p. 163-310.
45. Alonso A, Garcia-del Portillo F. Hijacking of eukaryotic functions by intracellular bacterial pathogens. *International Microbiol.* 2004;7(3):181-91.
46. Bamburg JR. *Listeria monocytogenes* cell invasion: a new role for cofilin in co-ordinating actin dynamics and membrane lipids. *Mol Microbiol.* 2011;81(4):851-4.
47. Ham H, Sreelatha A, Orth K. Manipulation of host membranes by bacterial effectors. *Nature Rev Microbiol.* 2011;9(9):635-46.
48. Croinin TO, Backert S. Host epithelial cell invasion by *Campylobacter jejuni*: trigger or zipper mechanism? *Front Cell Infect Microbiol.* 2012;2:25.
49. Ray K, Marteyn B, Sansonetti PJ, Tang CM. Life on the inside: the intracellular lifestyle of cytosolic bacteria. *Nature Rev Microbiol.* 2009;7(5):333-40.
50. Garai P, Gnanadhas DP, Chakravorty D. *Salmonella enterica* serovars Typhimurium and Typhi as model organisms: revealing paradigm of host-pathogen interactions. *Virulence.* 2012;3(4):377-88.
51. Perry J, Staley J, Lory S. *Microbiologie.* Éditions Dunod; 2004. p. 648-65.
52. Potempa J, Pike RN. Corruption of innate immunity by bacterial proteases. *J Innate Immun.* 2008;1(2):70-87.
53. Tokuda M, Duncan M, Cho MI, Kuramitsu HK. Role of *Porphyromonas gingivalis* protease activity in colonization of oral surfaces. *Infect Immun.* 1996;64(10):4067-73.
54. Lantz MS. Are bacterial proteases important virulence factors? *J Periodontal Res.* 1997;32(1 Pt 2):126-32.
55. Casadevall A, Pirofski L-a. The damage-response framework of microbial pathogenesis. *Nature Rev Microbiol.* 2003;1(1):17-24.

56. Rutherford ST, Bassler BL. Bacterial quorum sensing: its role in virulence and possibilities for its control. *Cold Spring Harb Perspect Med.* 2012;2(11).
57. Xiong J. *Essential bioinformatics*: Cambridge University Press; 2006.
58. Consortium U. The Universal Protein Resource (UniProt) 2009. *Nucleic Acids Res.* 2009;37(Database issue):D169-74.
59. Consortium U. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* 2014;42(Database issue):D191-8.
60. Schneider M, Lane L, Boutet E, Lieberherr D, Tognolli M, Bougueleret L, et al. The UniProtKB/Swiss-Prot knowledgebase and its Plant Proteome Annotation Program. *J Proteomics.* 2009;72(3):567-73.
61. Consortium GO. Creating the gene ontology resource: design and implementation. *Genome Res.* 2001;11(8):1425-33.
62. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genet.* 2000;25(1):25-9.
63. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, et al. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* 2004;32(Database issue):D258-61.
64. Thomas PD, Mi H, Lewis S. Ontology annotation: mapping genomic regions to biological function. *Curr Opin Chem Biol.* 2007;11(1):4-11.
65. Hill DP, Smith B, McAndrews-Hill MS, Blake JA. Gene Ontology annotations: what they mean and where they come from. *BMC Bioinformatics.* 2008;9(Suppl 5):S2.
66. Torto-Alalibo T, Collmer CW, Gwinn-Giglio M, Lindeberg M, Meng S, Chibucos MC, et al. Unifying themes in microbial associations with animal and plant hosts described using the gene ontology. *Microbiol Mol Biol Rev.* 2010;74(4):479-503.
67. Giglio MG, Collmer CW, Lomax J, Ireland A. Applying the Gene Ontology in microbial annotation. *Trends Microbiol.* 2009;17(7):262-8.
68. McCarthy FM, Gresham CR, Buza TJ, Chouvarine P, Pillai LR, Kumar R, et al. AgBase: supporting functional modeling in agricultural organisms. *Nucleic Acids Res.* 2011;39(Database issue):D497-506.
69. Bhatia VN, Perlman DH, Costello CE, McComb ME. Software tool for researching annotations of proteins: open-source protein annotation software with data visualization. *Anal Chem.* 2009;81(23):9819-23.
70. McCarthy FM, Wang N, Magee GB, Nanduri B, Lawrence ML, Camon EB, et al. AgBase: a functional genomics resource for agriculture. *BMC Genomics.* 2006;7(1):229.
71. I S. OralOme - O Contributo dos Microrganismos Dissertação apresentada à Universidade Católica Portuguesa para obtenção do grau de Mestre em Medicina Dentária Julho 2012.
72. M P. OralOme - O contributo dos microrganismos revelado por estudos de proteómica dirigida: Dissertação apresentada à Universidade Católica

Portuguesa para obtenção do grau de Mestre em Medicina Dentária; Julho 2013.

73. Ajdic D, McShan WM, McLaughlin RE, Savic G, Chang J, Carson MB, et al. Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. *Proc Natl Acad Sci USA*. 2002;99(22):14434-9.

74. Watanabe T, Maruyama F, Nozawa T, Aoki A, Okano S, Shibata Y, et al. Complete genome sequence of the bacterium *Porphyromonas gingivalis* TDC60, which causes periodontal disease. *J Bacteriol*. 2011;193(16):4259-60.

75. Solheim M, Aakra Å, Snipen LG, Brede DA, Nes IF. Comparative genomics of *Enterococcus faecalis* from healthy Norwegian infants. *BMC Genomics*. 2009;10(1):194.

76. Takeno A, Okamoto A, Tori K, Oshima K, Hirakawa H, Toh H, et al. Complete genome sequence of *Bacillus cereus* NC7401, which produces high levels of the emetic toxin cereulide. *J Bacteriol*. 2012;194(17):4767-8.

77. van der Ploeg JR. Characterization of *Streptococcus gordonii* prophage PH15: complete genome sequence and functional analysis of phage-encoded integrase and endolysin. *Microbiology*. 2008;154(10):2970-8.

78. Cliver DO, Potter M, Riemann HP. *Foodborne infections and intoxications*. Academic Press; 2011.p.129-64

79. Lubec G, Afjehi-Sadat L, Yang JW, John JP. Searching for hypothetical proteins: theory and practice based upon original data and literature. *Prog Neurobiol*. 2005;77(1-2):90-127.

80. Newman MG, Takei H, Klokkevold PR, Carranza FA. *Clinical Periodontology*. 10th Edition: Elsevier, Saunders; 2006. p. 100-275.

81. Cortelli J, Aquino D, Cortelli S, Roman-Torres C, Franco G, Gomez R, et al. *Aggregatibacter actinomycetemcomitans* serotypes infections and periodontal conditions: a two-way assessment. *Eur J Clin Microbiol Infect Dis*. 2012;31(7):1311-8.

82. Cortelli JR, Roman-Torres CV, Aquino DR, Franco GC, Costa FO, Cortelli SC. Occurrence of *Aggregatibacter actinomycetemcomitans* in Brazilians with chronic periodontitis. *Braz Oral Res*. 2010;24(2):217-23.

83. Kleinberg I. A mixed-bacteria ecological approach to understanding the role of the oral bacteria in dental caries causation: an alternative to *Streptococcus mutans* and the specific-plaque hypothesis. *Crit Rev Oral Biol Med*. 2002;13(2):108-25.

84. Xie H, Onsongo G, Popko J, de Jong EP, Cao J, Carlis JV, et al. Proteomics analysis of cells in whole saliva from oral cancer patients via value-added three-dimensional peptide fractionation and tandem mass spectrometry. *Mol Cell Proteomics*. 2008;7(3):486-98.

85. Baliban RC, Sakellari D, Li Z, DiMaggio PA, Garcia BA, Floudas CA. Novel protein identification methods for biomarker discovery via a proteomic analysis of periodontally healthy and diseased gingival crevicular fluid samples. *J Clin Periodontol*. 2012;39(3):203-12.

86. Bostanci N, Ramberg P, Wahlander A, Grossman J, Jonsson D, Barnes VM, et al. Label-free quantitative proteomics reveals differentially regulated proteins in experimental gingivitis. *J Proteom Res*. 2013;12(2):657-78.
87. McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, et al. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature*. 2001;413(6858):852-6.
88. Ganash M, Phung D, Sedelnikova SE, Lindback T, Granum PE, Artymiuk PJ. Structure of the NheA component of the Nhe toxin from *Bacillus cereus*: implications for function. *PLoS one*. 2013;8(9):e74748.
89. Tran SL, Guillemet E, Ngo-Camus M, Clybourn C, Puhar A, Moris A, et al. Haemolysin II is a *Bacillus cereus* virulence factor that induces apoptosis of macrophages. *Cellular Microbiol*. 2011;13(1):92-108.
90. Hardy SP, Lund T, Granum PE. CytK toxin of *Bacillus cereus* forms pores in planar lipid bilayers and is cytotoxic to intestinal epithelia. *FEMS Microbiol Lett*. 2001;197(1):47-51.
91. Senesi S, Ghelardi E. Production, secretion and biological activity of *Bacillus cereus* enterotoxins. *Toxins*. 2010;2(7):1690-703.
92. Bien J, Sokolova O, Bozko P. Characterization of Virulence Factors of *Staphylococcus aureus*: Novel Function of Known Virulence Factors That Are Implicated in Activation of Airway Epithelial Proinflammatory Response. *J Pathog*. 2011;2011:601905.
93. Hoskins J, Alborn WE, Jr., Arnold J, Blaszczyk LC, Burgett S, DeHoff BS, et al. Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J Bacteriol*. 2001;183(19):5709-17.
94. Tsai C-T, Huang W-L, Ho S-J, Shu L-S, Ho S-Y. Virulent-GO: prediction of virulent proteins in bacterial pathogens utilizing gene ontology terms. *Int J Biol Life Sci*. 2009;1(2):3.

## **7.ANEXOS**



São disponibilizadas, como anexos, as tabelas que apresentam os mecanismos de virulência associados às proteínas do microbioma oral presentes na base de dados OralCard. Os procedimentos usados para a realização destas tabelas são descritos na secção de Material e Métodos.



**Tabela I** - Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível da proteína. Toda a informação relativa aos mecanismos específicos foi retirada da base de dados Uniprot à exceção das proteínas assinaladas com\* cuja informação foi retirada da referência indicada.

		Uniprot	Nome da proteína	Bactéria	Mecanismo específico
Mecanismos ofensivos	Adesão	B2RH54	Major fimbrial subunit protein type-1 (Fimbrillin) (Fimbrilin) (Major fimbrial subunit protein type I)	<i>Porphyromonas gingivalis</i>	Medeia a adesão celular a células epiteliais.
		Q8ZRJ9	Outer membrane protein PagN (Adhesin/invasion) protein PagN)	<i>Salmonella typhimurium</i>	Hemaglutinina que facilita a adesão e invsão das células epiteliais usando o proteoglicano heparinado como recetor.
		P28784*	Gingipain R1 (EC 3.4.22.37) (Arg-gingipain)(Gingipain1) (RGP-1)	<i>Porphyromona gingivalis</i>	Promove a adesão através de domínios de adesão específicos. *Pubmed: PMID 8999833
		P72194*	Lys-gingipain 381 (EC 3.4.22.47)	<i>Porphyromonas gingivalis</i>	
		Q51817*	Lys-gingipain W83	<i>Porphyromonas gingivalis</i>	
	Mobili dade	Q8ZP57	Secreted effector kinase SteC (EC 2.7.-.-) (Salmonella translocated effector C)	<i>Salmonella typhimurium</i>	Cinase necessária à criação de feixes de filamentos de actina.
	Invasão celular do hospedeiro	E1WAC6	Cell invasion protein SipA (Effector protein sipA)	<i>Salmonella typhimurium</i>	Estimula a polimerização da actina e com SipC é necessária para a internalização da bactéria.
		P0CL52	Cell invasion protein SipB (Effector protein SipB)	<i>Salmonella typhimurium</i>	Necessária para a transferência e/ou apresentação de SipC e outras proteínas bacterianas para a célula do hospedeiro.
		P0CL47	Cell invasion protein SipC (Effector protein SipC)	<i>Salmonella typhimurium</i>	Promove a polimerização e condensação dos filamentos de actina (bundeling). É potenciada pela SipA.
		Q56026	Cell invasion protein SipD ( <i>Salmonella</i> invasion protein D)	<i>Salmonella typhimurium</i>	Provavelmente regula a secreção de SipA, B e C e é necessária à translocação de proteínas efetoras do sistema de secreção tipo III. Essencial à invasão.
O30916		Inositol phosphate phosphatase SopB (EC 3.1.3.-) (Effector protein SopB)	<i>Salmonella typhimurium</i>	Um dos efetores do sistema de injeção da <i>Salmonella</i> . Alteração da composição em fosfoinosítídeo da membrana promovendo “ruffling”.	
P40722		Secreted effector protein SopD (Salmonella outer protein D)	<i>Salmonella typhimurium</i>	Com SopB contribui para a fissão da membrana e formação de um macro pinosoma.	
P0COJ1		Streptopain (EC3.4.22.10) (Exotoxin typeB) (SPEB) (Streptococcal cysteine proteinase) (Streptococcus peptidase A) (SPP).	<i>Streptococcus pyogenes</i> serotype M1	Cliva fibronectina e degrada vitronectina, ativa o precursor de IL1B. Pode induzir a apoptose em monócitos e células epiteliais <i>in vitro</i> . Reduz a atividade fagocítica dos monócitos.	
Q8ZRJ9		Outer membrane protein PagN	<i>Salmonella typhimurium</i>	Hemaglutinina que facilita a invasão das células epiteliais usando o proteoglicano heparinado como recetor.	
P35672		Protein InvG	<i>Salmonella typhimurium</i>	Envolvida na invasão de células do epitélio intestinal	

**Tabela I (continuação).** Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível da proteína. Toda a informação relativa aos mecanismos específicos foi retirada da base de dados Uniprot à exceção das proteínas assinaladas com\* cuja informação foi retirada da referência indicada.

Mecanismos defensivos	Lise Celular	Uniprot	Nome da proteína	Bactéria	Mecanismo específico
		Q3Y6N6*	NheA	<i>Bacillus cereus</i>	Provoca lise de hemácias. Atua em conjunto com NheB e NheC para formar poros na membrana das hemácias. *Pubmed:PMID 24040335
Apoptose	O87120*	Cytolethal distending toxin protein A	<i>Aggregatibacter actinomycetemcomitans</i>	Inibem a divisão celular causando a apoptose das células do hospedeiro. * Pubmed: PMID 22069704 PMID 17123907	
	Q7DK12*	Cytolethal distending toxin protein B			
	Q7DK11*	Cytolethal distending toxin protein C			
Defesas asntifagocíticas	POC0J1	Streptopain (EC 3.4.22.10) (Exotoxin type B) (SPE B) (Streptococcal cysteine proteinase) (Streptococcus peptidase A) (SPP)	<i>Streptococcus pyogenes serotype M1</i>	Reduz a atividade fagocítica das células monocíticas.	
	Q8ZQ79	Secreted effector protein Ssel	<i>Salmonella typhimurium</i>	Inibe a migração dos macrófagos e células dentriticas interagindo com o fator IQGAP1, um importante regulador do citoesqueleto e migração celular.	
	P28784*	Gingipain R1 (EC 3.4.22.37) (Arg-gingipain) (Gingipain1) (RGP-1)	<i>Porphyromonas gingivalis</i>	Altera a função dos fagocitos clivando os recetores de formil-metionil-leucil-fenilalanina (FMLP). Inactiva os neutrófilos tornando-os incapazes de reconhecer os agentes patogénicos * Pubmed: PMID 15001226 PMID 24466435	
	P95493*	Gingipain R2 (EC 3.4.22.37) (Arg-gingipain) (Gingipain 2) (RGP-2)	<i>Porphyromonas gingivalis</i>		
	B2RLK2*	Lys-gingipain (EC 3.4.22.47) (Lysine-specific cysteine proteinase Kgp)	<i>Porphyromonas gingivalis</i>		
	P72194*	Lys-gingipain 381 (EC 3.4.22.47)	<i>Porphyromonas gingivalis</i>		
	Q51817*	Lys-gingipain W83	<i>Porphyromonas gingivalis</i>		
			Degrada o recetor C5a na superfície das células fagociticas resultando na supressão da migração dos neutrófilos para o local da infeção. * Pubmed: PMID 15001226 PMID 24466435		

**Tabela I (continuação).** Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível da proteína. Toda a informação relativa aos mecanismos específicos foi retirada da base de dados Uniprot à exceção das proteínas assinaladas com\* cuja informação foi retirada da referência indicada.

Mecanismos defensivos	Proteases para sistema imune do hospedeiro	Uniprot	Nome da proteína	Bactéria	Mecanismo específico
		B2RLK2*	Lys-gingipain (EC 3.4.22.47) (Lysine-specific cysteine proteinase Kgp)	<i>Porphyromonas gingivalis</i>	Hidrolisam a IgA e IgG humanas e o recetor C5a na superfície dos fagócitos. * Pubmed: PMID 15001226 PMID 24466435
		P72194*	Lys-gingipain 381 (EC 3.4.22.47)		
	Q51817*	Lys-gingipain W83			
	Proteínas de Stress	Q8DVR0	Biofilm regulatory protein A	<i>Streptococcus mutans serotype c (strain ATCC 700610 / UA159)</i>	Regulação da tolerância ao stress ácido e oxidativo.
		P66004	Dihydrolipoyl dehydrogenase (LPD) (EC 1.8.1.4) (Component of peroxynitrite reductase/peroxidase complex) (Component of PNR/P) (Dihydrolipoamide dehydrogenase) (E3 component of alpha-ketoacid dehydrogenase complexes)	<i>Mycobacterium tuberculosis</i>	Protege a bactéria do stress oxidativo gerado pelo sistema imune do hospedeiro pois tem funções de peroxinitrito redutase e degrada peroxidases de hidrogénio e alcalinas.
	Metabolismo celular	Q8ZQQ2	E3 ubiquitin-protein ligase SlrP (EC 6.3.2.-) (Secreted effector protein SlrP)	<i>Salmonella typhimurium</i>	Interfere com a via de ubiquitinação do hospedeiro. Promovendo a ubiquitinação da ubiquitina e da tioredoxina leva à redução da tioredoxina e portanto apoptose.
		P0CE12	E3 ubiquitin-protein ligase sspH2 (EC 6.3.2. (Secreted effect protein sspH2)	<i>Salmonella typhimurium</i>	Interfere com a via de ubiquitinação do hospedeiro pois liga-se à ubiquitina.

**Tabela I (continuação).** Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível da proteína. Toda a informação relativa aos mecanismos específicos foi retirada da base de dados Uniprot à exceção das proteínas assinaladas com\* cuja informação foi retirada da referência indicada. *S.* =*Salmonella*

Mecanismos defensivos	Sobrevivência intracelular	Uniprot	Nome da proteína	Bactéria	Mecanismo específico
		Q8ZMM8	Secreted effector protein pipB2 (Type III effector pipB2)	<i>S.typhimurium</i>	Reorganização doo endossoma/lisossoma para promover a sobrevivência da bactéria dento da célula do hospedeiro.
		O30916	Inositol phosphate phosphatase SopB (EC 3.1.3.-) (Effector protein SopB)	<i>S.typhimurium</i>	Participa na formação e manutenção dos vacuolos que contêm a <i>Salmonella</i> intracelularmente por ação sobre a PtdIns-3 mantendo os vacúolos que contêm a <i>Salmonella</i> fora da via endocítica levando à sobrevivência e reprodução da bactéria.
		P40722	Secreted effector protein SopD (Salmonella outer protein D)	<i>S.typhimurium</i>	Contribui para a replicação dentro dos macrófagos.
		Q8ZQC8	Secreted effector protein sopD2 (Salmonella outer protein D 2)	<i>S.typhimurium</i>	Contribui para a replicação dentro dos macrófagos participando na formação de <i>Salmonella induced fillaments</i> (Sifs) que são estruturas filamentosas contendo glicoproteínas das membranas dos lisossomas essenciais para a replicação da bactéria dentro das células do hospedeiro.
		Q56061	Secreted effector protein SifA	<i>S.typhimurium</i>	Responsável pela formação de Sifs.
		P74873	Secreted effector protein SptP [Includes: GTPase-activating protein (GAP); Tyrosine-protein phosphatase (EC 3.1.3.48).	<i>S.typhimurium</i>	Tem atividade de fosfatase de tirosina cuja função na internalização não é bem conhecida. Tem ainda função de ativação de GTPase e medeia a reversão das alterações do citoesqueleto provocadas pela SopE.
		Q7BVH7 O84947	Secreted effector protein SseB (Secretion system effector B) e Secreted effector protein SseC (Secretion system effector C)	<i>S.typhimurium</i>	Proteína translocadora da SPI-2 do Sistema de injeção que transfere proteínas da célula bacteriana dentro do fagossoma para a célula do hospedeiro.
		Q8ZQ79	Secreted effector protein Ssel	<i>S.typhimurium</i>	Inibe a migração normal dos macrófagos primários e das células dendríticas interagindo com o fator IQGAP1 do hospedeiro que regula a migração celular. Interage com o recetor do hospedeiro TRIP6.
		Q8ZP57	Secreted effector kinase SteC (EC 2.7.-.-) (Salmonella translocated effector C)	<i>S.typhimurium</i>	Cinase que promove a formação de redes de actina nas células infetadas.
P41783 P41784	Protein PrgH Protein PrgI	<i>S.typhimurium</i>	Promove a sinalização em células hospedeiras que levam a internalização da bactéria.		
Q8ZM36	Virulence protein STM3117	<i>S.typhimurium</i>	Interfere na modificação da camada de peptidoglicano da parede bacteriana interferindo desta forma com a replicação da bactéria dentro dos macrófagos.		

**Tabela I (continuação).** Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível da proteína. Toda a informação relativa aos mecanismos específicos foi retirada da base de dados Uniprot à exceção das proteínas assinaladas com\* cuja informação foi retirada da referência indicada.

	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	
Fatores não específicos	Sistemas de captação de Fe	P28784*	Gingipain R1 (EC 3.4.22.37) (Arg-gingipain)(Gingipain1) (RGP-1)	<i>Porphyromonas gingivalis</i>	Estas proteases de cisteína contribuem para a aquisição de ferro por parte da bactéria por contribuírem para a lise das hemácias e para a destruição dos tecidos do hospedeiro que levam à hemorragia. * Pubmed: PMID 15001226 PMID 24466435
		P95493*	Gingipain R2 (EC 3.4.22.37) (Arg-gingipain) (Gingipain 2) (RGP-2)	<i>Porphyromonas gingivalis</i>	
		B2RLK2*	Lys-gingipain (EC 3.4.22.47) (Lysine-specific cysteine proteinase Kgp)	<i>Porphyromonas gingivalis</i> (strain ATCC 33277 / DSM 20709 / JCM 12257)	
		P72194*	Lys-gingipain 381 (EC 3.4.22.47)	<i>Porphyromonas gingivalis</i>	
		Q51817*	Lys-gingipain W83	<i>Porphyromonas gingivalis</i>	
	Exoenzimas	Q8DPS0	Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate desidratase)	<i>Streptococcus pneumoniae</i> (strain ATCC BAA -255 / R6=)	Catalisa a conversão reversível de 2-fosfoglicerato para fosfoenolpiruvato
		Q8DTS9	Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate desidratase)	<i>Streptococcus mutans serotype c</i> (strain ATCC 700610/UA159)	Liga-se ao plasmidogénio e mucina MG2 do hospedeiro podendo contribuir para a degradação da matriz extracelular e transmigração da bactéria nos tecidos do hospedeiro.
		Q97QS2	Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate desidratase)	<i>Streptococcus pneumoniae serotype 4</i> (strain ATCC BAA-334/TIGR4)	
	Formação do biofilme	Q8DVR0	Biofilm regulatory protein A	<i>Streptococcus mutans serotype c</i> (strain ATCC 700610 / UA159)	Envolvida na formação do biofilme
		B2RLK2	Lys-gingipain (EC 3.4.22.47)	<i>Porphyromonas gingivalis</i> (strain ATCC 33277 / DSM 20709 / JCM 12257)	Envolvida na coagregação de <i>P.gingivalis</i> com outras bactérias orais.
P72194		Lys-gingipain 381 (EC 3.4.22.47)	<i>Porphyromonas gingivalis</i>		
Regulação dos genes de virulência	PODM78	Virulence transcriptional regulatory protein PhoP (Acid shock protein 29)	<i>Salmonella typhimurium</i>	Regula genes de virulência, adaptação ao stress ácido baixo Mg(+2) e resistência às defesas do hospedeiro.	
	P40676	Transcriptional regulator SlyA (Cytolysin SlyA) (Salmolysin)	<i>Salmonella typhimurium</i>	Regula genes de virulência e sobrevivência dentro dos macrófagos	

**Tabela II -** Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene

Mecanismos ofensivos	Adesão	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia
		C6AM75	Autotransporter adhesin	<i>Aggregatibacter aphrophilus</i> (strain NJ8700)	Medeia a ligação à superfície da mucosa oral.	Pubmed PMID:20688817
		G4BCZ5				
		G4BD01				
		B2RH54	Major fimbrial subunit protein (Fimbrillin)	<i>Porphyromonas gingivalis</i> (strain ATCC 33277 / DSM 20709 / JCM 12257)	Fibrilina é a subunidade estrutural das fímbrias, que são apêndices filamentosos sobre a superfície das células bacterianas.	Uniprot
		P59914				
		G4BD44	Extracellular matrix protein adhesin A	<i>Aggregatibacter aphrophilus</i> (strain NJ8700)	Responsável pela ligação ao colagénio.	Pubmed PMID:22958387
		C6AJR3				
		C2HR32	Surface protein	<i>Lactobacillus acidophilus</i> ATCC 4796	Medeia a ligação ao colagénio I e a fibronectina.	Science direct PMID: 11786253
	Q51845	Hemagglutinin A	<i>Porphyromonas gingivalis</i>	Hemaglutininas foram identificadas como importantes moléculas de adesão permitindo a <i>Porphyromonas gingivalis</i> aderir a células de tecido gengival por exemplo.	Uniprot Pubmed PMID:17220621	
Invasão celular	P10520	Streptokinase A	<i>Streptococcus pyogenes</i> serotype M1	Activa plasminogénio complexando-se com ele. Evita a formação de barreiras eficazes de fibrina em redor do local da infeção, o que contribui para a capacidade de invasão das células bacterianas.	Uniprot	

**Tabela II (continuação)**- Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene.

Mecanismos ofensivos	Sistema de secreção	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia
		H9L496	Secretion system apparatus	<i>Salmonella typhimurium (strain LT2 / SGSC1412 / ATCC 700720)</i>	Permite a injeção de proteínas efectoras directamente para o citoplasma da célula hospedeira eucariótica.	Uniprot Pubmed PMID:19390620
H9L4C5						
Lise	P0C0I3	Streptolysin O (Thiol-activated cytolysin)	<i>Streptococcus pyogenes serotype M1</i>	Toxina activada por sulfidrilo causando a citólise por formação de poros no colesterol das membranas do hospedeiro.	Uniprot	
	Q7ZAK5	Pneumolysin (Thiol-activated cytolysin)	<i>Streptococcus pneumoniae (strain ATCC BAA-255 / R6)</i>	Toxina activada por sulfidrilo causando a citólise por formação de poros no colesterol das membranas do hospedeiro	Uniprot	
	Q2TM57	Non-hemolytic enterotoxin A	<i>Bacillus cereus</i>	"Nhe " é uma abreviatura de " <i>non-hemolytic enterotoxin</i> ". Esta toxina é capaz de lisar os eritrócitos, embora em menor grau do que Hbl. Nhe é uma toxina de formação de poros constituído por três proteínas: NheA, NheB e NheC. Separadamente, estas proteínas não apresentam toxicidade.	Pubmed PMID:24040335	
	Q09K11					
	Q09K15					
	Q73A18					
	I6RD62	Nonhemolytic enterotoxin A (Fragment)	<i>Bacillus cereus</i>	NheC é obrigatório na etapa de preparação da formação de poros. NheA é obrigatório na etapa de final e dispara toxicidade por um mecanismo até agora desconhecido. NheB liga-se as membranas celulares independentemente dos outros componentes.		
	I6QU99					
	I6QQV3					
	I6RD71					
	I6UZK3					
	I6UGQ2					
	I6QQV6					
	I6RH94					
	I6RD68					
	I6V921					
	I6U705					
	I6QQV4					
I6QUA2						
I6VBD4						
I6RH90						
I6QQV5						
I6QUA6						
I6UGQ6						

**Tabela II-(continuação).** Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene.

Mecanismos ofensivos	Lise (continuação)	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia
		I6VBD9 I6UGP8 I6RH86 I6R786 I6V926 I6UZL0 I6U718 I6VBC7 I6R790 I6R794 I6UZJ7 I6V916	Nonhemolytic enterotoxin A (Fragment)	<i>Bacillus cereus</i>	"Nhe " é uma abreviatura de " <i>non-hemolytic enterotoxin</i> ". Esta toxina é capaz de lisar os eritrócitos de vários organismos, incluindo o ser humano, embora em menor grau do que Hbl. Nhe é uma toxina de formação de poros constituído por três proteínas: NheA, NheB e NheC. Separadamente, estas proteínas não apresentam toxicidade. NheC é obrigatório na etapa de preparação da formação de poros. NheA é obrigatório na etapa final e dispara toxicidade por um mecanismo até agora desconhecido. NheB liga-se as membranas celulares independentemente dos outros componentes.	Pubmed PMID:24040335
Q2TM56 Q73A17 B9IX99 Q3Y6N5 Q09K10 Q09K14	Non-hemolytic enterotoxin B	<i>Bacillus cereus</i> <i>Bacillus cereus</i> (strain ATCC 10987) <i>Bacillus cereus</i> (strain Q1) <i>Bacillus cereus</i>				
I6RD74 I6QUB6 I6QUB3 I6VBB6 I6RHA9 I6RHA3 I6RH98 I6RD79 I6QUA9 I6QQV8 I6UGN9 I6U6Z2 I6R7A3 I6U700	Nonhemolytic enterotoxin B (Fragment)	<i>Bacillus cereus</i>				

**Tabela II (continuação).** Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene.

Mecanismos ofensivos	Lise (continuação)	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia
		I6R798	Nonhemolytic enterotoxin B (Fragment)	<i>Bacillus cereus</i>	<p>“Nhe ” é uma abreviatura de “<i>non-hemolytic enterotoxin</i>”. Esta toxina é capaz de lisar os eritrócitos de vários organismos, incluindo o ser humano, embora em menor grau do que Hbl. Nhe é uma toxina de formação de poros constituído por três proteínas: NheA, NheB e NheC. Separadamente, estas proteínas não apresentam toxicidade.</p> <p>NheC é obrigatório na etapa de preparação da formação de poros. NheA é obrigatório na etapa de final e dispara toxicidade por um mecanismo até agora desconhecido. NheB liga-se as membranas celulares independentemente dos outros componentes.</p>	Pubmed PMID:24040335
		I6UZH7				
		I6QQV7				
		I6UGP3				
		I6UGN4				
		I6V910				
		I6R7A7				
		I6RD84				
		I6UZJ1				
I6UZI6						
I6VBC1						
I6V905						
I6R798						
I6V900						
Q51845	Hemagglutinin A	<i>Porphyromonas gingivalis</i>	Liga-se e lisa aos eritrócitos a fim de captar os iões de ferro essenciais para a nutrição das bactérias.	Uniprot Pubmed PMID:17220621		
Q7BYC7	Hemolysin BL binding component	<i>Bacillus cereus</i> (strain ATCC 14579 / DSM 31)	Hemolisina BL (HBL) é um sistema lítico na membrana composto por três proteínas distintas designada por B, L1, e L2. Foi proposto que as três proteínas HBL ligam-se aos eritrócitos de forma independente e formam um complexo de ataque à membrana que causa lise.	Pubmed PMID:22069656		
Q81BP9						
Q7BYC6	Hemolysin BL lytic component L1	<i>Bacillus cereus</i>				
Q81EZ7						
Q9L4L8	Hemolysin BL lytic component L2	<i>Bacillus cereus</i>				
Q81BP7		<i>Bacillus cereus</i> (strain ATCC 14579 / DSM 31)				
O05491		<i>Bacillus cereus</i>				
Q81EZ8		<i>Bacillus cereus</i> (strain ATCC 14579 / DSM 31)				
Q63CS4		<i>Bacillus cereus</i> (strain ZK / E33L)				

Tabela II (continuação). Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene.

Mecanismos ofensivos	Lise (continuação)	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia
		I6U6X0	Hemolysin BL component L2 (Fragment)	<i>Bacillus cereus</i>	Hemolisina BL (HBL) é um sistema lítico na membrana composto por três proteínas distintas designada por B, L1, e L2. Foi proposto que as três proteínas HBL ligam-se aos eritrócitos de forma independente e formam um complexo de ataque à membrana que causa lise.	Pubmed PMID:22069656
I6QQW3						
I6R7D1						
I6V8X9						
I6RHC4						
I6RDB0						
I6UZG0						
I6VB91						
I6V8Y4						
I6UGL3						
I6RHC7						
I6QUD5						
I6UZF4						
I6U6W4						
I6QQW4						
I6RDA7						
Q9K5F8						
Q9REG8						
B9IX98						
Q32TH6	Tripartite hemolysin BL component L2	<i>Bacillus cereu</i>				
Q32TH0						
Q32TH3						
Q32TI2						
Q32TH9						
G8U6L9	HbIB protein	<i>Bacillus cereus F837/76</i>				
Q9REG5					<i>Bacillus cereus</i>	

**Tabela II (continuação)**- Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene.

Mecanismos ofensivos	Lise (continuação)	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia
		Q09KJ1	Cytotoxin K	<i>Bacillus cereus</i>	CytK é uma toxina citotóxica, necrótica e hemolítica do <i>B. cereus</i> , e esta associada a intoxicação alimentar. CytR foi demonstrado possuir actividade enterotóxica contra o epitélio intestinal e tem a capacidade de formar poros. A Citotoxina K existe em duas formas diferentes, CYTK 1 e CYTK 2.	Pubmed PMID:22069656
		Q20L99				
		Q9EVA1				
		Q81GS6				
		Q63EQ2	CytK-2 protein	<i>Bacillus cereus</i>		
		Q937V2				
		Q937V0				
		Q937V1	Hemolysin A	<i>Aggregatibacter actinomycetemcomitans</i> serotype C (strain D11S-1)	HlyA pode penetrar na membrana das células do hospedeiro, criando um poro e levando-os a lise, que é um passo essencial para as bactérias iniciar o processo infeccioso. Esta toxina rompe membranas plasmáticas, causando a morte celular por necrose.	Pubmed PMID:22792337
		C9R5U4				
Q815P0	Perfringolysin	<i>Bacillus cereus</i> (strain ATCC 14579 / DSM 31)	Esta toxina rompe membranas plasmáticas, causando a morte celular por necrose.	Pubmed PMID:23748204		
C2DNR2	RTX-toxin family pore-forming toxin	<i>Escherichia coli</i> 83972	"RTX repeat domain" possui múltiplos de domínios de ligação Ca <sup>2+</sup> que são necessários para a actividade hemolítica.	Science direct PMID:19348784		
Q81AN8	Hemolysin II	<i>Bacillus cereus</i> (strain ATCC 14579 / DSM 31)	Hemolisina II (HlyII) de <i>Bacillus cereus</i> é responsável por causar a hemólise. HlyII tem propriedades hemolíticas contra os eritrócitos. HlyII forma poros através das membranas celulares. Esta formação de poros, eventualmente, leva a apoptose das células. HlyII induz também a lise das células fagocíticas.	Pubmed PMID:20731668 PMID:22522892		

**Tabela II (continuação)**- Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene. *A. actinomycetemcomitans* - *Aggregatibacter actinomycetemcomitans*

Mecanismos ofensivos	Apoptose	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia
		P16462	Leukotoxin (Lkt)	<i>A actinomycetemcomitans</i>	A leucotoxina (LtxA) induz a apoptose das células brancas do sangue de seres humanos. A via molecular que conduz à apoptose celular induzida e a paragem do ciclo celular por LtxA não é bem compreendido. Acredita-se que LtxA desempenha um papel crucial na evasão da resposta imune do hospedeiro pela bactéria. As evidências sugerem que a LtxA interrompe a resposta imune do hospedeiro, principalmente pela morte de células do sistema imunológico.	Pubmed PMID:23665198
		H0KFB7	Cytolethal distending toxin protein A	<i>A actinomycetemcomitans RhAA1</i>	Inibem a divisão celular causando a apoptose das células do hospedeiro. Todas as três subunidades são essenciais para conferir actividade plena da holotoxina.	Uniprot Pubmed PMID:22069704 PMID:17123907
		L8TZD5		<i>A actinomycetemcomitans serotype a str. A160</i>		
		G4AUG1		<i>A actinomycetemcomitans serotype b str. SCC1398</i>		
		L8UFU9		<i>A actinomycetemcomitans b str. S23A</i>		
		G4A1R5		<i>A actinomycetemcomitans serotype d str. I63B</i>		
		G3ZHS8		<i>A actinomycetemcomitans D17P-2</i>		
		G4AJD0		<i>A actinomycetemcomitans serotype e str. SCC393</i>		
		L8U8U9		<i>A actinomycetemcomitans serotype b str. SCC4092</i>		
G4B3V8	<i>A actinomycetemcomitans serotype b str. I23C</i>					
G4APJ0	<i>A actinomycetemcomitans serotype f str. D18P1</i>					

**Tabela II (continuação)**- Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene.

Mecanismos ofensivos		Apoptose (continuação)		Mecanismo específico	Bibliografia	
		Uniprot	Nome da proteína			Bactéria
Mecanismos ofensivos	Apoptose (continuação)	G4B6L8	Cytolethal distending toxin protein A	<i>A actinomycescomitans</i> serotype c str. SCC2302	Inibem a divisão celular causando a apoptose das células do hospedeiro. Todas as três subunidades são essenciais para a actividade plena da holotoxina.	Uniprot Pubmed PMID:22069704 PMID: 17123907
		G3ZYI7		<i>A actinomycescomitans</i> serotype a str. H5P1		
		G3ZAS6		<i>Aggregatibacter actinomycescomitans</i> D17P-3		
		L8U4U3		<i>A actinomycescomitans</i> serotype c str. AAS4A		
		C9R3B7		<i>A actinomycescomitans</i> serotype C (strain D11S-1)		
		G8MRB3		<i>A actinomycescomitans</i> ANH9381		
		I1XUC5		<i>A actinomycescomitans</i> D7S-1		
		H0KFB6	Cytolethal distending toxin protein B	<i>A actinomycescomitans</i> RhAA1		
		L8U170		<i>A actinomycescomitans</i> serotype a str. A160		
		G4AUG0		<i>A actinomycescomitans</i> serotype b str. SCC1398		
		L8UGT8		<i>A actinomycescomitans</i> serotype b str. S23A		
		G3ZHS9		<i>A actinomycescomitans</i> D17P-2		
		G4A1R4		<i>A actinomycescomitans</i> serotype d str. I63B		
		G4AJC9		<i>A actinomycescomitans</i> serotype e str. SCC393		
		L8UA25		<i>A actinomycescomitans</i> serotype b str. SCC4092		
		G4API9		<i>A actinomycescomitans</i> serotype f str. D18P1		

**Tabela II (continuação)**- Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com grau de evidência ao nível do gene.

Mecanismos ofensivos		Apoptose (continuação)		Mecanismo específico	Bibliografia	
		Uniprot	Nome da proteína			Bactéria
		G4B6L9	Cytolethal distending toxin protein B	<i>A actinomycescomitans serotype c str. SCC2302</i>	Inibem a divisão celular causando a apoptose das células do hospedeiro. Todas as três subunidades são essenciais para a actividade plena da holotoxina.	Uniprot Pubmed PMID:22069704 PMID: 17123907
		G3ZYI6		<i>A actinomycescomitans serotype a str. H5P1</i>		
		G3ZAS7		<i>A actinomycescomitans D17P-3</i>		
		Q9ZAB7		<i>A actinomycescomitans</i>		
		L8TYW2		<i>A actinomycescomitans serotype c str. AAS4A</i>		
		C9R3B8		<i>A actinomycescomitans serotype C (strain D11S-1)</i>		
		G8MRB4		<i>A actinomycescomitans ANH9381</i>		
		I1XUC6		<i>A actinomycescomitans D7S-1</i>		
		Q9S676		<i>A actinomycescomitans</i>		
		L1N4Q7		Putative cytolethal distending toxin subunit B		
		L1N4H4	Cytolethal distending toxin A/C family protein			
		L1N438				
		H0KFB5	Cytolethal distending toxin protein C	<i>A actinomycescomitans RhAA1</i>		
		G4AUF9		<i>A actinomycescomitans serotype b str. SCC1398</i>		
		L8UBE2		<i>A actinomycescomitans serotype b str. S23A</i>		
		G4A1R3		<i>A actinomycescomitans serotype d str. I63B</i>		

**Tabela II (continuação)**- Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene.

Mecanismos ofensivos	Apoptose (continuação)	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia
		G3ZHT0	Cytolethal distending toxin protein C	<i>A actinomycetemcomitans D17P-2</i>	Inibem a divisão celular causando a apoptose das células do hospedeiro. Todas as três subunidades são essenciais para a actividade plena da holotoxina.	Uniprot Pubmed PMID:22069704 PMID: 17123907
G4AJC8	<i>A actinomycetemcomitans serotype e str. SCC393</i>					
L8U4I2	<i>A actinomycetemcomitans serotype b str. SCC4092</i>					
G4B3W0	<i>A actinomycetemcomitans serotype b str. I23C</i>					
G4API8	<i>A actinomycetemcomitans serotype f str. D18P1</i>					
G4B6M0	<i>A actinomycetemcomitans serotype c str. SCC2302</i>					
G3ZYI5	<i>A actinomycetemcomitans serotype a str. H5P1</i>					
G3ZAS8	<i>A actinomycetemcomitans D17P-3</i>					
Q9F280	<i>A actinomycetemcomitans</i>					
L8U0L3	<i>A actinomycetemcomitans serotype c str. AAS4A</i>					
C9R3B9	<i>A actinomycetemcomitans serotype C (strain D11S-1) (Actinobacillus actinomycetemcomitans)</i>					
G8MRB5	<i>A actinomycetemcomitans ANH9381</i>					
I1XUC7	<i>A actinomycetemcomitans D7S-1</i>					
Q81AN8	Hemolysin II	<i>Bacillus cereus (strain ATCC 14579 / DSM 31)</i>	HlyII forma poros através das membranas celulares. Esta formação de poros, eventualmente, leva a apoptose das células. HlyII induz também a lise das células fagocíticas.	Pubmed PMID:20731668 PMID:22522892		

**Tabela II (continuação)**- Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene.

Mecanismos Defensivos	Defesas antifagocíticas	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia
		C2HR32	Surface protein	<i>Lactobacillus acidophilus</i> ATCC 4796	As proteínas das camadas superficiais conferem proteção contra o sistema do complemento e contra a fagocitose	Pubmed PMID:23677442
	P0C0H1	Hyaluronan synthase (EC 2.4.1.212) (Hyaluronate synthase) (Hyaluronic acid synthase) (HA synthase)	<i>Streptococcus pyogenes</i> serotype M1	A cápsula de ácido hialurônico esta implicada na patogenicidade dos <i>Streptococcus</i> do grupo A.	Uniprot	
	Proteases para Ig do hospedeiro	Q97QP7	Immunoglobulin A1 protease (IgA1 protease) (EC 3.4.24.13) (IgA-specific zinc metalloproteinase)	<i>Streptococcus pneumoniae</i> serotype 4 (strain ATCC BAA-334 / TIGR4)	Zinco metalloproteinase que cliva a IgA humana. Pode ser responsável pela infecção pneumocócica e é potencialmente envolvida em estágios distintos da doença pneumocócica.	Uniprot
		P72197	Lys-gingipain HG66 (EC 3.4.22.47) [Cleaved into: Lys-gingipain catalytic subunit; 39 kDa adhesin; 15 kDa adhesin; 44 kDa adhesin]	<i>Porphyromonas gingivalis</i>	Cisteína proteinase que hidrolisa a IgA e IgG. Perturba as funções dos leucócitos polimorfonucleares	Uniprot
	Protease para o sistema de complemento	P58099	C5a peptidase (EC 3.4.21.110) (SCP)	<i>Streptococcus pyogenes</i> serotype M1	Este fator de virulência de <i>S. pyogenes</i> cliva o C5a, destruindo a sua capacidade de servir como um quimioatratante.	Uniprot

**Tabela II (continuação)**- Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene.

		Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia
<b>Mecanismos defensivos</b>	<b>Sobrevivência intracelular</b>	P23988	Virulence membrane protein PagC	<i>Salmonella typhimurium</i> (strain LT2 / SGSC1412 / ATCC 700720)	Essencial para a sobrevivência dentro dos macrófagos.	Uniprot
		Q8ZQ59	Secreted effector protein PipB	<i>Salmonella typhimurium</i> (strain LT2 / SGSC1412 / ATCC 700720)	Promove a sobrevivência bacteriana nos tecidos do hospedeiro.	Uniprot
		Q56029	Virulence protein PagD	<i>Salmonella typhimurium</i> (strain LT2 / SGSC1412 / ATCC 700720)	Pode estar envolvido na sobrevivência do <i>S.typhimurium</i> dentro dos macrófagos.	Uniprot
		Q9KIB9	Secreted effector protein SifB	<i>Salmonella typhimurium</i> (strain LT2 / SGSC1412 / ATCC 700720)	Permite a sobrevivência bacteriana nos tecidos do hospedeiro.	Uniprot
		Q8ZPD7	Secreted effector protein SteA (Salmonella translocated effector A)	<i>Salmonella typhimurium</i> (strain LT2 / SGSC1412 / ATCC 700720)	Permite a sobrevivência bacteriana nos tecidos do hospedeiro.	Uniprot
		Q8ZPA6	Secreted effector protein SteB (Salmonella translocated effector B)	<i>Salmonella typhimurium</i> (strain LT2 / SGSC1412 / ATCC 700720)	Permite a sobrevivência bacteriana nos tecidos do hospedeiro.	Uniprot
		E1WFA1	Virulence transcriptional regulatory protein PhoP	<i>Salmonella typhimurium</i> (strain SL1344)	Essencial para a sobrevivência intracelular de <i>S. typhimurium</i> .	Uniprot

Tabela II (continuação) - Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com evidência a nível do gene.

	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia	
<b>Mecanismos defensivos</b>	<b>Proteínas de stress</b>	P0C0H3	DNA-binding protein HU	<i>Streptococcus pyogenes</i> serotype M1	Capaz de estabilizar o ADN e assim impedir a sua desnaturação, em condições ambientais extremas.	Uniprot
		Q9XB21		<i>Streptococcus mutans</i> serotype c (strain ATCC 700610 / UA159)		
		Q7CQN4	Major outer membrane lipoprotein 1 (Murein-lipoprotein 1)	<i>Salmonella typhimurium</i> (strain LT2 / SGSC1412 / ATCC 700720)	Permite que as bactérias resistem aos vários componentes tóxicos.	Uniprot
		Q8ZPP9	Major outer membrane lipoprotein 2 (Murein-lipoprotein 2)			
		E1WFA1	Virulence transcriptional regulatory protein PhoP	<i>Salmonella typhimurium</i> (strain SL1344)	Regula a expressão de genes envolvidos na virulência, adaptação ao ambiente ácido e de baixo teor em Mg (2 +) assim como a resistência a péptidos de defesa antimicrobianos do hospedeiro.	Uniprot
<b>Fatores de virulência não específicos</b>	<b>Sistema de captação de ferro</b>	Q51845	Hemagglutinin A	<i>Porphyromonas gingivalis</i>	Responsável pela lise dos eritrócitos para absorção dos iões de ferro necessários para a nutrição da bactéria.	Uniprot Pubmed PMID: 17220621
		P25806	Thiol protease (EC 3.4.22.-)	<i>Porphyromonas gingivalis</i> (strain ATCC BAA-308 / W83)	Estas proteases de cisteína contribuem para a aquisição de ferro por parte da bactéria por contribuírem para a lise das hemácias e para a destruição dos tecidos do hospedeiro que levam à hemorragia.	Pubmed PMID: 15001226 PMID:24466435
		P72197	Lys-gingipain HG66 (EC 3.4.22.47) [Cleaved into: Lys-gingipain catalytic subunit; 39 kDa adhesin; 15 kDa adhesin; 44 kDa adhesin]	<i>Porphyromonas gingivalis</i>		
	<b>Formação de biofilme</b>	P72197	Lys-gingipain HG66 (EC 3.4.22.47) [Cleaved into: Lys-gingipain catalytic subunit; 39 kDa adhesin; 15 kDa adhesin; 44 kDa adhesin]	<i>Porphyromonas gingivalis</i>	Envolvidas na coagregação de P.gingivalis com outras bactérias.	Uniprot
		A8AVS2	Serine protease challisin (EC 3.4.21.-)	<i>Streptococcus gordonii</i> (strain Challis / ATCC 35105 / CH1 / DL1 / V288)	A serina protease chalasina é capaz de inativar a molécula do quórum sensing CSP (competence-stimulating peptide) da <i>Streptococcus mutans</i> .	Pubmed PMIMD:21088146

**Tabela II (continuação)**- Resumo dos mecanismos de virulência associados às proteínas do microbioma oral com grau de evidência a nível do gene.

Regulação dos genes associados a virulência	Uniprot	Nome da proteína	Bactéria	Mecanismo específico	Bibliografia
	Q2FH47	Conserved virulence factor B	<i>Salmonella typhimurium</i> (strain LT2 / SGSC1412 / ATCC 700720)	Membro do sistema regulação de dois componentes PhoP / PhoQ que regula a expressão de genes envolvidos na virulência, adaptação ao ambiente ácido e de baixo teor em Mg (2 +) assim como a resistência a péptidos de defesa antimicrobianos do hospedeiro	Uniprot
	E1WFA1	Virulence transcriptional regulatory protein PhoP	<i>Salmonella typhimurium</i> (strain SL1344)	Contribui para a expressão de fatores de virulência. Envolvido na produção de hemolisina, ADNase, protease e proteína A	Uniprot

**Tabela III -** Proteínas para as quais não há informação na bibliografia acerca dos seus mecanismos específicos associados a patogénese.

Proteínas para os quais não há informação na bibliografia com as palavras-chaves definidas acerca dos seus mecanismos específicos associados a patogénese	Uniprot	Nome da proteína	Bactéria
	Q82ZX2	Cold shock-like protein CspLA (Cold-shock domain family protein)	<i>Enterococcus faecalis</i> (strain ATCC 700802 / V583)
	POA2N2	Virulence protein vsdE	<i>Salmonella typhimurium</i> (strain LT2 / SGSC1412 / ATCC 700720)
	C8MGI5	Exotoxin 15	<i>Staphylococcus aureus</i> A9635
	D3A8N8*	YadA-like domain protein (Fragment)	<i>Neisseria subflava</i> NJ9703
	B3WLE9	Putative toxin B	<i>Escherichia coli</i> B171
	Q20L98	Enterotoxin A	<i>Bacillus cereus</i>
	Q9S3N4		<i>Bacillus cereus</i>
	Q1XFK2		<i>Bacillus cereus</i>
	Q1XFK5		<i>Bacillus cereus</i>
	Q1XFK6		<i>Bacillus cereus</i>
	Q1XFJ6		<i>Bacillus cereus</i>
	Q1XFK0		<i>Bacillus cereus</i>
	Q1XFK1		<i>Bacillus cereus</i>
	Q1XFJ3		<i>Bacillus cereus</i>
	Q1XFJ2		<i>Bacillus cereus</i>
	Q1XFJ1		<i>Bacillus cereus</i>
	Q1XFJ9		<i>Bacillus cereus</i>
	Q1XFI9		<i>Bacillus cereus</i>
	Q1XFL0		<i>Bacillus cereus</i>
	Q1XFJ4		<i>Bacillus cereus</i>
	Q1XFJ5		<i>Bacillus cereus</i>
	Q1XFJ7		<i>Bacillus cereus</i>
	Q1XFJ0		<i>Bacillus cereus</i>
	Q1XFK9		<i>Bacillus cereus</i>
	Q1XFJ8		<i>Bacillus cereus</i>
	Q1XFK8	<i>Bacillus cereus</i>	
Q9S3N3	Enterotoxin B	<i>Bacillus cereus</i>	
Q1XFH6	Enterotoxin B (Fragment)	<i>Bacillus cereus</i>	
Q1XFH8		<i>Bacillus cereus</i>	
Q1XFH7		<i>Bacillus cereus</i>	
Q1XFH2		<i>Bacillus cereus</i>	
Q1XF10		<i>Bacillus cereus</i>	
Q1XFH5		<i>Bacillus cereus</i>	
Q1XF16		<i>Bacillus cereus</i>	
Q1XFH4		<i>Bacillus cereus</i>	
Q1XF15		<i>Bacillus cereus</i>	
Q1XFH3		<i>Bacillus cereus</i>	
D3A1B7*	Hep/Hag repeat protein (Fragment)	<i>Neisseria subflava</i> NJ9703	
Q1XBU9	Hemolysin B	<i>Bacillus cereus</i>	
Q1XBV2	Hemolysin C	<i>Bacillus cereus</i>	
C4FMC5*	Hep/Hag repeat protein	<i>Veillonella dispar</i> ATCC 17748	
C4FMC7*	Hep/Hag repeat protein	<i>Veillonella dispar</i> ATCC 17748	
Q813Q6	Non-expressed Enterotoxin C	<i>Bacillus cereus</i> (strain ATCC 14579 / DSM 31)	

\*O nome destas proteínas foi alterado na ferramenta bioinformática UniprotKB, ou seja, não corresponde aos nomes das proteínas que estão na base de dados Oral Card, versão de 15-12-2013. No Oral Card, as proteínas com os seguintes códigos Uniprot D3A1B7 e D3A8N8 estão associadas ao nome “*Hemagglutinin family*”, enquanto que C4FMC5 e C4FMC7 ao nome “*Putative uncharacterized protein*”.

Duas proteínas estão indicadas como “*putative*” (assinaladas com a cor azul), esse termo é usado para proteínas que exibem limitada similaridade de sequência com proteínas caracterizadas.

